Protein citrullination by *Porphyromonas gingivalis* and its implications for autoimmunity in rheumatoid arthritis

Natalia Wegner

Imperial College London
Faculty of Medicine
Kennedy Institute of Rheumatology Division

A thesis submitted for the degree of Doctor of Philosophy
March 2011

Abstract

Autoantibodies to citrullinated proteins are characteristic for rheumatoid arthritis (RA). RA has been associated with periodontitis in epidemiologic studies. Chronic periodontal infection with *Porphyromonas gingivalis* (*P. gingivalis*) is a possible risk factor for developing RA, as this bacterium expresses a peptidylarginine deiminase enzyme (PPAD) which could generate antigenic citrullinated proteins within an infectious context. The aim of this project was to investigate protein citrullination by *P. gingivalis* and its potential for breaking tolerance to citrullinated proteins in RA.

Endogenous citrullinated proteins were abundant in cell extracts of *P. gingivalis* but lacking in other tested oral bacteria. Deletion of the PPAD gene resulted in abrogation of protein citrullination. Inactivation of arginine-gingipains but not lysine-gingipains, the major proteolytic virulence factors, led to decreased citrullination. Incubation of wildtype *P. gingivalis* with fibrinogen or α-enolase caused degradation of the proteins by bacterial proteinases and peptidases and citrullination of the resulting peptides at carboxy-terminal arginine residues. No statistically significant antibody reactivity was found in RA serum with two of these peptides tested. The PPAD enzyme was cloned and expressed in *E. coli* in an enzymatically active form and used for *in vitro* enzymatic assays to confirm substrate specificity for carboxy-terminal arginine residues and demonstrate inhibition by 2-chloroacetamidine, but not by conventional RA drugs. Site-directed mutagenesis identified five amino acids crucial for catalysis. A polyclonal antibody was developed and demonstrated that PPAD is localised in the bacterial outer membrane.

The results demonstrate that *P. gingivalis* generates citrullinated endogenous and host peptides by proteolytic cleavage at arginine-X peptide bonds by arginine-gingipains followed by citrullination of carboxy-terminal arginines by PPAD. Mutagenesis and inhibition studies point to a critical cysteine (Cys-351) in the catalytic centre. The work provides the basis for future studies of the role of PPAD in bacterial virulence and tolerance breakdown in RA.

Table of Contents

Acknowledgements

I thank all the people who helped me to achieve this, be it through scientific, personal or financial help, guidance, advice and support. I am very grateful for all the generousity I have received and the experiences I collected during the last three years.

Declaration of Originality

Everything that is presented in this thesis is my own work and help from other people or the use of external sources has been appropriately acknowledged.

List of Tables and Figures

Chapter 1 Introduction

Chapter 2 Materials and Methods

Chapter 5 Generation and evaluation of an anti-PPAD antibody and analysis of native PPAD and citrullination *in vivo*

Chapter 6 Discussion and future directions

Abbreviations

ACPA	anti-citrullinated protein antibodies
ACR	American College of Rheumatology
ADI	arginine deiminase
ADMA	asymmetric dimethylarginine
AT	amidinotransferase
ATCC	American type culture collection
AMPD	2-amino-2-methyl-1,3-propanediol
BAEE	benzoyl-L-arginine ethylester
BCA	bicinchoninic acid
BLAST	basic local alignment search tool
bp	base pairs
CAL	clinical attachment loss
CCP	cyclic citrullinated peptides
CD	cluster of differentiation
CEP-1	citrullinated enolase peptide-1
CHES	2-(cyclohexylamino) ethanesulfonic acid
CI	confidence interval
CIA	collagen-induced arthritis
CRP	C-reactive protein
CXCL	CXC chemokine ligand
Da	Dalton
DAS28	disease activity score 28
DC	detergent-compatible
DDAH	dimethyl-L-arginine deiminase
DMARDs	disease-modifying anti-rheumatic drugs
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
ds	double-stranded
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDTA	ethylenediamine tetraacetic acid

ELISA	enzyme-linked immunosorbent assay
EULAR	The European League Against Rheumatism
ESR	erythrocyte sedimentation rate
E-value	expect value
Fc	fragment crystallizable region (antibody)
FcR	Fc receptor
g	gravitational constant
GCF	gingival crevicular fluid
GME	guanidino group modifying enzymes
GST	glutathione-S-transferase
HAQ-DI	Health Assessment Questionnaire Disability Index
His	histidin
HLA	human leukocyte antigen
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
IFN	interferon
Ig	immunoglobulin
IL	interleukin
IPTG	isopropyl β-D-1-thiogalactopyranoside
IUPAC	International Union of Pure and Applied Chemistry
kb	kilo base pairs
kDa	kilo Dalton
Kgp	lysine-gingipains
LB	Luria Bertani
LDS	lithium dodecyl sulfate
L-NMMA	N^G-monomethyl-L-arginine
LPS	lipopolysaccharid
MHC	major histocompatibility complex
MOPS	3-(N-morpholino) propanesulfonic acid
MS/MS	tandem mass spectrometry
NETs	neutrophil extracellular traps
NSAIDs	nonsteroidal anti-inflammatory drugs
OA	Osteoarthritis

OD	optical density
OR	odds ratio
PAD	peptidylarginine deiminase
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with Tween
PCR	polymerase chain reaction
PD	periodontitis
PGE_2	prostaglandin E_2
pI	isoelectric point
PPAD	*P. gingivalis* peptidylarginine deiminase
PTPN22	protein tyrosine phosphatase, non-receptor type 22
RA	rheumatoid arthritis
REP-1	arginine enolase peptide-1
RF	rheumatoid factor
Rgp	arginine- gingipains
RNA	ribonucleic acid
rpm	rounds per minute
SDS-PAGE	sodium dodecylsulfate – polyacrylamid gel electrophoresis
SE	shared epitope
SNP	single nucleotide polymorphism
SOC	super optimal broth including glucose
TBS	Tris buffered saline
TBS-T	Tris buffered saline with Tween
TIGR	The Institute for Genomic Research
TNF-α	tumor necrosis factor – alpha
TLR	toll-like receptor
Tris	tris(hydroxymethyl)aminomethane
Trx	thioredoxin

Chapter 1 Introduction

This project investigates the mechanism of protein citrullination by the oral bacterium *Porphyromonas gingivalis* (*P. gingivalis*) and evaluates its potential to prime or perpetuate the autoimmune response to citrullinated proteins in rheumatoid arthritis.

Rheumatoid arthritis and autoimmunity to citrullinated proteins

Rheumatoid arthritis - classification, aetiology, pathophysiology and current treatment

Rheumatoid arthritis (RA) is a systemic autoimmune disease of unknown aetiology, affecting 0.5-1.0% of the adult population (Symmons *et al.*, 2002; Helmick *et al.*, 2008). It is a complex, heterogenous disease in which chronic inflammation of synovial joints and erosion of bone result in joint destruction, pain, disability, and premature mortality (reviewed by Firestein 2003).

Classification

To aid uniform classification of RA, the American College of Rheumatology (ACR) published the former 'ACR 1987 classification criteria for rheumatoid arthritis' (Arnett *et al.* 1988) (Table 1.1).

Table 1.1 1987 ACR criteria for the classification of rheumatoid arthritis

1	Morning stiffness in and around joints lasting at least 1 hour before maximal improvement
2	Soft tissue swelling (arthritis) of 3 or more joint areas observed by a physician
3	Swelling (arthritis) of the proximal interphalangeal, metacarpophalangeal, or wrist joints
4	Symmetric swelling (arthritis)
5	Rheumatoid nodules
6	The presence of rheumatoid factor
7	Radiographic erosions and/or periarticular osteopenia in hand and/or wrist joints

Criteria 1 to 4 were required to have been present for at least 6 weeks, and RA was defined by the presence of 4 or more criteria. These criteria demonstrated 91-94% sensitivity and 89% specificity for RA when compared with non-RA rheumatic disease control subjects. However, over time, they proved insufficient to discriminate early RA from self-limiting or other inflammatory arthritides, and therefore to identify those patients who would benefit from early treatment. Further, the emerging findings of the importance of antibodies to citrullinated proteins in patients with early and established RA as a predictive marker for disease progression and severe disease triggered a concerted effort towards defining updated classification criteria, which were published recently as the '2010 ACR/EULAR (European League Against Rheumatism) classification criteria for RA' (Aletaha *et al.*, 2010).

Table 1.2 2010 ACR/EULAR criteria for the classification of RA

Criteria can be applied to individuals who have at least one joint with synovitis which is not better explained by another disease. RF = rheumatoid factor; ACPA = anti-citrullinated protein antibodies; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate. Definitions and detailed explanations for fulfilment of each of the features can be found in (Aletaha et al., 2010).

Category	**Variables**	**Score**
Joint involvement	1 large joint	0
	2-10 large joints	1
	1-3 small joints (with or without large joints)	2
	4-10 small joints (with or without large joints)	3
	>10 joints (at least 1 small joint)	5
Serology (at least 1 test)	Negative RF and negative ACPA	0
	Low-positive RF or low-positive ACPA	2
	High-positive RF or high-positive ACPA	3
Acute-phase reactants	Normal CPR and normal ESR	0
(at least 1 measure)	Abnormal CPR or abnormal ESR	1
Duration of symptoms	<6 weeks	0
	≥6 weeks	1

According to these new criteria, ‘definite RA’ is classified based on synovitis in at least 1 joint, no alternative diagnosis that better explains the synovitis, and a total score of at least 6 out of 10, added up from 4 categories (Table 1.2). Additionaly, patients can be classified as having ‘definite RA’ if they present with obvious erosions typical for RA or a longstanding disease history where they have previously fulfilled the criteria. Additionaly, patients with early disease that are being treated might not fulfil the criteria at initial presentation but at a later stage as the condition evolves.

These new criteria allow a standardised identification of patients who first present with an undifferentiated inflammatory synovitis and are at a sufficiently high risk of developing RA to be considered for therapeutic intervention. In turn, this means that the RA literature to date has to be viewed under the aspect of the old classification criteria, where subjects with early RA would not had been routinely included in studies. This applies in particular to clinical trials, genetic association and serological studies.

In light of the work presented in this thesis, the inclusion of anti-citrullinated protein antibody status into the classification criteria is particularly welcomed, as it formally underscores the importance of these autoantibodies in a large group of patients with RA.

Aetiology

Despite the heterogeneity of RA, it is generally considered an autoimmune disease. Evidence of autoimmunity includes the presence of autoantibodies, the composition of the cellular infiltrate in the inflamed joint featuring activated B cells, the linkage of RA to certain MHC class II alleles and the benefits of B-cell depleting therapy.

Genetic influences are estimated to be responsible for around 50% of the risk of developing RA (MacGregor *et al.,* 2000; van der Woude *et al.,* 2009) with hormonal, environmental, and other factors (epigenetics, post-translational modification, stochastics) explaining the remaining risk.

Amongst the major and best studied genetic risk factors identified so far is a group of MHC class II gene variants collectively called the ‘shared epitope’ (SE). MHC class II molecules

are glycoproteins expressed on the surface B cells, dendritic cells, macrophages and other cells and are involved in the presentation of peptide antigens to T cells. Each molecule consists of a noncovalent complex of two chains, α and β, each of which is made up of two domains (Figure 1.1).

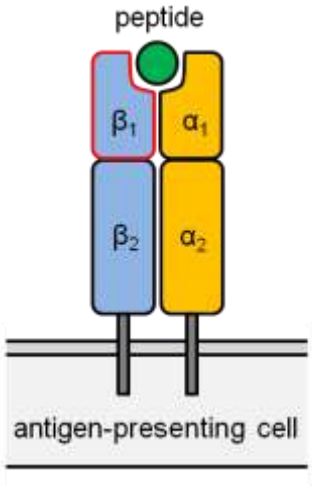


Figure 1.1 Schematic structure of an MHC class II molecule. The MHC class II molecule is composed of two transmembrane glycoprotein chains (α and β). Each chain consistes of two domains (1 and 2). Domains β_1 and α_1 form the peptide-binding cleft, which binds peptides (green) for presentation to T cells. The β_1 domain (outlined in red) is the source of the majority of genetic polymorphism in the rheumatoid arhtitis-relevant HLA-DR subtypes of the MHC class II molecules. This figure was generated by myself using Microsoft Powerpoint.

The α_1 and β_1 domains form the walls of a cleft on the surface of the molecule (peptide-binding cleft) to which peptides bind for presentation and interaction with T cells. This is also the site of major polymorphism, allowing the binding of a wide range of peptides. There are three pairs of MHC class II α- and β-chain genes: HLA-DP, -DQ, and –DR. In HLA-DR, the β-chain is encoded by four different loci (HLA-DRB1, -DRB3, -DRB4, -DRB5), which show considerable polymorphism. In total, more than 400 different alleles are known for the genes encoding the HLA-DR β-chain (indicated by a star [*] followed by four digits). The shared epitope is made up of the alleles DRB1*0401, *0404, *0408, *0405, *0101, *0102, *1001 and *1402, which are expressed in the serotypes HLA-DR4, -DR1 and -DR10 (Zanelli *et al.* 2000b). All alleles share variants of the Q/R-K/R-R-A-A amino acid motif present in the peptide-binding cleft of the MHC class II β chain (Gregersen *et al.* 1987). There is much speculation in the literature regarding the underlying mechanism of SE-RA association. Hypotheses range from a direct role of the SE on increased affinity and presentation of autoantigenic peptides and activation of self-reactive T cells (Holoshitz 2010), to a decreased

activation of regulatory T cells (Zanelli *et al.* 2000a), an influence on thymic T cell repertoire selection (Bhayani *et al.* 1991), to a role of the SE as an innate immune system activator (Ling *et al.*, 2007; De Almeida *et al.*, 2010).

Women are three times as likely as men to develop RA (Pritchard 1992). The mechanism underlying this observation has been under intensive research and findings are controversial, but most studies show an increased risk of RA associated with pregnancy and breastfeeding and a decreased risk associated with the prolonged use of oral contraceptives (≥7 years) (Berglin *et al.* 2010).

Environmental factors are considered to contribute to the onset of RA in a genetically predisposed individual. Evidence mainly stems from the low disease concordance rate (15%) in monozygotic twins (Silman *et al.* 1993) and the declining incidence of RA in genetically predisposed populations such as the Pima Indians (Jacobsson *et al.*, 1994; Ferucci *et al.*, 2005). The best-studied environmental factors to date are smoking and infections.

Infection as a contributory or even sole aetiological factor in RA has been suggested and discussed for decades, but no convincing data has been produced pinpointing any specific organism (discussed by Carty *et al.* 2004). No major viable organism could be isolated specifically from the rheumatoid joint to date. However, the relationship between a microorganism and RA might be more complex and subtle, and trigger or contribute to the disease by one of its products rather than cause it directly by residing at the diseased site, the joint. Temporal aspects might also play a role, such as a time lag between the microbial triggering event(s) and the onset of RA, which would be difficult to account for in studies. Findings supporting a role of microorganisms in RA also come from arthritis animal models, where bacterial components, on their own (Kim *et al.* 2009) or combined with protein antigen such as collagen type II (Luross *et al.* 2001), are routinely used to induce or exacerbate arthritis (reviewed by Asquith *et al.* 2009). In view of the combined data, a direct, causal role between a specific microorganism and RA still seems unlikely, but a more indirect contributory/triggering function leading to RA in a subset of predisposed individuals appears possible and is explored in depth in this thesis, focussing on the possible role of the bacterium *P. gingivalis*.

Cigarette smoking is today the best-known environmental risk factor for the development of RA. The link to smoking was first recognized in 1987 (Vessey *et al.* 1987) as an unexpected finding in a study investigating the association between RA and the use of oral contraceptive, and later confirmed in a number of case-control and cohort studies (reviewed by Sugiyama *et al.* 2010). The most striking results were presented from the Arthritis and Rheumatism Council Twin Study, where 13 pairs of monozygotic twins, discordant for RA and smoking, were identified, and in 12 out of 13 cases the RA patient was also the smoker (Silman *et al.* 1996). It is currently unknown how smoking may contribute to the development of RA. Cigarette smoke contains thousands of toxic compounds (Pryor *et al.* 1993) which may cause tissue damage and inflammation and modulate the immune cell function and cytokine production (Costenbader *et al.,* 2006; Baka *et al.,* 2009). However, it is important to bear in mind that *in vitro* smoking studies use different compounds, cells, and experimental conditions to mimic smoking, and therefore afford only a limited reproducibility of the true short- and long-term physiological events accompanying tobacco smoking.[1]

Pathophysiology

The classic characteristic of RA is the chronically inflamed synovial joint. Experimental evidence to date suggests the following key players in mediating the pathology: Autoantibodies and cytokines, degradative enzymes, T and B cells, macrophage-like and fibroblast-like synovial cells, macrophages, dendritic cells, neutrophils, osteoclasts and chondrocytes, with both the innate and adaptive immune system playing a role (reviewed by Firestein 2003; Sidiropoulos *et al.* 2008). These components presumably interact in a highly complex system, with substantial variations between individuals.

Cytokines have been in the centre of interest lately due to the success of anti-tumour necrosis factor (TNF)-α therapy in RA. Besides TNF-α, several other cytokines are believed to play a role in pathophysiology, such as IL-1, IL-6, IL-15, IL-17, IL-18, IL-32, and IL-33, leading to activation and recruitment of inflammatory cells (reviewed in McInnes *et al.* 2007).

[1] This paragraph was originally published in Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables (2010a). "Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis." Immunol Rev **233**(1): 34-54.

T cells can be found in large numbers in the rheumatoid joint. Their definite role is unclear and seems to differ between animal models and human disease, but involves cytokine secretion and effector and regulatory function (reviewed by Lundy *et al.* 2007).
B cells moved back into the centre of RA research recently because of the advent of B cell-depleting therapeutics such as the CD20-specific monoclonal antibody rituximab, which depletes all B cell subsets except plasma cells. In addition to autoantibody production and immune complex formation, B cells are considered to contribute to RA pathology through the production of cytokines and chemokines, antigen presentation and by exerting regulatory effects on lymphoid neogenesis, follicular dendritic cells and T cell-macrophage and T cell-B cell interactions (reviewed by Mauri *et al.* 2007). The role of autoantibodies in RA pathology will be discussed separately.

In addition to these classical mediators of adaptive immunity, several other cell types have recently attracted interest in rheumatology research: macrophage-like and fibroblast-like synoviocytes are believed to be potent producers of pro-inflammatory cytokines and to invade cartilage, respectively (reviewed by Muller-Ladner *et al.* 2007). Activated osteaclasts contribute to bone erosion (reviewed by Schett 2007). Neutrophils synthesise a wide variety of cytokines and release chemokines, prostaglandins, reactive oxygen and nitrogen species and were therefore suggested to contribute to the inflammatory and hypoxic milieu in the RA joint (reviewed by Cascao *et al.* 2010). Macrophages are believed to mediate and amplify tissue distruction by chemokine-mediated recruitment of pro-inflammatory and antigen-presenting cells, cytokine-mediated activation of inflammatory cells and secretion of matrix-degrading enzymes (reviewed by Kinne *et al.* 2007). Dendritic cells' major role most likely lies in antigen presentation and MHC-restricted autoimmune response priming, and the production of innate immune inflammatory mediators (reviewed by Lutzky *et al.* 2007). Data on a direct involvement of chondrocytes in RA are sparse but suggest that chondrocytes secrete pro-inflammatory cytokines and release degradative enzymes, participating in the distruction of the cartilage matrix (reviewed by Otero *et al.* 2007).

Current treatment

The aim of current treatment is to achieve disease remission or a low state of disease activity and prevent long-term disabilities. Traditional drugs are nonsteroidal anti-inflammatory drugs (NSAIDs), corticosteroids and disease-modifying anti-rheumatic drugs (DMARDs)

(reviewed by O'Dell 2004; and Scott *et al.* 2010), and, more recently, biological agents were included in the spectrum.

NSAIDs were traditionally used in the first few weeks of symptoms for pain relief and to reduce stiffness until a definitive diagnosis of RA has been made. In the long-term, however, they do not slow progression of disease and may result in many adverse effects, such as gastrointestinal ulcer, perforation, haemorrhage and cardiovascular disease. At present, the use of NSAIDS as 'first line therapy' has been replaced by the early use of DMARDS and and corticosteroids. Corticosteroids are potent suppressors of inflammation and are given to reduce synovitis in the short-term and to decrease joint damage in the long-term. However, they come with a range of severe adverse effects such as infections, thinning of the skin, osteoporosis, hypertension, cataracts, and hyperlipidemia. DMARDs are a heterogenous group of drugs grouped together by use and convention. They reduce joint swelling and pain, decrease acute-phase markers, limit progressive joint damage and improve function. Their mechanisms of action are believed to be diverse and are incompletely understood. The major DMARDs are methotrexate, sulfasalazine, leflunomide, hydroxychloroquine and chloroquine. Gold and cyclosporine are additional DMARDs but their use is limited by toxic effects. Adverse effects include nausea and hepatotoxicity, blood dyscrasias (abnormal constitution of blood) and interstitial lung disease. Methotrexate is usually the first DMARD administered to people with RA and is often combined with other DMARDs, corticosteroids or biological agents to increase efficacy.

Biological agents are antibody-based drugs such as TNF-α inhibitors (etanercept, infliximab, adalimumab, certolizumab, golimumab), T cell co-stimulation blockers targeting the B7 receptor (abatacept), B cell depletors targeting CD20 (rituximab), and IL-6 inhibitors (tocilizumab). The discovery of TNF-α inhibitors, the first biological agents, as a highly effective treatment for RA at this institution (Elliott *et al.* 1993; reviewed by Feldmann 2002), has revolutionised RA treatment due to the halting and even repair of joint destruction (Lipsky *et al.* 2000). However, adverse effects exist, such as reactions and infections at drug infusion and injection sites and an increased risk of tuberculosis.

Despite these advances, not every patient can be treated successfully; life-long ongoing treatment is often required and complete remission is rare. Further, the overall treatment and societal cost is substantial. Therefore, focus is now shifting towards strategies to prevent the

disease by identifying people with a high risk of developing it on the basis of genetic and serological markers, and to cure the disease by treating patients with very early signs of arthritis intensively (discussed by Machold 2010). The strategy of prevention is particularly relevant to this project and will be discussed accordingly.

Protein citrullination - mechanism and its role in health and disease

Citrulline, in the context of a peptide backbone, is a non-standard amino acid that results from posttranslational modification of arginine residues. In humans it is catalysed by the peptidylarginine deiminase (PAD) family of enzymes in the presence of sufficient concentrations of Ca^{2+}. This conversion, termed citrullination or deimination, reduces the net charge of the protein by the loss of a positive charge per citrulline residue (Figure 1.2).

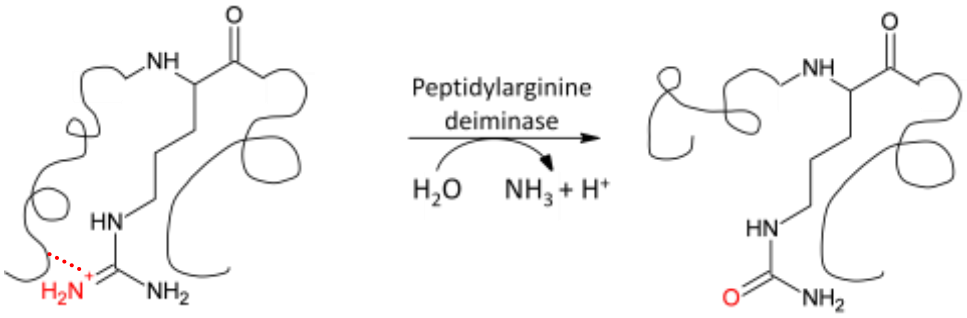


Figure 1.2 Enzymatic conversion of peptidylarginine into peptidylcitrulline, catalysed by the family of peptidylarginine deiminase enzymes. The positively charged arginine guanidino group is converted into the neutral citrulline ureido group in the presence of water, yielding ammonia and a hydrogen ion as by-products. The possible consequences on the three-dimensional protein structure are shown schematically. The red dotted line represents hydrogen bonds (with the protein backbone or with polar amino acid side chains) or ionic interactions (with negatively charged amino acid side chains). This figure was originally published in (Wegner *et al.* 2010a).

Arginine residues within polypeptides often play a central role in the structural integrity of a protein, due to their ability to participate in ionic interactions with negatively charged amino acid side chains, substrates and cofactors, and to form multiple hydrogen bonds to both the peptide backbone and other amino acid side chains (Borders *et al.* 1994). Arginine has got the most polar of all the common amino acid side chains and is therefore the amino acid that is most likely to be found on the surface of proteins in an aqueous environment. Citrullination

would be expected to destroy the ionic interactions, interfere with hydrogen bonds, and create new interactions. In fact, the unfolding of certain proteins upon citrullination was reported to be similar to that observed in the presence of high concentrations of urea (>4.5 M) (Tarcsa *et al.* 1996). Hence, the conversion of arginine into citrulline may result in an altered three-dimensional structure and function of a protein.[2]

Human peptidylarginine deiminases (PADs): Enzymes that catalyse protein citrullination

PAD enzymes are activated by higher Ca^{2+} concentrations than that present in intact cells. The Ca^{2+} concentration threshold for PAD activity is $>10^{-6}$ M, but this varies with the source of PAD enzyme and *in vitro* substrate used (Inagaki *et al.* 1989; Nakayama-Hamada *et al.* 2005). The estimates for the intracellular Ca^{2+} concentration range from 10^{-9} to 10^{-7} (Orlov *et al.* 2003). Five members of the PAD family (PAD1, -2, -3, -4, -6) have been found in humans, sharing 50-55% sequence identity overall. They have different tissue distribution, subcellular localisation and substrates (reviewed by Vossenaar *et al.* 2003b). Of particular relevance to RA are PAD2 and PAD4. PAD2 is distributed widely in various tissues, and is especially abundant in musle and brain. PAD4 is mainly located in bloodstream granulocytes. However, both their expression has been demonstrated in rheumatoid synovial membrane (Chang *et al.* 2005; De Rycke *et al.* 2005; Nakayama-Hamada *et al.* 2005; Foulquier *et al.* 2007), synovial fluid cells (Vossenaar *et al.* 2004b), and extracellularly in synovial fluid (Kinloch *et al.* 2008).[2]

Structure and classification of human PAD4

Human PAD4 is the only protein of the PAD family for which the three-dimensional structure has been solved (Arita *et al.* 2003). Based on X-ray crystallographic data, it provided an insight into the catalytic residues of PAD4 and its structural basis for Ca^{2+} dependency (Arita *et al.* 2004) and preference for arginine residues in certain amino acid contexts (Arita *et al.* 2006).

PAD4 is organised into two domains: The N-terminal domain (Met1-Pro300) consists of two immunoglobulin-like subdomains and the nuclear localisation signal (NLS), the C-terminal

[2] This paragraph was originally published in: Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables. 2010. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol Rev* 1:34-54

domain (Asn301-Pro663) is folded into an α/β propeller (five ββαβ modules arranged circularly in a pseudo-five fold symmetric structure), containing the active site (Figure 1.3).

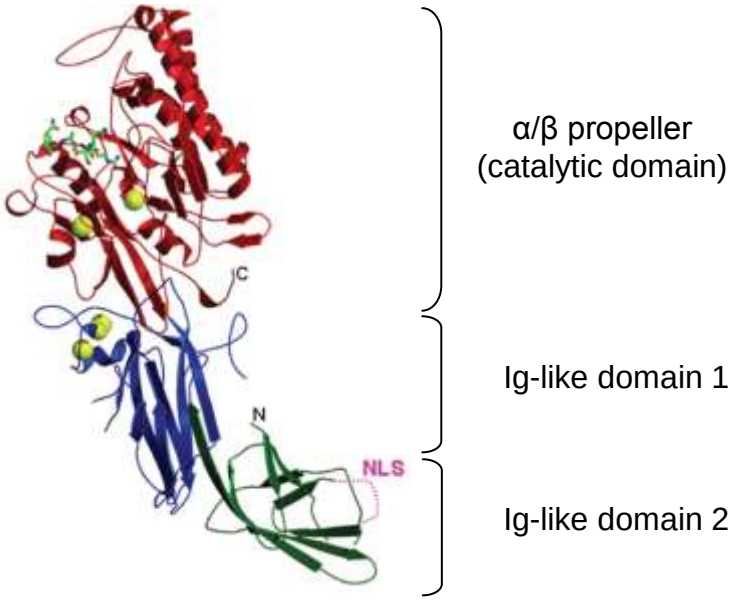


Figure 1.3 Ribbon representation of the structure of Ca^{2+}-bound PAD4 in complex with a histone peptide. Ca^{2+} ions are shown as yellow balls, the histone peptide is shown as a green stick model. Carboxy- and amino-termini are indicated (C and N, respectively). NLS: nuclear localisation sequence. Figure adapted from (Arita *et al.* 2006).

PAD4 exists as a dimer in the protein crystal and in solution, formed by head-to-tail contact between the N-terminal domain of one molecule and the C-terminal domain of the other. Each PAD4 monomer binds five Ca^{2+} ions, three in the N-terminal domain and two in the C-terminal active site domain. The Ca^{2+} ions in both domains induce a conformational change, but only the two Ca^{2+} ions in the active site domain are essential for recognition of substrate and catalysis by stabilising disordered regions around the active site cleft and organising it. The active site residues were studied by site-directed mutagenesis and were identified to be cysteine-645, histidine-471, aspartic acid-350 and -473. Cysteine and histidine are involved in catalysis, whereas the two aspartic acid residues are crucial for substrate binding and orientation by hydrogen bonding of the arginine side chain nitrogen atoms. Mutation of any of these residues to alanine abolished enzymatic activity (Arita *et al.* 2004).

The C-terminal domain of human PAD4, in its active Ca^{2+}-bound state, is structurally similar to other arginine-processing enzymes which are calcium-independent, such as

amidinotransferase (AT), dimethyl-L-arginine deiminase (DDAH), arginine deiminase (ADI) and *P. gingivalis* PAD (PPAD). Despite no overall sequence similarity between these enzymes, their (proposed) catalytic residues overlap (Arita *et al.* 2004; Shirai *et al.* 2006). Based on these findings, the enzymes have been assigned to a new superfamily, called the guanidino-group modifying enzymes (GME) (Shirai *et al.* 2006).

Substrate specificity and enzymatic mechanism of human PAD4

The following catalytic mechanism was proposed: Side chain carboxyl groups of Asp-350 and Asp-473 form hydrogen bonds and a salt bridge with the substrate peptidyl-L-arginine. A nucleophilic attack by the thiol group of Cys-645 on the Cζ atom of peptidyl-L-arginine cleaves the bond between this carbon atom and nitrogen η2, generating ammonia. A water molecule, activated by His-471, forms a second nucleophilic attack on Cζ, and hydrolysis yields peptidyl-L-citrulline (Arita *et al.* 2004).

Building up on this work, Arita *et al*. studied the interaction between PAD4 and peptide substrates from the N-terminal tail of histones H3 and H4 (Arita *et al.* 2006). They reported that PAD4 recognises five successive residues of the peptide substrate via hydrogen bonds to the backbone of the peptide. It was suggested that the substrate specificity is fairly broad, with the only restriction being at the N-2 position (two amino acids before arginine), where amino acids with small side-chains are favoured. The five residues were further observed to form a bent conformation within the active site cleft, and therefore flexible, unstructured, surface-exposed peptide regions in substrates might be favoured by PAD4, which is in line with experimental results (Takahara *et al.* 1985; Kurokawa *et al.* 1987; Tarcsa *et al.* 1996).

Human PAD4 does not catalyse conversion of free L-arginine to L-citrulline. This was suggested to be due to essential hydrogen bond interaction between PAD4 (via Arg-372, Arg-374, Arg-639) and atoms of the peptide bond formed between arginine and the preceding amino acid, which are absent in free L-arginine, but required for recognition of substrate (Arita *et al.* 2004). Early *in vitro* studies using rabbit PAD2 further showed that highest activity is observed towards substrates where both the amino- and carboxyl-groups of arginine are substituted with a peptide or ester bond, equivalent to internal rather than C- or N-terminal peptidylarginine (Sugawara *et al.* 1982).

Significance of citrullination in healthy physiology[3]

In healthy physiology, citrullinated proteins are present in a variety of cells and tissues. An overview of the type of citrullinated proteins, localisation, function and relevant references is given in Table 1.3. Citrullinated proteins play a major role in the generation of structural tissue, the proper functioning of the central nervous system, in transcriptional regulation and in cytoskeletal stability. Recently, a direct role for protein citrullination in immune system regulation has been suggested by demonstrating that stimulation of peripheral blood mononuclear cells with IFN-γ and dsRNA caused citrullination of the chemokines CXCL8 and CXCL10 with fundamental effects on their receptor usage, proteolytic processing and biological activities. In innate immunity, histone citrullination has been shown to result in the decondensation of chromatin and formation of neutrophil extracellular traps (NETs) as an innate response to infectious and inflammatory stimuli. *In vitro* experiments have demonstrated that fibrinogen can be citrullinated, resulting in defective thrombin-catalysed polymerisation, but the physiological role of this is unclear.

Table 1.3 Citrullinated proteins in healthy human physiology

Citrullinated protein	**Localisation**	**Function**	**References**
Keratin, Filaggrin	Epidermis	Cornification of epidermins, skin moisturisation	(Steinert *et al.* 1979) (Tarcsa *et al.* 1996)
Trichohyalin	Hair follicle	Mechanical strengthening of the hair follicle	(Tarcsa *et al.* 1997) (Steinert *et al.* 2003)
S100A3	Hair cuticle	Maturation of hair cuticle cells	(Kizawa *et al.* 2008)
Myelin basic protein	CNS (myelin sheaths)	CNS development and function	(Pritzker *et al.* 2000)

[3] This paragraph was partly published in: Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables. 2010. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol Rev* 1:34-54.

Citrullinated protein	**Localisation**	**Function**	**References**
Histones H2A, H3, H4	Various cells	Transcriptional regulation, formation of neutrophil extracellular traps	(Wang *et al.* 2004) (Cuthbert *et al.* 2004) (Hagiwara *et al.* 2005) (Yao *et al.* 2008) (Li *et al.* 2008) (Neeli *et al.* 2008) (Li *et al.* 2010a) (Li *et al.* 2010b)
Vimentin	Monocytes, Macrophages	? Apoptosis (*In vitro*: Unfolding of citrullinated proteins, diminished cytoskeletal stability)	(Inagaki *et al.* 1989) (Vossenaar *et al.* 2004b)
Chemokines (CXCL-8, -10)	Peripheral blood mononuclear cells	? Immune regulation (*In vitro*: citrullinated chemokines show altered receptor usage, proteolytic processing and biological activities)	(Proost *et al.* 2008) (Loos *et al.* 2008)
Fibrin(ogen)	Peripheral blood, tissue deposits	? Anti-inflammatory mechanism (*In vitro*: Inhibition of thrombin-catalysed fibrinogen polymerisation)	(Okumura *et al.* 2009) (Nakayama-Hamada *et al.* 2008)

Citrullinated proteins in disease[4]

A dysregulation of protein citrullination at the site of disease has been associated with inflammation and tissue injury in general, not just in RA but also in other forms of inflammatory arthritis, multiple sclerosis, glaucoma, psoriasis, myositis, Alzheimer's disease, and most recently in sporadic Creutzfeldt-Jakob disease (Table 1.4). It is unknown what causes the aberrant protein citrullination in these diseases. In inflammatory arthritis, including RA, it is assumed that inflammation and the resulting release of pro-inflammatory stimuli and increased cell death activates PAD in a calcium rich environment, which might account for the accumulation of citrullinated proteins seen at diseased sites. However, intracellular citrullinated proteins in inflammatory arthritides were also observed by a number of investigators (Baeten *et al.* 2001; Cantaert *et al.* 2005; De Rycke *et al.* 2005) and might result from a controlled increase of intracellular Ca^{2+} concentrations, for example in response to activation by chemokines.

Table 1.4 Citrullinated proteins in diseased human physiology

Disease	**Citrullinated proteins confirmed at diseased site**	**Localisation of citrullinated proteins**	**References**
Rheumatoid arthritis and other inflammatory arthritides	Fibrin(ogen), vimentin, collagen type II, α-enolase, fibronectin	Synovial joint	(Vossenaar *et al.* 2004c) (Makrygiannakis *et al.* 2006) (Chapuy-Regaud *et al.* 2005) (Chang *et al.* 2005) (Nakayama-Hamada *et al.* 2005) (De Rycke *et al.* 2005) (Foulquier *et al.* 2007) (Van Steendam *et al.* 2010)
Multiple sclerosis	Myelin basic protein	Brain	(Moscarello *et al.* 1994)
Glaucoma	Myelin basic protein	Optic nerve	(Bhattacharya *et al.* 2006)

[4] This paragraph was partly published in: Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables. 2010. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol Rev* 1:34-54.

Disease	Citrullinated proteins confirmed at diseased site	Localisation of citrullinated proteins	References
Psoriasis	Keratin (decreased levels of citrullination)	Epidermis (keratinocytes)	(Ishida-Yamamoto *et al.* 2000)
Myositis	?	Muscle	(Makrygiannakis *et al.* 2006)
Cancer (Adenocarcinoma)	Cytokeratin	Various	(Chang *et al.* 2006)
Alzheimer's	Vimentin, glial fibrillary acidic protein	Brain (hippocampus)	(Ishigami *et al.* 2005)
Creutzfeldt-Jakob disease	Vimentin, glial fibrillary acidic protein, enolase, aldolase A, myelin basic protein, phosphoglycerate kinase, cyclophilin A	Brain (frontal cortex)	(Jang *et al.* 2010)

Autoimmunity to citrullinated proteins in rheumatoid arthritis[5]

Since accumulation of citrullinated proteins occurs in a variety of conditions, it is now accepted that it is the autoantibody response rather than the presence of the citrullinated antigen that characterises RA. These autoantibodies are therefore most likely the result of an abnormal immune response to particular citrullinated proteins, where additional short- and long-term factors, such as the genetic background and environment, are required to induce and/or sustain antibody formation (Figure 1.4). This triad of factors - autoantibodies, genes and environment - is of course a gross simplification of the diverse and complex pathways

[5] This paragraph and Figure 1.4 were partly published in: Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables. 2010. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol Rev* 1:34-54.

that underlie the clinical manifestations seen in RA. Nevertheless, it serves as a useful template for addressing a fundamental issue; namely, what actually causes the disease.

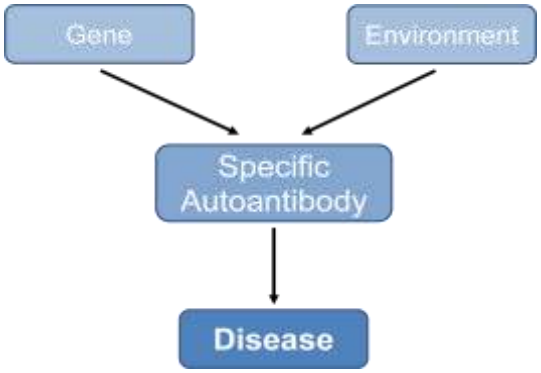


Figure 1.4 The gene/environment/autoimmunity triad. Schematic representation of the hypothesis for the involvement of genes and environment in the development of specific and pathogenic autoantibodies in chronic autoimmune diseases.

Rheumatoid factor

Historically, rheumatoid factor (RF), an antibody reactive with the Fc portion of IgG and elevated in up to 80% of patients with RA, has been the main serological marker for the diagnosis of RA, and is still used as one of the criteria for the classification of the disease (Table 1.2). However, RFs also occur in other autoimmune diseases, infections and in 5% of the healthy population, and thus may simply be a consequence of polyclonal B cell activation (Sutton *et al.* 2000).

Anti-citrullinated protein antibodies: Historical development[6]

Antibodies to citrullinated proteins as a more specific marker of RA were described over 40 years ago as the anti-perinuclear factor test; though at the time it was not known that it was detecting anti-citrullinated protein antibodies. In this test, human buccal mucosa cells were used as the substrate and antibodies reactive with granular structures around the nucleus were

[6] Text and figures from pages 32-39 were partly published in: Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables. 2010. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol Rev* 1:34-54.

detected by indirect immunofluorescence (Nienhuis *et al.* 1964). The assay had a high diagnostic specificity for RA but with variable sensitivity. This, combined with the cumbersome nature of the assay, precluded its use for routine diagnosis. An antibody reactive with the keratinised layer of rat oesophagus was subsequently reported as the so-called 'anti-keratin antibody' (Young *et al.* 1979) and correlated strongly with the anti-perinuclear factor. More recently, it was shown that both anti-perinuclear and anti-keratin antibodies reacted with mature filaggrin (Hoet *et al.* 1991; Simon *et al.* 1993; Sebbag *et al.* 1995), and a crucial discovery was that this reactivity was citrulline-specific (Schellekens *et al.* 1998; Girbal-Neuhauser *et al.* 1999).

Detection of anti-citrullinated protein antibodies for RA diagnostics today

Nowadays, enzymatic assays using cyclic citrullinated peptides (CCP) as surrogate target antigens are routinely used in the diagnosis of RA and form part of the revised RA classification criteria (Table 1.2). The first generation of CCP assay (CCP1) was based on citrulline-containing peptides from the arginine-rich filaggrin sequence, and had a diagnostic sensitivity of approximately 70% with a disease-specificity of 96% (Schellekens *et al.* 2000). Because filaggrin, a protein which is involved in epidermal differentiation and hydration, is found only in epithelial cells but not in the joint, it was concluded that citrullinated filaggrin acted as a surrogate antigen and that other citrullinated proteins were more likely to be driving the autoimmune response in RA. In order to improve the diagnostic sensitivity and specificity of the CCP1 test, the CCP2 assay was developed and is now widely used in diagnostic laboratories. The substrates for this test were selected from a large panel of randomly generated citrulline-containing peptides which were tested against RA and control serum, with the sequences giving the best discrimination in diagnostic sensitivity and specificity being adopted for a commercial assay. Using an enzyme-linked immunosorbent assay (ELISA), CCP2 antibodies have a diagnostic sensitivity that is at least equivalent to that of RF (up to 82%) and a specificity that is much higher (98%) (van Venrooij *et al.* 2008). CCP antibodies have been detected prior to the development of clinically apparent RA (Rantapaa-Dahlqvist *et al.* 2003) and are associated with more severe and erosive disease (reviewed by Zendman *et al.* 2006). There is also increasing evidence that anti-citrullinated protein antibodies contribute to RA pathogenesis, possibly by formation of immune complexes resulting in production of proinflammatory molecules, perpetuating the

inflammatory process and leading to tissue injury and erosion in humans (Clavel *et al.* 2008) and mice (Kuhn *et al.* 2006; Petkova *et al.* 2006) (also see Figure 1.5).

In spite of being a powerful diagnostic tool, the cyclic citrullinated peptides used in the CCP2 are of limited use for understanding the disease aetiology and pathogenesis of RA as they do not correspond to *in vivo* generated citrullinated proteins. However, four citrullinated proteins that are targeted by anti-citrullinated protein antibodies, and are present in the joint, are now well established: citrullinated fibrinogen/fibrin (Masson-Bessiere *et al.* 2001), vimentin (Vossenaar *et al.* 2004a), collagen type II (Burkhardt *et al.* 2005), and α-enolase (Kinloch *et al.* 2005), with further awaiting identification and characterisation (reviewed by Wegner *et al.* 2010a). Of particular relevance to this project are citrullinated fibrin(ogen) and α-enolase.

Citrullinated fibrin(ogen) as a target of anti-citrullinated protein antibodies

Fibrinogen is the precursor of fibrin, and citrullinated fibrin(ogen) is one of the best-established specific autoantigens in RA. Citrullinated fibrin(ogen) is abundantly expressed in inflammatory joints of humans (Masson-Bessiere *et al.* 2001; Vossenaar *et al.* 2004c; Chapuy-Regaud *et al.* 2005; Tilleman *et al.* 2005; Matsuo *et al.* 2006) and experimental animals (Vossenaar *et al.* 2003a). It was originally identified by its cross-reaction with anti-citrullinated filaggrin antibodies, which were affinity-purified from RA serum. By immunoblotting, these antibodies reacted mainly with the α- and β-chains of fibrin (Masson-Bessiere *et al.* 2001). Antibodies to whole citrullinated fibrinogen are equal, in terms of diagnostic sensitivity and specificity, to the CCP-assay (vander Cruyssen 2006). Fibrinogen-containing immune complexes were found in serum samples from about 50% of anti-CCP-positive RA patients and immunohistochemical staining demonstrated co-localisation of fibrinogen, immunoglobulin and complement component 3 (C3) in the pannus tissue. Evidence for citrullination of fibrinogen contained in immune complexes and pannus tissue was also shown (Zhao *et al.* 2008). The expression of citrullinated fibrin(ogen) in the joint, and the detection of disease-specific antibodies in the serum, makes immune complexes containing citrullinated fibrin(ogen) an attractive mechanism for mediating some of the immunopathology of RA. This was recently explored in *in vitro* model systems. Clavel *et al.* found that human monocyte-derived macrophages produced large amounts of TNF-α when incubated with immobilised immune complexes containing citrullinated fibrinogen, and that the effect was mediated by Fcγ receptor IIa (Clavel *et al.* 2008). These results were

supported and extended by Sokolove *et al.* (Sokolove *et al.* 2011), showing that immune complexes containing citrullinated fibrinogen induce TNF-α secretion by macrophages via dual engagement of Fcγ receptors and the innate immune toll-like receptor (TLR) 4, a well-known ligand for fibrinogen (Smiley *et al.* 2001; Kuhns *et al.* 2007). The levels of secreted TNF-α were significantly higher when citrullinated fibrinogen immune complexes were used, as compared to citrullinated or native fibrinogen alone, anti-fibrinogen antibody alone, or native fibrinogen-containing immune complexes, suggesting that the presence of anti-citrullinated fibrinogen antibodies mediates the pro-inflammatory response and that this is influenced by the specific antigen incorporated in the immune complex. The combined findings highlight how immune complexes, containing citrullinated fibrinogen, might be important in the pathogenesis of RA by driving inflammation via Fcγ, toll-like and complement receptors, resulting in inflammatory cytokine release (Figure 1.5).

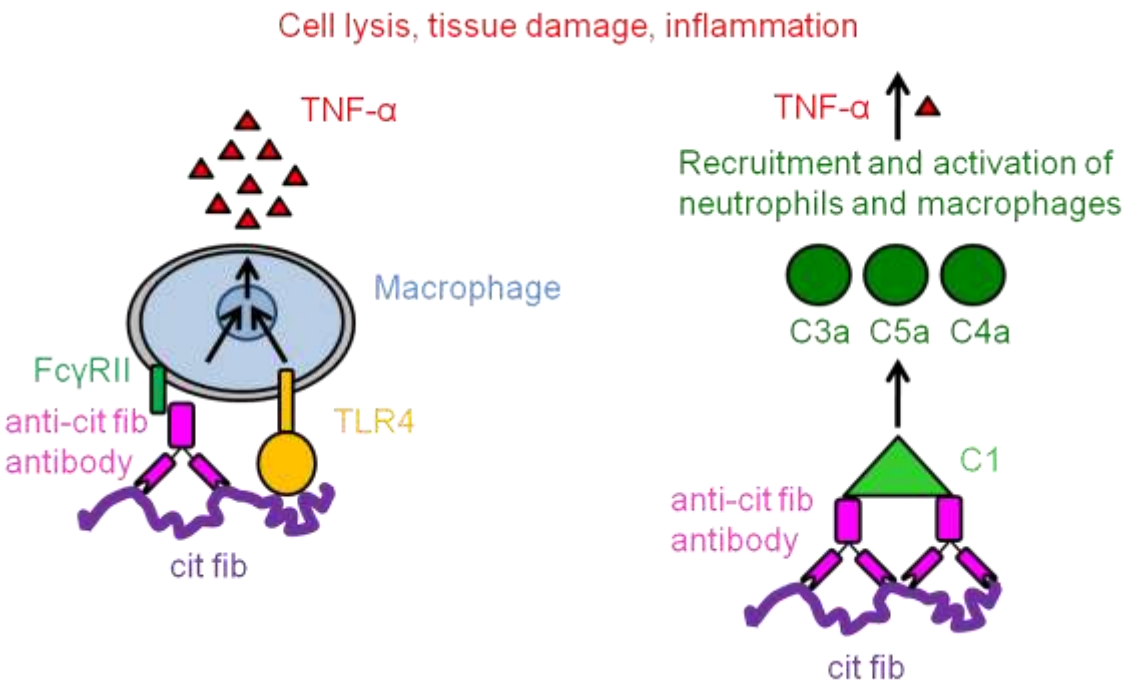


Figure 1.5 Schematic representation of potential pathways contributing to the pathogenic role of immune complexes containing citrullinated fibrinogen in the arthritic joint of RA patients. Antibodies specific for citrullinated fibrin(ogen) form immune complexes with citrullinated fibrin(ogen) deposits in the joint tissue. The immune complexes activate pro-inflammatory effector mechanisms by activating macrophages via Fcγ receptor (FcγR) II and toll-like receptor (TLR) 4, leading to the secretion of tumour necrosis factor-α (TNF-α). The immune complexes can also bind complement component 1 (C1), which eventually results in the release of C3a, C5a and C4a, leading to the recruitment and activation of neutrophils, macrophages and other pro-inflammatory components, and secretion of the cytokine TNF-α, amongst others, by these cells. The combined activity of all pathways results in cell lysis, tissue damage, and amplification of the inflammatory process. This figure was generated by myself using Microsoft Powerpoint.

Citrullinated α-enolase as a target of anti-citrullinated protein antibodies

Alpha-enolase was first described by Saulot and colleagues as a candidate autoantigen reactive with 25% of early RA sera, using human placenta and epithelial cell line extracts (Saulot *et al.* 2002). However, using recombinant α-enolase as the substrate, only eight of the 36 previously positive samples showed reactivity, strongly suggesting the presence of post-translational modifications in α-enolase from the placental/cell line tissue extracts. Work in our laboratory described the citrullinated form of α-enolase as a polypeptide in deiminated lysates of HL-60 cells reactive with serum from patients with RA (Kinloch *et al.* 2005). Alpha-enolase was found to be abundantly expressed in RA synovial membrane and antibodies reacting only with the citrullinated form of the molecule were specific for RA (sensitivity approximately 40%), compared to healthy serum. More recently, using cyclic citrullinated α-enolase peptides covering 15 out of 17 arginine residues present in human α-enolase, our laboratory identified a peptide containing the immunodominant B cell epitope (Lundberg *et al.* 2008). This region comprised amino acids 5-21 of α-enolase, with arginine-9 and arginine-15 replaced by citrulline (^{5}KIHAcitEIFDScitGNPTVE^{21}). Carboxy- and amino-terminal cysteine residues were added to enable the peptide to adopt a cyclic conformation, which increases antibody reactivity. The resultant peptide is referred to as citrullinated α-enolase peptide-1 (CEP-1). Using mass spectrometry, it had already been confirmed that both arginine-9 and arginine-15 can be citrullinated *in vitro* (Kinloch *et al.* 2005). A three-dimensional representation of the human α-enolase protein shows that the CEP-1 peptide is present near the amino-terminus on a pole of the molecule, comprising the first and part of the second beta strand, and is surface-accessible (Figure 1.6). Arginine-15, which is the crucial amino acid in CEP-1, is located in the loop region between the first and the second beta strand. Taken together, this could explain why CEP-1 might be a dominant epitope and why a cyclic version of the peptide may be required for the detection of antibodies. In our studies of over 300 serum samples, antibodies to CEP-1 were found in 37% of RA patients, 3% of disease controls and 2% of healthy controls (Lundberg *et al.* 2008). Similar frequencies of anti-CEP-1 antibodies (41%) were found in a large cohort of Swedish patients with RA (Snir *et al.* 2009). Approximately half of the anti-CCP-positive patients with RA also have antibodies to CEP-1 (Lundberg *et al.* 2008).

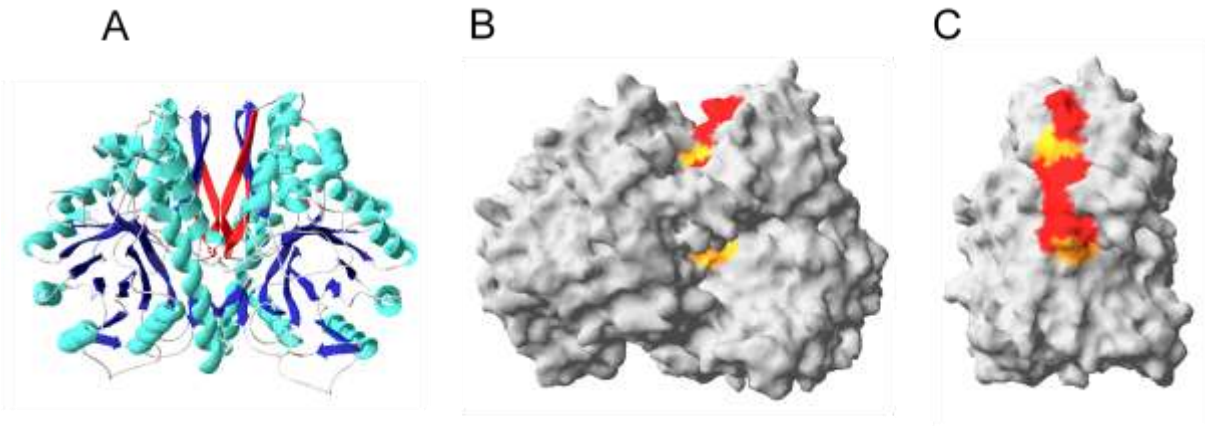


Figure 1.6 Molecular representation of the human α-enolase dimer (A, B) and monomer (C). **A:** Ribbon diagramm. The immunodominant B cell epitope (CEP-1) is highlighted in red. Secondary structure elements (helices, light blue; strands, dark blue; loops, grey) are labelled. **B:** Molecular surface model. The CEP-1 epitope in the dimeric molecule is highlighted in red, with the two arginines in orange. **C:** Molecular surface model of the monomer, looking onto the dimer interface. Colour coding as in B. This figure was generated by myself using the software DeepView v4.0 (http://spdbv.vital-it.ch/) and is based on the X-ray crystal structure published by Kang and co-workers (Kang *et al.* 2008). This figure was published in (Wegner *et al.* 2010a).

Similar to anti-CCP antibodies, expression of anti-CEP-1 antibodies precedes the onset of clinical RA by several years. In a study of 15 pre-RA serum samples selected on the basis of HLA-DRB1 SE alleles and anti-CCP antibodies being present, antibodies to a range of citrullinated enolase peptides were detected in 9 samples, 5 of which were directed specifically towards CEP-1 (Lundberg 2009). This finding suggests that the initial breakdown of immunological tolerance to citrullinated proteins, including α-enolase, might occur as a consequence of an inflammatory event outside the joint. Furthermore, the presence of IgA anti-CCP (73.9%) (dos Anjos *et al.* 2009) and IgA anti-CEP-1 (16%) (Snir *et al.* 2009) indicates the involvement of mucosal surfaces.

Similar to other established antigens in human autoimmune diseases, enolase is a highly conserved and ubiquitous protein (reviewed by Wegner *et al.* 2009). Mammals express three enolase isotypes (α, β and γ), which form homodimers or heterodimers. Alpha-enolase is the most ubiquitous isotype and is expressed in a variety of tissues, while β-enolase is exclusively found in muscle tissues and γ-enolase in neuronal and neuroendocrine tissues. Enolase is a multifunctional protein: it catalyses the penultimate step in glycolysis, but also plays a role in various processes, such as hypoxia tolerance, growth control, and in

fibrinolysis, extracellular matrix remodelling and cell migration. This is due to its ability to bind plasminogen on the cell surface of various eukaryotic cells (Miles *et al.* 1991; Redlitz *et al.* 1995; Lopez-Alemany *et al.* 2003; Wygrecka *et al.* 2009) and a number of bacteria (Pancholi *et al.* 1998; Bergmann *et al.* 2001; Ge *et al.* 2004; Yavlovich *et al.* 2007), fungi (Fox *et al.* 2001), and parasites (Jolodar *et al.* 2003; Vanegas *et al.* 2007; Mundodi *et al.* 2008). Using human cell lines and a monoclonal antibody directed towards the plasminogen-binding region of α-enolase, Lopez-Alemany *et al.* showed that α-enolase, despite being responsible for only around 15-20% of the total plasminogen binding, accounts for up to 90% of plasminogen's activation activity on the cell surface (Lopez-Alemany *et al.* 2003). It was also reported that plasminogen upregulates α-enolase gene expression in fibroblasts and peripheral blood mononuclear cells (Sousa *et al.* 2005), supporting an important physiological role of α-enolase in the plasminogen system. A recent study by Wygrecka and colleagues demonstrated a role for cell-surface α-enolase in inflammatory cell recruitment. They showed that cell-surface expression of α-enolase is upregulated in monocytes upon LPS stimulation and enhances plasmin generation, monocyte cell migration and matrix invasion in inflamed lungs (Wygrecka *et al.* 2009).

Previous work in the laboratory demonstrated increased expression of α-enolase in synovial membrane from patients with RA compared to osteoarthritis (OA) (Kinloch *et al.* 2005), and upregulation of extracellular α-enolase in cell-free synovial fluid in inflammatory arthritis compared to controls (Kinloch *et al.* 2008). In a synovial fluid sample from a patient with RA, citrullination of α-enolase was indirectly demonstrated through a shift of the molecule (using an antibody raised against CEP-1) towards a more acidic isoelectric point (pI) in the two-dimensional electrophoretic migration (Kinloch *et al.* 2008). This relative 'acidic shift' towards the anode is a characteristic of protein citrullination due to the resulting decreased net positive charge. These data provide some evidence that citrullinated α-enolase is involved in the pathogenesis of RA, and an animal model is currently under investigation.

Risk factors for developing anti-citrullinated protein antibodies

It is now beginning to emerge that two of the best described and strongest risk factors for RA, smoking and the presence of shared epitope alleles, are in fact specifically associated with CCP-positive RA (van der Helm-van Mil *et al.* 2006; Michou *et al.* 2008; reviewed by Baka

et al. 2009), suggesting a biological gene-environment interaction. More recently, a case–control analysis of 1,000 patients with RA and 872 healthy controls from the Karolinska Institute, Stockholm, and our laboratory found that smoking and the shared epitope is primarily associated with the anti-CEP-1 positive subset of CCP-positive RA. The subset of patients that (i) smoked, (ii) expressed shared epitope alleles and (iii) the susceptibility allele of *620W PTPN22* (a gene encoding a tyrosine phosphatase involved in T and B cell signaling (Kallberg *et al.* 2007)) and (iv) had antibodies to both CCP and CEP-1, demonstrated an OR of 39.0 (95% CI 15.6–98.0) for the development of RA, compared to an odds ratio (OR) of 4.3 (95% CI 2.2–8.6) in the corresponding subset that was negative for anti-CEP-1 (Mahdi *et al.* 2009). These data strongly suggest that CEP-1 is a peptide from a specific autoantigen which links smoking to the shared epitope in the development of RA.

Of further relevance is the finding that polymorphisms in the human *PADI4* gene have been associated with RA in Asian populations (Suzuki *et al.* 2003; Ikari *et al.* 2005; Kang *et al.* 2006; Takata *et al.* 2008; Bang *et al.* 2010), and weakly in some Caucasian cohorts (Plenge *et al.* 2005; Iwamoto *et al.* 2006; Lee *et al.* 2007; Gandjbakhch *et al.* 2009; Hoppe *et al.* 2009), but not in most other Caucasian cohorts (Barton *et al.* 2004; Caponi *et al.* 2005; Harney *et al.* 2005; Martinez *et al.* 2005; Burr *et al.* 2009). In the original study using a Japanese cohort, four single nucleotide polymorphisms (SNPs) (padi4_89: Serine-55 to Glycine; padi4_90: Alanine-82 to Valine; padi4_92: Alanine-112 to Glycine; padi4_104: Leucine-117, no amino acid residue replacement) were described to be associated with RA (OR 1.97 over the entire gene), and individuals harbouring these haplotypes showed higher anti-CCP antibody levels than controls ($p = 0.038$) (Suzuki *et al.* 2003). A large Korean case control study (1,313 patients and 1,004 controls) confirmed that these SNPs are associated with RA susceptibility, however regardless of anti-CCP status (Bang *et al.* 2010). Further experiments using RNA containing the two most common haplotypes of RA-susceptible SNPs showed that it is degraded more slowly by HL-60 cell supernatants, compared to non-susceptible RNA, suggesting that these nucleotide substitutions make *PAD4* mRNA more resistant to degradation, possibly resulting in increased levels of PAD4 enzyme and thereby citrullination (Suzuki *et al.* 2003). On the protein level, the resulting amino acid replacements are located in the N-terminal domain far from the active site and are therefore unlikely to affect catalytic activity directly, but might influence interaction with other proteins. A recent study determined the crystallographic structure of the mutant PAD4 enzyme harbouring all three amino acid replacement, confirming that the structure of the catalytic domain is identical to

that of the wildtype enzyme, and that the enzymatic activity is not affected by the mutations (Horikoshi *et al.* 2011).

Inhibition of protein citrullination as a possible therapeutical strategy in RA

Blockade of human PADs has been suggested by several investigators as a powerful novel therapy by inhibiting the generation of the antigens that sustain autoimmunity in RA. The approach employed for developing potential PAD inhibitors by most investigators is targeting the crucial cysteine (cysteine-645 in PAD4) in the active site (Arita *et al.* 2004).

Stone and colleagues investigated 2-chloroacetamidine as a potential inhibitor (Stone *et al.* 2005). 2-chloroacetamidine was found to inhibit both human PAD4 and dimethylarginase (also called dimethylarginine dimethylaminohydrolase; DDAH) from *Pseudomonas aeruginosa*. DDAH hydrolyses methylated arginines into citrulline and the corresponding alkylamine. In humans, DDAH enzymes regulate nitric oxide production by hydrolysing the endogenous inhibitors of nitric oxide synthase, L-NMMA (N^G-monomethyl-L-arginine) and ADMA (asymmetric dimethylarginine) (Vallance *et al.* 2002). Both DDAH and PAD4 belong to the GME superfamily and share a similar catalytic mechanism involving an active-site Cys residue (Shirai *et al.* 2006). 2-chloroacetamidine (Figure 1.7) consists of an amidinium group similar to the natural substrates of PAD, but also a chloromethylene moiety, proposed to bind covalently to the active site cysteine.

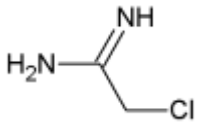


Figure 1.7 Chemical structure of 2-chloroacetamidine (IUPAC name: 2-chloroethanimidamide)

The reported K_i (inhibitor dissociation constant) was relatively high at 20 mM for PAD4, and 3.1 mM for DDAH, indicating weaker inhibitor potency of 2-chloroacetamidine with PAD4 compared to DDAH. Addition of excess substrate (BAEE, 10 mM) to the enzyme-inhibitor mix decreased the inactivation rate, indicating that 2-chloroacetamidine acts at the active site. Using DDAH as a model, the authors showed by employing mass spectrometry that the inhibitor binds covalently to a peptide containing the active site cysteine, and suggested that it inhibits PAD4 via a similar mechanism due to the analogous active site residues. The

authors concluded that 2-chloroacetamidine should be used as a chemical scaffold to develop bioavailable inhibitors with higher affinity and selectivity. Two such compounds were shown to be F-amidine [*N*-α-benzoyl- N^5-(2-fluoro-1-iminoethyl)-l-ornithine amide] and Cl-amidine [*N*-α-benzoyl-N^5-(2-chloro-1-iminoethyl)-l-ornithine amide], which inhibit PAD with an IC_{50}[7] of 21.6 μM (F-amidine) (Luo *et al.* 2006b) and 5.9 μM (Cl-amidine) (Luo *et al.* 2006a). Both molecules combine the basic structure of natural PAD substrates (peptidylarginine side chain), combined with a haloacetamidine warhead that is also found in 2-chloroacetamidine, and a benzoyl moiety (Figure 1.8). They act as irreversible inactivators of PAD4 by covalently modifying the active-site cysteine (Knuckley *et al.* 2010b). Their increased potency over 2-chloroacetamidine is likely to be due to their closer similarity of natural and synthetic PAD substrates such benzoyl-arginine ethylester (BAA) (Figure 1.8), targeting the compounds to the active site.

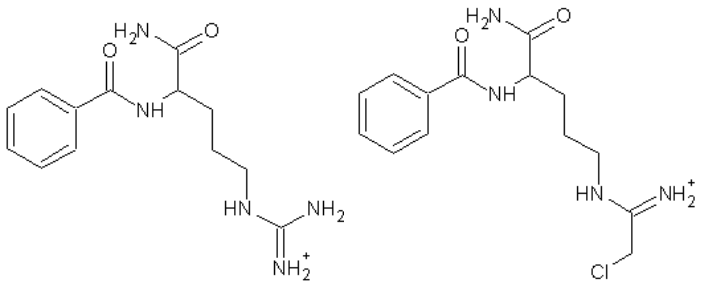


Figure 1.8 Chemical structures of the synthetic peptidylarginine deiminase substrate BAA (benzoyl-arginine ethylester) and the inhibitor Cl-amidine (*N*-α-benzoyl-N^5-(2-chloro-1-iminoethyl)-l-ornithine amide)

Inactivation by F-amidine and Cl-amidine is calcium-dependent. Both compounds preferentially inactivate the calcium-bound form >10-fold (Luo *et al.* 2006a; Luo *et al.* 2006b). *In vitro* studies with PAD4 have shown that calcium binding leads to a conformational change that moves Cys-645 and His-471 into a position that is competent for catalysis (Arita *et al.* 2004), and presumably reactive with F-amidine and Cl-amidine. This is of therapeutic importance as these compounds would be expected to inhibit PAD4 in its activated state only at sites of inflammatory activity such as the synovium and not the inactive enzyme at other sites in the body, limiting toxicity. *Ex vivo* studies with F-amidine and Cl-amidine, using a cell line and an assay measuring PAD4-mediated citrullination of a

[7] IC_{50} represents the concentration of an agent required to inhibit a reaction by 50%.

nuclear protein and the resulting enhancement in binding to another protein, indicated that these inhibitors can be taken up by cells (Luo *et al.* 2006a; Luo *et al.* 2006b). Another study investigating PAD4 inhibition in cancer therapy showed that treatment of osteosarcoma U2OS cells with 200 μM Cl-amidine for 3 days halted cell proliferation and lowered levels of citrullinated histone H3 citrullination by ~50%, compared to untreated cells (Li *et al.* 2008; Li *et al.* 2010b). Van Willis and colleagues published a paper demonstrating the potential of these agents as useful therapeutic inhibitors in a mouse model of arthritis (Willis *et al.* 2011). Cl-amidine treatment in the CIA mouse model of arthritis (daily treatment on days 0-35 after immunisation with collagen II) significantly reduced clinical disease activity scores by 55%, 53%, and 42% in the 50, 10, and 1 mg/kg per day groups, respectively. Histological severity scores, serum and synovial citrullination levels and C3 deposition in synovial tissue and cartilage paralleled the decrease in disease activity. In addition, mice receiving Cl-amidine (50 mg/kg) showed a reduced autoantibody response to native synovial antigens and citrullinated filaggrin peptides. In summary, in this model of arthritis, treatment with a PAD inhibitor does not prevent but at least ameliorates disease activity and synovial inflammation, probably through limiting the generation of anti-citrullinated protein and other autoantibodies, which in turn reduces pro-inflammatory effector mechanisms and thereby disease severity. A more indirect effect on inflammatory signalling cannot be excluded however, for example through the inhibition of chemokine citrullination.

The combined data suggest that Cl-amidine and related compounds may indeed represent a novel class of RA therapeutics that specifically target citrullination. A challenge will be to define not only PAD-specific but further isozyme-specific inhibitors, as the inhibition of all PAD activity with equal efficiency in the long-term might lead to severe side effects associated with disturbed skin and hair follicle cell differentiation, apoptosis, antimicrobial response etc. The study described above did not mention any side effects of Cl-amidine treatment on mice despite the fact that Cl-amidine is an unselective inhibitor of PADs, although the potential long-term effects cannot be deducted. A recent study determined the IC_{50} levels of Cl-and F-amidine for PAD1 and 3 as 0.8 μM (Cl-amidine+PAD1), 6.2 μM (Cl-amidine+PAD3), 29.5 μM (F-amidine+PAD1), and ~350 μM (F-amidine+PAD3) (Knuckley *et al.* 2010a). Comparing these to the IC_{50} values for PAD4 described earlier (21.6 μM with F-amidine and 5.9 μM for Cl-amidine), it becomes apparent that neither compound is isozyme-specific; Cl-amidine would inhibit PAD1, 3, and 4 while Fl-amidine would show a

preference for PAD1 and 4, if all enzymes were available at the site of inhibition at the same concentration.

On the basis of the therapeutic use of DMARDs and antibiotics (O'Dell *et al.* 1999) in RA, Knuckley and colleagues screened a range of DMARDs (5-aminosalicylic acid, leflunomide, methotrexate), tetracycline derivatives (minocycline, doxycycline, tetracycline, chlortetracycline), and other antibiotics (streptomycin, clindamycin, sulfamethoxazole, trimethoprim, sulfapyridine, azithromycin) for their potential to inhibit PAD4 activity (Knuckley *et al.* 2008). Chlortetracycline was identified as the most potent inhibitor (IC_{50} 100 μM) and was suggested to bind to a region distal from the active site (Knuckley *et al.* 2008). Minocycline, tetracycline and doxycycline demonstrated IC_{50} values of 620, 780 and 860 μM, respectively. Streptomycin, an aminoglycoside antibiotic, was tested because of its two guanidinium groups that could act as substrate inhibitors of PAD4. Streptomycin was not citrullinated, but found to inhibit PAD4, though with a lower potency (IC_{50} = approximately 1.8 mM) compared to tetracyclines, and was suggested to bind within or in close proximity to the active site. All other compounds were ineffective or only weakly inhibitory (IC_{50} between 2.4 and >10 mM). The data suggest that tetracyclines are indeed PAD4 inhibitors, but their relatively weak inhibitory potency suggests that the efficacy in RA is probably not only due to PAD inhibition mainly to other mechanism, such as inhibition of metalloproteinases, as has been demonstrated for a range of tetracyclines (Sorsa *et al.* 1998). Nevertheless, they could provide a valuable scaffold for engineering inhibitors with greater potency and selectivity.

Paclitaxel, a chemotherapeutic agent that interferes with microtubule function in cell mitosis, migration, chemotaxis, and intracellular transport, was shown to weakly inhibit PAD from bovine brain in the millimolar range (IC_{50} ~5 mM) (Pritzker *et al.* 1998). It prevented the induction of collagen-induced arthritis (CIA) and caused significant regression of existing CIA (Brahn *et al.* 1994), although it was not investigated what mediated this effect. An open-label multicenter phase II study of paclitaxel in patients with RA was completed in July 2008, although results of this are still pending[8].

[8] http://clinicaltrials.gov (ClinicalTrials.gov Identifier: NCT00055133)

Porphyromonas gingivalis, periodontitis and rheumatoid arthritis

Summary of previous work and state of knowledge leading to this project

The identification of citrullinated human α-enolase, and specifically the CEP-1 epitope within it, as a major target for autoantibodies to citrullinated proteins in RA, raised the question as to which factors trigger this specific response. As enolase is an evolutionary highly conserved enzyme and present in all domains of life, including bacteria, the investigation of the epitope mimicry hypothesis was prompted. According to the epitope mimicry hypothesis as it is commonly interpreted, an autoimmune response is induced by an infectious agent and initially directed against its microbial antigens. Owing to the structural resemblance between microbial and particular host antigens, the autoimmune response not only attacks the invading pathogen but also host tissue, which, in susceptible individuals, may tip the balance of immunological response versus immune tolerance and subsequently lead to autoimmune disease (discussed by Davies 1997; Oldstone 1998).

P. gingivalis, a major causative agent in periodontitis (PD), seemed a promising candidate as it was reported to expresses a unique bacterial PAD-like enzyme (McGraw *et al.* 1999), and because PD had been associated with RA in a number of epidemiologic studies (reviewed by de Pablo *et al.* 2009). Further, an amino acid sequence alignment revealed that full-length human α-enolase and *P. gingivalis* enolase are 50.9% identical overall (Figure 1.9), and share 82% sequence identity at the region corresponding to CEP-1 (highlighted in Figure 1.9). In addition, the nine amino acids at positions 13–21, which surround the crucial citrulline residue at position 15 in terms of antibody reactivity (Lundberg *et al.* 2008), are 100% identical. The role of enolase-specific autoimmunity in RA, evolutionary conservation of enolase and presence of a PAD-like enzyme in *P. gingivalis* therefore led to the initial hypothesis that *P. gingivalis* is involved in the aetiology of RA via citrullination of *P. gingivalis* enolase by the bacterial PAD enzyme, the triggering of antibodies to bacterial citrullinated enolase, which then cross-react with citrullinated human α-enolase in the joint. Of note, the citrullination link between RA and *P. gingivalis* had also previously been suggested in an opinion article by Rosenstein (Rosenstein *et al.* 2004), however without a focus on specific mediating molecules.

Figure 1.9 Amino acid sequence alignment of human α-enolase (ENOA_HUMAN, SwissProt ID P06733) and *P. gingivalis* enolase (ENO_PORGI, SwissProt ID Q7MTV8). Identical amino acids are indicated by vertical lines, similar amino acids are indicated by colons and dots. The CEP-1 sequence is highlighted in colour. Identical amino acids are blue, non-identical amino acids are red. The two arginine residues (Arg-9 and Arg-15) which are substituted by citrulline in CEP-1 are underlined. The alignment was created using the EMBOSS program[9] (Matrix: Blosom62, Gap penalty 10.0, Extend penalty 0.5).

This initial hypothesis was investigated in our laboratory prior to the start of this project: the antibody response in RA patients to the *P. gingivalis* version of the CEP-1 peptide was measured by ELISA, and demonstrated to correlate strongly with that to human CEP-1 (r^2 = 0.8, P <0.0001) (Lundberg *et al.* 2008). Antibody cross-reactivity between whole citrullinated *P. gingivalis* enolase and citrullinated human α-enolase was then tested by immunoblot using CEP-1-affinity-purified antibodies from a patient with RA, which cross-reacted with citrullinated *P. gingivalis* enolase (Lundberg *et al.* 2008).

[9] http://www.ebi.ac.uk/Tools/emboss/align/

The hypothesis that evolved from these initial experiments and formed the basis for this project was that infection with *P. gingivalis* reduces tolerance to citrullinated antigens, and that one antigen in particular, citrullinated α-enolase, is central to the initiation of the pathogenic pathway leading to RA (Figure 1.10).

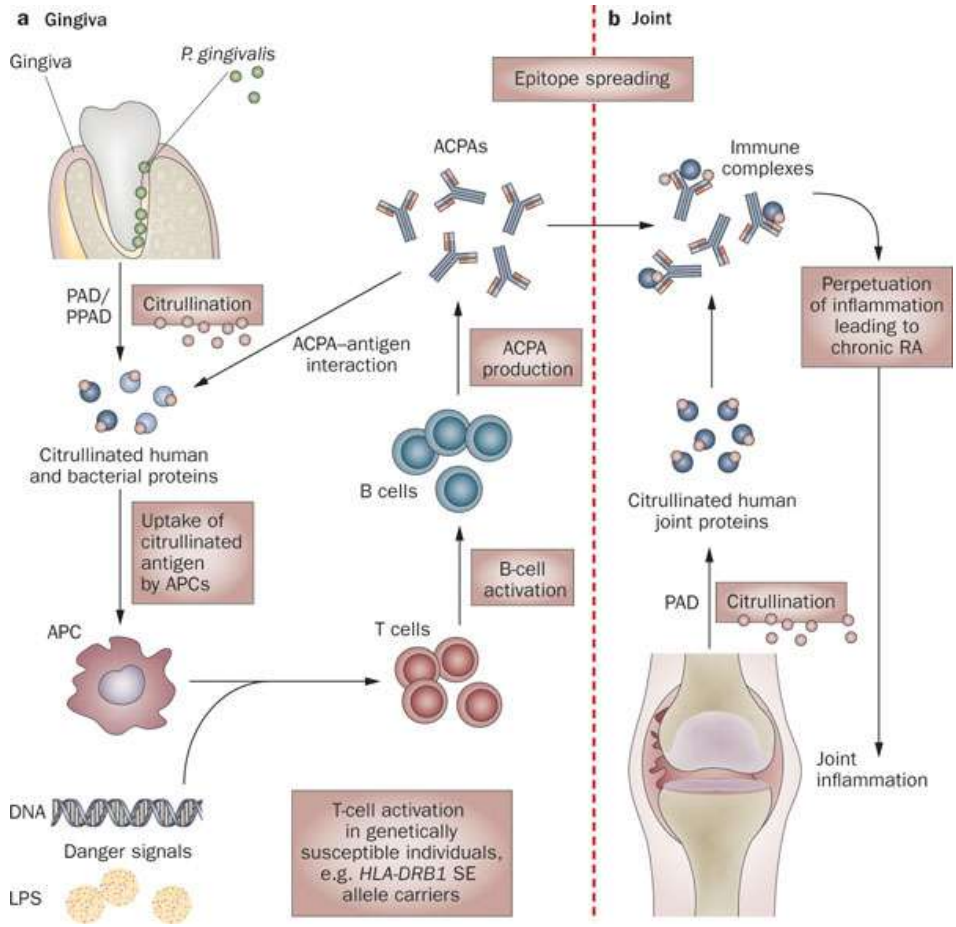


Figure 1.10 Schematic illustration of the aetiological hypothesis for *P. gingivalis* involvement in RA. a) Infection by *P. gingivalis* leads to citrullination of bacterial proteins (i.e. *P. gingivalis* enolase) by PPAD in the gingiva. In the presence of danger signals, such as LPS and DNA, pathogenic T cells are activated by APCs, which present citrullinated antigen in the context of the HLA-DRB1 shared epitope. T cell-mediated activation of pathogenic B cells results in the production of antibodies specific for citrullinated proteins generated in the gingiva. b) A second inflammatory event occurs in the joint, leading to citrullination of joint proteins (i.e. human α-enolase) by human PADs. Anti-citrullinated protein antibodies initially formed as a result of gingival infection and citrullination cross-react with citrullinated joint antigens, through molecular mimicry and epitope spreading, resulting in the formation of immune complexes in the joint and activation of pro-inflammatory effector mechanisms. The perpetuation of the inflammatory process eventually causes chronic RA. Abbreviations: ACPA, anti-citrullinated protein/peptide antibody; APC, antigen presenting cell; LPS, lipopolysaccharide; *P. gingivalis*, *Porphyromonas gingivalis*; PAD, peptidylarginine deiminases; PPAD, *P. gingivalis* PAD; RA, rheumatoid arthritis; SE, shared epitope. Figure was published in a review co-authored by myself (Lundberg *et al.* 2010).

Porphyromonas gingivalis - the organism and its role in human disease

P. gingivalis is an interesting organism that has been linked to RA not only through the expression of its bacterial PAD-like enzyme but also through epidemiological and serological studies and shared risk factors. In the following sections, the relevant literature is summarised to highlight its basic characteristics, role in periodontitis and links to RA.

Characteristics of Porphyromonas gingivalis

P. gingivalis is a gram-negative, anaerobic bacterium. It grows on blood-containing culture medium forming small glossy colonies of 1-2 mm diameter, with a varying degree of darkening because of the production of a heme-derived pigment (Figure 1.11).

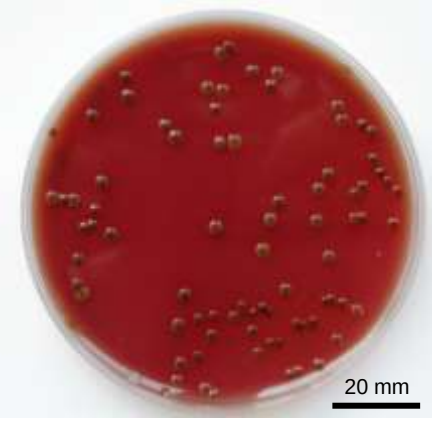


Figure 1.11 *P. gingivalis* colonies (strain W83) growing on sheep blood agar. Picture courtesy of Katarzyna Maresz, Jagiellonian University, Krakow, Poland.

The organism is present, together with other oral bacteria, as a biofilm in the gingival crevice and intracellularly in oral epithelial cells (Rudney *et al.* 2001; Colombo *et al.* 2007). Colonisation of the subgingival region is facilitated by the ability of *P. gingivalis* to adhere to available substrates such as adsorbed salivary molecules, matrix proteins, epithelial cells, and bacteria that are already established as a biofilm on tooth and epithelial surfaces (Lamont *et al.* 2000). *P. gingivalis* produces an array of virulence factors with the capacity to colonise the host, obtain nutrients from the host, perturb host defence mechanisms and disintegrate periodontal tissue and bone. These include in particular extracellular cysteine proteases called gingipains, a unique lipopolysaccharide (LPS), and adhesins such as fimbriae and hemagglutinins (Holt *et al.* 1999).

Gingipains are of particular relevance to this project, as at the start of this work it was known that they share physical properties with PPAD (they were repeatedly co-purified from *P. gingivalis* cell extracts) (Hayashi *et al.* 1993; McGraw *et al.* 1999), that they are active under the same conditions as PPAD (citrullinating activity was detected by coincidence when assaying for gingipain activity) (Hayashi *et al.* 1993; McGraw *et al.* 1999) and that they target arginine residues, raising the possibility of concerted activity. Gingipains are major contributors to the pathogenic potential of *P. gingivalis*, playing a role in all stages of destructive periodontitis: adherence and colonisation, nutrient acquisition, neutralisation of host defences, manipulation of the inflammatory response, tissue destruction, invasion and dissemination (reviewed by Curtis *et al.* 2001; Guo *et al.* 2010). They are potent proteases and cleave various proteins/peptides after either arginine residues (i.e., arginine gingipain A and B [RgpA and RgpB]) or lysine residues (i.e., lysine gingipain [Kgp] (Curtis *et al.* 1999)), resulting in peptides with carboxy-terminal arginines or lysines. All gingipains have a catalytic protease domain and an immunoglobulin-like domain. RgpA and Kgp further contain several hemagglutinin-adhesin domains. Gingipains are subjected to extensive post-translational proteolytic processing and glycosylation on their way to the cell surface, where they are attached to the outer membrane via sugar links or are secreted (reviewed by Potempa *et al.* 2003).

Periodontitis - classification and pathogenesis

The prevalence of severe generalised PD ranges between 4-15% (AAOP 1996; Borrell *et al.* 2005), although this varies depending on the definition criteria of PD and the applied examination protocol. Currently, two broad types of PD are distinguished: aggressive and chronic PD (Armitage 1999). Unlike aggressive PD, chronic PD is traditionally considered a slowly progressing disease; however some patients may also have short periods of rapid progression. Both chronic and aggressive PD can be further subclassified based on extent and severity. The extent is defined by the percentage of potential sites involved, resulting in localised (≤30% of sites involved) or generalised (>30% of sites involved) PD. The severity is characterised on the basis of the clinical attachment loss (CAL), which is a cumulative measure of the destruction of the tooth-supporting connective tissue and alveolar bone, measured as the distance between the cement-enamel junction and the bottom of the periodontal pocket. Slight PD is defined as 1-2 mm CAL, moderate PD as 3-4 mm CAL, and

severe PD as ≥5 mm CAL. Not all studies adhere to these classification guidelines though, often using other or additional measures such as gingival bleeding on probing, periodontal pocket probing depth, periodontal bone loss and tooth mobility, all of which are increased in PD. Confusingly, in the older literature various types of PD are treated as separate diseases that are now known as subcategories of chronic or aggressive PD, or both. For example, chronic PD now includes the old terms 'adult' and sometimes 'generalised juvenile' or 'rapidly progressing PD'. Aggressive PD may include 'early-onset', 'juvenile', and also 'rapidly progressing' PD (Armitage 1999).

PD is characterised by inflammation of the gums, connective tissue breakdown and alveolar bone destruction. The initial periodontal lesion usually begins with a self-sustaining inflammation of the periodontium as a result of plaque accumulation, genetic and environmental risk factors, and then progresses into a destructive disease, accompanied by the release or activation of complement, TNF-α, IL1-β, IL-6, IL-8, IL-17, prostaglandins E_2, and metalloproteinases (Gemmell *et al.* 1997). Persistent, uncontrolled inflammation leads to an accumulation of lymphocytes and monocytes and release of more degradative and proinflammatory molecules, resulting in soft tissue degradation and erosion of adjacent teeth and bone, and eventually in loss of teeth, if left untreated (Ohlrich *et al.* 2009).

The role of Porphyromonas gingivalis in periodontitis

P. gingivalis is a major aetiological agent in the initiation and progression of PD (Lamont *et al.* 1998). *P. gingivalis* is directly associated with PD: Higher numbers of *P. gingivalis* are found in subgingival plaque taken from diseased sites than those from healthy sites (Dzink *et al.* 1988), and *P. gingivalis* is prevalent more often in PD patients than in controls (OR 12.3) (van Winkelhoff *et al.* 2002). Using a PCR-based protocol, *P. gingivalis* could be detected in 80-95% of periodontitis patients and only 10-30% of healthy subjects (Griffen *et al.* 1998; Takeuchi *et al.* 2001; Yang *et al.* 2004). Implantation of *P. gingivalis* into the periodontal tissue of monkeys produced symptoms similar to human periodontitis (Holt *et al.* 1988). Other organisms often associated with periodontitis but with a less evident or unclear aetiological role are *Tannerella forsythensis*, *Aggregatibacter actinomycetemcomitans*, *Fusobacterium nucleatum*, *Prevotella intermedia*, *Treponema dendicola*, *Campylobacter*

rectus, *Eubacterium nodatum*, *Peptostreptococcus micros*, and *Staphylococcus intermedius* (reviewed by Lovegrove 2004).

In recent years it has emerged that PD is a complex disease, with a combination of individual susceptibility factors and the microbial ecosystem contributing to the occurrence and severity of disease. Genetic influences account for around 50% of the total risk for developing chronic periodontitis (Michalowicz *et al.* 2000). Gene polymorphisms associated with periodontitis were reported for a number of genes encoding cytokines, Fc receptors, MHC class II alleles, each of which confers a relatively small risk, and no consensus has been reached to date (reviewed by Kinane *et al.* 2005; and Yoshie *et al.* 2007).

Smoking has also been linked to PD with an odds ratio in the order of 2 to 6 (reviewed in Kinane *et al.* 2000). There is a dose-response relationship between the number of cigarettes smoked per day and the odds of developing disease (Tomar *et al.* 2000). Smokers have more severe PD, as measured by the degree of alveolar bone erosion and/or loss. In a large study of 1,361 subjects, smokers were at greater risk for severe bone loss than non-smokers, with odds ratios of 3.25 and 7.28 for light and heavy smokers, respectively (Grossi *et al.* 1995). Levels of cotinine, the principle metabolite of nicotine, have been shown to correlate directly with periodontal destruction (Yamamoto *et al.* 2005). The mechanistic role of smoking in the pathogenesis of periodontitis is uncertain, but studies have shown increased levels of TNF-α in gingival crevicular fluid of smoking as compared to non-smoking periodontitis patients (Bostrom *et al.* 1999).

Antibody response to Porphyromonas gingivalis in periodontitis

The oral cavity is a complex and diverse ecological reservoir for a large number of species, with more than 500 estimated species colonising the gingivae, teeth, tongue and cheek (Paster *et al.* 2001). A change in periodontal status is reflected by shifts in the subgingival oral community (Kumar *et al.* 2006). It remains unclear how the immune system differentiates between commensal and pathogenic species, but it is known that an immune response is mounted primarily against the major bacteria commonly associated with periodontal disease, such as *P. gingivalis*. A recent study investigated serum antibody levels against a number of pathogenic and commensal species in patients with periodontal destruction and healthy controls, showing that antibody levels to pathogenic species, in particular *P. gingivalis*, are

evelated in periodontal disease, whereas antibody levels to commensals are comparable between diseased subjects and controls. B cells and plasma cells are predominant in the periodontal lesion, representing about 68% of cells (Berglundh *et al.* 2005). The presence and levels of antibodies in periodontitis against pathogenic species seems to dependent on the combination of factors, such as the type of PD, smoking, gender, hygiene measurements and race/ethnicity (reviewed by Schenkein *et al.* 2007). For example, out of 18 different bacterial species tested using radioimmunoassay, only serum antibody levels to *P. gingivalis* were predictive for PD (Gunsolley *et al.* 1990). A more recent study of >5,000 dentate adults using checkerboard immunoblotting found similar results; out of 19 bacterial species tested, high serum IgG levels to *P. gingivalis* were most strongly associated with PD (OR ~2). All reviewed studies focussed on IgG responses; a few studies compared Ig subclasses and found that in general IgA are lower than IgG responses and that IgM responses do not discriminate between PD and healthy subjects (for example, Naito *et al.* 1987; De Nardin *et al.* 1991).

Most assays to measure antibody responses to *P. gingivalis* use whole cell bacterial lysates or whole formalin-fixed bacteria as antigen. So far, efforts to identify one or more dominant antigens (fimbrial proteins, LPS, gingipains, heat-shock proteins, mix of outer membrane proteins) were of limited success (reviewed by Kinane *et al.* 1999). This may be due to the loss of avidity when considering only one or a few antigens, variations in the makeup of the predominant antigens between different strains and different major autoantigens between individuals. For example, one study showed that six different *P. gingivalis* strains were required to adequately document the serum IgG response, and using a single or only a few strains missed a significant number of responses (Califano *et al.* 1999). Another study showed that the response to three different strains of *P. gingivalis* varies considerably between individuals with periodontal disease (De Nardin *et al.* 1991).

A number of studies reported a protective effect of higher serum antibody levels to whole cell extracts of *P. gingivalis* on periodontal disease severity (as measured by degree of bone loss) (Gunsolley *et al.* 1990; Rams *et al.* 2006), although other studies contradict this finding (reviewed by Kinane *et al.* 1999). Smoking seems to influence antibody levels depending on the context in which PD is presented (type of PD, race). Generally, lower IgG levels were reported in smokers compared to healthy controls (Quinn *et al.* 1998; Mooney *et al.* 2001).

The sum of these findings raises the question as to why the antibody response elicited by oral pathogenic bacteria, and in particular by *P. gingivalis*, does not clear the infection. Several mechanism are probably at play that allow bacterial immune evasion: the formation of biofilms in subgingival plaque acts as a physical barrier, potent proteases and peptidases (e.g. gingipains from *P. gingivalis*) inactivate immune mediators, for example by degrading antibodies and complement components, and a number of studies have suggested that the immune response against oral pathogens is immature in many patients, with a low antibody avidity and lowered protective capacity (Mooney *et al.* 1993; Benjamin *et al.* 1997).

Porphyromonas gingivalis PAD and citrullination

The most striking feature of this bacterium in the context of autoimmunity in RA is that *P. gingivalis* is the only prokaryote described to express a functional PAD enzyme (McGraw *et al.* 1999). The authors purified the enzymatically active form from the culture supernatant of *P. gingivalis* HG66. This strain is a mutant that arose naturally after repeated passage of the wildtype H66 strain. HG66 is characterised by spontaneous secretion of most of its cell-surface attached proteins, including PPAD, into the culture medium, and was therefore chosen for the comparably easy purification of cell-surface proteins from the culture supernatant (Jan Potempa, personal communication). The discovery of a peptidylarginine deiminase activity in *P. gingivalis* was a by-product of the initial aim to characterise peptide cleavage products using a gingipain preparation. Further fractionation of the gingipain preparation revealed an additional protein matching to a previously unassigned hypothetical protein in the *P. gingivalis* W83 genome, however starting at amino acid 44 (alanine) of the full-length hypothetical protein. This protein migrated as a single band with a molecular mass of 47 kDa, whereas the theoretical mass of the protein is 62 kDa, suggesting a truncated protein. The pH optimum was found at pH 9.3. Addition of flavin nucleotides appeared to increase the enzymatic activity, whereas Ca^{2+}, which is indispensable for the activity of human PAD enzymes, was not required for the activity of PPAD. Enzyme activity assays were performed using free L-arginine, several synthetic arginine derivates (such as benzoyl-L-arginine-ethylester), short peptidylarginine derivates (such as benzoyl-glycyl-L-arginine) and the short peptide bradykinin (contains nine amino acids, two of which are arginine: one on the amino-terminal and one on the carboxy-terminal end). The results showed that PPAD citrullinates free L-arginine, synthetic arginine derivatives, and bradykinin. Further analysis

revealed that bradykinin was citrullinated at the carboxy-terminal arginine, but not the amino-terminal arginine, suggesting a preference of PPAD towards arginine residues in a certain environment. Further, inhibition experiments using reagents that interfere with sulfhydryl groups (such as thiocitrulline, leupeptin and 4,4'-dithiopyridine) indicated the presence of a crucial cysteine residue in or near the active site of PPAD.

In retrospect, indications for the presence of a citrullinating enzyme can be found in a preceding publication (Hayashi *et al.* 1993). Insulin chain B was incubated with a semi-purified gingipain preparation (referred to as 'hemagglutinin with trypsin-like activity') and amino acid analysis of the resulting peptides by hydrolysis and thin-layer chromatography revealed that one peptide species contained a citrulline residue. This peptide species corresponded in molecular mass and remaining amino acid content to another peptide species harbouring one arginine, at its C-terminus. The arginine-containing peptide inhibited hemagglutination of red blood cells, whereas the citrulline-containing counterpart did not. To prove that the citrulline residue resulted from enzymatic conversion of arginine, free L-arginine was incubated with the gingipain preparation and the products analysed. 10-25% of the added L-arginine was recovered as L-citrulline, depending on incubation time, and no significant activity was detected using a boiled gingipain preparation. The authors therefore concluded in a side sentence that the gingipain preparation most likely contained an "arginine deiminase-like activity" and that this should be investigated further as a factor that plays a role in the hemagglutination by *P. gingivalis*.

Using a bioinformatics approach, two subsequent publications by Shirai *et al.* investigated the evolutionary relationship amongst a number of enzymes that catalyse the modification of arginine or arginine-derivatives. Based on the low level of amino acid sequence similarity between PPAD and human PADs, it was argued that these enzymes are not evolutionary related (Shirai *et al.* 2001). The crystal structure of PPAD has not yet been determined but structural prediction by Shirai *et al.* suggested an α/β propeller domain at its N-terminus and an immunoglobulin-like β sandwich domain at its C-terminus. Based on these structural similarities to human PAD4 (Figure 1.3), but also on the similar enzymatic function and (proposed) conserved catalytic residues (Figure 1.12), Shirai *et al.* concluded that human PADs, PPAD and five other guanidino group modifying enzyme families [arginine deiminase (ADI), aminotransferases (AT), dimethylarginine dimethylamino hydrolases (DDAH), agmatine iminohydrolase (AIH), and succinylarginine dihydrolase (AstB)] are actually

members of a novel superfamily, the guanidino group modifying enzymes (GME) (Shirai *et al.* 2006). In fact, on the amino acid sequence level PPAD is connected to the AIH family (Figure 1.12). AIHs deiminate agmatine (decarboxylated arginine) to *N*-carbamoylputrescine and ammonia but are often annotated in both genomic and structural databases as 'Porphyromonas-type peptidyl-arginine deiminases' or ‘putative peptidyl-arginine deiminases’. Several studies have now confirmed however that these enzymes are indeed AIHs and not PPADs (Nakada *et al.* 2003; Jones *et al.* 2010a; Jones *et al.* 2010b).

It is unknown what role protein citrullination might play in *P. gingivalis* and the pathogenesis of periodontitis. It was speculated that production of ammonia during PPAD-mediated deimination enhances the survival of *P. gingivalis* within the periodontal pocket (McGraw *et al.* 1999). Indeed, ADI- and AIH-catalysed ammonium production among oral bacterial species is known to act as a virulence factor, promoting the survival of microbial pathogens in the host environment (Takahashi *et al.* 1997). Ammonia neutralises acidic environments and might thereby optimise gingipain and PPAD function, and further has negative effects on neutrophil function by inhibiting phagocytosis, degranulation and oxygen metabolism (Niederman *et al.* 1990).

Additionally, based on the findings by Hayashi *et al.* described above, it can be speculated that PPAD acts as a virulence factor by generating citrullinated peptides which may assist the bacterium in spreading and circumventing the humoral immune response, for example by citrullinating peptides with naturally occurring C-terminal peptidylarginines or after gingipain-mediated protein cleavage, which can be efficient inhibitors of hemagglutination or unwelcome mediators of the host immune system (e.g. C5a).

A role for PPAD as a virulence factor is also supported by an *in vivo* study using the chamber model. In this model, a small metal coil (the chamber) is implanted subcutaneously into mice and inoculated with live *P. gingivalis*. Six hours after inoculation, the chamber fluid was removed and subjected to proteomic analysis. It was found that PPAD and 13 other proteins were upregulated, compared to the proteome of *P. gingivalis* grown in culture (Yoshimura *et al.* 2008).

Figure 1.12 Amino acid sequence alignment of the catalytic domains of PPAD (residues 86-363), AIH from *Dyadobacter fermentans* DSM 18053 (locus Dfer_2630, residues 60-352), and human PAD4 (residues 306-556). Residues identical in PPAD and AIH and/or PAD4 are highlighted. Guanidino-binding (#) and catalytic residues (*) which are conserved in all families of the guanidino-group modifying enzyme superfamily are indicated. The amino-terminal sequence of each enzyme is unique and not shown here. In PAD4 the amino-terminal portion is folded into two consecutive Ig-like β-sandwich domains preceding the catalytic domain harbouring the α/β-propeller fold (Arita *et al.* 2004). A long 200-residue carboxy-terminal extension of PPAD is predicted to adopt a Ig-like β-sandwich structure (Shirai *et al.* 2006). This figure was previously published in a review co-authored by myself (Mangat *et al.* 2010). PPAD: *P. gingivalis* peptidylarginine deiminase; AIH: agmatine iminohydrolase; PAD4: peptidylarginine deiminase 4.

Studies on the links between periodontitis and RA

Several studies, reviews and opinion articles have been published on possible associations between RA and PD or *P. gingivalis*, many of which are hampered by small study size, crude assays and/or uncritical evaluation. A critical summary shall be provided here

Epidemiological association between RA and PD

A number of studies have reported a significant association between RA and PD and/or tooth loss. The results are difficult to compare because widely heterogenous disease classification criteria for RA and PD were used, some even relying on self-reported health questionnaires, and only a few studies appropriately adjusted their results for important confounding factors such as smoking and age. Despite these limitations, it can be concluded that a trend emerges indicating that RA is more prevalent in individuals with periodontitis and *vice versa* in comparison to the general population. The most recent studies are summarised here (also see Table 1.5):

Mercado *et al.* studied 1,412 individuals attending a dental clinic in Brisbane, Australia (Mercado *et al.* 2000). Subjects were divided into two groups: those that were referred for periodontal treatment (defined as the PD group, n=809) and those that attended the clinic for other treatment (defined as the general group, n=603). Severity of periodontal bone loss (no, mild, moderate, severe) was further assessed in the PD group using dental X-rays. RA was assessed based on a self-reported health questionnaire and the self-reported prescription of drugs consistent with RA (e.g. NSAIDs, methotrexate). Based on these criteria, the prevalence of RA in the PD group was 32 out of 809 subjects (3.95%), compared to 4 out of 603 (0.66%) in the general group ($p<0.05$). Within the PD group, patients with RA were more likely to have moderate-to-severe PD (62.5% of all RA patients) than patients without RA (43.8%) ($p<0.05$). Obvious limitations to this study are the reliance on self-reported RA, the definition of PD using a non-validated parameter (i.e. referral), the failure to document and adjust for smoking, and the lack of a limiting time period between clinic attendance and self-reporting in the health questionnaire.

The authors therefore conducted a subsequent study (Mercado *et al.* 2001) involving 65 consecutive patients attending a tertiary referral rheumatology clinic in Brisbane. All subjects

fulfilled the ACR 1987 criteria for classification of RA and were clinically examined for their periodontal health. The control group consisted of an undisclosed number of age-, sex- and smoking status-matched individuals without RA that were selected from a pool of patients who presented at a general dental clinic in Brisbane and were examined for periodontal health in the same way as the RA patients. In comparing the RA with the non-RA group, statistically significant associations ($p<0.05$) were an increased number of missing teeth (11.6 in RA vs. 6.7 in non-RA), a higher percentage of subjects with deeper (>6.2 mm) pocketing (44.6 vs. 24.6%), and a higher percentage of subjects with moderate to severe bone loss (69.2% vs. 33.8%). The values for the latter parameter compare well with the previous study on self-reported RA (62.5% vs. 43.8%, see above).

Another study on data from 4,461 subjects from the American NHANES III (National Health and Nutrition Examination Survey) investigated the associations between RA and periodontal measures (de Pablo *et al.* 2008). All subjects were aged ≥60 years, as musculoskeletal examination had been carried out in this age group only in the initial investigation. RA was defined based on the presence of only 3 out of 7 ACR 1987 criteria (normally 4 are required) because radiographic data were not available. PD was defined on the basis of clinical parameters. 2.3% of all subjects were defined as having RA, which is considerably higher than the average (0.5-1%) and might result from the looser classification criteria for RA used in this study. This is in line with the finding that only 29% of RA subjects were positive for rheumatoid factor, indicating that a proportion of patients defined as having RA in fact had other rheumatic diseases. Compared with non-RA subjects, participants with RA had a higher prevalence of PD (16% vs 10%), edentulism (56% vs. 34%), and a higher number of missing teeth (20 vs. 16, $p=0.0001$). Across all subjects, participants with RA were more likely than individuals without RA to be edentulous (OR 2.27, 95% CI 1.56-3.31). Among dentate subjects, participants with RA were more likely to have PD, independent of age, sex, ethnicity and smoking habits (OR 1.82, 95% CI 1.04-3.20). Recognising the weaknesses of the study as a result of using 3 instead of 4 classification criteria for RA, the authors conducted a sensitivity analysis based on the presence of 4 criteria for the definition of RA. The results showed a stronger association between RA and endentulism (OR 3.34, 95% CI 1.16-9.64), and RA and PD (OR 4.13, 95% CI 1.30-13.15), independent of age, sex, race/ethnicity, and smoking.

The methodologically soundest study investigated 57 Caucasian subjects with RA and 52 non-RA controls (Pischon *et al.* 2008). RA subjects were recruited from a rheumatology clinic in Berlin, Germany, and RA was classified according to standard ARC 1987 criteria. Control subjects were recruited from a general dental clinic in Berlin, and were matched for age and sex. The exclusion criterion for both groups was a history of periodontal therapy. All subjects were clinically assessed for periodontal disease, which was defined as a mean CAL > 4 mm (moderate-to-severe PD). Questionnaires assessed smoking status (never, former, current), education, alcohol consumption and BMI. Statistical analysis showed an increased risk of PD in subjects with RA compared to controls (OR 8.05, 95% CI 2.93-22.09), independent of smoking, gender, education, alcohol consumption and BMI, and adjusted for age. When further adjusted for oral hygiene measures (plaque, gingival bleeding), the OR for the association of RA with PD was 6.09 (95% CI 1.72-21.58). Weaknesses of this study are the control population which is not representative of the general population and the relatively small sample size in both populations. However, the detailed documentation of smoking status, appropriate statististical adjustment and disease classification for both RA and PD are considerable strengths of this study.

The most recent study (Nesse *et al.* 2010) studied a large number of subjects (n=1,208), failed however on appropriate classification of PD and RA, and the sufficient documentation of smoking and the relevant statistical adjustment. The 1,208 subjects who met inclusion criteria had been randomly selected from dental records of patients attending a dental or periodontal clinic. Data on prevalence of RA, other autoimmune diseases and cardiovascular disease was obtained from a health questionaire within 1 year of date of periodontal screening. Sex, age, and smoking habits (currently yes or no) were determined from dental records; smoking status was available for only 50% of subjects, however. PD was defined based on a probing depth ≥ 4 mm. Probing depth is defined as the distance from the free gingival margin to the bottom of the periodontal pocket, and is therefore a fraction of CAL, the official measure. A probing depth of ≥ 4 mm therefore corresponds to a CAL of at least 4 mm (more likely ≥ 5 mm), which is defined as severe PD according to the official criteria (Armitage 1999). The control group consisted of patients attending the dental clinic without a diagnosis of PD (n=320). The PD group included subjects attending the dental or periontal clinic with a diagnosis of PD (n=888). In uncontrolled analysis, the prevalence of RA was significantly increased in PD patients [1 control (0.3%) had RA, and 26 PD subjects (2.9%) had RA]. The nature of the observed association between PD and RA remained unclear however because of

the low prevalence of RA (n=27) and the lack of smoking data for half of these subjects, reducing statistical power.

Table 1.5 Summary of epidemiological studies on the association between rheumatoid arthritis and periodontitis.

Cases	Controls	Adjustment for smoking	Reported associations	Limitations	Reference
809 subjects with PD PD definition: referred for periodontal treatment. Severity of PD assessed using dental X-rays RA definition: self-reported health questionnaire	603 non-PD subjects, unmatched Non-PD definition: subjects attending a clinic for treatment other than periodontitis RA definition: self-reported health questionnaire	No adjustment	PD vs non-PD: RA present in 3.95% of PD group and 0.66% of control group ($p<0.05$) Within PD group, moderate-to-severe PD present in 62.5% of all RA patients and 43.8% of non-RA patients ($p<0.05$)	Self-reported RA Definition of PD using non-validated parameter (referral) No documentation of time period between PD assessment and RA self-reporting	(Mercado *et al.* 2000)
65 subjects with RA RA definition: fulfilment of at least 4 out of 7 ACR 1987 criteria PD definition: clinical examination	n = ? non-RA subjects, matched for age, sex and smoking status Non-RA definition: no fulfilment of at least 4 out of 7 ACR 1987 criteria PD definition: clinical examination	Cases and controls matched for smoking	RA versus non-RA: More missing teeth (11.6 vs. 6.7); higher % of subjects with deep pockets (44.6 vs. 24.6%); higher % of subjects with moderate-to-severe bone loss (69.2 vs. 33.8%) ($p<0.05$ for all associations)	Cases and controls were matched for smoking but it is not clear according to which parameters this was done (ever/never/ current smoker, extend of smoking etc.) Small sample size	(Mercado *et al.* 2001)

Cases	Controls	Adjustment for smoking	Reported associations	Limitations	Reference
103 subjects with RA RA definition: fulfilment of at least 3 out of 7 ACR 1987 criteria PD definition: clinical examination	4,358 non-RA subjects Non-RA definition: no fulfilment of at least 3 out of 7 ACR 1987 criteria PD definition: clinical examination	Statistical adjustment based on questionnaire (never, former, current, number of cigarettes per day for current smokers)	RA vs. non-RA: more likely to have PD independent of age, sex, ethnicity and smoking habits (OR 1.82, 95% CI 1.04-3.20)	Adjustment for smoking based on simple questionnaire	(de Pablo *et al.* 2008)
57 subjects with RA RA definition: fulfilment of at least 4 out of 7 ACR 1987 criteria PD definition: clinical examination	52 non-RA subjects non-RA definition: no fulfilment of at least 4 out of 7 ACR 1987 criteria PD definition: clinical examination	Statistical adjustment based on questionnaire (never, former, current)	RA vs. non-RA: increased risk of PD when adjusted for smoking, gender, education, alcohol consumption, BMI, age and oral hygiene (OR 6.09, 95% CI 1.72-21.58)	Small sample size Adjustment for smoking based on simple questionnaire	(Pischon *et al.* 2008)
888 subjects with PD PD definition: clinical examination RA definition: health questionnaire	320 non- PD subjects Non-PD definition: no diagnosis of PD after clinical examination RA definition: health questionnaire	No adjustment	PD versus non-PD: Prevalence of RA increased (2.9% vs. 0.3%, $p<0.05$)	No adjustment for smoking RA defined by health questionnaire	(Nesse *et al.* 2010)

Common risk factors: Shared epitope and smoking

As cigarette smoking is a risk factor common to both RA and periodontitis, it also is an important confounding factor for epidemiological studies on the association between periodontitis and RA, as summarised in the previous section. Some recent studies (Mercado *et al.* 2001; de Pablo *et al.* 2008; Pischon *et al.* 2008) included adjustment for smoking in their analysis, showing that the association between RA and periodontitis is independent of smoking; however, the evaluation of smoking was based on questionnaires, sometimes containing a grossly simplified documentation of smoking variables (yes/no) (Mercado *et al.* 2001), resulting in the potential to inflate the incidence of RA in individuals with periodontitis in these particular studies (discussed by Hujoel *et al.* 2002).

Studies evaluating the 'shared epitope' as a risk factor for the development of PD are small in scale and heterogenous disease classification and patient/control selection criteria were used. Further, patients were often recruited from different ethnical backgrounds, and a varying number of HLA antigens was included in each study, making it difficult to interpret the results and to draw a conclusion. The relevant studies are summarised here:

The first study to report an association between HLA-DR4 and periodontitis was published in 1987 on 10 Caucasian patients with "rapidly progressing periodontitis" and 120 blood donor controls with unknown periodontal status (Katz *et al.* 1987). All subjects were typed for HLA-A, -B, -C, and -D. The only significant association was established for HLA-DR4, which was present in 80% of PD patients and in 38.3% of controls. Drawbacks to this study are the small number of sampled cases and the unknown periodontal status of the control group.

Alley *et al.* studied 15 Caucasian subjects with moderate to severe chronic PD and 15 healthy controls (no PD) by determining their HLA-D phenotype. HLA-DR4 was highly associated with PD ($p<0.001$) compared to healthy controls (Alley *et al.* 1993). The small sample size of both cases and controls makes this study less conclusive. The same results were published a second time in another paper, including an additional analysis on IgG serum responses to *Capnocytophaga* (Dyer *et al.* 1997), and is therefore not considered an independent study for this purpose.

Firatli et al. determined HLA-A, -B, -C, and -DR phenotypes in 30 Caucasian patients with "juvenile periodontitis", and 30 Caucasian patients with "rapidly progressive periodontitis" patients. Controls were 3,731 subjects of the general population (periodontal status unknown). In both forms of PD, DR4 was found "at a significantly higher level than the control group" (Firatli *et al.* 1996), although the detailed statistical values are not revealed, making an interpretation difficult, in addition to the small number of cases.

The study most often cited in this context was published by Bonfil *et al.*, who studied 48 cases with "severe periodontitis", and a subgroup of 12 patients characterised by extremely severe disease who met the criteria as applied in the study by Katz *et al.*, 1987 (see above). Control subjects (n = 55) were free of periodontal disease. DR4 alleles and subtypes were detected for all subjects and the results adjusted for ethno-geographic origin, as the study population was of mixed ethnic backgrounds. A trend, but not a significant association was found between the presence of the DR4 serotype and periodontitis (16% of controls were positive compared to 23% of severe PD and 42% of extremely severe PD). A statistically significant association was found between the presence of either of the subtypes *0401, *0404, *0405 and *0408 and extremely severe periodontitis (p=0.08) (Bonfil *et al.* 1999). The drawbacks to this study are a mixed population, although this was adjusted for statistically, and small number of cases. Although the same disease classification criteria were used to classify extremely severe PD by this study (n=12) and by Katz *et al.* (n=10) the DR4 frequency differs considerably. Katz *et al.* found that 80% of patients and 38.3% of controls were positive, while Bonfil *et al.* detected DR4 in 42% of patients and 16% of controls. The variations might stem from the small study sizes.

Marotte *et al.* investigated the link between RA, PD, and the 'shared epitope' by genotyping 147 patients with RA and assessing them for wrist and periodontal bone destruction using X-rays; no healthy controls were enrolled. Based on dental X-ray data, 56.5% of RA patients had periodontal bone destruction. Significant associations were found between the presence of the 'shared epitope' and wrist bone destruction, compared to the absence of the 'shared epitope' (OR 2.5, 95% CI 1.16 - 5.42, $p<0.05$); and further between the presence of the 'shared epitope' and periodontal bone destruction (OR 2.2, 95% CI 1.04 - 4.84, $p<0.05$). A stronger correlation was observed between patients with destruction at both sites and the presence of the 'shared epitope', compared to patients without any destructions (OR 3.9, 95% CI 1.53 - 9.96, $p<0.001$) (Marotte *et al.* 2006). However, on detailed analysis of the method

to determine the presence of the 'shared epitope' alleles, it becomes apparent that in fact only one primer pair was used, complementary to HLA-DRB4 (BLASTN analysis). While the DRB4 locus is linked to the DRB1 locus, which encodes the 'shared epitope' alleles, this is a rather confusing and indirect way of assessing the presence of the shared epitope.

Stein *et al.* performed a meta-analysis of case-control studies involving Caucasian subjects that were typed for HLA-A, -B, -Cw and DR and/or –DQ; where cases were defined as having aggressive (early-onset/rapidly progressive/localised/generalised juvenile) PD or chronic (adult) PD; where controls were defined as free of periodontal disease in the case of chronic PD (for aggressive PD blood donors were accepted as controls because aggressive PD is rare in the general population); and with consistent and reproducible data presentation. 12 studies fulfilled these criteria, including Katz *et al.* and Alley *et al.* The other 10 studies included however did not find a significant association between periodontitis and DR4 or specific 'shared epitope' alleles, so that the DR4 association lost its significance when all studies were combined. The only significant association in this meta-analysis was established between PD and the MHC class I antigens HLA-A9 and -B15 (Stein *et al.* 2008).

Shared antibody profiles

A recent study (Havemose-Poulsen *et al.* 2006) measured serum anti-CCP IgG levels in the serum of 23 patients with RA (no PD), 45 patients with generalised and localised aggressive periodontitis (no RA) and 25 healthy controls (no PD/RA). Periodontitis and RA was diagnosed after clinical examination and classified according to generally accepted criteria (ACR 1987 criteria for RA; American Academy of Periodontology classification criteria for PD). A somewhat higher than expected frequency of anti-CCP antibodies in patients with generalized aggressive periodontitis (7.4%; 2 out of 27) was found when compared with controls, though the difference was not statistically significant. Drawbacks to this study are the small sample size and the sampling of relatively young patients only (all subjects were ≤35 years old). Noticeable is also the fact that only 35% of RA patients had anti-CCP antibodies.

Another study measured anti-CCP antibodies in the serum of 33 individuals who developed RA up to 8 years after their serum was collected and their periodontal status assessed (Molitor *et al.* 2009). Patients were selected based on hospitalisation discharge codes of 5,413

participants in the Atherosclerosis Risk in Communities study. After adjustment for age, sex, and race, the hazard ratio (HR) of developing RA in subjects with moderate-to-severe PD (n=27) was 2.6 (95% CI 1.0-6.4), compared to those with no or mild PD (n=6). Among life-time non-smokers, the HR rose to 8.8 (95% CI 1.1-68.9). Anti-CCP levels were significantly higher (p=0.04) in patients with moderate-to-severe PD (225.5 U) than in those with no or mild PD (8.4 U). As these data are presented in the form of a conference abstract, a more detailed interpretation is not possible.

The study by Pischon *et al.*, presented in detail on p. 58, also measured anti-CCP antibodies in 57 subjects with RA and 52 non-RA controls that were clinically examined for periodontal disease. The mean anti-CCP levels did not differ significantly in RA subjects with or without PD (Pischon *et al.* 2008). Another study measuring anti-CCP antibody in serum from 22 patients with chronic periodontitis reported that only one patient was borderline positive (22.2 U/ml, cut-off is 20 U/ml), whereas Hendler *et al*. found anti-CCP antibodies in patients with aggressive periodontitis but not in patients with chronic periodontitis or gingivitis (Hendler *et al.* 2010).

Mikuls *et al.* studied serum of 78 RA patients (classified according to ACR 1987 criteria) with unknown periodontal status, 39 non-RA patients with moderate to severe PD and 40 healthy controls (Mikuls *et al.* 2009). Smoking status was recorded as either current smoker or not. Subjects were ≥19 years old. Antibody levels to whole cell extracts of *P. gingivalis* were measured by ELISA in all three groups, and anti-CCP was measured in PD and RA subjects. Antibody levels to *P. gingivalis* showed statistically significant differences between groups, with 40% of controls, 67% of RA, and 77% of PD being positive (p = 0.002). Further, 91% of RA subjects but only 5% of PD subjects (2 out of 39) were positive for anti-CCP antibodies. Antibody levels to *P. gingivalis* were correlated to clinical parameters in RA subjects, showing that total anti-CCP or RF did not correlate, but certain anti-CCP IgG subtypes did (IgG2, p = 0.01; IgM, p = 0.02), as did CRP (p = 0.003). The anti-*P. gingivalis* assay, which used a crude lysate of bacterial cells, could be criticised for a lack of specificity. Nevertheless, the test can differentiate between individuals infected with *P. gingivalis* and those who are not (Ebersole *et al.* 1987). The results are supported by a more recent study, where 82 patients with RA and 47 non-RA controls, all from a Native North American population genetically predisposed for RA, were assessed for antibody levels to *P. gingivalis* LPS and CCP, showing that anti-*P. gingivalis* antibody levels are higher in patients with RA

than controls (p=0.005) and associated with anti-CCP antibody positivity (p=0.04), but not rheumatoid factor-positivity (Hitchon *et al.* 2010).

Another commonly cited publication (Moen *et al.* 2003) investigated antibody responses to *P. gingivalis*, *P. intermedia*, *T. forsythensis*, and *Candida albicans*, in 116 RA serum, 52 non-matching RA synovial fluid, 43 non-RA (various arthritides) synovial fluid, 9 osteoarthritis synovial fluid and 100 blood donor (no history of RA; PD status unknown) samples. Antibody levels were measured by ELISA, however whole formalin-fixed cells instead of cell lysates were used. Further, the *P. gingivalis* strain used was different to the one used in Mikuls *et al*. The authors discuss various differences in antibody levels between groups and strain, for example that levels to *P. gingivalis* are comparable between all groups and that antibody levels to *T. forsythensis* are higher in synovial fluid of RA and non-RA compared to OA patients, but lower in the serum of RA patients compared to blood donor controls. Especially the latter is surprising, indicating inappropriate choice of study and control populations (non-matched serum/synovial fluid and unknown periodontal status of controls).

Finally, Ogrendik studied serum from 30 RA patients and 20 healthy blood-donor controls (Ogrendik *et al.* 2005). All subjects were free of periodontal disease. Antibody levels to *P. gingivalis* (same strain as Mikuls *et al.*), *Prevotella melaninogenica*, *T. forsythensis*, *P. intermedia*, and *A. actinomycetemcomitans* were measured by ELISA using whole bacteria. It is assumed here that the bacteria were formalin-fixed and not alive, although this is not explicitly stated in the paper. The results showed that IgG levels were significantly higher in RA patients compared to controls for all bacteria except *A. actinomycetemcomitans* (*P. gingivalis*, $p<0.001$; *P. intermedia*, $p< 0.001$; *P. melaninogenica*, $p<0.01$; *T. forsythensis*, $p<0.05$). Further, a positive correlation between antibody levels to *P. gingivalis* and serum CRP in RA patients was found ($r=0.809$). Although an interesting result, it is unclear what the cause of the increased antibody levels in RA patients is, as they were defined as free of current periodontal disease. Possible explanations is humoral immunity resulting from previous periodontal conditions or increased total serum IgG levels, which was not adjusted for in this study.

Presence of P. gingivalis DNA in serum or synovial fluid from RA patients

Moen *et al.* investigated serum and synovial fluid from 16 patients with RA, 14 with spondyloarthropathies, and 9 with osteoarthritis for *P. gingivalis* DNA (Moen *et al.* 2006). Total DNA was extracted and submitted to checkerboard DNA-DNA-hybridisation using 16S rRNA-based oligonucleotide probes for identification of 40 different oral bacteria species. The authors reported higher numbers of different bacterial species in the serum of RA (average of 6.2 different bacterial species) and spondyloarthropathies patients (5.4) over osteoarthritis patients (2.7). Higher numbers were also found in synovial fluid of RA (14.0) and spondyloarthropathies (19.4) over osteoarthritis (4.0) patients. *P. gingivalis* was detected in RA and spondyloarthropathies serum/synovial fluid but not in osteoarthritis patients. Unfortunately, as any technique based on affinity and avidity, checkerboard DNA-DNA-hybridisation, when used with DNA samples containing only a minimal fraction of bacterial DNA compared to human DNA, is prone to false positives and cross-reaction. If the results are to be trusted however, they indicate that synovial inflammation, which is characteristic for RA and spondyloarthropathies but generally not for osteoarthritis, traps oral bacterial DNA, possibly in leukocytes in the form of engulfed bacteria or bacterial products. However, it does not indicate that live bacteria in the joint are directly responsible for disease. A more recent study aimed at identifying periodontal bacterial DNA in subgingival plaque, serum, synovial fluid and leukocytes from 19 patients with both RA and periodontitis failed to support this view (Martinez-Martinez *et al.* 2009). DNA was extracted and subjected to PCR with species-specific primers to *P. gingivalis* and five other periodontal pathogens. DNA from at least one bacterium was detected in 100% of subgingival plaque samples, 100% of synovial fluid samples, and 83.5% of serum, but in none of the DNA samples extracted from blood leukocytes. No bacterial species could be grown on agar plates under aerobic or unaerobic conditions from any of the synovial fluid samples. The authors conclude that the most likely mode of transport of bacterial DNA from the mouth to the joints is in its free form.Unfortunately, the original PCR agarose gels are not shown to convince the reader of the specificity of the primers. If the results can be replicated, they would imply that free bacterial DNA may trigger inflammatory pathways in the joint.

Similar pathobiology

PD and RA are characterised by similar pathogenic findings and mechanisms: both diseases are characterised by destruction of tissue and erosion of periarticular bone and periodontal bone, respectively, as a result of persistent inflammation. In both diseases the pathogenesis includes B cells and proinflammatory cells, release or activation of proinflammatory mediators, notably TNF-α, antibody production and upregulation of matrix metalloproteinases (reviewed by Mercado *et al.* 2003; Bartold *et al.* 2005; de Pablo *et al.* 2009).

Common treatment effects

Another commonly cited link is a beneficiary effect on RA of treatment with certain antibiotics and of mechanical periodontal treatment.

Ogrendik evaluated the efficacy of ornidazole (Ogrendik *et al.* 2006), clarithromycin (Ogrendik 2007a), and levofloxacin in combination with methotrexate (Ogrendik 2007b) in RA. Ornidazole is a nitroimidazole antiprotozoal drug which also has antibacterial activity against anaerobes and is therefore used in periodontal therapy (Kamma *et al.* 2000); levofloxacin is a fluoroquinolone antibiotic used in periodontal therapy (Kleinfelder *et al.* 2000); clarithromycine is a macrolide antibiotic that is best known for its efficacy against *Helicobacter pylori*, but is also used in periodontal treatment (Burrell *et al.* 2008). All studies were randomised and placebo-controlled, enrolling 76-160 patients, with patients having RA (ornidazole), persistently active RA despite at least 6 months methotrexate therapy (levofloxacin), or early RA (clarithromycin). Studies ran for 3 months (ornidazole) or 6 months (levofloxacin, clarithromycin). Efficacy measures were the percentage of subjects showing an ACR 20/50/70 response, which requires a subject to have a 20/50/70% reduction in the number of swollen and tender joints, and a reduction of 20/50/70% in three of the following five parameters: physician global assessment of disease, patient global assessment of disease, patient assessment of pain, C-reactive protein or erythrocyte sedimentation rate, and degree of disability in Health Assessment Questionnaire (HAQ) score. It is reported in all three studies that the treatment group showed statistically significant ($p<0.001$) levels of ACR 20, 50 and 70 improvements, compared to placebo control, with ARC 20 being 50-62%

in cases and 18-33% in controls, ACR 50 34-38% in cases and 8-11% in controls, and ACR 70 18-20% in cases and 1-3% in controls for all drugs.

It is not clear which mechanisms lead to the observed efficacy of these antibiotics in RA. Although they are used to treat periodontal disease, these studies have not investigated whether efficacy mirrors improvement in periodontal health status. Additionally, other, non-antibiotic mechanisms might mediate their efficacy. This has been shown for example for minocycline, a tetracycline derivative antibiotic which shows efficacy in early rheumatoid factor-positive RA (O'Dell *et al.* 2001). Minocycline has been shown to have anti-inflammatory, immunomodulatory, and chondroprotective effects: it inhibits phospholipase A2 (Pruzanski *et al.* 1992) and metalloproteinases (Golub *et al.* 1998) activity, nitric oxide synthetase expression (Amin *et al.* 1996), and suppresses leukocyte-associated tissue damage (Gabler *et al.* 1991).

The benefit of mechanical periodontal therapy on RA disease activity, measured by the disease activity score (DAS) 28 or the HAQ Disability Index (HAQ-DI) score, was assesed in three studies. DAS28 is a numeric index and calculated using number of tender joints, swollen joints, patients' general assessment of their condition, and ESR. The HAQ-DI assesses a patient's level of functional ability and includes questions of fine movements of the upper extremity, locomotor activities of the lower extremity, and activities that involve both upper and lower extremities.

Al-Katma *et al.* enrolled 29 subjects with RA (classified using ACR 1987 criteria; DAS28 ≥ 2.5) and generalised mild-to-moderate chronic PD (≥3 years duration) (Al-Katma *et al.* 2007). All subjects were recruited from a rheumatology clinic in Cleveland, USA, were ≥30 years old, non-smokers and non-diabetic. 17 subjects received detailed oral hygiene instructions and periodontal treatment by full-mouth scaling and root planning (SRP), which is a cleaning of the root surfaces below the gum line. The other 12 subjects received neither treatment nor instructions. DAS28 and periodontal parameters were recorded at baseline and after 8 weeks. 13 out 17 subjects (76%) in the treatment group showed improvement in DAS28, compared to 2 out of 12 (17%) in the untreated group. A statistically significant difference was found between DAS28 and ESR between the treatment and control group. Further, all periodontal parameters were significantly correlated with improvement of DAS28 and ESR.

Similar results were obtained in another study in the same center (Ortiz *et al.* 2009). 40 subjects with RA and generalised severe chronic PD were enrolled. Subjects were ≥30 years old, non-smokers and non-diabetics. 20 subjects were randomly assigned to periodontal therapy (SRP) and oral hygiene instructions, the other 20 received neither therapy nor instructions. At baseline and after 6 weeks, periodontal clinical parameters, DAS28, ESR, and TNF-α levels in peripheral blood were measured. Patients receiving treatment showed a significant decrease in mean DAS28, ESR, periodontal clinical parameters and serum TNF-α at 6 weeks after treatment. No significant decrease was seen in subjects without periodontal treatment.

A third study (Ribeiro *et al.* 2005) enrolled 42 consecutive patients attending a rheumatology clinic in Brazil with RA (classified according to ACR 1987 criteria) and severe PD (clinical assessment). Patients were ≥40 years old and non-smokers. RA severity was determined based on the HAQ-DI, ESR, and serum IgM-RF. Subjects were divided into 2 groups, matched for age and gender. Group 1 (n=16) received oral hygiene instructions and professional supragingival tooth cleaning, whereas goup 2 (n=26) additionally submitted to SRP. Periodontal and laboratory measurements were taken 3 months after periodontal treatment. Group 2 showed significant improvement on periodontal clinical parameters; some improvement was recorded for group 1. No significant association between these periodontal clinical parameters and the HAQ-DI score was measured between groups and before and after therapy. RF levels were reduced for both groups but this was not statistically significant. However, the ESR was significantly reduced for group 2.

It becomes apparent that all three studies demonstrated a siginificant reduction in the ESR, which is a component of the DAS28 but not the HAQ score. This most likely reflects the reduction in periodontal inflammation as a result of periodontal treatment, and therefore lower levels of systemic inflammatory markers such as TNF-α. The time period between periodontal treatment and assessment for RA activity is relatively short in all three studies (6-12 weeks). An improvement in HAQ score might become apparent over a longer period of continous improved periodontal health, especially as this seems to result in significantly decreased systemic TNF-α levels.

Aims of this study

The aim of this project was to investigate the molecular basis of protein citrullination by *P. gingivalis* and the potential of this to contribute to the development of anti-citrullinated protein antibodies in RA.

This involved five major steps:

1. Examination of the presence of endogenous citrullinated proteins in *P. gingivalis* and other periodontopathogenic bacteria (Chapter 3).

2. Determination of the potential of *P. gingivalis* to citrullinate human proteins, in particular known RA autoantigens such as α-enolase and fibrinogen (Chapter 3).

3. Analysis of the antibody responses to *P. gingivalis*-generated citrullinated peptides in patients with RA (Chapter 3).

4. Cloning and expression of *P. gingivalis* PAD in an active form and investigation of its functional properties, activity towards human proteins, and its inhibition by candidate compounds (Chapter 4).

5. Generation of a polyclonal antibody to *P. gingivalis* PAD; examination of the native form and subcellular localisation of the enzyme in bacterial lysates; investigation of periodontal tissue for the presence of citrullinated proteins and *P. gingivalis* PAD (Chapter 5).

Chapter 2 Materials and Methods

List of materials and reagents

Reagent/Material	Supplier	Storage and Notes
Agar	BD	Difco™, granulated RT
Agarose, UltraPure™	Invitrogen	RT
Albumin	Sigma	From bovine serum 4˚C
2-Amino-2-methyl-1,3-propanediol (AMPD)	Sigma-Aldrich	RT
Ammonium ferric sulphate Dodecahydrate	Fluka	Protect from light RT
Ampicillin	Sigma	RT
Antarctic phosphatase	NEB	5,000 U/ml -20˚C
Anti-CEP-1 antibody	(Lundberg *et al.* 2008)	Rabbit polyclonal antibody Raised against the CEP-1 peptide from human α-enolase (CKIHA-Cit-EIFDS-Cit-GNPTVEC) -20˚C
Anti-citrulline (modified) detection kit	Upstate/Millipore	-20˚C
Anti-enolase (H-300) antibody	Santa-Cruz Biotechnology	Rabbit polyclonal antibody Raised against amino acids 1-300 mapping at the N-terminus of human α-enolase. Detects α-, β- and γ-enolase. 4˚C
Anti-fibrinogen α (C-20) antibody	Santa-Cruz Biotechnology	sc-18026 Goat polyclonal antibody Raised against a peptide mapping near the C-terminus of human fibrinogen-α. 4˚C
Anti-fibrinogen β (I-18) antibody	Santa-Cruz Biotechnology	sc-18027 Goat polyclonal antibody Raised against a peptide mapping within an internal region of human fibrinogen-β 4˚C
Anti-fibrinogen γ (C-20) antibody	Santa-Cruz Biotechnology	sc-18032 Goat polyclonal antibody Raised against a peptide mapping at the C-terminus of human fibrinogen-γ 4˚C
Anti-his antibody	Novagen	#70796-3 Mouse monoclonal antibody
Anti-goat Ig antibody, HRP-conjugated	DAKO	#P0449 Raised in rabbit 4˚C
Anti-GST antibody, HRP-conjugated	GE Healthcare	#27-4577-01 Goat polyclonal antibody 4˚C

Anti-mouse Ig antibody, HRP-conjugated	DAKO	#P0260 Raised in rabbit Dilution for WB: 1:5000 4˚C
Anti-rabbit Ig antibody, HRP-conjugate	DAKO	#P0448 Raised in goat Dilution for WB: 1:5000 4°C

BCA Protein Assay Kit	Pierce	RT
Benzonase Nuclease	Novagen	250 U/µl, purity >99% -20˚C
Biotaq DNA Polymerase	Bioline	-20˚C
BL21 (DE3) chemically competent cells	Merck	-80˚C
Bpu10I	NEB	5,000 U/ml (R0649S)
Bradford Coomassie Plus Assay Reagent	Pierce/Thermo Scientific	RT
Bromophenol Blue	Sigma	RT
BugBuster Protein Extraction Reagent	Novagen	RT
2,3-Butanedione oxime	Fluka	RT

C41 (DE3) OverExpress competent cells	Lucigen Corporation	-80˚C
C41 (DE3) pLysS OverExpress competent cells	Lucigen Corporation	-80˚C
C43 (DE3) OverExpress competent cells	Lucigen Corporation	-80˚C
C41 (DE3) pLysS OverExpress competent cells	Lucigen Corporation	-80˚C
Calcium chloride	Riedel-de Haën	97% powder, pure RT
Chloroacetamidine hydrochloride	Alfa Aesar	4˚C
2-(Cyclohexylamino) ethanesulfonic acid (CHES)	Sigma	RT

DC Protein Assay	Bio-Rad	RT
dNTP solution mix	NEB	10 mM each nucleotide avoid repeated freeze/thaw -20˚C
Doxycycline hyclate	VWR/Applichem	4˚C

ECL blocking agent	Amersham/GE Healthcare	4˚C
ECL and ECL PLUS Western blotting detection reagents	Amersham/GE Healthcare	4˚C
EDTA disodium salt	VWR/BDH	AnalaR NORMAPUR

Full-range rainbow molecular	Amersham/GE	-20˚C

weight markers	Healthcare	

Glutathione Sepharose 4B	GE Healthcare	4˚C
Glycerol	BDH/VWR	AnalaR Normapur, bidistilled, 99.5 % RT
GoTaq Hot Start Polymerase	Promega	5 U/µl -20˚C

High-range rainbow molecular weight markers	Amersham/GE Healthcare	-20˚C
HRV 3C protease	Novagen/Merck	2000 U/ml -20˚C
Hyperladder I	Bioline	4˚C

Imidazole, 99% pure	Sigma-Aldrich	RT
Instant Blue	Novexin/Triple Red	4˚C
IPTG, Dioxane-Free, High Purity	Calbiochem	2-8˚C protect from light

Kanamycin monosulfate	Sigma	from *Streptomyces kanamyceticus* RT
KOD Hot Start DNA polymerase	Novagen/Merck	1 U/µl -20˚C
KpnI	NEB	10,000 U/ml (R0142S) -20˚C

L-Arginine	Fluka	>99.0% RT
LB agar (Lennox)	Invitrogen	RT
LB broth (Miller)	Merck/VWR	RT
L-Citrulline	Fluka	>99.0% RT
L-Cysteine hydrochloride monohydrate	Sigma-Aldrich	4˚C
Low-range rainbow molecular weight markers	GE Healthcare	-20˚C

meta-Phosphoric acid	Sigma	33.5-36.5% RT
Methotrexate	VWR/Applichem	4˚C
Minocycline hydrochloride	VWR/Calbiochem	4˚C

NdeI	NEB	20,000 U/ml (R0111S)
NEB 5-alpha F'I^q chemically competent *E. coli*	NEB	-80˚C
Nitrocelllose membrane	1) GE Healthcare 2) Whatman	1) Hybond-ECL, 0.45 µm, 6 x 8 cm, 50 sheets 2) Protran, 0.1 µm, 30 cm x 3 m
NovaBlue Singles chemically competent cells	Merck	-80˚C
Novex Tricine Gels	Invitrogen	RT or 4˚C

NuPAGE Novex Bis-Tris Gels	Invitrogen	RT or 4˚C

PAD from rabbit skeletal muscle	Sigma	P1584 -80˚C
pGEM®-T Easy Vector System I	Promega	-20˚C
pET-49b(+) DNA	Novagen	-20˚C
pGEX-6P-3 vector DNA	GE Healthcare	-20˚C
P. gingivalis genomic DNA from strain W83	LGC Promochem	#ATCC BAA-308D-5 Lot Number 4726393
Phusion High-Fidelity DNA Polymerase	Finnzymes	-20˚C
Phusion Hot Start High-Fidelity DNA Polymerase	Finnzymes	-20˚C
PreScission Protease	GE Healthcare	-20˚C
PVDF Membrane	Invitrogen	0.2 µm RT

RC DC Protein Assay	Bio-Rad	RT
rLysozyme solution	Novagen	-20˚C
Rosetta Competent Cells	Novagen	-80˚C

SacI	NEB	20,000 U/ml (R0156S)
SnakeSkin Pleated Dialysis Tubing	Thermo Scientific	10,000 molecular weight cut off 22 mm x 35 feet dry diameter
Sodium chloride	BDH/VWR	RT
Sodium dihydrogen orthophosphate 1-hydrate (NaH_2PO_4 . H_20)	BDH/VWR	RT
Sodium hydrogen sulfite (sodium bisulfite)	Sigma	RT
StyI	NEB	10,000 U/ml (R0500S) -20˚C
Sucrose	BDH/VWR	RT
Sulfasalazine	VWR/Calbiochem	4˚C
swine anti-rabbit immunoglobulins, HRP-conjugated	DAKO	Dilution: 1:5000 4˚C

T4 DNA ligase	NEB	400,000 U/ml (M0202S) -20˚C
Talon Metal Affinity Resin	Clontech	4°C
Taq DNA Polymerase	NEB	5,000 U/ml (M0273G) -20˚C
Tergitol NP-40	Sigma-Aldrich	70% solution
Tetracycline hydrochlorid	VWR/AppliChem	RT
Thiosemicarbazide	Fluka	Store in poisons cabinet RT
TMB	KPL	4˚C

Tris-HCl (Trizma hydrochloride)	Sigma	Reagent grade RT
Tryptone	Fisher Scientific	large granules RT

Urea	VWR/BDH	AnalaR

Quick Ligation kit	NEB	-20˚C
QIAprep spin miniprep kit	Qiagen	RT
QIAquick gel extraction kit	Qiagen	RT
QIAquick PCR purification kit	Qiagen	RT

Virkon Tablets	Du Pont	Rely+On RT

X-Ray Film	Fuji/GRI Ltd	SuperRX, 100NIF, 18x24

Yeast Extract	Fisher Scientific	RT

Solutions and Buffers

Common molecular biology solutions and buffers

Tris-HCl, 1M

To make 50 ml of a 1 M stock solution, 7.88 g Tris-HCl (157.60 g/mol) were dissolved in double deionised water, pH adjusted to the desired value (7.0-9.2), and filter-sterilised using a 0.2 µm membrane. Store at RT or 4˚C.

1x PBS

1x PBS was made up from a commercially available 10x PBS stock by diluting the required volume with ddH_2O.

1x PBS-T

1x PBS-T was made up from a commercially available 10x PBS stock by diluting the required volume with ddH_2O and adding 500 µl Tween-20. Store at 4˚C.

1x TBS

50 mM Tris-HCl

150 mM NaCl

For 1 L of 1x TBS, 7.88 g of Tris-HCl and 8.77 g of NaCl were combined in 800 mL ddH_2O. pH was adjusted to 7.6 with HCl and volume was adjusted to 1L with ddH_2O.

1x TBS-T

50 mM Tris-HCl

150 mM NaCl

0.05% Tween-20

For 1 L of 1x TBS, 7.88 g of Tris-HCl and 8.77 g of NaCl were combined in 800 mL ddH_2O. pH was adjusted to 7.6 with HCl and volume was adjusted to 1L with ddH_2O. 500 µl Tween-20 were added. Store at 4˚C.

25x TAE

For a 25x stock of TAE buffer, 242 g of Tris base were dissolved in 500 ml of ddH_2O, 57.1 ml of glacial acetic acid and 100 ml of 0.5 M EDTA (pH 8.0) were added and the volume was adjusted to 1 L with ddH_2O. For electrophoresis, 1x TAE buffer was prepared by dilution in ddH_2O. Store at RT.

TE buffer

10 mM Tris-HCl, pH 8.0

1 mM EDTA, pH 8.0

For 50 ml TE buffer, 500 µl of 1 M Tris-HCl, pH 8.0, and 100 µl 0.5 M EDTA, pH 8.0 were added to 49.4 ml ddH_2O. The solution was filter-sterilised and stored at RT.

IPTG, 1M

To make 100 ml of a 1 M stock solution, 23.8 g IPTG (238.3 g/mol) were dissolved in double deionised water, sterile-filtered using a 0.2 µm membrane and aliquoted à 1 ml. Store at -20˚C.

DTT, 1M

To make 10 ml of a 1 M stock solution, 1.54 g DTT (154.25 g/mol) were dissolved in double deionised water, filter-sterilised using a 0.2 µm membrane and aliquoted à 0.5 ml. Store at -20˚C.

EDTA, 0.5 M

To make 100 ml, 18.61 g EDTA disodium salt were dissolved in 80 ml ddH_2O and the pH adjusted to ~8.0 with sodium hydroxid, until the solution cleared. The final volume was adjusted to 100 ml with ddH_2O. Store at RT.

LB agar, Lennox

For 250 mL of LB agar, sufficient for approximately 20 plates, 8 g of commercially available pre-mixed LB agar Lennox powder were dissolved in 250 mL ddH_2O and autoclaved. Antibiotics were added after the solution was allowed to cool down to <50˚C. Ampicillin was used at 100 µg/mL,kanamycin at15 µg/mL and chloramphenicol was used at 34 µg/mL.

LB broth, Miller

For 1 L of LB broth, 20g of commercially available pre-mixed LB broth granules were dissolved in 1L ddH_2O and autoclaved. Ampicillin was used at 100 µg/mL, kanamycin was used at 15 µg/mL and chloramphenicol was used at 34 µg/mL.

6x DNA loading dye

0.1% (w/v) Bromophenol Blue

40% (w/v) Sucrose

For 10 ml, 10 mg bromophenol blue was added to 4 g sucrose and the volume adjusted to 10 ml with double deionised water. Store at 4˚C.

Sample buffer for SDS-PAGE

4x LDS sample buffer and 2x Tricine sample buffer were purchased from Invitrogen and used according to the manufacturer's recommendations.

Buffers for purification of GST-enolase (human and *P. gingivalis*)

Lysis Buffer

50 mM Tris-HCl, pH 7.6

150 mM NaCl

5 mM $MgCl_2$

25 U/mL Benzonase

1 kU/mL rLysozyme

For 100 ml, mix 5 ml of a 1 M stock of Tris-HCl, pH 7.6, 15 ml of a 1 M stock of NaCl, and 500 µl of a 1 M stock of $MgCl_2$. Add 10 µl Benzonase (250 U/µl stock) and 3.3 µl rLysozyme (30 kU/µl stock). Add ddH_2O to 100 ml and filter-sterilise. Store at 4˚C.

Binding/Wash Buffer

50 mM Tris-HCl, pH 7.6

150 mM NaCl

5 mM $MgCl_2$

For 100 ml, mix 5 ml of a 1 M stock of Tris-HCl, pH 7.6, 15 ml of a 1 M stock of NaCl, and 500 µl of a 1 M stock of $MgCl_2$. Add ddH_2O to 100 ml and filter-sterilise. Store at 4˚C.

Glutathione elution buffer

50 mMTris-HCl

10 mM reduced Glutathione

pH 8.0

For 100 ml, dissolve 0.78 g Tris-HCl and 0.31 g reduced Glutathione in 80 ml ddH_2O, set pH to 8.0 using sodium hydroxid, and adjust to a final volume of 100 ml with ddH_2O. Filter-sterilise and aliquot. Store at -20˚C.

PreScission protease buffer

50 mMTris-HCl pH 7.6

150 mM NaCl

1 mM EDTA

1 mM DTT

For 50 ml, mix 2.5 ml of 1 M Tris-HCl, pH 7.6, 7.5 ml of 1 M NaCl, 100 µl of 0.5 M EDTA, and50 µl of 1 M DTT. Add ddH_2O to 50 ml and filter-sterilise.

Buffers for purification of GST-PPAD

Lysis Buffer

90% (v/v) BugBuster

10% (v/v) Glycerol

25 U/ml Benzonase

1 kU/ml rLysozyme
For 100 mL, mix 90 mL BugBuster and 10 mL glycerol. Add 10 µl Benzonase (250 U/µl stock) and 3.3 µl rLysozyme (30 kU/µl stock). Add ddH_2O to 100 ml. Store at 4ºC.

Binding/Wash Buffer
50 mM Tris-HCl, pH 7.6
150 mM NaCl
For 100 ml, mix 5 ml of a 1 M stock of Tris-HCl, pH 7.6 and 15 ml of a 1 M stock of NaCl. Add ddH_2O to 100 ml and filter-sterilise. Store at 4˚C.

Glutathione elution buffer
see "Buffers for purification of GST-enolase"

PreScission protease buffer
see "Buffers for purification of GST-enolase"

Buffers for purification of His-PPAD and Trx-His-PPAD

Lysis Buffer
see "Buffers for purification of GST-PPAD"

Binding/Wash Buffer
50 mM $NaH_2PO_4 \cdot H_20$
300 mM NaCl
10 mM Imidazole
pH 8.0
For 250 ml, dissolve 1.73 g $NaH_2PO_4 \cdot H_20$, 4.39 g NaCl and 0.17 g imidazole in 230 ml ddH_2O, adjust pH to 8.0 using hydrochloric acid and add double deionised water to a final volume of 250 ml. Keep at 4˚C.

Elution Buffer
50 mM $NaH_2PO_4 \cdot H_20$
300 mM NaCl
250 mM imidazole

pH 7.0

For 100 ml, dissolve 0.7 g NaH_2PO_4 . H_20, 1.75 g NaCl and 1.7 g imidazole in 90 ml double deionised water, adjust pH to 7.0 using hydrochloric acid and add double deionised water to a final volume of 100 ml. Keep at 4˚C.

HRV 3C protease buffer

50 mMTris-HCl pH 7.6

150 mM NaCl

For 50 ml, mix 2.5 ml of 1 M Tris-HCl, pH 7.6 and 7.5 ml of 1 M NaCl. Add ddH_2O to 50 ml and filter-sterilise.

50% Glutathione slurry (for purification of GST- and GST-His-proteins)

The slurry is provided as 75% (v/v) slurry. The "bed volume" equals the volume of 100% (v/v) slurry. 1 ml bed volume holds 5-8 mg protein. The working concentration for pull-down of GST-tag proteins is 50% (v/v) slurry in binding/wash buffer. Typically, for a 3-L culture, 0.5-1ml bed volume (=1-2 ml 50% slurry) was used. To make up 1 ml of 50% slurry, 0.665 ml of 75% slurry were used. To make up the working solution of slurry, the bottle was inverted a few times and the required volume transferred to a 15-ml Falcon using a wide-bore tip (1-ml tip with the lower third cut off). Binding/wash buffer was added to the tube and the tube inverted a few times until the slurry was resuspended. The slurry was then allowed to set on the bench top (10-15 min) or centrifuged at 500 g for 5 min at RT. The supernatant was decanted and the wash repeated three times. After the final wash, the required volume of binding/wash buffer for a 50% slurry was added and kept on ice until use. Prolonged storage should be avoided due to microbial growth.

50% Talon slurry (for purification of His- and Trx-His-proteins)

Talon slurry is provided as 50% (v/v) slurry. 1 ml bed volume holds 5-8 mg protein. The working concentration for pull-down of His-tag proteins is 50% (v/v) slurry in binding/wash buffer. Based on the expected amount of target protein present in the cell lysate, the required volume of well resuspended slurry was transferred to a 15-ml Falcon using a wide-bore tip. To remove ethanol, the slurry was washed four times in binding/wash buffer as described above for glutathione slurry.

PPAD activity buffer (for *P. gingivalis* PAD)

50 mM CHES, pH 9.5
10 mM DTT
For 100 ml, dissolve 1.04 g CHES in 90 ml double deionised water, adjust pH to 9.5 using sodium hydroxid and add double deionised water to a final volume of 100 ml. Keep at 4˚C. Add DTT to a final concentration of 10 mM from a 1M stock (1:100 dilution) just before use.

2x PAD activity buffer (for rabbit PAD)

200 mM Tris-HCl, pH 7.4
20 mM $CaCl_2$
10 mM DTT
For 100 ml, add 20 ml of a 1 M Tris-HCl, pH 7.4, stock solution to 80 ml double deionised water, dissolve 0.22 g $CaCl_2$, add double deionised water to a final volume of 100 ml. Keep at 4˚C. Add DTT to a final concentration of 10 mM from a 1M stock (1:100 dilution) just before use.

Reagents for colorimetric quantification of citrulline

Colour developing reagent (CDR), the working reagent, is assembled from 1 volume of solution A and 3 volumes of solution B just before use. It should be stored at 4˚C in the dark for no more than 1 hour.

Solution A
80 mM 2,3-Butanedione oxime (DAMO)
2 mM Thiosemicarbazide (TSC)
For 200 ml, dissolve 1.62 g DAMO and 36 mg TSC in 200 ml of double deionised water and aliquot à 1-5 ml. Store at -20˚C. Keep at 4˚C for no more than 1 month.

Solution B
3 M 85% H_3PO_4
6 M 96-98% H_2SO_4
2 mM $NH_4Fe(SO_4)_2 \cdot 12H_2O$ (Ammonium ferric sulphate Dodecahydrate)
For 200 ml, pour 90 ml double deionised water into a 250-mL glas bottle on ice, slowly add

40 ml H_3PO_4 under gentle stirring, then slowly add 66 ml H_2SO_4. Leave on ice for a few minutes to cool down and then dissolve 150 mg $NH_4Fe(SO_4)_2 \cdot 12H_2O$. Upon cooling to RT, adjust volume to 250 mL with double deionised water.

Buffers for ELISA analysis

Coating buffer (50 mM carbonate buffer, pH 9.5)
To make 1 L, combine 1.59 g Na_2CO_3, 2.93 g $NaHCO_3$ and 2 ml of a 10% NaN_3 stock. Adjust to 1 L with ddH_2O.

RIA buffer
1% (w/v) BSA
350 mM NaCl
10 mM Tris-HCl (pH 7.6)
1% (v/v) Triton X-100
0.5% (w/v) Na-deoxycholate
0.1% (W/v) SDS
To make 1 L, combine 10g BSA, 20.4 g NaCl, 10 ml of 1 M Tris-HCl (pH 7.6), 10 ml Triton-X-100, 5 g Na-deoxycholate, 10 ml of 10% SDS and ddH_2O to 1L.

DNA cloning and mutagenesis

Primer synthesis and DNA sequencing

DNA oligonucleotide primer synthesis and DNA sequencing was carried out by Eurofins MWG Operon (Ebersberg, Germany). The chosen purification procedure was HPSF (high purity salt free), unless otherwise stated. Primer melting temperatures (T_m, defined as the temperature at which half of all double-stranded DNA is present in a dissociated, single-stranded form) were determined with according to modified Breslauer's thermodynamics (Breslauer *et al.* 1986). An online calculator using this algorithm can be found on the Finnzymes website[10]. Non-matching primer bases (restriction sites, mutations) were excluded from the T_m calculation. Lyophilised primers were reconstituted with TE buffer or ddH_2O to a final concentration of 100 pmol/µL (=100 µM) and a small aliquot further diluted to give a

[10] https://www.finnzymes.fi/tm_determination.html

working concentration of 10 µM. Reconstituted primers were stored at -20°C. Repeated freeze-thaw cycles were avoided as this resulted in primer degradation and unsuccessful PCR reactions.

Agarose gel electrophoresis

Agarose gel electrophoresis was used for separation of DNA fragments. Typically, a 1% (w/v) agarose gel was prepared (separation range: 0.5 – 8 kb). The desired amount of agarose was added to 40-100 ml of water, depending on gel size, and heated in a glas bottle with the lid off in a microwave, until the agarose had fully melted. After cooling to a safe temperature for handling, ethidium bromide (final concentration: 0.5 µg/ml, from a 10 mg/ml stock) were added. The solution was then poured into a gel chamber, a plastic comb inserted, and left to solidify. The gel was then immersed in 1x TAE buffer in the gel chamber. DNA samples were mixed with 6x DNA loading dye and loaded into the gel wells, along with a suitable DNA ladder (Hyperladder I-V). The DNA was separated under constant voltage (80-130V, depending on gel size).

Preparation of glycerol stocks

For long-term storage of *E. coli* cells, 150 µl of autoclaved, sterile glycerol (100%) were added to 850 µl of a logarithmic-phase *E.coli* culture (OD_{600} 0.6-0.8). The vial was then vortexed vigorously to ensure even mixing of the glycerol and the bacterial culture and frozen in liquid nitrogen for storage at -80˚C. A small scraping from the surface of the frozen cells was later used to inoculate liquid medium and the vial returned to the freezer immediately to maintain viability of bacteria.

Quantification of DNA using NanoDrop

To measure the DNA concentration in samples containing purified DNA in buffer, 1 µl of the sample was used on the Nano-Drop Spectrophotometer (ND-1000, Thermo Fisher Scientific) according to the manufacturer's instructions. The ratio of absorbance at 260 nm (proteins) and 280 nm (nucleic acids) was recorded and used as an indicator for the purity of the sample. A value around 1.8 is generally considered as 'pure' for DNA. Lower values indicate protein contamination.

Polymerase chain reaction (PCR)

PCR reactions were performed using Taq Polymerase, KOD Polymerase, Phusion High-Fidelity Polymerase, or Phusion Hot Start High-Fidelity Polymerase, depending on amount of available template DNA and complexity of the amplification reaction. The reactions were set up according to the manufacturer's instructions. Parameters for individual reactions can be found in the respective sections. Thin-walled tubes and PCR thermal cyclers with heated lids were used for all reactions.

Purification of PCR product reactions

Following PCR, a small amount of the reaction product (1-5 µl) was analysed by agarose gel electrophoresis to assess the efficiency and specificity of amplification and confirm the size of the DNA product amplified. When only one product was present, the remaining volume was purified using the QIAquick PCR purification kit to remove DNA polymerase, primers, nucleotides, and salt. When more than one product was present, the complete reaction was loaded onto an agarose gel and, after electrophoresis, the desired product was cut out and gel-purified using the QIAquick gel extraction kit.

Restriction enzyme digestion of DNA

In order to insert a gene of interest into vector DNA, all DNA was digested using appropriate restriction endonuclease enzymes, generating compatible ends on gene inserts and vector DNA. In a typical reaction, between 1 and 5 µg of insert and vector DNA were combined with 1 µl of each restriction enzyme, 10% (v/v) of supplied BSA solution, the supplied buffer, and ddH_2O to 50 µl. Double digests were routinely perfomed, allowing directional insertion of the insert DNA. The reactions were mixed thoroughly and incubated at 37˚C for 3 hours. Vector DNA was also treated with Antarctic Phosphatase to remove the 5' phosphate group, thus preventing self-ligation. 5 units of Antarctic Phosphatase and 10x buffer (1/10 of the reaction volume) were added after 2.5 hr of the incubation period and the reaction then allowed to proceed at 37˚C for a further 30 min. The total reaction volume was then loaded onto an agarose gel and the desired DNA fragments gel-purified following electrophoresis.

Purification of DNA from agarose gels

The desired DNA fragments were excised from the gel in a slice using a clean scalpel and purified using the QIAquick Gel Extraction Kit, following the manufacturer's instructions. DNA was eluted using 30 µl buffer EB or nuclease-free water. To evaluate the purification and estimate the DNA concentration for ligation, 2 µl of the eluted DNA was analysed on an agarose gel.

Ligation of insert into linearised vector DNA

Constructs were generated by the ligation of purified insert and linearised vector DNA. This was done using either NEB T4 Ligase pr NEB Quick Ligase, following the manufacturer's instructions. It was important to mix all buffers thoroughly prior to use. Typically, in a total reaction volume of 20 µl, 50 ng vector DNA was mixed with the insert DNA in a molar vector-to-insert ratio of 1:3. The amount of required insert DNA was calculated using the following formula:

$$ng\ insert\ DNA = \frac{ng\ vector\ DNA\ \times\ kb\ insert\ DNA}{kb\ vector\ DNA} \times ratio\ insert{:}\ vector$$

As a negative control, water instead of insert DNA was added. After careful mixing the reactions were incubated overnight at 16˚C (T4 ligase) or 10 min - 2 hours at RT (Quick Ligase). The ligation reactions were used immediately for transformation or stored at -20˚C.

Ligation PCR

In cases where ligation proved a challenging step and several conditions had to be tested, PCR was used to evaluate the ligation reaction. For this, 1 µl of a 1:10 dilution of the ligation reaction was mixed with 10 µl 5x Phusion HF buffer, 5 µl 2 mM dNTPs, 2.5 µl each of 10 µM primers complementary to the vector DNA upstream and downstream of the insert, 0.5 µl Phusion High-Fidelity Polymerase and nuclease-free water to a total volume of 50 µl. The cycling conditions were adjusted to the particular template and primers and are specified in the relevant sections. The total volume of the PCR reaction was loaded onto an agarose gel. Successful ligation was indicated by the presence of a band corresponding to the expected combined size of insert and vector DNA located between the chosen primers and the absence of such band in the control ligation lacking insert DNA.

Transformation of competent *E. coli* cells with ligated DNA and evaluation

Chemically competent *E. coli* cells (prepared in our laboratory or commercially available cells) were transformed with ligated DNA with the aim of selecting *E. coli* clones harbouring the final DNA construct. Pre-chilled 14-ml BD Falcon polypropylene round-bottom tubes were used thoughout the procedure. Typically, 1-2 µl of the ligation reaction was used to transform 20-100 µl of cells, depending on transformation efficiency of the cells.

Specifically, the following protocols were employed:

Self-prepared NovaBlue cells: 2 µl ligation reaction was mixed with 100 µl cells and incubated on ice for 30 min. Heat-shock was carried out at 42˚C for 45 seconds. Cells were recovered on ice for 2 min. 900 µl SOC was added and the mixture incubated at 37˚C and 220 rpm for 1 hour. 100 µl were spread onto LB-agar plates containing the appropriate antibiotics for selection of cells containing the construct.

Commercial NovaBlue: 1 µl ligation reaction was mixed with 20 µl cells and incubated on ice for 15 min. Heat-shock was carried out at 42˚C for 30 seconds. Cells were recovered on ice for 2 min. 80 µl SOC was added and the mixture incubated at 37˚C and 220 rpm for 1 hour. 10 µl were spread in a pool of 90 µl LB medium onto LB-agar plates containing the appropriate antibiotics for selection of cells containing the construct.

Commercial NEB 5-alpha F'I^q cells: 2 µl ligation reaction was mixed with 50 µl cells and incubated on ice for 30 min. Heat-shock was carried out at 42˚C for 30 seconds. Cells were recovered on ice for 5 min. 950 µl SOC was added and the mixture incubated at 37˚C and 220 rpm for 1 hour. 100 µl were spread onto LB-agar plates containing the appropriate antibiotics for selection of cells containing the construct.

The plates were incubated overnight at 37˚C. The next morning, plates were examined and the number of colonies was estimated and compared to the control plate (*E. coli* cells transformed with vector DNA alone).

Confirmation of successful ligation by plasmid miniprep and analytical digestion

4-12 single clones were picked and used to inoculate 5 ml of LB medium containing the appropriate antibiotics. Cultures were grown at 37˚C and 220 rpm overnight. Cells were then

harvested by centrifugation at 3000 g for 3 min at RT. The supernatant was discarded and the plasmid DNA isolated from the cells using the QIAprep Spin Miniprep Kit in accordance with the manufacturer's instructions. Purified plasmid DNA was eluted in 50 µl EB or nuclease-free water.

To confirm correct ligation and transformation, an analytical digest was performed using restriction enzymes that would generate fragments of known size. To this end, 5 µl of miniprep DNA was mixed with 0.5 µl of each restriction enzyme and the appropriate buffer and incubated for 1 hour at 37˚C, in a total volume of 10 µl. The reaction was then loaded onto an agarose gel and the number and size of the generated DNA fragments was compared with the number and size of the hypothetical fragments for a successful ligation.

To unambiguously verify the identity of the construct and exclude the occurrence of mutations, the constructs were sequenced by a commercial service using primers complementary to the vector DNA ~50 nucleotides upstream and downstream of the insert.

An alternative method to confirm successful transformation of clones was colony PCR, as described in the protocol below.

Confirmation of successful ligation by colony PCR

PCR with primers complementary to the vector DNA upstream and downstream of the insert was performed directly with the colonies of transformed *E. coli* cells to determine which clones contain the insert. To do this, 4-12 colonies were picked from the plate using a sterile pipette tip, making sure to collect as many cells as possible, touched briefly onto a numbered grid on a fresh agar plate (to cultivate a reference colony for plasmid isolation later on), before transferring each colony to a tube containing 50 µl sterile water. The tube was vigorously vortexted and then placed in a heat block at ~99˚C for 5 min to lyse the cells and denature DNAse enzymes. The cell debris was pelleted at 12,000g for 2 min. 10 µl of the supernatant was transferred to a PCR tube and placed on ice. The remaining components were added to the PCR tube according to the recommendations for the DNA polymerase used, resulting in a total volume of 50 µl. As a positive control, 10 ng of uncut vector DNA was used instead of the cell lysates. The cycling conditions were adjusted to the primers and were in accordance with the recommendations for the DNA polymerase used.

After PCR, 15 µl were loaded onto an agarose gel. The presence of the desired construct DNA in a clone was indicated by the presence of a band corresponding to the expected combined size of insert and vector DNA located between the chosen primers. The positive clones were picked from the reference plate and used to inoculate 5 ml of LB medium containing the appropriate antibiotics. Cultures were grown at 37˚C and 220 rpm overnight. Cells were then harvested by centrifugation at 3000 g for 3 min at RT. The supernatant was discarded and the plasmid DNA isolated from the cells using the QIAprep Spin Miniprep Kit in accordance with the manufacturer's instructions. Purified plasmid DNA was eluted in 50 µl EB or nuclease-free water.

Transformation of competent *E. coli* cells with plasmid DNA

In order to amplify previously generated constructs for maintainance and further experiments, chemically competent *E. coli* cells were transformed with plasmid DNA according to the protocols described above, except that 10 ng plasmid DNA were used (1 µl of a 10 ng/µl diluted in nuclease-free water). Plasmid DNA was isolated from overnight cultures using the QIAprep Spin Miniprep Kit as described above.

Construction of recombinant expression vectors

This section describes protocols and procedures specific for the respective constructs. If not otherwise mentioned, the standard procedures as described above were followed. All expression constructs and the encoded recombinant proteins generated in this study are listed in Table 2.1.

Table 2.1 Expression constructs generated in this study. Pg = *Porphyromonas gingivalis*

Construct	**Vector**	**Insert**	**Restriction sites**	**Fusion Protein (tag-linker-protein)**
Peno/pGex	pGEX 6P-3	Pg enolase	*BamHI/XhoI*	GST-peno ^{2}SILKI...NPLAK^{434}
PPAD/pET48	pET-48b(+)	Pg PAD	*KpnI/SacI*	Thioredoxin-His-PPAD ^{2}KKLLQ...MKILK^{556}
mPPAD/pET48	pET-48b(+)	Pg PAD (truncated)	*KpnI/SacI*	Thioredoxin-His-mPPAD ^{44}AFQET...MKILK^{556}
PPAD/pET47	pET-47b(+)	Pg PAD	*KpnI/SacI*	His-PPAD ^{2}KKLLQ...MKILK^{556}
PPAD/pET49	pET-49b(+)	Pg PAD	*KpnI/SacI*	GST-His-PPAD ^{2}KKLLQ...MKILK^{556}
D130A/pET48	pET-48b(+)	Pg PAD with Asp-130 replaced by Ala	*KpnI/SacI*	Thioredoxin-His-PPAD ^{2}KKLLQ...^{130}A…MKILK^{556}
H236A/pET48	pET-48b(+)	Pg PAD with His-236 replaced by Ala	*KpnI/SacI*	Thioredoxin-His-PPAD ^{2}KKLLQ...^{238}A…MKILK^{556}
D238A/pET48	pET-48b(+)	Pg PAD with Asp-351 replaced by Ala	*KpnI/SacI*	Thioredoxin-His-PPAD ^{2}KKLLQ...^{238}A…MKILK^{556}
N297A/pET48	pET-48b(+)	Pg PAD with Asn-297 replaced by Ala	*KpnI/SacI*	Thioredoxin-His-PPAD ^{2}KKLLQ...^{297}A…MKILK^{556}
C351A/pET48	pET-48b(+)	Pg PAD with Cys-351 replaced by Ala	*KpnI/SacI*	Thioredoxin-His-PPAD ^{2}KKLLQ...^{351}A…MKILK^{556}

P. gingivalis enolase in pGEX 6P-3

Primer design and PCR

P. gingivalis genomic DNA from strain W83 was purchased in a dried form and rehydrated overnight at 4˚C with deionised, autoclaved and sterile-filtered water, resulting in a final concentration of 0.1 μg/μl. Primers were designed to add a *BamHI* restriction site at the 5' end and a *XhoI* restriction site at the 3' end of the PCR amplified full-length nucleotide sequence (coding for residues ^{2}SILKI...NPLAK^{434} of *P. gingivalis* enolase, referring to SwissProt entry Q7MTV8) for insertion into the pGEX 6P-3 expression vector, which contains a coding sequence for an amino-terminal GST tag. The expected PCR product size was 1.3 kilobases.

Forward primer (pg-eno_for)	BamHI 5′ AGTT-GGATCC-GAAATTGCAAAGATTATTGGCAG Tm 63°C
Reverse primer (pg-eno_rev)	XhoI 5′ ATC-CTCGAG-TTACACTTTCTTGTAGCCGTATAC Tm 58°C

P. gingivalis enolase nucleotide sequence (GeneID: 2551984)[11]

ATGGAAATTGCAAAGATTATTGGCAGGGAGATCCTCGATTCCAGAGGCAATCCGACGGTTGAGGTAGATGTACACCTGGCCTG
TGGCATAATAGGACGCGCGGCTGTACCCAGCGGAGCTTCCACCGGTGAGAATGAGGCTATCGAGCTTCGCGATCAGGATAAGG
CTCGCTACTGTGGCAAAGGTGTCCTCAAGGCTGTGAAGAACGTAAATGAAGTGATAGATCCCGCACTCTGCGGAATGTCTGTA
TTGGAACAAACAGCTATTGACCGCAAGCTCATCGAATTGGATGGCACCAAAACGAAAAGCAACCTCGGAGCAAATGCCATGCT
GGGCGTATCACTGGCCGTAGCCAAAGCAGCTGCAGCTTATTTGGATATTCCTCTCTACCGATATATCGGCGGTTCGAACACCT
ATGTCCTGCCTGTTCCTATGATGAATATCATCAACGGCGGTTCTCACTCGGACGCTCCTATCGCATTCCAGGAGTTCATGATC
CGCCCTGTGGGAGCATGCTGCTTCCGTGAGGGCTTGCGCATGGGTGCCGAAGTTTTCCATGCACTCAAAAAGGTGCTTCACGA
TCGTGGTCTGAGTACTGCAGTCGGGGATGAAGGAGGTTTTGCTCCTGCCCTGAATGGTACGGAAGATGCTATCGAATCAATCC
TCAAGGCTGTGGAAGCTGCCGGATACGTGCCCGGTAAGGATATTACCATCGCAATGGACTGCGCCTCTTCCGAGTTCTTCAAG
GATGGTATTTACGACTACACTAAGTTCGAAGGCGAAAAGGGCAAGAAACGCTCTATCGATGAGCAGGTGGCTTATCTGACCGA
ACTGGTTGGCAAGTATCCCATCGATTCTATCGAAGACGGTATGAGCGAAAATGACTGGGAAGGATGGAAGAAGCTGACCGTAG
CTTTGGGAGATAAGGTTCAGCTCGTCGGCGACGATCTCTTCGTTACCAACGTGGAATTCCTTCGCCGTGGTATCGCCGAGAAG
TGCGGCAACTCCATCCTTATCAAGGTAAACCAGATTGGTACGCTCACAGAGACCCTGAATGCCATCGAGATGGCACACCGTCA
CGGCTTTACCAGCGTTACTTCGCACCGCTCCGGAGAGACAGAAGACACCACTATCGCAGACATTGCTGTAGCAACCAACTCGG
GACAGATCAAGACGGGATCGCTCAGTCGTACCGACCGTATGGCTAAGTACAACCAACTGCTCCGCATCGAAGAGGAGTTAGGC
CCTTGCGCTGTATACGGCTACAAGAAAGTGTAA

[11] http://www.ncbi.nlm.nih.gov/gene/2551984

P. gingivalis enolase amino acid sequence (SwissProt: Q7MTV8)[12] (ProteinID: AAQ66821.1)[13]

MEIAKIIGREILDSRGNPTVEVDVHLACGIIGRAAVPSGASTGENEAIELRDQDKARYCGKGVLKAVKNVNEVIDPALCGMSV
LEQTAIDRKLIELDGTKTKSNLGANAMLGVSLAVAKAAAAYLDIPLYRYIGGSNTYVLPVPMMNIINGGSHSDAPIAFQEFMI
RPVGACCFREGLRMGAEVFHALKKVLHDRGLSTAVGDEGGFAPALNGTEDAIESILKAVEAAGYVPGKDITIAMDCASSEFFK
DGIYDYTKFEGEKGKKRSIDEQVAYLTELVGKYPIDSIEDGMSENDWEGWKKLTVALGDKVQLVGDDLFVTNVEFLRRGIAEK
CGNSILIKVNQIGTLTETLNAIEMAHRHGFTSVTSHRSGETEDTTIADIAVATNSGQIKTGSLSRTDRMAKYNQLLRIEEELG
PCAVYGYKKV

PCR was performed in a total reaction volume of 50 µl. The components were assembled on ice in this order: 10x NEB Taq Buffer (final concentration: 1x), ddH_2O, 0.2 µM of each primer, 200 µM of each deoxynucleotide (dATP, dGTP, dTTP, dCTP), 50 ng template DNA (1 µg/ml final concentration), 2.5 U of NEB Taq Polymerase (50 U/ml final concentration). Three different annealing temperatures were used: 55˚C, 59˚C, and 65˚C.

The cycling parameters were:

1 x	94˚C for	2 min
25 x	94˚C for	30 seconds
	55/59/65 ˚C	1 minute
	72˚C	2 minutes
1 x	72˚C	5 minutes
1 x	4˚C	∞

Ligation of PCR product into pGEM®-T Easy Vector

The purified PCR product was inserted via its single 5'-A overhangs into a TA-vector (pGEM®-T Easy Vector) containing complementary single 3'-T overhangs. In a total volume of 10 µl, 50 ng of the vector were mixed with 65 ng of the PCR product (desired molar ratio of vector:insert was 1:3), 3 U of the provided T4 ligase, 2x Rapid Ligation buffer (final concentration: 1x) and ddH_2O to 10 µl. After careful mixing the reaction was incubated overnight at 4˚C.

Transformation of competent *E. coli* cells with the enolase-pGEM construct

4 µl of ligation mix was added to 100 µl chemically competent DH5α E. coli cells in pre-chilled 14-ml Falcon round-bottom tubes. The tubes were incubated for 25 min on ice, then

[12] http://www.uniprot.org/uniprot/Q7MTV8
[13] http://www.ncbi.nlm.nih.gov/protein/AAQ66821.1

for 45 seconds in a 42˚C waterbath, and again on ice for 2 min. 900 µl of LB medium was added and incubated for 1 hour at 37˚C in a shaking incubator at 220 rpm. 150 µl were then plated on LB-Ampicillin plates and incubated overnight at 37˚C.

Isolation of plasmid DNA from transformed *E. coli* cells

6 colonies were picked from the plate and used to inoculate 5 ml of LB-Ampicillin medium each. After overnight incubation at 37˚C, 220 rpm, plasmid DNA was isolated from the cultures using the QIAprep spin miniprep kit according to the manufactorer's instructions.

Analytical restriction digest of isolated enolase-pGEM construct

To confirm successful insertion of the enolase PCR product into the pGEM vector, a small-scale restriction digest and subsequent agarose gel electrophoresis were performed. 5 µl of miniprep DNA was mixed with 1 µl NEB3 buffer, 1 µl BSA (final concentration: 1 µg/ml), 1 µl *NotI* and 2 µl ddH_2O, in a total volume of 10 µl. The reaction was incubated for 3 hours at 37˚C. 2 µl 6x DNA sample buffer were added to the reaction and visualized on a 0.8% agarose gel with ethidium bromide. Positive clones were sequenced to confirm the correct nucleotide sequence and reading frame. Sequencing primers were standard T7 forward (5'-TAATACGACTCACTATAGGG) and SP6 reverse (5'-CATTTAGGTGACACTATAG) primers.

Subcloning into the pGEX-6P-3 expression vector

1.5 µg of enolase-pGEM construct and 1.5 µg of pGEX-6P-3 DNA were digested separately with the restriction enzymes *XhoI* and *BamHI* according to the standard protocol described above.

All further procedures leading to the generation of the final expression constructs were carried out according to the standard protocols described above.

Constructs were sent for nucleotide sequencing using pGEX forward and reverse sequencing primers (standard primers provided by sequencing service; "pGex for": ATAGCATGGCCTTTGCAGG and "pGex rev": GAGCTGCATGTGTCAGAGG) to confirm correct amplification and insertion of the full-length enolase sequence.

Human α-enolase in pGEX 6P-3

A pGEX-6P3 expression construct containing the full nucleotide sequence for human α-enolase (GeneID: 2023)[14], inserted between the *BamHI* and *XhoI* restriction sites, was constructed by Dr. Andrew Kinloch (The Kennedy Institute of Rheumatology) and the experimental procedures have been published (Kinloch *et al.* 2008) (co-authored paper in Appendix).

Full and mature P. gingivalis PAD in pET-48b(+)

Primer design and PCR

Primers were designed to add a *KpnI* restriction site at the 5' end and a *SacI* restriction site at the 3' end of the PCR amplified full-length nucleotide sequence (coding for residues ^{2}KKLLQ...MKILK^{556} of *P. gingivalis* PAD, referring to SwissProt entry Q9RQJ2), for insertion into the pET-48b(+) expression vector coding for an amino-terminal thioredoxin and His tag. The expected PCR product size was 1.7 kilobases.

Forward primer (PAD_full_F):

```
            KpnI
5′ CATATC-GGTACC-TGAAAAAGCTTTTACAGGCTAAAGCCTTGATTC
Tm 72°C
```

Reverse primer (PAD_full_R):

```
              SacI
5′ TCAAATAA-GAGCTC-TTATTTGAGAATTTTCATTGTCTCACGGATTC
Tm 70°C
```

For amplification of the hypothesised "mature" form of *P. gingivalis* PAD (residues ^{44}AFQET...MKILK^{556}), the following forward primer was used, along with the same reverse primer as above (PAD_full_R):

Forward primer (PAD_mat_F):

```
                        KpnI
5′ GAAGAGA-GGTACC-TGGCATTCCAGGAAACGAATC
Tm 70°C
```

[14] http://www.ncbi.nlm.nih.gov/gene/2023

P. gingivalis PAD nucleotide sequence (GeneID: 2552184)[15]

ATGAAAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATCGCCCAAACGCAAATGCA
AGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGC
GTGCTATCGCTGAGTACGAACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAGCTG
GCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTTATAACCCAGTACACCCAAAGCGGTGT
GAATCTCTCTAATTGCGATTTCATCATTGCGAAAACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACG
ATACGAACAAAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAATACGAAGCACAA
TATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGCAACTACATGACGGACGGATATGGATCCGCTGTGCA
GTCACATATCGCATATACGGAGAACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC
ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAGTATTTGGCACCGAACAAAATC
CTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGC
ATGGGGAACGAAGTACGAGGTATATCGCGCTTTGGCCACCAATGAA**C**AACCGTACACGAACTCTCTGATTCTGAACAACAGGG
TATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACGTCTATAAGACGGCAATGCCCGGTTACGAAATT
ATAGGTGTCAAAGGGGCTTCAGGAACACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGG
CTATCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAAGCAGATGTCGTCTCATGCG
CCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATCAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAA
TCAACAGGTCACTATACTTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGACAATAG
TGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACGTGTATGAACGAAACCAATACATGTACTG
TGACCGGAGCTGCCAAAGCTCTTCGTGCATGGTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCC
GGCACATATCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTAGCAGGGACGAGTGT
CTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTGGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAA
TTCTCAAATAA

P. gingivalis PAD amino acid sequence (Protein ID: NP_905579.1)[16] (SwissProt: Q9RQJ2)[17]

MKKLLQAKALILALGLFQLPAIAQTQMQADRTNGQFATEEMQRAFQETNPPAGPVRAIAEYERSAAVLVRYPFGIPMELIKEL
AKNDKVITIVASESQKNTVITQYTQSGVNLSNCDFIIAKTDSYWTRDYTGWFAMYDTNKVGLVDFIYNRPRPNDDEFPKYEAQ
YLGIEMFGMKLKQTGGNYMTDGYGSAVQSHIAYTENSSLSQAQVNQKMKDYLGITHHDVVQDPNGEYINHVDCWGKYLAPNKI
LIRKVPDNHPQHQALEDMAAYFAAQTCAWGTKYEVYRALATNEQPYTNSLILNNRVFVPVNGPASVDNDALNVYKTAMPGYEI
IGVKGASGTPWLGTDALH**C**RTHEVADKGYLYIKHYPILGEQAGPDYKIEADVVSCANATISPVQCYYRINGSGSFKAADMTME
STGHYTYSFTGLNKNDKVEYYISAADNSGRKETYPFIGEPDPFKFTCMNETNTCTVTGAAKALRAWFNAGRSELAVSVSLNIA
GTYRIKLYNTAGEEVAAMTKELVAGTSVFSMDVYSQAPGTYVLVVEGNGIRETMKILK

PCR for both full-length and "mature" form was performed in a total reaction volume of 50 μl containing 50 ng W83 genomic template DNA (1 μg/ml final concentration), 500 nM of each primer, 200 μM of each dNTP, 1 unit Phusion High-Fidelity Polymerase (20 U/ml final concentration), 1x HF Buffer, 3% DMSO and deionised, autoclaved and sterile-filtered water to a total volume of 50 μl.

The cycling parameters were:

1 x	98˚C	30 sec
25 x	98˚C	10 seconds
	58˚C	30 seconds
	72˚C	60 seconds
1 x	72˚C	5 minutes
1 x	4˚C	∞

[15] http://www.ncbi.nlm.nih.gov/gene/2552184
[16] http://www.ncbi.nlm.nih.gov/protein/NP_905579.1
[17] http://www.uniprot.org/uniprot/Q9RQJ2

Restriction digestion of PCR product and expression vector

1 µg of pET-48b(+) vector, 2.3 µg of full-length PAD PCR poduct and 1.6 µg of "mature" PAD PCR poduct were digested separately with the restriction enzymes *KpnI* and *SacI* according to the following protocol: In a total volume of 50 µl, the DNA was mixed with NEB1 Buffer (final concentration 1x), BSA (final concentration 1x), 1 µl *SacI* (20 U), 1.5 µl *KpnI* (15U), and ddH_2O to 50 µl. The reactions were incubated for 3 hours at 37˚C.

Ligation

The digested and purified PCR products were ligated between *KpnI* and *SacI* restriction sites of the digested pET48b(+) vector with a molar insert:vector ratio of 4 (75 ng vector, 91 ng full-length PPAD, 80 ng "mature" PPAD). NEB T4 ligase was used. The ligation mix was incubated overnight at 16°C.

Ligation PCR with T7 promoter (5'-TAATACGACTCACTATAGGG) and T7 terminator primer (5'-CTAGTTATTGCTCAGCGGT) was used to control for successful ligation prior to proceeding with transformation. Phusion High-Fidelity Polymerase in a total volume of 50 µl was used according to the manufacturer's instructions.

The cycling parameters were:

1 x	98˚C	30 sec
30 x	98˚C	10 seconds
	55˚C	30 seconds
	72˚C	60 seconds
1 x	72˚C	5 minutes
1 x	4˚C	∞

The expected product size for full-length PPAD was 2.5 kb and 2.3 kb for "mature" PPAD. A lack of insert in a circular vector was expected to yield a band of 0.8 kb.

Transformation, colony PCR, and analytical digest

Supercompetent NEB 5-alpha F*'I*^{q} *E. coli* cells were used for transformation according to the standard protocol described above.

Colony PCR with T7 promoter (5'-TAATACGACTCACTATAGGG) and T7 terminator primer (5'-CTAGTTATTGCTCAGCGGT) was used to select clones containing the expression construct. All steps were performed as described for the standard protocol above. Phusion High-Fidelity Polymerase in a total volume of 40 µl was used according to the manufacturer's instructions.

The cycling parameters were:

1 x	98˚C	30 sec
35 x	98˚C	10 seconds
	55˚C	30 seconds
	72˚C	60 seconds
1 x	72˚C	5 minutes
1 x	4˚C	∞

Constructs were sent for nucleotide sequencing using a forward primer annealing ~ 50 bp upstream of the PPAD insert (5'-TAACAATCCTCCTACTCCTAC) and a reverse primer annealing ~ 50 bp downstream of the PPAD insert (5'-GTCCATGTGCTGGCGTTC).

P. gingivalis PAD in pET-47b(+)

The insert from the *P. gingivalis* PAD/pET-48b(+) construct was excised using restriction enzyme digestion with *KpnI* and *SacI* according to the standard protocol. The pET-47b(+) vector was linearised using the same enzymes. The purified insert and vector DNA was ligated using NEB T4 ligase in a molar vector-to-insert ration of 1:4. Supercompetent NEB 5-alpha F'I^q *E. coli* cells were used for transformation. Colony PCR was performed as described for the pET48b(+) construct above.

P. gingivalis PAD in pET-49b(+)

The insert from the *P. gingivalis* PAD/pET-48b(+) construct was excised using restriction enzyme digestion with *KpnI* and *SacI* according to the standard protocol. The pET-49b(+) vector was linearised using the same enzymes. The purified insert and vector DNA was ligated using NEB Quick Ligase in a molar vector-to-insert ration of 1:4. Supercompetent NEB 5-alpha F'I^q *E. coli* cells were used for transformation.

Site-directed mutagenesis of *P. gingivalis* PAD

Expression constructs encoding PPAD with single amino acid substitutions were created to assess the contribution of these amino acids to enzymatic activity. A list of the generated constructs is presented in Table 2.1.

Site-directed mutagenesis was performed in four steps:

1. PCR amplification: Point mutations were created by designing a mismatch in the forward primer. The reverse primer was designed to anneal back-to-back with its 5'-end to the 5'-end of the forward primer. A deletion was created by designing primers that border the deleted area on both sides with their 5'-ends (Fig. 2.1). The PCR resulted in an amplified linear target plasmid.

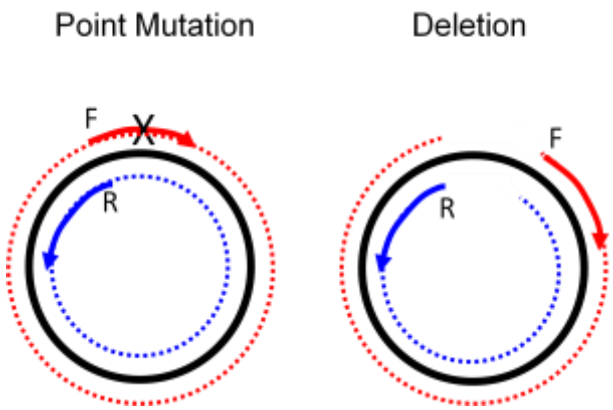


Figure 2.1 Schematic representation of primer annealing sites for point mutation and deletion. "F" and "R" stands for forward and reverse primer, respectively. "X" represents the mismatched nucleotide(s). The primers are drawn in the 5'-3' direction.

2. Ligation: The plasmid was re-circularised by ligation.

3. Transformation: *E.coli* cells were transformed with the re-circularised plasmid. Background from the parent plasmid was expected to be low due to the exponential PCR amplification of the mutation/truncated plasmid.

4. Sequencing: Successful mutagenesis was confirmed by DNA sequencing.

PCR amplification

Ultra-pure primers (commercially purified by polyacrylamid gel electrophoresis, HYPUR-grade from MWG Eurofins) with 5'-phosphorylation, which is crucial for re-circularisation of plasmids in step 2, were used (Table 2.2).

Table 2.2 Primers used for site-directed mutagenesis of PPAD.

Mismatched nucleotides are underlined. Primers are written in the 5'-3' direction.

Construct	**Forward primer**	**Reverse primer**
D130A/pET48	actggacacgcgcctataccggtt	aagagtcagttttcgcaatgatgaaatc
H236A/pET48	gcgaatatatcaacgccgtggactgttgg	cgttcggatcttgtaccacatcatg
D238A/pET48	caaccatgtggcctgttgggg	atatattcgccgttcggatcttgtacc
N297A/pET48	aacaaccgtacacggcctctctgattctg	cattggtggccaaagcgc
C351A/pET48	atgccctgcatgcccgtactcacgag	ctgttcctaaccaaggtgttcctgaag

PCR was performed using Phusion Hot Start High-Fidelity DNA Polymerase in a total volume of 50 µl, with HF buffer, 200 µM dNTPs, 0.5 µM of each primer, 1 ng template DNA (PPAD/pET48) and 0.5 µl Polymerase.

A two-step PCR protocol was employed for all constructs except D238A/pET48, as it proved unsuccessful for this construct. The cycling parameters were:

All except D238A/pET48:

1 x	98˚C	30 sec
25 x	98˚C	10 seconds
	72˚C	3.5 minutes
1 x	72˚C	10 minutes
1 x	4˚C	∞

D238A/pET48:

1 x	98˚C	30 sec
25 x	98˚C	10 seconds
	66˚C	10 seconds
	72˚C	3 minutes
1 x	72˚C	10 minutes
1 x	4˚C	∞

5 µl of the PCR product were loaded onto an agarose gel to verify successful amplification and estimate the DNA concentration for ligation. The expected product size was 7305 bp.

Ligation

25 ng PCR product were mixed with nuclease-free water to a final volume of 5 µl. 5 µl 2x NEB Quick Ligase buffer were added and mixed. 0.5 µl of NEB Quick Ligase was added, mixed thoroughly and incubated on the bench for 1 hour.

Transformation

20 µl super-competent NovaBlue cells and 1 µl of the ligation reaction were used according to the standard protocol as described above. Four colonies for each construct were picked and grown as described previously. Plasmid DNA was isolated as described previously.

Sequencing

Constructs were sent for nucleotide sequencing using a forward primer annealing ~ 50 bp upstream of the PPAD insert (5'-TAACAATCCTCCTACTCCTAC) and a reverse primer annealing ~ 50 bp downstream of the PPAD insert (5'-GTCCATGTGCTGGCGTTC).

Construction of single kgp and double rgp/rgpB *P. gingivalis* mutants

These mutants were a gift from Jan Potempa (Jagiellonian University, Krakow, Poland). The general procedure is described elsewhere (Simpsonv *et al.* 2004; Nguyen *et al.* 2007). Briefly, the 1-kb flanking regions surrounding the *rgpA* gene were amplified by PCR and ligated into the multiple cloning site in the pUC19 vector along with an intervening chloramphenicol resistance gene. A similar approach was used for *rgpB*, except that the final construct contained an erythromycin resistance cassette and a truncated version of the *rgpB* gene (Δ410, C-terminal 97 amino acids deleted, resulting in a complete loss of arginine-specific protease activity). For the *kgp* mutant (deletional inactivation) a chloramphenicol resistance cassette was used. The resulting plasmids were electroporated into competent W83 cells and mutants were selected on antibiotic selective media. Genetic integration via a double crossover event was confirmed by PCR using primers annealing outside the zone of integration. The absence of arg-specific proteolytic activity in the double *rgpA rgpB* mutant and lysine-specific proteolytic activity in the *kgp* mutant was confirmed using enzymatic assays (Simpsonv *et al.* 2004; Nguyen *et al.* 2007).

Construction of ppad *P. gingivalis* mutants

These mutants were constructed in collaboration with Ky-Anh Nguyen (Institute of Dental Research, Westmead Centre for Oral Health, Sydney, Australia and Faculty of Dentistry, University of Sydney, Sydney, Australia).

A 1-kb region 3′ to the *P. gingivalis ppad* gene (Genbank accession number 2552184; locus tag PG1424) was amplified by PCR (primers: 5'-GCTCTAGATGGAATCCGTGAGACAATG and 5'-TAAGCATGCGATATTTGTCGGAAGGACTC) for insertion into *XbaI* and *SphI* sites of the pUC19 plasmid (New England Biolabs Inc., USA). An erythromycin resistance cassette *ermF/ermAM* from plasmid pVA2198 was amplified and inserted it into *SmaI* and *XbaI* sites of the modified pUC19 plasmid. The resultant plasmid was modified further by incorporating (i) an amplified 1-kb region 5′ to the *ppad* gene (primers: 5'-AAGAGCTCAAGCACGTAATAAGGACAATGA and 5'-TTATCCCGGGTGTTCCTGAACATATGATAAGATCT) into *SacI* and *SmaI* sites to create the deletional inactivation plasmid construct (pΔ*ppad*) or (ii) the entire *ppad* gene and a 1-kb region 5′ to the gene (primers: 5'-AAGAGCTCAAGCACGTAATAAGGACAATGA and 5'-TTATCCCGGGTGTCTACCTGAGGAGTATTCT) into *SacI* and *SmaI* sites to create the control mutant construct (p*ppad*+) to control for possible polar effects. The correct placement and orientation of the DNA segments were confirmed by sequencing. The modified plasmid constructs were integrated into the *P. gingivalis* W83 or ATCC 33277 genome by a double crossover recombination event by electroporation using standard protocols (Nguyen *et al.* 2007). Erythromycin-resistant clones were subcultured on selective plates and genomic integration confirmed by PCR, using primers from outside of the cloned regions surrounding the *ppad* gene. The absence of the *ppad* gene in the resulting *P. gingivalis* cultures was confirmed using PCR with purified genomic DNA from these strains and primers annealing to the 5' and 3' ends of full-length PPAD (expected product size: 1.7 kb). As a positive control, primers for the *P. gingivalis* 16S gene were used (expected product size: 0.4 kb).

PPAD forward primer: 5' CATATCGGTACCTGAAAAAGCTTTTACAGGCTAAAGCCTTGATTC
PPAD reverse primer: 5' TCAAATAAGAGCTCTTATTTGAGAATTTTCATTGTCTCACGGATTC

P.g. 16S forward primer: 5' AGGCAGCTTGCCATACTGCG
P.g. 16S reverse primer: 5' AGCGAAAACTGTTAGCAACTACCGATGT

PCR was performed using 50 ng template DNA and KOD polymerase according to the manufacturer's recommendations. Annealing temperature for both primer pairs was 60˚C.

Expression and purification of recombinant proteins

Small-scale protein expression and solubility testing

The expression of the recombinant proteins encoded on the generated constructs was tested in a small scale first to determine whether the protein is expressed and whether it is expressed in a soluble form. Expression parameters (*E. coli* host strain, temperature, IPTG concentration, optical density at induction) were optimised if necessary.

Transformation of E. coli expression hosts

Transformation was performed according to these standard protocols:
Self-prepared BL21 and BL21 (DE3) cells: 1 µl of a 10 ng/µl plasmid solution was mixed with 100 µl cells and incubated on ice for 30 min. Heat-shock was carried out at 42˚C for 45 seconds. Cells were recovered on ice for 2 min. 900 µl SOC was added and the mixture incubated at 37˚C and 220 rpm for 1 hour. 100 µl were spread onto LB-agar plates containing the appropriate antibiotics for selection of cells containing the construct.

Commercial BL21, BL21 (DE3) and BL21 (DE3) pLysS cells: 1 µl of a 10 ng/µl plasmid solution was mixed with 20 µl cells and incubated on ice for 15 min. Heat-shock was carried out at 42˚C for 30 seconds. Cells were recovered on ice for 2 min. 80 µl SOC was added and the mixture incubated at 37˚C and 220 rpm for 1 hour. 10 µl were spread in a pool of 90 µl LB medium onto LB-agar plates containing the appropriate antibiotics for selection of cells containing the construct.

Commercial C41 (DE3), C41 (DE3) pLysS, C43 (DE3) and C43 (DE3) pLysS cells: 1 µl of a 10 ng/µl plasmid solution was mixed with 20 µl cells and incubated on ice for 15 min. Heat-shock was carried out at 42˚C for 45 seconds. Cells were recovered on ice for 2 min. 950 µl "Expression Recovery Medium" (provided with cells) was added and the mixture incubated at 37˚C and 220 rpm for 1 hour. 50 µl were spread onto YT-agar plates containing the

appropriate antibiotics for selection of cells containing the construct.

A single colony was picked from the plate and used to inoculate 3-6 ml LB medium containing the appropriate antibiotic for overnight growth as described previously.

Small-scale expression

250-ml concial flasks containing 50 ml LB medium+antibiotic were used, with the lid loosened, to allow sufficient aeration. 250 µl of the overnight culture (dilution 1:200) were transferred to into the flask and the culture grown at 37˚C and 220 rpm. The OD_{600} was monitored regularly with a spectrophotometer, using plain LB medium as blank. Once an OD_{600} value of 0.6-0.8 was reached (~3.5 hours), protein expression was induced by the addition of IPTG and the incubation temperature was set to the required value. As controls, uninduced cells (no IPTG added) and cells transformed with an empty vector (+/- IPTG) were used. The cultures were incubated for 3-5 hours to allow for sufficient protein expression. 250 µl of culture was centrifuged to harvest cells (8.000 rpm, tabletop centrifuge, 5 min, 4ºC) and the pellet resuspended in 25 µl lysis buffer (depending on protein to be expressed, see "*Solutions and Buffers*" and sonicated (Vibra-cell Sonicator, Sonics & Materials Inc) if required. The mix was rotated at RT for 10 min. 25 µl 4x LDS sample buffer was added, the sample heated to 70˚C for 10 min, centrifuged briefly and 10 µl per lane analysed on a SDS-PAGE gel. Expression of recombinant protein was detected by the presence of a prominent band in the IPTG-induced extract but not in the uninduced control.

Solubility testing

To determine whether the recombinant protein detected in the whole cell extract, as described above, can be conveniently purified from the soluble cell extract fraction, the solubility of the recombinant protein had to be evaluated. To this end, 1 g (ca. 25 ml culture) of cell pellet from the induced culture was resuspended in 5 ml lysis buffer and incubated on a rotary mixer for 20 min at RT or 4˚C, depending on protein. The mix was then sonicated, if required. The insoluble fraction was separated from the soluble fraction by centrifugation (16.000g, 20 min, 4˚C or RT). The supernatant was defined as the soluble fraction. The pellet was resuspended in 5 ml ice-cold denaturing lysis buffer (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 8.0) and rotated at RT for 20 min. The sample was then centrifuged (16.000g, 20

min, RT) and the resulting supernatant defined as the insoluble fraction, whereby "insoluble" refers to solubility in native buffers. 7.5 µl of each fraction were analysed on a SDS-PAGE gel. As the recombinant protein present in the soluble fraction typically is not as concentrated as in the total cell extract or insoluble fraction, immunoblotting was used to confirm presence of the recombinant protein.

Variation of parameters affecting expression and solubility of recombinant proteins

In cases were proteins were expressed for the first time and without previous references, such as all PPAD proteins, several *E. coli* strains were tested as host strains. These included BL21, BL21 (DE3), BL21 (DE3) pLysS, C41 (DE3), C41 (DE3) pLysS, C43 (DE3), C43 (DE3) pLysS. Once a suitable strain was found, factors resulting in maximum soluble protein yield were investigated, including expression temperature (25-30˚C), expression time (1-24 hours), IPTG concentration (0.1 - 1 mM) and optical density at induction (OD_{600} = 0.6 - 1.0). as a starting point, for pGEX vectors BL21 cells were used as hosts, with induction at OD 0.6-0.8 with 0.1 mM IPTG, expression at 28˚C for 5 hours. For pET vectors, BL21 (DE3) cells were used, with induction at OD 0.6-0.8 with 1 mM IPTG and expression at 30˚C for 5 hours. When the optimal or sufficient levels of soluble protein yield were achieved, the recombinant protein was expressed in large-scale cultures and purified for further studies.

Large-scale protein expression

A single colony harbouring the construct of interest or a scraping from a glycerol stock of the expression transformed with the construct of interest was used to inoculate 6 ml of LB medium containing the appropriate antibiotic. Overnight cultures were grown at 37˚C, 220 rpm. The following morning, 500 ml LB medium + antibiotic, in a 2-L conical flask, were inoculated with 1 ml overnight culture. Typically, a total of 3 L (6x500 ml) was cultured. The optical density (OD_{600}) of the cultures was monitored regularly. Once an OD_{600} value of 0.6-0.8 was reached, protein expression was induced by the addition of IPTG and the incubation temperature was set to the required value for a further 5 hours (see Table 2.3).

Table 2.3 Expression parameters for large-scale protein expression

E. coli host strains, induction temperature and IPTG concentration used for successful expression of recombinant proteins are listed for each construct. All cultures were induced at OD_{600} 0.6-0.8 and expressed for 5 hours."Heno/pGEX" stands for human α-enolase in pGEX 6P-3.

Construct	***E. coli* host**	**Induction Temperature (˚C)**	**IPTG concentration for induction (mM)**
Peno/pGex	BL21	28	0.1
Heno/pGex	Rosetta	28	0.1
PPAD/pET48	BL21 (DE3)	30	1
mPPAD/pET48	BL21 (DE3)	30	1
PPAD/pET47	BL21 (DE3)	30	1
PPAD/pET49	BL21 (DE3)	30	1
D130A/pET48	BL21 (DE3)	30	1
H236A/pET48	BL21 (DE3)	30	1
D238A/pET48	BL21 (DE3)	30	1
N297A/pET48	BL21 (DE3)	30	1
C351A/pET48	BL21 (DE3)	30	1

The cultures were then transferred to 500-ml centrifuge tubes (Corning Inc.) and centrifuged at 2500 g for 10 min at RT. The supernatants were discarded and the pellets weighted. The pellets were resuspended in 5 ml lysis buffer per g pellet. For recombinant enolase, this was followed by sonication according to the manufacturer's recommendations. For all PPAD proteins, the lysates was simply rotated at RT for 30 min. A 20-µl sample was taken as the "total cell extract (TCE)". The remaining sample was then centrifuged (Sorvall Evolution RC, SA600 rotor) at 16000g, 20 min, 4˚C (enolase) or RT (PPAD). The supernatant was filtered using a 50-ml syringe and 0.45 µm filter. A 20-µl sample was taken as the "soluble cell extract (SCE)". The recombinant protein was then purified from the filtered supernatant using a resin pull-down approach, in which the resin specifically captures the tag present in the recombinant protein.

Pull-down

For His- and Thioredoxin-His-tagged proteins Talon metal affinity resin was used, and for GST- and GST-His-tagged proteins Glutathione Sepharose 4B resin was used. Preparation of the resins prior to use and the required buffers are described in the section *"Solutions and Buffers"*.

The required amount of resin was added to the filtered extract and rotated for 30 min at RT to allow the recombinant protein to bind to the resin. The mix was then transferred to 10-ml Pierce centrifuge disposable columns with the end cap on and the resin allowed to settle. The end cap was removed and the flow-through collected. The resin was then washed 6 x 10 min with 10 ml binding/wash buffer. To elute, 1 bed volume (the volume equivalent to 100% resin) of elution buffer was added, incubated for 5 min at RT, and the flow though collected as eluate fraction 1. This was repeated to a total of 6 times, resltung in eluate fractions 2-6.

In cases were this procedure was carried out for the first time, all collected fractions (total cell extract, soluble cell extract, flow-though, washes (all 5 µl), eluates (2.5 µl) were analysed on a SDS-PAGE gel. For established protocols, flow-though and washes were not analysed.

Tag cleavage

All expression vectors harboured a cleavage site recognised by the HRV 3C protease, resulting in enzymatic cleavage of amino-terminal tag(s). For His- and Thioredoxin-His-tagged proteins His-tagged "HRV 3C protease" from Novagen was used. For GST- and GST-His-tagged proteins GST-tagged "PreScission protease" was used. Both recombinant enzymes recognise the same cleavage site but differ in the tag they are coupled to, allowing their pull-down along with uncleaved protein and fragment. All required buffers are described in the section *"Solutions and Buffers"*.

On-column cleavage with PreScission protease was used for all enolase constructs. Here, the protein was not eluted from the resin. Instead, after the last wash step, the resin was equilibrated in 10 ml PreScission protease buffer. Then, the PreScission protease mix was prepared by mixing 80 µl (160 U) of PreScission protease and 920 µl of PreScission protease buffer for each ml of resin bed volume. PreScission Protease buffer was drained from the column and the mix added to the resin. The column was sealed and incubated overnight at

4˚C on a rotary device. Following incubation, the cleaved protein was eluted in the flow through. To maximise yield, the column was eluted a total of 6 times with 1 bed volume PreScission protease buffer and the cleaved protein collected in the eluate.

Cleavage in solution was used for all PPAD constructs, as tag cleavage proved inefficient and on-column cleavage further educes efficiency. The eluted protein was dialysed into HRV 3C protease buffer. The dialyed protein was mixed with 200 U (100 µl) HRV 3C protease per mg protein in 1 mM DTT, 1 mM EDTA, 0.1% Triton X-100 and HRV-3C buffer to 2.5 ml. The mix was incubated at 4˚C overnight on a rotary device. To remove the His-tagged protease, his-tagged cleavage products and uncleaved protein, 500 µl of 50% Talon resin in HRV 3C buffer was added to the mix and incubated for 20 min at RT. The mix was then transferred to a 2-ml Pierce centrifuge disposable column. The column was placed into a 14-ml Falcon tube and the cleaved protein was collected by centrifugation at 1000g for 1 min at RT. To increase yield, the original volume of HRC 3C buffer (2.5 ml per mg protein) was added to the resin, mixed, and the flow-though collected as described. To recover all bound proteins and fragments, the same volume of imidazole-containing elution buffer (see "*Solutions and Buffers*") was added to the resin, mixed, and recovered by centrifugation as described.

Size exclusion chromatography

Size exclusion chromatography was used as a final purification step for Thioredoxin-His-PPAD for subsequent enzymatic analysis and inhibitor studies. A Sephacryl 200 (XK 16/100 200 cm^3) column (Pharmacia/GE Healthcare) was used coupled to an ÄKTA FPLC P-920 system with an UPC-900 UV monitor and an Frac-950 fraction collector. Unicorn 3.20 software was used. The storage buffer was washed out with deionised water and the column then equilibrated in PBS. The concentrated protein sample (500 µl at 2 mg/ml) was loaded directly onto the column using a syringe. The proteins were then eluted with 200 ml of PBS at 0.5 ml/min (linear flow rate: 15 cm/hour) and a maximum pressure of 0.5 mPa. 1-ml fractions were collected. All steps were performed at RT. UV (280 nm) was plotted against elution volume. The fractions corresponding to peaks were analysed by SDS-PAGE (10-20 µl per 1-ml fraction) and the presence of PPAD activity was assessed by an enzymatic activity assay (30 µl per 1-ml fraction) using BAEE as substrate. The fractions of interest were then concentrated and the protein concentration was determined using the BCA assay.

Protein analysis

Calculation of theoretical protein parameters

Protein parameters such as theoretical molecular weight, amino acid composition and extinction coefficient at 280 nm were determined using the Protparam tool at the Expasy Server (http://www.expasy.ch/tools/protparam.html).

Determination of protein concentration

Samples free of interfering substances (reducing agents, detergents) either Bradford Coomassie Plus Reagent was used, according to the standard microplate protocol, or the BCA Protein Assay Kit, according to the microplate procedure. For samples containing detergents, the BCA Protein Assay Kit was used. For samples containing reducing agents, the Bradford Coomassie Plus Reagent was used. For samples containing both detergents and reducing agents the RC DC Protein Assay was used. Standard curves were prepared with BSA following each of the manufacturer's instructions.

Concentration of protein samples

Proteins were concentrated using Vivaspin columns (GE Healthcare). The cut-off and column size was chosen according to the protein of interest and the sample volume. A 3 K cut-off was used to concentrate the Thioredoxin-His tag and whole bacterial lysates. A 10 K cut-off was used for all enolase proteins. A 30 K cut-off was used for all full-length PPAD proteins. All steps were performed according to the manufacturer's instructions.

Dialysis

Proteins were dialysed using either 3-ml Slide-A-Lyzer Dialysis cassettes, 10 K MWCO or, for larger volumes, SnakeSkin Pleated Dialysis Tubing, 10 K MWCO (both from Pierce). Typically, dialysis was performed in 3 changes of buffer with least 100x sample volume for 2 hours at RT and 1 change of buffer with at least 500x sample volume overnight at 4˚C.

In vitro citrullination of recombinant proteins

Citrullinated proteins were often required as a positive control in analysis of protein citrullination of unknown samples (for example in *P. gingivalis* cell lysates). To generate citrullinated proteins *in vitro*, equal volumes of recombinant protein and 2x PAD buffer (200 mM Tris-HCl, pH 7.6, 20 mM $CaCl_2$, 10 mM DTT) were mixed and incubated with rabbit skeletal PAD at a concentration of 8.75 U per mg substrate protein, for 3 hours, at 37°C. Citrullination was terminated by the addition of EDTA (final concentration 10 mM). As a control reaction, deionised water instead of PAD was added. All samples were stored at -20°C until further use.

SDS-Polyacrylamide gelectrophoresis (SDS-PAGE)

Pre-casted minigels (Invitrogen) were used for all PAGE applications. 4-12% or 12% NuPAGE Bis-Tris gels, with 10, 12, or 15 wells per gel, were used for resolution of standard protein samples (20-200 kDa), together with 4x LDS sample buffer, 20x MOPS running buffer (all from Invitrogen). 10-20% Novex Tricine gels, with 10 wells per gel, were used for resolution of peptide samples (2.5-50 kDa), together with 2x Tricine sample buffer and 10x Tricine running buffer. The sample buffers do not contain reducing agents; therefore, DTT was added to a final concentration of 50 mM to all samples just before use. Samples for Bis-Tris gels were heated at 70°C for 10 min, samples for Tricine gels were heated at 85°C for 3 min. When handling cell lysates of *P. gingivalis* and other oral bacteria, the samples were heated in non-reducing sample buffer, allowed to cool down on ice for 2 min, before adding DTT to a final concentration of 50 mM. This procedure was found to minimise protease activation and protein degradation. Bis-Tris gels were resolved at 180 V, Tricine gels were resolved at 80V. Protein molecular mass markers were Full-range rainbow molecular weight markers or low-range rainbow molecular weight markers.

Immunoblotting analysis

After electrophoresis, membranes were assembled in a sandwich using the XCell Sure Lock (Invitrogen) module, according to the manufacturer's instructions. Nitrocellulose membranes, pre-soaked in transfer buffer, were routinely used. The transfer buffer (1 L) was composed of 50 ml NuPAGE 20x transfer buffer (Invitrogen), 750 ml deionised water and 200 ml methanol. The transfers were performed either for 2 hours at 30 V, RT or overnight at 12 V,

4˚C. After transfer, membranes were transferred to platic containers and blocked in 5% (w/v) non-fat dried milk in TBS-T or PBS-T for 2 hours at RT or overnight at 4˚C, under constant agitation. The primary antibody was diluted in TBS-T or PBS-T (details for the individual antibodies can be found in the "*List of Materials and Reagents*"). The blocking milk was discarded and the diluted primary antibody added to the membrane and incubated for 1 hour at RT or overnight at 4˚C. The membrane was then thoroughly washed (1x 15 min then 3x 5 min, RT, agitation) with TBS-T or PBS-T. The secondary HRP-conjugated antibody was diluted in in TBS-T or PBS-T, added to the membrane and incubated for 1 hour at RT. The membrane was again thoroughly washed (1 x 15 min then 3 x 5 min, RT, agitation) with TBS-T or PBS-T. The bound secondary antibody was detected using ECL or ECL Plus (3 ml per minigel membrane) according to the manufacturer's instructions.

Experimental details for detection of individual proteins:

His-tagged proteins

Non-fat dried milk was used at 10%. Mouse anti-his antibody was diluted at 1:1000. Proteins were detected using HRP-conjugated goat anti-mouse IgG, diluted 1:5000. Membranes were developed using the ECL.

Citrullinated proteins

Citrullinated proteins were detected using the anti-citrulline (modified) detection kit in accordance with the manufacturer's instructions. Briefly, after electrophoresis and transfer of proteins onto nitrocellulose membranes, membranes were incubated in 3 ml solution A (25% v/v H_2SO_4, 20% v/v H_3PO_4, 0.025% w/v $FeCl_3$) mixed with 3 ml solution B (0.5% w/v 2,3-butanedione monoxime, 0.25% w/v antipyrine, 0.5 M acetic acid) at 37°C for 7 hours. This step modifies the guanidino group of the citrulline side-chain into a pyrimidine ring derivative. The membranes were then processed in a standard immunoblotting procedure. Non-fat dried milk was used at 5%. Rabbit anti-citrulline (modified) antibody was diluted at 1:1000 and incubated overnight at 4˚C. Proteins were detected using HRP-conjugated goat anti-rabbit IgG, diluted 1:5000. Membranes were developed using the ECL Plus. Controls were performed in which the modification step or the primary antibody were omitted, to control for non-specific binding by the primary antibody to structures other than modified citrulline side chains, or binding of the secondary antibody to proteins other than the primary antibody, respectively. For the modification control, membranes were incubated in 3 ml

solution A and 3 ml 0.5 M acetic acid, resulting in a mix lacking the guanidino-group modifying chemicals.

Fibrinogen

Non-fat dried milk was used at 5%. Membranes were incubated with a mixture of three anti-fibrinogen antibodies (goat anti-fibrinogen α, anti-fibrinogen β, anti-fibrinogen γ; each diluted 1:500). Proteins were detected using HRP-conjugated donkey anti-goat antibody, diluted 1:3000. Membranes were developed using ECL Plus.

PPAD

Non-fat dried milk was used at 5%. Rabbit anti-PPAD antibody (TEA eluate) was diluted at 1:500 (cell lysates) or 1:1000 (recombinant protein). Proteins were detected using HRP-conjugated swine anti-rabbit Ig, diluted 1:5000. Membranes were developed using the ECL Plus.

CEP-1

Non-fat dried milk was used at 5%. Rabbit anti-CEP-1 antibody was diluted at 1:200. Proteins were detected using HRP-conjugated goat anti-rabbit IgG, diluted 1:2000. Membranes were developed using the ECL Plus.

Dotblot analysis

For dot blotting, 10 μL of protein sample was spotted directly onto an nitrocellulose membrane equilibrated in transfer buffer. The sample was allowed to dry before proceeding with the standard immunoblotting protocol as outlines above (starting with the blocking step).

Protein visualisation

Coomassie

InstantBlue reagent, a ready-made, non-toxic Coomassie-based stain, was used for staining of protein bands on polyacrylamide gels. The bottle was inverted a few times, and 10 ml solution per gel were added to a plastic dish containing the gel and stained with gentle agitation at RT until the protein bands reached the desired intensity (15-60 min). To reduce

background staining, the gel was then washed in ddH_2O two to three times for 15 min each.

Silver staining

Silver staining was used for polyacrylamide gels when a lower detection limit was required (5-10 ng/band). The following solutions were made up as stock solutions: A) 50% (v/v) methanol, 5% (v/v) acetic acid; B) 50% (v/v) methanol; C) 0.02% (w/v) sodium thiosulfate pentahydrate ($Na_2S_2O_3 \cdot 5H_20$); F) 5% (v/v) acetic acid. The following solutions were made up fresh (50 ml/gel): D) 0.2% (w/v) silver nitrate ($AgNO_3$), 50 ml/gel, avoid exposure to light; E) 2% (w/v) sodium carbonate (Na2CO3), add 0.05 % (v/v) formaldehyde (37%) just before use, 150 ml/gel. The plastic dish to be used for staining was cleaned thoroughly with washing up liquid and then rinsed with methanol. It was important to avoid touching the gel or any materials without gloves as this leaves marks on the gel. Instead, forceps with flat ends were used for handling the gel. All steps were carried out with gentle agitation at RT, unless stated otherwise. The protein bands were fixed in A for at least 30 min. The gel was washed in B for 30 min, and then three times in ddH_2O for 10 min each. The gel was sensitised in C for 1 min, washed in ddH_2O twice for 1 min each, and then stained in D for 20 min, at 4°C in the dark. The gel was washed in ddH_2O twice for 1 min each. Protein bands were developed in E. It was important to discard solution E after it turned yellow (1-5 min) and add fresh solution E. When the bands had reached the desired intensity, the reaction was stopped with addition of F for 10 min. Finally, the gel was washed in ddH_2O for 15 min.

Ponceau S staining

Ponceau S staining was used to reversibly stain protein bands on nitrocellulose membranes after transfer. 0.1% (w/v) Ponceau S in 5% acetic acid was used. The membrane was incubated for 1-5 minutes in the solution until the bands became visible. The stain was removed by washing the membrane in deionised water with frequent changes.

Growth and preparation of oral bacterial strains

P. gingivalis and other oral bacteria cultures

Porphyromonas gingivalis wild-type strains (W83, W50, ATCC 33277), *P. gingivalis* clinical isolates (MaRL, D243, JH16, J430), obtained from patients with severe periodontitis, and *P. gingivalis* mutants (Δ*ppad*, *ppad+*, Δ*rgp*, Δ*kgp*, Δ*rgp+kgp*) were grown in Schaedler

anaerobe broth (Oxoid, UK), supplemented with 5% sheep blood, at 37°C in an anaerobic chamber (90% N_2, 5% CO_2, 5% H_2). Erythromycin or tetracycline was used at 5 μg/mL or 1 μg/mL, respectively, on solid media. The concentrations were doubled for selective growth in liquid culture. Other anaerobic oral bacteria (*Prevotella intermedia* H13 (clinical isolate), *Prevotella oralis* ATCC 33269, *Capnocytophaga gingivalis* ATCC 33624, *Capnocytophaga ochracea* ATCC 27872) were grown in Schaedler anaerobe broth, supplemented with 2.5 μg/L vitamin K, at 37°C in an anaerobic chamber (90% N_2, 5% CO_2, 5% H_2). *Fusobacterium nucleatum* ATCC 10953 was grown in Schaedler anaerobe broth in an anaerobic chamber with 80% N_2, 10% CO_2 and 10% H_2 at 37^{o}C. *Aggregatibacter actinomycetemcomitans* ATCC 43718 was grown in Tryptic soy broth (Sigma, UK), supplemented with 6% yeast extract and 8% glucose, in 5% CO_2 at 37^{o}C. Aerobic bacteria (*Streptococcus constellatus* ATCC 27823, *Streptococcus gordonii* ATCC 10558, *Streptococcus sanguinis* ATCC 10556, *Streptococcus salivarius* ATCC 7073) were grown on Columbia agar plates, supplemented with 8% defibrinated sheep blood or brain heart infusion broth.

Preparation of bacterial whole protein extracts

All oral bacterial extracts were kindly provided by Aneta Sroka from Jagiellonian University Krakow, Poland, Dr. Sigrun Eick from Bern University, Switzerland, and Shauna Culshaw from Glasgow University, UK.

Bacterial cultures were grown in liquid media until early stationary phase (24 hours). Twenty mL of culture was centrifuged at 10,000 × g for 15 min at 4°C. The resulting bacterial pellet was resuspended in phosphate buffered saline (PBS) and the optical density (at 600 nm) was measured and adjusted to 1.0 with PBS. Samples were sonicated on ice for 5 min (5 sec on – 10 sec off). Sodium azide (final concentration 0.02% v/v) was added to all samples as preservative.

Cellular compartment fractionation of *P. gingivalis*

The procedure was adapted from published protocols (Filip *et al.* 1973; Parker *et al.* 1993). 15 ml of *P. gingivalis* W83 culture was grown as described and the OD_{600} adjusted to 1.5. 15 ml of the adjusted culture were centrifuged at 4,500 x g for 30 min at 4˚C. 1 ml of supernatant was collected as the "vesicle" fraction. A further 5 ml of the supernatant were ultracentrifuged at 150,000 x g for 1 hour at 4˚C and 1 ml of the supernatant was collected as

the “medium” sample. The cell pellet was washed once with 15 ml PBS at 4˚C by resuspension and centrifugation (4,500 x g, 15 min, 4˚C).

From here on, two similar but not identical protocols were used by different people (myself and Yonghua Gho) in different laboratories (Krakow University and Louisville University).

Protocol used for analysis of the subcellular localization of citrullinated proteins *(samples prepared by myself)*

The resulting pellet was resuspended in 5 ml 0.25 M sucrose, 30 mM Tris, pH 7.6. Cells were left mixing gently for 15 min at 4˚C, and then pelleted at 12,500 x g for 15 min at 4˚C. The pellet was rapidly resuspended in 5 ml ice-cold deionised water. Cells were left mixing gently for 15 min at 4˚C. The suspension was then centrifuged at 12,500 x g for 15 min at 4˚C and the supernatant was designated the “periplasmic” sample. The remaining spheroblasts were washed once with 5 ml PBS, centrifuged at 10,000 x g for 15 min at 4˚C and the pellet was resuspended in 5 ml PBS and sonicated (10 times 5 sec pulses with 2 sec rest between each pulse) in an ice-water bath. Cellular debris and membranes were pelleted by ultracentrifugation at 150,000 x g for 1 hour at 4˚C. The supernatant was designated the “cytoplasm” sample. The pellet was washed once with 20 ml PBS and then dispersed by first scraping with the sonication tip in 5 ml ice-cold PBS and then by sonication (3 times 5 sec pulses). Ice-cold sodium-lauryl sarcosine (final concentration 1% v/v) was added and the mixture was left to mix gently for 20 min at 4˚C, and then ultracentrifuged at 150,000 x g for 1 hour at 4˚C. The supernatant was designated as the “inner membrane” sample. The remaining pellet was washed once with 20 ml ice-cold PBS prior to dispersion in 5 ml PBS by sonication as described above and was designated as the “outer membrane” sample. Sodium azide (final concentration 0.02% v/v) was added to all samples. Proteins in the “vesicle”, “medium”, “periplasm” and “cytoplasm” samples were precipitated using a standard trichloroacetic acid (TCA)/acetone protocol. Briefly, equal volumes of protein sample and ice-cold 20% TCA in acetone (v/v) were mixed (10% v/v final TCA concentration) and left at -20˚C overnight. Precipitated proteins were pelleted by centrifugation at 14,000 rpm (Eppendorf Microfuge) for 30 min at 4˚C and washed three times in 200 µl ice-cold acetone. After the final wash, the supernatant was discarded and the protein pellet was air-dried and resuspended in 4x LDS SDS-PAGE sample buffer (Invitrogen).

Protocol used for analysis of the subcellular localization of *P. gingivalis* PAD *(samples prepared by Yonghua Gho)*

The resulting pellet was resuspended in 5 ml PBS and ultrasonicated in an ice-water bath for 10 x 5 sec (17 W per pulse), 3 cycles, with 2 sec rest between each pulse. Cellular debris and membranes were pelleted by ultracentrifugation at 150.000g for 1 hour at 4˚C. The supernatant was designated as the "Periplasmic+Cytoplasmic" (PP/CP) fraction.
The pellet was washed once with 20 ml PBS and then dispersed by first scraping with the sonication tip in 5 ml ice-cold PBS and then by sonication (3 times 5 sec pulses). 0.5 ml were collected as the "Cell Envelope" fraction.
Cold 1% Triton X-100 in 200 mM $MgCl_2$ was added and the mixture was left to mix gently for 30 min at 4˚C to dossilve the inner membrane. Triton X-100 -resistant outer membrane was pelleted by ultracentrifugation at 150000g for 1 hour at 4˚C. The supernatant was designated as the "Inner Membrane" fraction. The remaining pellet was washed once with 20 ml PBS prior to resuspension in 2.5 ml PBS by sonication as described before and was designated as the "Outer Membrane" fraction. Protein content in all fractions, except in the "vesicles" and "medium" fractions due to interfering agents, was determined by BCA assay. All fractions were freeze-dried and subsequently resuspended using 4x LDS sample buffer + 1mM TLCK to a final protein concentration of 2 mg/ml. The "vesicles" and "medium" fractions, originally 1 ml, were resuspended with 300 µl 4x LDS buffer + 1 mM TLCK. Samples were heated for 5 min at 95˚C, aliquoted at stored at -20˚C. Prior to SDS-PAGE, 7 µl of sample was mixed with 7 µl of reducing sample buffer and 14 µl of water (0.5 mg/ml protein concentration), heated for 3 min at 95˚C. 10 µl (5 µg) were loaded per lane.

Analysis of *P. gingivalis*-mediated proteolysis and citrullination

Incubation of *P. gingivalis* with fibrinogen and α-enolase

The OD_{600} of 24 hour-old *P. gingivalis* W83 cultures was measured and noted. Bacteria were pelleted at 4,500 x g for 15 min at 4˚C, washed once in ice-cold PBS, and resuspend in assay buffer (10 mM HEPES, 150 mM NaCl, 1 mM $CaCl_2$, pH 7.5, 10 mM L-Cysteine) to yield a final OD_{600} of 1.0. Human fibrinogen (Sigma, F3879) and recombinant human α-enolase were diluted in assay buffer at a concentration of 0.5 mg/ml. Equal volumes of protein

solution and bacterial cell suspension were mixed and an aliquot immediately withdrawn (corresponds to time point t = 1 min). The cultures were then incubated at 37˚C on a shaking platform, and further aliquots withdrawn after 1.5 h, 3 h and 6 h. Bacterial cells were immediately removed from all aliquots by centrifugation (10,000 x g, 5 min, 4˚C). The resulting supernatant was used for analysis by SDS-PAGE gels and immuno-/dot-blotting as described and for protein precipitation with 15% *meta*-phosphoric acid, leaving small peptides in solution, and subsequent analysis by mass spectrometry.

Protein precipitation

Five-hundred µl of each supernatant was mixed with 80 µl 15% *meta*-phosphoric acid by vortexing and incubated on ice for 10 min. Proteins were removed by centrifugation (14000 rpm, tabletop centrifuge, 10 min, 4˚C). The resulting supernatant was freeze-dried and used for peptide analysis by mass spectrometry.

Mass spectrometry

Performed by Dr. Robin Wait, The Kennedy Institute

In-gel digestion of proteins

Protein bands were excised with a scalpel and in-gel digestion was performed using a robotic system (Investigator ProGest, Genomic Solutions, Huntingdon, UK). The bands were washed in 100 mM ammonium bicarbonate buffer and dehydrated in 100% acetonitrile. Cysteine residues were reduced with 10 mM DTT, then carboamidomethylated with 55 mM iodoacetamide. Digestion was performed for six hours at 37 ºC by addition of modified porcine trypsin (10 µL at 6.5 ng/µL in 25 mM ammonium bicarbonate) and peptides were recovered by sequential extraction with 25 mM ammonium bicarbonate buffer, 5% formic acid, and acetonitrile. Extracts were pooled, lyophilised and re-dissolved in 0.1% formic acid prior to mass spectrometry.

Mass spectrometry analysis

Tandem electrospray mass spectra were recorded with a Q-Tof hybrid quadrupole/orthogonal acceleration time-of-flight spectrometer (Waters, Manchester, UK) interfaced to a CapLC chromatograph. Freeze-dried peptide samples were redissolved in 0.1% formic acid, and 6 µL injected onto a Pepmap C_{18} column (300 µm × 0.5 cm; LC Packings, Amsterdam, The

Netherlands), and eluted with an acetonitrile/0.1% formic acid gradient at 1 µl per minute. The capillary voltage was set to 3,500 V, and data-dependent product ion scans were performed on precursor ions with charge states of 2, 3 or 4 over a survey mass range of 400 to 1,400. The raw spectra were smoothed, deisotoped, transformed onto a singly charged *m/z* axis using a maximum entropy method as implemented in the peptide auto module of MassLynx (Waters UK) and saved in the peaklist (pkl) format prior to database searching. Proteins were identified by correlation of uninterpreted spectra to entries in SwissProt / TrEMBL using ProteinLynx Global Server (version 1.1, Waters, Manchester, UK) and a local installation of Mascot (version 2.2: www.matrixscience.com). The database used was a FASTA format composite constructed in house by merging SwissProt, TrEMBL and associated splice variants (releases of 26/05/09; 8,413,758 sequences). Searches were run in error tolerant mode and no mass or taxomic constraints were applied. The initial enzyme specificity was set to trypsin, but subsequent searches of the *P. gingivalis* digestions of fibrinogen and enolase were repeated with no enzyme specificity in order to match peptides resulting from the combination of gingipain activity with other enzymes such as amino- and carboxy-peptidases. All spectra matching citrullinated peptides were reviewed manually by interpretation of sequence-specific fragment ions to confirm presence and location of the citrulline residue and to exclude other modifications such as deamidation of aspartic acid which also result in a mass increase of 1 Da.

Production and evaluation of a polyclonal antibody against *P. gingivalis* PAD (anti-PPAD)

A polyclonal antibody recognising *P. gingivalis* PAD was produced in rabbits in cooperation with Cambridge Research Biochemicals (CRB), Billingham, UK. The immunisation and evaluation procedure is outlined in Table 2.4.

Table 2.4 Overview of steps for production of a polyclonal antibody to *P. gingivalis* PAD.

"N/A" stands not "not applicable", as the date was not important for the respective steps.

Schedule	Step	Details	Carried out by
N/A	Preparation of the immunising protein antigen	Thioredoxin-His-PPAD (1.5 mg), expressed in *E. coli* BL21 (DE3) cells, purified using Talon resin, dialysed into PBS (1 mg/ml)	Natalia Wegner
N/A	Preparation of antigen for ELISA analysis	Thioredoxin-His-PPAD (1 mg) and Thioredoxin-His-tag (1.2 mg), expressed in *E. coli* BL21 (DE3) cells, purified using Talon resin, dialysed into PBS (1 mg/ml)	Natalia Wegner
Day 1	Pre-immune bleed (PI)		CRB
Day 1	Immunisation #1	2 rabbits Freund's complete adjuvant	CRB
Day 14	Immunisation #2	Freunds incomplete adjuvant	CRB
Day 21	Test Bleed #1 (TB1)		CRB
N/A	ELISA analysis on PI, TB1	Antigens: Thioredoxin-His-PPAD and Thioredoxin-His-tag	CRB
Day 28	Immunisation #3	Freunds incomplete adjuvant	CRB
Day 35	Test Bleed #2 (TB2)		CRB
N/A	ELISA analysis on TB2	Antigens: Thioredoxin-His-PPAD and Thioredoxin-His-tag	CRB
N/A	Immunblot analysis of TB2	Antigens: Thioredoxin-His-PPAD, cleaved PPAD (no tag), Thioredoxin-His-tag	Natalia Wegner
Day 42	Immunisation #4	Freunds incomplete adjuvant	CRB
Day 49	Test Bleed #3 (TB3)		CRB
Day 56	Immunisation #5	Freunds incomplete adjuvant	CRB

Day 63	Test Bleed #4 (TB4)		CRB
N/A	ELISA analysis on TB3, TB4	Antigens: Thioredoxin-His-PPAD and Thioredoxin-His-tag	CRB
Day 70	Immunisation #6	Freunds incomplete adjuvant	CRB
Day 77	Final Bleed (FB)	77 ml (pooled)	CRB
N/A	ELISA analysis on FB	Antigens: Thioredoxin-His-PPAD and Thioredoxin-His-tag	CRB
N/A	Immunblot analysis of FB	Antigens: cleaved PPAD (no tag), Thioredoxin-His-tag, *P. gingivalis* WT and Δ*ppad* lysates. Small-scale negative depletion studies with Thioredoxin-His-tag.	Natalia Wegner
N/A	Preparation of antigen for affinity-purification	GST-His-PPAD, expressed in *E. coli* BL21 (DE3) cells, purified using Glutathione resin, and dialysed into PBS.	Natalia Wegner
N/A	Affinity purification of FB	Serum pooled from both rabbits. Purification column: GST-PPAD coupled to thiopropyl sepharose. One purification pass. Two elutions (glycine and TEA).	CRB
N/A	Immunblot analysis of affinity-purified antibody (glycine and TEA eluates)	Antigens: cleaved PPAD (no tag), Thioredoxin-His-tag, *P. gingivalis* WT and Δ*ppad* lysates.	Natalia Wegner

The immunising antigen, Thioredoxin-His-PPAD, and all other antigens were prepared as outlined in the respective sections ("*Expression and purification of recombinant proteins*").

Immunoblot analysis was performed according to the standard protocol described in "*Protein analysis*". Specifically for this application, the following details are important: The serum from the two rabbits was evaluated separately. Serum dilutions of 1:1000, 1:10.000 and 1:100.000 were used. 5% milk in TBS-T was used as blocking and antibody buffer.

Secondary antibody was swine anti-rabbit iG, HRP-conjugated, used at a dilution of 1:5000. Blocking was performed for 1.5 hours at RT, primary antibody (=serum) incubation for 1 hour at RT and secondary antibody incubation for 1 hour at RT. ECL was used to detect binding.

For the small-scale negative depletion study, a 1:1000 serum dilution (in 5% milk and TBS-T) was pre-incubated with thioredoxin-his-tag at 1 μg/ml, 10 μg/ml and 100 μg/ml, for 1 hour at RT. All remaining steps were performed as described.

Affinity purification of the final bleeds pooled from the two rabbits resulted in 11 ml of "glycine eluate" at 1.31 mg/ml in PBS and 40 ml of "TEA eluate" at 0.19 mg/ml in PBS. These two eluates were evaluated for their sensitivity and specificity to cleaved PPAD (no tag) and native PPAD (in *P. gingivalis* cell lysates) by immunoblotting as described above.

Enzymatic studies with *P. gingivalis* PAD

Colorimetric quantification of citrulline

The citrulline side chain was detected using a colorimetric method (Knipp *et al.* 2000). A standard curve was created using free L-citrulline (0, 10, 50, 100, 200, 400 μM) in PPAD activity buffer (see "*Solutions and Buffers*"). The performance of this method was evaluated using various concentrations of free L-citrulline against free L-arginine.

60 μl of citrulline-containing sample or standard was mixed with 200 μl of freshly prepared "citrulline detection reagent" (see "*Solutions and Buffers*") in a 96-well ELISA plate. The plate was sealed with adhesive film and incubated at 95˚C for 15 min. The plate was then allowed to cool down on ice (5-10 min) and the absorbance read at 540 nm in an ELISA plate reader.

The standard curve was created after substracting the blank absorbance value (0 μM) from all standards and plotting a linear function with x = μM citrulline and y = A (540 nm). The slope was used to determine the concentration of citrulline in μM in the unknown sample wells, after adjusting for background from substrate alone (control). A sample containing enzyme but no substrate was run to control for possible background from the enzyme itself.

"Gold standard" PPAD activity assay using BAEE as synthetic substrate

In a 96-well microtitre plate the following components were combined: 30 µl 20 mM BAEE in PPAD activity buffer (see "*Solutions and Buffers*"), the desired amount of PPAD (0.5 - 2 µM), and PPAD activity buffer to a final volume of 60 µl. Control wells and standard curve were prepared as described in the section above. The plate was sealed with adhesive film and incubated at 37˚C for 15 min. Citrulline was detected and quantified as described above. Free L-arginine and L-benzoyl-arginine as substrates were assessed in an analogous manner.

Determination of enzymatic activity of PPAD

The enzymatic activity (v) of PPAD was calculated using equation (1):

(1) $$v = \frac{(A - A_0) * V}{B * T}$$

where,

v is the enzymatic activity, in µmol/min = U (unit),
A is the absorption of enzymatic reaction (PPAD+substrate+buffer), in absorption units,
A_0 is the absorption of blank (substrate+buffer), in absorption units,
V is the incubation volume, in L,
B is the slope of the calibration curve, in 1/µM, and
T is the time of enzymatic reaction, in min,

at 37°C and pH 9.5.

One unit of PPAD = Amount of enzyme that catalyses the production of 1 µmol of BCEE from BAEE per min at 37 °C at pH 9.5

The enzyme activity unit U can also be expressed in the SI unit katal (mol/s), which is a very large unit and therefore less practical (1 U = 16.67 nanokatals). Therefore, the U was used in this study.

Calibration curves were generated using L-citrulline as standard as described above. The measured absorption values at 540 nm, corrected for blank (buffer), were plotted on the y-axis against the L-citrulline concentration on the x-axis. A linear trendline was generated, intercepting at "0", and the resulting slope *B* was used in equation (1). BAEE was used at a final concentration of 10 mM as the substrate and PPAD (>95% pure as determined by SDS-PAGE) was used at a final concentration of 0.1, 0.5 and 1 μM. The incubation time was 10 min at 37˚C as this was time was within the linear period of enzyme activity versus substrate concentration, as determined in previous experiments.

Specific activity

The specific activity of PPAD was calculated based on the results gained from the experiment outlined above for a final concentration of 0.1, 0.5 and 1 μM PPAD.

The specific activity is the enzyme activity per milligram of total protein. The unit is μmol/min/mg = U/mg. The SI unit is katal/kg but will not be used for reasons mentioned above. In the case of a pure enzyme preparation, it measures enzyme processivity. In the case of an impure prepration, it indicates the purity of the sample (if the specific activity of a ~100% pure enzyme is known):

% purity = 100% × (specific activity of enzyme sample / specific activity of pure enzyme).

The impure sample has lower specific activity because some of the mass is not actually enzyme. If the specific activity of 100% pure enzyme is known, then an impure sample will have a lower specific activity, allowing purity to be calculated. This was done with PPAD preparations which were not further purified by size exclusion chromatography and therefore contained ~30% of a low-molecular-weight truncated protein, mostly made up of Thioredoxin-His-tag. This did not contribute to or interfere with PPAD enzyme catalysis, as determined in activity assays. The calculated purity of the respective PPAD preparations made it possible to use PPAD at a fixed concentration in all assays (e.g. 0.5 μM) and thereby eliminate variations in enzyme concentration.

PPAD activity assay using peptides

Synthetic peptides were were diluted in PPAD activity buffer to a concentration of 4 mM, 2 mM and 0.2 mM and 30 µl of each was mixed with 1 µM PPAD (final peptide concentration: 2 mM, 1 mM, 0.1 mM). Control wells and standard curve were prepared as described in the section above. The plate was sealed with adhesive film and incubated at 37˚C for 15 min. Citrulline was detected and quantified as described above.

PPAD activity assay with haemoglobin

Oxy-haemoglobin was used at a final concentration of 100, 40, and 4 µM (referring to the tetrameric state). PPAD was used at a final concentration of 1 µM. Control wells and standard curve were prepared as described in the section above. The plate was sealed with adhesive film and incubated at 37˚C for 60 min. Citrulline was detected and quantified as described above. Citrullination in the sample using 4 µM haemoglobin was also confirmed by immunoblotting. 16.25 µl of the reaction mix were run on a SDS-PAGE gel, transferred onto nitrocellulose membrane and citrullination detected using the anti-(modified)-citrulline kit, as described previously.

Time course of PPAD activity (progress curve)

BAEE was used as substrate at a final concentration of 10 mM. PPAD was used at 0.5 µM. All components were combined in an Eppendorf tube and 60 µl were removed immediately as the t = 0 min sample. The tube was placed in an 37˚C incubator and 60 µl samples were withdrawn after 1, 2, 5, 10, 20, 40, 50, and 60 min. 200 µl citrulline detection reagents was added to each aliquot immediately after withdrawing the sample in order to stop the enzymatic reaction. All samples were transferred to a 96-well plate. Control wells and standard curve were prepared as described in the section above. The plate was sealed with adhesive film and citrulline was detected and quantified as described above.

PPAD activity with different substrate concentrations (kinetics)

PPAD was used at 0.5 µM and BAEE was used as a substrate at 0, $1x10^{-4}$, $1x10^{-3}$, $5x10^{-3}$, $1x10^{-2}$, $1x10^{-1}$, $5x10^{-1}$, 1, 2, 5, and 10 mM. Control wells and standard curve were prepared as described in the section above. The plate was sealed with adhesive film and incubated at 37˚C

for 15 min. Citrulline was detected and quantified as described before. The enzyme activity for each substrate concentration was calculated according to equation (1). The enzyme activity was expressed as (µmol/min)x10^{-3}. The data were fit with GraphPad Prism 5 to determine the Michaelis-Menten-Konstant K_m and the maximal reaction rat v_{max} by non-linear regression.

Inhibition studies

The following compounds were tested for inhibition of PPAD enzymatic activity: 2-chloroacetamidine, tetracycline, doxycycline, minocycline, sulfasalazine, methotrexate. PPAD was used at 0.5 µM. The compounds were used at final concentrations 0.01, 0.1, 1, 10, 100 and 1000 µM. 30 µl of enzyme solution (0.5 µM PPAd in activity buffer) was mixed with 30 µl of inhibitor in assay buffer and pre-incubated for 10 min at 37˚C. 3 µl of substrate (20 mM BAEE, leading to a final concentration of 1 mM) was added and the mix incubated for 15 min at 37˚C. Control wells lacking inhibitor, substrate, or enzyme were included. Citrulline was detected and quantified as described previously. Absorbance values of the respective control wells lacking enzyme were substracted from the values gained with enzyme. The activity value gained from the contol lacking inhibitor was defined as "100 % activity". All other activity values were calculated as a fraction thereof.

Activity studies with PPAD mutants

The PPAD mutants D130A, H236A, D238A, N297A and C351A were used alongside the PPAD wild-type to assess whether the introduced amino acid substitutions affected PPAD activity. BAEE and fibrinogen-derived peptides were used as substrates, both at a final concentration of 1 mM. Each of the PPAD enzymes was used at a final concentration of 1 µM. Control wells and standard curve were prepared as described in the section above. The plate was sealed with adhesive film and incubated at 37˚C for 15 min. Citrulline was detected and quantified as described above.

Analysis of antibody responses by ELISA

Patients

Reactivity of RA and healthy control serum to fibrinogen peptides

Samples were kindly provided by Dr. Peter Charles, Kennedy Institute of Rheumatology

Serum samples from 48 RA patients were collected from patients attending rheumatology

clinics at Charing Cross Hospital following consent under approval from the local ethics committee and anonymised before storage at -80˚C. All patients met the American College of Rheumatology (ACR) 1987 classification criteria for definite or probable RA (Arnett *et al.* 1988). Serum samples from 33 healthy controls were obtained by National Blood Transfusion Service after consent and were anonymised before storage at -80˚C.

Peptide synthesis and sequences

All peptides were synthesised by Cambridge Research Biochemicals (Billingham, UK) and purified by HPLC to a purity of >70%. The C-terminal amino acid was routinely in the amide form, except for all peptides where the C-terminal amino acid was citrulline or arginine, in which case it was synthesised in the acid form so as to resemble the physiological state after proteolytic cleavage by gingipains (Table 2.5). All fibrinogen peptides were resuspended at 10 mg/ml in sterile, deionised water, aliquoted and stored at -20˚C. CEP-1/REP-1 peptides were resuspended in 10% DMSO in sterile, deionised water and stored as described above.

Table 2.5 Amino acid sequences of synthetic peptides used for ELISA analysis.
"Cit" stands for "citrulline". "P.g." stands for "P. gingivalis".

Name	**Sequence**
FibA-Cit	CESSSHHPGIAEFPS-Cit-acid
FibA-Arg	CESSSHHPGIAEFPS-R-acid
FibA-Cit-XX	CESSSHHPGIAEFPS-Cit-GK-amide
FibA-Arg-XX	CESSSHHPGIAEFPS-R-GK-amide
FibB-Cit	CPAPPPISGGGY-Cit-acid
FibB-Arg	CPAPPPISGGGY-R-acid
FibB-Cit-XX	CPAPPPISGGGY-Cit-AR-amide
FibB-Arg-XX	CPAPPPISGGGY-R-AR-amide
CEP-1 (human)	CKIHA-Cit-EIFDS-Cit-GNPTVEC-amide
REP-1 (human)	CKIHA-R-EIFDS-R-GNPTVEC-amide
CEP-1 (P.g.)	CKIIG-Cit-EILDS-Cit-GNPTVEC-amide
REP-1 (P.g.)	CKIIG-R-EILDS-R-GNPTVEC-amide

ELISA protocol

Peptides were diluted at 10 µg/ml in coating buffer (see "*Solutions and Buffers*"). 96-well EIA/RIA plates (Costar) were coated with 100 µl/well peptide solution or with coating buffer alone, and incubated overnight at 4˚C. Wells were washed 4x with PBS-Tween (0.05%) and blocked with 200 µl/well 2% BSA (diluted in PBS) for 3 hours at RT. Serum was diluted 1:100 in RIA buffer (see "*Solutions and Buffers*"), and 100 µl/well was added in duplicates and incubated for 1.5 hours at RT. Plates were washed as described above. Peroxidase-conjugated mouse anti-human IgG or IgA was diluted 1:1000 in RIA-buffer and added to the wells (100 µl/well). Plates were incubated for 1 hour at RT. After a final wash, bound antibodies were detected with TMB substrate. 100 µl/well were added and incubated for 2 minutes. The reaction was stopped by the addition of 1M H_2SO_4 and absorbance measured at 450 nm in a Multiscan Ascent microplate reader (Thermo Labsystems). A control serum was included on all plates to correct for plate-to-plate variation. The OD_{450} value of the arginine-containing control peptides was subtracted from the OD_{450} of the corresponding citrulline-containing peptides to calculate the citrulline-specific response. The positive cut-off value was defined as the 98^{th} percentile of the healthy control population.

Statistical analyses

The Mann-Whitney test (non-parametric, unmatched groups) was used to compare differences between antibody responses to specific peptides in serum samples from RA patients and healthy controls and to compare the response to different peptides within the RA group. The Fisher's exact test was used to compare the percentage of positive serum samples (above cut-off) to specific peptides from RA patients and healthy controls. Calculations were performed using GraphPad Prism. P-values below the value of 0.05 were considered significant.

Chapter 3

Investigation of protein citrullination by *P. gingivalis* and the associated host antibody response

Background

The results and some text presented in this chapter (p.124-142) were partly published in a paper authored by myself in the journal Arthritis & Rheumatism; please see attached paper (Wegner et al. 2010b).

A functional PPAD enzyme had previously been isolated and purified from the culture supernatant of a *P. gingivalis* mutant strain (McGraw *et al.* 1999), however this enzyme was poorly characterised and its role was unknown. For example, it was unclear whether *P. gingivalis* citrullinates its own and/or foreign (host) proteins using PPAD, which arginine residues can be citrullinated and whether this is the only citrullinating enzyme in the bacterium. Further, it had been speculated that gingipains, the major proteolytic virulence factors of *P. gingivalis*, were involved in processing of PPAD into an active form and/or in processing of PPAD substrates (McGraw *et al.* 1999). It has even been claimed that *P. gingivalis* citrullinates human fibrinogen in the periodontal pocket (Rosenstein *et al.* 2004), despite the lack of any experimental data, including the referenced article (Travis *et al.* 1997). Unfortunately, this claim has since been repeatedly cited (de Pablo *et al.* 2008; de Pablo *et al.* 2009; Liao *et al.* 2009; Mikuls *et al.* 2009; Berthelot *et al.* 2010; Routsias *et al.* 2011), adding to confusion and misperceptions in this field.

The preliminary data leading up to this project (p. 44-46) identified citrullinated *P. gingivalis* enolase as a cross-reactive antigen with RA serum. However, it was impossible to conclude whether this had biological significance, since it was unknown whether *P. gingivalis* expresses citrullinated enolase and whether in particular the equivalent of the CEP-1 epitope can be generated by PPAD. The aim of the work presented in the first part of this chapter was to investigate the first step of the hypothesis as highlighted in Figure 3.1: Does *P. gingivalis* generate citrullinated human and bacterial proteins and how does it do it?

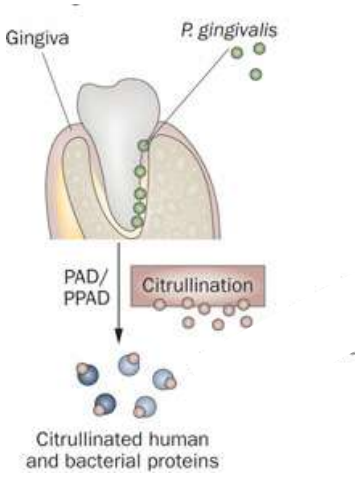


Figure 3.1 Schematic illustration of the first step of the aetiological hypothesis for *P. gingivalis* involvement in RA. For the full picture and a detailed explanation please refer to Figure legend 1.10.

Results

Expression of endogenous citrullinated proteins in *P. gingivalis* reference strain W83

In order to test whether *P. gingivalis* citrullinates its own proteins, a culture of the *P. gingivalis* wild-type reference strain W83 was grown under standard conditions for 24 hours (representing early logarithmic phase) or 72 hours (representing late logarithmic phase), and whole cell lysates or concentrated culture supernatant were analysed by immunoblotting (Figure 3.2).

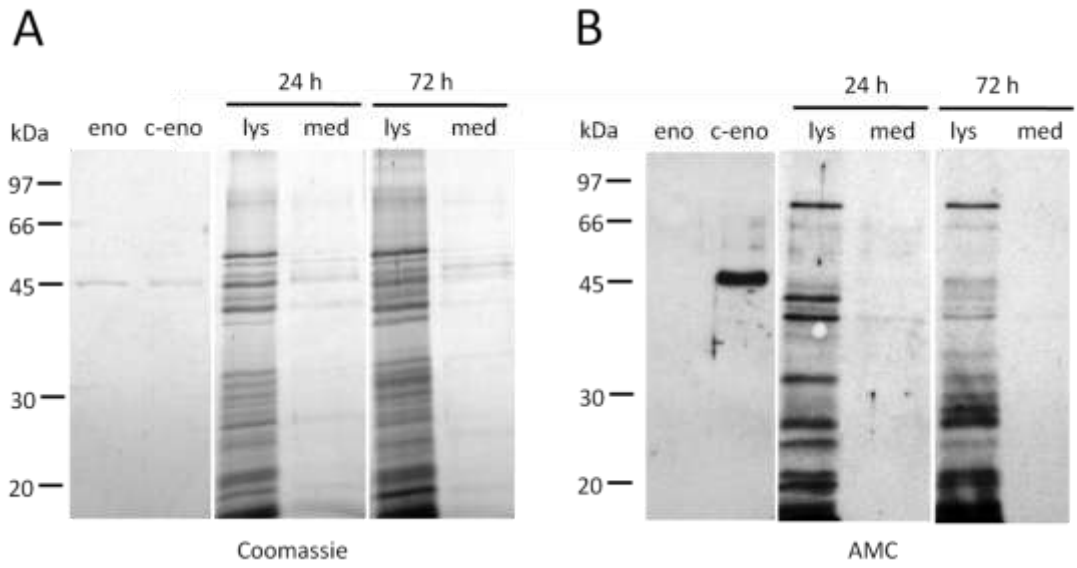


Figure 3.2 Endogenous citrullinated proteins in *P. gingivalis* W83. The soluble fractions of total cell lysates of *P. gingivalis* cultures ("lys") and 20-fold concentrated culture medium ("med"), after 24 hours and 72 hours of growth, were resolved on a 10% tris-glycine gel and either stained with Coomassie (A) or transferred onto nitrocellulose membranes and analysed for citrullinated proteins using the anti-modified citrulline antibody (AMC) (B). As a negative control, unmodified recombinant *P. gingivalis* enolase was used (eno). *In vitro* citrullinated recombinant *P. gingivalis* enolase was used as a positive control (c-eno). The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Strong, distinct bands were observed in the lysates and only weak bands in the concentrated culture supernatant. The citrullination patterns changed with growth stage, as there were observable differences in the number and intensity of citrullinated protein bands between cell lysates of a 24 hour and a 72 hour culture. Twenty-four hour cultures were chosen for all subsequent experiments.

To test reproducibility, *P. gingivalis* W83 cell lysates from three different sources (two different starter cultures from Jan Potempa, Jagiellonian University, Krakow, Poland and one from Sigrun Eick, University of Jena, Germany) were used. Similar patterns of citrullinated proteins were observed (Figure 3.3). Minor variability between the samples was observed and might result from different storage times and conditions.

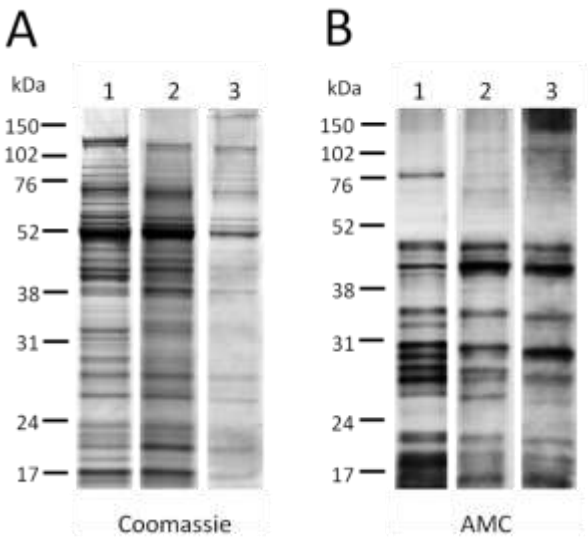


Figure 3.3 Comparison of citrullinated proteins from three separate preparations (1-3) of *P. gingivalis* W83 whole cell lysate. The soluble fractions of total cell lysates of *P. gingivalis* cultures were resolved on a 12% Bis-Tris gel and either stained with Coomassie (A) or transferred onto nitrocellulose membranes and analysed for citrullinated proteins using the anti-modified citrulline antibody (AMC) (B). The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Citrullinated proteins in other *P. gingivalis* reference strains and clinical isolates

Citrullinated proteins were found in various different *P. gingivalis* strains, including clinical isolates (Figure 3.4). The clinical isolates were obtained by Sigrun Eick (University of Jena, Germany) from diseased sites from periodontitis patients and, after initial cultivation and verification, grown in the same way as the wild-type reference strains.

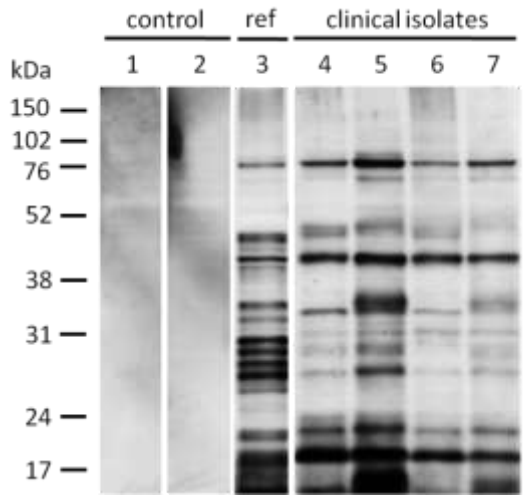


Figure 3.4 Expression of endogenous citrullinated proteins is ubiquitous in *P. gingivalis*. Protein citrullination in total cell extracts of the *P. gingivalis* wild-type reference strain W83 (1-3) and four clinical isolates (4-7; corresponding to strains MaRL, D243, JH16, J430) was analysed by immunoblotting with the anti-modified citrulline antibody. Lanes 1 and 2 are controls in which the modification step (1) or secondary antibody (2) was omitted. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

The gene encoding PPAD was detected in all tested *P. gingivalis* strains, including the clinical isolates, by PCR (Figure 3.5).

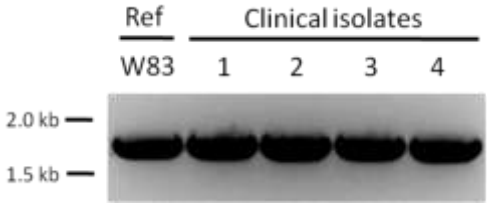


Figure 3.5 The PPAD gene is present in all tested *P. gingivalis* strains. Genomic DNA was isolated from *P. gingivalis* reference strain W83 (Ref) and four clinical isolates (1-4, corresponding to strains MaRL, D243, JH16, J430) and subjected to PCR analysis using *ppad*-specific primers amplifying the full-length gene (1.7 kb). The ethidium-bromide stained agarose gel of the PCR product is shown. The position of the nucleotide ladder in kilobases (kb) is indicated on the left.

Subcellular localisation of citrullinated proteins

The subcellular distribution of citrullinated proteins was studied using a cell fractionation procedure with *P. gingivalis* strain W83, resulting in six fractions: medium+vesicles (little outer membrane bubbles released from the cell surface, containing virulence factors (Haurat *et al.* 2010)), medium (culture supernatant without vesicles), periplasm (containing soluble periplasmic proteins), cytoplasm (containing soluble cytoplasmic proteins), inner membrane (containing proteins which can be solubilised by the detergent sodium lauryl sarcosine which applies to the majority of proteins attached to or integrated into the inner membrane), and outer membrane (all other proteins which cannot be solubilised by sodium lauryl sarcosine, which is mostly true for proteins from the outer membrane). Immunoblotting with the anti-modified citrulline antibody showed that the majority of citrullinated proteins are associated with the periplasm and the membrane fractions. Only faint bands were detected in the cytoplasm fraction, and one prominent band was detected in the vesicles+medium and medium fractions (Figure 3.6). These results imply that PPAD either citrullinates proteins directly on the cell surface or in the periplasm or that the proteins are translocated to these compartments or secreted/shedded after citrullination.

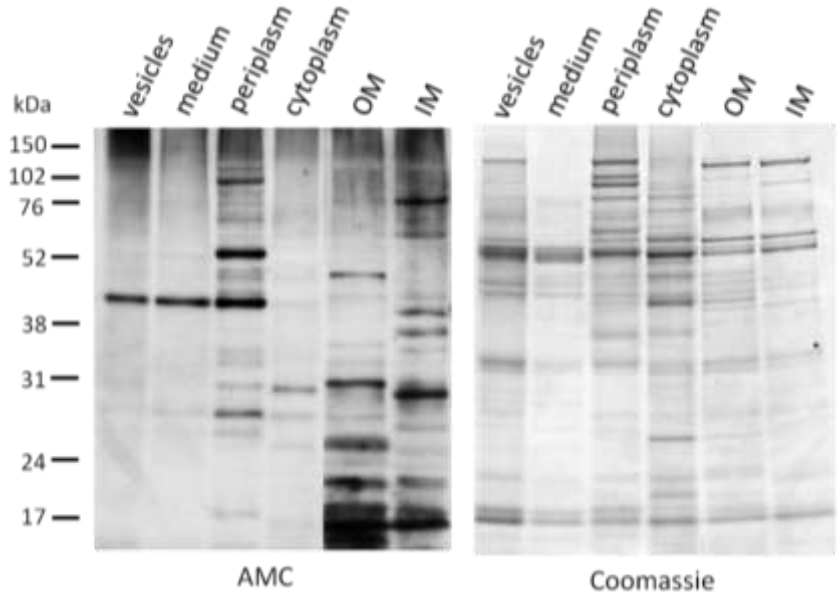


Figure 3.6 The majority of endogenous citrullinated proteins in *P. gingivalis* W83 is associated with the periplasm and membrane fractions. *P. gingivalis* strain W83 was fractionated into vesicles+medium ('vesicles'), medium, periplasm, cytoplasm, outer membrane (OM) and inner membrane (IM) and the fractions analysed for citrullinated proteins using the anti-modified citrullinated (AMC) antibody. The corresponding Coomassie-stained gel showing total protein in each of the fractions is shown on the right. The positions of the molecular mass markers in kilodalton (kDa) are indicated on the left.

Protein citrullination in other oral bacteria

To examine whether endogenous citrullination is a unique ability of *P. gingivalis* within the community of oral pathogens, whole cell lysates of ten other oral organisms were tested for the presence of citrullinated proteins. None were detected except in *P. gingivalis* (Figure 3.7), suggesting that functional PAD enzymes are absent from the other strains tested. Weak bands were noticeable in a number of strains, but were the result of non-specific antibody binding (see 'control').

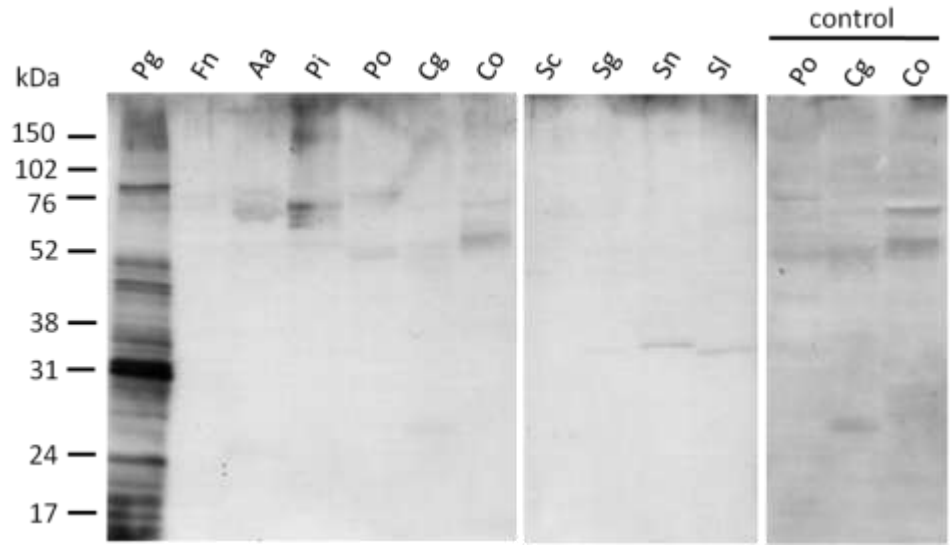


Figure 3.7 Lack of protein citrullination in ten other oral bacteria. Total cell extracts of *P. gingivalis* W83 and ten other prominent oral bacteria were tested for endogenous protein citrullination using the AMC antibody. Background signals were confirmed to stem from non-specific binding of the primary antibody ('control'). Pg: *Porphyromonas gingivalis*; Fn: *Fusobacterium nucleatum*; Aa: *Aggregatibacter actinomycetemcomitans*; Pi: *Prevotella intermedia*; Po: *Prevotella oralis*; Cg: *Capnocytophaga gingivalis*; Co: *Capnocytophaga ochracea*; Sc: *Streptococcus constellatus*; Sg: *Streptococcus gordonii*; Sn: *Streptococcus sanguinis*; Sl: *Streptococcus salivarius*. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

These findings are in accordance with PPAD homology searches in oral bacteria genome sequences. Amongst the bacteria tested here, the genome sequences are available for *Porphyromonas gingivalis* W83 (Nelson *et al.* 2003), *Porphyromonas gingivalis* ATCC 33277 (Naito *et al.* 2008), *Fusobacterium nucleatum* (Kapatral *et al.* 2002) and *Prevotella intermedia* (provided prior to publication by John Heidelberg at TIGR and by Margaret

Duncan at the Forsyth Institute, available on the Oral Pathogen Sequence Database[18]): Nucleotide sequence homology searches with the BLASTN programme (Altschul *et al.* 1990) using PPAD (GenBank ID 34541100) as input sequence show hits for both *P. gingivalis* strains but no significant matches (E-value < 0.0001) in *F. nucleatum* or *P. intermedia*. Similarity searches using BLAST against all published sequences available revealed a number of candidate orthologues, all originating from non-oral bacterial species (Table 3.1). Most share the predicted conserved catalytic residues of PPAD and other members of the guanidino-group modifying enzymes superfamily (for a representative alignment please see Appendix, p. 243). As these are similar between peptidylarginine deiminases and agmatine deiminases, the actual enzymatic activity might be that of agmatine deiminases, but should be investigated. If confirmed, the presence of a PAD in these bacteria could provide further clues to the physiological role of PADs in bacteria.

Table 3.1 Potential peptidylarginine deiminase enzymes in other bacterial species. *Hits were identified by BLAST using Porphyromonas gingivalis W83 peptidylarginine deiminase (Q9RQJ2) as query sequence. Entry names refer to the Uniprot database. Identity in % was calculated as the number of identical amino acids divided by the number of total aligned amino acids.*

Organism	**Database protein name**	**Entry name**	**% Identitiy to PPAD**
Cloacamonas acidaminovorans	Peptidylarginine deiminase	B0VH37_CLOAI	44%
unidentified eubacterium SCB49	Peptidylarginine deiminase	A6ES44_9BACT	29%
Flavobacteriales bacterium HTCC2170	Peptidylarginine deiminase	A4AR43_9FLAO	26%
Clostridium difficile (strain 630)	Putative peptidylarginine deiminase	Q188A5_CLOD6	24%

[18] http://www.oralgen.lanl.gov

The role of PPAD and arginine-gingipains in endogenous protein citrullination by *P. gingivalis*

To confirm that the observed endogenous protein citrullination in *P. gingivalis* is due to the enzymatic activity of PPAD, and to rule out the possibility of a second, uncharacterized, bacterial PAD enzyme, a *P. gingivalis* W83 knockout strain (Δ*ppad*) was created by replacement of the entire *ppad*-encoding DNA sequence with an antibiotic cassette. A strain in which the antibiotic cassette was inserted downstream of the *ppad* gene was used as a control against polar effects from genetic manipulations (*ppad*+). Figure 3.8 demonstrates successful genetic manipulation.

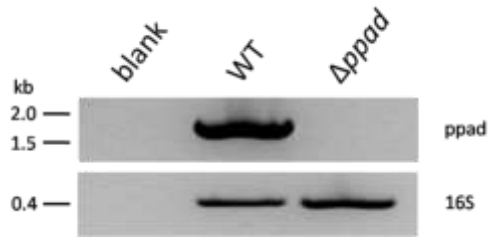


Figure 3.8 The *ppad* gene is absent in a deletional inactivation mutant of *P. gingivalis* Δ*ppad*. Genomic DNA was isolated from the *P. gingivalis* wildtype strain W83 (WT) and from a mutant strain lacking the peptidylarginine deiminase gene (Δ*ppad*). PCR analysis was performed using primers amplifying the full-length *ppad* gene (upper panel) or, as a positive control, a fragment from the *P. gingivalis* 16S rRNA gene (lower panel). As a negative control, template DNA was left out of the reaction mixture ('blank').

Immunoblotting of the total cell lysates for citrullinated proteins showed that the Δ*ppad* strain entirely lacks endogenous citrullinated proteins, while the wild-type strain (WT) and the control strain (*ppad*+) showed a similar pattern and intensity of citrullinated proteins (Figure 3.9). These data demonstrate that PPAD is essential for citrullination of endogenous proteins in *P. gingivalis* and that it possesses only one peptidylarginine deiminase.

It was then examined how citrullination depends on the activity of gingipains, the major virulence factors in *P. gingivalis.* As arginine-gingipains cleave various proteins/peptides after arginine residues and thereby generate peptides with carboxy-terminal arginines, their activity could explain the previously suggested preference of native PPAD for carboxy-terminal arginine residues *in vitro* (McGraw *et al.* 1999).

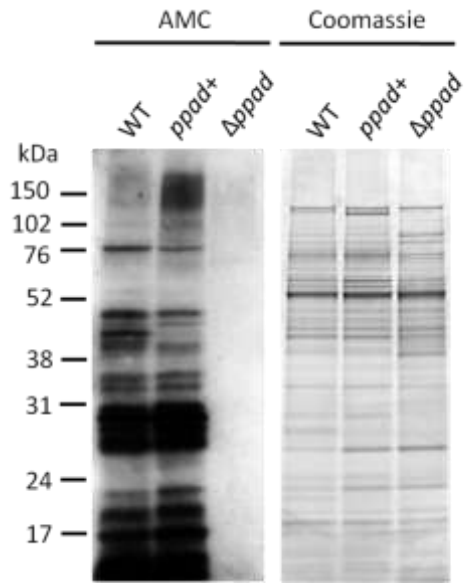


Figure 3.9 Citrullination in *P. gingivalis* depends on the bacterial peptidylarginine deiminase enzyme. A *P. gingivalis* mutant strain lacking the bacterial peptidylarginine deiminase gene (Δ*ppad*) was constructed and total cell extracts were analysed for the presence of citrullinated proteins (AMC). A control mutant containing the entire *ppad* gene and the antibiotic cassette (*ppad*+) was created to control for possible polar effects. The wild-type strain W83 was used as positive control (WT).

Hence, *P. gingivalis* mutants lacking functional arginine-gingipains (Δ*rgp*), lysine-gingipains (Δ*kgp*), or both types of gingipains (*Δrgp+kgp*) were analysed for endogenous citrullination. Immunoblotting of whole cell lysates showed that there is a significantly decreased level, but not complete abrogation, of citrullinated proteins in the Δ*rgp* and Δ*rgp+kgp* strains, but not in the Δ*kgp* strain (Figure 3.10), confirming that arginine-gingipains play a role in protein citrullination, probably by generating proteins with carboxy-terminal arginine residues that are subsequently citrullinated by PPAD. However, an indirect adverse effect on PPAD activity via altered processing of PPAD due to the lack of arginine-gingipains cannot be excluded. The residual citrullinated proteins seen in the Δ*rgp* and Δ*rgp+kgp* strains might be due to proteins naturally containing a carboxy-terminal arginine residue, which have not been proteolytically processed, and therefore appear at different molecular weights compared to the wildtype.

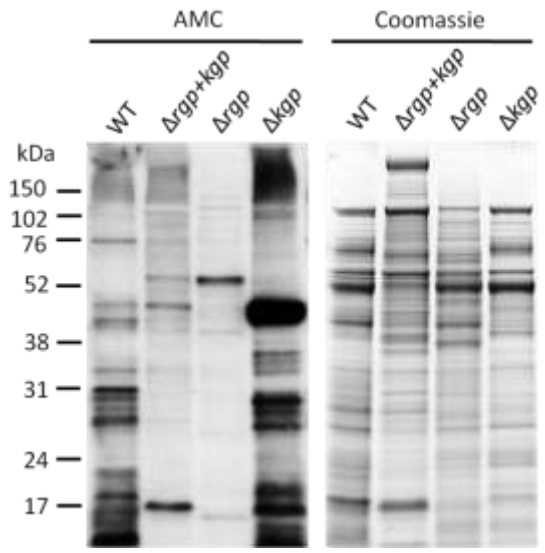


Figure 3.10 Citrullination in *P. gingivalis* is influenced by arginine-gingipain mediated proteolytic cleavage of substrate proteins. Total cell extracts of *P. gingivalis* wild-type W83 ('WT') and mutant strains lacking all proteolytic gingipain activity (Δ*rgp+kgp*) or either arginine-gingipain (Δ*rgp*) or lysine-gingipain (Δ*kgp*) were analysed for the presence of citrullinated proteins by immunoblotting with the AMC antibody.

Citrullination of human fibrinogen and α-enolase by *P. gingivalis*

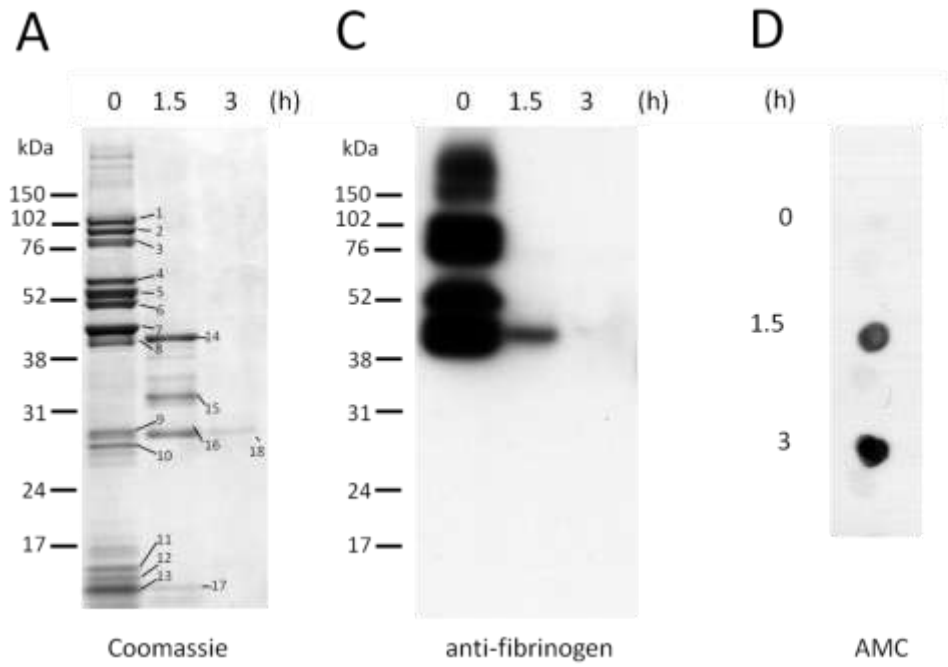
Having shown that citrullination of endogenous *P. gingivalis* proteins depends on the presence of PPAD and is influenced by arginine-gingipains, the question arose whether the same principles apply to human proteins.

Human fibrinogen was initially chosen for this study as it is a major RA autoantigen in its citrullinated form (Masson-Bessiere *et al.* 2001; Sebbag *et al.* 2006) and an important part of the inflammatory response in general, due to its function in the coagulation and platelet aggregation cascade. Fibrinogen is also involved in the pathogenesis of periodontitis, where it is found abundantly in the periodontal lesion, being an established target protein of gingipains (Imamura *et al.* 1995).

The potential of *P. gingivalis* to citrullinate human fibrinogen was initially investigated using intact, live wild-type W83 strain. The culture medium was removed by centrifugation and the bacteria were washed, fibrinogen was added, and the suspension incubated at 37˚C. Aliquots were removed immediately after mixing bacteria and fibrinogen ('0 hours'), and after an incubation time of 1.5 hours and 3 hours. Immediately after withdrawing each aliquot,

bacterial cells were removed from the samples by centrifugation and discarded. The supernatant was then analysed on SDS-PAGE. Figure 3.11 A shows that fibrinogen was rapidly cleaved by *P. gingivalis*, which is consistent with previous publications (Ally *et al.* 2003). To confirm that these bands were indeed fibrinogen products and not contaminants, all protein bands were excised and analysed by MS/MS (done in collaboration with Robin Wait, The Kennedy Institute of Rheumatology). All bands that could be identified are marked in Figure 3.11 A (numbers 1-18) and are listed in Figure 3.11 B. In summary, no proteins from *P. gingivalis* were detected and only one contaminating protein from the fibrinogen preparation (human Igκ chain C region) was detected; all other protein bands are derived from one or more of the three (α, β, γ) fibrinogen chains. The expected molecular masses of the fibrinogen chains were quoted by the manufacturer (Sigma) as 63.5 kD (α chain), 56 kD (β chain) and 47 kD (γ chain). Comparing this information to the experimental results, it seems that the γ chain forms a dimer (bands 1-3 in Figure 3.11 A), and that the α chain is rapidly degraded, as the estimated molecular mass of the largest band assigned to fibrinogen α chain at '0 h' is less than 31 kD (band 9 in Figure 3.11 A). The presence of fibrinogen was also confirmed by immunoblot using a mixture of three antibodies, each raised against a fibrinogen-derived peptide within fibrinogen α, β, or γ chain, therefore recognising only fibrinogen fragments that contain the respective peptide sequence (Figure 3.11 C).

To test for citrullination of the fibrinogen peptides, samples were applied as a spot directly onto a nitrocellulose membrane and the membrane was then processed in a standard immunoblot procedure using the anti-modified citrulline antibody. Figure 3.11 D shows that the samples after 1.5 hours and after 3 hours incubation contained citrullinated fragments. To identify these citrullinated peptides by MS/MS, the larger protein fragments were removed from the samples by precipitation using using *meta*-phosphoric acid, leaving small peptides in solution. *Meta*-phosphoric acid was used because it is compatible with the subsequent high-perfomance liquid chromatography (HPLC) analysis. The resulting soluble fractions from samples '1.5 hours' and '3 hours' were then analysed by HPLC on a reverse-phase column (performed by myself with technical help from Marysia and Andrzej Kozik, Jagiellonian University Krakow, Poland). After an initial run to determine the amount and number of peaks, the eluted fractions from the '3 hour' sample were collected, as there were more peaks of higher intensity, and hence probably more peptides in higher amounts.



B

No.	Identity	No.	Identity
1	Fibrinogen γ chain	**10**	Fibrinogen α chain
2	Fibrinogen γ chain	**11**	Fibrinogen α chain
3	Fibrinogen γ chain	**12**	Fibrinogen α chain
4	Fibrinogen β chain	**13**	Fibrinogen α chain
5	Fibrinogen β chain Fibrinogen γ chain	**14**	Fibrinogen β chain
6	Fibrinogen γ chain	**15**	Fibrinogen γ chain
7	Fibrinogen β chain	**16**	Fibrinogen γ chain Human Igκ chain C region
8	Fibrinogen γ chain	**17**	Fibrinogen α chain
9	Fibrinogen α chain	**18**	Fibrinogen γ chain Human Igκ chain C region

Figure 3.11 ***P. gingivalis*** **cleaves and citrullinates small peptides from human fibrinogen.** Whole, live *P. gingivalis* W83 cells were washed and mixed with human fibrinogen, and incubated at 37˚C. Samples were removed immediately after mixing (0 h), after 1.5 hours incubation (1.5 h) and after 3 hours incubation (3 h). Bacterial cells were removed by centrifugation and the resulting supernatant was resolved on a 12% Bis-Tris SDS-PAGE for detection of total protein by Coomassie (A) and for immunoblotting using anti-fibrinogen (α, β and γ chain) antibody (C), or applied directly onto a nitrocellulose antibody for detection of citrullinated peptides using the anti-(modified) citrulline (AMC) antibody (D). Protein bands in gel A were excised and analysed using MS/MS (B). The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Equal volumes of all collected peak fractions were then analysed by MS/MS (without trypsin digestion) to confirm the identity of the fragments and the presence and position of citrullination (performed in collaboration with Robin Wait, The Kennedy Institute, London). 16 fibrinogen-derived peptides could be unambiguously identified, 14 of which harboured a C-terminal lysine residue, and two contained a C-terminal citrulline in place of an arginine: ^{237}GGSTSYGTGSETESPCit^{252} and ^{524}ESSSHHPGIAEFPSCit^{538} (please see Appendix for mass spectrum and detailed analysis, p. 244). Peptides with internal citrulline or internal or C-terminal arginines were not detected.

Based on this encouraging pilot experiment, the setup was repeated including fibrinogen and α-enolase as protein substrates, control strains (Δ*ppad* and Δ*rgp* knockout strains), control reactions (incubate protein alone, bacteria alone), and more time points (pre-incubation, 1 min, 1.5 hours, 3 hours, 6 hours).

Analogous to the previous experiment, fibrinogen was rapidly cleaved by wild-type *P. gingivalis* (Figure 3.12 A). A similar degradation pattern was observed with the Δ*ppad* strain, indicating that citrullination of substrate proteins, including gingipains themselves, is not essential for the proteolytic potency of gingipains. As expected, fibrinogen samples incubated with the Δ*rgp* strain showed a considerably decreased proteolytic cleavage and a different pattern of cleaved peptides, with the residual proteolytic activity mainly due to cleavage by lysine-gingipains and other proteinases and peptidases from *P. gingivalis* (Jagels *et al.* 1996; Nelson *et al.* 1999; Banbula *et al.* 2000; Masuda *et al.* 2002). Analysis of the identity of the protein bands by mass spectrometry confirmed that the majority were derived from any of the three fibrinogen chains (Figure 3.12 A). To exclude cleavage of protein by plasma-derived proteases which may have been present as contaminants in the fibrinogen preparation, control reactions in which protein alone was incubated in assay buffer were performed, and no cleavage was detected (Figure 3.12 B). Similarly, *P. gingivalis* incubated without protein did not secrete signigficant amounts of protein which could interfere with analysis (Figure 3.12 C).

Next, it was examined whether the cleaved fibrinogen fragments visible in Figure 3.12 A had been citrullinated by *P. gingivalis*. Immunoblotting of all samples that contained protein bands detected two citrullinated peptide bands around 8.5 kDa, mapping to the amino-terminal region of the fibrinogen α-chain, in samples incubated with *P. gingivalis* wild-type but not Δ*ppad* or Δ*rgp* (Fig. 3.12 D), confirming that Arg-X proteolytic cleavage of

fibrinogen is a prerequisite for subsequent citrullination by PPAD. Further, a weak positive signal in fibrinogen samples taken before incubation with *P. gingivalis* ('pre') and in the Δ*rgp* '1 min' sample was observed, while the corresponding controls ('pre ctrl1': modification control; 'pre ctrl2': conjugate control) were negative (Fig. 3.12 D), suggesting that purified human fibrinogen, as used in these experiments, is already endogenously citrullinated by human PADs to some degree.

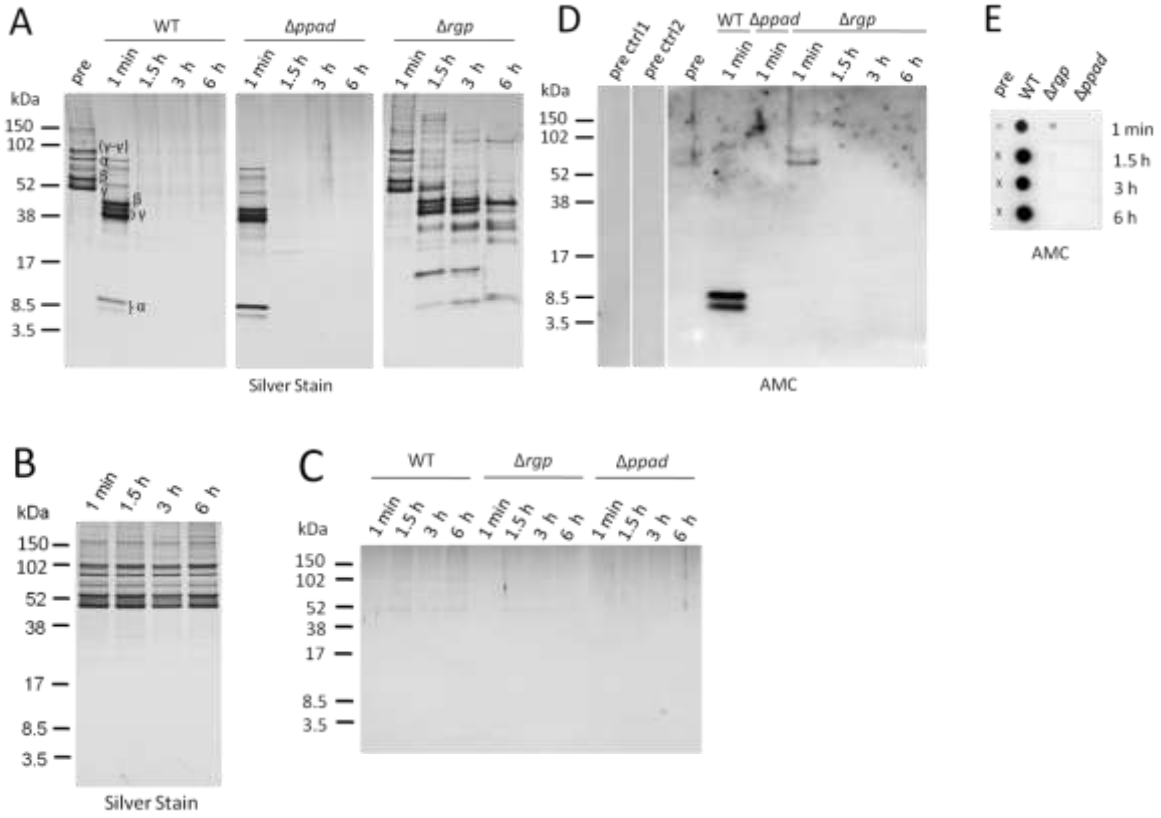


Figure 3.12 ***P. gingivalis*** **rapidly cleaves human fibrinogen through arginine-gingipain activity and the resulting peptides are citrullinated at the carboxy-terminus by bacterial peptidylarginine deiminase.** **(A)** Fibrinogen fragments after incubation with *P. gingivalis* wild-type (WT), mutants lacking peptidylarginine deiminase (Δ*ppad*), or arginine-gingipains (Δ*rgp*), for 1 min, 1.5, 3 and 6 hours, were resolved by SDS-PAGE and visualised using silver staining. A fibrinogen sample prior to incubation with *P. gingivalis* served as control (pre). Protein bands were analysed by mass spectrometry (α, β, γ for fibrinogen α-, β-, and γ- chains). **(B)** Fibrinogen was incubated without bacteria and visualised by silver staining. **(C)** *P. gingivalis* wild-type, Δ*rgp* and Δ*ppad* strains were incubated without fibrinogen and the supernatant visualised by silver staining. **(D)** All samples with visible protein bands from **A** were tested for citrullination (AMC) by immunoblotting. Controls for the 'pre' sample were performed in which the modification step (pre ctrl1) or secondary antibody (pre ctrl2) were omitted. (**E)** All samples from **A** were analysed for the presence of citrullinated peptides by dot blot. Crosses indicate areas on the membrane where samples were not applied.

The fact that the majority of fibrinogen had been degraded within minutes, and that only two citrullinated proteins/peptides could be detected by immunoblotting, suggested that the majority of generated peptides was smaller than the size limit of the peptide gels used for this analysis (around 3 kDa). Thus, a dot blot technique was applied that confirmed that wild-type *P. gingivalis* cells rapidly degrade and citrullinate fibrinogen into small citrullinated peptides (Fig. 3.12 E), and that both arginine-gingipains and PPAD are required, as no positive signals were observed with the Δ*ppad* and Δ*rgp* strains.

In order to identify the amino acid sequence of the citrullinated fibrinogen peptides and determine the position of the citrulline residue, peptides derived from the wild-type and Δ*ppad* strains were analysed by LC-MS/MS. A total of 30 peptides were identified, all derived from fibrinogen (see Appendix, p. 248). The majority of identified peptides were the product of proteolytic cleavage after either an arginine or lysine residue, which is consistent with the results described above, but further proteolytic processing by other peptidases, in particular at the amino-terminus after glycine-, alanine- and serine-residues, was also evident. In samples incubated with *P. gingivalis* wild-type but not in Δ*ppad*, four peptides were found that contained a carboxy-terminal citrulline residue (Figure 3.13):

(1) ^{1}ADSGEGDFLAEGGGV<u>Cit</u>^{16},
(2) ^{31}PAPPPISGGGY<u>Cit</u>^{42},
(3) ^{253}GGSTSYGTGSETESP<u>Cit</u>^{268},
(4) ^{540}ESSSHHPGIAEFPS<u>Cit</u>^{554}.

The corresponding carboxy-terminal arginine-containing peptides were detected only in the Δ*ppad*, but not in the wild-type (Table 3.2), suggesting that citrullination of fibrinogen is tightly linked to cleavage by arginine-gingipains. The combined data support the concept that target proteins such as human fibrinogen are cleaved by arginine-gingipains, generating suitable peptide substrates for subsequent citrullination at the exposed carboxy-terminal arginine residue by *P. gingivalis* PAD.

Figure 3.13 Sequences of citrullinated peptides from human fibrinogen generated after incubation with *P. gingivalis*. Amino acid sequences of human fibrinogen α-chain (Fib A, SwissProt entry P02671) and β-chain (Fib B, SwissProt entry P02675) are shown. Peptides were detected using tandem mass spectrometry. Citrullinated peptides detected after incubation of fibrinogen with *P. gingivalis* wild-type are underlined. Fibrinopeptide A and B and the thrombin cleavage sites are indicated. Cit = citrulline.

To further test this concept, analogous experiments were performed using recombinant human α-enolase. Similar to fibrinogen, enolase was rapidly degraded by the WT, Δ*ppad* and less so by the Δ*rgp* strain (Figure 3.14). Using immunoblotting on peptide SDS-PAGE gels (Figure 3.14 B) and dot blot (Figure 3.14 C), no citrullination could be detected. Analysis of the WT and Δ*ppad* derived samples by mass spectrometry revealed only one citrullinated peptide in the WT (^{41}TGIYEALELCit^{50}, Fig. 3.14 D and Table 3.2), amongst a total of 17 detected peptides (for a full list please see Appendix, p. 250). The arginine-containing counterpart of this citrullinated peptide (^{41}TGIYEALELR^{50}) was detected in the samples incubated with Δ*ppad*. Analogous to fibrinogen, no peptides with carboxy-terminal arginine were detected in the WT samples. The proportion of peptides cleaved at residues other than arginine and lysine was higher than that found in fibrinogen, suggesting extensive cleavage by non-argine/lysine specific peptidases. Combined with the higher relative number of lysine residues in α-enolase (8.8 % versus 6.9 % in fibrinogen), this might result in the generation of short peptides, some of which might be citrullinated but too short to be detectable using the methods employed. Therefore, *P. gingivalis* Δ*kgp* was incubated with α-enolase (Figure 3.14 A-C), and five citrullinated peptides were detected by mass spectrometry (Figure 3.14 D and Appendix. p. 251), confirming that PPAD is able to citrullinate α-enolase peptides. Using the

AMC dot blot, which relies on long peptides that are hydrophobic enough to bind to the membrane, weak positive signals were observed with the Δ*kgp* strain, which decreased with time (Figure 3.14 C), again suggesting extensive proteolytic degradation by other proteinases.

Table 3.2: Mass spectrometry analysis of citrullinated peptides generated after incubation of human fibrinogen or α-enolase with *P. gingivalis* wild-type (WT) and their respective arginated peptides generated with the Δ*ppad* strain.

Sequence[1]	Protein	Location (amino acids)	Strain	*m/z* ratio[2]	Mascot Score
-.ADSGEGDFLAEGGGVCit.G	Fibrinogen α-chain	1-16	WT	769.29 (+2)	56
-.ADSGEGDFLAEGGGVR.G			Δ*ppad*	768.77 (+2)	99
R.GGSTSYGTGSETESPCit.N	Fibrinogen α-chain	253-268	WT	546.91 (+3)	51
R.GGSTSYGTGSETESPR.N			Δ*ppad*	786.82 (+2)	78
K.ESSSHHPGIAEFPSCit.G	Fibrinogen α-chain	540-554	WT	819.83 (+2)	98
S.SHHPGIAEFPSR.G		543-554	Δ*ppad*	445.54 (+3)	39
R.PAPPPISGGGYCit.A	Fibrinogen β-chain	31-42	WT	585.30 (+2)	26
R.PAPPPISGGGYR.A			Δ*ppad*	584.77 (+2)	65
S.TGIYEALELCit.D	α-enolase	41-50	WT	583.30 (+2)	24
S.TGIYEALELR.D			Δ*ppad*	582.79 (+2)	56

[1] Amino acids in the uncleaved proteins located carboxy- and amino-terminal to the identified peptides are indicated to demonstrate sites of proteolytic cleavage. Identified citrulline residues (Cit) are underlined.

[2] *m/z*: mass-to-charge; numbers in brackets: peptide ion charge state

For detailed analysis and explanations please see Appendix (p. 244).

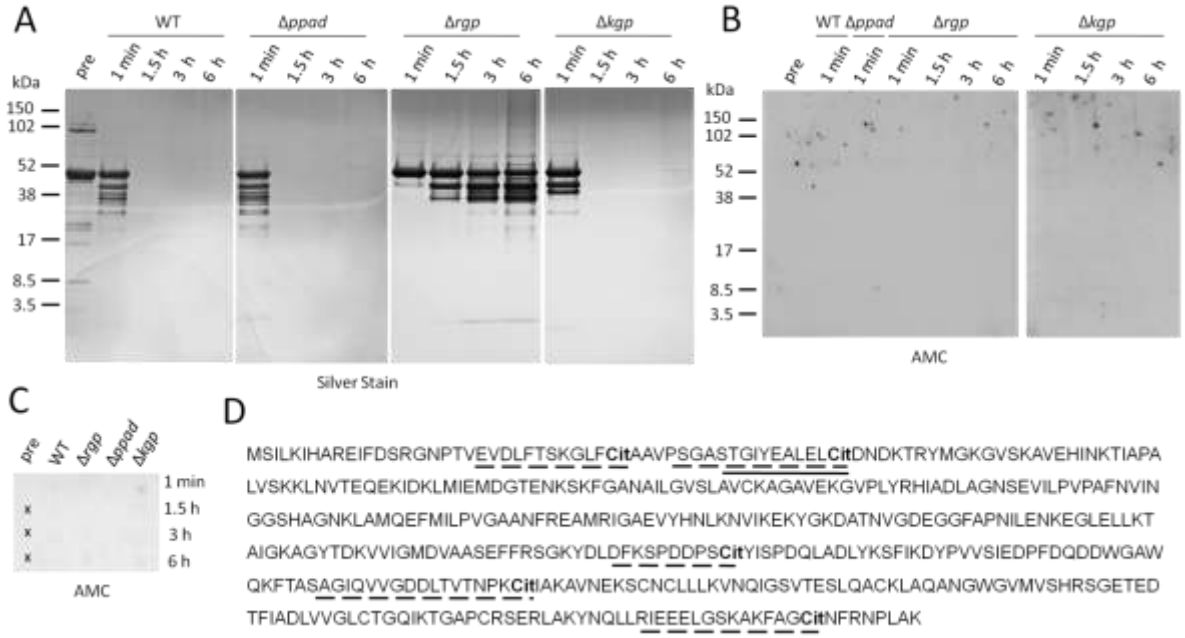


Figure 3.14 Human α-enolase is rapidly cleaved by *P. gingivalis* gingipains and citrullinated peptides are detectable by mass spectrometry. (A) Experiments were performed analogous to fibrinogen (Figure 3.11). Fragments of human α-enolase after incubation with *P. gingivalis* wild-type (WT), mutants lacking peptidylarginine deiminase (Δ*ppad*), arginine-gingipains (Δ*rgp*), or lysine-gingipains (Δ*kgp*), for 1 min, 1.5, 3 and 6 hours, were resolved by SDS-PAGE and visualised using silver staining. An enolase sample prior to incubation with *P. gingivalis* served as control (pre). **(B)** All samples with visible protein bands from **A** were tested for citrullination (AMC) by immunoblotting. **(C)** All samples from **A** were analysed for citrullinated peptides by dot blot. Crosses indicate areas on the membrane where samples were not applied. **(D)** Citrullinated peptides detected by mass spectrometry after incubation of enolase with *P. gingivalis* wild-type and Δ*kgp* are underlined with a continuous and dashed line, respectively. Amino acid sequence of human α-enolase: SwissProt entry P06733. Cit = citrulline.

These results were further confirmed using the colorimetric method to detect citrulline. Figures 3.15 A and B show that the citrulline levels are elevated in both fibrinogen and α-enolase samples incubated with WT and Δ*kgp* and increase with incubation time, but are comparably low in Δ*ppad* and Δ*rgp*. To exclude false-positive signals stemming from free L-citrulline as a result of normal *P. gingivalis* metabolism, controls were performed in which bacteria were incubated alone without protein; these showed no positive signals in either strain in the citrullination assay (Figure 3.15 C).

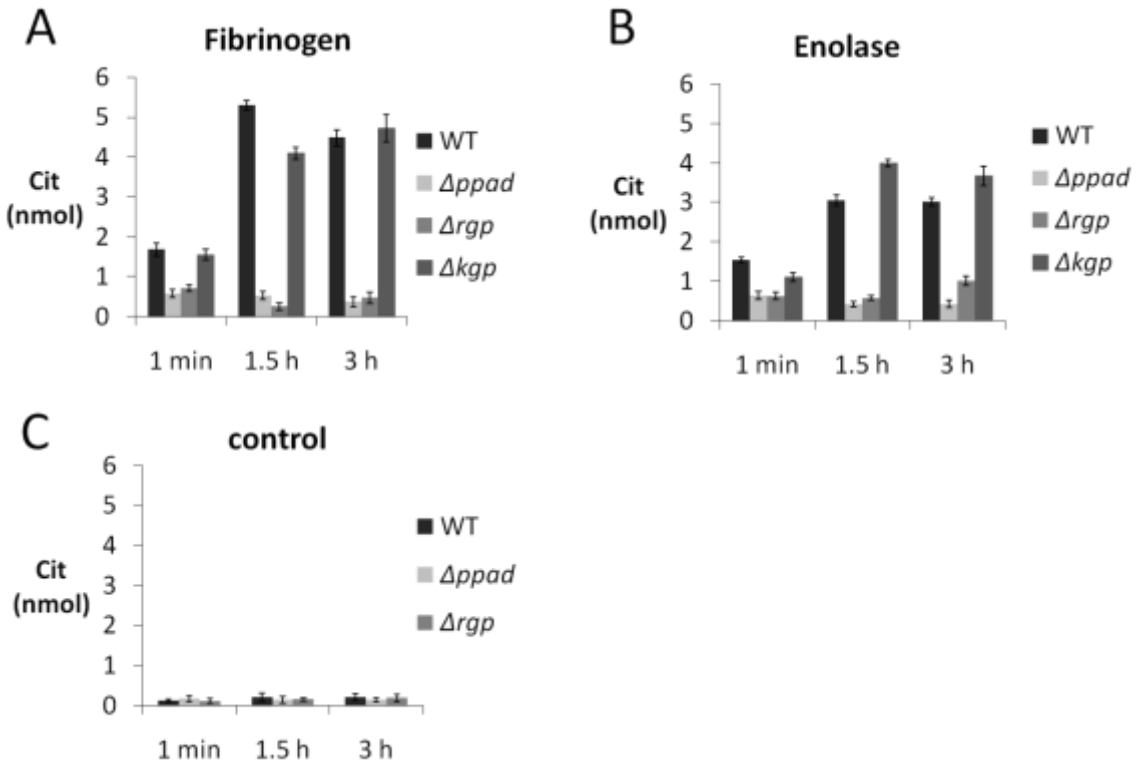


Figure 3.15 Short citrullinated fibrinogen and α-enolase peptides are generated by *P. gingivalis* in the presence of arginine-gingipains and peptidylarginine deiminase. Human fibrinogen (A), α-enolase (B), or bacteria alone without protein (C) were incubated with *P. gingivalis* wild-type (WT), mutants lacking peptidylarginine deiminase (Δ*ppad*), arginine-gingipains (Δ*rgp*), or lysine-gingipains (Δ*kgp*) for 1 min, 1.5 and 3 hours. Bacteria were removed and 60 µl of the supernatant were analysed for citrulline content in nmol by the colorimetric method as described in 'Materials and Methods'. A brief spin prior to photometric read-out ensured precipitated protein did not interfere with analysis. Due to sample volume restrictions Δ*kgp* samples are not present in the controls.

Analysis of antibody titres to *P. gingivalis*-generated citrullinated fibrinogen peptides

Having identified the sequence of citrullinated peptides that are generated by *P. gingivalis* from fibrinogen and α-enolase, it was of interest to investigate whether an antibody response to these peptides exists in RA patients (Figure 3.16), as all identified human citrullinated epitopes that are targeted by anti-citrullinated protein antibodies in RA harbour internal and not C-terminal citrulline residues.

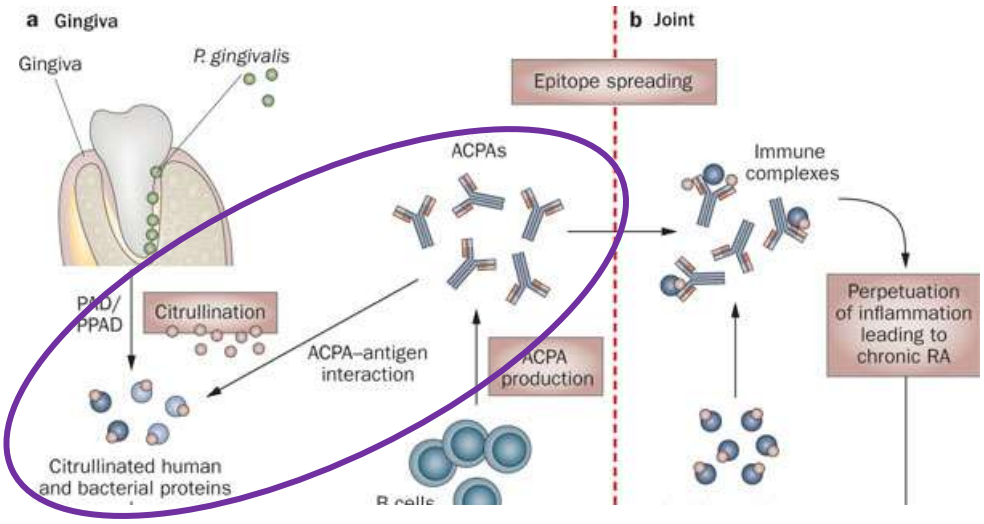


Figure 3.16 Schematic illustration of the aetiological hypothesis for *P. gingivalis* involvement in RA, with the focus on antibody reactivity (purple circle). For the full picture and a detailed explanation please refer to Figure legend 1.10.

A pilot study using 48 RA and 33 healthy control serum samples was conducted using peptide ^{540}ESSSHHPGIAEFPSCit^{554} derived from the fibrinogen α-chain, which is generated after incubation of fibrinogen with *P. gingivalis*, as identified by mass spectrometry (see Table 3.2). To determine the influence of the position of the citrulline residue within the peptide, the same cohort was evaluated using a peptide with two additional amino acid residues at the C-terminus (^{540}ESSSHHPGIAEFPSCitGK^{556}), representing an epitope that would be generated by human PADs. An N-terminally truncated version of this peptide (^{542}SSHHPGIAEFPSCitGK^{556}) had been described in a previous publication as one of the 11 most frequent fibrinogen target epitopes for anti-citrullinated protein antibodies in RA, reactive with 25% (5 out of 20) of RA serum samples (Sebbag *et al.* 2006). The same epitope was also identified by another study as part of one of the three major epitopes reactive with RA serum and it was further confirmed that it is generated by the activity of human PADs (Van Beers *et al.* 2010). In a mass spectrometry-based study of fibrinogen isolated from rheumatoid arthritis synovial tissue, two citrullinated peptides were identified, both harbouring regions that were shown to be citrullinated C-terminally by *P. gingivalis* in this thesis (^{540}ESSSHHPGIAEFPScitGK^{556} and ^{22}KcitEEAPSLcitPAPPPISGGGYcitAcitPAK^{47} (Hermansson *et al.* 2010); underlined regions are peptides generated by *P. gingivalis* with only the C-terminal citrulline present, as identified in this study).

Both peptides ('XXX-Cit' = ^{540}ESSSHHPGIAEFPSCit^{554}; 'XXX-Cit-XX' = ^{540}ESSSHHPGIAEFPSCitGK^{556}), along with their arginine-containing counterparts as controls ('XXX-Arg' = ^{540}ESSSHHPGIAEFPSR^{554}; 'XXX-Arg-XX' = ^{540}ESSSHHPGIAEFPSRGK^{556}) were tested for binding to IgG and IgA antibodies in RA and healthy control serum by indirect ELISA. Bound serum IgA was evaluated to account for the possibility of a mucosa-driven immune response. As only citrulline-specific responses were of interest, the OD values of the analogous arginine control peptides were substracted from the citrulline-peptide OD values (Figures 3.17-19).

Using the Fisher's Exact test, the results show that a statistically significant difference in the number of positive (above cut-off) samples between RA and healthy controls was found only with the IgG response to the XXX-Cit-XX peptide (46% vs. 0%; $p<0.0001$), but not with the C-terminal citrulline peptide XXX-Cit, or the IgA response to either peptide (Figure 3.17).

Comparing all data points (as opposed to the numbers of positives) in the RA versus healthy control groups, using the Mann-Whitney U test, statistically significant differences were found only with the XXX-Cit-XX IgG response (***: $p<0.0001$) and marginally the IgA response (*: $p=0.0487$). Analysing the graphs visually, however, it seems surprising that the difference between RA versus healthy controls is significant for XXX-Cit-XX (IgA) but not XXX-Cit (IgG). This can be explained by the employed statistical test, highlighting the limitations of statistical tests. The test applied here, the Mann-Whitney U test, is used for non-Gaussian (and non-parametric in general) population and assesses the null hypothesis that two independent samples of observation have equally large values. To calculate statistical values, it uses the ranks of the data from both groups rather than their absolute values. As many values in the healthy control population with XXX-Cit (IgG) are marginally higher than the ones in RA, despite all of them being below threshold, and are therefore assigned higher ranks even if the differences in absolute values are small, the difference is calculated as not significant. In contrast, most values between RA and healthy controls in XXX-Cit-XX (IgA) are clustered equally around zero, with the exeption of the three values above threshold, resulting in a marginally statistically significant result ($p=0.0487$). The test does not take into account that the absolute positive values in XXX-Cit (IgG) are higher than the absolute positive values in XXX-Cit-XX (IgA) as it compares relative ranks alone.

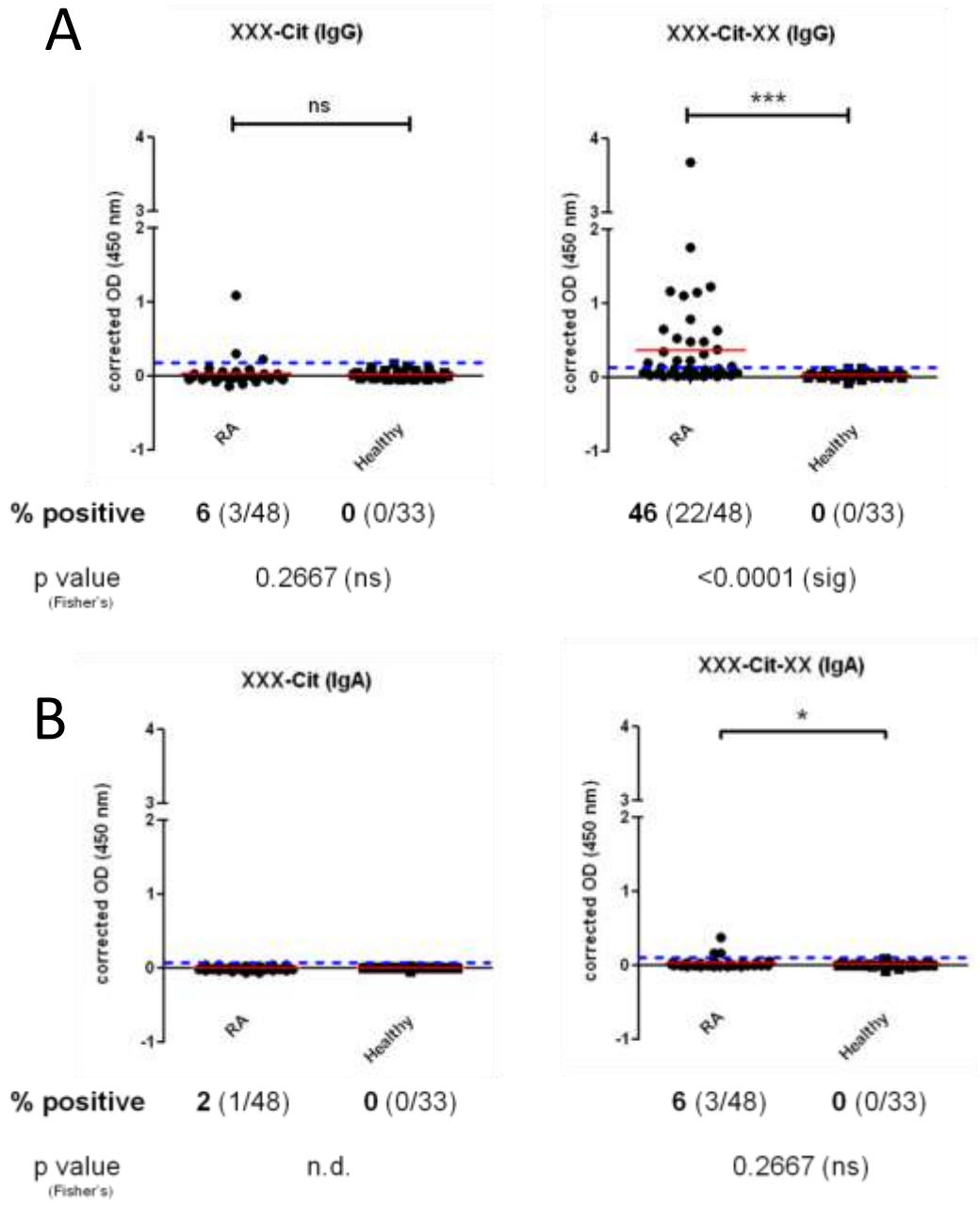


Figure 3.17 Analysis of citrulline-specific antibody titres to two citrullinated fibrinogen-α peptides (FibA). Peptides were synthesised and purified by a commercial supplier. IgG (A) and IgA responses (B) to citrullinated peptides and the analogous arginine-containing control peptides were analysed in serum samples from rheumatoid arthritis patients (n=48) and healthy controls (n=33). Bound serum IgG/IgA was measured by indirect ELISA and data presented as citrullinated peptide $OD_{450\ nm}$ - arginine peptide $OD_{450\ nm}$. One dot represents one sample. The red line indicates mean OD values. Cut-offs were calculated based on the 98th percentile of the healthy control sample values for the respective peptides and are indicated with a blue dashed line. The Mann-Whitney U test was used to calculate p values for all RA versus Healthy values. The Fisher's Exact test was used to compare % positives in RA versus Healthy samples, where indicated. ns = not significant; sig = significant; * = $p<0.05$; *** = $p<0.0001$; n.d. = not done.

The selection of cut-off points is another controversial issue. Here, the 98th percentile of each healthy control population for each peptide was selected as a cut-off point. As the sample size was relatively small, this equalled the highest value in the healthy control group. This approach was chosen as the purpose was to establish whether there is a specific response in RA patients compared to healthy controls. Therefore, by definition the proportion of positives in each healthy control group is 0%.

A second peptide was evaluated for IgG reactivity in an analogous manner. Again, the peptide sequence was chosen from the list of identified *P. gingivalis*-mediated citrullinated fibrinogen peptides (^{31}PAPPPISGGGYCit^{42}; Table 3.2). The C-terminally extended analogue was ^{31}PAPPPISGGGYCitAR^{44}. A similar peptide containing two additional citrulline residues (^{30}CitPAPPPISGGGYCitACit^{44}) was described as the most frequent fibrinogen target epitope for anti-citrullinated protein antibodies in RA, reactive with 70% (14 out of 20) of RA serum samples (Sebbag *et al.* 2006). Two other publications confirmed this finding. Van Beers *et al.* showed that this epitope, containing at least two (Cit-42 and Cit-44) but ideally all three (Cit-42, -44, -30) citrulline residues constitutes one of the three major citrullinated epitope regions in human fibrinogen (Van Beers *et al.* 2010). Zhao *et al.* identified the peptide sequence ^{24}EEAPSLcitPAPPPISGGGYRA^{43} by mass spectrometry from fibrinogen-containing immune complexes isolated from the serum of a patient with RA (Zhao *et al.* 2008).

The data, presented in Figure 3.18, give a similar picture to the previous results: a statistically significant difference in the number of positive samples between RA and healthy controls was found with the XXX-Cit-XX peptide (23% vs. 0%; p=0.0040), but not with XXX-Cit peptide. Comparing all data points in the RA versus healthy control groups using the Mann-Whitney U test, a statistically significant difference was found only with the XXX-Cit-XX peptide (*: p=0.0345).

In summary, these data show that the position of the citrulline residue in these two peptides is crucial. Moving the citrulline residue towards the N-terminus by just two positions significantly increases IgG antibody reactivity (Figure 3.19). Further, there is some overlap in the reactivity of a particular serum samples to the XXX-Cit and XXX-Cit-XX peptides within this peptide pair (colour-matched dots in Figure 3.19). To determine the extent of cross-reactivity between the peptide pairs, cross-absoprtion studies should be performed.

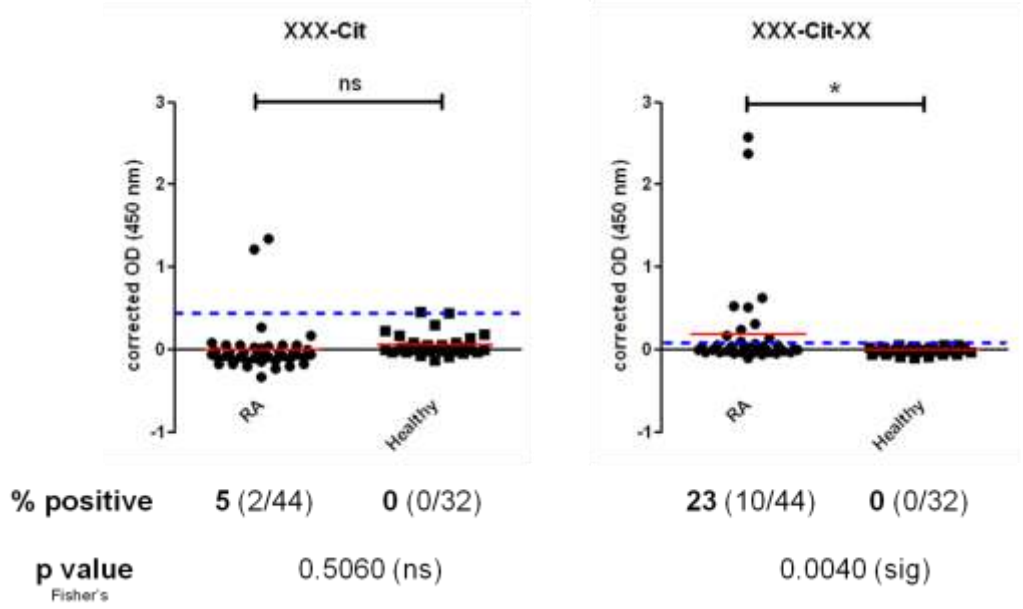


Figure 3.18 Analysis of citrulline-specific antibody titres to two citrullinated fibrinogen-β peptides (FibB). IgG responses to citrullinated peptides and the analogous arginine-containing control peptides were analysed in serum samples from rheumatoid arthritis patients (n=44) and healthy controls (n=32). Bound serum IgG was measured by indirect ELISA and data presented as citrullinated peptide $OD_{450\ nm}$ - arginine peptide $OD_{450\ nm}$. One dot represents one sample. The red line indicates mean OD values. Cut-offs were calculated based on the 98th percentile of the healthy control sample values for the respective peptides and are indicated with a blue dashed line. The Mann-Whitney U test was used to calculate p values for all RA versus Healthy values. The Fisher's Exact test was used to compare % positives in RA versus Healthy samples, where indicated. ns = not significant; sig = significant; * = $p<0.05$.

A much lower reactivity of RA serum with peptide ^{31}PAPPPISGGGYCitAR^{44} was observed (23%), compared to the previously published reactivity of 70% with peptide ^{30}CitPAPPPISGGGYCitACit^{44}. This is most likely due to the two additional citrulline residues present in this peptide. As the numbers in this pilot study were small, it cannot be concluded whether the findings have a biological meaning or not, but it would be worth testing a higher sample size, including cross-absorption studies, and a higher number of peptides to evaluate the possibility that a small proportion of RA patients indeed has specific antibodies to peptides with C-terminal citrullines, as such peptides are to date not known to result from human physiological activities (PADs citrullinate internal arginines and there are no proteases known to cleave after citrulline residues). Further, possible technical limitations should be investigated, such as suboptimal conformation or binding of the C-terminal

citrulline peptides to the plate. This could be evaluated by linking the peptides to biotin at the amino-terminus and binding to streptavidin-coated ELISA wells.

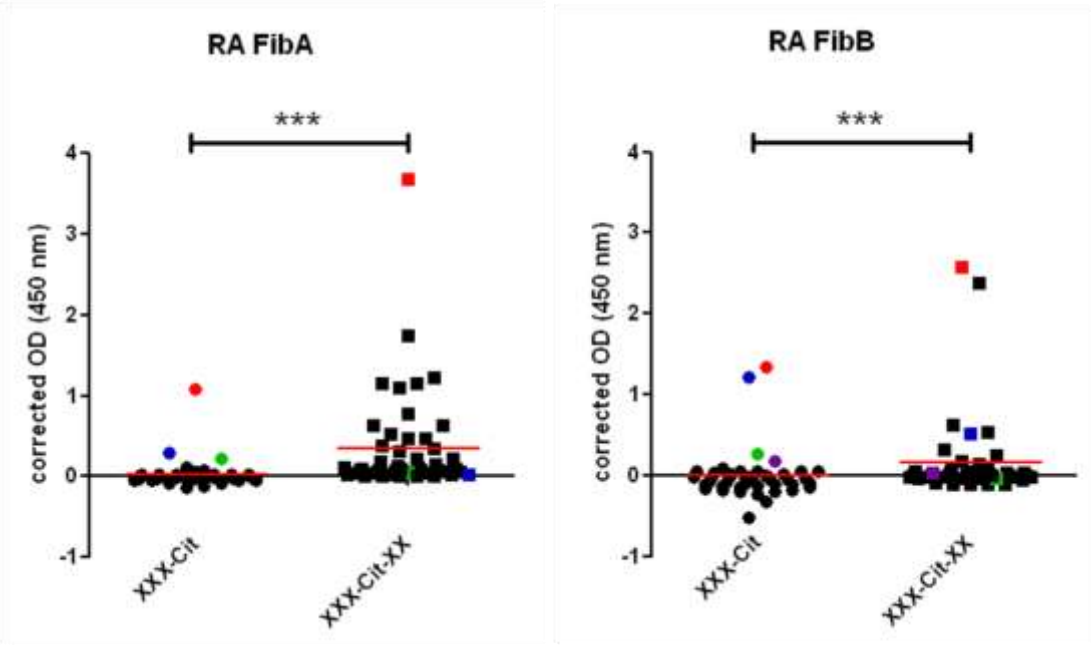


Figure 3.19 Comparison of IgG antibody reactivity of RA serum to two sets (FibA, FibB) of C-terminal (XXX-Cit) and internal (XXX-Cit-XX) citrullinated peptides. Analysis was performed as described in Figures 3.17 and 3.18. One dot represents one serum sample. Dots of the same colour within each panel represent the same serum sample. The red line indicates mean OD values. The Mann-Whitney U test was used to calculate p values for XXX-Cit versus XXX-Cit-XX reactivities. *** = $p<0.0001$ for FibA and $p=0.0009$ for FibB.

Summary

Endogenous protein citrullination in P. gingivalis and other oral bacteria

- PPAD is functional in *P. gingivalis* and unique amongst the bacterial strains tested: Endogenous protein citrullination was abundant in *P. gingivalis* wild-type and clinical isolates but lacking in ten other oral bacterias strains tested; PPAD gene was detected by PCR in all tested *P. gingivalis* strains; pattern of expressed citrullinated proteins varied slightly with growth stage and between strains
- Majority of citrullinated proteins in *P. gingivalis* is associated with the periplasm and membrane fractions

- Deletion of the PPAD gene resulted in complete abrogation of protein citrullination, providing proof that only one functional peptidylarginine deiminase enzyme is present in *P. gingivalis*
- Inactivation of arginine-gingipains, but not lysine-gingipains, led to significantly decreased levels of endogenous citrullination, indicating that arginine-gingipains are involved in the process

P. gingivalis-mediated citrullination of human proteins

- Incubation of *P. gingivalis* with human fibrinogen or α-enolase resulted in degradation of these proteins and, in wild-type but not PPAD or arginine-gingipain knockout strains, citrullination of the resulting peptides at carboxy-terminal arginine residues only
- Arginine-gingipains are required for citrullination by *P. gingivalis*, presumably by proteolytic cleavage of proteins at arginine-X peptide bonds, followed by citrullination of the carboxy-terminal arginines by PPAD

Antibody reactivity of RA serum with C-terminal citrulline peptides

- 5-6% of RA serum samples contain IgG antibodies that recognise the tested C-terminal citrulline peptides, compared to 23-46% recognising the corresponding internal citrulline peptide
- Statistically significant differences between RA and healthy controls could only be demonstrated with peptides containing internal citrulline
- Results need to be confirmed with a larger number of samples and peptides but might indicate that antibodies towards C-terminal citrulline peptides exist in a subset of patients with RA where *P. gingivalis* plays a role in disease development

Chapter 4
Expression and characterisation of recombinant *P. gingivalis* PAD

Background

The results presented in the previous chapter indicated that PPAD is crucial for citrullination of endogenous bacterial and exogenous human proteins and that citrullination occurs primarily at C-terminal arginines. To unambiguously answer the question as to which epitopes can be created by PPAD, the active recombinant enzyme was cloned and expressed in *E. coli*. Mutants were created by site-directed mutagenesis to pin down the catalytic residues. Further, a number of candidate compounds were tested as potential inhibitors of PPAD.

When this work was started, no publications or protocols had been available for the cloning and expression of enzymatically active PPAD. Further, it was not known what the requirements for catalytic activity of PPAD are. In the first paper on PPAD, the authors isolated a truncated form from the supernatant of a mutant *P. gingivalis* strain that secretes all membrane-bound proteins into the medium. Therefore, it was unclear whether this represents the wild-type native enzyme or whether it is an artefact resulting from the mutated secretion apparatus or the co-purification with gingipains (McGraw *et al.* 1999).

Results

Cloning and expression of PPAD

A strategy to clone both the full-length PPAD gene and the truncated form was developed (Figure 4.1). The pET48b(+) expression vector was chosen as host vector as it contains an N-terminal cleavable His-tag for easy detection and purification and a thioredoxin (Trx)-tag (see Appendix). Thioredoxin-tagged proteins were shown to possess greater solubility (LaVallie *et al.* 1993). Thioredoxin also facilitates the reduction of proteins through the reversible oxidation of its own active center dithiol to a disulfide, thereby catalysing dithiol-disulfide exchange reactions. Hence, this vector was chosen as it was speculated that active PPAD had a reduced catalytic cysteine, and a greater amount of stable protein could potentially be

achieved using this strategy.

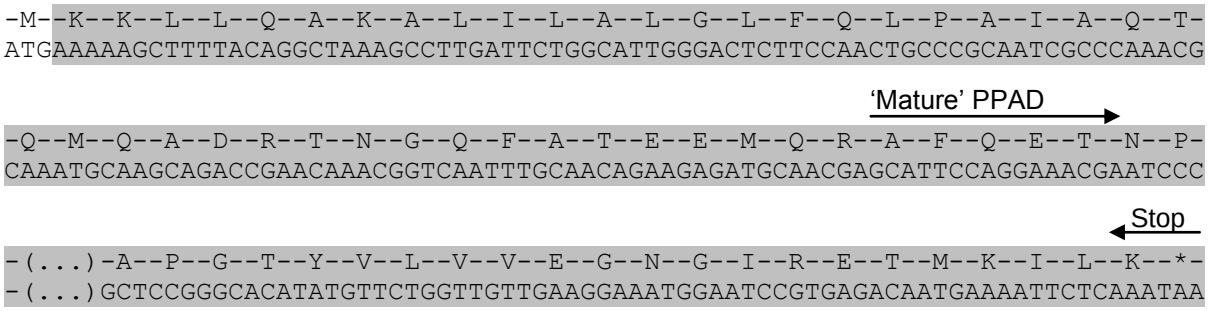


Figure 4.1 **Cloning strategy for PPAD.** Amino acids 1-50 and 534-556 together with the corresponding nucleotides are shown. Section highlighted in grey indicates the full-length PPAD construct (amino acids 2-556), arrow labelled 'mature PPAD' indicates the truncated construct (amino acids 44-556).

Initial attemps to clone PPAD suffered from a high frequency of multiple non-synonymous mutations, especially within repeat regions (for example ^{289}CAAAAAAAC^{297}, or ^{604}TCCTCTCTGTCTCAA^{618}). Site-directed mutagenesis to correct these mutations was of limited success, as the procedure introduced new mutations. Hence, several high-fidelity proof-reading polymerases, primer combinations, and PCR conditions were tested, eventually leading to a successful method using Finnzymes Phusion High-Fidelity DNA polymerase (Figure 4.2).

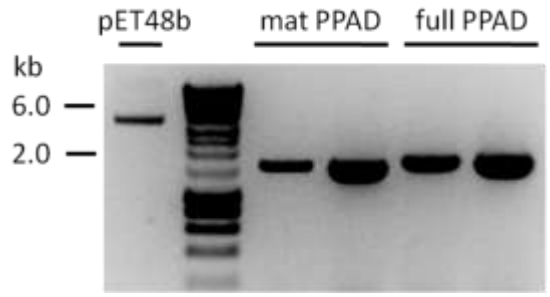


Figure 4.2 **Amplification PPAD inserts and vector.** Agarose gel after PCR amplification of the *P. gingivalis* W83 PAD gene (locus PG_1424), either the full-length ('full PPAD', 1.67 kb) or the truncated form ('mat PPAD', 1.54 kb), from genomic DNA. 3 µl and 6 µl from a total reaction volume of 100 µl were loaded. 30 ng of purified circular pET48b vector (5.6 kb) were loaded. The positions of the relevant DNA size markers (in kilobases, kb) are shown on the left.

Following restriction digest and ligation, progress was controlled by ligation PCR using primers annealing up- and downstream of the vector-insert junction. As demonstrated in Figure 4.3, ligation was successful for both PPAD inserts.

As it was unclear whether the PPAD gene product might cause toxicity in *E. coli*, a standard cloning strain (Novagen NovaBlue) and a strain with a tighter control of low-level gene expression by the *lacI*^{q} promoter (NEB 5-alpha F*'I*^{q}) were used for transformation. The number of colonies obtained was slightly higher (~30%) with NEB 5-alpha F*'I*^{q}. However, fewer colonies were obtained with the construct containing the 'mature PPAD' insert compared to the construct with full-length PPAD (factor 10), despite equal DNA amounts and all steps performed in parallel.

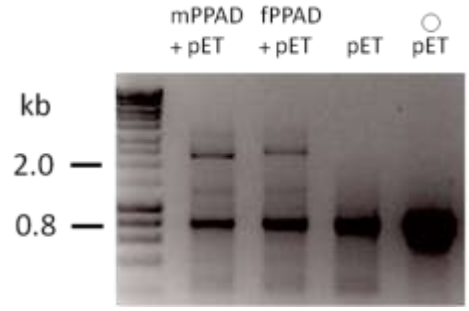


Figure 4.3 Ligation PCR of PPAD-pET48 constructs and controls. Agarose gel after ligation PCR using primers annealing up- and downstream of the insert-vector junction sites. Successful insertion of the PPAD inserts results in product of 2.3 (mature PPAD+pET48) and 2.5 kb (full-length PPAD+pET48). Vector lacking insert results in a 0.8 kb product. 'pET' stands for ligation control where linearised vector alone was used. 'pET ○' stands for a PCR control where circular vector was used as a positive control for the 0.8 kb product. The positions of the relevant DNA size markers (in kilobases, kb) are shown on the left.

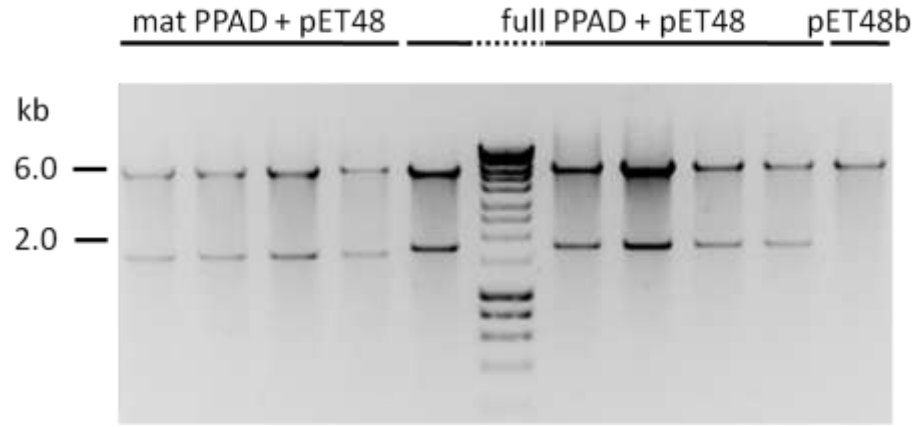


Figure 4.4 Successful cloning of PPAD-pET48 constructs. Agarose gel after diagnostic digest of PPAD-pET48 constructs containing either the 'mature PPAD' insert (1.54 kb) or the full-length PPAD insert (1.67 kb). On the far right, linearised pET48 vector (5.6 kb) was loaded as a control. The positions of the relevant DNA size markers (in kilobases, kb) are shown on the left.

Colony PCR was used to select *E. coli* clones that contained the correct constructs. For both constructs the success rate was 50%. A diagnostic restriction enzyme digest of the isolated plasmid DNA confirmed successful insertion of the insert DNA (Figure 4.4).

DNA sequencing confirmed the correct reading frame and the absence of mutations in one out of four constructs containing the 'mature PPAD' insert and in five out of five constructs containing the full-length PPAD insert (see Appendix). These were used for all subsequent experiments.

To increase chances of success, several transformation hosts were initially used to test for recombinant protein expression: Novagen BL21 (DE3), Lucigen C41 (DE3), and Lucigen C43 (DE3). Although transformation was successful with all three strains and both constructs, undegraded recombinant protein could only be obtained with the Novagen BL21 (DE3) host transformed with the full-length PPAD vector (Figure 4.5).

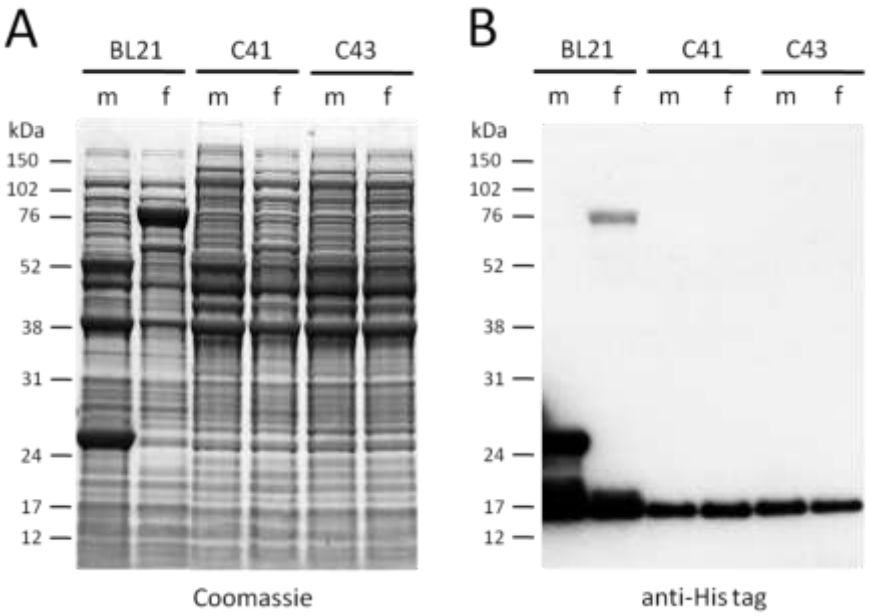


Figure 4.5 Screening for expression of recombinant PPAD. *E.coli* BL21 (DE3), C41 (DE3), or C43 (DE3) cells were transformed with constructs encoding Trx-His-mature PPAD ('m'; expected size 72 kDa) or Trx-His-full length PPAD ('f'; expected size 77 kDa) and recombinant protein expression was induced with IPTG. Cell lysates were then resolved on SDS-PAGE gels and total protein stained with Coomassie (A) or transferred onto membranes and immunblotted with anti-His tag antibody (B). The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

‘Mature’ PPAD with a predicted size of 72 kDa (including tags) could not be detected in any of the strains. It is possible that it was degraded, as a major His-tagged product at ~25 kDa and several minor products are visible on the anti-His tag immunoblot in the BL21 (DE3) strain (Figure 4.5 B). A protein of ~15 kDa is visible in all samples, which could be explained by aborted expression, resulting in expression of the tag alone (15 kDa). As a result, it was speculated that ‘mature’ PPAD is the enzymatically active form, causing toxicity to the *E. coli* cell and is therefore degraded. Hence, the BL21 (DE3) pLysS strain was transformed with both the ‘mature’ and full-length PPAD constructs and tested for expression. pLysS strains are recommended for the expression of potentially toxic proteins as the pLysS plasmid produces a T7 lysozyme that inhibits transcription by the T7 RNA polymerase, thereby reducing basal level expression of the gene of interest. However, no overexpression of either full-length or ‘mature’ recombinant PPAD could be detected with this strain (Figure 4.6).

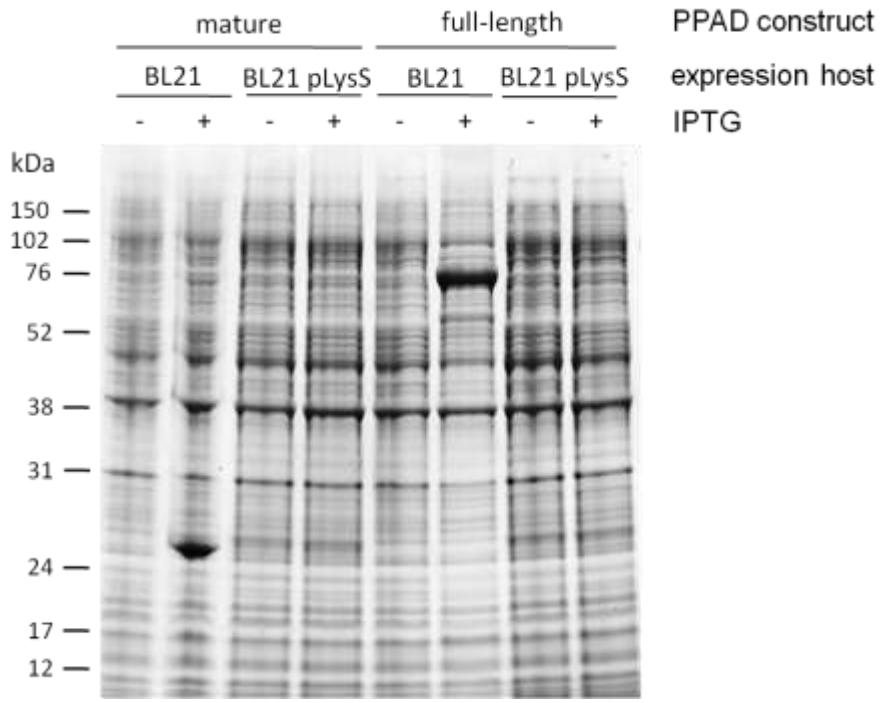


Figure 4.6 Comparison of the strains BL21 (DE3) and BL21 (DE3) pLysS for the expression of recombinant PPAD. *E.coli* BL21 (DE3) or BL21 (DE3) pLysS cells were transformed with constructs encoding Trx-His-mature PPAD (‘mature’; expected size 72 kDa) or Trx-His-full length PPAD (‘full-length’; expected size 77 kDa) and recombinant protein expression was induced with IPTG (‘+). No IPTG was added as a control (‘-‘). Cell lysates were resolved on an SDS-PAGE gel and total protein stained with Coomassie. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

BL21 (DE3) cells transformed with the full-length PPAD constructs were therefore chosen for subsequent experiments to assess the solubility of the recombinant protein, tag cleavage, and enzymatic activity.

For determination of protein solubility, two cell lysis methods were initially evaluated: standard sonication and a milder, detergent-assisted cell lysis ('BugBuster' reagent). The total cell lysates was further fractionated into native and denaturing fraction by centrifugation and all three samples were analysed by SDS-PAGE and Coomassie staining. The majority of the recombinant protein was present in the insoluble (denaturing) fraction. However, a thin overexpressed band corresponding in size to recombinant full-length PPAD was observed in the soluble (native) fraction lysed with detergent (Figure 4.7, arrow), but not in the soluble fraction lysed by sonication.

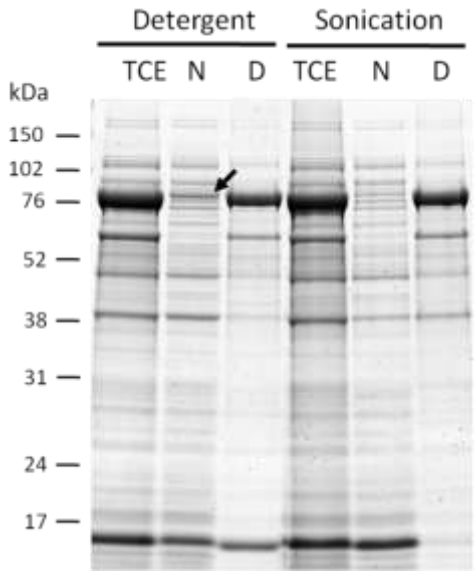


Figure 4.7 Solubility testing of recombinant PPAD. *E.coli* BL21 (DE3) cells were transformed with a construct encoding Trx-His-full length PPAD (expected size 77 kDa). Cells were lysed using a detergent-containing buffer or sonication. The resulting total cell extracts ('TCE') were further fractionated into native/soluble ('N') and denaturing/insoluble ('D') fraction by centrifugation. Samples were resolved on an SDS-PAGE gel and total protein stained with Coomassie. The arrow indicates the overexpressed protein in the native fraction lysed using detergents. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

To confirm this result, a protein pull-down of each native fraction using His-tag affinity resin (Talon) was performed. Figure 4.8 clearly demonstrates that more recombinant PPAD could

be extracted with the detergent method compared to sonication. The detergent method was therefore chosen for cell lysis in subsequent experiments.

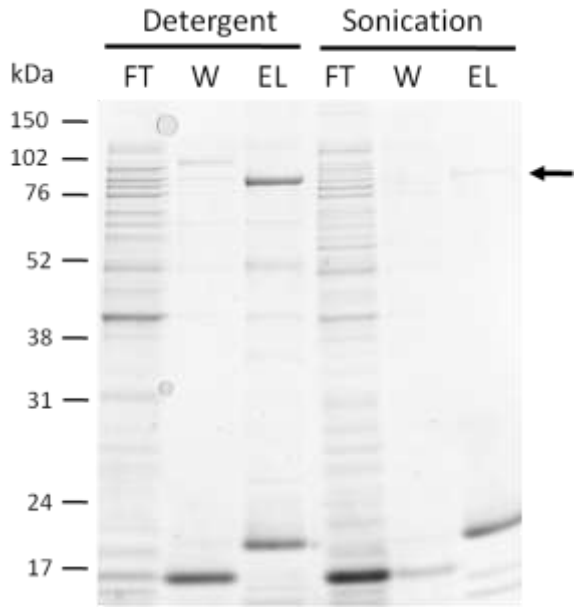


Figure 4.8 Pull-down of soluble PPAD. *E.coli* BL21 (DE3) overexpressing Trx-His-full length PPAD (expected size 77 kDa) were lysed using a detergent-containing buffer or sonication. The native/soluble fraction of each lysis method was subjected to protein pull-down using His-tag affinity resin. Flow-through ('FT'), wash ('W') and eluate fractions ('EL', containing recombinant PPAD) were collected and and visualised on an SDS-PAGE gel by Coomassie staining. The arrow indicates the size of recombinant PPAD. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

In an attempt to further increase the amount of soluble protein, the detergent-based extraction buffer (from hereon called BugBuster) was supplemented with either 10% glycerol, 5 mM β-mercaptoethanol, or both. Based on the results in Figure 4.9 it was concluded that 10% glycerol might increase the yield slightly.

The recombinant PPAD contained N-terminal tags (Trx and His) and had been generated with the aim of performing activity assays and producing a polyclonal anti-PPAD antibody. As it is known that tags can interfere with activity, and that immunisation with a tagged antigen may lead to antibodies reactive with the tag rather than the antigen, a protocol for efficient tag cleavage was attempted. The pET48b vector contains a human rhinovirus (HRV) 3C protease cleavage site, resulting in tag-free PPAD (predicted size 62 kDa) and only four additional vector-derived N-terminal amino acids (Gly-Pro-Gly-Tyr).

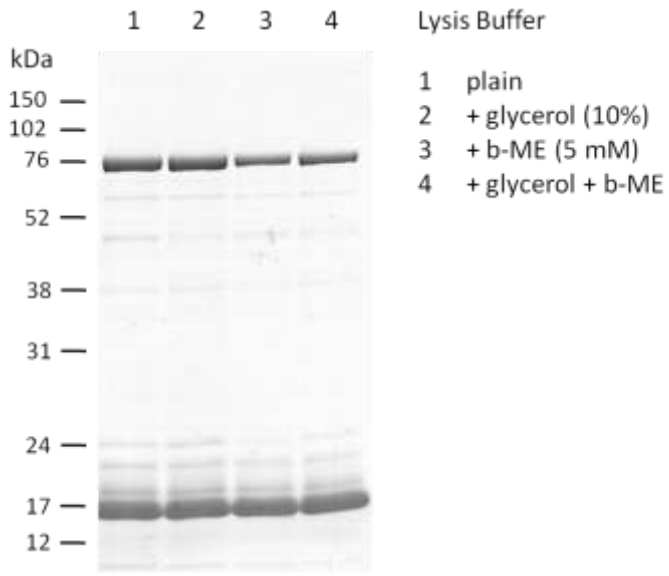


Figure 4.9 Optimisation of detergent-based buffer conditions for extraction of soluble PPAD. *E.coli* BL21 (DE3) overexpressing Trx-His-full length PPAD were lysed using a detergent-containing buffer without any supplements (lane 1), or with 10% glycerol (lane), 5 mM β-mercaptoethanol (lane 3), or both 10% glycerol and 5 mM β-mercaptoethanol (lane 4). Trx-His-PPAD was then pulled down using His-tag affinity resin and analysed on an SDS-PAGE gel and visualised by Coomassie staining. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

In a small-scale experiment, purified Trx-His-PPAD was incubated with His-HRV 3C protease with various protease-to-protein ratios and incubation times (24 hours). However, only a small fraction of the tagged protein was cleaved (Figure 4.10 A), even with a high protease-to-protein ratio (1:5) and prolonged incubation time. Therefore, various buffer conditions were tested, as recommended by the manufacturer (Figure 4.10 B). 100% cleavage could not be achieved under any conditions. The highest cleavage efficiency was attained with the standard buffer (50 mM Tris, 150 mM NaCl) supplemented with DTT (1 mM), EDTA (1 mM) and Triton X-100 (0.1%) (Figure 4.10 B lane 14).

These optimised buffer conditions were used to re-evaluate the protease-to-protein ratio, since a ratio of 1:5, as used for the experiment in Figure 4.10 B is too costly for large-scale production. However, as shown in Figure 4.11 A, satisfactory results (>50% cleavage) were achieved only with the 1:5 ratio; hence, this condition was used to generate cleaved PPAD for activity measurements (Figure 4.11 B, 'FT1').

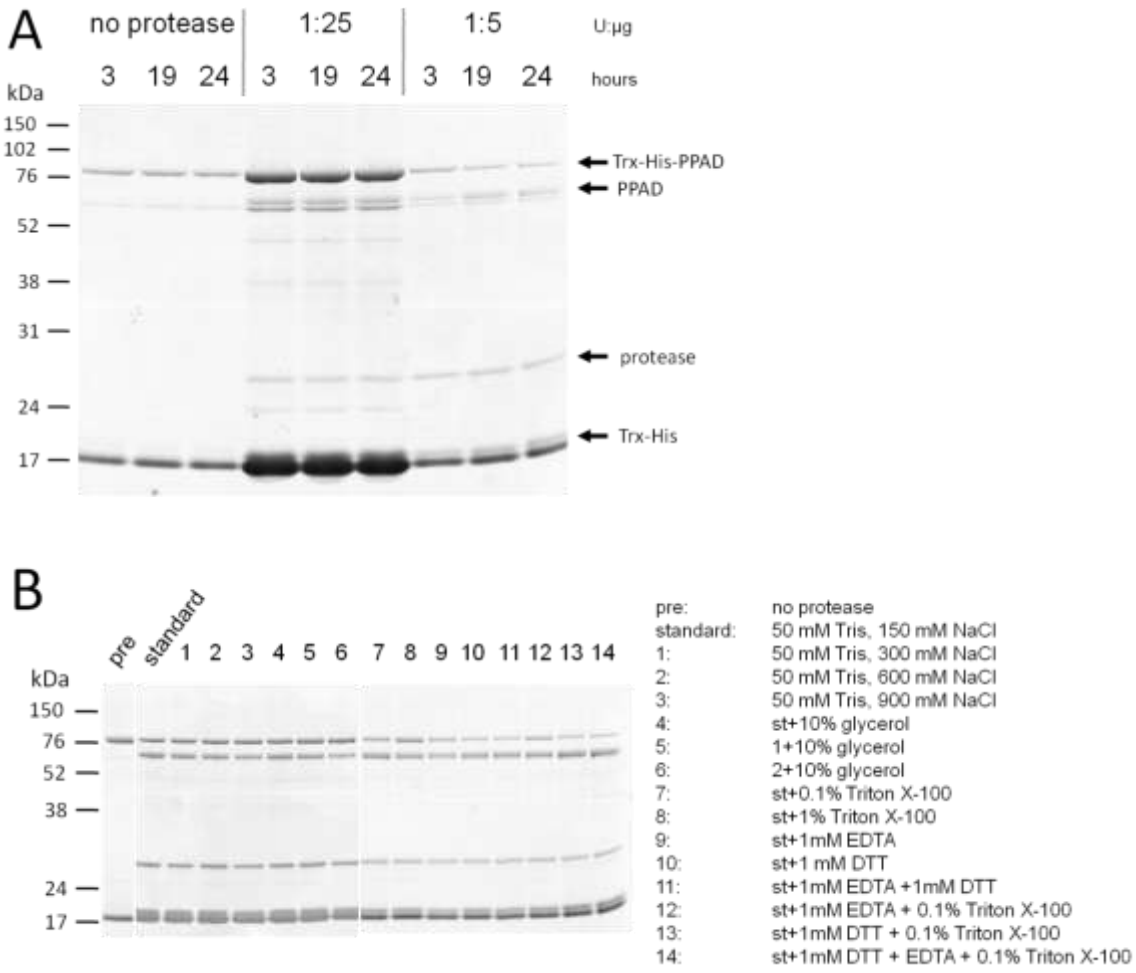


Figure 4.10 Optimisation of tag cleavage of Trx-His-PPAD. (A) Purified Trx-His-PPAD was incubated for 3, 19, or 24 hours without protease (8 µg PPAD, no protease), with a 1:25 ratio of protease-to-protein (2 U protease, 50 µg PPAD), or with a 1:5 ratio of protease-to-protein (2 U protease, 8 µg protein) in standard cleavage buffer (50 mM Tris-HCl, 150 mM NaCl) and visualised on an SDS-PAGE gel by Coomassie staining. The arrows indicate the expected size of the reaction components. **(B)** Purified Trx-His-PPAD (10 µg) was incubated for 24 hours without protease ('pre'), or with a 1:5 ratio of protease-to-protein (2 U protease) in various buffer conditions, as indicated on the right, and visualised on an SDS-PAGE gel by Coomassie staining. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

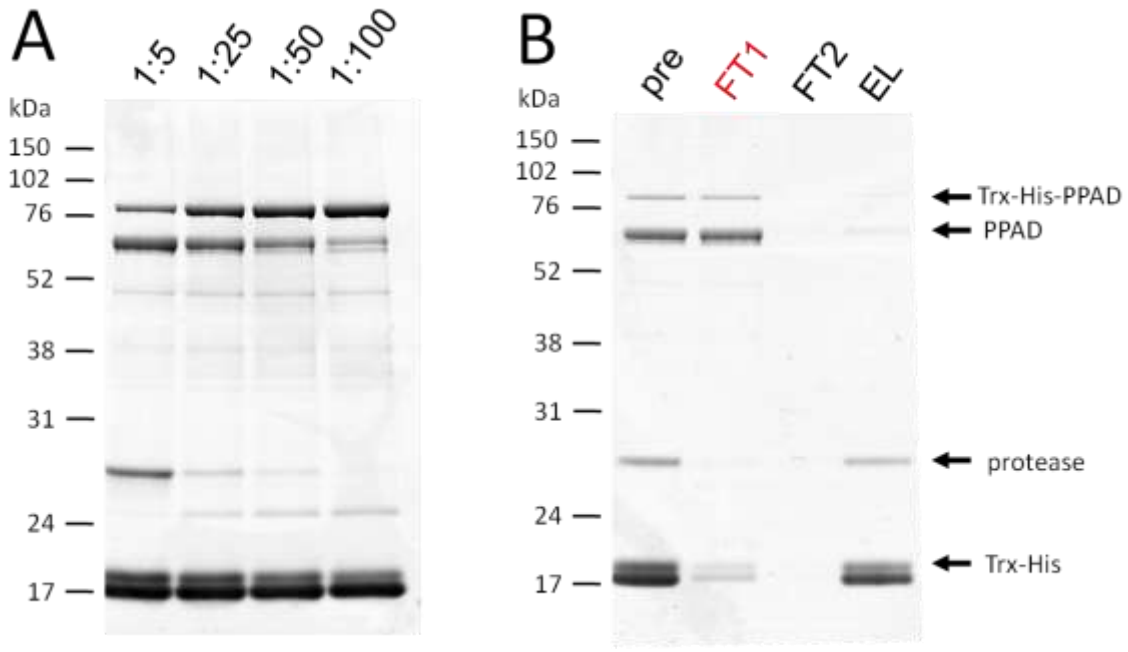


Figure 4.11 Generation of cleaved PPAD for activity assays. (A) Purified Trx-His-PPAD (10 µg) was incubated for 24 hours with various protease-to-protein (U:µg) and visualised on an SDS-PAGE gel by Coomassie staining. **(B)** Purified Trx-His-PPAD (10 µg) was incubated for 48 hours and a 1:5 ratio of protease-to-protein in a buffer containing 50 mM Tris, 150 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.1% Triton X-100 ('pre'). His-tagged protease and protein truncations were then removed using His-tag affinity resin, and cleaved PPAD collected in the flow-through ('FT1'). A second wash ensured complete collection of cleaved PPAD ('FT2'). Bound entities were subsequently eluted from the resin ('EL'). The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Activity studies with recombinant PPAD

In *vitro* enzymatic activity of native PPAD isolated from the culture supernatant of *P. gingivalis* has been previously evaluated (McGraw *et al.* 1999) using free L-arginine, synthetic arginine derivates (benzoyl-L-arginine [BA], benzoyl-L-arginine-ethylester [BAEE], benzoyl-L-arginine amide [BAA]), a short peptidylarginine derivates (benzoyl-glycyl-L-arginine) and the short peptide bradykinin (Figure 4.12). Enzyme activity was expressed as the Vmax/Km quotient, which serves as an index for the potential enzyme activity. The highest activity was reported for benzoyl-L-arginine (35.9×10^{-3}), followed by benzoyl-glycyl-L-arginine (6.6×10^{-3}) and bradykinin (2.5×10^{-3}), and lower activities with BAA, L-arginine, and BAEE ($0.1 - 0.5 \times 10^{-3}$), demonstrating a preference of this enzyme for peptidylarginine substrates with an unsubstituted carboxy-terminal arginine. Nonetheless,

BAEE had been used as the standard substrate in all subsequent assays as it is easy to work with (solubility, pH) and inexpensive. Activity had been demonstrated within a wide pH range (3-11), with optimum activity at ~ pH 9.3. McGraw *et al.* also reported that addition of flavin nucleotides was essential for enzyme stability and increased enzymatic activity.

Figure 4.12 Structural formulae for a selection of synthetic PPAD substrates.

Based on this information, BAEE was chosen as a substrate for initial activity assays in this thesis, in a CHES buffer system (pH 9.5). No additional components except purified, cleaved PPAD were added to the assay mix. Citrulline was quantified using a colorimetric method (Knipp *et al.* 2000), which was first evaluated using L-arginine and L-citrulline as it had not been used in this laboratory before. It was found that absorbance at 540 nm is linear with L-citrulline concentration (0-400 μM) and that L-citrulline can be differentiated from L-arginine at a concentration >10 μM (0.6 nmol in a total reaction volume of 60 μl), which was satisfactory for the purposes of this study (Figure 4.13).

In a second evaluation experiment, commercially available rabbit skeletal PAD was used with BAEE as substrate. A standard curve was created using L-citrulline to quantify the obtained citrulline levels. The results (Figure 4.14) show a near-linear correlation between substrate concentration and PAD-catalysed citrulline formation within the range examined (0-500 μM BAEE), and the absence of significant positive signals in samples where PAD had not been added. Next, three different substrates were compared (Figure 4.15), demonstrating highest activity towards the double substituted arginine derivative BAEE and lowest towards free L-arginine, as is to be expected for mammalian PADs (Sugawara *et al.* 1982).

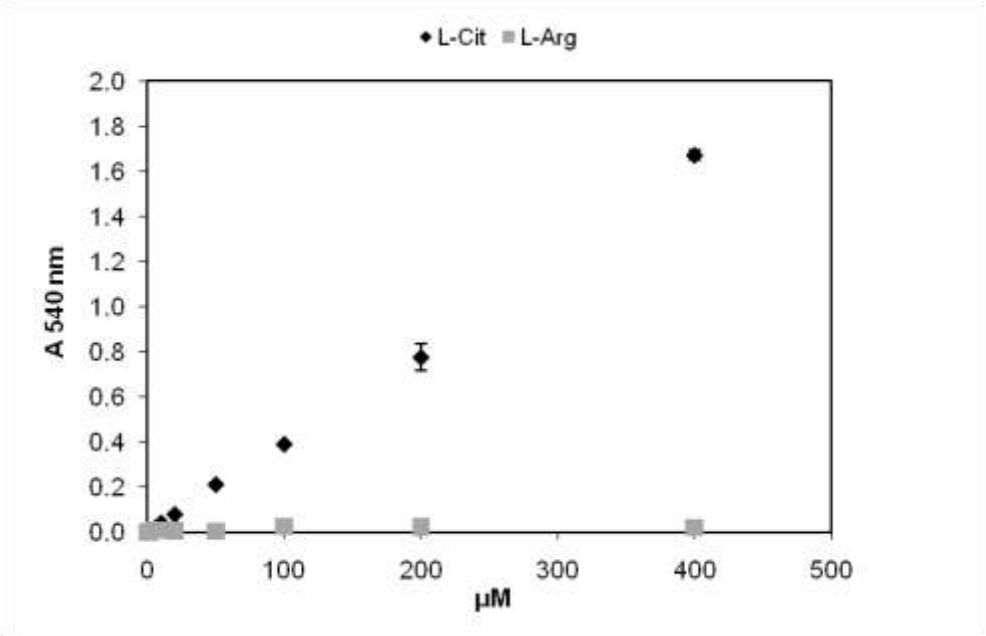


Figure 4.13 Evaluation of the colorimetric method for detection of citrulline. Various concentrations of either L-citrulline or L-arginine (0-400 µM) were mixed with PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5). Citrulline detection reagent (see Materials and Methods) was added and the mixture heated for 15 min at 95˚C. Absorbance was read at 540 nm and blank values (buffer alone) substracted from each value for analysis. All samples were measured in triplicates. Error bars indicate standard deviation of triplicate measurements.

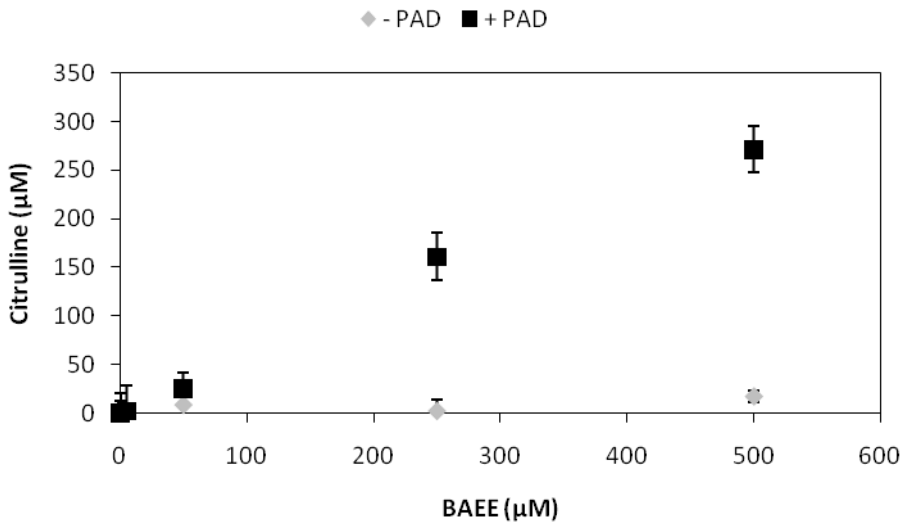


Figure 4.14 Quantification of citrullination of benzoyl-arginine ethylester (BAEE) by rabbit skeletal PAD. Various concentrations of BAEE were incubated in the presence ('+PAD') or absence ('-PAD') of 0.07 U rabbit skeletal PAD (Sigma) for 1 hour at 52˚C in rabbit PAD activity buffer (100 mM Tris-HCl, 10 mM $CaCl_2$, 5 mM DTT, pH 7.4). After photometric readout, blank values (buffer alone) were substracted from each value for analysis. All measurements were performed in triplicates. Error bars indicate standard deviation of triplicate measurements.

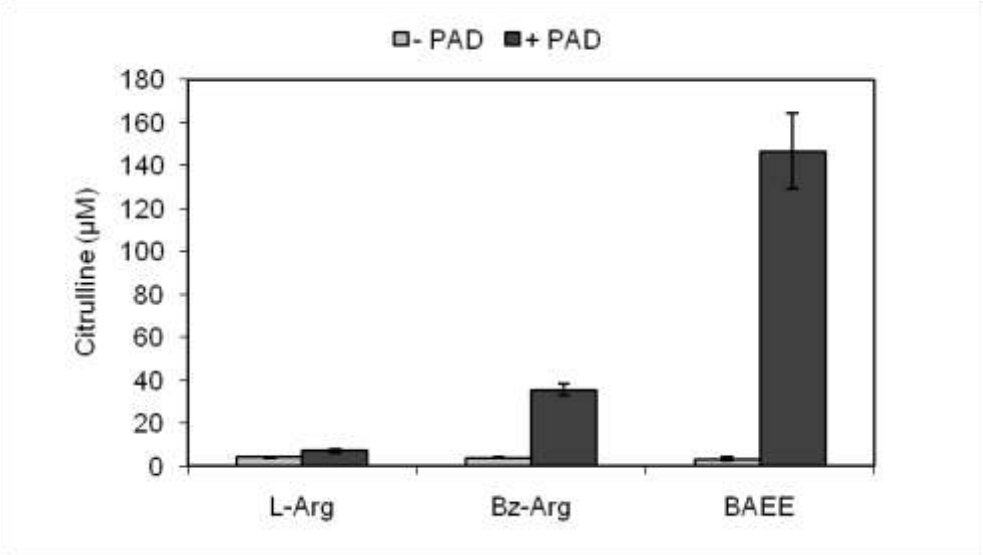


Figure 4.15 Comparison of substrates for rabbit skeletal PAD. Three substrates (5 mM each) were incubated in the presence ('+PAD') or absence ('-PAD') of 0.07 U rabbit skeletal PAD in PAD activity buffer (100 mM Tris-HCl, 10 mM $CaCl_2$, 5 mM DTT, pH 7.4) for 30 min at 52˚C. After photometric readout, blank values (buffer alone) were substracted from each value for analysis. All measurements were performed in triplicates. Error bars indicate standard deviation of triplicate measurements. L-Arg: L-arginine; Bz-Arg: Benzoyl-L-arginine; BAEE: Benzoyl-l-arginine ethylester.

The combined results were regarded as promising and prompted the investigation of PPAD activity using an analogous protocol. In an initial experiment to assess whether the full-length enzyme (preparation from Figure 4.11 B 'FT1') was enzymatically active or not, a large amount of the PPAD preparation (10 µg) was incubated in the presence or absence of substrate (10 mM BAEE) in CHES activity buffer (pH 9.5) supplemented with DTT (10 mM), for 1 hour at 37˚C. As a control, substrate alone was incubated. Figure 4.16 shows that full-length PPAD was active, and that either PPAD or substrate incubated on their own gave only low background readings.

To verify these results, a time course (progress curve) was performed and boiled PPAD was additionally included in the setup. Figure 4.17 shows a typical enzymatic progress curve for PPAD+substrate, and a flat line for boiled PPAD+substrate and substrate or PPAD alone, confirming that the observed citrulline readout stems from enzymatic activity of PPAD. It further confirms that BAEE contributes to a background reading, and a 'substrate alone/-PPAD' control was therefore included in all subsequent experiments and corrected for by substracting the value from all readings containing substrate.

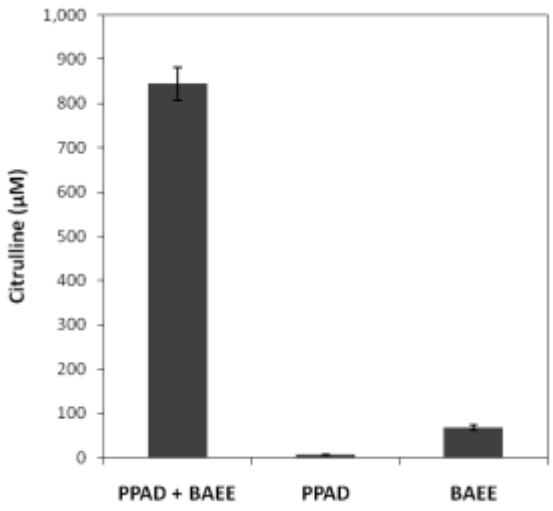


Figure 4.16 Recombinant PPAD is enzymatically active. 10 µg PPAD preparation was incubated in the presence ('PPAD+BAEE') or absence ('PPAD') of substrate (10 mM BAEE) in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) for 1 hour at 37˚C. Substrate alone was used as control. After photometric readout, blank values (buffer alone) were substracted from each value for analysis. All measurements were performed in triplicates. Error bars indicate standard deviation of triplicate measurements. BAEE: Benzoyl-l-arginine ethylester.

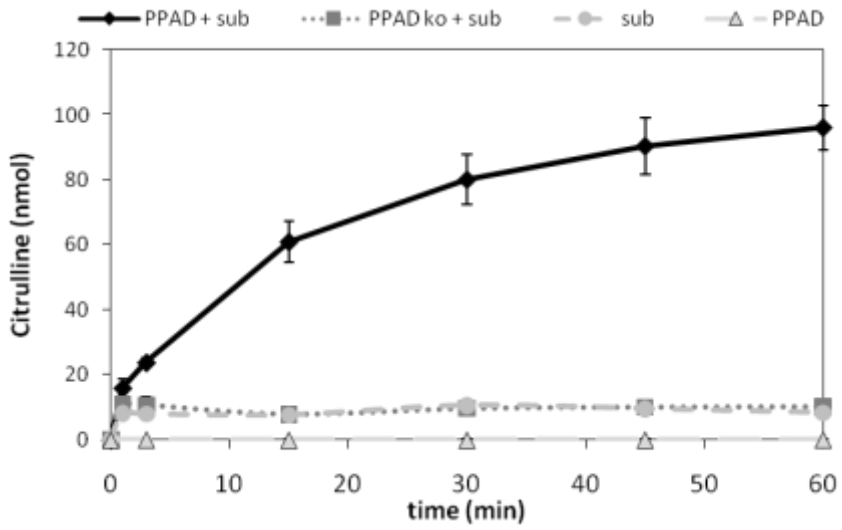


Figure 4.17 Time course experiment for recombinant PPAD activity. 10 µg PPAD preparation was incubated at 37˚C in the presence ('PPAD+sub') or absence ('PPAD') of substrate (10 mM BAEE) in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5). Substrate alone ('sub) and boiled PPAD in the presence of substrate ('PPAD ko + sub') were used as control. Samples were removed before addition of substrate (0 min) and after 1, 3, 15, 30, 45, and 60 min. After photometric readout, blank values (buffer alone) were substracted from each value for analysis. All measurements were performed in dublicates. The average of two independent experiments is shown. Error bars represent the standard error of the mean (SEM). Citrulline is expressed as nmol in a total reaction volume of 60 µl. BAEE: Benzoyl-l-arginine ethylester.

Next, analogous to Figure 4.15, L-arginine, benzoyl-arginine and BAEE were compared as substrates for PPAD. The results qualitatively reflected the data from native PPAD published previously (McGraw *et al.* 1999), in that L-arginine is citrullinated by PPAD, but to a lesser extent compared to benzoyl-arginine or BAEE (V_{max} values from McGraw *et al.*: L-arginine 26 nmol/min/mg, Bz-arginine 104 nmol/min/mg, BAEE 90 nmol/min/mg). The combined results establish that full-length PPAD is enzymatically active; hence cleavage of the 43 N-terminal amino acids, as suggested in the publication by McGraw *et al.*, is not essential for "activation" of PPAD.

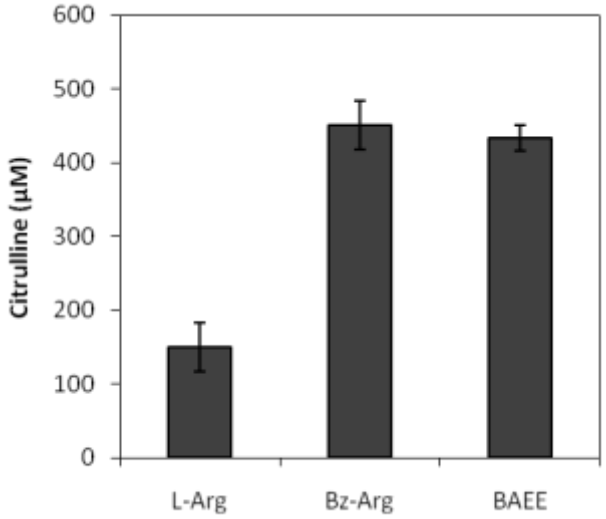


Figure 4.18 Comparison of substrates for PPAD. L-arginine, benzoyl-L-arginine, and BAEE (benzoyl-l-arginine ethylester) (10 mM each) were incubated with PPAD for 1 hour at 37˚C in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5). After photometric readout, blank values (substrate alone) were substracted from each value for analysis. The average of three independent experiments is shown. Error bars represent the standard error of the mean (SEM).

As described above, cleavage of the Trx-His tag in recombinant PPAD was only satisfactorily efficient in the presence of reducing agents and detergent. This was not only costly and time-consuimg, but further posed an obstacle for some applications of recombinant PPAD, for example when used as an immunising antigen in animals and also in downstream application such as protein determination. Therefore, it was evaluated whether the tagged form would give similar results in terms of activity and could therefore be used in all subsequent experiments instead. Equimolar amounts of both enzymes were compared for activity with BAEE as substrate and their activity was almost identical (Figure 4.19), indicating that the presence of the tag does not interfere with activity and therefore most likely the three-dimensional structure, which is important for antibody development.

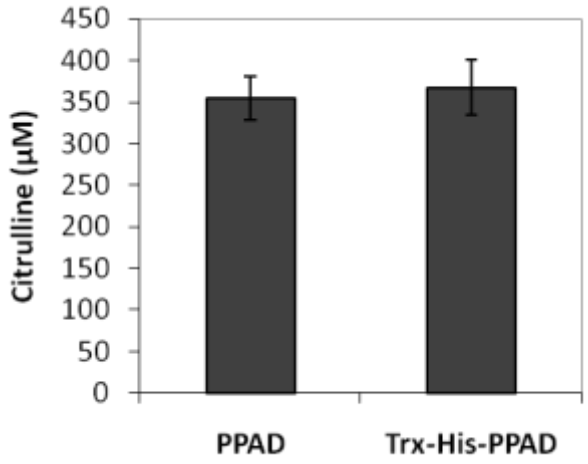


Figure 4.19 Comparison of the activity of cleaved and tagged PPAD. 1 µM of tag-free ('PPAD') and tagged PPAD ('Trx-His-PPAD') were incubated with substrate (BAEE, 10 mM) in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) for 45 min at 37˚C and formed citrullinated product was quantified. The average of two independent experiments is shown. Error bars represent the standard error of the mean (SEM).

However, ~30% of the Trx-His-PPAD preparation contained a ~15 kDa Trx-His truncation product (by weight; estimated from SDS-PAGE). In order to obtain a reference value for enzymatic activity of 100% Trx-His-PPAD, the enzyme preparation was further purified by size exclusion chromatography (gel filtration), by which molecules are sorted based on their hydrodynamic radius (simplified: shape and molecular mass). Molecules with the biggest hydrodynamic radius exit the column first. Figure 4.20 A+C shows that pure Trx-His-PPAD could be successfully separated from the truncation. The enzyme activity of the Trx-His-PPAD fractions mirrored their respective protein concentration (mAU reading) (Figure 4.20 B). Fractions containing pure Trx-His-PPAD were pooled, concentrated and used for subsequent experiments.

The influence of reducing agents on enzyme activity was first assessed. 0-10 mM of DTT or L-cysteine were added to the assay mix and PPAD activity was quantified. The data in Figure 4.21 demonstrate that some activity can be observed in the absence of reducing agent. However, activity increases significantly with the concentration of reducing agent, and this is comparable between DTT and L-cysteine, supporting the view that a reduced cysteine side chain is present in the active catalytic centre.

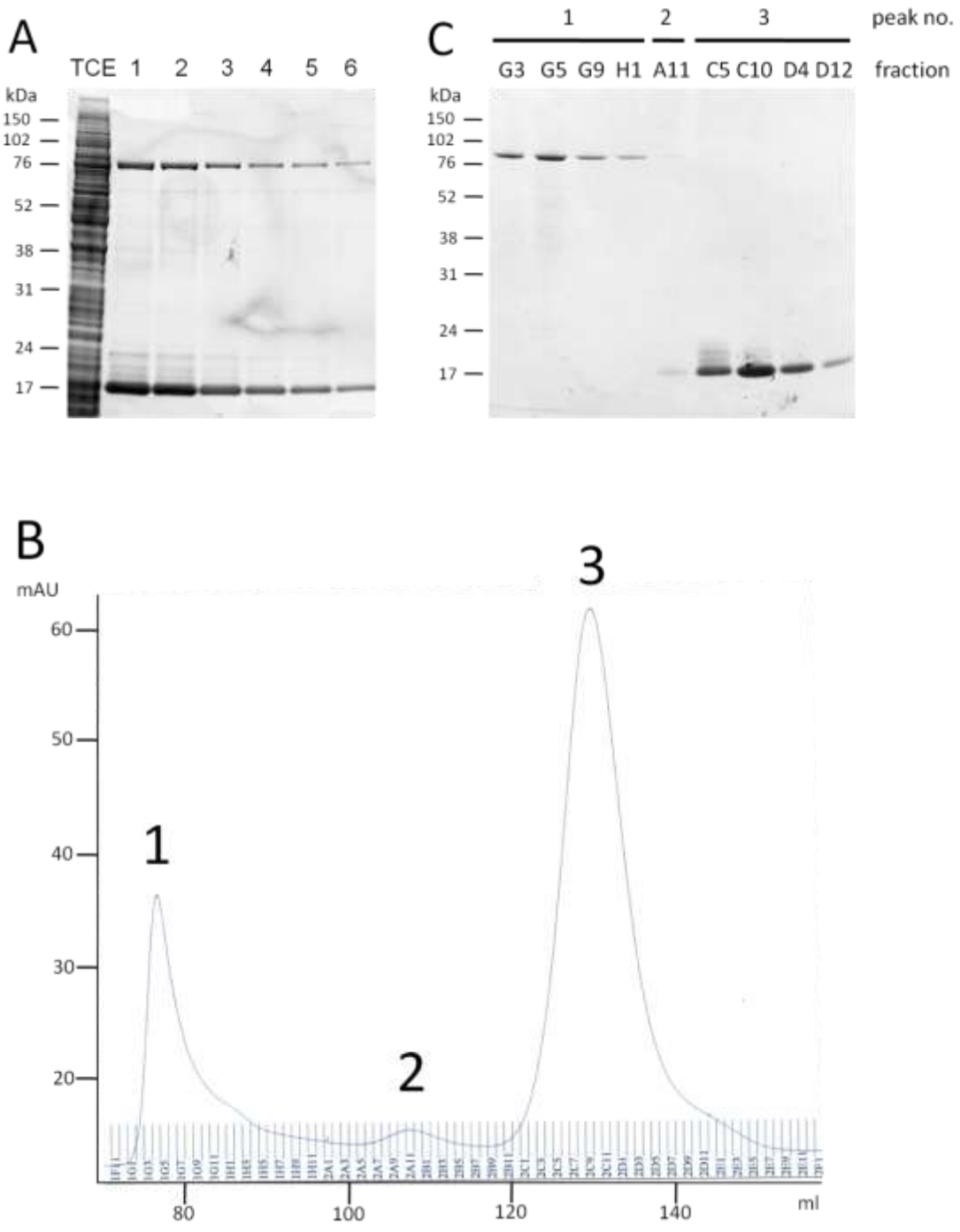


Figure 4.20 Purification of Trx-His-PPAD by size exclusion chromatography. (A) Representative SDS-PAGE gel showing purification of Trx-His-PPAD by His-tag affinity chromatography. TCE: Total cell extract of E. coli BL21 (DE3) overexpressing Trx-His-PPAD. 1-6: Eluate fractions of affinity-purified Trx-His-PPAD. The average yield of recombinant Trx-His-PPAD was 2 mg per Liter of culture. (B) Size exclusion chromatogram after fractionation of the pooled Trx-His-PPAD eluates. Peaks are labelled 1-3. Collected fractions (1 ml each) are indicated at the bottom. (C) SDS-PAGE gel of a selection of the eluted fractions from (B). mAU: milli absorbance units.

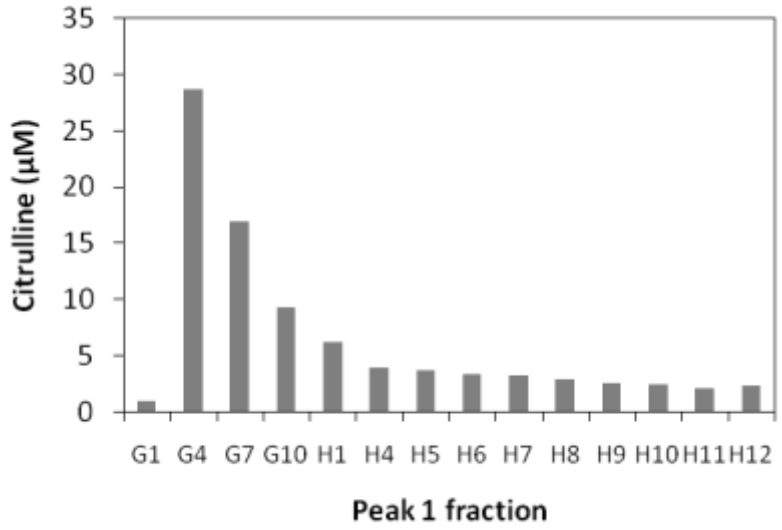


Figure 4.21 Enzymatic activities of Trx-His-PPAD fractions after size exclusion chromatography. 30 µl of several 1-ml fraction corresponding to peak 1 in Figure 3.34 B were tested for citrullinating activity with 10 mM BAEE as substrate. Samples were incubated for 30 min at 37˚C in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) and citrullinated product was quantified. Single measurements were performed due to available sample volume restrictions.

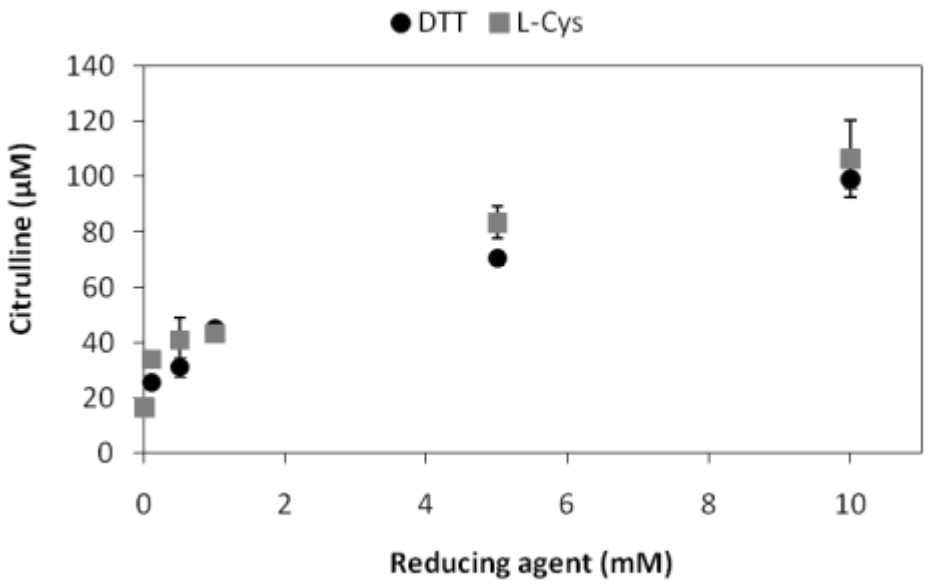


Figure 4.22 Influence of reducing agents on the enzymatic activity of PPAD. 1 µM Trx-His-PPAD was incubated in the presence of varying amounts of reducing agent (DTT or L-Cysteine) with 10 mM BAEE in PPAD activity buffer (50 mM CHES, pH 9.5) as substrate for 15 min at 37˚C and citrullinated product was quantified. Measurements were performed in triplicates. Error bars indicate standard deviation of triplicate measurements. DTT: Dithiotheitol.

In order to determine a reference value for PPAD activity in a pure preparation, different molar amounts of the gel-purified Trx-His-PPAD were incubated with substrate and the specific activity was calculated as a mean of the values obtained for each setup. Based on the data presented in Figure 4.23, the specific activity of pure Trx-His-PPAD was calculated as 23.8 U/mg, under the specified conditions (pH 9.5, 37˚C, 10 mM BAEE).

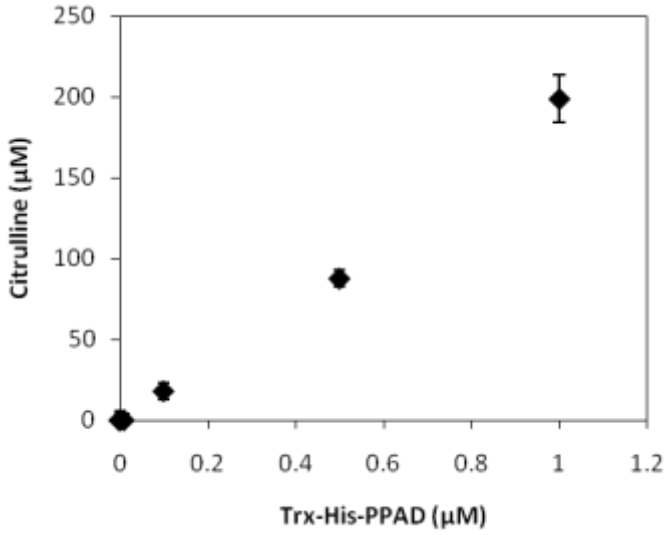


Figure 4.23 Determination of the specific activity of pure PPAD. Various concentrations of gel-purified Trx-His-PPAD (0, 0.1, 0.5, 1 µM) were incubated with substrate (10 mM BAEE) in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) for 10 min at 37˚C and the formed citrullinated product quantified. Measurements were performed in dublicate. Error bars indicate standard deviation of dublicate measurements.

The background to the calculations is explained in detail in 'Materials and Methods' (p.118/119). For the purpose of clarity, the calculation for 1 µM PPAD is reproduced here:

First, the enzymatic activity *v* in Units (U = µmol/min) is calculated:

$$v = \frac{(A - A_0) * V}{B * T}$$

A is the absorption of enzymatic reaction (PPAD+substrate+buffer), in absorption units:

→ 0.695

A_0 is the absorption of blank (substrate+buffer), in absorption units:

→ 0

V is the incubation volume, in L:

→ 0.006

B is the slope of the calibration curve, in 1/µM:

→ 0.0035

T is the time of enzymatic reaction, in min:

→ 10

$$v = \frac{(0.695 - 0) * 0.006}{0.0035 * 10}\ \mu mol/\min = 0.12\ U$$

The specific activity is the enzyme activity in U per milligram of total protein.

$$\frac{0.12\ U}{4.66 \cdot 10^{-3}\ mg} = 25.7\ U/mg$$

This procedure was repeated for the other enzyme concentrations and the average was calculated as 23.8 U/mg. This allowed the calculation of % enzyme purity in subsequent preparations that had not been gel-purified and thereby proper adjustment for active enzyme concentration.

The pure enzyme preparation was then used for the study of Michaelis-Menten kinetics with the substrate BAEE. Briefly, Michaelis-Menten kinetics describes how the initial reaction rate of an enzymatic reaction (time period in which product formation is linear with time) depends on substrate concentration, applying a number of assumptions (eg. the reverse reaction is negligible during the measurement period). Using non-linear regression algorithms to fit the Michaelis-Menten equation, the constants K_M and V_{max} can be accurately calculated.

Michaelis-Menten equation: $v_0 = \frac{v_{max} \cdot [S]}{K_M + [S]}$

[S] stands for substrate concentration, v_0 for the initial reaction rate, and K_M is the substrate concentration concentration at which the rate of the enzyme reaction is half the maximal reaction rate V_{max}.

Previous time course experiments showed that even with relatively high enzyme concentration (2.7 µM) the reaction rate is linear at least for 15 min (for a representative graph see Figure 4.17). With an enzyme concentration of 0.5 µM, the linear period lasted for 40 min (not shown). Therefore, an incubation time of 15 min (and 0.5 µM enzyme) was chosen as this should guarantee to be in the linear region. Various concentrations of substrate (BAEE) were incubated with 0.5 µM Trx-His-PPAD and the enzyme activity quantified and plotted against substrate concentration. Using non-linear regression (Graphpad Prism), a graph was fitted to the data points (r^2 goodness of fit: 0.9817), showing decreasing activity after a maximum at 1 mM substrate, as shown in Figure 4.24. This is called 'substrate inhibition'. Based on this data and the Michaelis-Menten equation, the Michaelis-Menten constants were calculated as V_{max} 0.78 ± 0.04 (µmol/min)x10^{-3} and K_M 0.10 ± 0.02 mM under the applied experimental conditions. These values are useful when comparing different substrates, as they are a measure for the affinity of the enzyme to a substrate (the lower K_M the higher the affinity) and the maximum activity that can be achieved under conditions of substrate saturation (the higher V_{max} the higher the activity at substrate saturation).

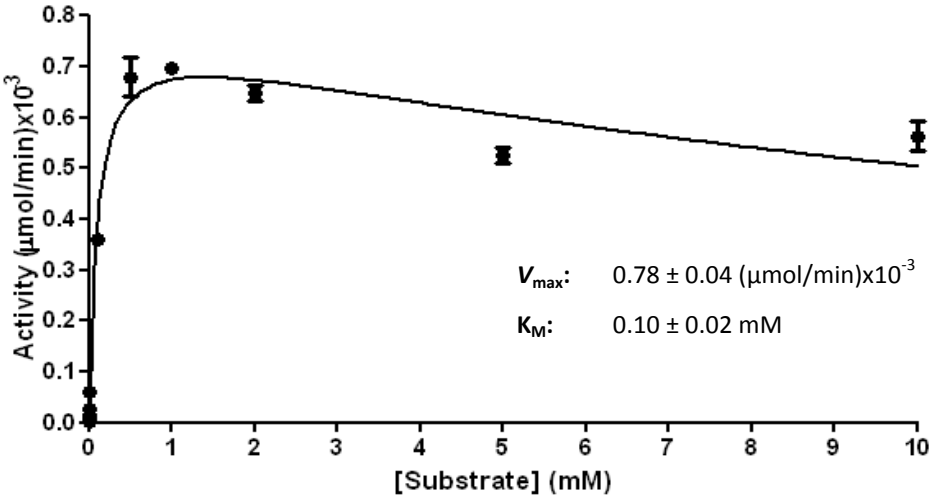


Figure 4.24 Determination of Michaelis-Menten kinetic constants for PPAD and the substrate BAEE. 0.5 µM Trx-His-PPAD was incubated with 0-10 mM BAEE for 15 min at 37˚C in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) and enzymatic activity quantified. The average of two independent experiments is shown. Error bars represent the standard error of the mean (SEM). The calculated vaues for V_{max} and K_m are displayed.

The analogous kinetic data from native PPAD from *P. gingivalis* culture supernatants and BAEE as substrate (McGraw *et al.* 1999) were K_M=874 µM (versus 100 µM here) and V_{max}=90 nmol/min/mg (versus 339 nmol/min/mg here). The results are different but lie in the same range. Deviations are probably attributable to the different experimental conditions: McGraw *et al.* used native, truncated PPAD (versus recombinant full-length PPAD), pH 8.0 (versus pH 9.5), and a completely different assay buffer composition (200 mM Tris, 1 mM EDTA, 25 µM FMN, 25 µM FAD, 100 µM NADPH, 10 mM cysteine, versus 50 mM CHES, 10 mM DTT).

Citrullination of fibrinogen and enolase peptides by recombinant PPAD

A preference of PPAD towards C-terminal arginine residues was previously indirectly indicated based on the lack of detected peptides containing internal citrulline residues after incubation with *P. gingivalis*. An *in vitro* assay using recombinant Trx-His-PPAD and synthetic peptides containing arginine residues in three different positions (C-terminus, near C-terminus, internal) was performed to directly investigate the issue. The peptide substrate containing a C-terminal arginine ('Fib-R') corresponded in sequence to a peptide detected by MS/MS after incubation of fibrinogen with *P. gingivalis* (Table 3.2). Peptide 'Fib-R-XX' corresponds to the same sequence but with two amino acids added after the arginine. Peptide 'Eno-REP1' corresponds to the CEP-1 peptide with arginine instead of citrulline.

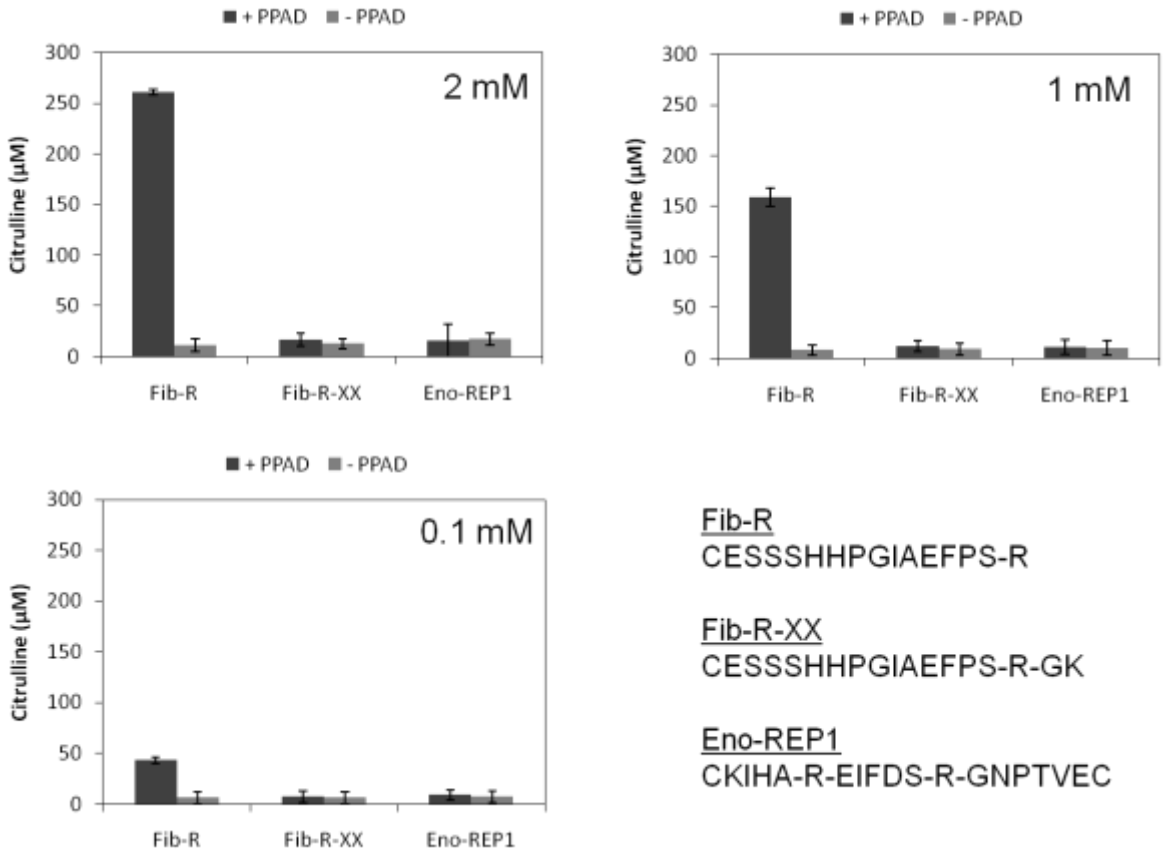


Figure 4.25 Citrullination of fibrinogen and enolase peptides by PPAD. 1 µM Trx-His-PPAD was incubated with 2, 1, or 0.1 mM peptide substrate ('+PPAD') in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5). Samples without enzyme were used as controls ('-PPAD'). The sequences of the peptides are shown. 10 mM BAEE was used as positive control substrate. The reactions were incubated for 15 min at 37˚C and enzymatic activity quantified. The average of two independent experiments is shown. Error bars represent the standard error of the mean (SEM). The activity for 10 mM BAEE was 125 µM (+PPAD) and 22 µM (-PPAD).

Figure 4.25 shows that indeed only the peptide containing a C-terminal arginine residue is citrullinated. The addition of just two amino acids to the C-terminus of this peptide reduced the citrulline signal to nonsignificant background levels, as was the case with the REP-1 peptide. A further experiment supported this finding: equimolar amounts, referring to the number of arginine residues, of human α-enolase and fibrinogen were either directly incubated with PPAD or first digested with trypsin and then incubated with PPAD. Trypsin, like gingipain, cleaves after arginine and lysine residues. Figure 4.26 shows that whole enolase or fibrinogen was not citrullinated by PPAD; in contrast pre-incubation of the same amount of protein with trypsin resulted in substantial levels of citrullination for both substrates. This data unequivocally supports previous indications for strong substrate specificity of PPAD towards C-terminal arginine over internal arginine.

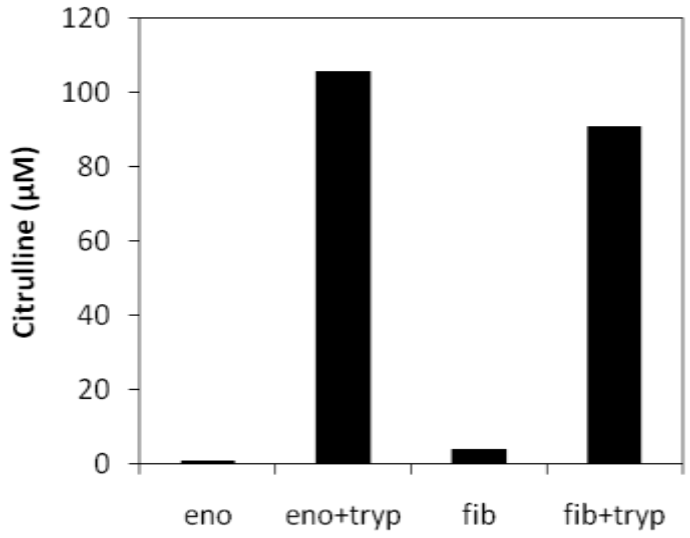


Figure 4.26 Citrullination of whole and trypsin-digested enolase and fibrinogen by PPAD. Equimolar amounts (26.4 nmol, referring to the number of arginine residues) of human α-enolase or fibrinogen were incubated in the presence ('+tryp') or absence of trypsin. Samples were freeze-dried, resuspended in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) with 2.7 µM Trx-His-PPAD, and incubated for 30 min at 37˚C and enzymatic activity quantified. Single measurements were performed due to sample volume restriction.

Citrullination of hemoglobin by PPAD

As presented in the previous section, the *P. gingivalis* mutant strains Δ*rgp* and Δ*rgp+kgp* contained residual citrullinated proteins, despite the absence of arginine-gingipains that could generate C-terminal arginines as a substrate for PPAD. It was discussed that this might be due

to proteins naturally containing a carboxy-terminal arginine residue. To test this hypothesis, a protein with a naturally occurring C-terminal arginine, hemoglobin, was chosen to be studied as a substrate for PPAD. Hemoglobin plays an important physiological role for *P. gingivalis* as a source of iron for nutrition and of iron protoporphyrin IX (FePPIX) dimer as oxidative protective shield, giving the bacterium its distinctive black pigmentation (see Figure 1.11) as it accumulates on the cell surface. Gingipains and the outer membrane proteins HmuR and HmuY are critical for binding and degradation of hemoglobin and acquisition of the heme group (Smalley *et al.* 1998; Olczak *et al.* 2006).

Hemoglobin A is the most common type in adult humans and comprises two α chains (14 kDa) and two β chains (17 kDa), each binding a heme group. Hemoglobin can be saturated with oxygen or desaturated with oxygen, and is accordingly called oxy-hemoglobin and deoxy-hemoglobin, respectively. Met-hemoglobin is another form, in which the iron of the heme group is in the Fe^{3+} state and cannot bind oxygen. It is known that *P. gingivalis* uses oxy-hemoglobin, which is converted into met-hemoglobin by arginine gingipains (possibly indirectly/non-proteolytically) and then degraded by lysine-gingipains to form the FePPIX pigment (Smalley *et al.* 2004). The hemoglobin α chain contains a C-terminal arginine residue, the citrullination of which by PPAD might affect the susceptibility of oxy-hemoglobin to be converted into met-hemoglobin and therefore its degradation by lysine-gingipains.

For this experiment, horse oxy-hemoglobin was kindly provided by John Smalley, University of Liverpool. Horse and human hemoglobin α chains are 89% identical, including the C-terminal arginine residue (Figure 4.27 A). An analysis of the three-dimensional structure of the human oxy-hemoglobin A tetramer shows that the two C-terminal arginine residues are facing inwards and a computational analysis of the accessible surface indicated that only 20% of the surface of these two residues is accessible (calculated with DeepView) (Figure 4.27 B).

Citrullination of hemoglobin by PPAD was first assessed using a standard PPAD activity assay with varying concentration of hemoglobin. DTT was excluded from the assay mix in order to avoid any direct effects caused by its reducing activity on hemoglobin. Further, two different buffers with pH 7.5 or 9.5 were tested because the pH, along with the tetrameric hemoglobin concentration, affects the dimer/tetramer ratio present in solution. The data presented in Figure 4.28 A show that hemoglobin is citrullinated under these conditions, with

2-3 fold higher levels of citrullination observed at pH 7.5 compared to 9.5, despite an expected ~2 fold higher level of PPAD activity at pH 9.5 compared to 7.5 (McGraw *et al.* 1999).

A

```
sp|P69905|HBA_HUMAN     MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHG 60
sp|P01958|HBA_HORSE     MVLSAADKTNVKAAWSKVGGHAGEYGAEALERMFLGFPTTKTYFPHFDLSHGSAQVKAHG 60
                        ****.**********.***.***************.**********************.**

sp|P69905|HBA_HUMAN     KKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTP 120
sp|P01958|HBA_HORSE     KKVGDALTLAVGHLDDLPGALSNLSDLHAHKLRVDPVNFKLLSHCLLSTLAVHLPNDFTP 120
                        ***.**** **.*:**:*.*** ************************ ***.*** :***

sp|P69905|HBA_HUMAN     AVHASLDKFLASVSTVLTSKYR 142
sp|P01958|HBA_HORSE     AVHASLDKFLSSVSTVLTSKYR 142
                        **********:**********
```

B

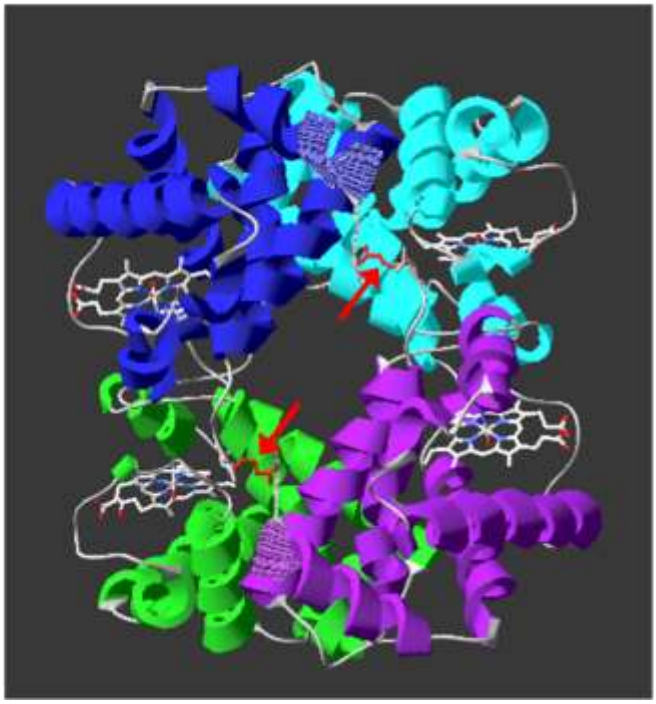


Figure 4.27 **Haemoglobin. (A)** Amino acid alignment of the α-chains of human and horse haemoglobin. The C-terminal arginine residue is highlighted in red. **(B)** Three-dimensional X-ray model of human oxy-hemoglobin (1GZX)[19]. The C-terminal arginine residue of the two α chains is highlighted in red and with red arrows. Dark-blue and purple: α chains; turquoise and green: β chains; grey: loops. This figure was generated by myself using the software DeepView v4.0 (http://spdbv.vital-it.ch/).

[19] www.rcsb.org/pdb/explore.do?structureId=1GZX

This could be explained by the aforementioned effect of the pH on the tetramer dissociation constant. The tetramer-dimer dissociation constant of oxy-hemoglobin (K_{diss} = $[Dimer]^2/[Tetramer] \approx 10^{-6}$ M) decreases as the pH goes up, therefore there is a higher proportion of dimer compared to tetramer at pH 7.5. Combined with a higher accessibility of the C-terminal arginine residue in the α-β dimeric form, this might explain the higher citrullination levels at pH 7.5 compared to 9.5 despite lower PPAD activity.
Controls were performed in which substrate (hemoglobin) alone was incubated without PPAD and the absorption values substracted from the substrate+PPAD value. This was especially critical in this experiment as hemoglobin absorbs light in the 540 nm range, which is the wavelength used for the readout of the citrulline assay. To confirm the results, citrullination was also assessed by anti-citrulline immunoblot, which showed a band at the size expected for hemoglobin α chain (14 kDa) (Figure 4.28 B), confirming that PPAD citrullinates hemoglobin.

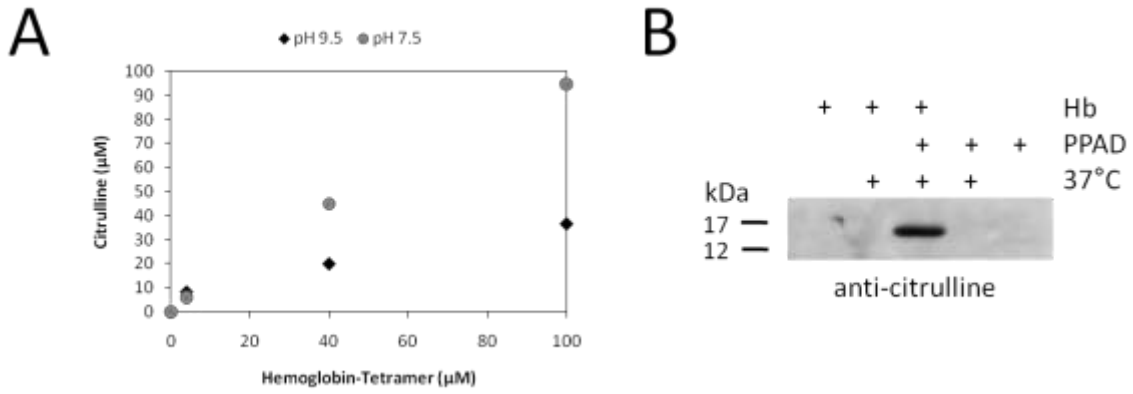


Figure 4.28 Citrullination of haemoglobin by PPAD. (A) 1 µM Trx-His-PPAD was incubated with 4, 40, or 100 µM horse oxy-haemoglobin tetramer for 60 min at 37˚C at pH 9.5 (50 mM CHES) or pH 7.5 (50 mM Tris-HCl). Enzymatic activity was quantified. Single measurements were performed due to sample volume restrictions. **(B)** 1 µM Trx-His-PPAD was incubated with 4 µM horse oxy-haemoglobin tetramer for 60 min at 37˚C at pH 7.5 (50 mM Tris-HCl). As controls, either haemoglobin or PPAD were left out. PPAD and haemoglobin without incubation served as pre-reaction controls. The samples were resolved by SDS-PAGE and citrullinated protein detected by anti-modified citrulline immunoblot. The constituents of each reaction mix are indicated by '+': Hb = horse oxy-haemoglobin; PPAD = Trx-His-PPAD; 37˚C = reaction was incubated at this temperature. The positions of the relevant molecular mass markers are indicated on the left. Hemoglobin protein was kindly provided by John Smalley, University of Liverpool.

Inhibition of PPAD

Three groups of compounds were tested for their potential to inhibit PPAD activity: (1) tetracyclines and derivatives; (2) sulfhydryl-reactive compounds; (3) disease-modifying anti-rheumatic drugs (Figure 4.29).

Group 1 comprised tetracycline, doxycycline, and minocycline and was included based on the reported efficacy of tetracycline in human PAD inhibition (see p. 43) and the reported inhibitory effects on gingipains, where doxycycline was found to most effective compared to tetracycline and minocycline with an IC_{50} of 3 and 20 μM for arginine-gingipains and lysine-gingipain, respectively (Imamura *et al.* 2001). The authors suggested that inhibition is due to interference with the oxidative state of the proteinases by doxycycline as the inhibitory activity was reduced to IC_{50} ~100 μM for arginine-gingipain in the presence of reducing agent (10 mM L-Cysteine).

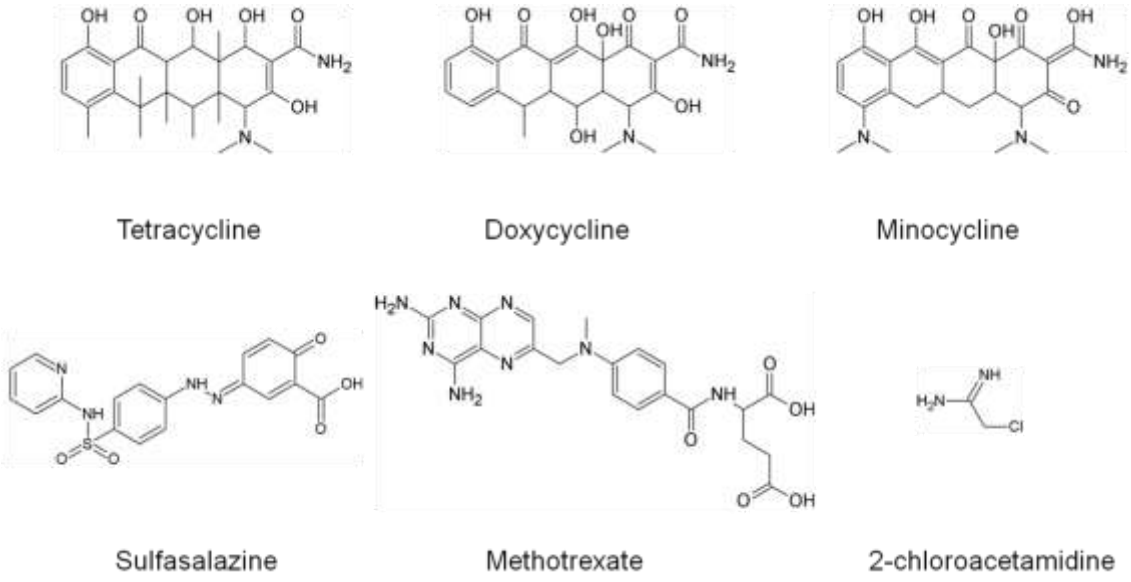


Figure 4.29 Chemical structures of compounds tested for inhibition of PPAD.

Group 2 was represented by 2-chloroacetamidine, a modest inhibitor of human PAD4 and bacterial DDAH (see p. 40). As this compound was suggested to block the catalytic cysteine residue in both PAD4 and DDAH, and PPAD contains a predicted catalytic cysteine, it was

chosen as a potential inhibitor. Other PAD4 inhibitors acting via the same mechanism such as Cl-amidine and F-amidine were not tested as they are not commercially available.
Group 3 contained sulfasalazine and methotrexate, two common DMARDs, whose detailed mode of action in RA is unclear. They were included to test whether they might inihibit PPAD and possibly thereby be effective in RA. Metabolites of these drugs might also be important, in particular in the case of sulfasalazine which is metabolised into 5-aminosalicylic acid and the antibiotic sulfapyridine. However, these were not tested.

Varying concentrations of the compounds were pre-incubated with PPAD and then substrate (BAEE) was added and the formation of citrulline quantified. The data in Figure 4.30 show that no inhibition occurred with tetracycline/tetracycline-derivatives DMARDs, but efficient inhibition occurred with 2-choloracetamidine.

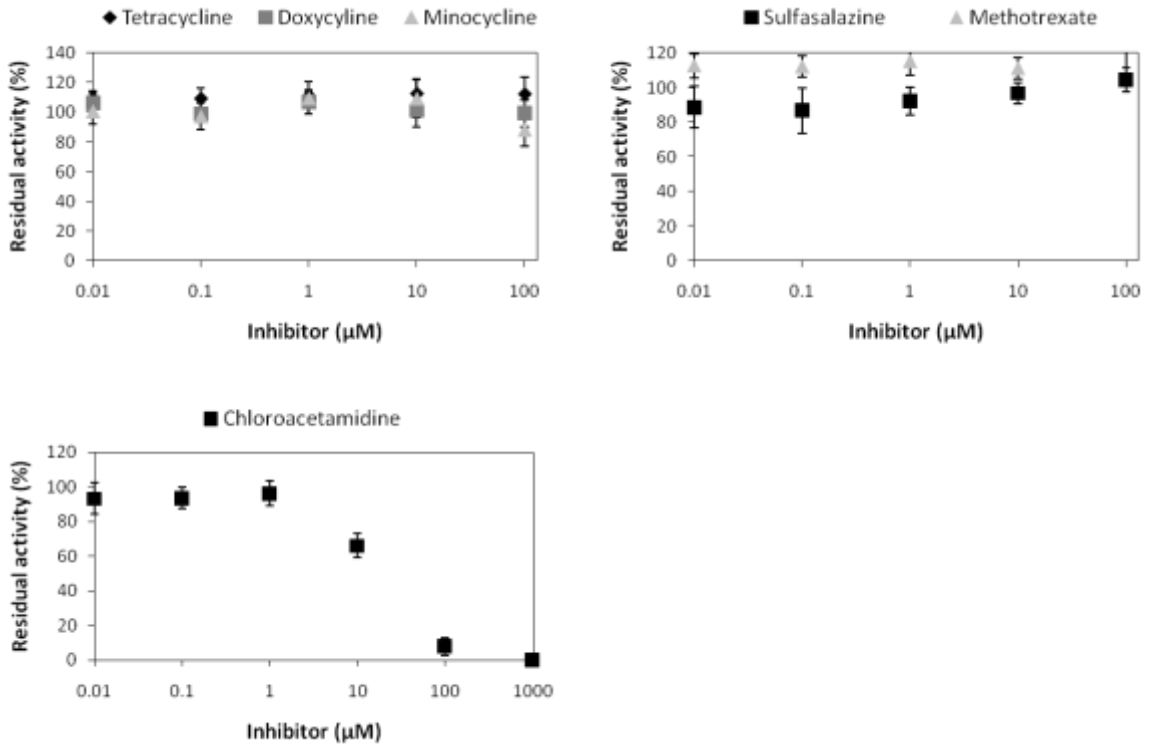


Figure 4.30 Relative residual PPAD activity in the presence of various compounds. 0.5 µM Trx-His-PPAD was preincubated with varios concentrations of inhibitor for 10 min at 37˚C in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5). Substrate (1mM BAEE) was added and the reaction incubated for a further 15 min at 37˚C. Citrulline was quantified according to the standard protocol. A control reaction to which no inhibitor was added was used as a reference value to calculate 100% PPAD activity. The activity in the presence of the described compounds was then expressed as a fraction relative to 100% activity. Error bars indicate standard deviation of triplicate measurements.

Further analysis of the inhibition data of PPAD with 2-choloracetamidine using non-linear regression (dose response, four-parameter sigmoidal; GraphPad Prism) calculated an IC_{50} value of 21.5 ± 1.3 μM. The inhibitor dissociation constant K_i, which is an absolute value and does not depend on the specific reaction conditions, can be calculated from IC_{50}, K_M and the substrate concentration based on the equation (Cheng *et al.* 1973):

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_M}} = \frac{21.5\ \mu M}{1 + \frac{1000\ \mu M}{100\ \mu M}} = 1.95\ \mu M$$

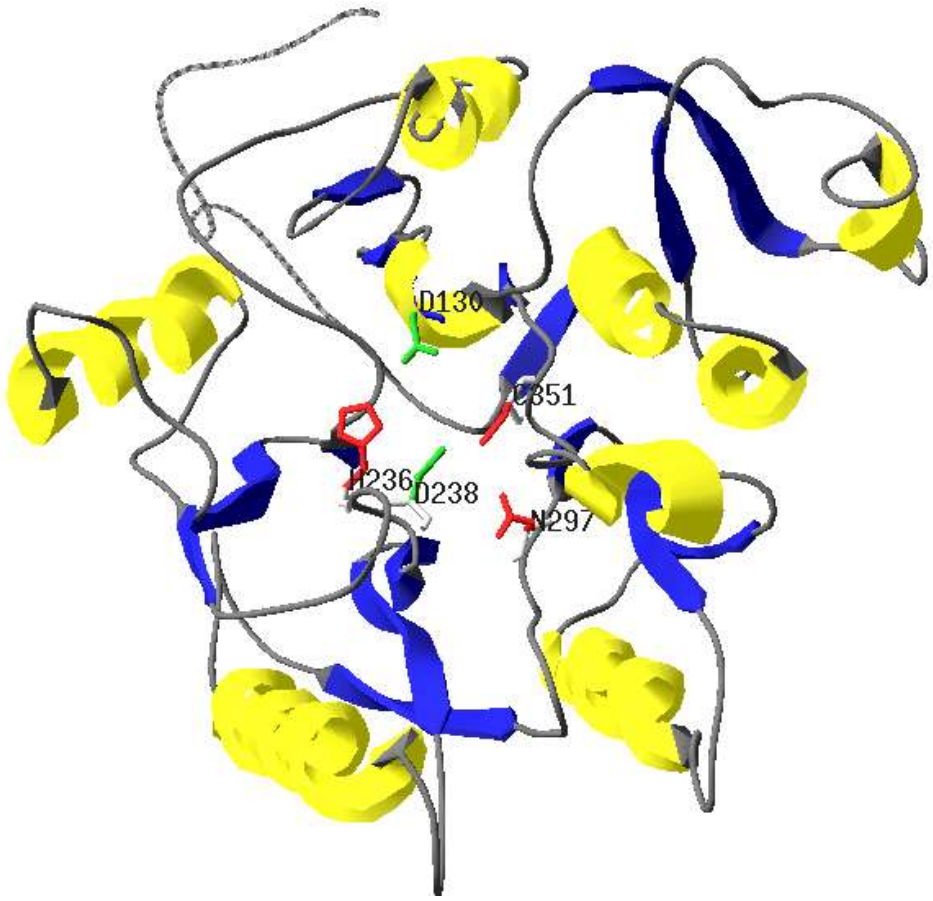
which gives a value well below that found with human PAD4 or bacterial DDAH (K_i = 20 mM and 3.1 mM respectively). K_i equals the concentration of competing substrate in a competition assay which would occupy 50% of the enzyme if no substrate was present. Here it was assumed that competitive inhibiton (inhibitor and substrate compete for binding to the active site) is taking place as this was shown to be the case for human PAD4 and DDAH with 2-chloroacetamidine. In the case of noncompetitive inhibition (inhibitor reacts with enzyme-substrate complex) or uncompetitive inhibition (inhibitor binds to the enzyme and enhances substrate binding but reduces the rate of product formation) the K_i would calculate as and 21.5 and 19.5 μM, respectively, which is still well below the values for PAD4 and DDAH.

Site-directed mutagenesis of proposed PPAD catalytic residues

The functional residues of PPAD are not known, however a bioinformatics analysis of all members of the guanidino-group modifying enzymes superfamily identified five residues predicted to be crucial for catalysis (H-236, N-297, C-351) or substrate binding/stabilisation of the catalytic centre (D-130 and D-238) (Figure 4.31). In order to experimentally map the crucial residues, site-directed mutagenesis was used to replace each of these five amino acids with alanine in the recombinant Trx-His-PPAD protein.

After the successful amino acid code substitution in the generated constructs was confirmed by nucleotide sequencing (see Appendix for alignments), expression in the total cell lysates

A



B

MKKLLQAKALILALGLFQLPAIAQTQMQADRTNGQFATEEMQRAFQETNPPAGPVRAIAE
YERSAAVLVRYPFGIPMELIKELAKNDKVITIVASESQKNTVITQYTQSGVNLSNCDFII
AKTDSYWTRDYTGWFAMYDTNKVGLVDFIYNRPRPNDDEFPKYEAQYLGIEMFGMKLKQT
GGNYMTDGYGSAVQSHIAYTENSSLSQAQVNQKMKDYLGITHHDVVQDPNGEYINHVDCW
GKYLAPNKILIRKVPDNHPQHQALEDMAAYFAAQTCAWGTKYEVYRALATNEQPYTNSLI
LNNRVFVPVNGPASVDNDALNVYKTAMPGYEIIGVKGASGTPWLGTDALHCRTHEVADKG
YLYIKHYPILGEQAGPDYKIEADVVSCANATISPVQCYYRINGSGSFKAADMTMESTGHY
TYSFTGLNKNDKVEYYISAADNSGRKETYPFIGEPDPFKFTCMNETNTCTVTGAAKALRA
WFNAGRSELAVSVSLNIAGTYRIKLYNTAGEEVAAMTKELVAGTSVFSMDVYSQAPGTYV
LVVEGNGIRETMKILK

Figure 4.31 Site-directed mutagenesis of predicted functional residues in PPAD. (A) Predicted three-dimensional structure of PPAD (residues 44-357) with proposed functional residues highlighted. The structure was computed based on the X-ray structure of agmatine deiminase from *Streptococcus mutans* (PDB ID Q8DW17) (sequence identity 19%, please see Appendix for an amino acid alignment) as a template. Helices are represented in yellow, strands in blue and loops in grey. This figure was generated by myself using the software DeepView v4.0. **(B)** Amino acid sequence of full-length PPAD (residues 1-556) with the residues represented in (A) underlined. Proposed functional residues are highlighted and colour-coded according to (A).

was tested in a small-scale experiment. All five mutant proteins could be expressed in *E. coli* BL21 (DE3) (Figure 4.32 A). In a larger-scale setup (1L culture), the mutant proteins were expressed along with the unmutated wild-type PPAD, taking extreme care as not to cross-contaminate the cultures, and the soluble proteins were purified with His-affinity resin. The yield was comparable between proteins (Figure 4.32 B). The proteins were then assessed for enzyme activity, using three different substrates: BAEE and two fibrinogen peptides containing a C-terminal arginine residue, both corresponding to originally identified peptides in *P. gingivalis*-fibrinogen digests (Fib-A: ESSSHHPGIAEFPSR; Fib-B: PAPPPISGGGYR). No detectable activity was observed with any of the five mutants. This was a surprising result, as it was expected that the bioinformatic prediction was not 100% correct and that at least some activity would be present with some of the mutants, as only one amino acid had been replaced in each mutant. Therefore, the experiment was repeated entirely beginning with recombinant protein expression, but the same result was found in repeated analysis (combined results in Figure 4.32 C). As all steps had been performed in parallel with the wild-type enzyme, and the wild-type enzyme showed activity in the expected range with all three substrates, but none of the mutants did with any of substrates, this was considered to be a true result, i.e. residues C-351, N-297, H-236, D-130 and D-238 are indeed crucial functional residues in the enzymatic mechanism of PPAD. Table 4.1 presents the data from Figure 4.32 C relative to the wild-type enzyme along with the proposed function of the respective mutated residue (Shirai *et al.* 2006)

Table 4.1 Relative activity of PPAD mutants. *Calculations are based on the data presented in Figure 4.32 C using Fib-A as substrate.*

Enzyme	Relative activity (%)	Proposed function of mutated residue
wild-type	100.0	-
C351A	1.0±0.8	Catalytic residue
N297A	0.8±0.4	Catalytic residue
H236A	0.9±0.5	Catalytic residue
D130A	0.9±0.3	Substrate binding, catalytic center stabilisation
D238A	1.3±0.2	Substrate binding, catalytic center stabilisation

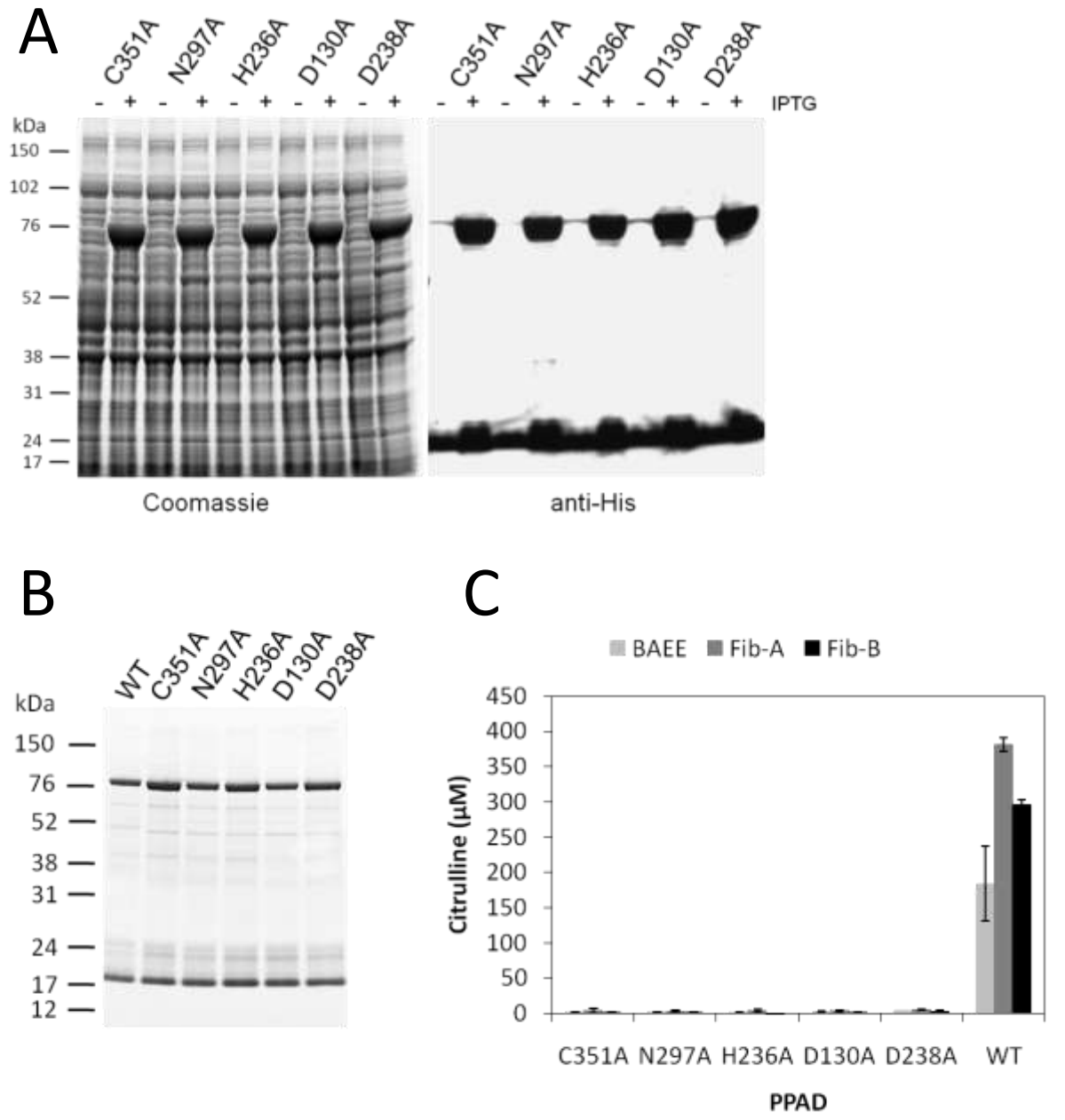


Figure 4.32 Expression, purification, and activity of PPAD mutants. (A) *E.coli* BL21 (DE3) cells were transformed with a constructs encoding mutated versions of Trx-His-full length PPAD, as indicated (expected size 77 kDa). Protein expression was initiated with the addition of IPTG ('+'). No IPTG was added as control ('-'). Total cell extracts were resolved on an SDS-PAGE gel and total protein stained with Coomassie or transferred onto nitrocellulose membranes and immunoblotted for the his-tag. **(B)** Wild-type and mutant PPAD was expressed as described in (A) and then purified using his-tag affinity resin and analysed by SDS-PAGE and Coomassie staining. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left. **(C)** 1 µM wild-type PPAD and mutants were tested for enzyme activity using 1 mM BAEE, Fib-A (ESSSHHPGIAEFPSR) or Fib-B peptide (PAPPPISGGGYR) as substrate. The reactions were incubated for 15 min at 37˚C in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) and citrullinated product was quantified spectrophotometrically. Error bars indicate the SEM from two independent experiments.

Auto-citrullination of PPAD

During expression of recombinant wild-type PPAD and mutants, it was observed that it took consistently approximately twice as long for *E. coli* cultures expressing wild-type PPAD to reach a certain OD_{600} value compared to cultures expressing mutant proteins or the empty expression vector. It was concluded that the wild-type enzyme might exert a stress on *E. coli* through its enzymatic activity by depleting energy or nutrient resources. Further, previous experiments where reactions containing recombinant PPAD and protein substrate were tested for citrullination using anti-citrulline immunoblot, showed an anti-citrulline signal in the molecular mass region of the PPAD enzyme. Taken together, this prompted the investigation of citrullination taking place in *E. coli* lysates expressing recombinant PPAD. The availability of the inactive mutant enzymes represented a negative control.

Figure 4.33 A demonstrates that *E. coli* total cell lysates expressing the wild-type PPAD enzyme show strong citrullinated protein bands, but not those expressing the mutant PPADs or the tag alone. Next, purified wild-type Trx-His-PPAD was immunoblotted with anti-citrulline antibody with and without prior incubation at 37˚C (Figure 4.33 B). The resulting pattern of citrullinated bands was similar but not identical to the one found in the whole cell lysates in Figure 4.33 A, where a few additional bands in the 15-30 kDa region were visible. Taken together, these data show that under the artificial conditions of recombinant protein overexpression, where the recombinant protein represents the majority of the total cellular protein and lacks an excess of natural substrate, citrullination of the full-length Trx-His-PPAD itself, and of one or two N-terminal ~45 kDa PPAD truncations, and possibly also *E. coli*-derived proteins, occurs. Post-expression incubation of the purified enzyme does not increase citrullination levels, indicating that citrullination takes place primarily during the described artificial environment of recombinant protein overexpression.

Of note, not in any of the colorimetric citrulline detection experiments where a control reaction containing recombinant PPAD alone was tested, a positive signal above background level was detected. This can be explained however by the lower detection level of the colorimetric assay and the much lower amount of enzyme used compared to blotting. Assuming a sensitivity of 1 ng protein for immunoblot, and the calculated sensitivity of the colorimetric assay, the immunoblot would be more than 3,000-fold more sensitive per

citrulline residue. An open question is the citrullination of full-length PPAD despite the lack of a C-terminal arginine residues in its sequence. A possible explanation is that internal arginine residues may become a substrate for PPAD under such unphysiological conditions. However, this would have to be evaluated in further experiments.

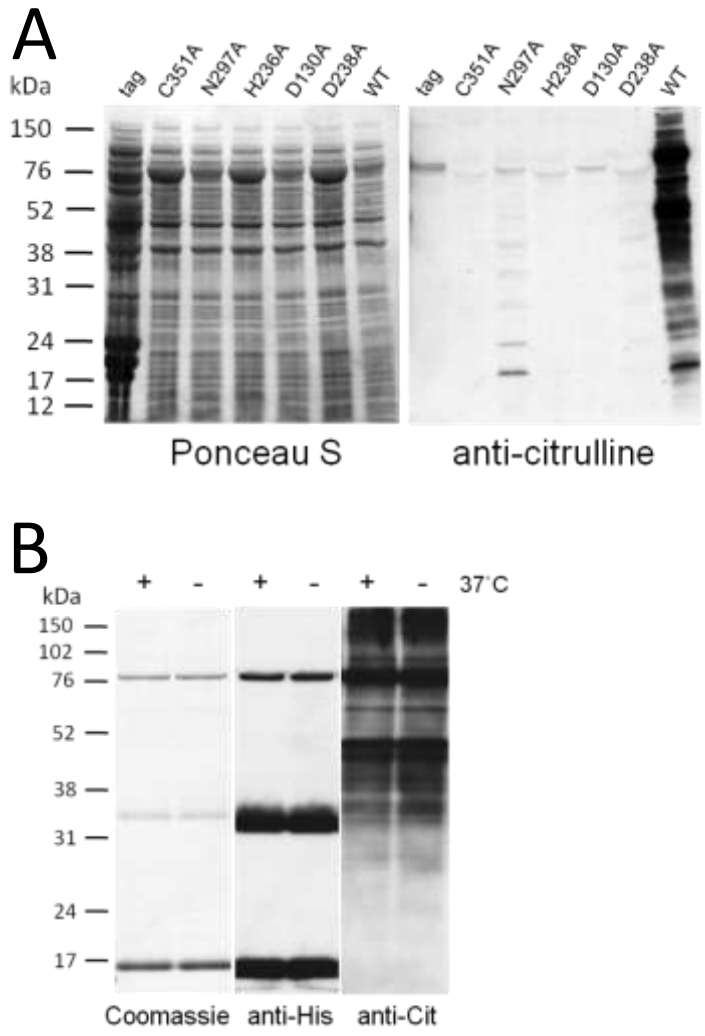


Figure 4.33 Auto-citrullination of PPAD during protein overexpression in *E. coli*. (A) Total cell lysates of *E.coli* BL21 (DE3) cells overexpressing the Trx-His-tag alone, PPAD mutants, or wild-type were resolved by SDS-PAGE, transferred onto a nitrocellulose membrane and total protein stained by Ponceau S or citrullinated proteins detected with an anti-modified citrulline antibody. **(B)** 1.5 µg purified wild-type Trx-His-PPAD was incubated in PPAD activity buffer (50 mM CHES, 10 mM DTT, pH 9.5) without additional substrate at 37˚C for 1 hour ('+') or stored as as usual after purification and not incubated ('-'). Coomassie stained total protein, anti-His immunoblot detected the His-tagged full-length and truncated PPAD and anti-modified citrulline immunoblot detected citrullinated PPAD. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Summary

Expression and purification of PPAD

- Full-length PPAD can be expressed in *E. coli* BL21 (DE3) and purified most efficiently using a detergent-based cell lysis method
- The hypothesised 'mature' form of PPAD lacking the N-terminal 43 amino acids is degraded during expression/purification
- Cleavage of the Thioredoxin-His tag is efficient only in the presence of the reducing agent dithiothreitol and the detergent Triton X-100, which may interfere with downstream applications

Activity studies, site-directed mutagenesis and self-citrullination of PPAD

- Recombinant full-length PPAD is enzymatically active
- Cofactors, particular ions, reducing agents or other additives are not required for enzymatic activity, as opposed to human PADs which are Ca^{2+}-dependant
- Reducing agents (L-Cysteine, dithiothreitol) increase enzymatic activity
- Enzymatic activity is not influenced by the presence of the Thioreoxin-His tag
- Synthetic substrates: L-arginine, benzoyl-L-arginine (N-terminally substituted L-arginine) and benzoyl-L-arginine ethylester (N- and C-terminally substituted L-arginine) are substrates for PPAD; activity is lowest with L-arginine
- Peptide substrates: Peptides with C-terminal arginines only are substrates for PPAD. Peptides with internal arginines are not citrullinated.
- Protein substrates: Proteins with C-terminal arginines only (e.g. hemoglobin A) are substrates for PPAD. Proteins with internal arginines (e.g. enolase, fibrinogen) are not citrullinated unless cleaved into peptides harbouring C-terminal arginines.
- Using benzoyl-arginine ethylester as substrate the following constants were determined:

  Specific activity: 23.8 U/mg[20] enzyme

  K_m = 0.10 mM

  V_{max} = 0.78 nmol/min

[20] see p. 118/119 for definition of 'U' and 'specific activity'

- Site-directed mutagenesis: Cysteine-351, Asparagine-297, Histidine-236, Aspartic acid-130 and -238 are each essential for PPAD activity. Substitution of any of these amino acids with alanine results in enzymatic activity of ≤1.3% of that of the wild-type enzyme
- Full-length recombinant wild-type PPAD is auto-citrullinated in *E. coli*. PPAD mutants are not citrullinated.

Inhibition of PPAD

- Tetracycline and the derivatives doxycycline and minocycline, or the DMARDS sulfasalazine and methotrexate do not inhibit PPAD activity at 10 nmol-0.1 mM. This demonstates that these agents cannot derive their therapeutic efficiency in RA from the direct inhibition of PPAD, but does not exclude a role of PPAD further upstream in the disease process.
- The sulfhydryl-reactive compound 2-chloroacetamidine inhibits PPAD with an IC_{50} = 21.5 μM. Complete inhibition accurs at 1 mM 2-chloroacetamidine. This agent could be further explored as a basis for developing specific PPAD inhibitors.

Chapter 5

Generation and evaluation of an anti-PPAD antibody and analysis of native PPAD and citrullination *in vivo*

Background

Having shown that *P. gingivalis* generates endogenous and host citrullinated proteins and peptides *in vitro*, and that recombinant PPAD is enzymatically active in its full-length form, it was of interest to test whether citrullinated proteins are also generated *in vivo* at site of periodontal inflammation and investigate the native form of PPAD. To this end, an anti-PPAD antibody was required, the production, evaluation and application of which is described in this chapter.

Results

Antigen preparation

The ideal immunising antigen does not contain protein tags that could result in the generation of tag-reactive antibodies. As the Trx-His-PPAD proved to be difficult to cleave with high efficiency, it was attempted to re-clone PPAD into a His-tag vector (pET-47b(+), see Appendix for detailed information) and use His-tag PPAD as immunising antigen. This would add only 19 additional amino acids to the N-terminus of PPAD, possibly increasing cleavage efficiency or if that proved to be unsuccessful, result in a shorter, less immunogenic tag compared to Trx-His-PPAD.

The PPAD insert was cut from the Trx-His-PPAD construct using the same restriction enzyme sites as before, and inserted into the His-vector. Successful and correct insertion was confirmed by colony PCR, analytical digest, and nucleotide sequencing (please see Appendix for alignments).

Expression of the recombinant protein was tested in BL21 (DE3) and BL21 (DE3) pLysS cells. The protein expressed successfully, but only in BL21 (DE3) cells (Figure 5.1), as was also the case for Trx-His-PPAD (see Figure 4.6).

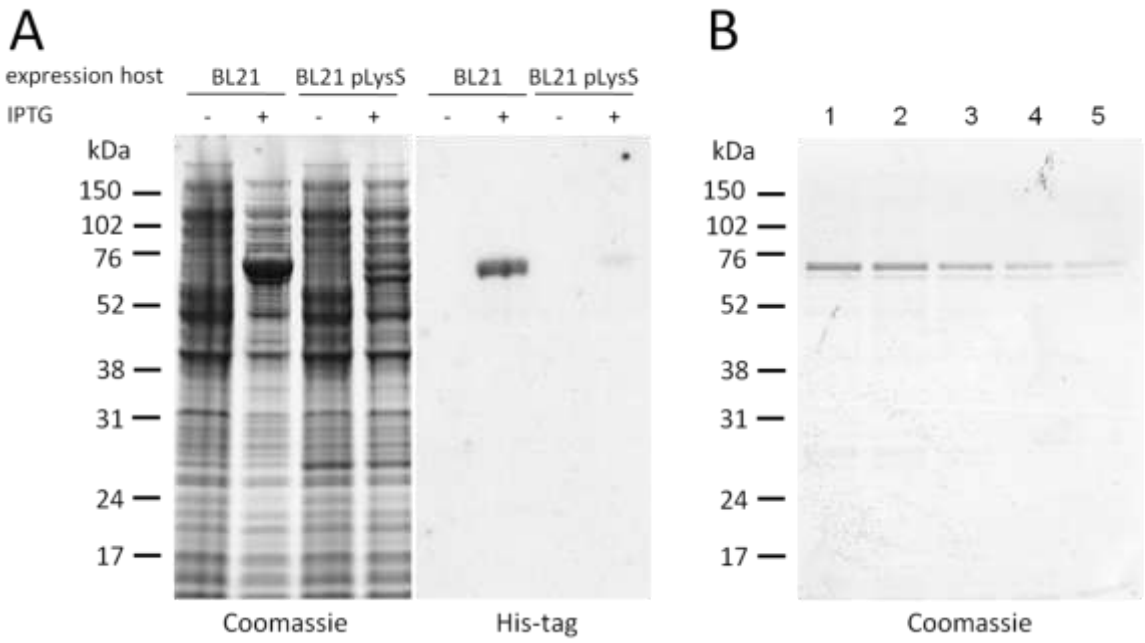


Figure 5.1 Expression and purification of His-PPAD. (A) *E.coli* BL21 (DE3) or BL21 (DE3) pLysS cells were transformed with a construct encoding His-PPAD (expected size 64 kDa) and recombinant protein expression was induced with IPTG ('+). No IPTG was added as a control ('-'). Cell lysates were resolved on an SDS-PAGE gel and total protein stained with Coomassie or transferred onto nitrocellulose membrane and recombinant protein detected with anti-his-tag antibody. **(B)** His-PPAD was purified from BL21 (DE3) transformed with His-PPAD construct using his-tag affinity resin. Elution fractions (1-5) are shown. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Large-scale expression (3-6 L) was performed twice, but the average recombinant protein yield was very low (66 µg/L culture; compared to Trx-His-PPAD: ~750 µg/L culture). As ~40 L of culture would have been required to produce the required amount of antigen (~2.5 mg), this approach was deemed ineffectual.

The alternative approach was therefore chosen, in which Trx-His-PPAD was used as an antigen, and tag-reactive antibodies would then be removed from the serum by positive selection towards a differently tagged PPAD (e.g. GST-PPAD) and/or negative selection against the purified Trx-His-tag. For this purpose, Trx-His-PPAD and Trx-His-tag were expressed and purified as described previously.

An expression vector encoding GST-PPAD was constructed by transferring the PPAD insert from the Trx-His-PPAD construct into a GST-encoding construct (pET49b(+)), using the same restriction sites. A GST-tag was chosen as it is a large solubility-enhancing tag and easy to detect with antibodies. BL21(DE3) and BL21 (DE3) pLysS cells were transformed with the expression construct and two clones from each strain were picked to assess recombinant protein expression in total cell lysates. Figure 5.2 indicates that full-length GST-PPAD was expressed only in BL21 (DE3) cells, as was also the case for Trx-His-PPAD and His-PPAD. It should be mentioned at this point that the BL21 (DE3) pLysS cells from the same batch were used successfully for expression of other proteins, therefore it can be excluded that they were faulty.

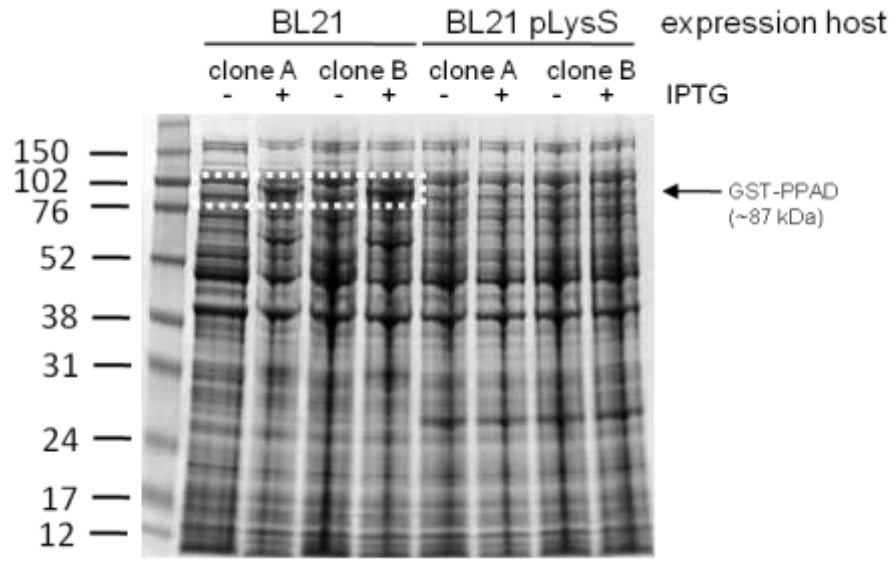


Figure 5.2 Expression of GST-PPAD. *E.coli* BL21 (DE3) or BL21 (DE3) pLysS cells were transformed with a construct encoding GST-PPAD (expected size 87 kDa) and two clones each were chosen to assess recombinant protein expression after induction with IPTG (‘+) or in the absence of IPTG as control (‘-‘). Cell lysates were resolved on an SDS-PAGE gel and total protein stained with Coomassie. Probable recombinant full-length GST-PPAD is highlighted with a white box and arrow. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Next, it was determined whether soluble GST-PPAD can be purified. The total cell extract of BL21 (DE3) cells was split into soluble and insoluble fraction by centrifugation and the soluble fraction analysed by SDS-PAGE and anti-GST immunoblot. Figure 5.3 A shows that soluble full-length GST-PPAD could be purified, along with a truncation of ~31 kDa. In a next step, it was assessed whether cleaving the GST-tag is an option. However, as shown in

Figure 5.3 B, cleavage was inefficient compared to simple elution of the tagged protein from the resin, as had also been the case for Trx-His-PPAD as described previously. Full-length GST-PPAD was therefore expressed and purified in large scale (6 L) to be used as an antigen for positive affinity purification of anti-PPAD antibodies. The yield was relatively high with ~2.4 mg/L culture (~3-fold higher compared to Trx-His-PPAD).

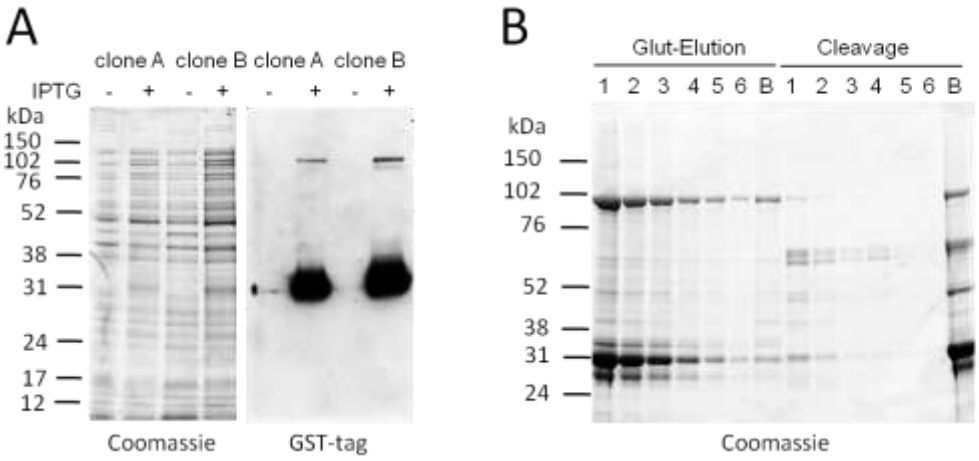


Figure 5.3 Purification of GST-PPAD. (A) *E.coli* BL21 (DE3) cells were transformed with a construct encoding GST-PPAD (expected size 87 kDa)) and two clones (A and B) were chosen to assess recombinant protein expression in the soluble cell fraction after induction with IPTG ('+) or in the absence of IPTG as control ('-'). Total protein was stained with Coomassie or transferred onto nitrocellulose membrane and recombinant protein detected with anti-GST-tag antibody. **(B)** Soluble GST-PPAD was purified from BL21 (DE3) transformed with GST-PPAD construct using GST-tag affinity resin. GST-PPAD was either eluted from the resin using glutathione ('Glut-Elution') or PPAD was cleaved from the GST tag by PreScission protease and the cleaved PPAD eluted from the resin (expected size: 62 kDa). Elution fractions (1-6) and non-eluted proteins from the resin beads ('B') are shown. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Screening of rabbit antiserum for anti-PPAD reactivity

Test bleed #2 (after the initial and two booster immunisations) and final bleed (after a further three booster immunisations) were screened for reactivity against Trx-His-PPAD and Trx-His-tag by ELISA, performed by the manufacturer, and by immunoblot, carried out by myself (see Material & Methods for the immunisation timetable). The ELISA results of the serum response can be found in the Appendix. In summary, after two booster immunisations, the immune response against both Trx-His-PPAD and tag was so strong that the OD values

exceeded the maximum value on the standardised chart, therefore appearing as a straight line at the top. This showed that the immune response is very effective, but also that there is reactivity towards the tag alone.

For the first screening with test bleed #2, serum from the two rabbits that were used for this experiment was assessed individually, even though the ELISA results indicated that the response is comparable between animals. The serum was assessed at dilutions 1:1,000 (=10^3), 1:10,000 (=10^4) and 1:100,000 (=10^5). The test antigen was a cleaved PPAD preparation which had not been further purified by size-exclusion chromatography and therefore contained some uncleaved Trx-His-PPAD and Trx-His-tag (please see Figure 4.11 B).

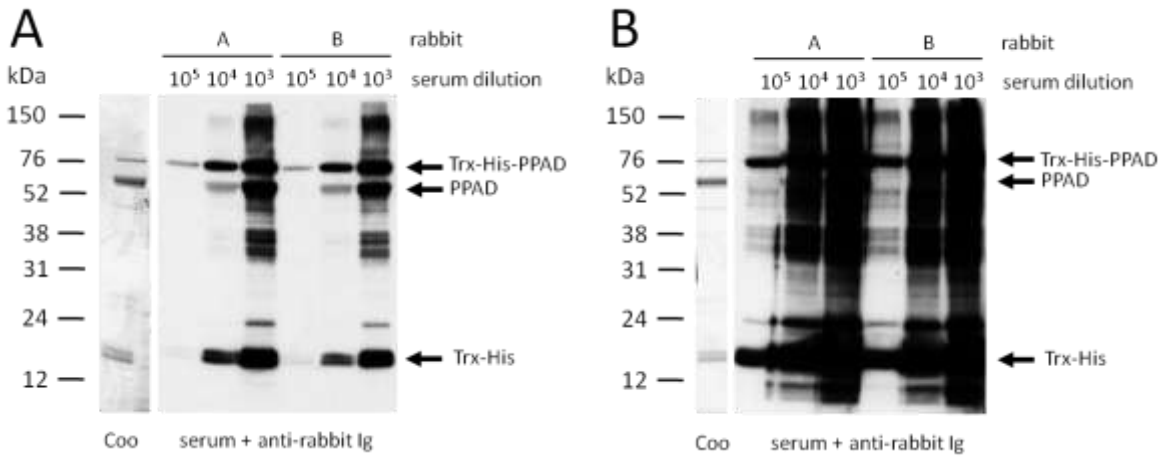


Figure 5.4 Screening of rabbit antiserum for reactivity towards PPAD. Two rabbits (No. 4488 = 'A'; No. 4489 = 'B') were immunised with Trx-His-PPAD in complete Freund's complete adjuvant. After two booster immunisations bleed #2 **(A)** was collected; the final bleed **(B)** was collected after another three booster immunisations. Serum was diluted as indicated and the reactivity towards 2.5 µg/lane of a PPAD preparation was evaluated using an immunoblot procedure. Bound antibodies were detected using HRP-conjugated swine anti-rabbit-Ig antibodies (1:5000). Exposure time was 20 seconds for both blots. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Figure 5.4 A reflects the ELISA results, showing that reactivity is strong, but is directed towards both PPAD and tag. This reactivity increased in the course of the immunisation experiment, as stronger signals were obtained at lower dilutions with the final bleed in Figure 5.4 B.

From these results it was clear that the reactivity towards the tag must be eliminated to reduce potential cross-reactivity with thioredoxin, especially in bacterial samples, as thioredoxin is a conserved protein. A small-scale experiment aimed at negative depletion of tag-reactive antibodies by pre-incubating the immune serum with Trx-His-tag did not give the desired results, showing that the net effect is a much weaker reactivity towards PPAD compared to Trx-His-PPAD or Trx-His tag (Figure 5.5).

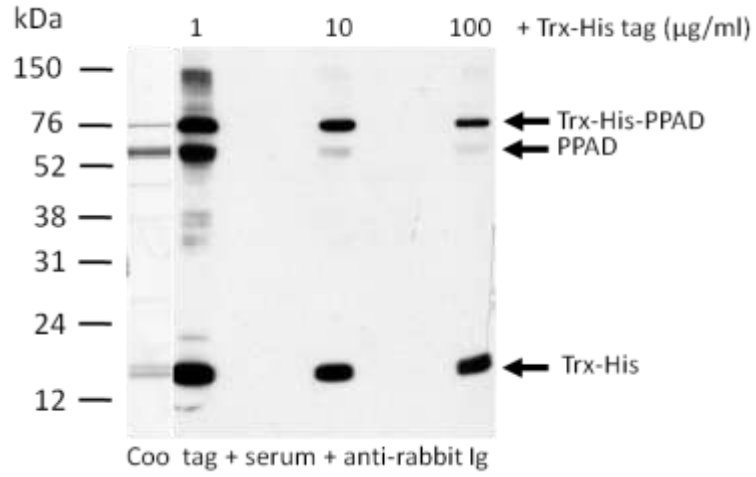


Figure 5.5 **Small-scale negative depletion of antiserum with Trx-His-tag.** Final bleed serum from rabbit B was pre-incubated with purified Trx-His-tag at 1, 10, 100 µg/ml for 1 hour at room temperature. The mix was then diluted to 1:1,000 and reactivity assessed towards 2.5 µg/lane of a PPAD preparation by immunoblotting. Bound antibodies were detected using HRP-conjugated swine anti-rabbit-Ig antibody. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

This result was confirmed using *P. gingivalis* wild-type, Δ*ppad*, and *E. coli* total cell lysates. As seen in Figure 5.6, without pre-incubation with tag, the antiserum recognised a doublet band in both *P. gingivalis* wild-type strains around 76 kDa, but not in Δ*ppad* or *E. coli*. A second strong band is visible around 12 kDa, most likely representing cross-reactivity with bacterial thioredoxin (predicted molecular weight 11.4, 11.8, and 15.5 kDa for *P. gingivalis* thioredoxin and *E. coli* thioredoxin-1 and -2, respectively). Upon pre-incubation of the antiserum with 10 µg/ml Trx-His-tag, the intensity of the majority of cross-reactive bands is reduced or eliminated, but so is the PPAD signal. Therefore the alternative approach, positive affinity purification using a GST-PPAD affinity resin column, was chosen.

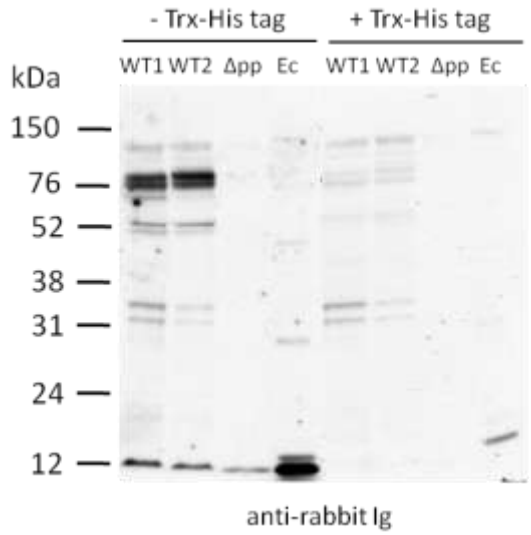


Figure 5.6 Negative depletion of antiserum on *P. gingivalis* cell lysates. Final bleed serum from rabbit B was pre-incubated with 10 μg/ml purified Trx-His-tag for 1 hour at room temperature, diluted 1:1000, and reactivity assessed towards total cell lysates of two *P. gingivalis* wild-types (clinical isolates D245 and MaRL), a *P. gingivalis* Δppad knockout strain, and an *E. coli* strain (DH5α). Bound antibodies were detected using HRP-conjugated swine anti-rabbit-Ig antibody. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Affinity purification was performed by the antibody manufacturer; GST-PPAD for derivatisation of the column was prepared by myself (12 mg). 77 ml harvest bleed from the two rabbits (pooled) resulted in two eluate fractions: glycine eluate (11 ml at 1.31 mg/ml protein) and TEA eluate (40 ml at 0.19 mg/ml protein). These two eluate fractions were then evaluated for reactivity towards recombinant PPAD and native PPAD (in *P.gingivalis* cell lysates) by myself using immunoblotting. Figure 5.7 shows that affinity purification of the antiserum, most evident with the TEA eluate at a dilution of 1:1000, resulted in a stronger reactivity towards PPAD compared to Trx-His-PPAD, tag, or truncated versions. *P. gingivalis* wild-type and Δ*ppad* cell lysates were then tested using the TEA eluate at different dilutions to establish the optimal concentration for studies of native PPAD. The result was positive, as a clean blot without thioredoxin crossreactivity was obtained at all dilutions (Figure 5.8).

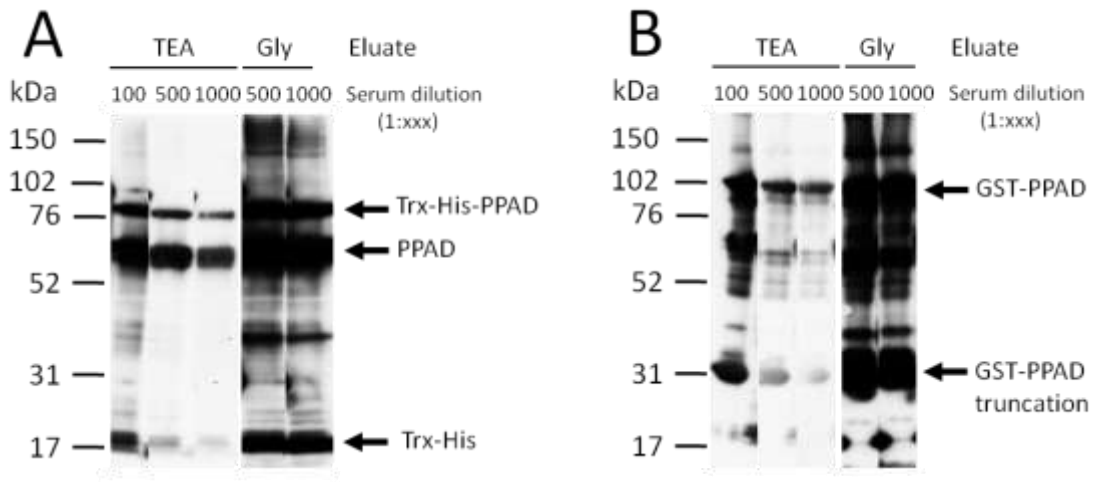


Figure 5.7 **Testing of affinity-purified antiserum.** 2.5 µg of a cleaved PPAD preparation (A) or a GST-PPAD preparation (B) were resolved by SDS-PAGE, transferred onto nitrocellulose membrane and immunoblotted with different dilutions of either TEA or glycine eluate. Bound antibodies were detected using HRP-conjugated swine anti-rabbit-Ig antibody. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

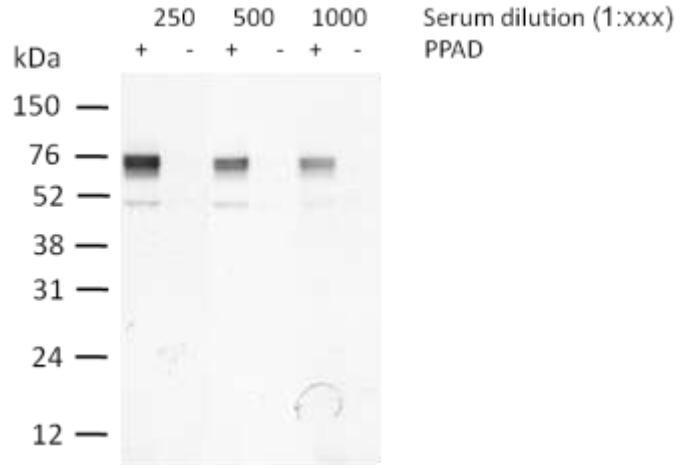


Figure 5.8 **Testing of TEA antiserum eluate on lysates of *P. gingivalis* WT and Δ*ppad* strain.** *P. gingivalis* total cell lysates of a wild-type (ATCC 33277) or the corresponding Δ*ppad* strain (indicated by '+/- PPAD') were resolved by SDS-PAGE, transferred onto nitrocellulose membrane and immunoblotted with different dilutions of TEA antiserum eluate. Bound antibodies were detected using HRP-conjugated swine anti-rabbit-Ig antibody. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Next, native PPAD was detected in a number of *P. gingivalis* wildtype, clinical isolate and Δ*rgp* strains, and compared to recombinant full-length PPAD. Figure 5.9 shows that native PPAD appears as a doublet band with approximately the same molecular mass as the full-length recombinant PPAD (no tag), in all strains except Δ*ppad*. A band that would indicate the proposed 'mature' form of PPAD (~46 kDa) was not detected. A weak band in this molecular weight range appeared in wildtype strain 33277 but was not reproducible and not found in other wild-type or clinical isolate strains. Another finding that supports a full-length versus 'mature' native PPAD is that deletion of arginine-gingipains did not affect the molecular mass of the detected PPAD. Arginine-gingipains were suggested to process PPAD. The lack of any effect of deletion of arginine-gingipains at least on the molecular mass of PPAD indicates that PPAD is not proteolytically processed by arginine-gingipains.

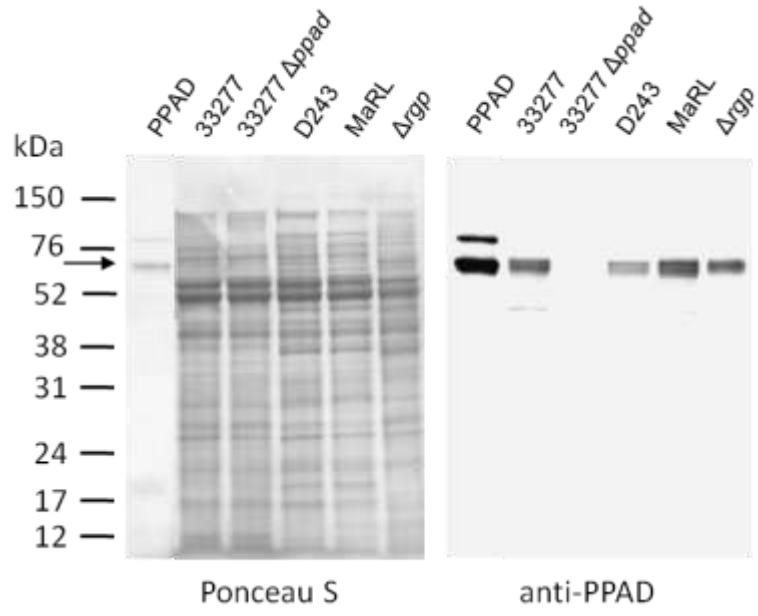


Figure 5.9 Detection of native PPAD in *P. gingivalis* strains. Total cell lysates of a wild-type strain (ATCC 33277), the corresponding Δ*ppad* strain, two clinical isolates (D243, MaRL), and the Δ*rgp* strain were resolved by SDS-PAGE, transferred onto nitrocellulose membrane and total protein stained with Ponceau S and immunoblotted with the TEA anti-PPAD antiserum (1:1000). A recombinant cleaved PPAD preparation was used as positive control ('PPAD'). Bound antibodies were detected using HRP-conjugated swine anti-rabbit-Ig antibody. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left.

Subcellular localisation of PPAD

The subcellular localisation of PPAD was examined using fractions prepared from *P. gingivalis* wild-type strain W83. From Figure 5.10 it becomes evident that PPAD is associated with fractions containing membrane, but not with the periplasm+cytoplasm fraction. The strongest band was observed in the outer membrane fraction. In conclusion, PPAD is most likely located on the bacterium's cell surface. No conclusions can be made about the vesicle+medium and medium fractions as the protein content was too low.

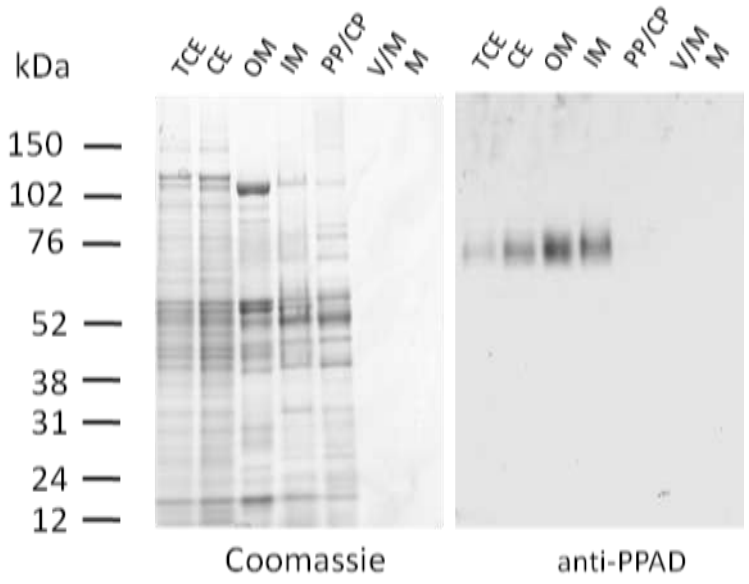


Figure 5.10 Subcellular localisation of PPAD in *P. gingivalis*. *P. gingivalis* strain W83 total cell extract (TCE) was fractionated into cell envelope (CE), outer membrane (OM), inner membrane (IM), periplasm+cytoplasm (PP/CP), vesicles+medium (V/M) and medium (M) fractions, which were then analysed with the anti-PPAD antibody by immunoblotting. The corresponding Coomassie-stained gel showing total protein in each of the fractions is shown on the left. The positions of the molecular mass markers in kilodalton (kDa) are indicated on the left. Fractions were kindly provided by Yonghua Gho, Louisville University, Kentucky.

These results support previous indirect data pointing to a cell surface location of PPAD: (1) In Chapter 3 it is described that incubation of whole *P. gingivalis* cells with fibrinogen/enolase results in the citrullination of the proteolytically processed peptides - assuming that the majority of bacterial cells are intact, citrullination would require cell-surface exposed PPAD; (2) In a previous publication (McGraw *et al.* 1999) PPAD activity was determined using either whole cells, supernatant or vesicles, and highest activity was

found with whole cells and vesicles, indicating that PPAD is membrane-attached; (3) Figure 3.6 in Chapter 3 shows that the majority of citrullinated proteins are found in the periplasm and membrane fraction, but not in the cytoplasm. Assuming that extensive translocation of citrullinated proteins between different subcellular compartments does not take place, this indicates a periplasm/membrane location of PPAD.

Cross-reactivity with human PADs

The availability of the anti-PPAD antibody made it possible to study the cross-reactivity with human PADs as a measure of homology. Figure 5.11 shows that the anti-PPAD antibody specifically recognises PPAD but not human PAD2 or PAD4, despite the previously suggested structural homology between PPAD and human PADs (Shirai *et al.* 2006). This is surprising as the antibody was raised against the whole protein, and not a specific peptide sequence, which would have explained the lack of reactivity due to the non-homologous nature of the amino acid sequences.

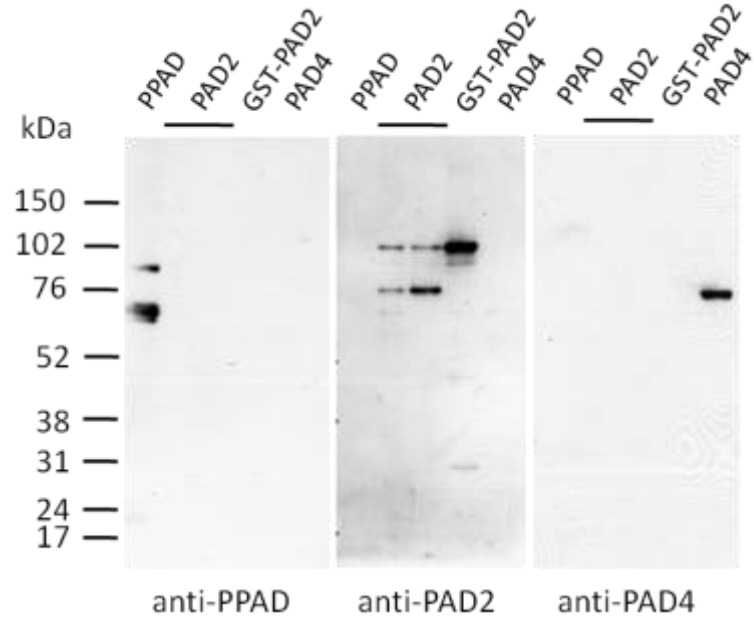


Figure 5.11 Cross-reactivity between *P. gingivalis* PAD and human PADs. Equivalent amounts of cleaved recombinant PPAD, three recombinant PAD2 preparations (two cleaved and one uncleaved GST-PAD2) and PAD4 (cleaved) were resolved by SDS-PAGE, transferred onto nitrocellulose membranes and probed with either anti-PPAD, anti-PAD2, or anti-PAD4 antibody. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left. Figure was contributed by our post-doc Dr. Elena Lugli.

On the other hand, immunoblotting is a denaturing technique and the results might simply reflect the different primary structure of PPAD and human PADs. Therefore, it is possible that the structural homology between PPAD and human PADs is not as significant as predicted, or that a sufficient amount of antibodies in the polyclonal anti-PPAD preparation recognise linear epitopes that are not present in human PADs due to the divergent amino acid sequences. A non-denaturing technique such as immunoprecipitation would be better suited to investigate this point further.

Detection of PPAD and citrullinated proteins in periodontal tissue and cell culture

Having shown that *P. gingivalis* generates endogenous and host citrullinated proteins and peptides *in vitro*, it was of interest to test whether these are also generated *in vivo*. First, gingival crevicular fluid (GCF) from 12 subjects with periodontitis was analysed. GCF is an inflammatory exudate containing bacterial and host products and can be collected by washing the gingival crevice with a saline solution using a micropipette. As the fluid has a serum-like constitution, occasionally mixed with blood, and contains large amounts of immunoglobulins and albumin, secondary antibody cross-reactivity in immunoblotting procedure is a point of concern and was always controlled for.

Figure 5.12 shows that discrete citrullinated protein bands were detected in samples 1, 3, 7, 10, 11, 12 and possibly also in 8 and 9 (difficult to judge because of the background). The citrullinated proteins are either in the ~15 kDa or ~50 kDa molecular mass range. Their identity was not investigated, but the ~15 kDa protein could be hemoglobin α chain (14 kDa), which is present in GCF and was shown to be citrullinated by PPAD (Figure 4.28). Human PADs could also contribute to citrullination in GCF as a result of inflammation and tissue damage. Unfortunately, these GCF samples are not well-characterised, and a correlation of protein citrullination to *P. gingivalis* load or inflammatory state is not possible. The anti-PPAD antibody was not available yet at this stage of the project. GCF from healthy subjects cannot be obtained using the same method, as they normally do not have enlarged periodontal pockets holding significant amounts of GCF.

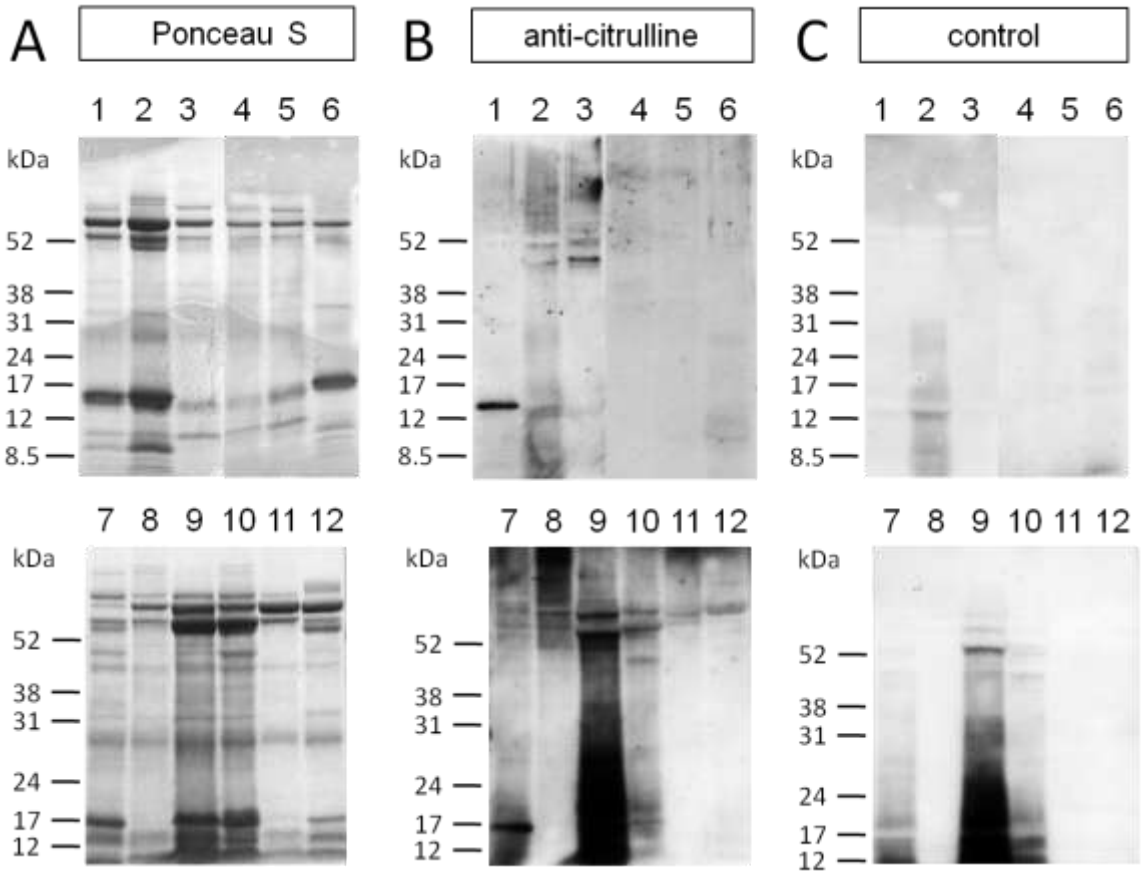


Figure 5.12 Analysis of protein citrullination in gingival crevicular fluid from periodontitis patients. Gingival crevicular fluid was collected from 12 periodontitis patients and resolved by SDS-PAGE (peptide gels). Total protein was visualised with Coomassie **(A)** or transferred onto nitrocellulose membranes for detection of citrullinated proteins with the anti-modified citrulline antibody **(B)** and positive signals due to cross-reactive secondary antibodies were control by omitting the primary antibody **(C)**. The positions of the protein molecular mass markers in kilodalton (kDa) are indicated on the left. Gingival crevicular fluid samples were kindly provided by Dr. Sigrun Eick, Bern University, Switzerland.

Next, periodontal tissue was analysed from 8 periodontitis patients and 4 controls for the presence of citrullinated proteins and PPAD. The control subjects were adults with no signs of periodontitis who were undergoing gingivectomy, which is a surgical aesthetic procedure including the removal of undesired gingival tissue. Some inflammation can therefore be expected in these tissues, even though much less severe than in periodontitis tissue. Figure 5.13 shows that PPAD could not be detected in any of the samples (B). In periodontitis patients, the tissue is taken from sites that show evidence of periodontal tissue inflammation (Figure 5.14) but may not directly harbour *P. gingivalis*, which is most exclusively found deep in the pocket near the root of the tooth (see O'Brien-Simpson *et al.* 2009).

This could explain the lack of detectable PPAD in the periodontitis samples. Further, *P. gingivalis* might be completely absent in a subset of periodontitis patients. Alternatively, *P. gingivalis* might be present in these tissues, but the protein amount analysed was not high enough resulting in insufficient bacterial cell mass for detection.

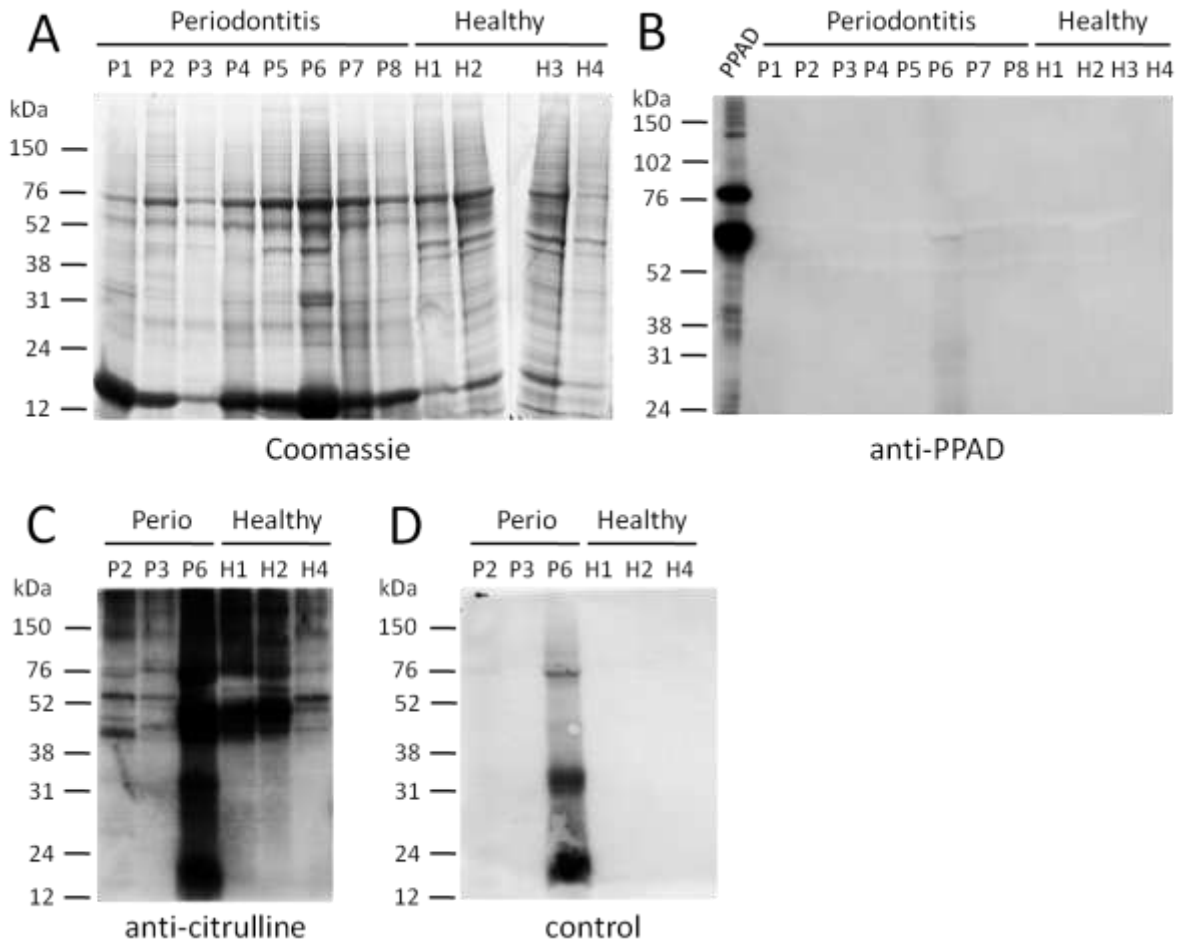


Figure 5.13 Analysis of protein citrullination and PPAD expression in gingival tissue from periodontitis patients and healthy controls. Gingival tissue from subjects with periodontitis (P1-P8) and non-periodontitis controls undergoing gingivectomy (H1-H4) was lysed and total protein resolved by SDS-PAGE. **(A)** Total protein was stained with Coomassie; **(B)** Total protein was transferred onto nitrocellulose membrane and the presence of PPAD analysed using anti-PPAD antibody. Recombinant PPAD was used a control; **(C)** Representative anti-citrulline immunoblot showing the presence of citrullinated proteins in a selection of periodontitis (P2, P3, P6) and healthy (H1, H2, H4) tissue samples; **(D)** Conjugate control to C. Tissue samples were kindly provided by Dr. Shauna Culshaw, Glasgow University, UK. Figure panel B was contributed by our research assistant Dr. Anne-Marie Quirke.

A representative anti-citrulline blot shows that citrullinated proteins in the 40-80 kDa range were detected in all samples, irrespective of periodontal health status (Figure 5.13 C). The presence of citrullinated proteins in tissue from both periodontitis patients and controls indicates that protein citrullination does occur regardless of disease, and it is not clear what the source of the PAD activity is (PPAD, human PADs) and which physiological process the citrullinated proteins indicate (normal physiology, general inflammation, periodontitis). In summary, the presence of citrullinated proteins does not seem to be a marker for periodontal disease or the presence of *P. gingivalis*.

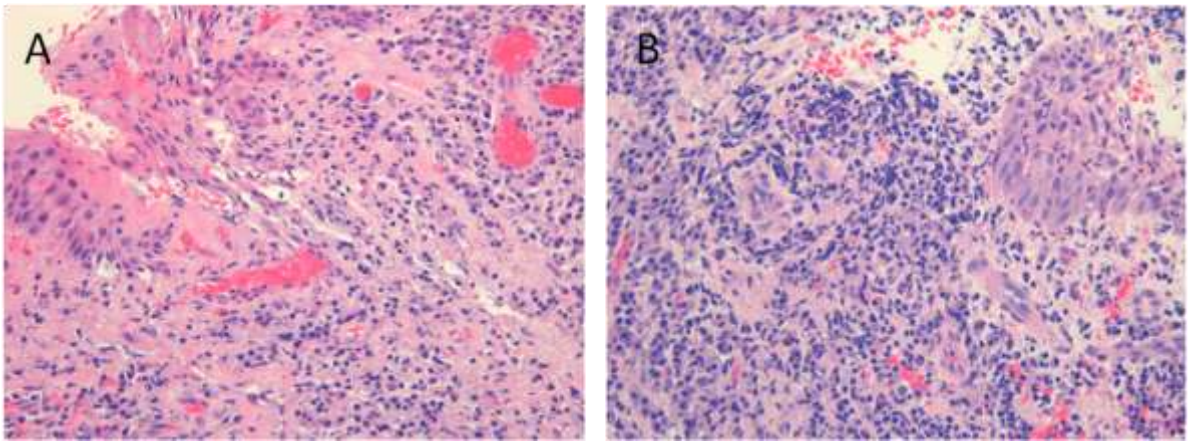


Figure 5.14 Representative hematoxylin & eosin - stained periodontal tissue from two periodontitis patients. Inflamed periodontal tissue was removed from two patients (A, B) with periodontitis during a surgical treatment procedure and stained with hematoxylin & eosin (H&E). Nuclei are stained blue and eosiniphilic structures (mostly cytoplasm/proteins) is stained red. Inflammatory cells (blue) infiltrating the gingival connective tissue (red) can be seen in both samples. Magnification: x200. Pictures were kindly provided by Dr. Shauna Culshaw, Glasgow University, UK.

To verify these results further, the anti-PPAD antibody was used to test for the presence of *P. gingivalis* by immunohistochemistry of periodontal tissue from a periodontitis patient. An *in vitro* cell culture using *P. gingivalis*-infected oral epithelial cells demonstrated that the antibody binds to *P. gingivalis* and does not cross-react with epithelial cells (Figure 5.15). No convincing evidence for the presence of *P. gingivalis* was obtained with a periodontal tissue sample (Figure 5.16). A small clump of stained objects was noticeable and could represent a cluster of *P. gingivalis* cells, but this would have to be confirmed in a larger number of samples.

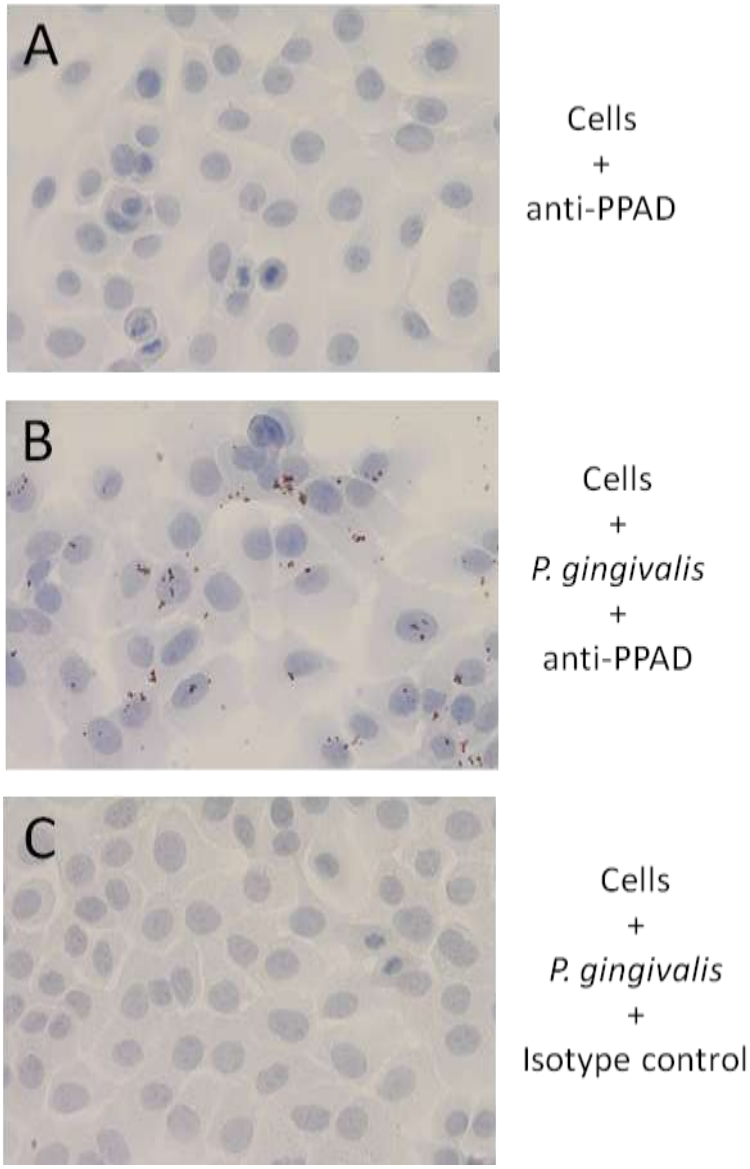


Figure 5.15 PPAD-staining of an oral epithelial cell line infected with *P. gingivalis*. An immortalised oral keratinocyte cell line (OKF6) was cultured under anaerobic conditions and stained with anti-PPAD antibody (A). As a positive control, cells were infected with *P. gingivalis* stained with anti-PPAD antibody (B) or isotype control (C). Cell nuclei were counterstained blue. Magnification : x400. Pictures were kindly provided by Dr. Shauna Culshaw, Glasgow University, UK.

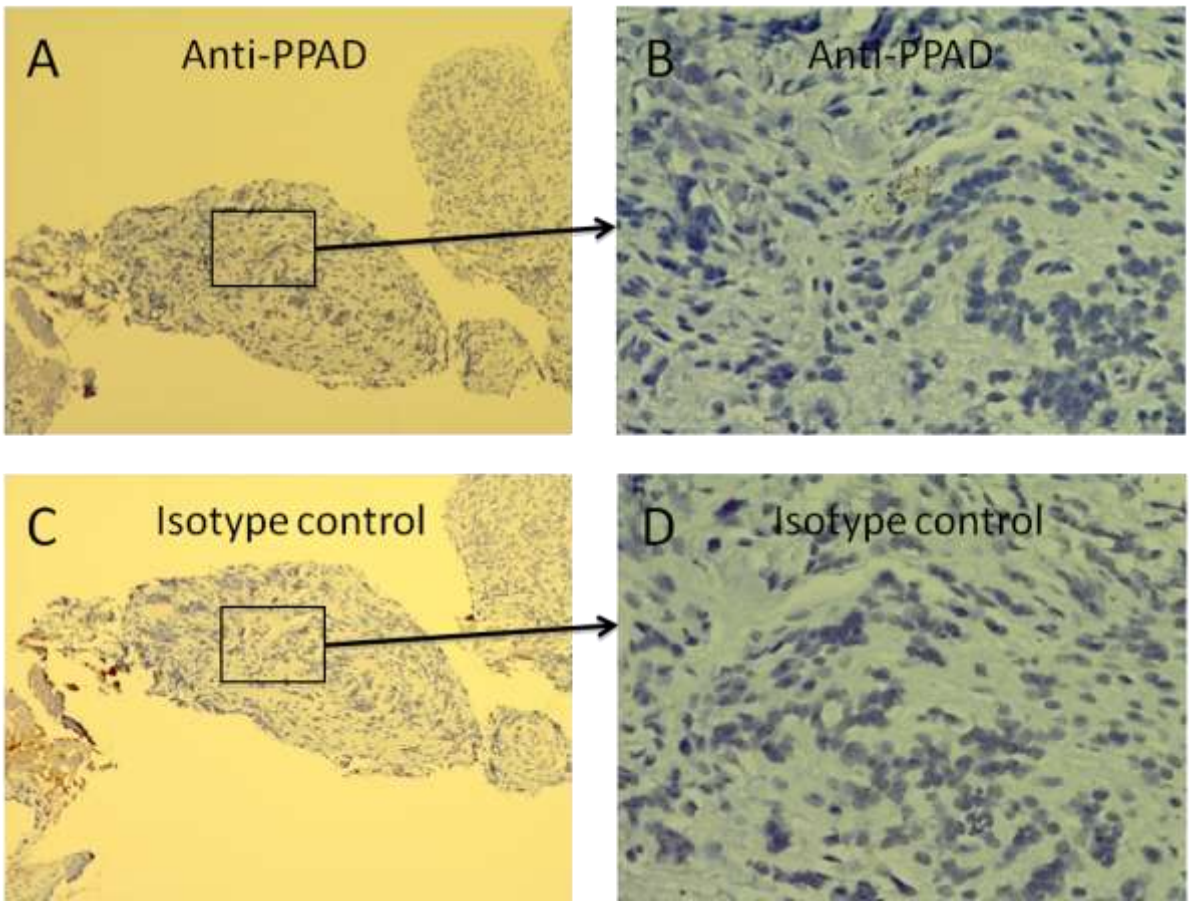


Figure 5.16 Periodontal tissue from a periodontitis patient stained with anti-PPAD antibody. (A, B) Inflamed periodontal tissue stained with anti-PPAD antibody conjugate (brown). (C, D) Isotype staining to control for unspecific signals. Nuclei were counterstained blue. Original magnification : x100 (A, C) and x400 (B, D). Pictures were kindly provided by Dr. Shauna Culshaw, Glasgow University, UK.

Summary

Anti-PPAD antibody generation and evaluation

- Trx-His-PPAD was used as immunising antigen as cleavage of Trx-His-tag would involve the use agents interfering with immunisation and His-PPAD expression levels were insufficient
- Rabbits developed a very strong immune response towards Trx-His-PPAD, with a considerable reactivity directed towards the Trx-His tag
- Trx-His tag reactivity was eliminated by positive affinity purification with GST-PPAD

- There is no cross-reactivity between PPAD and human PAD2/4: anti-PPAD antibody does not bind human PAD2 or PAD4; anti-PAD2 and anti-PAD4 antibodies do not bind PPAD

Native PPAD in P. gingivalis

- Native PPAD in *P. gingivalis* wild-type, clinical isolates and Δ*rgp* appears as a doublet band corresponding in size to the full-length form, ruling out the necessity for proteolytic cleavage into a 'mature' form by Rgp
- Native PPAD in *P. gingivalis* wild-type is associated with the membrane fraction, suggesting a cell-surface location

Citrullinated protein and PPAD expression in vivo

- Evidence for citrullinated proteins in gingival crevicular fluid from periodontitis patients
- Citrullinated proteins were also detected in gingival tissue from periodontitis patients and controls
- Whole *P. gingivalis* cells can be visualised in an oral epithelial cell culture using the generated anti-PPAD antibody
- PPAD could not be detected by immunoblot or immunohistochemistry in gingival tissue from periodontitis patients or controls

Chapter 6 Discussion and future directions

The aim of this study was to investigate the mechanism of protein citrullination by PPAD, a novel bacterial peptidylarginine deiminase enzyme from the periodontal pathogen *P. gingivalis*, and its potential to contribute to anti-citrulline autoimmunity in the development of rheumatoid arthritis.

The combined results provide evidence that *P. gingivalis* is an alternative source in the human host for generating citrullinated proteins and peptides. The underlying mechanism - proteolytic cleavage and subsequent citrullination at carboxy-terminal arginine residues - differs from that of the human PAD enzymes, which preferentially citrullinate internal arginine residues in whole proteins. Hence, PPAD creates (i) different epitopes as compared to established citrullinated RA autoantigens, and (ii) epitopes to which immunological tolerance may not exist, not only due to the presence of foreign citrullinated proteins from the bacterium, but also through a foreign mode of proteolytic processing and post-translational modification of host antigens.

Using fibrinogen and α-enolase as model antigens, it has been demonstrated that *P. gingivalis* rapidly generates small peptides with carboxy-terminal citrulline residues (Chapter 3). In RA, fibrinogen that is citrullinated at internal arginines has at least two possible pathogenic roles, serving as an autoantigen and disturbing the coagulation cascade and linked pathways (Nakayama-Hamada *et al.* 2008). In periodontitis, *P. gingivalis* gingipain-mediated degradation of human fibrinogen inhibits fibrinogen polymerization and results in the localised bleeding tendency which is typical for chronic periodontitis (Imamura *et al.* 1995), providing valuable nutrients to *P. gingivalis* and facilitating spreading. However, the role of fibrinogen peptides with carboxy-terminal citrulline residues, generated by the combined action of gingipain and PPAD, is yet unknown and opens up a novel area for future investigations. Arginine-gingipains cleave a number of other human proteins, releasing biologically active peptides with C-terminal arginines that have important roles in immunity and inflammation, such as C5a (Popadiak *et al.* 2007) and bradykinin (Imamura *et al.* 1994; Monteiro *et al.* 2009). Simultaneous citrullination of these peptides by PPAD might have a previously unappreciated role in human disease. Chapter 4 shows that proteins with naturally occuring C-terminal arginines such as hemoglobin α chain are citrullinated by PPAD *in vitro*

and this might play a role in hemoglobin degradation and heme acquisition by *P. gingivalis*. An important next step would be to identify citrullinated peptides/proteins from *in vivo* sources such as gingival crevicular fluid and investigate their function in the pathophysiology of *P. gingivalis* and to explore their potency for triggering a specific pathogenic immune response in animal models of RA.

In fact, the identification of *in vivo* citrullinated proteins was one of the initial aims of this study, but unfortunately could not be carried out because the only reliable antibody available on the market to detect citrullinated proteins (the anti-modified citrulline antibody) became unavailable in 2009 due to a problem in its manufacturing process. Alternative antibodies were tested but did not fulfil the requirements for these experiments, as they did not detect citrulline irrespective of neighbouring amino acids. The colorimetric method was subsequently chosen as a method of choice for detection of citrullination, which is suitable for the analysis of enzyme activity, but not sensitive enough to detect small amounts of citrullinated proteins. If the antibody or an alternative becomes available in the future, it would be interesting to investigate, in addition to the aims mentioned above, whether for example in a mouse model of *P. gingivalis* infection, citrullination at the site of infection differs between mice infected with wildtype *P. gingivalis* and those with the Δ*ppad* strain. This would allow a conclusion as to what extent citrullination by host PADs takes place at the site of infection.

The pathophysiologic role of bacterial citrullinated proteins from *P. gingivalis* is unclear. The results from Chapter 5 demonstrate a cell-surface localisation of PPAD. In gram-negative bacteria, the cell surface is a complex mixture of carbohydrates, lipopolysaccharides, lipids and proteins. It protects from and interacts with the environment. Both protein and peptide citrullination causes a change in the charge state and thereby is expected to directly affect interaction with binding partners, or indirectly through a change in structure in the case of proteins. Citrullinated proteins/peptides could therefore have a role in biofilm formation, adherence to host cells or to other bacteria, in spreading, recruitment of inflammatory cells and inactivation of immunological mediators. For example, arginine and short peptides harbouring exposed arginines inhibit the hemagglutinating activity of *P. gingivalis* (Hayashi *et al.* 1992; Murakami *et al.* 1992; Hayashi *et al.* 1993), and peptides harbouring C-terminal arginines in particular were shown to inhibit the binding of *P. gingivalis* fimbriae to fibronectin (Kontani *et al.* 1997). Hence, citrullination of cell-surface attached or released

bacterial peptides and proteins with C-terminal arginines by PPAD could play a role in the inactivation of potential endogenous inhibitors of hemagglutination and adhesion to the extracellular matrix and thereby facilitate colonisation and virulence. Of note, the reverse reaction of citrullination has not been described to date, although it is possible that *P. gingivalis* proteolytically degrades citrullinated peptides/proteins and thereby reverses the effect of citrullination of these targets. Another pathogenic role of protein/peptide citrullination might be the production of ammonia, which increases the local pH in the favour of *P. gingivalis*. However, citrullination is not the only source of ammonia, as it is also generated by agmatine and arginine deiminases during catalysis by *P. gingivalis* and other oral bacteria.

Apart from the specificity towards C-terminal arginine residues, it is unclear whether PPAD has stringent substrate specificity. Assuming a random selection of short peptides with a C-terminal arginine residue, it is unknown whether PPAD would citrullinate all peptides with equal efficiency, or whether certain anchor residues are favourable. In this study, two fibrinogen peptides with C-terminal arginine were probed as PPAD substrates, resulting in a comparable degree of citrullination (Figure 4.32 C). However, both peptides were originally identified in their citrullinated form from fibrinogen+*P. gingivalis* cultures, and are therefore suitable PPAD substrates by definition, limiting the power of conclusion. Comparing all identified citrullinated peptides (Table 3.2), it becomes apparent that serine, proline, and glycine are prominent, which might indicate that the presence of these amino acids at certain positions within the peptide is favourable for citrullination by PPAD.

The results of this study show that native PPAD in *P. gingivalis* cells is present in the full-length form, and not a truncated 'mature' form as has been suggested (McGraw *et al.* 1999). Chapter 4 shows that full-length recombinant PPAD is stable and active, whereas the 'mature' version is degraded during expression/purification. Five amino acids residues crucial for catalysis were identified: Cysteine-351, Asparagine-297, Histidine-236, Aspartic acid-130 and -238. Towards the final stages of this project, another research group published a paper (Rodriguez *et al.* 2009) presenting data that partly overlap with the work in this thesis (please see Table 6.1 for a comparison).

Table 6.1 Comparative summary of the resuls on PPAD expression and characterisation presented in this thesis and in a recently published paper (Rodriguez *et al.* 2009)

Feature	This thesis	Rodriguez *et al.* 2009
Expression and purification of 'mature' PPAD	Cloned into pET48b(+) vector encoding N-terminal thioredoxin-his tag, expressed in BL21 (DE3) cells at 30˚C, no expression in BL21 (DE3) pLysS cells, degraded during expression/purification	Cloned into pCRT7/NT-TOPO vector encoding N-terminal His-tag and Xpress epitope, expressed in BL21 (DE3) pLysS cells at 37˚C, no degradation
Expression and purification of full-length PPAD	Cloned into pET48b(+) vector encoding N-terminal thioredoxin-his tag or GST-tag, expressed in BL21 (DE3) cells at 30˚C, successful purification, no expression in BL21 (DE3) pLysS cells	Cloned into pCRT7/NT-TOPO vector encoding N-terminal His-tag and Xpress epitope, expressed in BL21 (DE3) pLysS cells at 37˚C, no degradation
Growth kinetics of *E. coli* host cells containing PPAD-encoding expression plasmid	Retarded growth observed, slow-growing	Retarded growth observed, slow-growing
Stability of purified PPAD	Full-length form: temperature-insensitive	'Mature' and full-length forms: cold-sensitive (aggregation at 4˚C)
Enzymatic activity of purified 'mature' PPAD	not tested because enzyme unavailable	100% = 0.175 U/mg K_M 18 mM
Enzymatic activity of purified full-length PPAD[1]	23.8 U/mg K_M 0.10 mM	40% of 'mature' PPAD

Feature	This thesis	Rodriguez *et al.* 2009
Co-factor requirements for PPAD activity	none	none
Substrate specificity of purified PPAD	Full-length form: specificity for C-terminal arginines	'Mature' form: specificity for C-terminal arginines
Auto-citrullination of PPAD	Full-length form: Active wild-type PPAD overexpressed in *E. coli* cells is auto-citrullinated, no additional measurable citrullination upon incubation in activity assay buffer or during storage	'Mature' form: PPAD overexpressed in *E. coli* cells is not citrullinated, but auto-citrullination occurs during prolonged storage, resulting in decreased activity (see comment in text)

[1] Same substrate, buffer, reaction conditions and calculations were used

The biological meaning behind the observed higher activity of the 'mature' form compared to the full-length form in the paper by Rodriguez *et al.* is unclear, as to date there is no evidence of such a 'mature' form of PPAD *in vivo*. The findings with regard to substrate specificity are comparable to the results in Chapter 4, in that peptide substrates with a C-terminal arginine are preferred and the activity with substrates containing internal arginines is insignificant.

One drawback to the paper is that the unspecific colour formation stemming from the substrate is not controlled for when using the colorimetric assay (buffer alone instead of buffer+substrate alone is used as blank), resulting in the interpretation of consistently low absorbance values as low but significant levels of enzyme activity. Notably, this is observed when using substrates containing internal arginine residues or lacking important elements of the arginine amino acid (e.g. agmatine, dimethyl-L-arginine), which supports the view that the "low activity" in fact is background colour formation stemming from the substrate.

Auto-citrullination of PPAD was also addressed by Rodriguez *et al.*, however with a different outcome and less convincing results. It is stated that recombinant PPAD, as freshly purified

from *E. coli*, does not contain citrulline residues, but slowly auto-citrullinates itself upon prolonged storage. Using quantitative amino acid analysis and the colorimetric citrulline detection assay, it is reported that zero citrulline residues are found with freshly purified PPAD, but 2 are found with PPAD that had been stored for 21 days, and that the activity of the enzyme decreases as a result. Unfortunately, it is not said whether this was reproducible and error bars are absent from all graphs, limiting the power to draw conclusions from this data. The results on PPAD auto-citrullination as presented in Chapter 4 of this thesis are more robust and better controlled, showing that freshly purified wild-type PPAD is already citrullinated and enzymatically active, whereas the functional mutants are not citrullinated. Here, the colorimetric assay did not give a significant positive signal in any of the experiments when used on wild-type PPAD alone, indicating that the absolute level of citrullination is low. Prolonged storage of the enzyme and re-assaying with the colorimetric method did not result in a change of signal. It is possible that the overall low enzymatic activity of PPAD in the study by Rodriguez *et al.* could have impacted on the ability of PPAD to citrullinate itself during expression. On the other hand, it seems surprising that 21 days of incubation time at room temperature are required to catalyse auto-citrullination. It is still unclear whether PPAD auto-citrullination has any physiological significance or whether it is an artefact. A next step would be to identify the citrullinated residues of recombinant PPAD by mass spectrometry and compare this to native PPAD immunoprecipitated from *P. gingivalis* using the anti-PPAD antibody obtained in this project.

A second paper was published by the same group in September 2010, identifiying cysteine-351 as a crucial amino acid for catalysis (Rodriguez *et al.* 2010). Site-directed mutagenesis was used to generate a PPAD C351S mutant by replacing cysteine-351 with serine in the 'mature' PPAD. A similar approach was used in Chapter 4 of this thesis, except that cysteine was replaced with alanine in full-length PPAD. The results are comparable, as the PPAD C351S mutant showed less than 0.01% activity of the wildtype enzyme, confirming that cysteine-351 is a crucial residue for analysis as shown in the present project. In this thesis, four additional residues have been identified that are required for catalysis: asparagine-297, histidine-236, aspartic acid-130 and -238. It is unknown whether the loss of enzymatic activity in these mutants is due to a direct loss of function because of missing functional elements or due to an indirect loss of function because of an altered three-dimensional structure. Spectroscopic analysis should help clarify this question. It has been suggested that these amino acids are directly involved in catalysis; therefore, a direct loss of function seems

probable. This is also supported by the fact that the mutant proteins did not show any obvious alterations in physical properties, such as solubility or temperature sensitivity.

The study by Rodriguez *et al.* (2010) also investigated iodoacetamide and iodoacetic acid as inhibitors of PPAD. Both compounds, like 2-chloroacetamidine which was used in Chapter 4 of this thesis, are sulfydryl-reactive alkylating agents (Figure 6.1), again suggesting a crucial cysteine residue in the catalytic centre of PPAD. This is further supported by the fact that the presence of substrate protected from inactivation, as reported by the authors. Mass spectrometry analysis of iodoacetamide-treated PPAD showed modification of cysteine-351, providing further evidence that cysteine-351 is essential for catalysis and modified by alkylating agents, resulting in loss of activity.

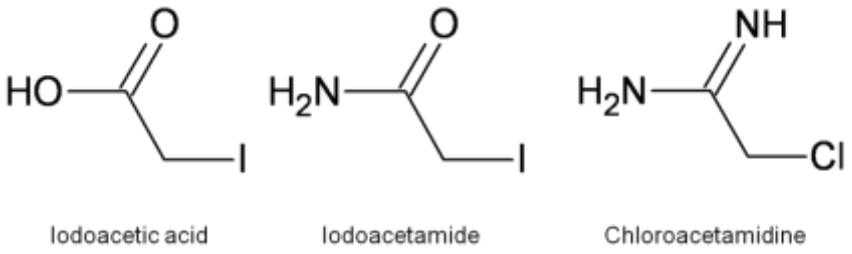


Figure 6.1 Chemical structures of PPAD inhibitors. The structures were created by myself using ACD/Chemsketch.

PPAD inhibition is a potential strategy for the development of an antibacterial agent and also an anti-rheumatic drug, providing the link between *P. gingivalis* infection and RA is substantiated. Its role as a definite virulence factor for *P. gingivalis* has not yet been published, but preliminary studies with our collaborator Jan Potempa, Jagiellonian University Krakaw, Poland, convincingly show that PPAD is indeed a virulence factor in the subcutaneous chamber model of *P. gingivalis* infection. In this model, live *P. gingivalis* cells are injected into a small coil-shaped chamber that has been subcutaneously implanted into mice. The Δ*ppad* knockout strain is rapidly cleared by the immune system, whereas the wild-type strain results in high levels of local and systemic inflammation. This is reflected by lower numbers of local pro-inflammatory cells but a higher proportion of viable pro-inflammatory cells, and less weight loss in mice infected with the Δ*ppad* knockout strain as compared to wild-type (Jan Potempa and Katarzyna Maresz, personal communication). Indirect evidence comes from a recent proteomics study reporting that PPAD is one of 14

proteins that are upregulated *in vivo* using the same model, when compared to *P. gingivalis* grown in culture (Yoshimura *et al.* 2008).

Targeting PPAD could be achieved by (i) direct inhibition with small chemical compounds; (ii) direct inhibition using monoclonal antibodies (passive vaccination); (iii) indirect inhibition using a PPAD vaccine (active vaccination). Chapter 5 shows that PPAD is surface-exposed, making it a good drug target. Its catalytic mechanism is, in principle, analogous to human PADs and other members of the guanidine-group modifying enzymes superfamily, raising the question of specificity and potential side effects when using small chemical compounds that target the catalytic centre. Simple alkylating agents such as chloroacetamidine or iodoacetamide are unspecific and would be expected to be toxic due to their general alkylating activity. More sophisticated and targeted inhibitors would be required, such as those currently in development for human PADs (Knuckley *et al.* 2010b; Willis *et al.* 2011). A potential strategy could be to capitalise on the substrate specificity of PPAD towards C-terminal arginines, by combining an alkylating warhead with a suitable peptide substrate containing a C-terminal arginine.

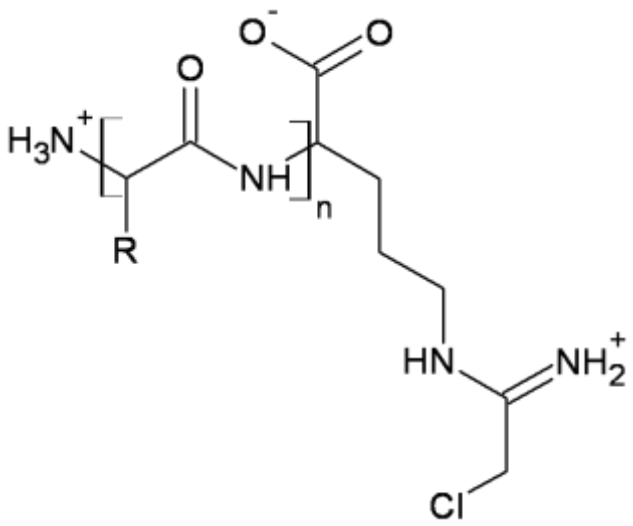


Figure 6.2 Chemical structure of a potential peptide-based PPAD inhibitor. Depicted is a peptide consisting of n amino acids and a carboxy-terminal arginine derivative in which the ω amino group is replaced by a chloromethyl group. Alternatively, a fluoromethyl group could be used. R = amino acid side chain. This structure was created by myself using ACD/Chemsketch.

A drawback to using peptide-based inhibitors is their bioavailability with the risk of degradation by proteases/peptidases. An alternative approach would be direct inhibition by

monoclonal antibodies. Monoclonal antibodies have the advantage of being highly specific, but are associated with higher development and production costs. Although having a similar catalytic mechanism, the amino acid sequence of PPAD has no significant similarity to any known human proteins. The three-dimensional structure of the catalytical domains of PPAD and human PADs were suggested to be homologous based on bioinformatic prediction (Shirai *et al.* 2006), but the overall surface-accessible structure might be significantly different, enabling the development of an inhibitory specific monoclonal antibody to PPAD without crossreactivity to human PADs or other members of the GME superfamily.

The question arises, however, whether such an antibody would bind PPAD in a way that prevents enzymatic activity and whether the antibody would resist the proteolytic attack by gingipains. This is also important for the third approach, vaccination against PPAD. PPAD is surface-exposed and probably highly immunogenic due to its pathogenic origin. The extraordinarily high antibody response to PPAD in rabbits after immunisation, as reported in Chapter 5, supports this view. However, for a vaccine or an antibody to be effective, the antibodies must reach the site of infection, bind to their target(s) with high affinity and avidity and activate downstream effector mechanisms. A degradation of antibodies by gingipains would preclude efficient binding of the antibody and recruitment of downstream effectors. Therefore, a solution for both the monoclonal antibody and vaccine approach could be the combination of gingipain and PPAD antibodies/vaccine, targeting the two potentially most potent virulence factors in *P. gingivalis*. In fact, topical application of a monoclonal antibody (Mab 61BG1.3) that recognises the hemagglutinin domain of gingipains onto the cleaned tooth surfaces has been shown to prevent recolonisation by *P. gingivalis* for up to 6 months. The association with periodontal disease activity was less evident (Booth *et al.* 1997). Vaccination efforts so far used heat- or formalin-killed whole *P. gingivalis* cells, or specific antigens, namely gingipains, fimbriae, and capsular polysaccharide (reviewed by Choi *et al.* 2010), showing that immunised animals (mice, rats, guiny pigs, macaque monkey) produce a significant antibody response to *P. gingivalis*, compared to controls (reviewed by Persson 2005). Protection from colonisation with *P. gingivalis* and lower levels of alveolar bone loss following immunisation were reported only in a few studies, most of them using whole cells as immunogen (reviewed by Persson 2005), further supporting the view that a combination of antigens that target multiple virulence pathways, such as gingipains and PPAD, might be the way forward. A major obstacle in the development of a vaccine is the fact that experimental animal models of periodontal disease are not representative of the multifactorial and

polymicrobial nature of human periodontitis, and trials in humans have not yet been performed. As periodontitis is a polymicrobial disease, prevention of colonisation by or the virulent activity of *P. gingivalis* alone might not eradicate the disease in all cases, but nevertheless would be expected to significantly lower incidence when used as a preventive strategy, and decrease disease activity when used as a therapeutic strategy. A major effort to develop a vaccine is underway at the Cooperative Research Centre for Oral Health Science from the University of Melbourne, in cooperation with the biopharmaceutical companies CSL Ltd and Sanofi Pasteur, reporting that clinical trials in humans are scheduled to start in three to four years.[21]

Summary and Outlook

The initial simple hypothesis on which this project and several other studies and reviews at that time were based, is that periodontal inflammation in the presence of *P. gingivalis* results in protein citrullination by PPAD, which - in the context of genetic and environmental susceptibility factors - triggers a defensive antibody response against bacterial citrullinated proteins. These antibodies then cross-react with conserved human antigens in the joint (e.g. enolase), which are citrullinated by human PADs as a result of subclinical tissue injury, causing more tissue injury and contributing to the development of chronic joint inflammation.

The discovery in the course of this project that PPAD citrullinates C-terminal arginines only was an important step towards evaluating this hypothesis. This indicated that citrullination of bacterial proteins at internal arginines by PPAD, as a potential mechanism for triggering autoantibodies to conserved human homologues (i.e. enolase) via molecular mimicry, which initially was the hypothesis behind this work (Lundberg *et al.* 2008 and Figure 1.10), is unlikely. Citrullination of bacterial proteins at internal arginines could still be the result of the action of human PAD enzymes present at the site of oral inflammation. To investigate this alternative scenario, it would be necessary to test whether human PADs are expressed and active at the site of oral inflammation, whether oral inflammation triggers citrullination, whether citrullinated (at internal arginines) bacterial antigens can be detected at the site, and whether an antibody response is triggered in periodontitis patients to these antigens. A recent study presented as an abstract at the American College of Rheumatology Meeting in 2009 (Nesse *et al.* 2009) reported the presence of citrullinated proteins in periodontal tissue from

[21] http://www.crcoralhealthscience.org.au

periodontitis patients (6 out of 15) using immunohistochemistry. Citrullinated proteins and human PAD2 were also demonstrated in gingival crevicular fluid of controls taken from sites with low levels of gingival inflammation (4 and 6 out of 8, respectively), using immunoblotting. Although healthy controls were not investigated in this study, the results show that protein citrullination is widespread in inflamed tissue and that human PAD2 is expressed at sites of gingival inflammation. These results support the findings in Chapter 5: analysis of *in vivo* citrullination in periodontal tissue showed that citrullinated proteins can be found in tissue from both healthy subjects and patients with periodontitis. Taken together, these results show that human PADs are active in healthy periodontal tissue as well as bacterial periodontal inflammation, which means that citrullination in gingival tissue is not in itself an indicator of periodontitis or infection with *P. gingivalis* and that citrullination by of proteins from pathogenic bacteria by human PADs during periodontal inflammation could contribute to the generation of bacterial antigenic epitopes, such as citrullinated *P. gingivalis* enolase. It must be noted that the potential contribution of other pathogenic oral bacteria or commensals is not taken into account here.

The finding that PPAD and human PADs produce different citrullinated epitopes prompted the view that antibodies that are generated against PPAD-generated epitopes might not recognise proteins/peptides generated by human PADs. Thus, the antibody response in serum from RA patients against citrullinated fibrinogen peptides that are generated by *P. gingivalis* PAD (C-terminal citrulline) versus analogous peptides that are generated by human PADs (internal citrulline) was investigated. The results showed that reactivity to these PPAD-generated peptides is infrequent (5-6%) and not statistically significant compared to healthy controls, whereas the reactivity to PAD-generated epitopes is frequent (23-46%) and statistically significant. Although only two such peptide pairs were tested, the results suggest that in active RA, reactivity to these PPAD-generated peptides is indeed absent or low.

The findings and discussion presented here led to an amendment and refinement of the initial hypothesis: there are two sources for citrullinated proteins/peptides in periodontal tissue: (i) human citrullinated proteins generated by human PADs are constitutively expressed in health and/or induced temporarily as a result of local tissue injury (for example during surgery or gingival inflammation), an immune response to these proteins is not mounted as they are the result of normal physiology; (ii) human and *P. gingivalis* citrullinated proteins are generated during periodontal infection with *P. gingivalis* by both human PADs and *P. gingivalis* PAD.

Proteins/peptides citrullinated by *P. gingivalis* PAD contain C-terminal citrulline residues. Proteins/peptides citrullinated by human PADs contain internal citrulline residues. Human and bacterial peptides citrullinated by *P. gingivalis* PAD contain epitopes that cannot be generated by the activity of human PADs and are therefore likely to have been previously hidden from immune surveillance. In the context of bacterial infection and periodontal inflammation in a genetically susceptible host, this may trigger an immune response. Bacterial proteins citrullinated by human PADs at internal arginines (e.g. *P. gingivalis* enolase) may also trigger an immune response due to their foreign origin and presentation in an infectious context. Antibodies bind to internally citrullinated epitopes on joint proteins (e.g. fibrinogen, enolase) as a result of cross-reactivity and epitope spreading, causing more tissue injury and contributing to the development of chronic joint inflammation. Figure 6.3 depicts the amended hypothesis based on the results and discussion in this thesis.

There are still several uncertainties in this hypothesis, which requires further careful testing. One important step will be to investigate a large and well-characterised cohort of serum from periodontitis patients for antibodies to *P. gingivalis*-generated citrullinated epitopes, as it is unknown whether the citrullinated proteins/peptides generated by *P. gingivalis* are immunogenic. If indeed an effective antibody response is mounted to citrullinated epitopes in periodontitis, longitudinal studies would be required to investigate whether there is an association with the development of seropositive RA.

The potential role of T cells is another important factor to consider. In periodontal disease, different T cell subsets seem to predominate at different stages of disease (progressive vs. stable) and with this different cytokine profiles contribute to the disease course, with a bias towards a Th1-response at diseased sites, although this is much debated, and evidence for an involvement of Th17 cells exists (reviewed by Gemmell *et al.* 2007). However, studies also show that gingipains degrade a range of Th1 and Th2 cytokines (IL-12, INF-γ, IL-4, IL-5) (reviewed by Pathirana *et al.* 2010). The result might therefore be general immune dysregulation rather than a specific bias towards a certain T cell response. It is unknown how citrullinated peptides and proteins generated by *P. gingivalis* are processed and presented by antigen-presenting cells.

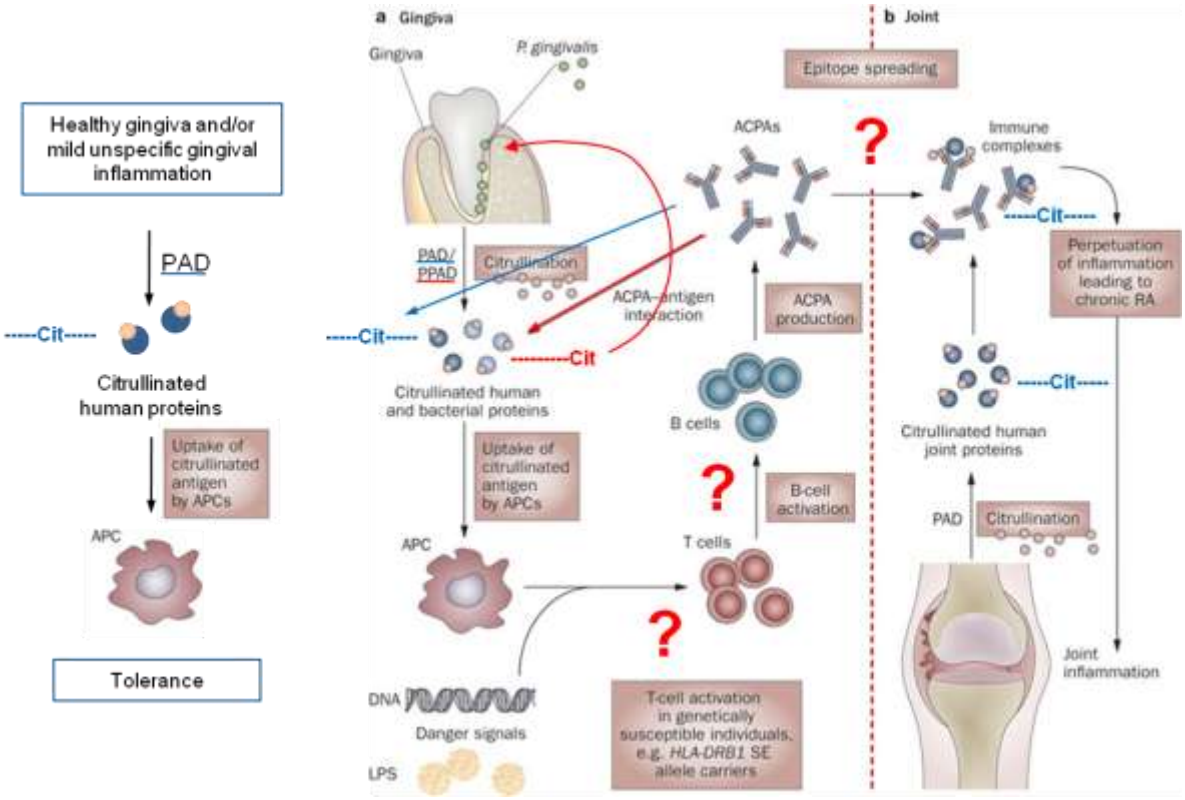


Figure 6.3 Schematic illustration of the amended hypothesis for the role of *P. gingivalis*-mediated citrullination in the development of RA. a) Gingiva: human proteins/peptides are citrullinated at internal arginines in healthy gingiva or during mild unspecific inflammation through the activity of human PADs (blue peptide). Citrullinated proteins are recognised as 'self' and an immune response is not mounted. During periodontal inflammation in the presence of *P. gingivalis*, human and bacterial proteins/peptides are citrullinated at internal arginines by human PADs (blue peptides) and at C-terminal arginines by PPAD (red peptides). In the context of danger signals, the immune system is activated and antibodies against these citrullinated antigens are produced. Host and *P. gingivalis* citrullinated proteins/peptides generated by PPAD may assist the bacterium in adhesion to host cells, immune evasion and nutrient acquisition (symbolised by the red loop). b) Joint: a second inflammatory event occurs in the joint, leading to citrullination of joint proteins at internal arginines by human PADs. Anti-citrullinated protein/peptide antibodies initially formed as a result of *P. gingivalis*-mediated periodontal inflammation recognise citrullinated joint antigens as a result of epitope spreading and cross-reactivity. The formation of immune complexes in the joint perpetuates the inflammatory process and contributes to the development of RA. Abbreviations: ACPA, anti-citrullinated protein/peptide antibody; APC, antigen presenting cell; LPS, lipopolysaccharide; *P. gingivalis*, *Porphyromonas gingivalis*; PAD, peptidylarginine deiminases; PPAD, *P. gingivalis* PAD; RA, rheumatoid arthritis; SE, shared epitope. 'Bacterial' refers to *P. gingivalis*. The role of other oral pathogenic bacteria or commensals is not taken into account here. Question marks indicate areas for future investigation. This figure is based on a published review article co-authored by myself (Lundberg *et al.* 2010).

Further, the activation of citrulline-specific B cells would rely on the presentation of the peptide in the peptide-binding groove of MHC molecules in such a manner than the C-terminal citrulline residue is specifically recognised by T cell receptors. Studies aimed at demonstrating the activation of citrulline-specific T cells in human RA or animal models have been of only limited success, with only two studies demonstrating specific activation of T cells: peripheral blood mononuclear cells from RA patients but not from healthy controls proliferated upon stimulation with a citrullinated aggrecan peptide, but not when stimulated with the arginine-containing controls peptide (von Delwig *et al.* 2010); however, citrullinated aggrecan is not a known target of autoantibodies in RA. Further, a citrullinated vimentin-derived peptide was shown to activate T cells from a DR4-IE transgenic mouse (Hill *et al.* 2003). Studies using the CEP-1 peptide from enolase have been unfruitful so far (Patrick Venables, personal communication). These results may indicate that T cells do not play a major role in the induction of citrulline-specific autoantibodies in RA. Before any prediction for periodontitis can be made, more studies on the T cell and antibody response to citrullinated peptides in periodontitis patients are indicated.

Despite the fact that bacterial citrullination offers an attractive hypothesis for the explanation of the association between periodontitis and RA on a molecular basis, this association might also be indirect, in that periodontal inflammation generates a chronic systemic pro-inflammatory state where bacterial products (e.g.LPS) dysregulate the immune system, exacerbating joint inflammation and the formation of an autoimmune response (discussed by Mercado *et al.* 2003; de Pablo *et al.* 2009). A recent study might support this view. Sponges impregnated with heat-killed *P. gingivalis* were implanted into the backs of rats, establishing a chronic inflammatory lesion, and adjuvant arthritis was induced 35 days later. The authors reported that severe arthritis developed more rapidly in animals with pre-implanted *P. gingivalis* (Bartold *et al.* 2010). While the mechanism of arthritis enhancement cannot be deducted from the data, it shows that the presence of inactivated *P. gingivalis* has the potential to exacerbate the development of arthritis. It is unclear whether enzymatic mechanisms such as citrullination by PPAD could have partly mediated the effect, as the bacteria were heat-killed at 60˚C for 5 min, and it is not known to what extent PPAD resists such conditions. However, the contribution of human PADs at the site of inflammation cannot be excluded. Studies at a collaborator's laboratory have shown that CIA-induced arthritis in mice is more severe in the presence of (live) *P. gingivalis* wild-type bacteria as

compared to the Δ*ppad* strain, indicating an active role of PPAD in the exacerbation of arthritis (Jan Potempa and Katarzyna Maresz, personal communication).

Whether direct or indirect, should a link between *P. gingivalis* and RA be confirmed, the prevention and treatment of periodontal disease would become an even more important health focus and drugs targeting important virulence factors such as PPAD and gingipains would be expected to be a promising strategy.

References

AAOP (1996). "Position paper: epidemiology of periodontal diseases. American Academy of Periodontology." *J Periodontol* **67**(9): 935-945.

Al-Katma, M. K., N. F. Bissada, J. M. Bordeaux, J. Sue and A. D. Askari (2007). "Control of periodontal infection reduces the severity of active rheumatoid arthritis." *J Clin Rheumatol* **13**(3): 134-137.

Alley, C. S., R. A. Reinhardt, C. A. Maze, L. M. DuBois, T. O. Wahl, W. C. Duckworth, J. K. Dyer and T. M. Petro (1993). "HLA-D and T lymphocyte reactivity to specific periodontal pathogens in type 1 diabetic periodontitis." *J Periodontol* **64**(10): 974-979.

Ally, N., J. C. Whisstock, M. Sieprawska-Lupa, J. Potempa, B. F. Le Bonniec, J. Travis and R. N. Pike (2003). "Characterization of the specificity of arginine-specific gingipains from Porphyromonas gingivalis reveals active site differences between different forms of the enzymes." *Biochemistry* **42**(40): 11693-11700.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." *J. Mol. Biol.* **215**(3): 403-410.

Amin, A. R., M. G. Attur, G. D. Thakker, P. D. Patel, P. R. Vyas, R. N. Patel, I. R. Patel and S. B. Abramson (1996). "A novel mechanism of action of tetracyclines: effects on nitric oxide synthases." *Proc Natl Acad Sci U S A* **93**(24): 14014-14019.

Arita, K., H. Hashimoto, T. Shimizu, K. Nakashima, M. Yamada and M. Sato (2004). "Structural basis for Ca(2+)-induced activation of human PAD4." *Nat Struct Mol Biol* **11**(8): 777-783.

Arita, K., H. Hashimoto, T. Shimizu, M. Yamada and M. Sato (2003). "Crystallization and preliminary X-ray crystallographic analysis of human peptidylarginine deiminase V." *Acta Crystallogr D Biol Crystallogr* **59**(Pt 12): 2332-2333.

Arita, K., T. Shimizu, H. Hashimoto, Y. Hidaka, M. Yamada and M. Sato (2006). "Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4." *Proc Natl Acad Sci U S A* **103**(14): 5291-5296.

Armitage, G. C. (1999). "Development of a classification system for periodontal diseases and conditions." *Ann Periodontol* **4**(1): 1-6.

Arnett, F. C., S. M. Edworthy, D. A. Bloch, D. J. McShane, J. F. Fries, N. S. Cooper, L. A. Healey, S. R. Kaplan, M. H. Liang, H. S. Luthra and et al. (1988). "The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis." *Arthritis Rheum* **31**(3): 315-324.

Asquith, D. L., A. M. Miller, I. B. McInnes and F. Y. Liew (2009). "Animal models of rheumatoid arthritis." *Eur J Immunol* **39**(8): 2040-2044.

Baeten, D., I. Peene, A. Union, L. Meheus, M. Sebbag, G. Serre, E. M. Veys and F. De Keyser (2001). "Specific presence of intracellular citrullinated proteins in rheumatoid arthritis synovium: relevance to antifilaggrin autoantibodies." *Arthritis Rheum* **44**(10): 2255-2262.

Baka, Z., E. Buzas and G. Nagy (2009). "Rheumatoid arthritis and smoking: putting the pieces together." *Arthritis Res Ther* **11**(4): 238.

Banbula, A., M. Bugno, J. Goldstein, J. Yen, D. Nelson, J. Travis and J. Potempa (2000). "Emerging Family of Proline-Specific Peptidases of Porphyromonas gingivalis: Purification and Characterization of Serine Dipeptidyl Peptidase, a Structural and Functional Homologue of Mammalian Prolyl Dipeptidyl Peptidase IV." *Infect. Immun.* **68**(3): 1176-1182.

Bang, S. Y., T. U. Han, C. B. Choi, Y. K. Sung, S. C. Bae and C. Kang (2010). "Peptidyl arginine deiminase type IV (PADI4) haplotypes interact with shared epitope regardless of anti-cyclic citrullinated peptide antibody or erosive joint status in rheumatoid arthritis: a case control study." *Arthritis Res Ther* **12**(3): R115.

Bartold, P. M., V. Marino, M. Cantley and D. R. Haynes (2010). "Effect of Porphyromonas gingivalis-induced inflammation on the development of rheumatoid arthritis." *J Clin Periodontol* **37**(5): 405-411.

Bartold, P. M., R. I. Marshall and D. R. Haynes (2005). "Periodontitis and rheumatoid arthritis: a review." *J Periodontol* **76**: 2066-2074.

Barton, A., J. Bowes, S. Eyre, K. Spreckley, A. Hinks, S. John and J. Worthington (2004). "A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population." *Arthritis Rheum* **50**(4): 1117-1121.

Benjamin, P. A., P. A. Rogers, S. U, N. W. Johnson, M. F. Cole and M. A. Curtis (1997). "Increased titre and avidity of IgG antibodies to Porphyromonas gingivalis whole cells and a cell surface protein in subjects with adult periodontitis." *J Periodontal Res* **32**(1 Pt 1): 31-39.

Berglin, E., H. Kokkonen, E. Einarsdottir, A. Agren and S. Rantapaa Dahlqvist (2010). "Influence of female hormonal factors, in relation to autoantibodies and genetic markers, on the development of rheumatoid arthritis in northern Sweden: a case-control study." *Scand J Rheumatol*.

Berglundh, T. and M. Donati (2005). "Aspects of adaptive host response in periodontitis." *J Clin Periodontol* **32 Suppl 6**: 87-107.

Bergmann, S., M. Rohde, G. S. Chhatwal and S. Hammerschmidt (2001). "alpha-Enolase of Streptococcus pneumoniae is a plasmin(ogen)-binding protein displayed on the bacterial cell surface." *Mol Microbiol* **40**(6): 1273-1287.

Berthelot, J. M. and B. Le Goff (2010). "Rheumatoid arthritis and periodontal disease." *Joint Bone Spine* **77**(6): 537-541.

Bhattacharya, S. K., J. S. Crabb, V. L. Bonilha, X. Gu, H. Takahara and J. W. Crabb (2006). "Proteomics implicates peptidyl arginine deiminase 2 and optic nerve citrullination in glaucoma pathogenesis." *Invest Ophthalmol Vis Sci* **47**(6): 2508-2514.

Bhayani, H. R. and S. M. Hedrick (1991). "The role of polymorphic amino acids of the MHC molecule in the selection of the T cell repertoire." *J Immunol* **146**(4): 1093-1098.

Bonfil, J. J., F. L. Dillier, P. Mercier, D. Reviron, B. Foti, R. Sambuc, J. M. Brodeur and C. Sedarat (1999). "A "case control" study on the role of HLA DR4 in severe periodontitis and rapidly progressive periodontitis." *Journal of Clinical Periodontology* **26**(2): 77-84.

Booth, V. and T. Lehner (1997). "Characterization of the Porphyromonas gingivalis antigen recognized by a monoclonal antibody which prevents colonization by the organism." *J Periodontal Res* **32**(1 Pt 1): 54-60.

Borders, C. L., Jr., J. A. Broadwater, P. A. Bekeny, J. E. Salmon, A. S. Lee, A. M. Eldridge and V. B. Pett (1994). "A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens." *Protein Sci* **3**(4): 541-548.

Borrell, L. N., B. A. Burt and G. W. Taylor (2005). "Prevalence and Trends in Periodontitis in the USA: from the NHANES III to the NHANES, 1988 to 2000." *Journal of Dental Research* **84**(10): 924-930.

Bostrom, L., L. E. Linder and J. Bergstrom (1999). "Smoking and crevicular fluid levels of IL-6 and TNF-alpha in periodontal disease." *J Clin Periodontol* **26**(6): 352-357.

Brahn, E., C. Tang and M. L. Banquerigo (1994). "Regression of collagen-induced arthritis with taxol, a microtubule stabilizer." *Arthritis Rheum* **37**(6): 839-845.

Breslauer, K. J., R. Frank, H. Blocker and L. A. Marky (1986). "Predicting DNA duplex stability from the base sequence." *Proc Natl Acad Sci U S A* **83**(11): 3746-3750.

Burkhardt, H., B. Sehnert, R. Bockermann, A. Engstrom, J. R. Kalden and R. Holmdahl (2005). "Humoral immune response to citrullinated collagen type II determinants in early rheumatoid arthritis." *Eur J Immunol* **35**(5): 1643-1652.

Burr, M. L., H. Naseem, A. Hinks, S. Eyre, L. Gibbons, J. Bowes, B. Consortium, Y. Consortium, A. G. Wilson, J. Maxwell, A. W. Morgan, P. Emery, S. Steer, L. Hocking, D. M. Reid, B. P. Wordsworth, P. Harrison, W. Thomson, J. Worthington and A. Barton (2009). "PADI4 genotype is not associated with rheumatoid arthritis in a large UK Caucasian Population." *Ann Rheum Dis*.

Burrell, R. C. and J. D. Walters (2008). "Distribution of systemic clarithromycin to gingiva." *J Periodontol* **79**(9): 1712-1718.

Califano, J. V., R. E. Schifferle, J. C. Gunsolley, A. M. Best, H. A. Schenkein and J. G. Tew (1999). "Antibody reactive with Porphyromonas gingivalis serotypes K1-6 in adult and generalized early-onset periodontitis." *J Periodontol* **70**(7): 730-735.

Cantaert, T., P. Coucke, L. De Rycke, E. M. Veys, F. De Keyser and D. Baeten (2005). "Functional haplotypes of PADI4: relevance for rheumatoid arthritis specific synovial intracellular citrullinated proteins and anticitrullinated protein antibodies." *Ann Rheum Dis* **64**(9): 1316-1320.

Caponi, L., E. Petit-Teixeira, M. Sebbag, F. Bongiorni, S. Moscato, F. Pratesi, C. Pierlot, J. Osorio, S. Chapuy-Regaud, M. Guerrin, F. Cornelis, G. Serre and P. Migliorini (2005). "A family based study shows no association between rheumatoid arthritis and the PADI4 gene in a white French population." *Ann Rheum Dis* **64**(4): 587-593.

Carty, S. M., N. Snowden and A. J. Silman (2004). "Should infection still be considered as the most likely triggering factor for rheumatoid arthritis?" *Ann Rheum Dis* **63 Suppl 2**: ii46-ii49.

Cascao, R., H. S. Rosario, M. M. Souto-Carneiro and J. E. Fonseca (2010). "Neutrophils in rheumatoid arthritis: More than simple final effectors." *Autoimmun Rev* **9**(8): 531-535.

Chang, X. and J. Han (2006). "Expression of peptidylarginine deiminase type 4 (PAD4) in various tumors." *Mol Carcinog* **45**(3): 183-196.

Chang, X., R. Yamada, A. Suzuki, Y. Kochi, T. Sawada and K. Yamamoto (2005). "Citrullination of fibronectin in rheumatoid arthritis synovial tissue." *Rheumatology (Oxford)* **44**(11): 1374-1382.

Chapuy-Regaud, S., M. Sebbag, D. Baeten, C. Clavel, C. Foulquier, F. De Keyser and G. Serre (2005). "Fibrin deimination in synovial tissue is not specific for rheumatoid arthritis but commonly occurs during synovitides." *J. Immunol.* **174**(8): 5057-5064.

Cheng, Y. and W. H. Prusoff (1973). "Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction." *Biochem Pharmacol* **22**(23): 3099-3108.

Choi, J. I. and G. J. Seymour (2010). "Vaccines against periodontitis: a forward-looking review." *J Periodontal Implant Sci* **40**(4): 153-163.

Clavel, C., L. Nogueira, L. Laurent, C. Iobagiu, C. Vincent, M. Sebbag and G. Serre (2008). "Induction of macrophage secretion of tumor necrosis factor alpha through Fcgamma receptor IIa engagement by rheumatoid arthritis-specific autoantibodies to citrullinated proteins complexed with fibrinogen." *Arthritis Rheum.* **58**(3): 678-688.

Colombo, A. V., C. M. da Silva, A. Haffajee and A. P. Colombo (2007). "Identification of intracellular oral species within human crevicular epithelial cells from subjects with chronic periodontitis by fluorescence in situ hybridization." *J Periodontal Res* **42**(3): 236-243.

Curtis, M. A., J. Aduse-Opoku and M. Rangarajan (2001). "Cysteine proteases of Porphyromonas gingivalis." *Crit Rev Oral Biol Med* **12**(3): 192-216.
Curtis, M. A., H. K. Kuramitsu, M. Lantz, F. L. Macrina, K. Nakayama, J. Potempa, E. C. Reynolds and J. Aduse-Opoku (1999). "Molecular genetics and nomenclature of proteases of Porphyromonas gingivalis." *J Periodontal Res* **34**(8): 464-472.
Cuthbert, G. L., S. Daujat, A. W. Snowden, H. Erdjument-Bromage, T. Hagiwara, M. Yamada, R. Schneider, P. D. Gregory, P. Tempst, A. J. Bannister and T. Kouzarides (2004). "Histone deimination antagonizes arginine methylation." *Cell* **118**(5): 545-553.
Davies, J. M. (1997). "Molecular mimicry: can epitope mimicry induce autoimmune disease?" *Immunol Cell Biol* **75**(2): 113-126.
De Nardin, A. M., H. T. Sojar, S. G. Grossi, L. A. Christersson and R. J. Genco (1991). "Humoral immunity of older adults with periodontal disease to Porphyromonas gingivalis." *Infect Immun* **59**(12): 4363-4370.
de Pablo, P., I. L. C. Chapple, C. D. Buckley and T. Dietrich (2009). "Periodontitis in systemic rheumatic diseases." *Nat. Rev. Rheumatol.* **5**(4): 218-224.
de Pablo, P., T. Dietrich and T. E. McAlindon (2008). "Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population." *J. Rheumatol.* **35**(1): 70-76.
De Rycke, L., A. P. Nicholas, T. Cantaert, E. Kruithof, J. D. Echols, B. Vandekerckhove, E. M. Veys, F. De Keyser and D. Baeten (2005). "Synovial intracellular citrullinated proteins colocalizing with peptidyl arginine deiminase as pathophysiologically relevant antigenic determinants of rheumatoid arthritis-specific humoral autoimmunity." *Arthritis Rheum* **52**(8): 2323-2330.
dos Anjos, L. M., I. A. Pereira, E. d 'Orsi, A. P. Seaman, R. W. Burlingame and E. F. Morato (2009). "A comparative study of IgG second- and third-generation anti-cyclic citrullinated peptide (CCP) ELISAs and their combination with IgA third-generation CCP ELISA for the diagnosis of rheumatoid arthritis." *Clin Rheumatol* **28**(2): 153-158.
Dyer, J. K., M. A. Peck, R. A. Reinhardt, W. C. Duckworth, S. J. Booth, G. J. Seymour and K. D. Patil (1997). "HLA-D types and serum IgG responses to Capnocytophaga in diabetes and periodontitis." *J Dent Res* **76**(12): 1825-1832.
Dzink, J. L., S. S. Socransky and A. D. Haffajee (1988). "The predominant cultivable microbiota of active and inactive lesions of destructive periodontal diseases." *J Clin Periodontol* **15**(5): 316-323.
Ebersole, J. L., M. A. Taubman, D. J. Smith, D. E. Frey, A. D. Haffajee and S. S. Socransky (1987). "Human serum antibody responses to oral microorganisms. IV. Correlation with homologous infection." *Oral Microbiol Immunol* **2**(2): 53-59.
Elliott, M. J., R. N. Maini, M. Feldmann, A. Long-Fox, P. Charles, P. Katsikis, F. M. Brennan, J. Walker, H. Bijl, J. Ghrayeb and et al. (1993). "Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor alpha." *Arthritis Rheum* **36**(12): 1681-1690.
Feldmann, M. (2002). "Development of anti-TNF therapy for rheumatoid arthritis." *Nat Rev Immunol* **2**(5): 364-371.
Filip, C., G. Fletcher, J. L. Wulff and C. F. Earhart (1973). "Solubilization of the cytoplasmic membrane of Escherichia coli by the ionic detergent sodium-lauryl sarcosinate." *J Bacteriol* **115**(3): 717-722.
Firatli, E., A. Kantarci, I. Cebeci, H. Tanyeri, G. Sonmez, M. Carin and O. Tuncer (1996). "Association between HLA antigens and early onset periodontitis." *J Clin Periodontol* **23**(6): 563-566.

Firestein, G. S. (2003). "Evolving concepts of rheumatoid arthritis." *Nature* **423**(6937): 356-361.

Foulquier, C., M. Sebbag, C. Clavel, S. Chapuy-Regaud, R. Al Badine, M. C. Mechin, C. Vincent, R. Nachat, M. Yamada, H. Takahara, M. Simon, M. Guerrin and G. Serre (2007). "Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation." *Arthritis Rheum* **56**(11): 3541-3553.

Fox, D. and A. G. Smulian (2001). "Plasminogen-binding activity of enolase in the opportunistic pathogen Pneumocystis carinii." *Med Mycol* **39**(6): 495-507.

Gabler, W. L. and H. R. Creamer (1991). "Suppression of human neutrophil functions by tetracyclines." *J Periodontal Res* **26**(1): 52-58.

Gandjbakhch, F., I. Fajardy, B. Ferre, S. Dubucquoi, R. M. Flipo, N. Roger and E. Solau-Gervais (2009). "A functional haplotype of PADI4 gene in rheumatoid arthritis: positive correlation in a French population." *J Rheumatol* **36**(5): 881-886.

Ge, J., D. M. Catt and R. L. Gregory (2004). "Streptococcus mutans surface alpha-enolase binds salivary mucin MG2 and human plasminogen." *Infect Immun* **72**(11): 6748-6752.

Gemmell, E., R. I. Marshall and G. J. Seymour (1997). "Cytokines and prostaglandins in immune homeostasis and tissue destruction in periodontal disease." *Periodontol 2000* **14**: 112-143.

Gemmell, E., K. Yamazaki and G. J. Seymour (2007). "The role of T cells in periodontal disease: homeostasis and autoimmunity." *Periodontol 2000* **43**: 14-40.

Girbal-Neuhauser, E., J. J. Durieux, M. Arnaud, P. Dalbon, M. Sebbag, C. Vincent, M. Simon, T. Senshu, C. Masson-Bessiere, C. Jolivet-Reynaud, M. Jolivet and G. Serre (1999). "The epitopes targeted by the rheumatoid arthritis-associated antifilaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deimination of arginine residues." *J Immunol* **162**(1): 585-594.

Golub, L. M., H. M. Lee, M. E. Ryan, W. V. Giannobile, J. Payne and T. Sorsa (1998). "Tetracyclines inhibit connective tissue breakdown by multiple non-antimicrobial mechanisms." *Adv Dent Res* **12**(2): 12-26.

Gregersen, P. K., J. Silver and R. J. Winchester (1987). "The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis." *Arthritis Rheum* **30**(11): 1205-1213.

Griffen, A. L., M. R. Becker, S. R. Lyons, M. L. Moeschberger and E. J. Leys (1998). "Prevalence of Porphyromonas gingivalis and periodontal health status." *J Clin Microbiol* **36**(11): 3239-3242.

Grossi, S. G., R. J. Genco, E. E. Machtei, A. W. Ho, G. Koch, R. Dunford, J. J. Zambon and E. Hausmann (1995). "Assessment of risk for periodontal disease. II. Risk indicators for alveolar bone loss." *J Periodontol* **66**(1): 23-29.

Gunsolley, J. C., J. G. Tew, C. Gooss, D. R. Marshall, J. A. Burmeister and H. A. Schenkein (1990). "Serum antibodies to periodontal bacteria." *J Periodontol* **61**(7): 412-419.

Guo, Y., K. A. Nguyen and J. Potempa (2010). "Dichotomy of gingipains action as virulence factors: from cleaving substrates with the precision of a surgeon's knife to a meat chopper-like brutal degradation of proteins." *Periodontol 2000* **54**(1): 15-44.

Hagiwara, T., Y. Hidaka and M. Yamada (2005). "Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes." *Biochemistry* **44**(15): 5827-5834.

Harney, S. M., C. Meisel, A. M. Sims, P. Y. Woon, B. P. Wordsworth and M. A. Brown (2005). "Genetic and genomic studies of PADI4 in rheumatoid arthritis." *Rheumatology (Oxford)* **44**(7): 869-872.

Haurat, M. F., J. Aduse-Opoku, M. Rangarajan, L. Dorobantu, M. R. Gray, M. A. Curtis and M. F. Feldman (2010). "Selective sorting of cargo proteins into bacterial membrane vesicles." *J Biol Chem*.

Havemose-Poulsen, A., J. Westergaard, K. Stoltze, H. Skjodt, B. Danneskiold-Samsoe, H. Locht, K. Bendtzen and P. Holmstrup (2006). "Periodontal and hematological characteristics associated with aggressive periodontitis, juvenile idiopathic arthritis, and rheumatoid arthritis." *J. Periodontol.* **77**(2): 280-288.

Hayashi, H., M. Morioka, S. Ichimiya, K. Yamato, D. Hinode, A. Nagata and R. Nakamura (1993). "Participation of an arginyl residue of insulin chain B in the inhibition of hemagglutination by Porphyromonas gingivalis." *Oral Microbiol Immunol* **8**(6): 386-389.

Hayashi, H., A. Nagata, D. Hinode, M. Sato and R. Nakamura (1992). "Survey of a receptor protein in human erythrocytes for hemagglutinin of Porphyromonas gingivalis." *Oral Microbiol Immunol* **7**(4): 204-211.

Hendler, A., T. K. Mulli, F. J. Hughes, D. Perrett, M. Bombardieri, Y. Houri-Haddad, E. I. Weiss and A. Nissim (2010). "Involvement of autoimmunity in the pathogenesis of aggressive periodontitis." *J Dent Res* **89**(12): 1389-1394.

Hermansson, M., K. Artemenko, E. Ossipova, H. Eriksson, J. Lengqvist, D. Makrygiannakis, A. I. Catrina, A. P. Nicholas, L. Klareskog, M. Savitski, R. A. Zubarev and P. J. Jakobsson (2010). "MS analysis of rheumatoid arthritic synovial tissue identifies specific citrullination sites on fibrinogen." *Proteomics Clin Appl* **4**(5): 511-518.

Hill, J. A., S. Southwood, A. Sette, A. M. Jevnikar, D. A. Bell and E. Cairns (2003). "Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule." *J Immunol* **171**(2): 538-541.

Hitchon, C. A., F. Chandad, E. D. Ferucci, A. Willemze, A. Ioan-Facsinay, D. van der Woude, J. Markland, D. Robinson, B. Elias, M. Newkirk, R. M. Toes, T. W. Huizinga and H. S. El-Gabalawy (2010). "Antibodies to porphyromonas gingivalis are associated with anticitrullinated protein antibodies in patients with rheumatoid arthritis and their relatives." *J Rheumatol* **37**(6): 1105-1112.

Hoet, R. M., A. M. Boerbooms, M. Arends, D. J. Ruiter and W. J. van Venrooij (1991). "Antiperinuclear factor, a marker autoantibody for rheumatoid arthritis: colocalisation of the perinuclear factor and profilaggrin." *Ann Rheum Dis* **50**(9): 611-618.

Holoshitz, J. (2010). "The rheumatoid arthritis HLA-DRB1 shared epitope." *Curr Opin Rheumatol* **22**(3): 293-298.

Holt, S. C., J. Ebersole, J. Felton, M. Brunsvold and K. S. Kornman (1988). "Implantation of Bacteroides gingivalis in nonhuman primates initiates progression of periodontitis." *Science* **239**(4835): 55-57.

Holt, S. C., L. Kesavalu, S. Walker and C. A. Genco (1999). "Virulence factors of Porphyromonas gingivalis." *Periodontol 2000* **20**: 168-238.

Hoppe, B., T. Haupl, K. Egerer, R. Gruber, H. Kiesewetter, A. Salama, G. R. Burmester and T. Dorner (2009). "Influence of peptidylarginine deiminase type 4 genotype and shared epitope on clinical characteristics and autoantibody profile of rheumatoid arthritis." *Ann Rheum Dis* **68**(6): 898-903.

Horikoshi, N., H. Tachiwana, K. Saito, A. Osakabe, M. Sato, M. Yamada, S. Akashi, Y. Nishimura, W. Kagawa and H. Kurumizaka (2011). "Structural and biochemical analyses of the human PAD4 variant encoded by a functional haplotype gene." *Acta Crystallogr D Biol Crystallogr* **67**(Pt 2): 112-118.

Hujoel, P. P., M. Drangsholt, C. Spiekerman and T. A. DeRouen (2002). "Periodontitis-systemic disease associations in the presence of smoking--causal or coincidental?" *Periodontol 2000* **30**: 51-60.

Ikari, K., M. Kuwahara, T. Nakamura, S. Momohara, M. Hara, H. Yamanaka, T. Tomatsu and N. Kamatani (2005). "Association between PADI4 and rheumatoid arthritis: a replication study." *Arthritis Rheum* **52**(10): 3054-3057.

Imamura, T., K. Matsushita, J. Travis and J. Potempa (2001). "Inhibition of trypsin-like cysteine proteinases (gingipains) from Porphyromonas gingivalis by tetracycline and its analogues." *Antimicrob Agents Chemother* **45**(10): 2871-2876.

Imamura, T., R. N. Pike, J. Potempa and J. Travis (1994). "Pathogenesis of periodontitis: a major arginine-specific cysteine proteinase from Porphyromonas gingivalis induces vascular permeability enhancement through activation of the kallikrein/kinin pathway." *J Clin Invest* **94**(1): 361-367.

Imamura, T., J. Potempa, R. N. Pike, J. N. Moore, M. H. Barton and J. Travis (1995). "Effect of free and vesicle-bound cysteine proteinases of Porphyromonas gingivalis on plasma clot formation: implications for bleeding tendency at periodontitis sites." *Infect Immun* **63**(12): 4877-4882.

Inagaki, M., H. Takahara, Y. Nishi, K. Sugawara and C. Sato (1989). "Ca2+-dependent deimination-induced disassembly of intermediate filaments involves specific modification of the amino-terminal head domain." *J Biol Chem* **264**(30): 18119-18127.

Ishida-Yamamoto, A., T. Senshu, H. Takahashi, K. Akiyama, K. Nomura and H. Iizuka (2000). "Decreased deiminated keratin K1 in psoriatic hyperproliferative epidermis." *J Invest Dermatol* **114**(4): 701-705.

Ishigami, A., T. Ohsawa, M. Hiratsuka, H. Taguchi, S. Kobayashi, Y. Saito, S. Murayama, H. Asaga, T. Toda, N. Kimura and N. Maruyama (2005). "Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease." *J Neurosci Res* **80**(1): 120-128.

Iwamoto, T., K. Ikari, T. Nakamura, M. Kuwahara, Y. Toyama, T. Tomatsu, S. Momohara and N. Kamatani (2006). "Association between PADI4 and rheumatoid arthritis: a meta-analysis." *Rheumatology (Oxford)* **45**(7): 804-807.

Jagels, M. A., J. Travis, J. Potempa, R. Pike and T. E. Hugli (1996). "Proteolytic inactivation of the leukocyte C5a receptor by proteinases derived from Porphyromonas gingivalis." *Infect Immun* **64**(6): 1984-1991.

Jang, B., J. K. Jin, Y. C. Jeon, H. J. Cho, A. Ishigami, K. C. Choi, R. I. Carp, N. Maruyama, Y. S. Kim and E. K. Choi (2010). "Involvement of peptidylarginine deiminase-mediated post-translational citrullination in pathogenesis of sporadic Creutzfeldt-Jakob disease." *Acta Neuropathol* **119**(2): 199-210.

Jolodar, A., P. Fischer, S. Bergmann, D. W. Buttner, S. Hammerschmidt and N. W. Brattig (2003). "Molecular cloning of an alpha-enolase from the human filarial parasite Onchocerca volvulus that binds human plasminogen." *Biochim Biophys Acta* **1627**(2-3): 111-120.

Jones, J. E., C. P. Causey, L. Lovelace, B. Knuckley, H. Flick, L. Lebioda and P. R. Thompson (2010a). "Characterization and inactivation of an agmatine deiminase from Helicobacter pylori." *Bioorg Chem* **38**(2): 62-73.

Jones, J. E., C. J. Dreyton, H. Flick, C. P. Causey and P. R. Thompson (2010b). "Mechanistic studies of agmatine deiminase from multiple bacterial species." *Biochemistry* **49**(43): 9413-9423.

Kallberg, H., L. Padyukov, R. M. Plenge, J. Ronnelid, P. K. Gregersen, A. H. van der Helm-van Mil, R. E. Toes, T. W. Huizinga, L. Klareskog and L. Alfredsson (2007). "Gene-

gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis." *Am J Hum Genet* **80**(5): 867-875.
Kamma, J. J., M. Nakou and F. J. Mitsis (2000). "The clinical and microbiological effects of systemic ornidazole in sites with and without subgingival debridement in early-onset periodontitis patients." *J Periodontol* **71**(12): 1862-1873.
Kang, C. P., H. S. Lee, H. Ju, H. Cho, C. Kang and S. C. Bae (2006). "A functional haplotype of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Koreans." *Arthritis Rheum* **54**(1): 90-96.
Kang, H. J., S. K. Jung, S. J. Kim and S. J. Chung (2008). "Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme." *Acta Crystallogr D Biol Crystallogr* **64**(Pt 6): 651-657.
Kapatral, V., I. Anderson, N. Ivanova, G. Reznik, T. Los, A. Lykidis, A. Bhattacharyya, A. Bartman, W. Gardner, G. Grechkin, L. Zhu, O. Vasieva, L. Chu, Y. Kogan, O. Chaga, E. Goltsman, A. Bernal, N. Larsen, M. D'Souza, T. Walunas, G. Pusch, R. Haselkorn, M. Fonstein, N. Kyrpides and R. Overbeek (2002). "Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586." *J Bacteriol* **184**(7): 2005-2018.
Katz, J., J. Goultschin, R. Benoliel and C. Brautbar (1987). "Human leukocyte antigen (HLA) DR4. Positive association with rapidly progressing periodontitis." *J Periodontol* **58**(9): 607-610.
Kim, E. Y. and K. D. Moudgil (2009). "The determinants of susceptibility/resistance to adjuvant arthritis in rats." *Arthritis Res Ther* **11**(4): 239.
Kinane, D. F. and I. G. Chestnutt (2000). "Smoking and periodontal disease." *Crit Rev Oral Biol Med* **11**(3): 356-365.
Kinane, D. F., J. Mooney and J. L. Ebersole (1999). "Humoral immune response to Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis in periodontal disease." *Periodontol 2000* **20**: 289-340.
Kinane, D. F., H. Shiba and T. C. Hart (2005). "The genetic basis of periodontitis." *Periodontol 2000* **39**: 91-117.
Kinloch, A., K. Lundberg, R. Wait, N. Wegner, N. H. Lim, A. J. Zendman, T. Saxne, V. Malmstrom and P. J. Venables (2008). "Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis." *Arthritis Rheum* **58**(8): 2287-2295.
Kinloch, A., V. Tatzer, R. Wait, D. Peston, K. Lundberg, P. Donatien, D. Moyes, P. C. Taylor and P. J. Venables (2005). "Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis." *Arthritis Res Ther* **7**(6): R1421-1429.
Kinne, R. W., B. Stuhlmuller and G. R. Burmester (2007). "Cells of the synovium in rheumatoid arthritis. Macrophages." *Arthritis Res Ther* **9**(6): 224.
Kizawa, K., H. Takahara, H. Troxler, P. Kleinert, U. Mochida and C. W. Heizmann (2008). "Specific citrullination causes assembly of a globular S100A3 homotetramer: a putative Ca2+ modulator matures human hair cuticle." *J Biol Chem* **283**(8): 5004-5013.
Kleinfelder, J. W., R. F. Mueller and D. E. Lange (2000). "Fluoroquinolones in the treatment of Actinobacillus actinomycetemcomitans-associated periodontitis." *J Periodontol* **71**(2): 202-208.
Knipp, M. and M. Vasak (2000). "A Colorimetric 96-Well Microtiter Plate Assay for the Determination of Enzymatically Formed Citrulline." *Analytical Biochemistry* **286**: 257-264.
Knuckley, B., C. P. Causey, J. E. Jones, M. Bhatia, C. J. Dreyton, T. C. Osborne, H. Takahara and P. R. Thompson (2010a). "Substrate specificity and kinetic studies of

PADs 1, 3, and 4 identify potent and selective inhibitors of protein arginine deiminase 3." *Biochemistry* **49**(23): 4852-4863.
Knuckley, B., C. P. Causey, P. J. Pellechia, P. F. Cook and P. R. Thompson (2010b). "Haloacetamidine-based inactivators of protein arginine deiminase 4 (PAD4): evidence that general acid catalysis promotes efficient inactivation." *Chembiochem* **11**(2): 161-165.
Knuckley, B., Y. Luo and P. R. Thompson (2008). "Profiling Protein Arginine Deiminase 4 (PAD4): a novel screen to identify PAD4 inhibitors." *Bioorg Med Chem* **16**(2): 739-745.
Kontani, M., S. Kimura, I. Nakagawa and S. Hamada (1997). "Adherence of Porphyromonas gingivalis to matrix proteins via a fimbrial cryptic receptor exposed by its own arginine-specific protease." *Mol Microbiol* **24**(6): 1179-1187.
Kuhn, K. A., L. Kulik, B. Tomooka, K. J. Braschler, W. P. Arend, W. H. Robinson and V. M. Holers (2006). "Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis." *J Clin Invest* **116**(4): 961-973.
Kuhns, D. B., D. A. Priel and J. I. Gallin (2007). "Induction of human monocyte interleukin (IL)-8 by fibrinogen through the toll-like receptor pathway." *Inflammation* **30**(5): 178-188.
Kumar, P. S., E. J. Leys, J. M. Bryk, F. J. Martinez, M. L. Moeschberger and A. L. Griffen (2006). "Changes in periodontal health status are associated with bacterial community shifts as assessed by quantitative 16S cloning and sequencing." *J Clin Microbiol* **44**(10): 3665-3673.
Kurokawa, T., S. Hara, H. Takahara, K. Sugawara and T. Ikenaka (1987). "Conversion of peanut trypsin-chymotrypsin inhibitor B-III to a chymotrypsin inhibitor by deimination of the P1 arginine residues in two reactive sites." *J Biochem* **101**(6): 1361-1367.
Lamont, R. J. and H. F. Jenkinson (1998). "Life below the gum line: Pathogenic mechanisms of Porphyromonas gingivalis." *Microbiol Mol Biol Rev* **62**(4): 1244-1263.
Lamont, R. J. and H. F. Jenkinson (2000). "Subgingival colonization by Porphyromonas gingivalis." *Oral Microbiol Immunol* **15**(6): 341-349.
LaVallie, E. R., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel and J. M. McCoy (1993). "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm." *Biotechnology (N Y)* **11**(2): 187-193.
Lee, Y. H., Y. H. Rho, S. J. Choi, J. D. Ji and G. G. Song (2007). "PADI4 polymorphisms and rheumatoid arthritis susceptibility: a meta-analysis." *Rheumatol Int* **27**(9): 827-833.
Li, P., M. Li, M. R. Lindberg, M. J. Kennett, N. Xiong and Y. Wang (2010a). "PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps." *J Exp Med* **207**(9): 1853-1862.
Li, P., D. Wang, H. Yao, P. Doret, G. Hao, Q. Shen, H. Qiu, X. Zhang, Y. Wang and G. Chen (2010b). "Coordination of PAD4 and HDAC2 in the regulation of p53-target gene expression." *Oncogene* **29**(21): 3153-3162.
Li, P., H. Yao, Z. Zhang, M. Li, Y. Luo, P. R. Thompson, D. S. Gilmour and Y. Wang (2008). "Regulation of p53 target gene expression by peptidylarginine deiminase 4." *Mol Cell Biol* **28**(15): 4745-4758.
Liao, F., Z. Li, Y. Wang, B. Shi, Z. Gong and X. Cheng (2009). "Porphyromonas gingivalis may play an important role in the pathogenesis of periodontitis-associated rheumatoid arthritis." *Med Hypotheses* **72**(6): 732-735.
Lipsky, P. E., D. M. van der Heijde, E. W. St Clair, D. E. Furst, F. C. Breedveld, J. R. Kalden, J. S. Smolen, M. Weisman, P. Emery, M. Feldmann, G. R. Harriman and R. N. Maini

(2000). "Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group." *N Engl J Med* **343**(22): 1594-1602.

Loos, T., A. Mortier, M. Gouwy, I. Ronsse, W. Put, J. P. Lenaerts, J. Van Damme and P. Proost (2008). "Citrullination of CXCL10 and CXCL11 by peptidylarginine deiminase: a naturally occurring posttranslational modification of chemokines and new dimension of immunoregulation." *Blood* **112**(7): 2648-2656.

Lopez-Alemany, R., C. Longstaff, S. Hawley, M. Mirshahi, P. Fabregas, M. Jardi, E. Merton, L. A. Miles and J. Felez (2003). "Inhibition of cell surface mediated plasminogen activation by a monoclonal antibody against alpha-Enolase." *Am J Hematol* **72**(4): 234-242.

Lovegrove, J. M. (2004). "Dental plaque revisited: bacteria associated with periodontal disease." *J N Z Soc Periodontol*(87): 7-21.

Lundberg, K. (2009). "The pathobiology underpinning autoantibodies to citrullinated protein antigens [abstract SP0064]." *Ann. Rheum. Dis.* **68**(21): Suppl. 3.

Lundberg, K., A. Kinloch, B. A. Fisher, N. Wegner, R. Wait, P. Charles, T. R. Mikuls and P. J. Venables (2008). "Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase." *Arthritis Rheum* **58**(10): 3009-3019.

Lundberg, K., N. Wegner, T. Yucel-Lindberg and P. J. Venables (2010). "Periodontitis in RA-the citrullinated enolase connection." *Nat Rev Rheumatol*.

Lundy, S. K., S. Sarkar, L. A. Tesmer and D. A. Fox (2007). "Cells of the synovium in rheumatoid arthritis. T lymphocytes." *Arthritis Res Ther* **9**(1): 202.

Luo, Y., K. Arita, M. Bhatia, B. Knuckley, Y. H. Lee, M. R. Stallcup, M. Sato and P. R. Thompson (2006a). "Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization." *Biochemistry* **45**(39): 11727-11736.

Luo, Y., B. Knuckley, Y. H. Lee, M. R. Stallcup and P. R. Thompson (2006b). "A fluoroacetamidine-based inactivator of protein arginine deiminase 4: design, synthesis, and in vitro and in vivo evaluation." *J Am Chem Soc* **128**(4): 1092-1093.

Luross, J. A. and N. A. Williams (2001). "The genetic and immunopathological processes underlying collagen-induced arthritis." *Immunology* **103**(4): 407-416.

Lutzky, V., S. Hannawi and R. Thomas (2007). "Cells of the synovium in rheumatoid arthritis. Dendritic cells." *Arthritis Res Ther* **9**(4): 219.

Machold, K. P. (2010). "Prevention and cure of rheumatoid arthritis: is it possible?" *Best Pract Res Clin Rheumatol* **24**(3): 353-361.

Mahdi, H., B. A. Fisher, H. Kallberg, D. Plant, V. Malmstrom, J. Ronnelid, P. Charles, B. Ding, L. Alfredsson, L. Padyukov, D. P. Symmons, P. J. Venables, L. Klareskog and K. Lundberg (2009). "Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis." *Nat Genet* **41**(12): 1319-1324.

Makrygiannakis, D., E. af Klint, I. E. Lundberg, R. Lofberg, A. K. Ulfgren, L. Klareskog and A. I. Catrina (2006). "Citrullination is an inflammation-dependent process." *Ann. Rheum. Dis.* **65**(9): 1219-1222.

Mangat, P., N. Wegner, P. J. Venables and J. Potempa (2010). "Bacterial and human peptidylarginine deiminases: targets for inhibiting the autoimmune response in rheumatoid arthritis?" *Arthritis Res Ther* **12**(3): 209.

Marotte, H., P. Farge, P. Gaudin, C. Alexandre, B. Mougin and P. Miossec (2006). "The association between periodontal disease and joint destruction in rheumatoid arthritis extends the link between the HLA-DR shared epitope and severity of bone destruction." *Ann Rheum Dis* **65**(7): 905-909.

Martinez-Martinez, R. E., C. Abud-Mendoza, N. Patino-Marin, J. C. Rizo-Rodriguez, J. W. Little and J. P. Loyola-Rodriguez (2009). "Detection of periodontal bacterial DNA in serum and synovial fluid in refractory rheumatoid arthritis patients." *J Clin Periodontol* **36**(12): 1004-1010.

Martinez, A., A. Valdivia, D. Pascual-Salcedo, J. R. Lamas, M. Fernandez-Arquero, A. Balsa, B. Fernandez-Gutierrez, E. G. de la Concha and E. Urcelay (2005). "PADI4 polymorphisms are not associated with rheumatoid arthritis in the Spanish population." *Rheumatology (Oxford)* **44**(10): 1263-1266.

Masson-Bessiere, C., M. Sebbag, E. Girbal-Neuhauser, L. Nogueira, C. Vincent, T. Senshu and G. Serre (2001). "The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin." *J. Immunol.* **166**(6): 4177-4184.

Masuda, K., M. Yoshioka, D. Hinode and R. Nakamura (2002). "Purification and Characterization of Arginine Carboxypeptidase Produced by Porphyromonas gingivalis." *Infect. Immun.* **70**(4): 1807-1815.

Matsuo, K., Y. Xiang, H. Nakamura, K. Masuko, K. Yudoh, K. Noyori, K. Nishioka, T. Saito and T. Kato (2006). "Identification of novel citrullinated autoantigens of synovium in rheumatoid arthritis using a proteomic approach." *Arthritis Res Ther* **8**(6): R175.

Mauri, C. and M. R. Ehrenstein (2007). "Cells of the synovium in rheumatoid arthritis. B cells." *Arthritis Res Ther* **9**(2): 205.

McGraw, W. T., J. Potempa, D. Farley and J. Travis (1999). "Purification, characterization, and sequence analysis of a potential virulence factor from Porphyromonas gingivalis, peptidylarginine deiminase." *Infect. Immun.* **67**(7): 3248-3256.

McInnes, I. B. and G. Schett (2007). "Cytokines in the pathogenesis of rheumatoid arthritis." *Nat Rev Immunol* **7**(6): 429-442.

Mercado, F., R. I. Marshall, A. C. Klestov and P. M. Bartold (2000). "Is there a relationship between rheumatoid arthritis and periodontal disease?" *Journal of Clinical Periodontology* **27**(4): 267-272.

Mercado, F. B., R. I. Marshall and P. M. Bartold (2003). "Inter-relationships between rheumatoid arthritis and periodontal disease. A review." *Journal of Clinical Periodontology* **30**(9): 761-772.

Mercado, F. B., R. I. Marshall, A. C. Klestov and P. M. Bartold (2001). "Relationship between rheumatoid arthritis and periodontitis." *J Periodontol* **72**(6): 779-787.

Michalowicz, B. S., S. R. Diehl, J. C. Gunsolley, B. S. Sparks, C. N. Brooks, T. E. Koertge, J. V. Califano, J. A. Burmeister and H. A. Schenkein (2000). "Evidence of a substantial genetic basis for risk of adult periodontitis." *J Periodontol* **71**(11): 1699-1707.

Michou, L., V. H. Teixeira, C. Pierlot, S. Lasbleiz, T. Bardin, P. Dieude, B. Prum, F. Cornelis and E. Petit-Teixeira (2008). "Associations between genetic factors, tobacco smoking and autoantibodies in familial and sporadic rheumatoid arthritis." *Ann Rheum Dis* **67**(4): 466-470.

Mikuls, T. R., J. B. Payne, R. A. Reinhardt, G. M. Thiele, E. Maziarz, A. C. Cannella, V. M. Holers, K. A. Kuhn and J. R. O'Dell (2009). "Antibody responses to Porphyromonas gingivalis (P. gingivalis) in subjects with rheumatoid arthritis and periodontitis." *Int. Immunopharmacol.* **9**(1): 38-42.

Miles, L. A., C. M. Dahlberg, J. Plescia, J. Felez, K. Kato and E. F. Plow (1991). "Role of cell-surface lysines in plasminogen binding to cells: identification of alpha-enolase as a candidate plasminogen receptor." *Biochemistry* **30**(6): 1682-1691.

Moen, K., J. G. Brun, T. M. Madland, T. Tynning and R. Jonsson (2003). "Immunoglobulin G and A Antibody Responses to Bacteroides forsythus and Prevotella intermedia in

Sera and Synovial Fluids of Arthritis Patients." *Clin. Diagn. Lab. Immunol.* **10**(6): 1043-1050.

Moen, K., J. G. Brun, M. Valen, L. Skartveit, E. K. R. Eribe, I. Olsen and R. Jonsson (2006). "Synovial inflammation in active rheumatoid arthritis and psoriatic arthritis facilitates trapping of a variety of oral bacterial DNAs." *Clinical and Experimental Rheumatology* **24**(6): 656-663.

Molitor, J. A., A. Alonso, M. H. Wener, B. S. Michalowicz, J. D. Beck and A. R. Folsom (2009). "Moderate to severe adult periodontitis is associated with increased risk of seropositive rheumatoid arthritis in non-smokers: the ARIC study." *Ann Rheum Dis* **68**(Suppl3): 399 (Abstract FRI0129).

Monteiro, A. C., A. Scovino, S. Raposo, V. M. Gaze, C. Cruz, E. Svensjo, M. S. Narciso, A. P. Colombo, J. B. Pesquero, E. Feres-Filho, K. A. Nguyen, A. Sroka, J. Potempa and J. Scharfstein (2009). "Kinin danger signals proteolytically released by gingipain induce Fimbriae-specific IFN-gamma- and IL-17-producing T cells in mice infected intramucosally with Porphyromonas gingivalis." *J Immunol* **183**(6): 3700-3711.

Mooney, J., E. Adonogianaki and D. F. Kinane (1993). "Relative avidity of serum antibodies to putative periodontopathogens in periodontal disease." *J Periodontal Res* **28**(6 Pt 1): 444-450.

Mooney, J., P. J. Hodge and D. F. Kinane (2001). "Humoral immune response in early-onset periodontitis: influence of smoking." *J Periodontal Res* **36**(4): 227-232.

Moscarello, M. A., D. D. Wood, C. Ackerley and C. Boulias (1994). "Myelin in multiple sclerosis is developmentally immature." *J Clin Invest* **94**(1): 146-154.

Muller-Ladner, U., C. Ospelt, S. Gay, O. Distler and T. Pap (2007). "Cells of the synovium in rheumatoid arthritis. Synovial fibroblasts." *Arthritis Res Ther* **9**(6): 223.

Mundodi, V., A. S. Kucknoor and J. F. Alderete (2008). "Immunogenic and plasminogen-binding surface-associated alpha-enolase of Trichomonas vaginalis." *Infect Immun* **76**(2): 523-531.

Murakami, Y., H. Tamagawa, S. Shizukuishi, A. Tsunemitsu and S. Aimoto (1992). "Biological role of an arginine residue present in a histidine-rich peptide which inhibits hemagglutination of Porphyromonas gingivalis." *FEMS Microbiol Lett* **77**(1-3): 201-204.

Naito, M., H. Hirakawa, A. Yamashita, N. Ohara, M. Shoji, H. Yukitake, K. Nakayama, H. Toh, F. Yoshimura, S. Kuhara, M. Hattori and T. Hayashi (2008). "Determination of the genome sequence of Porphyromonas gingivalis strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in P. gingivalis." *DNA Res* **15**(4): 215-225.

Naito, Y., K. Okuda and I. Takazoe (1987). "Detection of specific antibody in adult human periodontitis sera to surface antigens of Bacteroides gingivalis." *Infect Immun* **55**(3): 832-834.

Nakada, Y. and Y. Itoh (2003). "Identification of the putrescine biosynthetic genes in Pseudomonas aeruginosa and characterization of agmatine deiminase and N-carbamoylputrescine amidohydrolase of the arginine decarboxylase pathway." *Microbiology* **149**(Pt 3): 707-714.

Nakayama-Hamada, M., A. Suzuki, H. Furukawa, R. Yamada and K. Yamamoto (2008). "Citrullinated Fibrinogen Inhibits Thrombin-catalysed Fibrin Polymerization." *J Biochem* **144**(3): 393-398.

Nakayama-Hamada, M., A. Suzuki, K. Kubota, T. Takazawa, M. Ohsaka, R. Kawaida, M. Ono, A. Kasuya, H. Furukawa, R. Yamada and K. Yamamoto (2005). "Comparison of enzymatic properties between hPADI2 and hPADI4." *Biochem Biophys Res Commun* **327**(1): 192-200.

Neeli, I., S. N. Khan and M. Radic (2008). "Histone deimination as a response to inflammatory stimuli in neutrophils." *J Immunol* **180**(3): 1895-1902.

Nelson, D., J. Potempa, T. Kordula and J. Travis (1999). "Purification and characterization of a novel cysteine proteinase (periodontain) from Porphyromonas gingivalis. Evidence for a role in the inactivation of human alpha1-proteinase inhibitor." *J Biol Chem* **274**(18): 12245-12251.

Nelson, K. E., R. D. Fleischmann, R. T. DeBoy, I. T. Paulsen, D. E. Fouts, J. A. Eisen, S. C. Daugherty, R. J. Dodson, A. S. Durkin, M. Gwinn, D. H. Haft, J. F. Kolonay, W. C. Nelson, T. Mason, L. Tallon, J. Gray, D. Granger, H. Tettelin, H. Dong, J. L. Galvin, M. J. Duncan, F. E. Dewhirst and C. M. Fraser (2003). "Complete genome sequence of the oral pathogenic Bacterium porphyromonas gingivalis strain W83." *J Bacteriol* **185**(18): 5591-5601.

Nesse, W., P. U. Dijkstra, F. Abbas, F. K. Spijkervet, A. Stijger, J. A. Tromp, J. L. van Dijk and A. Vissink (2010). "Increased prevalence of cardiovascular and autoimmune diseases in periodontitis patients: a cross-sectional study." *J Periodontol* **81**(11): 1622-1628.

Nesse, W., J. Westra, J. E. Van Der Wal, J. Balsma, F. Abbas, E. Brouwer and A. Vissink (2009). "The periodontium contains citrullinated proteins, PAD-2 enzymes and HC Gp-39." *Arthritis & Rheum* **60**(Suppl 10): 1165.

Nguyen, K.-A., J. Travis and J. Potempa (2007). "Does the importance of the C-terminal residues in the maturation of RgpB from Porphyromonas gingivalis reveal a novel mechanism for protein export in a subgroup of gram-negative bacteria?" *J. Bacteriol.* **189**(3): 833-843.

Niederman, R., B. Brunkhorst, S. Smith, R. N. Weinreb and M. I. Ryder (1990). "Ammonia as a potential mediator of adult human periodontal infection: inhibition of neutrophil function." *Arch Oral Biol* **35 Suppl**: 205S-209S.

Nienhuis, R. L. and E. Mandema (1964). "A New Serum Factor in Patients with Rheumatoid Arthritis; the Antiperinuclear Factor." *Ann Rheum Dis* **23**: 302-305.

O'Brien-Simpson, N. M., R. D. Pathirana, G. D. Walker and E. C. Reynolds (2009). "Porphyromonas gingivalis RgpA-Kgp proteinase-adhesin complexes penetrate gingival tissue and induce proinflammatory cytokines or apoptosis in a concentration-dependent manner." *Infect Immun* **77**(3): 1246-1261.

O'Dell, J. R. (2004). "Therapeutic strategies for rheumatoid arthritis." *N Engl J Med* **350**(25): 2591-2602.

O'Dell, J. R., K. W. Blakely, J. A. Mallek, P. J. Eckhoff, R. D. Leff, S. J. Wees, K. M. Sems, A. M. Fernandez, W. R. Palmer, L. W. Klassen, G. A. Paulsen, C. E. Haire and G. F. Moore (2001). "Treatment of early seropositive rheumatoid arthritis: a two-year, double-blind comparison of minocycline and hydroxychloroquine." *Arthritis Rheum* **44**(10): 2235-2241.

O'Dell, J. R., G. Paulsen, C. E. Haire, K. Blakely, W. Palmer, S. Wees, P. J. Eckhoff, L. W. Klassen, M. Churchill, D. Doud, A. Weaver and G. F. Moore (1999). "Treatment of early seropositive rheumatoid arthritis with minocycline: four-year followup of a double-blind, placebo-controlled trial." *Arthritis Rheum* **42**(8): 1691-1695.

Ogrendik, M. (2007a). "Effects of clarithromycin in patients with active rheumatoid arthritis." *Curr Med Res Opin* **23**(3): 515-522.

Ogrendik, M. (2007b). "Levofloxacin treatment in patients with rheumatoid arthritis receiving methotrexate." *South Med J* **100**(2): 135-139.

Ogrendik, M., A. Hakguder and N. Keser (2006). "Treatment of rheumatoid arthritis with ornidazole. A randomized, double-blind, placebo-controlled study." *Rheumatology (Oxford)* **45**(5): 636-637.

Ogrendik, M., S. Kokino, F. Ozdemir, P. S. Bird and S. Hamlet (2005). "Serum antibodies to oral anaerobic bacteria in patients with rheumatoid arthritis." *MedGenMed* **7**(2): 2.
Ohlrich, E. J., M. P. Cullinan and G. J. Seymour (2009). "The immunopathogenesis of periodontal disease." *Aust Dent J* **54 Suppl 1**: S2-10.
Okumura, N., A. Haneishi and F. Terasawa (2009). "Citrullinated fibrinogen shows defects in FPA and FPB release and fibrin polymerization catalyzed by thrombin." *Clinica Chimica Acta* **401**(1-2): 119-123.
Olczak, T., K. Siudeja and M. Olczak (2006). "Purification and initial characterization of a novel Porphyromonas gingivalis HmuY protein expressed in Escherichia coli and insect cells." *Protein Expr Purif* **49**(2): 299-306.
Oldstone, M. B. (1998). "Molecular mimicry and immune-mediated diseases." *FASEB J* **12**(13): 1255-1265.
Orlov, S. N., R. Grygorczyk and S. V. Kotelevtsev (2003). "Do we know the absolute values of intracellular free calcium concentration?" *Cell Calcium* **34**(6): 511-515.
Ortiz, P., N. F. Bissada, L. Palomo, Y. W. Han, M. S. Al-Zahrani, A. Panneerselvam and A. Askari (2009). "Periodontal therapy reduces the severity of active rheumatoid arthritis in patients treated with or without tumor necrosis factor inhibitors." *J Periodontol* **80**(4): 535-540.
Otero, M. and M. B. Goldring (2007). "Cells of the synovium in rheumatoid arthritis. Chondrocytes." *Arthritis Res Ther* **9**(5): 220.
Pancholi, V. and V. A. Fischetti (1998). "alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci." *J Biol Chem* **273**(23): 14503-14515.
Parker, A. C. and C. J. Smith (1993). "Genetic and biochemical analysis of a novel Ambler class A beta-lactamase responsible for cefoxitin resistance in Bacteroides species." *Antimicrob Agents Chemother* **37**(5): 1028-1036.
Paster, B. J., S. K. Boches, J. L. Galvin, R. E. Ericson, C. N. Lau, V. A. Levanos, A. Sahasrabudhe and F. E. Dewhirst (2001). "Bacterial diversity in human subgingival plaque." *J Bacteriol* **183**(12): 3770-3783.
Pathirana, R. D., N. M. O'Brien-Simpson and E. C. Reynolds (2010). "Host immune responses to Porphyromonas gingivalis antigens." *Periodontol 2000* **52**(1): 218-237.
Persson, G. R. (2005). "Immune responses and vaccination against periodontal infections." *J Clin Periodontol* **32 Suppl 6**: 39-53.
Petkova, S. B., K. N. Konstantinov, T. J. Sproule, B. L. Lyons, M. A. Awwami and D. C. Roopenian (2006). "Human antibodies induce arthritis in mice deficient in the low-affinity inhibitory IgG receptor Fc gamma RIIB." *J Exp Med* **203**(2): 275-280.
Pischon, N., T. Pischon, J. Kroger, E. Gulmez, B. M. Kleber, J. P. Bernimoulin, H. Landau, P. G. Brinkmann, P. Schlattmann, J. Zernicke, F. Buttgereit and J. Detert (2008). "Association among rheumatoid arthritis, oral hygiene, and periodontitis." *J Periodontol* **79**(6): 979-986.
Plenge, R. M., L. Padyukov, E. F. Remmers, S. Purcell, A. T. Lee, E. W. Karlson, F. Wolfe, D. L. Kastner, L. Alfredsson, D. Altshuler, P. K. Gregersen, L. Klareskog and J. D. Rioux (2005). "Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4." *Am J Hum Genet* **77**(6): 1044-1060.
Popadiak, K., J. Potempa, K. Riesbeck and A. M. Blom (2007). "Biphasic effect of gingipains from Porphyromonas gingivalis on the human complement system." *J Immunol* **178**(11): 7242-7250.

Potempa, J., A. Sroka, T. Imamura and J. Travis (2003). "Gingipains, the major cysteine proteinases and virulence factors of Porphyromonas gingivalis: structure, function and assembly of multidomain protein complexes." *Curr. Protein Pept. Sci.* **4**(6): 397-407.

Pritchard, M. H. (1992). "An examination of the role of female hormones and pregnancy as risk factors for rheumatoid arthritis, using a male population as control group." *Br J Rheumatol* **31**(6): 395-399.

Pritzker, L. B., S. Joshi, J. J. Gowan, G. Harauz and M. A. Moscarello (2000). "Deimination of myelin basic protein. 1. Effect of deimination of arginyl residues of myelin basic protein on its structure and susceptibility to digestion by cathepsin D." *Biochemistry* **39**(18): 5374-5381.

Pritzker, L. B. and M. A. Moscarello (1998). "A novel microtubule independent effect of paclitaxel: the inhibition of peptidylarginine deiminase from bovine brain." *Biochim Biophys Acta* **1388**(1): 154-160.

Proost, P., T. Loos, A. Mortier, E. Schutyser, M. Gouwy, S. Noppen, C. Dillen, I. Ronsse, R. Conings, S. Struyf, G. Opdenakker, P. C. Maudgal and J. Van Damme (2008). "Citrullination of CXCL8 by peptidylarginine deiminase alters receptor usage, prevents proteolysis, and dampens tissue inflammation." *J Exp Med* **205**(9): 2085-2097.

Pruzanski, W., R. A. Greenwald, I. P. Street, F. Laliberte, E. Stefanski and P. Vadas (1992). "Inhibition of enzymatic activity of phospholipases A2 by minocycline and doxycycline." *Biochem Pharmacol* **44**(6): 1165-1170.

Pryor, W. A. and K. Stone (1993). "Oxidants in cigarette smoke. Radicals, hydrogen peroxide, peroxynitrate, and peroxynitrite." *Ann N Y Acad Sci* **686**: 12-27; discussion 27-18.

Quinn, S. M., J. B. Zhang, J. C. Gunsolley, H. A. Schenkein and J. G. Tew (1998). "The influence of smoking and race on adult periodontitis and serum IgG2 levels." *J Periodontol* **69**(2): 171-177.

Rams, T. E., M. A. Listgarten and J. Slots (2006). "Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis subgingival presence, species-specific serum immunoglobulin G antibody levels, and periodontitis disease recurrence." *J Periodontal Res* **41**(3): 228-234.

Rantapaa-Dahlqvist, S., B. A. de Jong, E. Berglin, G. Hallmans, G. Wadell, H. Stenlund, U. Sundin and W. J. van Venrooij (2003). "Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis." *Arthritis Rheum.* **48**(10): 2741-2749.

Redlitz, A., B. J. Fowler, E. F. Plow and L. A. Miles (1995). "The role of an enolase-related molecule in plasminogen binding to cells." *Eur J Biochem* **227**(1-2): 407-415.

Ribeiro, J., A. Leao and A. B. Novaes (2005). "Periodontal infection as a possible severity factor for rheumatoid arthritis." *J Clin Periodontol* **32**(4): 412-416.

Rodriguez, S. B., B. L. Stitt and D. E. Ash (2009). "Expression of peptidylarginine deiminase from Porphyromonas gingivalis in Escherichia coli: Enzyme purification and characterization." *Arch. Biochem. Biophys.* **488**(1): 14-22.

Rodriguez, S. B., B. L. Stitt and D. E. Ash (2010). "Cysteine 351 is an essential nucleophile in catalysis by Porphyromonas gingivalis peptidylarginine deiminase." *Arch Biochem Biophys* **504**(2): 190-196.

Rosenstein, E. D., R. A. Greenwald, L. J. Kushner and G. Weissmann (2004). "Hypothesis: The humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis." *Inflammation* **28**(6): 311-318.

Routsias, J. G., J. D. Goules, A. Goules, G. Charalampakis and D. Pikazis (2011). "Autopathogenic correlation of periodontitis and rheumatoid arthritis." *Rheumatology (Oxford)*.

Rudney, J. D., R. Chen and G. J. Sedgewick (2001). "Intracellular Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis in buccal epithelial cells collected from human subjects." *Infect Immun* **69**(4): 2700-2707.

Saulot, V., O. Vittecoq, R. Charlionet, P. Fardellone, C. Lange, L. Marvin, N. Machour, X. Le Loet, D. Gilbert and F. Tron (2002). "Presence of autoantibodies to the glycolytic enzyme alpha-enolase in sera from patients with early rheumatoid arthritis." *Arthritis Rheum* **46**(5): 1196-1201.

Schellekens, G. A., B. A. de Jong, F. H. van den Hoogen, L. B. van de Putte and W. J. van Venrooij (1998). "Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies." *J Clin Invest* **101**(1): 273-281.

Schellekens, G. A., H. Visser, B. A. de Jong, F. H. van den Hoogen, J. M. Hazes, F. C. Breedveld and W. J. van Venrooij (2000). "The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide." *Arthritis Rheum* **43**(1): 155-163.

Schenkein, H. A., S. E. Barbour and J. G. Tew (2007). "Cytokines and inflammatory factors regulating immunoglobulin production in aggressive periodontitis." *Periodontol 2000* **45**: 113-127.

Schett, G. (2007). "Cells of the synovium in rheumatoid arthritis. Osteoclasts." *Arthritis Res Ther* **9**(1): 203.

Scott, D. L., F. Wolfe and T. W. Huizinga (2010). "Rheumatoid arthritis." *Lancet* **376**(9746): 1094-1108.

Sebbag, M., N. Moinard, I. Auger, C. Clavel, J. Arnaud, L. Nogueira, J. Roudier and G. Serre (2006). "Epitopes of human fibrin recognized by the rheumatoid arthritis-specific autoantibodies to citrullinated proteins." *Eur. J. Immunol.* **36**(8): 2250-2263.

Sebbag, M., M. Simon, C. Vincent, C. Masson-Bessiere, E. Girbal, J. J. Durieux and G. Serre (1995). "The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies." *J Clin Invest* **95**(6): 2672-2679.

Shirai, H., T. L. Blundell and K. Mizuguchi (2001). "A novel superfamily of enzymes that catalyze the modification of guanidino groups." *Trends Biochem. Sci.* **26**(8): 465-468.

Shirai, H., Y. Mokrab and K. Mizuguchi (2006). "The guanidino-group modifying enzymes: structural basis for their diversity and commonality." *Proteins* **64**(4): 1010-1023.

Sidiropoulos, P. I., G. Goulielmos, G. K. Voloudakis, E. Petraki and D. T. Boumpas (2008). "Inflammasomes and rheumatic diseases: evolving concepts." *Ann Rheum Dis* **67**(10): 1382-1389.

Silman, A. J., A. J. MacGregor, W. Thomson, S. Holligan, D. Carthy, A. Farhan and W. E. Ollier (1993). "Twin concordance rates for rheumatoid arthritis: results from a nationwide study." *Br J Rheumatol* **32**(10): 903-907.

Silman, A. J., J. Newman and A. J. MacGregor (1996). "Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins." *Arthritis Rheum* **39**(5): 732-735.

Simon, M., E. Girbal, M. Sebbag, V. Gomes-Daudrix, C. Vincent, G. Salama and G. Serre (1993). "The cytokeratin filament-aggregating protein filaggrin is the target of the so-called "antikeratin antibodies," autoantibodies specific for rheumatoid arthritis." *J Clin Invest* **92**(3): 1387-1393.

Simpsonv, W., T. Olczak and C. A. Genco (2004). "Lysine-specific gingipain K and heme/hemoglobin receptor HmuR are involved in heme utilization in Porphyromonas gingivalis." *Acta Biochim Pol* **51**(1): 253-262.

Smalley, J. W., J. Silver, P. J. Marsh and A. J. Birss (1998). "The periodontopathogen Porphyromonas gingivalis binds iron protoporphyrin IX in the mu-oxo dimeric form:

an oxidative buffer and possible pathogenic mechanism." *Biochem J* **331 (Pt 3)**: 681-685.

Smalley, J. W., M. F. Thomas, A. J. Birss, R. Withnall and J. Silver (2004). "A combination of both arginine- and lysine-specific gingipain activity of Porphyromonas gingivalis is necessary for the generation of the micro-oxo bishaem-containing pigment from haemoglobin." *Biochem J* **379**(Pt 3): 833-840.

Smiley, S. T., J. A. King and W. W. Hancock (2001). "Fibrinogen stimulates macrophage chemokine secretion through toll-like receptor 4." *J Immunol* **167**(5): 2887-2894.

Snir, O., M. Widhe, C. von Spee, J. Lindberg, L. Padyukov, K. Lundberg, A. Engstrom, P. J. Venables, J. Lundeberg, R. Holmdahl, L. Klareskog and V. Malmstrom (2009). "Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles." *Ann. Rheum. Dis.* **68**(5): 736-743.

Sokolove, J., X. Zhao, P. E. Chandra and W. H. Robinson (2011). "Immune complexes containing citrullinated fibrinogen costimulate macrophages via Toll-like receptor 4 and Fcgamma receptor." *Arthritis Rheum* **63**(1): 53-62.

Sorsa, T., N. S. Ramamurthy, A. T. Vernillo, X. Zhang, Y. T. Konttinen, B. R. Rifkin and L. M. Golub (1998). "Functional sites of chemically modified tetracyclines: inhibition of the oxidative activation of human neutrophil and chicken osteoclast pro-matrix metalloproteinases." *J Rheumatol* **25**(5): 975-982.

Sousa, L. P., B. M. Silva, B. S. Brasil, S. V. Nogueira, P. C. Ferreira, E. G. Kroon, K. Kato and C. A. Bonjardim (2005). "Plasminogen/plasmin regulates alpha-enolase expression through the MEK/ERK pathway." *Biochem Biophys Res Commun* **337**(4): 1065-1071.

Steen, H. and M. Mann (2004). "The ABC's (and XYZ's) of peptide sequencing." *Nat Rev Mol Cell Biol* **5**(9): 699-711.

Stein, J. M., H. K. Machulla, R. Smeets, F. Lampert and S. Reichert (2008). "Human leukocyte antigen polymorphism in chronic and aggressive periodontitis among Caucasians: a meta-analysis." *J Clin Periodontol* **35**(3): 183-192.

Steinert, P. M. and W. W. Idler (1979). "Postsynthetic modifications of mammalian epidermal alpha-keratin." *Biochemistry* **18**(25): 5664-5669.

Steinert, P. M., D. A. Parry and L. N. Marekov (2003). "Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root shealth." *J Biol Chem* **278**(42): 41409-41419.

Stone, E. M., T. H. Schaller, H. Bianchi, M. D. Person and W. Fast (2005). "Inactivation of two diverse enzymes in the amidinotransferase superfamily by 2-chloroacetamidine: dimethylargininase and peptidylarginine deiminase." *Biochemistry* **44**(42): 13744-13752.

Sugawara, K., Y. Oikawa and T. Ouchi (1982). "Identification and properties of peptidylarginine deiminase from rabbit skeletal muscle." *J. Biochem.* **91**(3): 1065-1071.

Sugiyama, D., K. Nishimura, K. Tamaki, G. Tsuji, T. Nakazawa, A. Morinobu and S. Kumagai (2010). "Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies." *Ann Rheum Dis* **69**(1): 70-81.

Sutton, B., A. Corper, V. Bonagura and M. Taussig (2000). "The structure and origin of rheumatoid factors." *Immunology Today* **21**(4): 177-183.

Suzuki, A., R. Yamada, X. Chang, S. Tokuhiro, T. Sawada, M. Suzuki, M. Nagasaki, M. Nakayama-Hamada, R. Kawaida, M. Ono, M. Ohtsuki, H. Furukawa, S. Yoshino, M. Yukioka, S. Tohma, T. Matsubara, S. Wakitani, R. Teshima, Y. Nishioka, A. Sekine, A. Iida, A. Takahashi, T. Tsunoda, Y. Nakamura and K. Yamamoto (2003).

"Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis." *Nat Genet* **34**(4): 395-402.
Takahara, H., H. Okamoto and K. Sugawara (1985). "Specific modification of the functional arginine residue in soybean trypsin inhibitor (Kunitz) by peptidylarginine deiminase." *J Biol Chem* **260**(14): 8378-8383.
Takahashi, N., K. Saito, C. F. Schachtele and T. Yamada (1997). "Acid tolerance and acid-neutralizing activity of Porphyromonas gingivalis, Prevotella intermedia and Fusobacterium nucleatum." *Oral Microbiol Immunol* **12**(6): 323-328.
Takata, Y., H. Inoue, A. Sato, K. Tsugawa, K. Miyatake, D. Hamada, F. Shinomiya, S. Nakano, N. Yasui, T. Tanahashi and M. Itakura (2008). "Replication of reported genetic associations of PADI4, FCRL3, SLC22A4 and RUNX1 genes with rheumatoid arthritis: results of an independent Japanese population and evidence from meta-analysis of East Asian studies." *J Hum Genet* **53**(2): 163-173.
Takeuchi, Y., M. Umeda, M. Sakamoto, Y. Benno, Y. Huang and I. Ishikawa (2001). "Treponema socranskii, Treponema denticola, and Porphyromonas gingivalis are associated with severity of periodontal tissue destruction." *J Periodontol* **72**(10): 1354-1363.
Tarcsa, E., L. N. Marekov, J. Andreoli, W. W. Idler, E. Candi, S. I. Chung and P. M. Steinert (1997). "The fate of trichohyalin. Sequential post-translational modifications by peptidyl-arginine deiminase and transglutaminases." *J Biol Chem* **272**(44): 27893-27901.
Tarcsa, E., L. N. Marekov, G. Mei, G. Melino, S. C. Lee and P. M. Steinert (1996). "Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin." *J Biol Chem* **271**(48): 30709-30716.
Tilleman, K., K. Van Beneden, A. Dhondt, I. Hoffman, F. De Keyser, E. Veys, D. Elewaut and D. Deforce (2005). "Chronically inflamed synovium from spondyloarthropathy and rheumatoid arthritis investigated by protein expression profiling followed by tandem mass spectrometry." *Proteomics* **5**(8): 2247-2257.
Tomar, S. L. and S. Asma (2000). "Smoking-attributable periodontitis in the United States: findings from NHANES III. National Health and Nutrition Examination Survey." *J Periodontol* **71**(5): 743-751.
Travis, J., R. Pike, T. Imamura and J. Potempa (1997). "Porphyromonas gingivalis proteinases as virulence factors in the development of periodontitis." *J. Periodontal Res.* **32**(1 Pt 2): 120-125.
Vallance, P. and J. Leiper (2002). "Blocking NO synthesis: how, where and why?" *Nat Rev Drug Discov* **1**(12): 939-950.
Van Beers, J. J., R. Raijmakers, L.-E. Alexander, J. Stammen-Vogelzangs, A. Lokate, A. Heck, R. Schasfoort and G. J. Pruijn (2010). "Mapping of citrullinated fibrinogen B-cell epitopes in rheumatoid arthritis by imaging surface plasmon resonance." *Arthritis Res Ther* **12**(6): R219.
van der Helm-van Mil, A. H., K. N. Verpoort, F. C. Breedveld, T. W. Huizinga, R. E. Toes and R. R. de Vries (2006). "The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis." *Arthritis Rheum.* **54**(4): 1117-1121.
Van Steendam, K., K. Tilleman, M. De Ceuleneer, F. De Keyser, D. Elewaut and D. Deforce (2010). "Citrullinated vimentin as an important antigen in immune complexes from synovial fluid of rheumatoid arthritis patients with antibodies against citrullinated proteins." *Arthritis Res Ther* **12**(4): R132.

van Venrooij, W. J. and A. J. Zendman (2008). "Anti-CCP2 Antibodies: An Overview and Perspective of the Diagnostic Abilities of this Serological Marker for Early Rheumatoid Arthritis." *Clin Rev Allergy Immunol* **34**(1): 36-39.

van Winkelhoff, A. J., B. G. Loos, W. A. van der Reijden and U. van der Velden (2002). "Porphyromonas gingivalis, Bacteroides forsythus and other putative periodontal pathogens in subjects with and without periodontal destruction." *J Clin Periodontol* **29**(11): 1023-1028.

Vanegas, G., W. Quiñones, C. Carrasco-López, J. Concepción, F. Albericio and L. Avilán (2007). "Enolase as a plasminogen binding protein in Leishmania mexicana." *Parasitology Research* **101**(6): 1511-1516.

Vessey, M. P., L. Villard-Mackintosh and D. Yeates (1987). "Oral contraceptives, cigarette smoking and other factors in relation to arthritis." *Contraception* **35**(5): 457-464.

von Delwig, A., J. Locke, J. H. Robinson and W. F. Ng (2010). "Response of Th17 cells to a citrullinated arthritogenic aggrecan peptide in patients with rheumatoid arthritis." *Arthritis Rheum* **62**(1): 143-149.

Vossenaar, E. R., N. Despres, E. Lapointe, A. van der Heijden, M. Lora, T. Senshu, W. J. van Venrooij and H. A. Menard (2004a). "Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin." *Arthritis Res Ther* **6**(2): R142-150.

Vossenaar, E. R., S. Nijenhuis, M. M. Helsen, A. van der Heijden, T. Senshu, W. B. van den Berg, W. J. van Venrooij and L. A. Joosten (2003a). "Citrullination of synovial proteins in murine models of rheumatoid arthritis." *Arthritis Rheum* **48**(9): 2489-2500.

Vossenaar, E. R., T. R. Radstake, A. van der Heijden, M. A. van Mansum, C. Dieteren, D. J. de Rooij, P. Barrera, A. J. Zendman and W. J. van Venrooij (2004b). "Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages." *Ann Rheum Dis* **63**(4): 373-381.

Vossenaar, E. R., T. J. Smeets, M. C. Kraan, J. M. Raats, W. J. van Venrooij and P. P. Tak (2004c). "The presence of citrullinated proteins is not specific for rheumatoid synovial tissue." *Arthritis Rheum.* **50**(11): 3485-3494.

Vossenaar, E. R., A. J. Zendman, W. J. van Venrooij and G. J. Pruijn (2003b). "PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease." *Bioessays* **25**(11): 1106-1118.

Wang, Y., J. Wysocka, J. Sayegh, Y. H. Lee, J. R. Perlin, L. Leonelli, L. S. Sonbuchner, C. H. McDonald, R. G. Cook, Y. Dou, R. G. Roeder, S. Clarke, M. R. Stallcup, C. D. Allis and S. A. Coonrod (2004). "Human PAD4 regulates histone arginine methylation levels via demethylimination." *Science* **306**(5694): 279-283.

Wegner, N., K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables (2010a). "Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis." *Immunol Rev* **233**(1): 34-54.

Wegner, N., R. Wait, A. Sroka, S. Eick, K. A. Nguyen, K. Lundberg, A. Kinloch, S. Culshaw, J. Potempa and P. J. Venables (2010b). "Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and alpha-enolase: implications for autoimmunity in rheumatoid arthritis." *Arthritis Rheum* **62**(9): 2662-2672.

Wegner, N., R. Wait and P. J. Venables (2009). "Evolutionarily conserved antigens in autoimmune disease: implications for an infective aetiology." *Int J Biochem Cell Biol* **41**(2): 390-397.

Willis, V. C., A. M. Gizinski, N. K. Banda, C. P. Causey, B. Knuckley, K. N. Cordova, Y. Luo, B. Levitt, M. Glogowska, P. Chandra, L. Kulik, W. H. Robinson, W. P. Arend, P. R. Thompson and V. M. Holers (2011). "N-alpha-Benzoyl-N5-(2-Chloro-1-

Iminoethyl)-L-Ornithine Amide, a Protein Arginine Deiminase Inhibitor, Reduces the Severity of Murine Collagen-Induced Arthritis." *J Immunol* **186**(7): 4396-4404.

Wygrecka, M., L. M. Marsh, R. E. Morty, I. Henneke, A. Guenther, J. Lohmeyer, P. Markart and K. T. Preissner (2009). "Enolase-1 promotes plasminogen-mediated recruitment of monocytes to the acutely inflamed lung." *Blood* **113**(22): 5588-5598.

Yamamoto, Y., N. Nishida, M. Tanaka, N. Hayashi, R. Matsuse, K. Nakayama, K. Morimoto and S. Shizukuishi (2005). "Association between passive and active smoking evaluated by salivary cotinine and periodontitis." *J Clin Periodontol* **32**(10): 1041-1046.

Yang, H. W., Y. F. Huang and M. Y. Chou (2004). "Occurrence of Porphyromonas gingivalis and Tannerella forsythensis in periodontally diseased and healthy subjects." *J Periodontol* **75**(8): 1077-1083.

Yao, H., P. Li, B. J. Venters, S. Zheng, P. R. Thompson, B. F. Pugh and Y. Wang (2008). "Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis." *J Biol Chem* **283**(29): 20060-20068.

Yavlovich, A., H. Rechnitzer and S. Rottem (2007). "Alpha-enolase resides on the cell surface of Mycoplasma fermentans and binds plasminogen." *Infect Immun* **75**(12): 5716-5719.

Yoshie, H., T. Kobayashi, H. Tai and J. C. Galicia (2007). "The role of genetic polymorphisms in periodontitis." *Periodontol 2000* **43**: 102-132.

Yoshimura, M., N. Ohara, Y. Kondo, M. Shoji, S. Okano, Y. Nakano, Y. Abiko and K. Nakayama (2008). "Proteome analysis of Porphyromonas gingivalis cells placed in a subcutaneous chamber of mice." *Oral Microbiol Immunol* **23**(5): 413-418.

Young, B. J., R. K. Mallya, R. D. Leslie, C. J. Clark and T. J. Hamblin (1979). "Anti-keratin antibodies in rheumatoid arthritis." *Br Med J* **2**(6182): 97-99.

Zanelli, E., F. C. Breedveld and R. R. de Vries (2000a). "HLA association with autoimmune disease: a failure to protect?" *Rheumatology (Oxford)* **39**(10): 1060-1066.

Zanelli, E., F. C. Breedveld and R. R. de Vries (2000b). "HLA class II association with rheumatoid arthritis: facts and interpretations." *Hum Immunol* **61**(12): 1254-1261.

Zendman, A. J. W., W. J. van Venrooij and G. J. M. Pruijn (2006). "Use and significance of anti-CCP autoantibodies in rheumatoid arthritis." *Rheumatology* **45**(1): 20-25.

Zhao, X., N. L. Okeke, O. Sharpe, F. M. Batliwalla, A. T. Lee, P. P. Ho, B. H. Tomooka, P. K. Gregersen and W. H. Robinson (2008). "Circulating immune complexes contain citrullinated fibrinogen in rheumatoid arthritis." *Arthritis Res. Ther.* **10**(4): R94.

Appendix

Alignment of *P. gingivalis* PAD with *Cloacamonas acidaminovorans* B0VH37

PG = *Porphyromonas gingivalis* W83 peptidylarginine deiminase

CA = Cloacamonas acidaminovorans B0VH37

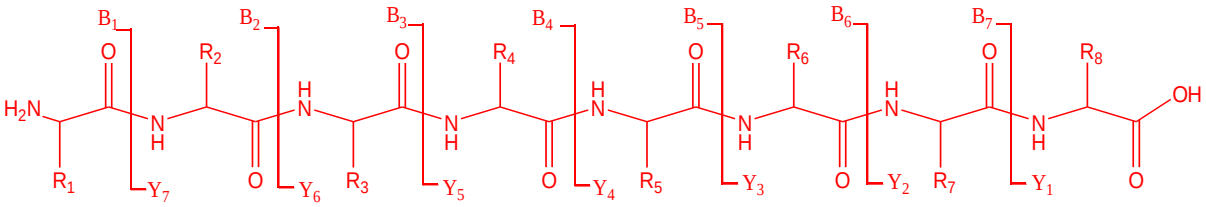
Predited catalytic residues in PPAD are highlighted in grey.

```
PG 2   KKLLQAKALILALGLFQLPAIAQTQMQADRTNGQFATEEMQRAFQETNPPAGPVRAIAEY  61
CA 4   QRILRILLLCSVLCILSSTAFAHHNLPADKANAEIR-------FTETNPPVAPVRPIAEF  56
          *    *   *      * *     **  *           * *****  *** ***

PG 62  ERSAAVLVRYPFGIPMELIKELAKNDKVITIVASESQKNTVITQYTQSGVNLSNCDFIIA 121
CA 57  EPASDVLIRYPLGIPVSLVVQLANTANVICIVSS-SQQSSAISTFTNAGVNMERVSFLNA 115
       *    ** *** ***  *   **    ** ** * **    *   *  ***     *  *

PG 122 KTDSYWTRDYTGWFAMYDTNKVGLVDFIYNRPRPNDDEFPKYEAQYLGIEMFGMKLKQTG 181
CA 116 ATDSYWTRDYGPWFIFDGNDEYGVVDFIYNRPRPNDNLIPQVFANHFALNYYGMNLQQTG 175
        *********  **        * ************   *   *        ** * ***

PG 182 GNYMTDGYGSAVQSHIAYTENSSLSQAQVNQKMKDYLGITHHDVVQDPNGEYINHVDCWG 241
CA 176 GNYMSDGINTAAQTTLVYTENGN-NQTNVNTKMQQYLGITNYLVMTDPNNTYIEHIDCWA 234
       **** **   * *    ****    *  ** **  *****   *  ***  ** * ***

PG 242 KYLAPNKILIRKVPDNHPQHQALEDMAAYFAAQTCAWGTKYEVYRALATNEQPYTNSLIL 301
CA 235 KFLAPDKVMIRSVPASHSQYNAIENAANYFATHNCAWGYPYRVYRVYTPSNEPYTNSLIL 294
       * *** *  ** **  * *  * *  * ***   ****  * ***       ********

PG 302 NNRVFVPVNGPASVDNDALNVYKTAMPGYEIIGVKGASGTPWLGTDALHCRTHEVADKGY 361
CA 295 NKKVFVPIVG-SSNDNSALQAYREAMPGYEVIGVSQISSAPWESTDALHCRTHEIPDKNM 353
       *  ****  *  * ** **  *  ****** ***   *  **  **********  **

PG 362 LYIKHYPILG-EQAGPDYKIEADVVSCANATISPVQCY--YRINGSGSFKAADMTMESTG 418
CA 354 LNIVHTPWHSIVPVGTDIIINTEIIAHSGQPLYTDSLFVCYKVN-SGAWQRNYLQPLTRN 412
       * * * *       * *  *                    *  * **

PG 419 HYTYSFTGLNKNDKVEYYISAADNSGRKETYPFIGEPDPFKF 460
CA 413 YYTVTLNGFAYGDTIRYFIHAADQSGRSVDHPVFAALDPHLF 454
        **    *    *   * * *** ***    *     **  *
```

Identification of citrullinated peptides by tandem mass spectrometry

A good introduction to general aspects of peptide sequencing by mass spectrometry is given in a recent review by Steen and Mann (Steen *et al.* 2004). Here, I will focus on the relevant information regarding identification of citrulline residues in peptides.

Peptides can fragment in the mass spectrometer via several different pathways. Cleavage between the carbonyl group and the nitrogen of the peptide bond (shown in the diagram below) is the most important fragmentation process in collision-induced dissociation (CID) of peptides. The fragments produced may retain either the original carboxyl end of the peptide, or the original amino end. Fragments containing the original amino terminus are known as B ions, and fragments containing the original carboxyl-terminal end of the peptide are known as Y ions. If trypsin or other arg/lysine-specific proteases (such as gingipains in this project) are used for the generation of peptides, the formed peptides will harbour a positive charge on the carboxyl-terminal end. Therefore, after CID, Y type fragment ions are predominantly (but not exclusively) detected, as they retain the positive charge conferred by arginine/lysine at the carboxy-terminus. The diagram below illustrates the conventional nomenclature for peptide fragment ions. The residue cleaved is indicated by a superscripted number, counting from the appropriate end.



The amino acid sequence of the parent ion (injected peptide) can be concluded from the mass difference between the Y-ion or B-ion series. Conversion of arginine to citrulline results in a mass increase of ~1 Da. This is done automatically by a search algorithm in the first instance, highlighting potential hits. Each candidate peptide is then carefully inspected to rule out another amino acid modification (deamidation of asparagines and glutamine residues) which results in the same mass increase of ~1 Da.

Example

In Chapter 3, citrullination of the carboxyl terminal arginine residue of the fibrinogen-derived peptide ESSSHHPGIAEFPSc<u>it</u> was identified. The first fragmentation diagram below shows the theoretical mass differences between the B- and Y-series ions of the arginine-containing peptide (ESSSHHPGIAEFPS<u>R</u>). In the case of a C-terminal citrulline residue (see second diagram below), the mass of all Y-series ions will be shifted (+1 Da) compared to the Y-series ions of the arginine-containing peptide, beginning with the Y_1 ion as this is the ion containing the C-terminal amino acid only. All other Y ions retain, by definition, the original C-terminal amino acid, and will therefore be shifted as well. The B ions contain the original amino-terminal end but not the original carboxy-terminal end and will therefore not be shifted.

BIOLYNX - PROTEIN REPORT Page 1

Data file: Fibrinogen peptide noncitt.pep
Description: Untitled
Printed: Thu Jan 29 17:42:05 2009

Average Mass = 1637.7296, Monoisotopic Mass = 1636.7593
Residues 1-15
N-Terminus = H, C-Terminus = OH
Fragment ions: Monoisotopic/Average (3000) m/z ratios with 1 positive charge(s).

b	130.05	217.08	304.11	391.15	528.21	665.26	762.32	819.34	932.42	1003.46	1132.50
	1	2	3	4	5	6	7	8	9	10	11
	Glu	Ser	Ser	Ser	His	His	Pro	Gly	Ile	Ala	Glu
	15	14	13	12	11	10	9	8	7	6	5
y''	-	1508.72	1421.69	1334.66	1247.63	1110.57	973.51	876.46	819.44	706.35	635.32

b	1279.57	1376.62	1463.66	-
	12	13	14	15
	Phe	Pro	Ser	Arg
	4	3	2	1
y''	506.27	359.20	262.15	175.12

BIOLYNX - PROTEIN REPORT Page 1

Data file: Fibrinogen peptide citt.pep
Description: Untitled
Printed: Thu Jan 22 20:02:36 2009

Associated Spectrum Datafile: 23December08Q_08
Mass Window = 0.60 amu, Threshold = 1.50%
Average Mass = 1638.7143, Monoisotopic Mass = 1637.7433
Residues 1-15
N-Terminus = H, C-Terminus = OH
Modified amino acids: Cit(J) = Citrulline
Fragment ions: Monoisotopic/Average (3000) m/z ratios with 1 positive charge(s).

b	130.05	217.08	304.11	391.15	528.21	665.26	762.32	819.34	932.42	1003.46	1132.50
	1	2	3	4	5	6	7	8	9	10	11
	Glu	Ser	Ser	Ser	His	His	Pro	Gly	Ile	Ala	Glu
	15	14	13	12	11	10	9	8	7	6	5
y''	-	1509.71	1422.68	1335.64	1248.61	1111.55	974.49	877.44	820.42	707.34	636.30

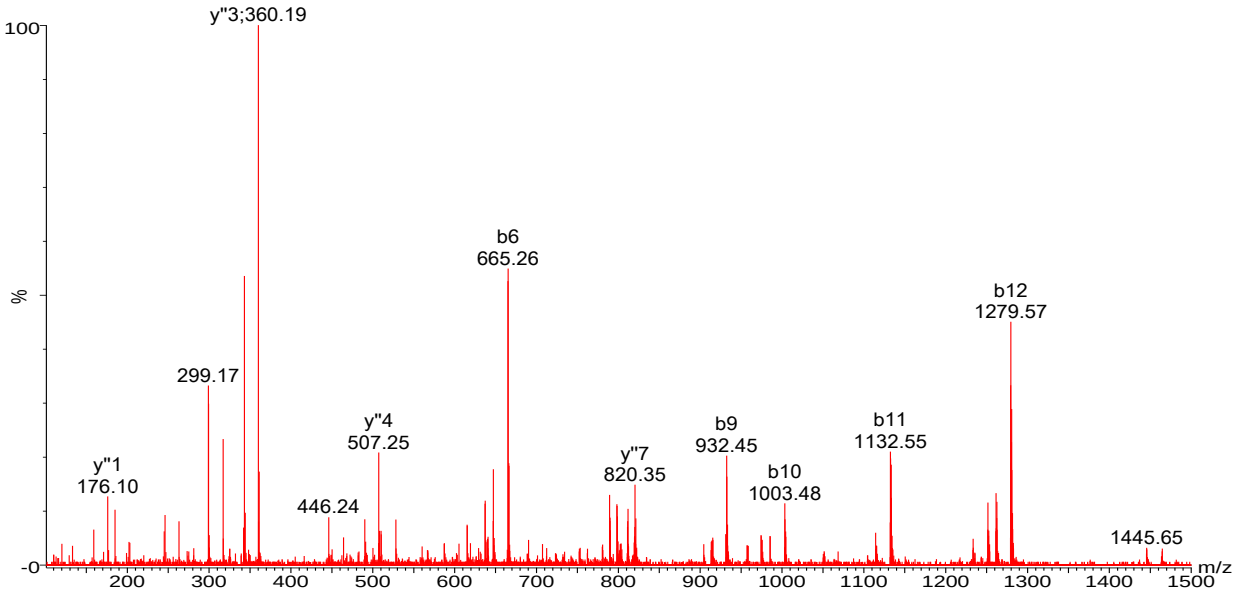
b	1279.57	1376.62	1463.66	-
	12	13	14	15
	Phe	Pro	Ser	Cit
	4	3	2	1
y''	507.26	360.19	263.14	176.10

Usually, not all possible ions of a series are detected, but the combination of the available Y- and B-ion data enables the deduction of the amino acid sequence of the peptide.

Tandem mass spectrum of ESSSHHPGIAEFPS(Cit):

x-axis: mass-to-charge ratio (m/z) of the detected fragments

y-axis: frequency of the detected fragment ions in % relative to the most frequently detected fragment ion (100%, here y''3)



List of peptides generated after incubation of human fibrinogen and α-enolase with *P. gingivalis*

Comments:

The *m/z* ratio is the mass-to-charge ratio of the precursor peptide (before fragmentation), followed by the number and type of elementary charges on the ion (peptide ion charge state, number in brackets). The mass-to-charge ratio is a dimensionless, physical quantity of the peptide ions that determines their separation in the electromagnetic field of the mass spectrometer. The same peptide can exist in several different charge states, giving rise to different m/z ratios. To deduct the original mass of the peptide, the m/z ratio needs to be multiplied by the peptide ion charge state.

The Mascot score is a probability-based score from the Mascot search engine, which is used to identify the source proteins based on the mass spectrometry data. The total score is the absolute probability that the observed match is a random event. The score is reported as -10*LOG_{10}(P), where P is the probability for a match to be a random effect. A probability of P= 0.01 thus becomes a score of 20. The higher the score, the lower is the probability that the match is random. The threshold value for a significant match depends on the number of entries of the database searched and was around 67 ($p<0.05$). However, the experimentally obtained score highly depends on the frequency of the individual detected peptides, which is expected to be low in a relatively complex and dilute sample such as analysed here. In this case, a manual and careful inspection of the raw sequence data is required, which was mainly performed by Dr. Robin Wait.

Table 1 Mass spectrometry analysis of peptides generated after incubation of human fibrinogen with *P. gingivalis* wild-type (WT) or Δ*ppad* strains.

Amino acids that would be present in the uncleaved proteins located carboxy- and amino-terminal to the identified peptides are indicated (separated by a full stop) to demonstrate sites of proteolytic cleavage. Identified citrulline residues (Cit) are underlined.

Sequence	Protein	Location (amino acids)	Strain	*m/z* ratio	Mascot Score
-.ADSGEGDFLAEGGGVCit.G	Fibrinogen α-chain	1-16	WT	769.29 (+2)	56
-.ADSGEGDFLAEGGGVR.G			Δ*ppad*	768.77 (+2)	99
R.GGSTSYGTGSETESPCit.N	Fibrinogen α-chain	253-268	WT	546.91 (+3)	51
R.GGSTSYGTGSETESPR.N			Δ*ppad*	786.82 (+2)	78
R.NPGSSGTGGTATWK.P	Fibrinogen α-chain	290-303	WT	660.78 (+2)	105
			Δ*ppad*	660.77 (+2)	75
P.GSSGTGGTATWK.P		292-303	Δ*ppad*	555.25 (+2)	77
G.SSGTGGTATWK.P		293-303	Δ*ppad*	556.73 (+2)	72
T.LSGIGTLDGFR.H	Fibrinogen α-chain	481-491	Δ*ppad*	568.28 (+2)	46
R.HPDEAAFFDTASTGK.T	Fibrinogen α-chain	494-508	WT	797.32 (+2)	115
K.ESSSHHPGIAEFPSCit.G	Fibrinogen α-chain	540-554	WT	819.83 (+2)	98
S.SHHPGIAEFPSR.G		543-554	Δ*ppad*	445.54 (+3)	39
K.MADEAGSEADHEGTHST.K	Fibrinogen α-chain	584-600	WT	582.19 (+3)	85
K.MADEAGSEADHEGTHSTK.R		584-601	Δ*ppad*	624.90 (+3)	72

Sequence	**Protein**	**Location (amino acids)**	**Strain**	***m/z* ratio**	**Mascot Score**
A.DEAGSEADHEGTHST.K		586-600	WT	771.78 (+2)	81
R.PAPPPISGGGYCit.A	Fibrinogen β-chain	31-42	WT	585.30 (+2)	26
R.PAPPPISGGGYR.A			Δ*ppad*	584.77 (+2)	65
K.DNENVVNEYSSELEK.H	Fibrinogen β-chain	134-148	WT	884.83 (+2)	106
			Δ*ppad*	884.86 (+2)	74
K.GGETSEMYLIQPDSSVK.P	Fibrinogen β-chain	218-234	WT	920.87 (+2)	106
Y.LIQPDSSVK.P		226-234	Δ*ppad*	493.78 (+2)	48
R.QDGSVDFGR.K	Fibrinogen β-chain	256-264	Δ*ppad*	490.70 (+2)	51
K.QGFGNVATNTDGK.N	Fibrinogen β-chain	271-283	WT	654.74 (+2)	117
G.FGNVATNTDGK.N		273-283	WT	562.25 (+2)	62
K.EDGGGWWYNR.C	Fibrinogen β-chain	397-406	Δ*ppad*	620.22 (+2)	40
K.PNMIDAATLK.S	Fibrinogen γ-chain	76-85	WT	537.29 (+2)	62
K.DTVQIHDITGK.D	Fibrinogen γ-chain	141-151	WT	613.79 (+2)	99
K.EGFGHLSPTGTTEFW.L	Fibrinogen γ-chain	213-227	Δ*ppad*	883.38 (+2)	52
R.VELEDWNGR.T	Fibrinogen γ-chain	248-256	Δ*ppad*	559.24 (+2)	41
F.AGGDAGDAFDGFDFGDDPSDK.F	Fibrinogen γ-chain	282-302	Δ*ppad*	1038 (+2)	72
R.LTIGEGQQHHLGGA.K	Fibrinogen γ-chain	392-405	WT	709.31 (+2)	71

Sequence	Protein	Location (amino acids)	Strain	*m/z* ratio	Mascot Score
T.IGEGQQHHLGGA.K		394-405	WT	401.84 (+3)	74
			Δ*ppad*	602.28 (+2)	65

Table 2 Mass spectrometry analysis of peptides generated after incubation of human α-enolase with *P. gingivalis* wild-type (WT) or Δ*ppad* strains.

Sequence	Location (amino acids)	Strain	*m/z* ratio	Mascot Score
G.ASTGIYEALELR.D	39-50	Δ*ppad*	661.83 (+2)	78
S.TGIYEALELCit.D	41-50	WT	583.30 (+2)	24
S.TGIYEALELR.D	41.50	Δ*ppad*	582.79 (+2)	56
K.LNVTEQEK.I	82-89	WT	480.76 (+2)	43
		Δ*ppad*	480.75 (+2)	39
R.HIADLAGNSEVI.L	133-144	Δ*ppad*	619.81 (+2)	65
R.HIADLAGNSEVILPVPA.F	133-149	WT	858.44 (+2)	86
R.HIADLAGNSEVILPVPAF.N	133-150	WT	931.99 (+2)	26
		Δ*ppad*	931.98 (+2)	50
R.HIADLAGNSEVILPVPAFN.V	133-151	WT	989.02 (+2)	61
		Δ*ppad*	988.99 (+2)	28
K.DATNVGDEGGFAPNILEN.K	203-220	WT	916.90 (+2)	69
		Δ*ppad*	916.89 (+2)	32
G.MDVAASEFFR.S	244-253	Δ*ppad*	586.77 (+2)	23
K.DYPVVSIEDPFDQDDWGAWQ.K	286-305	WT	1191.49 (+2)	27
S.IEDPFDQDDWGAWQ.K	292-305	WT	861.34 (+2)	77

Sequence	Location (amino acids)	Strain	*m/z* ratio	Mascot Score
S.AGIQVVGDDLTVTNPK.R	311-326	WT	813.90 (+2)	84
		Δ*ppad*	813.94 (+2)	62
A.GIQVVGDDLTVTNPK.R	312-326	WT	778.39 (+2)	98
		Δ*ppad*	778.38 (+2)	98
G.IQVVGDDLTVTNPK.R	313-326	WT	749.88 (+2)	106
		Δ*ppad*	749.89 (+2)	81
I.QVVGDDLTVTNPK.R	314-326	Δ*ppad*	693.34 (+2)	83
Q.VVGDDLTVTNPK.R	315-326	WT	629.33 (+2)	42

Table 3 Mass spectrometry analysis of citrullinated peptides generated after incubation of human α-enolase with *P. gingivalis* Δ*kgp*.

Sequence	Location (amino acids)	Strain	*m/z* ratio	Mascot Score
V.EVDLFTSKGLFCit.A	21-32	Δ*kgp*	706.85 (+2)	88
P.SGASTGIYEALELCit.D	37-50	Δ*kgp*	734.36 (+2)	39
D.FKSPDDPSCit.Y	261-269	Δ*kgp*	525.23 (+2)	28
S.AGIQVVGDDLTVTNPKCit.I	311-327	Δ*kgp*	892.47 (+2)	62
R.IEEELGSKAKFAGCit.N	413-426	Δ*kgp*	768.40 (+2)	20

Alignment of PPAD with Streptococcus mutans agmatine deiminase

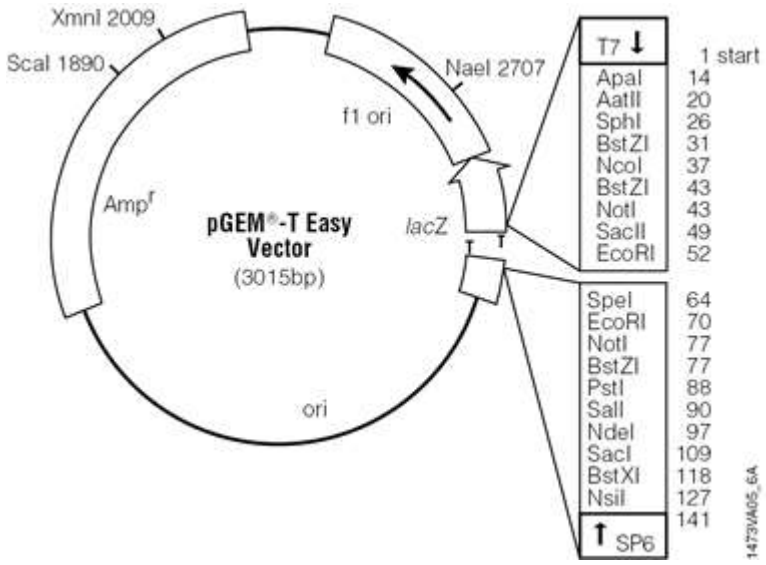
Target (PPAD) - Template (*S. mutans* agmatine deiminase, 2ewoA) alignment

Identical (id.) residues are marked in line three of each row, with the predicted catalytic residues of PPAD highlighted in grey.

```
PPAD            -RAFQETNPPAGPVRAIAE---------Y-------ERSAAVLVRYPFGIPMELIKELAK
2ewoA           AKRIKNTTPKQDGFRXPGEFEKQKQIWXLWPWRNDNWRLGAKPAQ---KAFLEVAEAISE
id. residues    ......T.P.....R...E..................R..A...........E.......

PPAD            NDKVITIVASESQKNTVITQYTQSGVNLSNCDFIIAKTDSYWTRDYTGWFAMYDTNKVGL
2ewoA           FEPVSLCVPP---LQYENALARVSELGSHNIRIIEXTNDDAWIRDCGPTFLVNDKGDLRA
id. residues    ...V...V..............S.....N...I....D..W.RD....F...D......

PPAD            V---------DFIYNRP-RPNDDEFPKYEAQYLGIE-MFGMKLKQTGGNYMTDGYGSAVQ
2ewoA           VDWEFNAWGGLVDGLYFPWDQDALVARKVCEIEGVDSYKTKDFVLEGGSIHVDGEGTVLV
id. residues    V...................D...........G............GG....DG.G....

PPAD            SHI---AYTENSSLSQAQVNQKMKDYLGITHHDVVQDPNGEY----INHVDCWGKYLAPN
2ewoA           TEXCLLHPSRNPHLTKEDIEDKLKD--YLNCVKVLWVKDGIDPYETNGHIDDVACFIRPG
id. residues    ..........N..L.......K.KD.........V.....G........H.D........P.

PPAD            KILIR-KVPDNHPQHQALEDMAAYFAAQTCAWGTKYEVYRALA---------T------N
2ewoA           EVACIYTDDKEHPFYQEAKAAYDFLSQQTDAKGRPLKVHKXCVTKEPCYLQEAATIDYVE
id. residues    ...........HP..Q...........QT.A.G....V......................

PPAD            ----EQPYTNSLILNNRVFVPVNGPASVDNDALNVYKTAMPGYEIIGVKGASGTPWLGTD
2ewoA           GEXAIASYLNFLIVNGGIILPQ-YGDENDQLAKQQVQEXFPDRK--VVGVRTEEIAYGGG
id. residues    .......Y.N.LI.N.....P.......D..A.........P......V.........G..

PPAD            ALHCRTHEV---
2ewoA           NIHCITQQQPAT
id. residues    ..HC.T......
```

Vector Maps and Sequences

pGEM®-T-Easy Vector Sequence



Source: www.promega.com

The pGEM(R)-T Easy Vector has been linearized with EcoRV at Base 60 of this sequence (indicated by an asterisk *) and a T added to both 3' -ends. The added T is not included in this sequence.

```
   1  GGGCGAATTG GGCCCGACGT CGCATGCTCC CGGCCGCCAT GGCGGCCGCG

  51  GGAATTCGAT* ATCACTAGTG AATTCGCGGC CGCCTGCAGG TCGACCATAT

 101  GGGAGAGCTC CCAACGCGTT GGATGCATAG CTTGAGTATT CTATAGTGTC

 151  ACCTAAATAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT

 201  TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG

 251  TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC

 301  GCTCACTGCC CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA

 351  TGAATCGGCC AACGCGCGGG GAGAGGCGGT TTGCGTATTG GGCGCTCTTC

 401  CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTCGTTCGG CTGCGGCGAG

 451  CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC AGAATCAGGG

 501  GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA

 551  CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCCTG

 601  ACGAGCATCA CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA

 651  GGACTATAAA GATACCAGGC GTTTCCCCCT GGAAGCTCCC TCGTGCGCTC

 701  TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC TTTCTCCCTT

 751  CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG
```

```
 801  GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA
 851  GCCCGACCGC TGCGCCTTAT CCGGTAACTA TCGTCTTGAG TCCAACCCGG
 901  TAAGACACGA CTTATCGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC
 951  AGAGCGAGGT ATGTAGGCGG TGCTACAGAG TTCTTGAAGT GGTGGCCTAA
1001  CTACGGCTAC ACTAGAAGAA CAGTATTTGG TATCTGCGCT CTGCTGAAGC
1051  CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC
1101  ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG
1151  AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG
1201  CTCAGTGGAA CGAAAACTCA CGTTAAGGGA TTTTGGTCAT GAGATTATCA
1251  AAAAGGATCT TCACCTAGAT CCTTTTAAAT TAAAAATGAA GTTTTAAATC
1301  AATCTAAAGT ATATATGAGT AAACTTGGTC TGACAGTTAC CAATGCTTAA
1351  TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT
1401  GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC
1451  TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG
1501  ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT
1551  CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAATTGTT GCCGGGAAGC
1601  TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG
1651  CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC
1701  TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA
1751  AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG
1801  CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT
1851  GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA
1901  GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT
1951  CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
2001  ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT
2051  GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT
2101  CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT
2151  GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT
2201  CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA
2251  GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG
2301  CGCACATTTC CCCGAAAAGT GCCACCTGAT GCGGTGTGAA ATACCGCACA
2351  GATGCGTAAG GAGAAAATAC CGCATCAGGA AATTGTAAGC GTTAATATTT
2401  TGTTAAAATT CGCGTTAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA
2451  TAGGCCGAAA TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT
2501  AGGGTTGAGT GTTGTTCCAG TTTGGAACAA GAGTCCACTA TTAAAGAACG
2551  TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA
2601  CTACGTGAAC CATCACCCTA ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA
2651  AGCACTAAAT CGGAACCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG
```

2701 GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG

2751 GGCGCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCAC

2801 ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCATT CGCCATTCAG

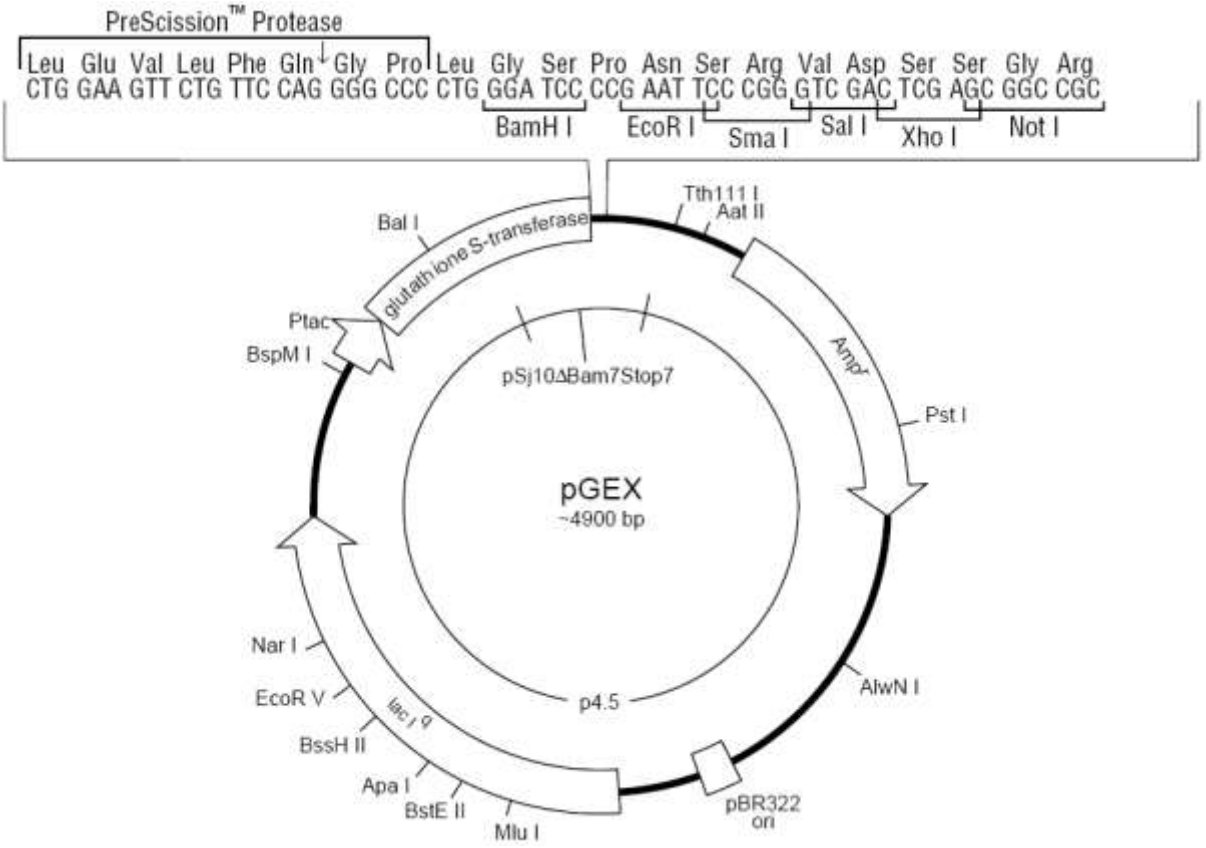
2851 GCTGCGCAAC TGTTGGGAAG GGCGATCGGT GCGGGCCTCT TCGCTATTAC

2901 GCCAGCTGGC GAAAGGGGGA TGTGCTGCAA GGCGATTAAG TTGGGTAACG

2951 CCAGGGTTTT CCCAGTCACG ACGTTGTAAA ACGACGGCCA GTGAATTGTA

3001 ATACGACTCA CTATA

pGEX-6P-3 Vector Sequence



Source: www.gelifesciences.com

GenBank Accession Number: U78874

```
   1 acgttatcga ctgcacggtg caccaatgct tctggcgtca ggcagccatc ggaagctgtg
  61 gtatggctgt gcaggtcgta aatcactgca taattcgtgt cgctcaaggc gcactcccgt
 121 tctggataat gtttttgcg ccgacatcat aacggttctg gcaaatattc tgaaatgagc
 181 tgttgacaat taatcatcgg ctcgtataat gtgtggaatt gtgagcggat aacaatttca
 241 cacaggaaac agtattcatg tcccctatac taggttattg gaaaattaag ggccttgtgc
 301 aacccactcg acttcttttg gaatatcttg aagaaaaata tgaagagcat ttgtatgagc
 361 gcgatgaagg tgataaatgg cgaaacaaaa agtttgaatt gggtttggag tttcccaatc
 421 ttccttatta tattgatggt gatgttaaat taacacagtc tatggccatc atacgttata
 481 tagctgacaa gcacaacatg ttgggtggtt gtccaaaaga gcgtgcagag atttcaatgc
 541 ttgaaggagc ggttttggat attagatacg gtgtttcgag aattgcatat agtaaagact
 601 ttgaaactct caaagttgat tttcttagca agctacctga aatgctgaaa atgttcgaag
 661 atcgtttatg tcataaaaca tatttaaatg gtgatcatgt aacccatcct gacttcatgt
 721 tgtatgacgc tcttgatgtt gttttataca tggacccaat gtgcctggat gcgttcccaa
 781 aattagtttg ttttaaaaaa cgtattgaag ctatcccaca aattgataag tacttgaaat
 841 ccagcaagta tatagcatgg cctttgcagg gctggcaagc cacgtttggt ggtggcgacc
 901 atcctccaaa atcggatctg gaagttctgt tccaggggcc cctgggatcc ccgaattccc
 961 gggtcgactc gagcggccgc atcgtgactg actgacgatc tgcctcgcgc gtttcggtga
1021 tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc
```

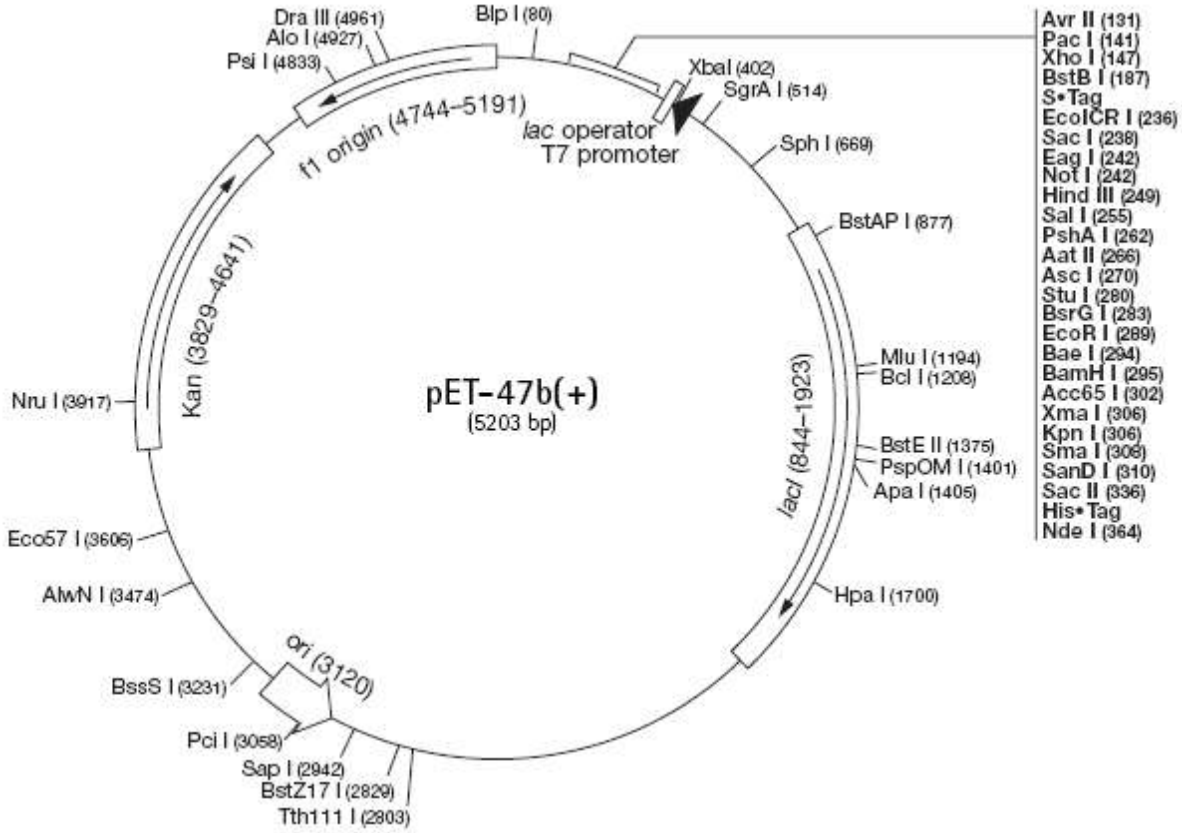
1081 ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg
1141 cgcagccatg acccagtcac gtagcgatag cggagtgtat aattcttgaa gacgaaaggg
1201 cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt cttagacgtc
1261 aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca
1321 ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa
1381 aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt
1441 ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca
1501 gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag
1561 ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc
1621 ggtattatcc cgtgttgacg ccgggcaaga gcaactcggt cgccgcatac actattctca
1681 gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt
1741 aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct
1801 gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt
1861 aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga
1921 caccacgatg cctgcagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact
1981 tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc
2041 acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga
2101 gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt
2161 agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga
2221 gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact
2281 ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga
2341 taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt
2401 agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca
2461 aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct
2521 ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta
2581 gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct
2641 aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc
2701 aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca
2761 gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga
2821 aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg
2881 aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt
2941 cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag
3001 cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt
3061 tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt
3121 tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga
3181 ggaagcggaa gagcgcctga tgcggtattt tctccttacg catctgtgcg gtatttcaca
3241 ccgcataaat tccgacacca tcgaatggtg caaaaccttt cgcggtatgg catgatagcg
3301 cccggaagag agtcaattca gggtggtgaa tgtgaaacca gtaacgttat acgatgtcgc
3361 agagtatgcc ggtgtctctt atcagaccgt ttcccgcgtg gtgaaccagg ccagccacgt
3421 ttctgcgaaa acgcgggaaa aagtggaagc ggcgatggcg gagctgaatt acattcccaa
3481 ccgcgtggca caacaactgg cgggcaaaca gtcgttgctg attggcgttg ccacctccag
3541 tctggccctg cacgcgccgt cgcaaattgt cgcggcgatt aaatctcgcg ccgatcaact
3601 gggtgccagc gtggtggtgt cgatggtaga acgaagcggc gtcgaagcct gtaaagcggc
3661 ggtgcacaat cttctcgcgc aacgcgtcag tgggctgatc attaactatc cgctggatga

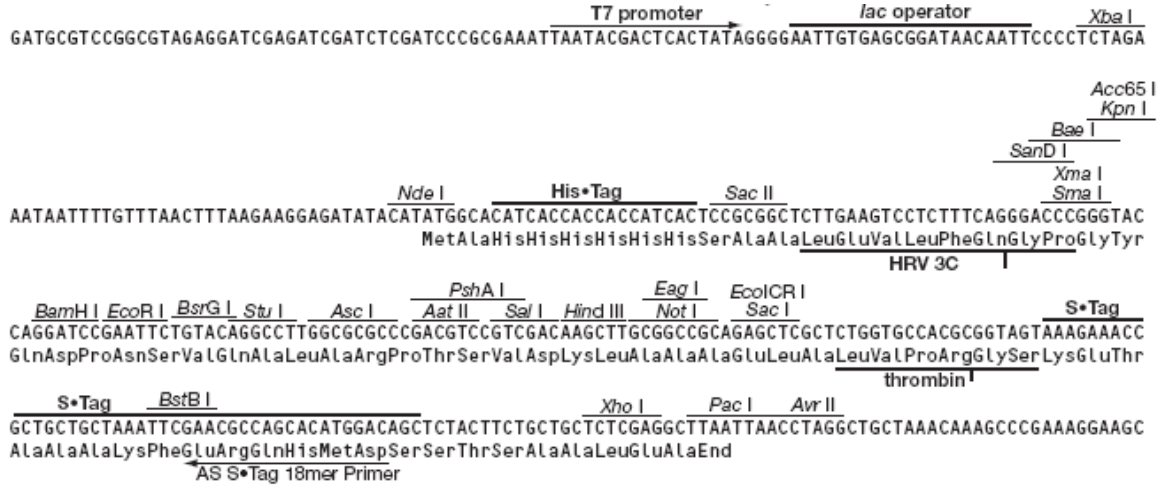
3721 ccaggatgcc attgctgtgg aagctgcctg cactaatgtt ccggcgttat ttcttgatgt
3781 ctctgaccag acacccatca acagtattat tttctcccat gaagacggta cgcgactggg
3841 cgtggagcat ctggtcgcat tgggtcacca gcaaatcgcg ctgttagcgg gcccattaag
3901 ttctgtctcg gcgcgtctgc gtctggctgg ctggcataaa tatctcactc gcaatcaaat
3961 tcagccgata gcggaacggg aaggcgactg gagtgccatg tccggttttc aacaaaccat
4021 gcaaatgctg aatgagggca tcgttcccac tgcgatgctg gttgccaacg atcagatggc
4081 gctgggcgca atgcgcgcca ttaccgagtc cgggctgcgc gttggtgcgg atatctcggt
4141 agtgggatac gacgataccg aagacagctc atgttatatc ccgccgtcaa ccaccatcaa
4201 acaggatttt cgcctgctgg ggcaaaccag cgtggaccgc ttgctgcaac tctctcaggg
4261 ccaggcggtg aagggcaatc agctgttgcc cgtctcactg gtgaaaagaa aaaccaccct
4321 ggcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc
4381 acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc
4441 tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa
4501 ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac ggattcactg
4561 gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttacccaact taatcgcctt
4621 gcagcacatc ccccttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct
4681 tcccaacagt tgcgcagcct gaatggcgaa tggcgctttg cctggtttcc ggcaccagaa
4741 gcggtgccgg aaagctggct ggagtgcgat cttcctgagg ccgatactgt cgtcgtcccc
4801 tcaaactggc agatgcacgg ttacgatgcg cccatctaca ccaacgtaac ctatcccatt
4861 acggtcaatc cgccgtttgt tcccacggag aatccgacgg gttgttactc gctcacattt
4921 aatgttgatg aaagctggct acaggaaggc cagacgcgaa ttatttttga tggcgttgga
4981 att

pET-47b(+) Vector Sequence

The pET-47b(+) plasmid features the ability to express fusion proteins with an N-terminal His•Tag® coding sequence that is cleavable with the human rhinovirus (HRV) 3C protease. The plasmid contains a strong T7*lac* promoter, an optimized RBS, the coding sequence for the HRV 3C protease cleavage site (LeuGluValLeuPheGlnGlyPro), and a multiple cloning site that contains restriction enzyme sites found in many other Novagen expression vectors to facilitate insert transfer. An optional C-terminal S•Tag™ coding sequence is compatible with purification, detection, and quantification.

pET-47b(+) sequence landmarks	
T7 promoter	437–453
T7 transcription start	436
His•Tag coding sequence	341–359
Multiple cloning sites (*San*D I – *Avr* II)	131–315
S•Tag coding sequence	168–212
T7 terminator	26–73
lacI coding sequence	844–1923
pBR322 ori	3120
Kan coding sequence	3829–4641
f1 origin	4744–5191






```
LOCUS       pET-47b(+)  5203 bp    DNA   circular            17-MAY-2004
DEFINITION pET-47b(+)
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.informaxinc.com/
COMMENT     VNTDATE|333903583|
COMMENT     VNTDBDATE|338806447|
COMMENT     VNTNAME|pET-47b(+)|
COMMENT     VNTAUTHORNAME|Bob Novy|
COMMENT     VNTREPLTYPE|Plasmid
COMMENT     VNTEXTCHREPL|Bacteria
FEATURES             Location/Qualifiers
     rep_origin      complement(4744..5191)
                     /vntifkey="33"
                     /label=f1\ori
     misc_RNA        complement(3120..3675)
                     /vntifkey="53"
                     /label=RNA\II
     rep_origin      complement(3120..3120)
                     /vntifkey="33"
                     /label=ColE1,\pBR\ori
     misc_RNA        3560..3667
                     /vntifkey="53"
                     /label=RNA\I
     CDS             2498..2689
                     /vntifkey="4"
                     /label=ROP\gene
     CDS             844..1926
                     /vntifkey="4"
                     /label=lacI
                     /note="lac repressor"
     promoter        complement(437..453)
                     /vntifkey="30"
                     /label=T7\promoter
     protein_bind    complement(412..432)
                     /vntifkey="31"
                     /label=lacO
     terminator      complement(26..73)
                     /vntifkey="43"
                     /label=T7\terminator
     CDS             3829..4644
```

```
                     /vntifkey="4"
                     /label=Kan
     misc_feature    complement(213..230)
                     /vntifkey="21"
                     /label=Thrombin\site
     protein_bind    complement(168..212)
                     /vntifkey="31"
                     /label=S-Tag
     misc_binding    complement(341..359)
                     /vntifkey="20"
                     /label=His-Tag
     protein_bind    complement(310..332)
                     /vntifkey="31"
                     /label=3c
     primer          complement(300..336)
                     /vntifkey="27"
                     /label=S-SacII/Bam\+SanDI\-dcm
     CDS             complement(141..365)
                     /vntifkey="4"
                     /label=MCS\ORF


BASE COUNT     1256 a      1328 c      1394 g      1225 t
ORIGIN
        1 atccggatat agttcctcct ttcagcaaaa aacccctcaa gacccgttta gaggccccaa
       61 ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt
      121 tgtttagcag cctaggttaa ttaagcctcg agagcagcag aagtagagct gtccatgtgc
      181 tggcgttcga atttagcagc agcggtttct ttactaccgc gtggcaccag agcgagctct
      241 gcggccgcaa gcttgtcgac ggacgtcggg cgcgccaagg cctgtacaga attcggatcc
      301 tggtacccgg gtccctgaaa gaggacttca agagccgcgg agtgatggtg gtggtgatgt
      361 gccatatgta tatctccttc ttaaagttaa acaaaattat ttctagaggg gaattgttat
      421 ccgctcacaa ttcccctata gtgagtcgta ttaatttcgc gggatcgaga tcgatctcga
      481 tcctctacgc cggacgcatc gtggccggca tcaccggcgc cacaggtgcg gttgctggcg
      541 cctatatcgc cgacatcacc gatggggaag atcgggctcg ccacttcggg ctcatgagcg
      601 cttgtttcgg cgtgggtatg gtggcaggcc ccgtggccgg gggactgttg ggcgccatct
      661 ccttgcatgc accattcctt gcggcggcgg tgctcaacgg cctcaaccta ctactgggct
      721 gcttcctaat gcaggagtcg cataagggag agcgtcgaga tcccggacac catcgaatgg
      781 cgcaaaacct ttcgcggtat ggcatgatag cgcccggaag agagtcaatt cagggtggtg
      841 aatgtgaaac cagtaacgtt atacgatgtc gcagagtatg ccggtgtctc ttatcagacc
      901 gtttcccgcg tggtgaacca ggccagccac gtttctgcga aaacgcggga aaaagtggaa
      961 gcggcgatgg cggagctgaa ttacattccc aaccgcgtgg cacaacaact ggcgggcaaa
     1021 cagtcgttgc tgattggcgt tgccacctcc agtctggccc tgcacgcgcc gtcgcaaatt
     1081 gtcgcggcga ttaaatctcg cgccgatcaa ctgggtgcca gcgtggtggt gtcgatggta
     1141 gaacgaagcg gcgtcgaagc ctgtaaagcg gcggtgcaca atcttctcgc gcaacgcgtc
     1201 agtgggctga tcattaacta tccgctggat gaccaggatg ccattgctgt ggaagctgcc
     1261 tgcactaatg ttccggcgtt atttcttgat gtctctgacc agacacccat caacagtatt
     1321 attttctccc atgaagacgg tacgcgactg ggcgtggagc atctggtcgc attgggtcac
     1381 cagcaaatcg cgctgttagc gggcccatta agttctgtct cggcgcgtct gcgtctggct
     1441 ggctggcata aatatctcac tcgcaatcaa attcagccga tagcggaacg ggaaggcgac
     1501 tggagtgcca tgtccggttt tcaacaaacc atgcaaatgc tgaatgaggg catcgttccc
     1561 actgcgatgc tggttgccaa cgatcagatg gcgctgggcg caatgcgcgc cattaccgag
     1621 tccgggctgc gcgttggtgc ggacatctcg gtagtgggat acgacgatac cgaagacagc
     1681 tcatgttata tcccgccgtt aaccaccatc aaacaggatt ttcgcctgct ggggcaaacc
     1741 agcgtggacc gcttgctgca actctctcag ggccaggcgg tgaagggcaa tcagctgttg
     1801 cccgtctcac tggtgaaaag aaaaaccacc ctggcgccca atacgcaaac cgcctctccc
     1861 cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg
     1921 cagtgagcgc aacgcaatta atgtaagtta gctcactcat taggcaccgg gatctcgacc
     1981 gatgcccttg agagccttca acccagtcag ctccttccgg tgggcgcggg gcatgactag
     2041 catgatcgtg ctcctgtcgt tgaggacccg gctaggctgg cggggttgcc ttactggtta
     2101 gcagaatgaa tcaccgatac gcgagcgaac gtgaagcgac tgctgctgca aaacgtctgc
     2161 gacctgagca acaacatgaa tggtcttcgg tttccgtgtt tcgtaaagtc tggaaacgcg
```

```
2221 gaagtcagcg ccctgcacca ttatgttccg gatctgcatc gcaggatgct gctggctacc
2281 ctgtggaaca cctacatctg tattaacgaa gcgctggcat tgaccctgag tgatttttct
2341 ctggtcccgc cgcatccata ccgccagttg tttaccctca caacgttcca gtaaccgggc
2401 atgttcatca tcagtaaccc gtatcgtgag catcctctct cgtttcatcg gtatcattac
2461 ccccatgaac agaaatcccc cttacacgga ggcatcagtg accaaacagg aaaaaaccgc
2521 ccttaacatg gcccgcttta tcagaagcca gacattaacg cttctggaga aactcaacga
2581 gctggacgcg gatgaacagg cagacatctg tgaatcgctt cacgaccacg ctgatgagct
2641 ttaccgcagc tgcctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct
2701 cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg
2761 cgcgtcagcg ggtgttggcg ggtgtcgggg cgcagccatg acccagtcac gtagcgatag
2821 cggagtgtat actggcttaa ctatgcggca tcagagcaga ttgtactgag agtgcaccat
2881 atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc aggcgctctt
2941 ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag
3001 ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca
3061 tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt
3121 tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc
3181 gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct
3241 ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg
3301 tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca
3361 agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact
3421 atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta
3481 acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta
3541 actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct
3601 tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt
3661 tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga
3721 tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca
3781 tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat gagccatatt
3841 caacgggaaa cgtcttgctc taggccgcga ttaaattcca acatggatgc tgatttatat
3901 gggtataaat gggctcgcga taatgtcggg caatcaggtg cgacaatcta tcgattgtat
3961 gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt tgccaatgat
4021 gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct tccgaccatc
4081 aagcatttta tccgtactcc tgatgatgca tggttactca ccactgcgat ccccggcaaa
4141 acagcattcc aggtattaga agaatatcct gattcaggtg aaaatattgt tgatgcgctg
4201 gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt taacagtgat
4261 cgcgtatttc gtctcgctca ggcgcaatca cgaatgaata acggtttggt tgatgcgagt
4321 gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaaga aatgcataaa
4381 cttttgccat tctcaccgga ttcagtcgtc actcatggtg atttctcact tgataacctt
4441 atttttgacg aggggaaatt aataggttgt attgatgttg gacgagtcgg aatcgcagac
4501 cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc ttcattacag
4561 aaacggcttt ttcaaaaata tggtattgat aatcctgata tgaataaatt gcagtttcat
4621 ttgatgctcg atgagttttt ctaagaatta attcatgagc ggatacatat ttgaatgtat
4681 ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgaaat
4741 tgtaaacgtt aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt
4801 taaccaatag gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg
4861 gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt
4921 caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat caccctaatc
4981 aagtttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag ggagcccccg
5041 atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa
5101 aggagcgggc gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc
5161 cgccgcgctt aatgcgccgc tacagggcgc gtcccattcg cca
```

Source: http://www.merck-chemicals.co.uk/

pET-48b(+) Vector Sequence

The pET-48b(+) vector is designed for cloning and high-level expression of target proteins fused with the 109aa Trx•Tag™ thioredoxin protein that is cleavable with the human rhinovirus (HRV) 3C protease. Cloning sites are available for producing fusion proteins also containing a cleavable N- terminal His•Tag® sequence for detection and purification. The plasmid contains a strong T7*lac* promoter, optimized RBS, the coding sequence for the rhinovirus 3C protease cleavage site (LeuGluValLeuPheGlnGlyPro) for N-terminal fusion tag removal and a multiple cloning site that contains restriction enzyme sites found in many other Novagen expression vectors to facilitate insert transfer. An optional C-terminal S•Tag™ coding sequence is compatible with purification, detection, and quantification.

Thioredoxin: A ubiquitous protein with two redox-active half-cysteine residues. Functions as a general disulfide reductant. Works efficiently on model compounds (e.g. Ellman's reagent) and protein disulfides. Plays an essential role in cellular protection against oxidative stress and cell death. Inhibits apoptosis in Jurkat T cells and human PBL.

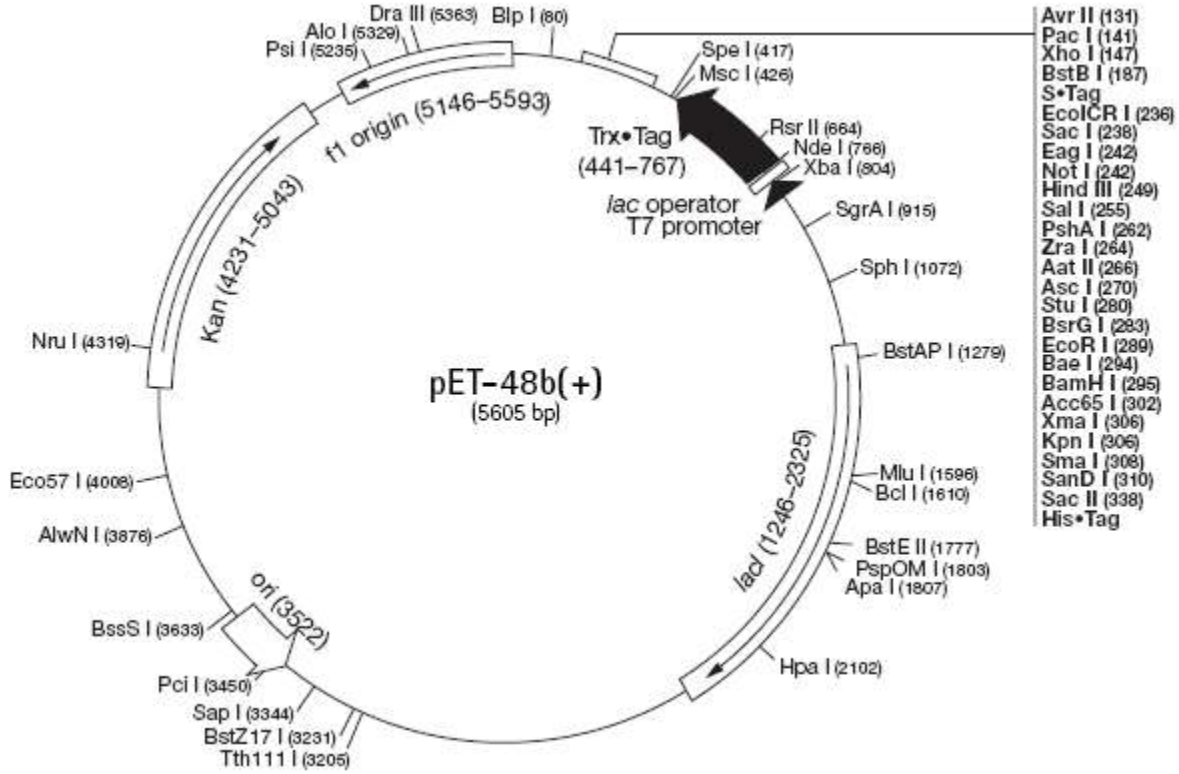
Here: *E.coli* thioredoxin (trxA, SwissProt accession number P0AA27) (LaVallie *et al.* 1993).

>sp|P0AA27|THIO_ECO57
MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVA
KLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

Thioredoxin is highly conserved amongst bacteria. The active domain is similar to some human proteins, such as mitochondrial thioredoxin (see alignment below) and protein disulfide isomerise. (Alignment created with uniprot alignment).

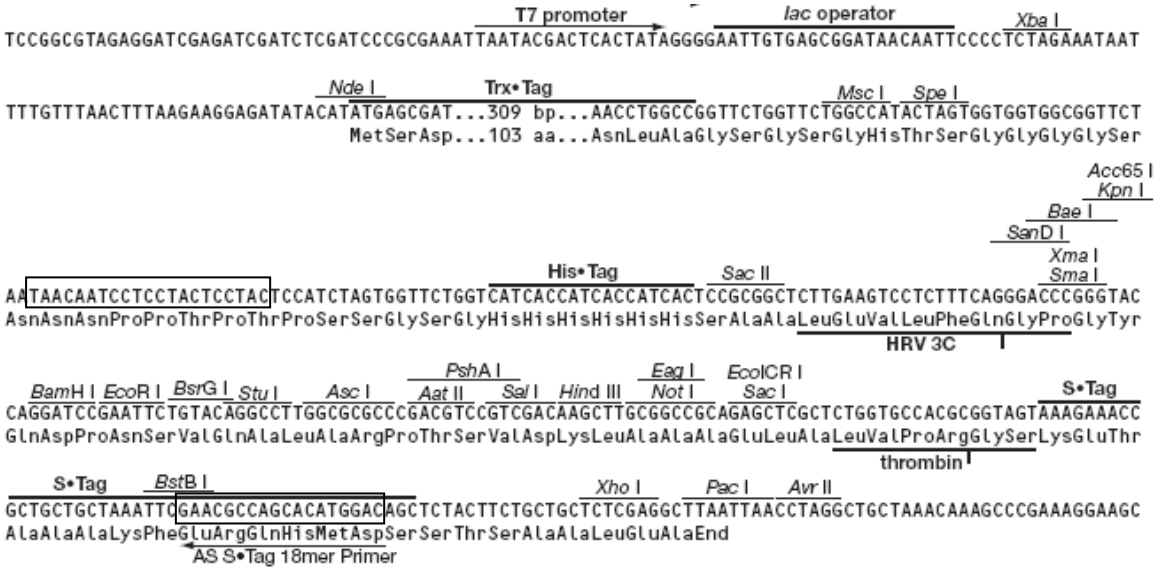
First line: *E. coli* Trx (P0AA27)
Second line: Human mitochondrial thioredoxin (Q99757)

```
MAQRLLLRRFLASVISRKPSQGQWPPLTSRALQTPQCSPGGLTVTPNPARTIYTTRISLT  60
-------------------------------------------------------MSDK  4
                                                           :* .

TFNIQDGPDFQDRVVNSETPVVVDFHAQWCGPCKILGPRLEKMVAKQHGKVVMAKVDIDD  120
IIHLTDDS-FDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQ  63
 ::: *.. *:  *:::: .::*** *:*******::.* *:::. : :**:.:**::**:

HTDLAIEYEVSAVPTVLAMKNGDVVDKFVGIKDEDQLEAFLKKLIG  166
NPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA  109
:.. * :* : .:**:* :***:*. . **  .:.**: **.  :.
```



pET-48b(+) sequence landmarks

T7 promoter	839–855
T7 transcription start	838
Trx•Tag coding sequence	441–767
His•Tag coding sequence	342–359
Multiple cloning sites (*San*D I – *Avr* II)	131–315
S•Tag coding sequence	168–212
T7 terminator	26–73
lacI coding sequence	1246–2325
pBR322 ori	3522
Kan coding sequence	4231–5043
f1 origin	5146–5593




```
LOCUS       pET-48b(+)    5605 bp    DNA    circular            17-MAY-2004
DEFINITION  pET-48b(+)
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.informaxinc.com/
COMMENT     VNTDATE|333905404|
COMMENT     VNTDBDATE|338806447|
COMMENT     VNTNAME|pET-48b(+)|
COMMENT     VNTAUTHORNAME|Bob Novy|
COMMENT     VNTREPLTYPE|Plasmid
COMMENT     VNTEXTCHREPL|Bacteria
FEATURES             Location/Qualifiers
     rep_origin      complement(5146..5593)
                     /vntifkey="33"
                     /label=f1\ori
     misc_RNA        complement(3522..4077)
                     /vntifkey="53"
                     /label=RNA\II
     rep_origin      complement(3522..3522)
                     /vntifkey="33"
                     /label=ColE1,\pBR\ori
     misc_RNA        3962..4069
                     /vntifkey="53"
                     /label=RNA\I
     CDS             2900..3091
                     /vntifkey="4"
                     /label=ROP\gene
     CDS             1246..2328
                     /vntifkey="4"
                     /label=lacI
                     /note="lac repressor"
     promoter        complement(839..855)
                     /vntifkey="30"
                     /label=T7\promoter
     protein_bind    complement(814..834)
                     /vntifkey="31"
                     /label=lacO
     protein_bind    complement(342..359)
                     /vntifkey="31"
                     /label=His-Tag
     terminator      complement(26..73)
```

```
                     /vntifkey="43"
                     /label=T7\terminator
     CDS             4231..5046
                     /vntifkey="4"
                     /label=Kan
     misc_feature    complement(213..230)
                     /vntifkey="21"
                     /label=Thrombin\site
     protein_bind    complement(168..212)
                     /vntifkey="31"
                     /label=S-Tag
     CDS             complement(441..767)
                     /vntifkey="4"
                     /label=Trx-Tag
     protein_bind    complement(310..332)
                     /vntifkey="31"
                     /label=3c
     primer          complement(300..336)
                     /vntifkey="27"
                     /label=S-SacII/Bam\+SanDI\-dcm
     misc_feature    complement(234..315)
                     /vntifkey="21"
                     /label=MCS
     primer          complement(374..421)
                     /vntifkey="27"
                     /label=S-Tag\Spacer-Spe\I

BASE COUNT     1355 a      1434 c      1492 g      1324 t
ORIGIN
        1 atccggatat agttcctcct ttcagcaaaa aacccctcaa gacccgttta gaggccccaa
       61 ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt
      121 tgtttagcag cctaggttaa ttaagcctcg agagcagcag aagtagagct gtccatgtgc
      181 tggcgttcga atttagcagc agcggtttct ttactaccgc gtggcaccag agcgagctct
      241 gcggccgcaa gcttgtcgac ggacgtcggg cgcgccaagg cctgtacaga attcggatcc
      301 tggtacccgg gtccctgaaa gaggacttca agagccgcgg agtgatggtg atggtgatga
      361 ccagaaccac tagatggagt aggagtagga ggattgttat tagaaccgcc accaccacta
      421 gtatggccag aaccagaacc ggccaggtta gcgtcgagga actctttcaa ctgaccttta
      481 gacagtgcac ccactttggt tgccgccact tcaccgtttt tgaacagcag cagagtcggg
      541 ataccacgga tgccatattt cggcgcagtg ccagggtttt gatcgatgtt cagttttgca
      601 acggtcagtt tgccctgata ttcgtcagcg atttcatcca gaatcggggc gatcattttg
      661 cacggaccgc accactctgc ccagaaatcg acgaggatcg ccccgtccgc tttgagtaca
      721 tccgtgtcaa aactgtcgtc agtcaggtga ataattttat cgctcatatg tatatctcct
      781 tcttaaagtt aaacaaaatt atttctagag gggaattgtt atccgctcac aattccccta
      841 tagtgagtcg tattaatttc gcgggatcga gatcgatctc gatcctctac gccggacgca
      901 tcgtggccgg catcaccggc gccacaggtg cggttgctgg cgcctatatc gccgacatca
      961 ccgatgggga agatcgggct cgccacttcg ggctcatgag cgcttgtttc ggcgtgggta
     1021 tggtggcagg ccccgtggcc gggggactgt tgggcgccat ctccttgcat gcaccattcc
     1081 ttgcggcggc ggtgctcaac ggcctcaacc tactactggg ctgcttccta atgcaggagt
     1141 cgcataaggg agagcgtcga gatcccggac accatcgaat ggcgcaaaac ctttcgcggt
     1201 atggcatgat agcgcccgga agagagtcaa ttcagggtgg tgaatgtgaa accagtaacg
     1261 ttatacgatg tcgcagagta tgccggtgtc tcttatcaga ccgtttcccg cgtggtgaac
     1321 caggccagcc acgtttctgc gaaaacgcgg gaaaaagtgg aagcggcgat ggcggagctg
     1381 aattacattc ccaaccgcgt ggcacaacaa ctggcgggca aacagtcgtt gctgattggc
     1441 gttgccacct ccagtctggc cctgcacgcg ccgtcgcaaa ttgtcgcggc gattaaatct
     1501 cgcgccgatc aactgggtgc cagcgtggtg gtgtcgatgg tagaacgaag cggcgtcgaa
     1561 gcctgtaaag cggcggtgca caatcttctc gcgcaacgcg tcagtgggct gatcattaac
     1621 tatccgctgg atgaccagga tgccattgct gtggaagctg cctgcactaa tgttccggcg
     1681 ttatttcttg atgtctctga ccagacaccc atcaacagta ttatttctc ccatgaagac
     1741 ggtacgcgac tgggcgtgga gcatctggtc gcattgggtc accagcaaat cgcgctgtta
     1801 gcgggcccat taagttctgt ctcggcgcgt ctgcgtctgg ctggctggca taaatatctc
     1861 actcgcaatc aaattcagcc gatagcggaa cgggaaggcg actggagtgc catgtccggt
```

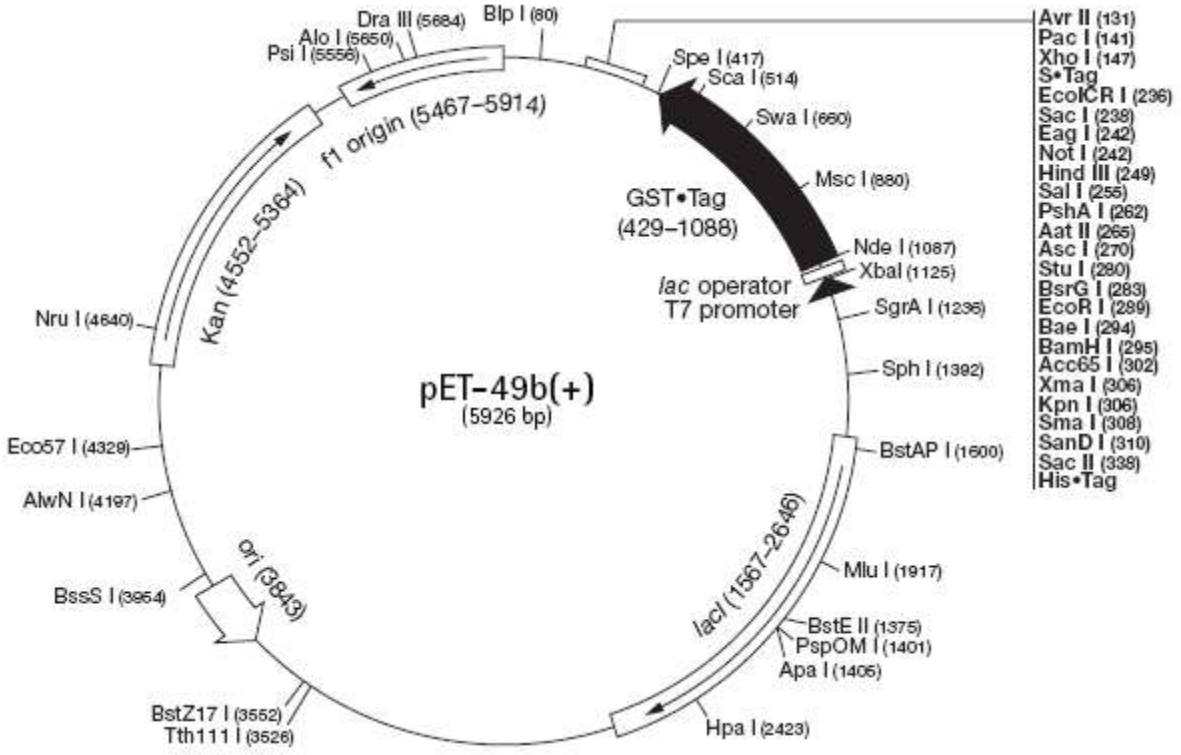
1921 tttcaacaaa ccatgcaaat gctgaatgag ggcatcgttc ccactgcgat gctggttgcc
1981 aacgatcaga tggcgctggg cgcaatgcgc gccattaccg agtccgggct gcgcgttggt
2041 gcggacatct cggtagtggg atacgacgat accgaagaca gctcatgtta tatcccgccg
2101 ttaaccacca tcaaacagga ttttcgcctg ctggggcaaa ccagcgtgga ccgcttgctg
2161 caactctctc agggccaggc ggtgaagggc aatcagctgt tgcccgtctc actggtgaaa
2221 agaaaaacca ccctggcgcc caatacgcaa accgcctctc cccgcgcgtt ggccgattca
2281 ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc gcaacgcaat
2341 taatgtaagt tagctcactc attaggcacc gggatctcga ccgatgccct tgagagcctt
2401 caacccagtc agctccttcc ggtgggcgcg gggcatgact agcatgatcg tgctcctgtc
2461 gttgaggacc cggctaggct ggcggggttg ccttactggt tagcagaatg aatcaccgat
2521 acgcgagcga acgtgaagcg actgctgctg caaaacgtct gcgacctgag caacaacatg
2581 aatggtcttc ggtttccgtg tttcgtaaag tctggaaacg cggaagtcag cgccctgcac
2641 cattatgttc cggatctgca tcgcaggatg ctgctggcta ccctgtggaa cacctacatc
2701 tgtattaacg aagcgctggc attgaccctg agtgattttt ctctggtccc gccgcatcca
2761 taccgccagt tgtttaccct cacaacgttc cagtaaccgg gcatgttcat catcagtaac
2821 ccgtatcgtg agcatcctct ctcgtttcat cggtatcatt accccccatga acagaaatcc
2881 cccttacacg gaggcatcag tgaccaaaca ggaaaaaacc gcccttaaca tggcccgctt
2941 tatcagaagc cagacattaa cgcttctgga gaaactcaac gagctggacg cggatgaaca
3001 ggcagacatc tgtgaatcgc ttcacgacca cgctgatgag ctttaccgca gctgcctcgc
3061 gcgtttcggt gatgacggtg aaaacctctg acacatgcag ctcccggaga cggtcacagc
3121 ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag cgggtgttgg
3181 cgggtgtcgg ggcgcagcca tgacccagtc acgtagcgat agcggagtgt atactggctt
3241 aactatgcgg catcagagca gattgtactg agagtgcacc atatatgcgg tgtgaaatac
3301 cgcacagatg cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc tcgctcactg
3361 actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa
3421 tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc
3481 aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc
3541 ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat
3601 aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc
3661 cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct
3721 cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg
3781 aacccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc
3841 cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga
3901 ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa
3961 ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta
4021 gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttgtt tgcaagcagc
4081 agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg
4141 acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgaacaat aaaactgtct
4201 gcttacataa acagtaatac aaggggtgtt atgagccata ttcaacggga aacgtcttgc
4261 tctaggccgc gattaaattc caacatggat gctgatttat atgggtataa atgggctcgc
4321 gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt atgggaagcc cgatgcgcca
4381 gagttgtttc tgaaacatgg caaaggtagc gttgccaatg atgttacaga tgagatggtc
4441 agactaaact ggctgacgga atttatgcct cttccgacca tcaagcattt tatccgtact
4501 cctgatgatg catggttact caccactgcg atccccggca aaacagcatt ccaggtatta
4561 gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc tggcagtgtt cctgcgccgg
4621 ttgcattcga ttcctgtttg taattgtcct tttaacagtg atcgcgtatt tcgtctcgct
4681 caggcgcaat cacgaatgaa taacggtttg gttgatgcga gtgattttga tgacgagcgt
4741 aatggctggc ctgttgaaca agtctggaaa gaaatgcata aacttttgcc attctcaccg
4801 gattcagtcg tcactcatgg tgatttctca cttgataacc ttatttttga cgaggggaaa
4861 ttaataggtt gtattgatgt tggacgagtc ggaatcgcag accgatacca ggatcttgcc
4921 atcctatgga actgcctcgg tgagttttct ccttcattac agaaacggct ttttcaaaaa
4981 tatggtattg ataatcctga tatgaataaa ttgcagtttc atttgatgct cgatgagttt
5041 ttctaagaat taattcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat
5101 aggggttccg cgcacatttc cccgaaaagt gccacctgaa attgtaaacg ttaatatttt
5161 gttaaaattc gcgttaaatt tttgttaaat cagctcattt tttaaccaat aggccgaaat
5221 cggcaaaatc ccttataaat caaaagaata gaccgagata gggttgagtg ttgttccagt
5281 ttggaacaag agtccactat taaagaacgt ggactccaac gtcaaagggc gaaaaaccgt
5341 ctatcagggc gatggcccac tacgtgaacc atcaccctaa tcaagttttt tggggtcgag
5401 gtgccgtaaa gcactaaatc ggaaccctaa agggagcccc cgatttagag cttgacgggg
5461 aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc
5521 gctggcaagt gtagcggtca cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc

```
5581 gctacagggc gcgtcccatt cgcca
```

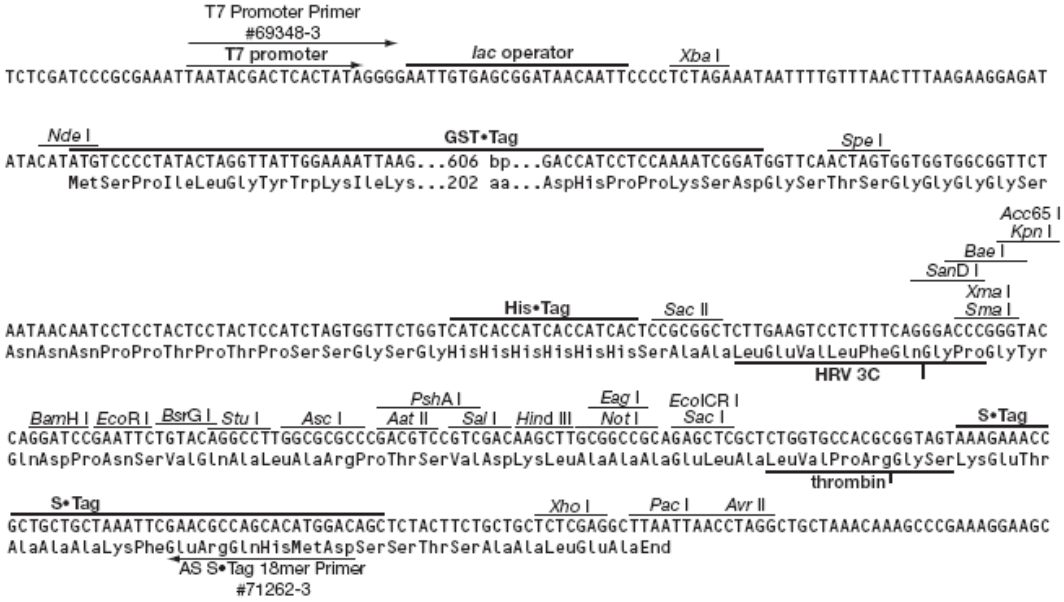
Source: http://www.merck-chemicals.co.uk/

pET-49b(+) Vector Sequence

The pET-49b(+) vector is designed for cloning and high-level expression of target proteins fused with the 220 aa schistosomal glutathione-S-transferase (GST•Tag™) protein that is cleavable with the human rhinovirus (HRV) 3C protease. Schistosomal glutathione-S-transferase (GST) is commonly used as a fusion partner when expressing proteins in *E. coli*. The GST•Tag sequence has been reported to enhance the production and in some cases the solubility of its fusion partners. When expressed in a soluble, properly folded form, GST•Tag fusion proteins can be purified with immobilized glutathione. Gentle elution is achieved with buffers containing reduced glutathione. Quantification of soluble GST fusions is also possible by assaying the glutathione-S-transferase activity. Cloning sites are available for producing fusion proteins also containing a cleavable N- terminal His•Tag® sequence for detection and purification. The plasmid contains a strong T7*lac* promoter, optimized RBS, the coding sequence for the HRV 3C protease cleavage site (LeuGluValLeuPheGlnGlyPro) for N-terminal fusion tag removal and a multiple cloning site that contains restriction enzyme sites found in many other Novagen expression vectors to facilitate insert transfer. An optional C-terminal S•Tag™ coding sequence is compatible with purification, detection, and quantification.



	Cat. No.
pET-49b(+) DNA	71463-3
pET-49b(+) sequence landmarks	
T7 promoter	1160–1176
T7 transcription start	1159
GST•Tag coding sequence	429–1088
His•Tag coding sequence	342–359
Multiple cloning sites (*San*D I – *Avr* II)	131–315
S•Tag coding sequence	168–212
T7 terminator	26–73
lacI coding sequence	1567–2646
pBR322 ori	3843
Kan coding sequence	4552–5364
f1 origin	5467–5914



pET-49b(+) cloning/expression region

```
LOCUS       pET-49b(+)    5926 bp    DNA    circular            17-MAY-2004
DEFINITION  pET-49b(+).
SOURCE
  ORGANISM
COMMENT     This file is created by Vector NTI
            http://www.informaxinc.com/
COMMENT     VNTDATE|333905385|
COMMENT     VNTDBDATE|338806447|
COMMENT     VNTNAME|pET-49b(+)|
COMMENT     VNTAUTHORNAME|Bob Novy|
COMMENT     VNTREPLTYPE|Plasmid
COMMENT     VNTEXTCHREPL|Bacteria
FEATURES             Location/Qualifiers
     rep_origin      complement(5467..5914)
                     /vntifkey="33"
                     /label=f1\ori
     misc_RNA        complement(3843..4398)
```

```
                     /vntifkey="53"
                     /label=RNA\II
     rep_origin      complement(3843..3843)
                     /vntifkey="33"
                     /label=ColE1,\pBR\ori
     misc_RNA        4283..4390
                     /vntifkey="53"
                     /label=RNA\I
     CDS             3221..3412
                     /vntifkey="4"
                     /label=ROP\gene
     CDS             1567..2649
                     /vntifkey="4"
                     /label=lacI
                     /note="lac repressor"
     promoter        complement(1160..1176)
                     /vntifkey="30"
                     /label=T7\promoter
     protein_bind    complement(1135..1155)
                     /vntifkey="31"
                     /label=lacO
     CDS             complement(429..1088)
                     /vntifkey="4"
                     /label=GST
     protein_bind    complement(342..359)
                     /vntifkey="31"
                     /label=His-Tag
     terminator      complement(26..73)
                     /vntifkey="43"
                     /label=T7\terminator
     CDS             4552..5367
                     /vntifkey="4"
                     /label=Kan
     misc_feature    complement(213..230)
                     /vntifkey="21"
                     /label=Thrombin\site
     protein_bind    complement(168..212)
                     /vntifkey="31"
                     /label=S-Tag
     protein_bind    complement(310..332)
                     /vntifkey="31"
                     /label=3c
     primer          complement(300..336)
                     /vntifkey="27"
                     /label=S-SacII/Bam\+SanDI\-dcm
     primer          complement(374..421)
                     /vntifkey="27"
                     /label=S-Tag\Spacer-Spe\I
BASE COUNT     1485 a      1486 c      1520 g      1435 t
ORIGIN
        1 atccggatat agttcctcct ttcagcaaaa aacccctcaa gacccgttta gaggccccaa
       61 ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt
      121 tgtttagcag cctaggttaa ttaagcctcg agagcagcag aagtagagct gtccatgtgc
      181 tggcgttcga atttagcagc agcggtttct ttactaccgc gtggcaccag agcgagctct
      241 gcggccgcaa gcttgtcgac ggacgtcggg cgcgccaagg cctgtacaga attcggatcc
      301 tggtacccgg gtccctgaaa gaggacttca agagccgcgg agtgatggtg atggtgatga
      361 ccagaaccac tagatggagt aggagtagga ggattgttat tagaaccgcc accaccacta
      421 gttgaaccat ccgattttgg aggatggtcg ccaccaccaa acgtggcttg ccagccctgc
      481 aaaggccatg ctatatactt gctggatttc aagtacttat caatttgtgg gatagcttca
      541 atacgttttt taaaacaaac taattttggg aacgcatcca ggcacattgg gtccatgtat
      601 aaaacaacat caagagcgtc atacaacatg aagtcaggat gggttacatg atcaccattt
```

661 aaatatgttt tatgacataa acgatcttcg aacattttca gcatttcagg tagcttgcta
721 agaaaatcaa ctttgagagt ttcaaagtct ttactatatg caattctcga aacaccgtat
781 ctaatatcca aaaccgctcc ttcaagcatt gaaatctctg cacgctcttt tggacaacca
841 cccaacatgt tgtgcttgtc agctatataa cgtatgatgg ccatagactg tgttaattta
901 acatcaccat caatataata aggaagattg ggaaactcca aacccaattc aaacttttg
961 tttcgccatt tatcaccttc atcgcgctca tacaaatgct cttcatattt ttcttcaaga
1021 tattccaaaa gaagtcgagt gggttgcaca aggcccttaa ttttccaata acctagtata
1081 ggggacatat gtatatctcc ttcttaaagt taaacaaaat tatttctaga ggggaattgt
1141 tatccgctca caattcccct atagtgagtc gtattaattt cgcgggatcg agatcgatct
1201 cgatcctcta cgccggacgc atcgtggccg gcatcaccgg cgccacaggt gcggttgctg
1261 gcgcctatat cgccgacatc accgatgggg aagatcgggc tcgccacttc gggctcatga
1321 gcgcttgttt cggcgtgggt atggtggcag gccccgtggc cgggggactg ttgggcgcca
1381 tctccttgca tgcaccattc cttgcggcgg cggtgctcaa cggcctcaac ctactactgg
1441 gctgcttcct aatgcaggag tcgcataagg gagagcgtcg agatcccgga caccatcgaa
1501 tggcgcaaaa cctttcgcgg tatggcatga tagcgcccgg aagagagtca attcagggtg
1561 gtgaatgtga aaccagtaac gttatacgat gtcgcagagt atgccggtgt ctcttatcag
1621 accgtttccc gcgtggtgaa ccaggccagc cacgtttctg cgaaaacgcg ggaaaaagtg
1681 gaagcggcga tggcggagct gaattacatt cccaaccgcg tggcacaaca actggcgggc
1741 aaacagtcgt tgctgattgg cgttgccacc tccagtctgg ccctgcacgc gccgtcgcaa
1801 attgtcgcgg cgattaaatc tcgcgccgat caactgggtg ccagcgtggt ggtgtcgatg
1861 gtagaacgaa gcggcgtcga agcctgtaaa gcggcggtgc acaatcttct cgcgcaacgc
1921 gtcagtgggc tgatcattaa ctatccgctg gatgaccagg atgccattgc tgtggaagct
1981 gcctgcacta atgttccggc gttatttctt gatgtctctg accagacacc catcaacagt
2041 attattttct cccatgaaga cggtacgcga ctgggcgtgg agcatctggt cgcattgggt
2101 caccagcaaa tcgcgctgtt agcgggccca ttaagttctg tctcggcgcg tctgcgtctg
2161 gctggctggc ataaatatct cactcgcaat caaattcagc cgatagcgga acgggaaggc
2221 gactggagtg ccatgtccgg ttttcaacaa accatgcaaa tgctgaatga gggcatcgtt
2281 cccactgcga tgctggttgc caacgatcag atggcgctgg gcgcaatgcg cgccattacc
2341 gagtccgggc tgcgcgttgg tgcggacatc tcggtagtgg gatacgacga taccgaagac
2401 agctcatgtt atatcccgcc gttaaccacc atcaaacagg attttcgcct gctggggcaa
2461 accagcgtgg accgcttgct gcaactctct cagggccagg cggtgaaggg caatcagctg
2521 ttgcccgtct cactggtgaa aagaaaaacc accctggcgc ccaatacgca aaccgcctct
2581 ccccgcgcgt tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc
2641 gggcagtgag cgcaacgcaa ttaatgtaag ttagctcact cattaggcac cgggatctcg
2701 accgatgccc ttgagagcct tcaacccagt cagctccttc cggtgggcgc ggggcatgac
2761 tagcatgatc gtgctcctgt cgttgaggac ccggctaggc tggcggggtt gccttactgg
2821 ttagcagaat gaatcaccga tacgcgagcg aacgtgaagc gactgctgct gcaaaacgtc
2881 tgcgacctga gcaacaacat gaatggtctt cggtttccgt gtttcgtaaa gtctggaaac
2941 gcggaagtca gcgccctgca ccattatgtt ccggatctgc atcgcaggat gctgctggct
3001 accctgtgga acacctacat ctgtattaac gaagcgctgg cattgaccct gagtgatttt
3061 tctctggtcc cgccgcatcc ataccgccag ttgtttaccc tcacaacgtt ccagtaaccg
3121 ggcatgttca tcatcagtaa cccgtatcgt gagcatcctc tctcgtttca tcggtatcat
3181 taccccccatg aacagaaatc ccccttacac ggaggcatca gtgaccaaac aggaaaaaac
3241 cgcccttaac atggcccgct ttatcagaag ccagacatta acgcttctgg agaaactcaa
3301 cgagctggac gcggatgaac aggcagacat ctgtgaatcg cttcacgacc acgctgatga
3361 gctttaccgc agctgcctcg cgcgtttcgg tgatgacggt gaaaacctct gacacatgca
3421 gctcccggag acggtcacag cttgtctgta agcggatgcc gggagcagac aagcccgtca
3481 gggcgcgtca gcgggtgttg gcgggtgtcg gggcgcagcc atgacccagt cacgtagcga
3541 tagcggagtg tatactggct taactatgcg gcatcagagc agattgtact gagagtgcac
3601 catatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgct
3661 cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat
3721 cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga
3781 acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt
3841 ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt
3901 ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc
3961 gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa
4021 gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct
4081 ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatccggta
4141 actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca gcagccactg
4201 gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtggtggc
4261 ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta

```
4321 ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg
4381 gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt
4441 tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg
4501 tcatgaacaa taaaactgtc tgcttacata aacagtaata caaggggtgt tatgagccat
4561 attcaacggg aaacgtcttg ctctaggccg cgattaaatt ccaacatgga tgctgattta
4621 tatgggtata aatgggctcg cgataatgtc gggcaatcag gtgcgacaat ctatcgattg
4681 tatgggaagc ccgatgcgcc agagttgttt ctgaaacatg gcaaaggtag cgttgccaat
4741 gatgttacag atgagatggt cagactaaac tggctgacgg aatttatgcc tcttccgacc
4801 atcaagcatt ttatccgtac tcctgatgat gcatggttac tcaccactgc gatccccggc
4861 aaaacagcat tccaggtatt agaagaatat cctgattcag gtgaaaatat tgttgatgcg
4921 ctggcagtgt tcctgcgccg gttgcattcg attcctgttt gtaattgtcc ttttaacagt
4981 gatcgcgtat ttcgtctcgc tcaggcgcaa tcacgaatga ataacggttt ggttgatgcg
5041 agtgattttg atgacgagcg taatggctgg cctgttgaac aagtctggaa agaaatgcat
5101 aaacttttgc cattctcacc ggattcagtc gtcactcatg gtgatttctc acttgataac
5161 cttatttttg acgaggggaa attaataggt tgtattgatg ttggacgagt cggaatcgca
5221 gaccgatacc aggatcttgc catcctatgg aactgcctcg gtgagttttc tccttcatta
5281 cagaaacggc tttttcaaaa atatggtatt gataatcctg atatgaataa attgcagttt
5341 catttgatgc tcgatgagtt tttctaagaa ttaattcatg agcggataca tatttgaatg
5401 tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga
5461 aattgtaaac gttaatattt tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt
5521 ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agaccgagat
5581 agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa
5641 cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcacccta
5701 atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaacccta aagggagccc
5761 ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc
5821 gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac
5881 acccgccgcg cttaatgcgc cgctacaggg cgcgtcccat tcgcca
```

Source: http://www.merck-chemicals.co.uk/

Construct sequences

Mature PPAD in pET48b (mPPAD/pET48)

Forward

```
M1-F_pET48-F      ATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGGCATTCC 60
PPAD              ------------------------------------------------------GCATTCC 7
                                                                        *******

M1-F_pET48-F      AGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAACGCTCTGCAG 119
PPAD              AGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAACGCTCTGCAG 67
                  ************************************************************

M1-F_pET48-F      CCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAGCTGGCCAAGA 179
PPAD              CCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAGCTGGCCAAGA 127
                  ************************************************************

M1-F_pET48-F      ACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTTATAACCCAGT 239
PPAD              ACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTTATAACCCAGT 187
                  ************************************************************

M1-F_pET48-F      ACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAAACTGACTCTT 299
PPAD              ACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAAACTGACTCTT 247
                  ************************************************************

M1-F_pET48-F      ACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAAGTAGGTCTCG 359
PPAD              ACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAAGTAGGTCTCG 307
                  ************************************************************

M1-F_pET48-F      TGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAATACGAAGCAC 419
PPAD              TGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAATACGAAGCAC 367
                  ************************************************************

M1-F_pET48-F      AATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGCAACTACATGA 479
PPAD              AATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGCAACTACATGA 427
                  ************************************************************

M1-F_pET48-F      CGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAACTCCTCTCTGT 539
PPAD              CGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAACTCCTCTCTGT 487
                  ************************************************************

M1-F_pET48-F      CTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACATCATGATGTGG 599
PPAD              CTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACATCATGATGTGG 547
                  ************************************************************

M1-F_pET48-F      TACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAGTATTTGGCAC 659
PPAD              TACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAGTATTTGGCAC 607
                  ************************************************************

M1-F_pET48-F      CGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCCCTGGAAG 719
PPAD              CGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCCCTGGAAG 667
                  ************************************************************

M1-F_pET48-F      ATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTACGAGGTATATC 779
PPAD              ATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTACGAGGTATATC 727
                  ************************************************************

M1-F_pET48-F      GCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAACAACAGGGTAT 839
PPAD              GCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAACAACAGGGTAT 787
                  ************************************************************

M1-F_pET48-F      TTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACGTCTATAAGACGG 899
PPAD              TTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACGTCTATAAGACGG 847
                  ************************************************************

M1-F_pET48-F      CAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACA------------  946
PPAD              CAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACACCTTGGTTAGGA  906
                  ************************************************************
```

Reverse

```
M1-R_pET48-R      ----------------------CAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCC 28
PPAD              TTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCC 660
                             ***********************************

M1-R_pET48-R      CTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTACGAG 88
PPAD              CTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTACGAG 720
                  ************************************************************

M1-R_pET48-R      GTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAACAAC 148
PPAD              GTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAACAAC 780
                  ************************************************************

M1-R_pET48-R      AGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACGTCTAT 208
PPAD              AGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACGTCTAT 840
                  ************************************************************

M1-R_pET48-R      AAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACACCTTGG 268
PPAD              AAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACACCTTGG 900
                  ************************************************************

M1-R_pET48-R      TTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTATCTCTAT 328
PPAD              TTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTATCTCTAT 960
                  ************************************************************

M1-R_pET48-R      ATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAAGCAGAT 388
PPAD              ATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAAGCAGAT 1020
                  ************************************************************

M1-R_pET48-R      GTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATCAATGGT 448
PPAD              GTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATCAATGGT 1080
                  ************************************************************

M1-R_pET48-R      TCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATACTTATAGC 508
PPAD              TCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATACTTATAGC 1140
                  ************************************************************

M1-R_pET48-R      TTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGACAATAGT 568
PPAD              TTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGACAATAGT 1200
                  ************************************************************

M1-R_pET48-R      GGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACGTGTATG 628
PPAD              GGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACGTGTATG 1260
                  ************************************************************

M1-R_pET48-R      AACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATGGTTCAAC 688
PPAD              AACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATGGTTCAAC 1320
                  ************************************************************

M1-R_pET48-R      GCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATATCGGATA 748
PPAD              GCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATATCGGATA 1380
                  ************************************************************

M1-R_pET48-R      AAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTAGCAGGG 808
PPAD              AAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTAGCAGGG 1440
                  ************************************************************

M1-R_pET48-R      ACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTGGTTGTT 868
PPAD              ACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTGGTTGTT 1500
                  ************************************************************

M1-R_pET48-R      GAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTCTGGTGCCA 928
PPAD              GAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA------------------ 1542
                  ****************************************
```

Full PPAD in pET48b (PPAD/pET48)

Forward

```
PPAD              ------------------------------------------------------------AA 2
F16-F_pET48-F     TCACCATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGAA 60
                                                                             **

PPAD              AAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAAT 62
F16-F_pET48-F     AAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAAT 120
                  ************************************************************

PPAD              CGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCA 122
F16-F_pET48-F     CGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCA 180
                  ************************************************************

PPAD              ACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGA 182
F16-F_pET48-F     ACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGA 240
                  ************************************************************

PPAD              ACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGA 242
F16-F_pET48-F     ACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGA 300
                  ************************************************************

PPAD              GCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGT 302
F16-F_pET48-F     GCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGT 360
                  ************************************************************

PPAD              TATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAA 362
F16-F_pET48-F     TATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAA 420
                  ************************************************************

PPAD              AACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAA 422
F16-F_pET48-F     AACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAA 480
                  ************************************************************

PPAD              AGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAA 482
F16-F_pET48-F     AGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAA 540
                  ************************************************************

PPAD              ATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGG 542
F16-F_pET48-F     ATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGG 600
                  ************************************************************

PPAD              CAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAA 602
F16-F_pET48-F     CAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAA 660
                  ************************************************************

PPAD              CTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACA 662
F16-F_pET48-F     CTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACA 720
                  ************************************************************

PPAD              TCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAA 722
F16-F_pET48-F     TCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAA 780
                  ************************************************************

PPAD              GTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCA 782
F16-F_pET48-F     GTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCA 840
                  ************************************************************

PPAD              AGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTA 842
F16-F_pET48-F     AGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTA 900
                  ************************************************************

PPAD              CGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 901
F16-F_pET48-F     CGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 960
                  ************************************************************

PPAD              ACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAAC 961
F16-F_pET48-F     ACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACA-------------- 1004
                  *******************************************
```

Reverse

```
F16-R_pET48-R      ----------------------------------------TGACAATCACCCTCAGCA 18
PPAD               AAGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCA 779
                                                           ******************

F16-R_pET48-R      CCAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAA 78
PPAD               CCAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAA 839
                   ************************************************************

F16-R_pET48-R      GTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCT 138
PPAD               GTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCT 899
                   ************************************************************

F16-R_pET48-R      GAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAA 198
PPAD               GAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAA 959
                   ************************************************************

F16-R_pET48-R      CGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAAC 258
PPAD               CGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAAC 1019
                   ************************************************************

F16-R_pET48-R      ACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTA 318
PPAD               ACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTA 1079
                   ************************************************************

F16-R_pET48-R      TCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGA 378
PPAD               TCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGA 1139
                   ************************************************************

F16-R_pET48-R      AGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTAT 438
PPAD               AGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTAT 1199
                   ************************************************************

F16-R_pET48-R      CAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATAC 498
PPAD               CAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATAC 1259
                   ************************************************************

F16-R_pET48-R      TTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGA 558
PPAD               TTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGA 1319
                   ************************************************************

F16-R_pET48-R      CAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTAC 618
PPAD               CAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTAC 1379
                   ************************************************************

F16-R_pET48-R      GTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATG 678
PPAD               GTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATG 1439
                   ************************************************************

F16-R_pET48-R      GTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATA 738
PPAD               GTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATA 1499
                   ************************************************************

F16-R_pET48-R      TCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGT 798
PPAD               TCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGT 1559
                   ************************************************************

F16-R_pET48-R      AGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCT 858
PPAD               AGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCT 1619
                   ************************************************************

F16-R_pET48-R      GGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTCT 918
PPAD               GGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA----------- 1668
                   ***********************************************
```

Full PPAD in pET47b (PPAD/pET47)

Forward

```
Ppad-F_T7       CATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGAAAAAGCTTTTA 120
PPAD            ----------------------------------------------AAAAAGCTTTTA 12
                                                              ************

Ppad-F_T7       CAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATCGCCCAAACG 180
PPAD            CAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATCGCCCAAACG 72
                ************************************************************

Ppad-F_T7       CAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAACGAGCATTC 240
PPAD            CAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAACGAGCATTC 132
                ************************************************************

Ppad-F_T7       CAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAACGCTCTGCA 300
PPAD            CAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAACGCTCTGCA 192
                ************************************************************

Ppad-F_T7       GCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAGCTGGCCAAG 360
PPAD            GCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAGCTGGCCAAG 252
                ************************************************************

Ppad-F_T7       AACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTTATAACCCAG 420
PPAD            AACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTTATAACCCAG 312
                ************************************************************

Ppad-F_T7       TACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAAACTGACTCT 480
PPAD            TACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAAACTGACTCT 372
                ************************************************************

Ppad-F_T7       TACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAAGTAGGTCTC 540
PPAD            TACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAAGTAGGTCTC 432
                ************************************************************

Ppad-F_T7       GTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAATACGAAGCA 600
PPAD            GTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAATACGAAGCA 492
                ************************************************************


Ppad-F_T7       CAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGCAACTACATG 660
PPAD            CAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGCAACTACATG 552
                ************************************************************

Ppad-F_T7       ACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAACTCCTCTCTG 720
PPAD            ACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAACTCCTCTCTG 612
                ************************************************************


Ppad-F_T7       TCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACATCATGATGTG 780
PPAD            TCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACATCATGATGTG 672
                ************************************************************

Ppad-F_T7       GTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAGTATTTGGCA 840
PPAD            GTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAGTATTTGGCA 732
                ************************************************************

Ppad-F_T7       CCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCCCTGGAA 900
PPAD            CCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAAGCCCTGGAA 792
                ************************************************************
```

Reverse

```
Ppad-R_s-tag-rev      CGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAAC 363
PPAD                  CGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAAC 1019
                      ************************************************************

Ppad-R_s-tag-rev      ACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTA 423
PPAD                  ACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTA 1079
                      ************************************************************

Ppad-R_s-tag-rev      TCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGA 483
PPAD                  TCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGA 1139
                      ************************************************************

Ppad-R_s-tag-rev      AGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTAT 543
PPAD                  AGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTAT 1199
                      ************************************************************

Ppad-R_s-tag-rev      CAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATAC 603
PPAD                  CAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATAC 1259
                      ************************************************************

Ppad-R_s-tag-rev      TTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGA 663
PPAD                  TTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGA 1319
                      ************************************************************

Ppad-R_s-tag-rev      CAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTAC 723
PPAD                  CAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTAC 1379
                      ************************************************************

Ppad-R_s-tag-rev      GTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATG 783
PPAD                  GTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATG 1439
                      ************************************************************

Ppad-R_s-tag-rev      GTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATA 843
PPAD                  GTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATA 1499
                      ************************************************************

Ppad-R_s-tag-rev      TCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGT 903
PPAD                  TCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGT 1559
                      ************************************************************

Ppad-R_s-tag-rev      AGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCT 963
PPAD                  AGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCT 1619
                      ************************************************************

Ppad-R_s-tag-rev      GGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTCT 1023
PPAD                  GGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA---------- 1668
                      **********************************************
```

Mutated PPADs in pET48b

C351A

Forward

```
C351_D-F_pET-48-F      CACCATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGAAA 60
PPAD                   ------------------------------------------------------ATGAAA 6
                                                                                *****

C351_D-F_pET-48-F      AAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATC 120
PPAD                   AAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATC 66
                       ************************************************************

C351_D-F_pET-48-F      GCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAA 180
PPAD                   GCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAA 126
                       ************************************************************

C351_D-F_pET-48-F      CGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAA 240
PPAD                   CGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAA 186
                       ************************************************************

C351_D-F_pET-48-F      CGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAG 300
PPAD                   CGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAG 246
                       ************************************************************

C351_D-F_pET-48-F      CTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTT 360
PPAD                   CTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTT 306
                       ************************************************************

C351_D-F_pET-48-F      ATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAA 420
PPAD                   ATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAA 366
                       ************************************************************

C351_D-F_pET-48-F      ACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAA 480
PPAD                   ACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAA 426
                       ************************************************************

C351_D-F_pET-48-F      GTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAA 540
PPAD                   GTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAA 486
                       ************************************************************

C351_D-F_pET-48-F      TACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGC 600
PPAD                   TACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGC 546
                       ************************************************************

C351_D-F_pET-48-F      AACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAAC 660
PPAD                   AACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAAC 606
                       ************************************************************

C351_D-F_pET-48-F      TCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACAT 720
PPAD                   TCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACAT 666
                       ************************************************************

C351_D-F_pET-48-F      CATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAG 780
PPAD                   CATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAG 726
                       ************************************************************

C351_D-F_pET-48-F      TATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAA 840
PPAD                   TATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAA 786
                       ************************************************************

C351_D-F_pET-48-F      GCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTAC 900
PPAD                   GCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTAC 846
                       ************************************************************

C351_D-F_pET-48-F      GAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAAC 960
PPAD                   GAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAAC 906
                       ************************************************************
```

Reverse

```
C351_D-RevCom      AAGTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATT 100
PPAD               AAGTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATT 900
                   ************************************************************

C351_D-RevCom      CTGAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTG 160
PPAD               CTGAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTG 960
                   ************************************************************

C351_D-RevCom      AACGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGA 220
PPAD               AACGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGA 1020
                   ************************************************************

C351_D-RevCom      ACACCTTGGTTAGGAACAGATGCCCTGCATGCCCGTACTCACGAGGTAGCGGATAAGGGC 280
PPAD               ACACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGC 1080
                   ******************************  **************************

C351_D-RevCom      TATCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATC 340
PPAD               TATCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATC 1140
                   ************************************************************

C351_D-RevCom      GAAGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGT 400
PPAD               GAAGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGT 1200
                   ************************************************************

C351_D-RevCom      ATCAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTAT 460
PPAD               ATCAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTAT 1260
                   ************************************************************

C351_D-RevCom      ACTTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCT 520
PPAD               ACTTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCT 1320
                   ************************************************************

C351_D-RevCom      GACAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTT 580
PPAD               GACAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTT 1380
                   ************************************************************

C351_D-RevCom      ACGTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCA 640
PPAD               ACGTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCA 1440
                   ************************************************************

C351_D-RevCom      TGGTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACA 700
PPAD               TGGTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACA 1500
                   ************************************************************

C351_D-RevCom      TATCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTA 760
PPAD               TATCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTA 1560
                   ************************************************************

C351_D-RevCom      GTAGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTT 820
PPAD               GTAGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTT 1620
                   ************************************************************

C351_D-RevCom      CTGGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCT 880
PPAD               CTGGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA--------- 1671
                   ************************************************
```

N297A

Forward

```
N297_1_F.       ---ACCATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGA 57
PPAD.           ------------------------------------------------------------A 1
                                                                             *

N297_1_F.       AAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAA 117
PPAD.           AAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAA 61
                ************************************************************

N297_1_F.       TCGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGC 177
PPAD.           TCGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGC 121
                ************************************************************

N297_1_F.       AACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACG 237
PPAD.           AACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACG 181
                ************************************************************

N297_1_F.       AACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAG 297
PPAD.           AACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAG 241
                ************************************************************

N297_1_F.       AGCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCG 357
PPAD.           AGCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCG 301
                ************************************************************

N297_1_F.       TTATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGA 417
PPAD.           TTATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGA 361
                ************************************************************

N297_1_F.       AAACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACA 477
PPAD.           AAACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACA 421
                ************************************************************

N297_1_F.       AAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCA 537
PPAD.           AAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCA 481
                ************************************************************

N297_1_F.       AATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTG 597
PPAD.           AATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTG 541
                ************************************************************

N297_1_F.       GCAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGA 657
PPAD.           GCAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGA 601
                ************************************************************

N297_1_F.       ACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC 717
PPAD.           ACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC 661
                ************************************************************

N297_1_F.       ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCA 777
PPAD.           ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCA 721
                ************************************************************

N297_1_F.       AGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACC 837
PPAD.           AGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACC 781
                ************************************************************

N297_1_F.       AAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGT 897
PPAD.           AAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGT 841
                ************************************************************

N297_1_F.       ACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGGCCTCTCTGATTCTGA 957
PPAD.           ACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 901
                ******************************************   **************

N297_1_F.       ACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACG 1017
PPAD.           ACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACG 961
                ******************************************
```

Reverse

```
N297_1_F.          -----------------AGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCAC 43
PPAD.              AACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCAC 659
                              ****************************************

N297_1_F.          ACATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGG 103
PPAD.              ACATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGG 718
                   ************************************************************

N297_1_F.          GCAAGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGC 163
PPAD.              GCAAGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGC 778
                   ************************************************************

N297_1_F.          ACCAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGA 223
PPAD.              ACCAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGA 838
                   ************************************************************

N297_1_F.          AGTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGGCCTCTCTGATTC 283
PPAD.              AGTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTC 898
                   ********************************************  ************

N297_1_F.          TGAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGA 343
PPAD.              TGAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGA 958
                   ************************************************************

N297_1_F.          ACGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAA 403
PPAD.              ACGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAA 1018
                   ************************************************************

N297_1_F.          CACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCT 463
PPAD.              CACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCT 1078
                   ************************************************************

N297_1_F.          ATCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCG 523
PPAD.              ATCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCG 1138
                   ************************************************************

N297_1_F.          AAGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTA 583
PPAD.              AAGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTA 1198
                   ************************************************************

N297_1_F.          TCAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATA 643
PPAD.              TCAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATA 1258
                   ************************************************************

N297_1_F.          CTTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTG 703
PPAD.              CTTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTG 1318
                   ************************************************************

N297_1_F.          ACAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTA 763
PPAD.              ACAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTA 1378
                   ************************************************************

N297_1_F.          CGTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCAT 823
PPAD.              CGTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCAT 1438
                   ************************************************************

N297_1_F.          GGTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACAT 883
PPAD.              GGTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACAT 1498
                   ************************************************************

N297_1_F.          ATCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAG 943
PPAD.              ATCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAG 1558
                   ************************************************************

N297_1_F.          TAGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTC 1003
PPAD.              TAGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTC 1618
                   ************************************************************

N297_1_F.          TGGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTC 1063
PPAD.              TGGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA---------- 1668
                   ************************************************
```

H236A

Forward

```
H236_1_F.       ------ATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGA 54
PPAD.           ------------------------------------------------------------A 1
                                                                             *

H236_1_F.       AAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAA 114
PPAD.           AAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAA 61
                ************************************************************

H236_1_F.       TCGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGC 174
PPAD.           TCGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGC 121
                ************************************************************

H236_1_F.       AACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACG 234
PPAD.           AACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACG 181
                ************************************************************

H236_1_F.       AACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAG 294
PPAD.           AACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAG 241
                ************************************************************

H236_1_F.       AGCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCG 354
PPAD.           AGCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCG 301
                ************************************************************

H236_1_F.       TTATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGA 414
PPAD.           TTATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGA 361
                ************************************************************

H236_1_F.       AAACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACA 474
PPAD.           AAACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACA 421
                ************************************************************

H236_1_F.       AAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCA 534
PPAD.           AAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCA 481
                ************************************************************

H236_1_F.       AATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTG 594
PPAD.           AATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTG 541
                ************************************************************

H236_1_F.       GCAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGA 654
PPAD.           GCAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGA 601
                ************************************************************

H236_1_F.       ACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC 714
PPAD.           ACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC 661
                ************************************************************

H236_1_F.       ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACGCCGTGGACTGTTGGGGCA 774
PPAD.           ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCA 721
                ****************************************   ****************

H236_1_F.       AGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACC 834
PPAD.           AGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACC 781
                ************************************************************

H236_1_F.       AAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGT 894
PPAD.           AAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGT 841
                ************************************************************

H236_1_F.       ACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 954
PPAD.           ACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 901
                ************************************************************
```

Reverse

```
H236_1_R_Com.       ---------------------------GGCGAATATATCAACGCCGTGGACTGTTGGGGC 33
PPAD.               CATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGC 720
                                               ***************   ***************

H236_1_R_Com.       AAGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCAC 93
PPAD.               AAGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCAC 780
                    ************************************************************

H236_1_R_Com.       CAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAG 153
PPAD.               CAAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAG 840
                    ********************** ***********************************

H236_1_R_Com.       TACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTG 213
PPAD.               TACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTG 900
                    ************************************************************

H236_1_R_Com.       AACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAAC 273
PPAD.               AACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAAC 960
                    ************************************************************

H236_1_R_Com.       GTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACA 333
PPAD.               GTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACA 1020
                    ************************************************************

H236_1_R_Com.       CCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTAT 393
PPAD.               CCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTAT 1080
                    ************************************************************

H236_1_R_Com.       CTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAA 453
PPAD.               CTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAA 1140
                    ************************************************************

H236_1_R_Com.       GCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATC 513
PPAD.               GCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATC 1200
                    ************************************************************

H236_1_R_Com.       AATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATACT 573
PPAD.               AATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATACT 1260
                    ************************************************************

H236_1_R_Com.       TATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGAC 633
PPAD.               TATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGAC 1320
                    ************************************************************

H236_1_R_Com.       AATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACG 693
PPAD.               AATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACG 1380
                    ************************************************************

H236_1_R_Com.       TGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATGG 753
PPAD.               TGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATGG 1440
                    ************************************************************

H236_1_R_Com.       TTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATAT 813
PPAD.               TTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATAT 1500
                    ************************************************************

H236_1_R_Com.       CGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTA 873
PPAD.               CGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTA 1560
                    ************************************************************

H236_1_R_Com.       GCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTG 933
PPAD.               GCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTG 1620
                    ************************************************************

H236_1_R_Com.       GTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTCTG 993
PPAD.               GTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA------------ 1668
                    ********************************************
```

D130A

Forward

```
D130_1_F.       CTCACCATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGA 60
PPAD.           ------------------------------------------------------------A 1
                                                                             *

D130_1_F.       AAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAA 120
PPAD.           AAAAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAA 61
                ************************************************************

D130_1_F.       TCGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGC 180
PPAD.           TCGCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGC 121
                ************************************************************

D130_1_F.       AACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACG 240
PPAD.           AACGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACG 181
                ************************************************************

D130_1_F.       AACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAG 300
PPAD.           AACGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAG 241
                ************************************************************

D130_1_F.       AGCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCG 360
PPAD.           AGCTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCG 301
                ************************************************************

D130_1_F.       TTATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGA 420
PPAD.           TTATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGA 361
                ************************************************************

D130_1_F.       AAACTGACTCTTACTGGACACGCGCCTATACCGGTTGGTTCGCAATGTACGATACGAACA 480
PPAD.           AAACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACA 421
                ********************** ***********************************

D130_1_F.       AAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCA 540
PPAD.           AAGTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCA 481
                ************************************************************

D130_1_F.       AATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTG 600
PPAD.           AATACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTG 541
                ************************************************************

D130_1_F.       GCAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGA 660
PPAD.           GCAACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGA 601
                ************************************************************

D130_1_F.       ACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC 720
PPAD.           ACTCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACAC 661
                ************************************************************

D130_1_F.       ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCA 780
PPAD.           ATCATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCA 721
                ************************************************************

D130_1_F.       AGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACC 840
PPAD.           AGTATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACC 781
                ************************************************************

D130_1_F.       AAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGT 900
PPAD.           AAGCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGT 841
                ************************************************************

D130_1_F.       ACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 960
PPAD.           ACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGA 901
                ************************************************************

D130_1_F.       ACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACG 1019
PPAD.           ACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACG 961
                ************************************************************
```

Reverse

```
D130_1_R.       GTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCT 256
PPAD.           GTACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCT 899
                ************************************************************

D130_1_R.       GAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAA 316
PPAD.           GAACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAA 959
                ************************************************************

D130_1_R.       CGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAAC 376
PPAD.           CGTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAAC 1019
                ************************************************************

D130_1_R.       ACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTA 436
PPAD.           ACCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTA 1079
                ************************************************************

D130_1_R.       TCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGA 496
PPAD.           TCTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGA 1139
                ************************************************************

D130_1_R.       AGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTAT 556
PPAD.           AGCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTAT 1199
                ************************************************************

D130_1_R.       CAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATAC 616
PPAD.           CAATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATAC 1259
                ************************************************************

D130_1_R.       TTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGA 676
PPAD.           TTATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGA 1319
                ************************************************************

D130_1_R.       CAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTAC 736
PPAD.           CAATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTAC 1379
                ************************************************************

D130_1_R.       GTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATG 796
PPAD.           GTGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATG 1439
                ************************************************************

D130_1_R.       GTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATA 856
PPAD.           GTTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATA 1499
                ************************************************************

D130_1_R.       TCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGT 916
PPAD.           TCGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGT 1559
                ************************************************************

D130_1_R.       AGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCT 976
PPAD.           AGCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCT 1619
                ************************************************************

D130_1_R.       GGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTCT 1036
PPAD.           GGTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA---------- 1668
                ********************************************
```

D238A

Forward

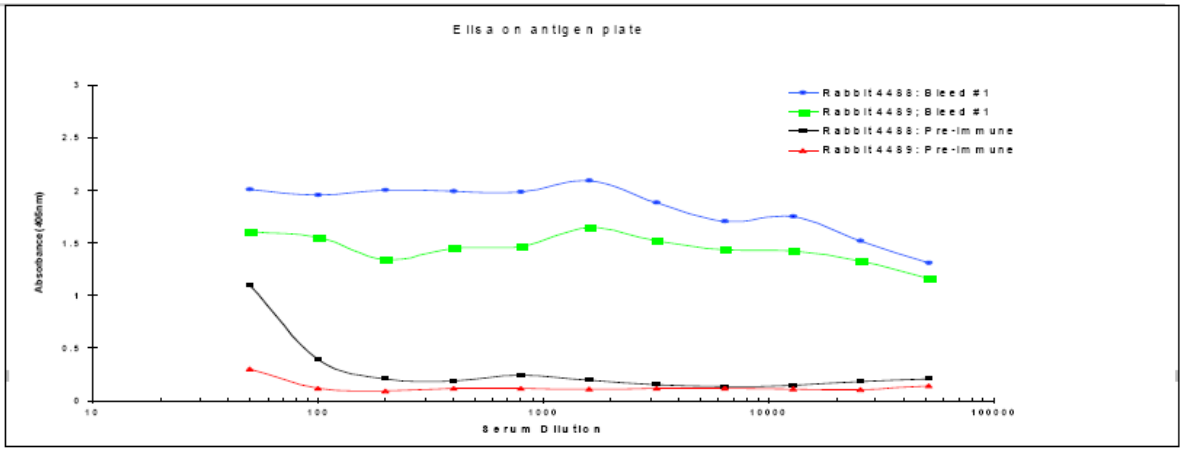
```
D238_1_F.       -ACCATCACCATCACTCCGCGGCTCTTGAAGTCCTCTTTCAGGGACCCGGGTACCTGAAA 59
PPAD.           ------------------------------------------------------AAA 3
                                                                        ***

D238_1_F.       AAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATC 119
PPAD.           AAGCTTTTACAGGCTAAAGCCTTGATTCTGGCATTGGGACTCTTCCAACTGCCCGCAATC 63
                ************************************************************

D238_1_F.       GCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAA 179
PPAD.           GCCCAAACGCAAATGCAAGCAGACCGAACAAACGGTCAATTTGCAACAGAAGAGATGCAA 123
                ************************************************************

D238_1_F.       CGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAA 239
PPAD.           CGAGCATTCCAGGAAACGAATCCCCCTGCAGGTCCTGTGCGTGCTATCGCTGAGTACGAA 183
                ************************************************************

D238_1_F.       CGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAG 299
PPAD.           CGCTCTGCAGCCGTTTTGGTACGCTACCCGTTCGGTATCCCGATGGAATTGATCAAAGAG 243
                ************************************************************

D238_1_F.       CTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTT 359
PPAD.           CTGGCCAAGAACGACAAGGTGATTACCATTGTGGCGAGTGAAAGCCAAAAAAACACCGTT 303
                ************************************************************

D238_1_F.       ATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAA 419
PPAD.           ATAACCCAGTACACCCAAAGCGGTGTGAATCTCTCTAATTGCGATTTCATCATTGCGAAA 363
                ************************************************************

D238_1_F.       ACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAA 479
PPAD.           ACTGACTCTTACTGGACACGCGACTATACCGGTTGGTTCGCAATGTACGATACGAACAAA 423
                ************************************************************

D238_1_F.       GTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAA 539
PPAD.           GTAGGTCTCGTGGACTTTATTTATAACCGCCCTCGTCCTAACGATGATGAATTCCCCAAA 483
                ************************************************************

D238_1_F.       TACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGC 599
PPAD.           TACGAAGCACAATATCTGGGCATCGAGATGTTCGGGATGAAGCTCAAGCAGACCGGTGGC 543
                ************************************************************

D238_1_F.       AACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAAC 659
PPAD.           AACTACATGACGGACGGATATGGATCCGCTGTGCAGTCACATATCGCATATACGGAGAAC 603
                ************************************************************

D238_1_F.       TCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACAT 719
PPAD.           TCCTCTCTGTCTCAAGCTCAAGTAAATCAAAAGATGAAAGACTATCTCGGCATCACACAT 663
                ************************************************************

D238_1_F.       CATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGCCTGTTGGGGCAAG 779
PPAD.           CATGATGTGGTACAAGATCCGAACGGCGAATATATCAACCATGTGGACTGTTGGGGCAAG 723
                ******************************************** **************

D238_1_F.       TATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAA 839
PPAD.           TATTTGGCACCGAACAAAATCCTCATCAGGAAAGTGCCTGACAATCACCCTCAGCACCAA 783
                ************************************************************

D238_1_F.       GCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTAC 899
PPAD.           GCCCTGGAAGATATGGCAGCCTACTTCGCAGCACAGACCTGCGCATGGGGAACGAAGTAC 843
                ************************************************************

D238_1_F.       GAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAAC 959
PPAD.           GAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTGAAC 903
                ************************************************************

D238_1_F.       AACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTG---------------------- 998
PPAD.           AACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAACGTC 963
                ************************************
```

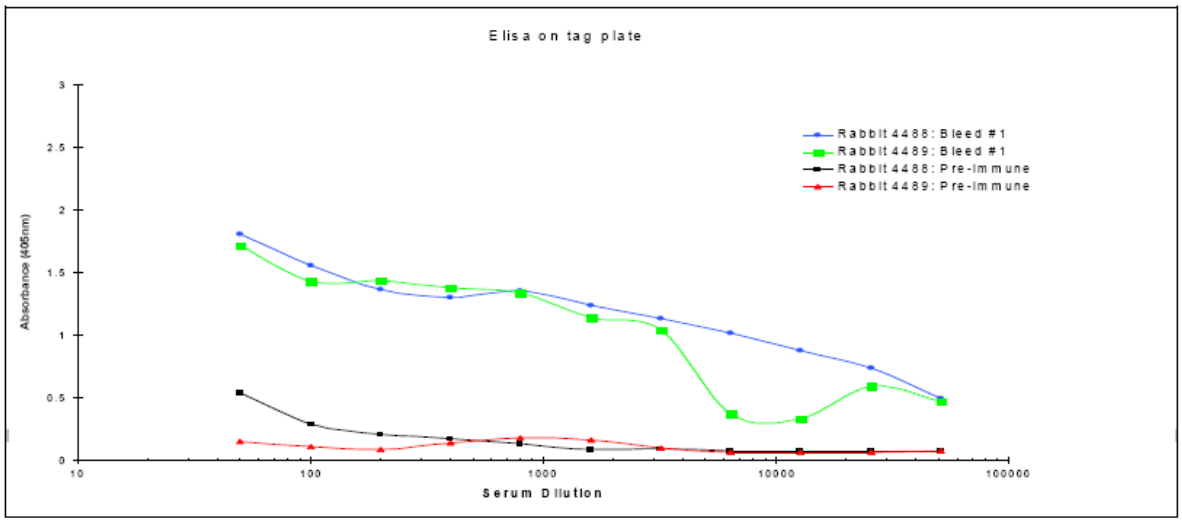
Reverse

```
D238_1_R.       TACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTG 219
PPAD.           TACGAGGTATATCGCGCTTTGGCCACCAATGAACAACCGTACACGAACTCTCTGATTCTG 900
                ************************************************************

D238_1_R.       AACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAAC 279
PPAD.           AACAACAGGGTATTTGTTCCTGTCAATGGCCCCGCCTCCGTGGACAACGATGCTCTGAAC 960
                ************************************************************

D238_1_R.       GTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACA 339
PPAD.           GTCTATAAGACGGCAATGCCCGGTTACGAAATTATAGGTGTCAAAGGGGCTTCAGGAACA 1020
                ************************************************************

D238_1_R.       CCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTAT 399
PPAD.           CCTTGGTTAGGAACAGATGCCCTGCATTGTCGTACTCACGAGGTAGCGGATAAGGGCTAT 1080
                ************************************************************

D238_1_R.       CTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAA 459
PPAD.           CTCTATATCAAGCACTACCCGATACTGGGCGAACAGGCAGGCCCTGATTATAAGATCGAA 1140
                ************************************************************

D238_1_R.       GCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATC 519
PPAD.           GCAGATGTCGTCTCATGCGCCAATGCTACTATCTCGCCGGTACAATGTTACTATCGTATC 1200
                ************************************************************

D238_1_R.       AATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATACT 579
PPAD.           AATGGTTCCGGTAGCTTTAAGGCTGCTGATATGACGATGGAATCAACAGGTCACTATACT 1260
                ************************************************************

D238_1_R.       TATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGAC 639
PPAD.           TATAGCTTTACAGGTCTTAACAAGAATGATAAGGTAGAATACTATATCTCTGCCGCTGAC 1320
                ************************************************************

D238_1_R.       AATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACG 699
PPAD.           AATAGTGGTCGCAAAGAGACTTATCCCTTTATCGGCGAACCTGATCCTTTCAAGTTTACG 1380
                ************************************************************

D238_1_R.       TGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATGG 759
PPAD.           TGTATGAACGAAACCAATACATGTACTGTGACCGGAGCTGCCAAAGCTCTTCGTGCATGG 1440
                ************************************************************

D238_1_R.       TTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATAT 819
PPAD.           TTCAACGCCGGTCGTTCAGAACTGGCTGTTTCGGTAAGTTTGAATATTGCCGGCACATAT 1500
                ************************************************************

D238_1_R.       CGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTA 879
PPAD.           CGGATAAAGCTTTATAACACCGCAGGAGAAGAAGTCGCTGCAATGACCAAGGAATTAGTA 1560
                ************************************************************

D238_1_R.       GCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTG 939
PPAD.           GCAGGGACGAGTGTCTTCAGTATGGATGTGTATTCTCAGGCTCCGGGCACATATGTTCTG 1620
                ************************************************************

D238_1_R.       GTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAAGAGCTCGCTCTG 999
PPAD.           GTTGTTGAAGGAAATGGAATCCGTGAGACAATGAAAATTCTCAAATAA------------ 1668
                ********************************************
```

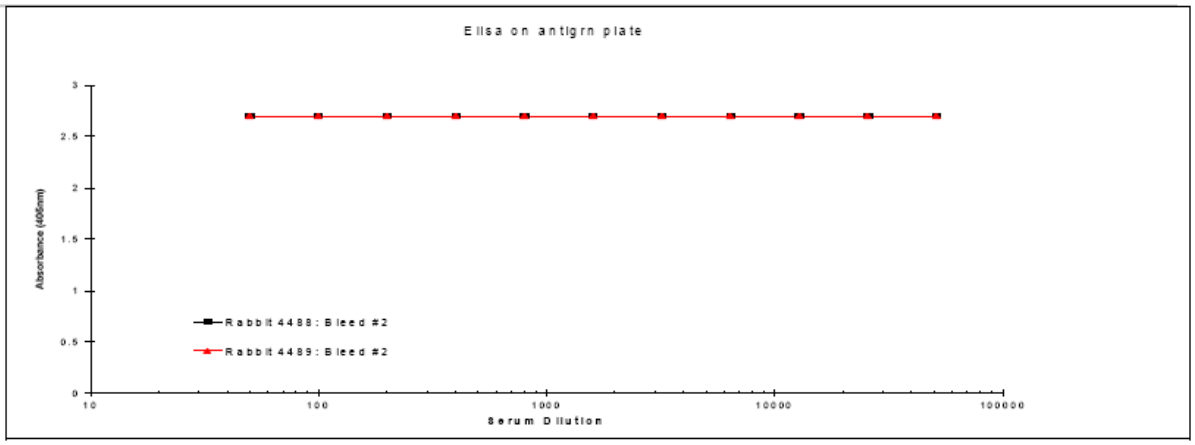
Anti-PPAD antibody production: ELISA analysis

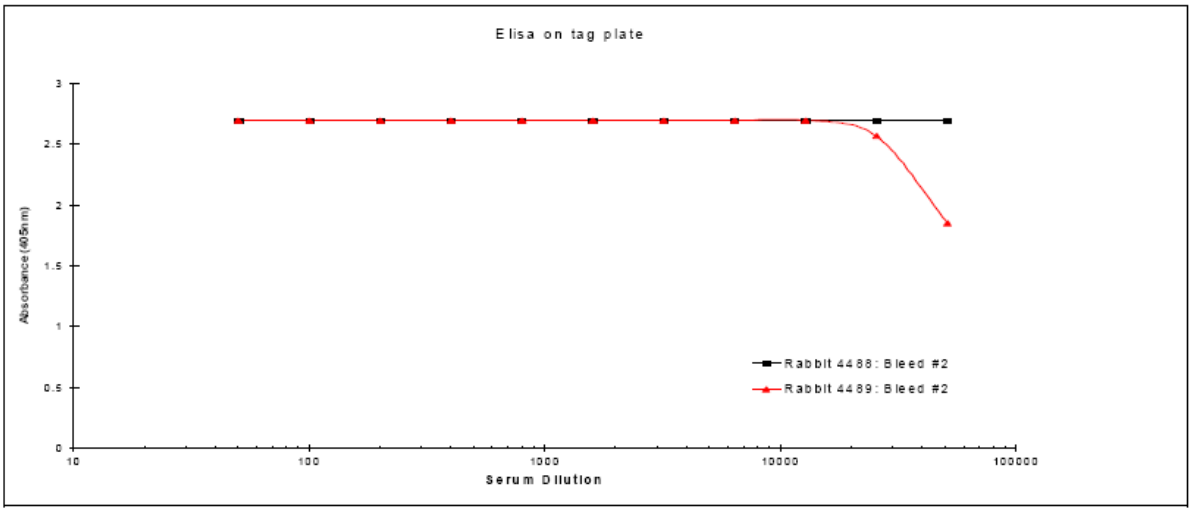
Pre-immune and bleed # 1



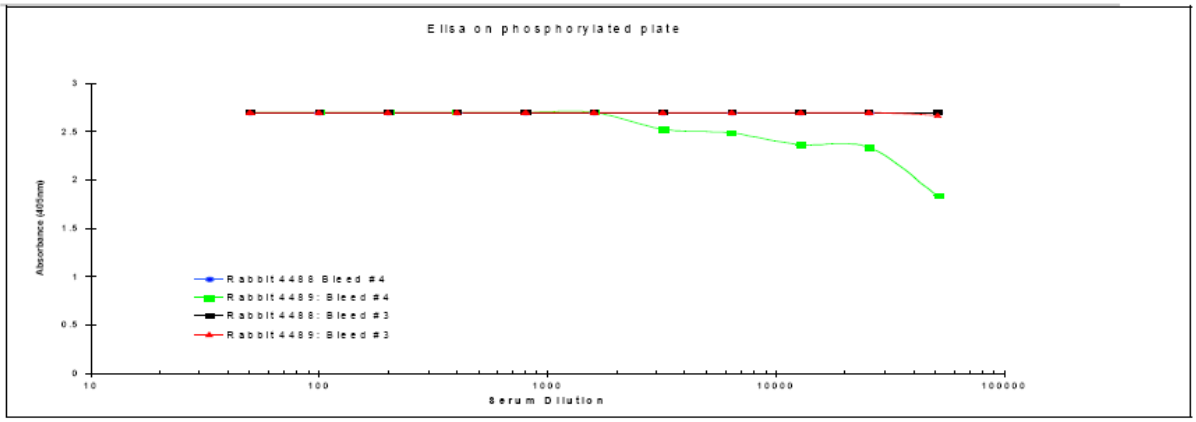


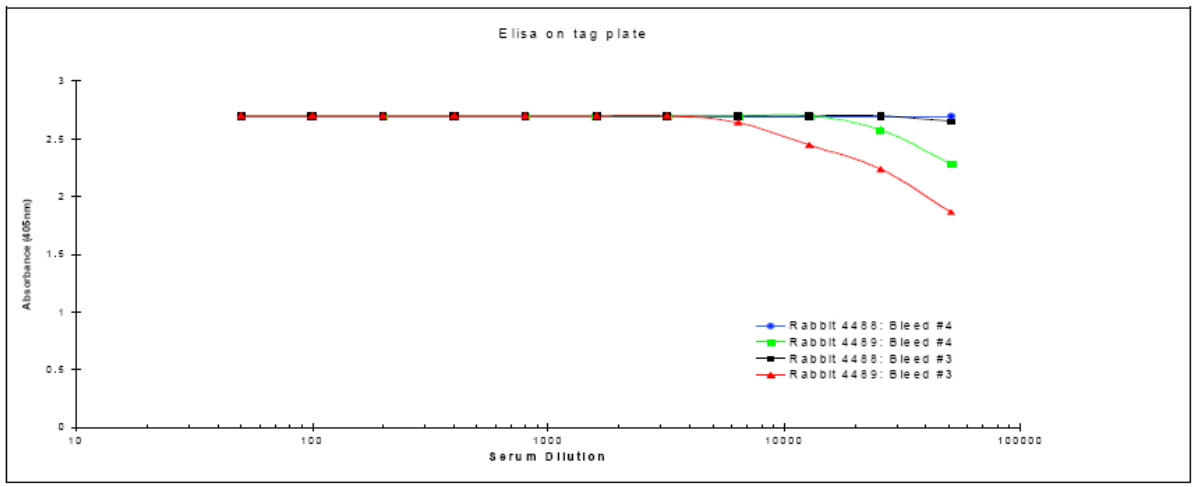
Bleed # 2




Elisa on tag plate
Absorbance (405nm)
Serum Dilution
Rabbit 4488: Bleed #2
Rabbit 4489: Bleed #2

Bleed # 3 and #4


Elisa on phosphorylated plate
Absorbance (405nm)
Serum Dilution
Rabbit 4488 Bleed #4
Rabbit 4489: Bleed #4
Rabbit 4488: Bleed #3
Rabbit 4489: Bleed #3


Elisa on tag plate
Absorbance (405nm)
Serum Dilution
Rabbit 4488: Bleed #4
Rabbit 4489: Bleed #4
Rabbit 4488: Bleed #3
Rabbit 4489: Bleed #3

Published papers

1. Kinloch, A., K. Lundberg, R. Wait, N. **Wegner, N.** H. Lim, A. J. Zendman, T. Saxne, V. Malmstrom and P. J. Venables. 2008. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. *Arthritis Rheum* 8:2287-2295.

Primary research article demonstrating extracellular protein citrullination by human PADs in RA synovial fluid. My contribution was the assistance in experimental work and writing of the article. The results are not presented in this thesis as it was not directly connected to my main project.

2. Lundberg, K., A. Kinloch, B. A. Fisher, **N. Wegner,** R. Wait, P. Charles, T. R. Mikuls and P. J. Venables. 2008. Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. *Arthritis Rheum* 10:3009-3019.

Primary research article describing the characterisation of the CEP-1 epitope and demonstrating cross-reactivity of anti-CEP-1 antibodies to full-length human α-enolase and P. gingivalis enolase. My contribution was the cloning and expression of P. gingivalis enolase, assistance in experimental work relating to the cross-reactivity studies and writing of the article. The results are presented in the Introduction of this thesis as they were the basis for my main project but not directly part of it and my intellectual contribution was minor.

3. **Wegner, N.**, R. Wait and P. J. Venables. 2009. Evolutionarily conserved antigens in autoimmune disease: Implications for an infective aetiology. *Int J Biochem Cell Biol* 2:390-397.

Review article focussing on the evolutionary conservation of a number of autoantigens in autoimmune disease and providing a critical evaluation of their potential involvement in an infective aetiology of the respective diseases. As a lead author I was responsible for the majority of the literature review and the writing. The article is feautured in the Introduction of this thesis.

4. **Wegner, N.**, K. Lundberg, A. Kinloch, B. Fisher, V. Malmstrom, M. Feldmann and P. J. Venables. 2010. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol Rev* 1:34-54.

Review article focussing on the aetiological components, in particular risk genes, smoking, and P. gingivalis, in the development of RA. As a lead author I was responsible for the majority of the literature review and the writing. The article and figures are feautured in the Introduction of this thesis.

5. **Wegner, N.**, R. Wait, A. Sroka, S. Eick, K. A. Nguyen, K. Lundberg, A. Kinloch, S. Culshaw, J. Potempa and P. J. Venables. 2010. Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and alpha-enolase: Implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum* 9:2662-2672.

Primary research article demonstrating endogenous citrullination in P. gingivalis, citrullination of human proteins bt P. gingialis PAD, and the requirement for PAD and arginine-gingipains in bacterial citrullination. It is entirely the result of my own work, unless where collaborations are indicated. I also wrote the article myself. The results are presented in Chapter 3 of this thesis.

6. Mangat, P., **N. Wegner**, P. J. Venables and J. Potempa. 2010. Bacterial and human peptidylarginine deiminases: Targets for inhibiting the autoimmune response in rheumatoid arthritis? *Arthritis Res Ther* 3:209.

Review article summarising and evaluating the current knowledge regarding protein citrullination by human and P. gingivalis PAD with a focus on possible therapeutical interventions. I contributed a substantial proportion to the text and figures (ca. 50%) and was highly involved in critical proofreading. The article and figures are feautured in the Introduction of this thesis.

7. Lundberg, K., **N. Wegner**, T. Yucel-Lindberg and P. J. Venables. 2010. Periodontitis in RA-the citrullinated enolase connection. *Nat Rev Rheumatol*.

Review article evaluating the experimental evidence for a specific link between periodontitis and P. gingivalis infection and anti-citrullinated enolase immunity in the development of rheumatoid arthritis. I contributed ca. 40% of text and figures and a substantial proportion of the overall intellectual input. The article and figures are feautured in the Introduction of this thesis.

ARTHRITIS & RHEUMATISM
Vol. 58, No. 8, August 2008, pp 2287–2295
DOI 10.1002/art.23618

Synovial Fluid Is a Site of Citrullination of Autoantigens in Inflammatory Arthritis

Andrew Kinloch,[1] Karin Lundberg,[1] Robin Wait,[1] Natalia Wegner,[1] Ngee Han Lim,[1] Albert J. W. Zendman,[2] Tore Saxne,[3] Vivianne Malmström,[4] and Patrick J. Venables[1]

Objective. **To examine synovial fluid as a site for generating citrullinated antigens, including the candidate autoantigen citrullinated α-enolase, in rheumatoid arthritis (RA).**

Methods. **Synovial fluid was obtained from 20 patients with RA, 20 patients with spondylarthritides (SpA), and 20 patients with osteoarthritis (OA). Samples were resolved using sodium dodecyl sulfate–polyacrylamide gel electrophoresis, followed by staining with Coomassie blue and immunoblotting for citrullinated proteins, α-enolase, and the deiminating enzymes peptidylarginine deiminase type 2 (PAD-2) and PAD-4. Proteins from an RA synovial fluid sample were separated by 2-dimensional electrophoresis, and each protein was identified by immunoblotting and mass spectrometry. Antibodies to citrullinated α-enolase peptide 1 (CEP-1) and cyclic citrullinated peptide 2 were measured by enzyme-linked immunosorbent assay.**

Results. **Citrullinated polypeptides were detected in the synovial fluid from patients with RA and patients with SpA, but not in OA samples. Alpha-enolase was detected in all of the samples, with mean levels of 6.4 ng/μl in RA samples, 4.3 ng/μl in SpA samples, and <0.9 ng/μl in OA samples. Two-dimensional electrophoresis provided evidence that the α-enolase was citrullinated in RA synovial fluid. The citrullinating enzyme PAD-4 was detected in samples from all 3 disease groups. PAD-2 was detected in 18 of the RA samples, in 16 of the SpA samples, and in none of the OA samples. Antibodies to CEP-1 were found in 12 of the RA samples (60%), in none of the SpA samples, and in 1 OA sample.**

Conclusion. **These results highlight the importance of synovial fluid for the expression of citrullinated autoantigens in inflammatory arthritis. Whereas the expression of citrullinated proteins is a product of inflammation, the antibody response remains specific for RA.**

Antibodies to citrullinated proteins (ACPAs) are highly specific for rheumatoid arthritis (RA) and are a powerful tool for its diagnosis and for prediction of disease severity. In addition, the study of ACPAs enables incorporation of risk factors, including smoking and presence of the shared epitope, into a common etiopathogenic model in which citrullination and the generation of antibodies play an intimate role in the pathogenesis of RA (for review, see ref. 1).

Citrullinated proteins are formed when arginine residues are deiminated by peptidylarginine deiminase (PAD). Five PADs have been identified in humans (2), of which PAD-2 and PAD-4 are found in rheumatoid synovial fluid cells (3) and in synovial membrane (4,5). PADs are calcium dependent (2), and are thus more likely to be active in the extracellular compartment, where calcium concentrations are higher. Cell death may result in increased PAD activity, because the loss of membrane integrity will both increase intracellular calcium concentration and enable extracellular leakage of PAD enzymes.

Citrullinated proteins have been detected in the synovial membrane of patients with various forms of arthritis (6) and in other inflamed tissues (7), suggesting

Supported by the Arthritis Research Campaign, UK and the European Union's Sixth Framework Programme (FP6) project AutoCure. Dr. Saxne's work was supported by the Swedish Medical Research Council.

[1]Andrew Kinloch, BSc (Hons), Karin Lundberg, MSc, PhD, Robin Wait, PhD, Natalia Wegner, MRES, Ngee Han Lim, BSc, Patrick J. Venables, MD, FRCP: Imperial College London, London, UK; [2]Albert J. W. Zendman, PhD: Radboud University Nijmegen, Nijmegen, The Netherlands; [3]Tore Saxne, MD, PhD: Lund University Hospital, Lund, Sweden; [4]Vivianne Malmström, PhD: Karolinska Institutet, Stockholm, Sweden.

Address correspondence and reprint requests to Patrick J. Venables, MD, FRCP, Professor of Rheumatology, Kennedy Institute of Rheumatology, 65 Aspenlea Road, Imperial College London, London W6 8LH, UK. E-mail: p.venables@imperial.ac.uk.

Submitted for publication March 28, 2007; accepted in revised form April 4, 2008.

that, whereas citrullination is associated with inflammation in general, the development of ACPAs is specific to RA. In patients with RA, ACPA-producing plasma cells have been detected in the synovial membrane (8), and higher concentrations of ACPAs in the rheumatoid joint compared with the serum have also been reported (8,9). This suggests that synovial citrullinated proteins are driving a local production of antibodies, and that the resulting immune complexes contribute to the chronic inflammation in the rheumatoid joint.

Previous studies of citrullinated antigens have tended to focus on synovial membrane (10–12). In the present study we investigated synovial fluid as a site of extracellular deimination, focusing on the candidate autoantigen citrullinated α-enolase (13). We recently mapped the autoantibody response to an immunodominant peptide, which we have termed citrullinated α-enolase peptide 1 (CEP-1). Antibodies to CEP-1 are closely correlated with antibodies to cyclic citrullinated peptide 2 (CCP-2), are highly specific for RA, and have a diagnostic sensitivity of ~50% depending on the cohort of patients studied (14). Having identified a major epitope, we now have reagents to study the distribution of native and citrullinated α-enolase and its antibodies in clinical samples. In addition, CEP-1 is derived from a real protein, as opposed to the purely artificial peptides that comprise the CCP-2 assay, and, as such, may reflect the pathologic processes occurring in the joints of patients with RA.

PATIENTS AND METHODS

Patient samples. Synovial fluid was removed from the knees of patients at the time of therapeutic arthrocentesis. Samples were obtained, following the patients' provision of informed consent and approval from the local ethics committee, from 20 patients with RA and 20 patients with spondylarthritides (SpA). These patients were attending the Rheumatology Clinic at Karolinska University Hospital (Stockholm, Sweden). Synovial fluid samples from 20 patients with osteoarthritis (OA), selected for the absence of cellularity and thus serving as noninflammation controls, were obtained from the knees of patients attending the Department of Rheumatology at Lund University Hospital (Lund, Sweden), following the provision of informed consent and approval from the local ethics committee. For 2-dimensional electrophoresis, an additional rheumatoid synovial fluid sample was obtained from the knee of a patient attending Charing Cross Hospital in London, UK.

After centrifugation to remove cells, the synovial fluid samples were stored at −70°C until used. After thawing, the samples were digested with hyaluronidase type IV-S (50 μg/ml; Sigma, St. Louis, MO), vortexed, and passed through a 0.2-μm Ministart filter unit (Sartorius, Hannover, Germany). Protease inhibitor cocktail (Sigma) and EDTA were added to final concentrations of 10 μl/ml and 50 m*M*, respectively, and the samples were stored at −20°C.

Generation of rabbit anti–CEP-1 antibodies. A rabbit polyclonal anti–CEP-1 antibody was generated at Cambridge Research Biochemicals (Ely, UK). Briefly, 2 rabbits were immunized subcutaneously every 2 weeks, for 10 weeks, with 200 μg keyhole limpet hemocyanin (KLH)–conjugated CEP-1 (peptide:KLH ratio 1:1) in Freund's incomplete adjuvant per booster. Blood was collected 7 days after each injection, and sera were analyzed for the presence of anti–CEP-1 antibodies by enzyme-linked immunosorbent assay (ELISA) (14).

When antibody titers had reached significant levels, animals were killed and blood samples were harvested. The crude antisera were depleted of cross-reactive antibodies by chromatography on a thiopropyl-Sepharose column conjugated to a control peptide in which the citrulline residues had been replaced with arginine residues. The unbound fraction was affinity purified on a second thiopropyl-Sepharose column, conjugated to CEP-1. Anti–CEP-1–specific antibodies were eluted and further depleted of nonspecific antibodies in 3 subsequent passages through the depleting column. Bound antibodies were eluted from the column with 27 ml of 0.1*M* glycine/HCl, pH 2.5, and the flow-through was collected as the unbound fraction.

Both antibody fractions demonstrated similar preferential reactivity for citrullinated α-enolase, with the glycine eluate having a greater sensitivity for citrullinated α-enolase (as demonstrated by immunoblotting of uncitrullinated and in vitro citrullinated α-enolase at the same dilution, 1:400 [data not shown]). Both fractions were stored at −20°C until further used.

Cloning and expression of recombinant enolase. The full-length human α-enolase coding sequence was amplified, by polymerase chain reaction (PCR), from the complementary DNA of HL60 cells differentiated with vitamin D_3 (13). The PCR forward and reverse primers contained the *Bam* HI and *Xho* I restriction sites, respectively (AGTTGGATCCTCTATTCTCAAGATCCATGCCA and ATCCTCGAGTTACTTGGCCAAGGGRTTTCTGAAGTTCCTG, respectively). The PCR product was ligated in-frame into the multiple cloning site of the plasmid expression vector pGEX 6P3 (GE Healthcare, Bucks, UK), 3′ to the glutathione S-transferase (GST) coding site.

Expression of GST α-enolase was induced in the protease-deficient BL21 strain of *Escherichia coli*. Briefly, the bacteria were grown at 28°C. Once an optical density at 600 nm ($OD_{600\ nm}$) of 0.6–0.8 had been reached, IPTG was added to the culture to a final concentration of 0.1 μ*M* to induce GST α-enolase expression. Cultures were incubated for a further 18 hours, and fusion protein was purified from bacterial lysates according to the plasmid manufacturer's instructions. PreScission Protease (GE Healthcare), used to cleave the GST from the α-enolase, was pelleted with the excised GST, using glutathione–Sepharose 4B (GE Healthcare). Purified α-enolase, as determined by Coomassie blue staining and tandem electrospray mass spectrometry (MS), was dialyzed against phosphate buffered saline (PBS), and the protein concentration was measured by bicinchoninic acid assay (Pierce, Rockford, IL).

Immunoblotting. Citrullinated proteins were identified using an Anti–Modified Citrulline (AMC) detection kit (Up-

state Biotechnology, Lake Placid, NY) according to the manufacturer's instructions. Rabbit anti–α-enolase (H300; Santa Cruz Biotechnology, Santa Cruz, CA), rabbit anti–PAD-2 (Abcam, Cambs, UK), and goat anti–PAD-4 (Abcam) were used at 1:200 in incubations overnight at 4°C.

Synovial fluid proteins were resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) using 10-well 4–12% Bis-Tris precast gels (Invitrogen, Paisley, UK). Gels were either stained with Coomassie blue or transferred to nitrocellulose membranes for immunoblotting. Following incubations with primary and secondary antibodies, the membranes were blocked in 5% milk in PBS/0.1% Tween, blotted with antibodies diluted in blocking buffer, and washed in PBS/0.1% Tween. The preparations were then developed using enhanced chemiluminescence. Anti-goat and anti-rabbit secondary antibodies conjugated to horseradish peroxidase (Dako, Glostrup, Denmark) were diluted to 1:5,000 and 1:3,000, respectively, and reacted with membranes for 1 hour at room temperature.

To prevent masking by abundant comigrating proteins, synovial fluid samples were depleted of albumin and IgG, using the Proteoextract Albumin/IgG Removal Kit (Calbiochem, Notts, UK), prior to immunoblotting for the quantification of α-enolase and the detection of PAD-2 and PAD-4. Briefly, synovial fluid, diluted 1:9 in the supplied binding buffer, was added at 560 μl per column. Proteins in the flow-through were immunoblotted for α-enolase without concentration or were immunoblotted at a 10-fold concentration, attained using a 0.5-ml concentrator with a molecular weight cutoff of 10,000 kd (Vivascience, Stonehouse, UK). Proteins in the flow-through were also immunoblotted for PAD-2 or PAD-4. Poly-His–tagged recombinant human PAD-2 and PAD-4 (each 50 ng/well) (3) were used as positive controls. In addition, nondepleted synovial fluid samples (diluted 1:30 in PBS) from 20 patients with RA, 20 patients with SpA, and 20 patients with OA were applied directly (2 μl/dot) onto a single nitrocellulose membrane for detection of citrullinated proteins. Immunoblotting was performed using the AMC detection kit in accordance with the manufacturer's instructions.

Measurement of synovial fluid soluble α-enolase concentration. Synovial fluid samples depleted of IgG and albumin and boiled in 2× Laemmli buffer were loaded in duplicate onto 15-well 4–12% Bis-Tris precast gels (Invitrogen). On each gel, a range of purified α-enolase standards was also loaded. Following SDS-PAGE and electrotransferesis, nitrocellulose membranes were blocked in 5% milk/PBS/0.1% Tween overnight, incubated in rabbit anti–α-enolase, diluted 1:200 in blocking buffer, washed in PBS/0.1% Tween, incubated with anti-rabbit secondary antibodies (Dako), washed in PBS/0.1% Tween, and developed using enhanced chemiluminescence. Concentrations of α-enolase were calculated by densitometry.

Films (Amersham Hyperfilm; GE Healthcare) were scanned using a scanning densitometer (GS-710 Calibrated Imaging Densitometer; Bio-Rad, Hercules, CA). Fifty-kilodalton bands were quantified for total pixel volumes (corresponding to total amounts of α-enolase) using Phoretix software (Nonlinear Dynamics, Ltd., Newcastle, UK). Pixel volumes of bands at 50 kd in lanes containing recombinant α-enolase were used to create a standard curve of total α-enolase levels relative to pixel volume, achieved using Prism software (GraphPad Software, San Diego, CA). From this curve, total α-enolase levels in each of the wells loaded with synovial fluid samples were deduced, and a mean α-enolase level for each patient sample was calculated from duplicate lanes. This assay was performed for each individual 15-well gel, to avoid generation of artefacts resulting from intergel and interfilm variation. Each gel was loaded with 4 samples, added in duplicate. Because of the dilutions and volumes of the synovial fluid samples, the concentration of α-enolase in the fluid was multiplied by 1.5. Thus, the densitometric value representing the lowest value on the standard curve, 0.6 ng/μl, equated to 0.9 ng/μl of α-enolase in the synovial fluid.

Liquid-phase isoelectric focusing (IEF). To characterize citrullinated proteins by tandem MS, an RA synovial fluid sample with an abundance of citrullinated proteins was fractionated by liquid-phase IEF (Zoom IEF; Invitrogen). Briefly, synovial fluid was mixed with 40 μl carrier ampholytes (Invitrogen) and 80 μl of 1*M* dithiothreitol, and made up to a volume of 4 ml in IEF buffer, giving a final concentration of 0.5 mg/ml protein. Polyacrylamide Zoom disks (Invitrogen), each with the appropriate pH, were added to the assembly, to yield 5 liquid-phase protein fractions (pH 3.0–4.5, 4.6–5.3, 5.4–6.1, 6.2–6.9, and 7.0–10.0). Anode (pH 2.5–2.9) and cathode (pH 3.0–10.0) buffers were added to their respective electrodes at a volume of 17.5 ml. Samples were added to each of the chambers (650 μl/chamber), and IEF was performed according to the manufacturer's guidelines, consisting of 100V for 20 minutes, 200V for 80 minutes, and 800V for 80 minutes. Each fraction was further separated by SDS-PAGE, followed by either staining with Coomassie blue or immunoblotting.

Coomassie blue–stained proteins were excised, digested with trypsin, and characterized by tandem MS as previously described (13). For these experiments, semipurified α-enolase (Hytest, Turku, Finland), with or without prior digestion with PAD, was used as the positive control.

ELISA. Ninety-six–well plates (Maxisorp; Nunc, Roskilde, Denmark) were coated with CEP-1 (CKIHA-X-EIFDS-X-GNPTVEC, where X represents citrulline) or the arginine-containing control peptide (CKIHAREIFDSRGNPTVEC) at 10 μg/ml (diluted in a 50-m*M* carbonate buffer, pH 9.6), or with 2% bovine serum albumin (BSA) in carbonate buffer, and incubated overnight at 4°C. Wells were washed with PBS/0.05% Tween and blocked with 2% BSA (diluted in PBS) for 3 hours at room temperature. Sera were diluted 1:100 in radioimmunoassay (RIA) buffer (10 m*M* Tris, 1% BSA, 350 m*M* NaCl, 1% Triton-X, 0.5% sodium deoxycholate, 0.1% SDS), added in duplicate, and incubated for 1.5 hours at room temperature. Plates were washed as described above and incubated for 1 hour at room temperature with peroxidase-conjugated mouse anti-human IgG (Hybridoma Reagent Laboratory, Baltimore, MD), diluted 1:1,000 in RIA buffer.

After a final wash (PBS/0.05% Tween), bound antibodies were detected with tetramethylbenzidine substrate (KPL, Gaithersburg, MD). The reaction was stopped by the addition of 1*M* H_2SO_4, and absorbance was measured at 450 nm in a Multiscan Ascent microplate reader (ThermoLabsystems, Helsinki, Finland). A control serum was included on all plates to correct for plate-to-plate variation.

For each serum tested, background $OD_{450\ nm}$ values (i.e., in wells coated with 2% BSA in carbonate buffer) were subtracted from peptide $OD_{450\ nm}$ values. OD values higher than 0.2 were considered to be peptide positive. Anti-CCP antibody status was

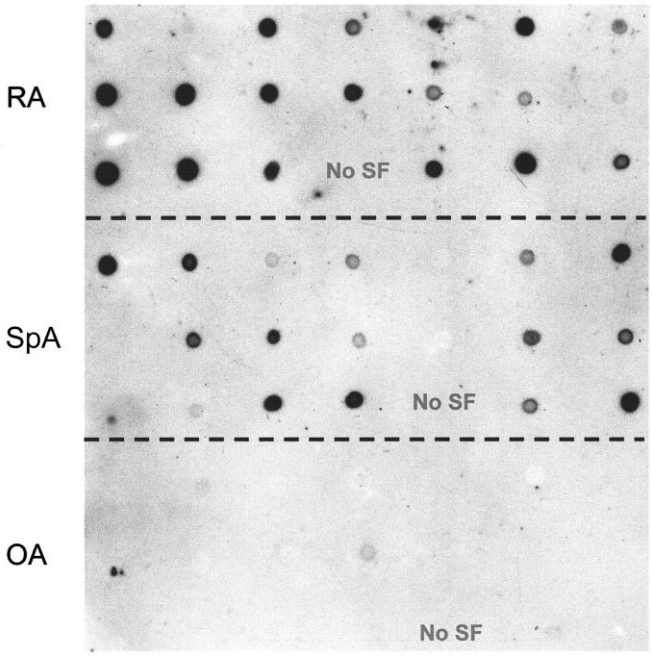


Figure 1. Dot-blots of synovial fluid samples from the 3 different disease groups (n = 20 per group), showing citrullinated proteins on a single membrane. "No SF" refers to a spot to which no synovial fluid was added. RA = rheumatoid arthritis; SpA = spondylarthritides; OA = osteoarthritis.

analyzed using the CCP-2 kit (Eurodiagnostica, Malmo, Sweden) according to the manufacturer's instructions.

RESULTS

Increased levels of citrullinated proteins and α-enolase in synovial fluid from patients with inflammatory arthritis. Immunoblotting with the AMC detection kit demonstrated that both the RA samples and the SpA samples contained citrullinated proteins across the entire range of molecular mass, although none were detected in the OA samples (individual immunoblotting results are available upon request from the corresponding author). These findings were also confirmed by dot-blot analysis of synovial fluid samples from all patients (Figure 1), assessed on a single membrane to eliminate possible variability between separate membranes. Citrullinated proteins were detected by dot-blot in 19 of 20 RA samples, in 17 of 20 SpA samples, and in only 2 OA specimens.

Immunoblotting demonstrated the presence of α-enolase in both the RA samples and the SpA samples;

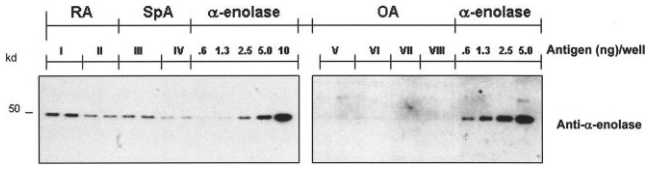


Figure 2. Immunoblots of synovial fluid samples from the 3 different disease groups (n = 20 per group), showing higher levels of α-enolase in synovial fluid from patients with RA (lanes I and II) and patients with SpA (lanes III and IV) compared with patients with OA (lanes V–VIII); standard curves of recombinant α-enolase are also shown. The OA synovial fluid required longer exposures to demonstrate the much lower levels of the reactive polypeptide. Results from representative duplicate samples are shown. See Figure 1 for definitions.

however, α-enolase was barely detectable in the OA samples (Figure 2). We used densitometry and a standard curve of recombinant protein to quantify the α-enolase in all 60 samples. As shown in Figures 2 and 3, higher levels of α-enolase were detected in samples from both the patients with RA (mean 6.4 ng/μl, range 1.2–15.4) and the patients with SpA (mean 4.3 ng/μl, range 1.1–9.2) as compared with the OA controls (<0.9 ng/μl in all samples).

RA specificity of increased levels of antibodies to CEP-1 in synovial fluid. Antibodies to CEP-1 were detected in 12 of the 20 RA synovial fluid samples (60%) but in only 1 OA sample (5%) and in none of the SpA samples (Figure 4). None of the RA patients had antibodies to the arginine-containing control peptide, while 1 OA synovial fluid sample and 1 from a patient

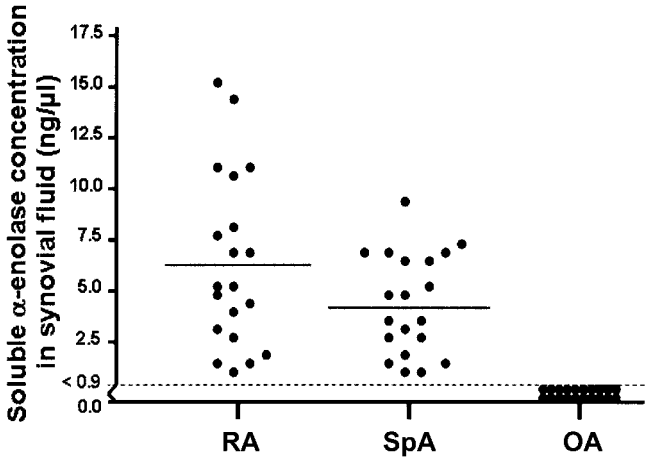


Figure 3. Levels of α-enolase in synovial fluid samples from the 3 different disease groups (n = 20 per group), as measured using immunoblotting and densitometry analyses. Broken line indicates the cutoff for positivity; solid lines indicate the mean. See Figure 1 for definitions.

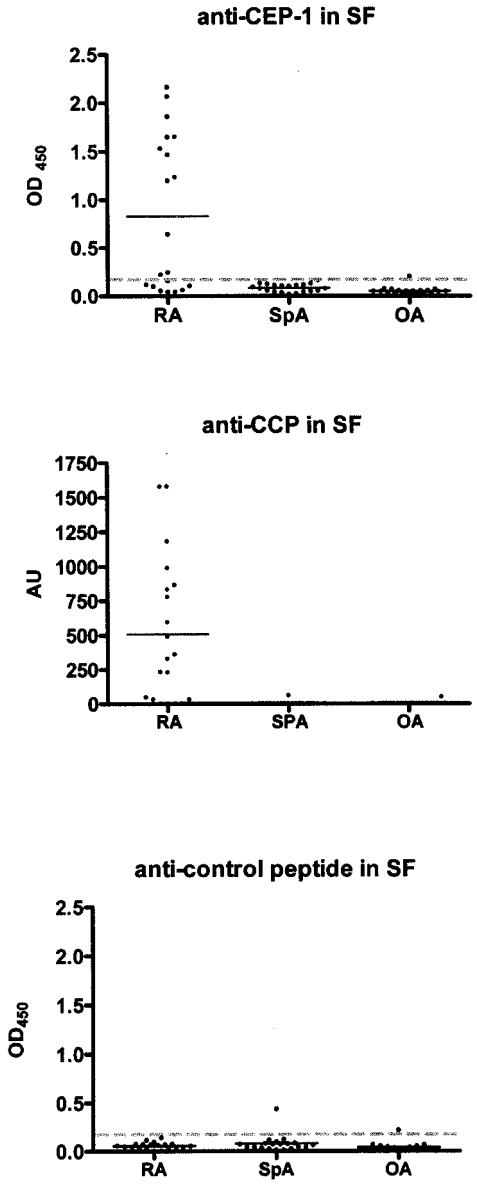


Figure 4. Antibodies to citrullinated α-enolase peptide 1 (anti–CEP-1) and cyclic citrullinated peptide 2 (anti–CCP-2) in synovial fluid (SF) samples from the 3 different disease groups (n = 20 per group), with elevated levels specifically observed in the synovial fluid from patients with RA compared with the patients with SpA and patients with OA. There was no reaction with the arginine-containing control peptide in the RA samples, confirming that the anti–CEP-1 reaction is citrulline-specific. Broken line indicates the cutoff for positivity; solid lines indicate the mean. OD_{450} = optical density at 450 nm; AU = arbitrary units (see Figure 1 for other definitions).

with SpA were positive for the control peptide. All 12 RA patients who were anti–CEP-1 positive were also anti–CCP-2 positive.

Characterization of citrullinated proteins in synovial fluid from a patient with RA. To characterize the citrullinated proteins present in rheumatoid synovial fluid, we selected an RA synovial fluid sample containing heavily citrullinated proteins. This sample was resolved by liquid-phase IEF, followed by SDS-PAGE. Bands from Coomassie blue–stained gels were identified by tandem MS (Figure 5a). Immunoblotting with the AMC antibody demonstrated the presence of citrullinated proteins with molecular masses between 35 kd and 250 kd (Figure 5b). The concentration of these proteins was greatest in the more acidic fractions, particularly between pH 3.0 and pH 4.6, whereas Coomassie blue–stained material was most abundant in fractions with a pH higher than 5.5. This is consistent with the acidic shift associated with deimination of arginine.

We identified α-enolase with 2 different antibodies: the rabbit anti–α-enolase H300, which reacted with purified α-enolase equally in its uncitrullinated and citrullinated forms, and the anti–CEP-1 antibody, which preferentially reacted with the in vitro citrullinated form. The H300 antibody reacted with a 50-kd polypeptide, which was most abundant in the IEF fractions with pH of 6.2–6.9 and 7.0–10.0 (Figure 5c). This is consistent with the characteristics of native α-enolase, which has a calculated pH of 6.9. The reactivity of the anti–CEP-1 antibody showed a shift toward more acidic forms of the molecule, including a stronger reaction with the 5.4–6.1 fraction and a weaker reaction with the 7.0–10.0 fraction (Figure 5d), compared with the reactions with the H300 antibody. This is consistent with the presence of some partially citrullinated α-enolase.

All bands visible by Coomassie blue staining were excised and digested in gel with trypsin, and the resulting peptides were sequenced by tandem MS; the proteins identified are shown in Figure 5a. No citrulline residues were observed with the use of tandem MS in peptides derived from any of the synovial fluid proteins, whereas peptides containing deiminated arginine residues were readily detected in a PAD-treated sample of α-enolase analyzed in parallel, suggesting that the stoichiometry of in vivo citrullination is relatively low.

Presence of PAD-2 and PAD-4 in synovial fluid from RA, SpA, and OA patients. We investigated the presence of enzymes capable of deimination of arginine, using antibodies against PAD-2 and PAD-4 in immunoblots of synovial fluid cultures that had been depleted of IgG and albumin and subsequently concentrated. Prominent bands were seen when blotting for PAD-4 in all of the synovial fluid samples. However, reactivity was gen-

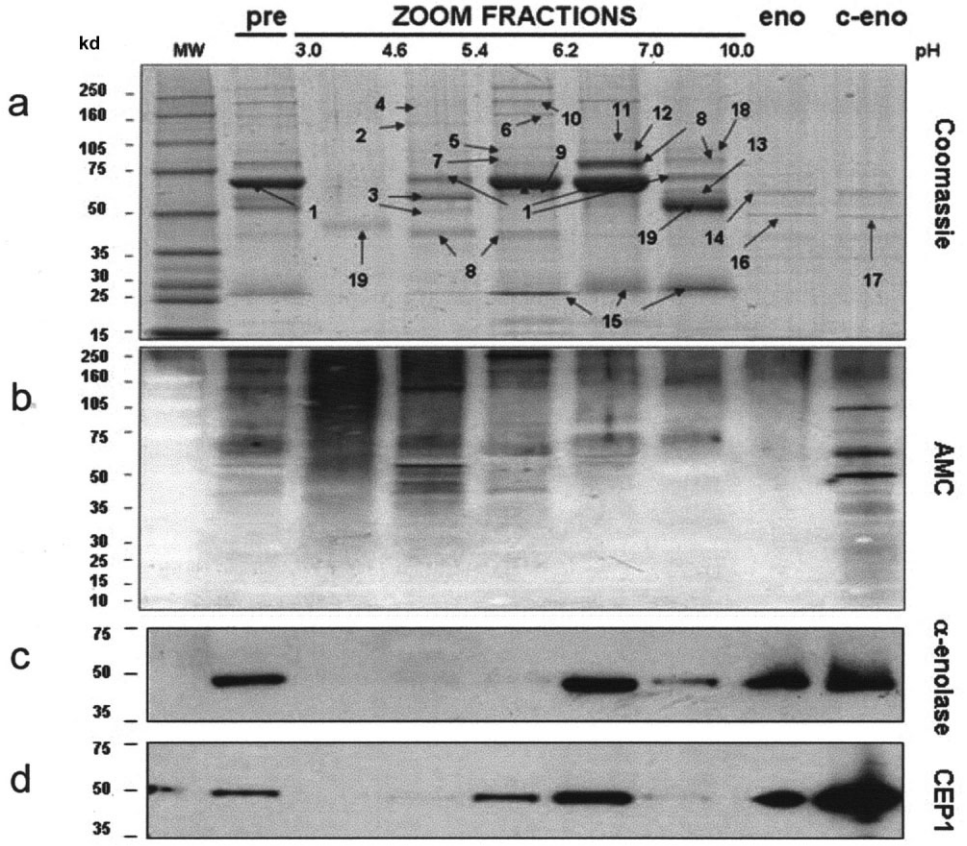


Figure 5. Resolution and identification of synovial fluid proteins by 2-dimensional electrophoresis. A synovial fluid sample from a patient with rheumatoid arthritis was analyzed by liquid-phase isoelectric focusing (without Zoom disks [pre] or with Zoom disks at 5 liquid-phase protein fractions of pH 3.0–4.5, 4.6–5.3, 5.4–6.1, 6.2–6.9, and 7.0–10) followed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. **a,** Proteins were identified by tandem electrospray mass spectrometry (MS) on Coomassie blue–stained gels: 1 = albumin, 2 = inter–α-trypsin inhibitor H-chain, 3 = α_1-antitrypsin, 4 = fibronectin, 5 = fibrinogen γ-chain, 6 = ceruplasmin, 7 = IgM H-chain, 8 = haptoglobin, 9 = IgMα_1 H-chain, 10 = α_2-macroglobulin, 11 = major histocompatibility complex factor B, 12 = gelsolin, 13 = IgMγ_3 H-chain, 14 = M1/M2 pyruvate kinase, 15 = apolipoprotein, 16 = native α-enolase (eno), 17 = citrullinated α-enolase (c-eno), 18 = IgMγ_3, and 19 = α_1-acid glycoprotein. Only in vitro citrullinated α-enolase could be demonstrated to contain citrulline residues by MS. **b,** Immunoblotting of Zoom fractions with the anti–modified citrulline antibody (AMC) demonstrated that citrullinated proteins were present in all Zoom fractions, but were more abundant in the most acidic ones. **c,** Immunoblotting for α-enolase with the H300 anti–α-enolase antibody, which reacts equally with native α-enolase and citrullinated α-enolase, demonstrated that α-enolase was most abundant in the 2 most basic fractions. **d,** Immunoblotting with the citrullinated α-enolase peptide 1 antibody (anti–CEP-1), which reacts preferentially with citrullinated α-enolase, showed reactivity with a 50-kd polypeptide that was relatively abundant in more acidic fractions, indicating that a proportion of the α-enolase is citrullinated in rheumatoid synovial fluid. MW indicates the molecular weight standard.

erally lower in the OA samples. In comparison, reactivity with the anti–PAD-2 antibody was detectable in 18 of 20 RA samples, in 16 of 20 SpA samples, and in none of the OA samples. Both PAD-2 and PAD-4 migrated at a molecular weight similar to the 75-kd molecular weight standard, just below the recombinant PAD-2– and PAD-4–positive controls, the difference in mass being accounted for by the poly-His tag (Figure 6).

Although the antibodies showed good specificity for their respective controls, there were widespread cross-reactions with other proteins in the synovial fluid. Some of the samples (see OA sample 3 in Figure 6) showed a reaction of the PAD-2 antibody with a polypeptide that migrated just above the 75-kd marker and the positive control. This was interpreted as not being PAD-2, because of the increased molecular mass. There was no reaction with the conjugate controls (results not shown). These findings are consistent with the notion that the extracellular compartment of synovial fluid in patients with inflammatory arthritis is a site for generation of citrullinated proteins by either or both of these enzymes.

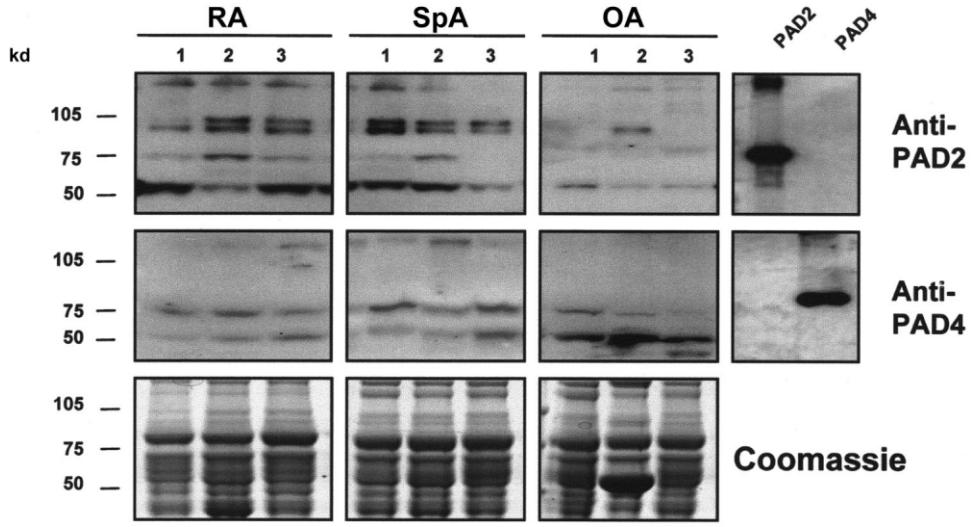


Figure 6. Representative immunoblots of peptidylarginine deiminase type 2 (PAD-2) and PAD-4 in synovial fluid of patients with rheumatoid arthritis (RA), patients with spondylarthritides (SpA), and patients with osteoarthritis (OA). PAD-2 was detected in all 3 RA samples and in 2 of the SpA samples (SpA1 and SpA2), but in none of the OA samples. PAD-4 was found in all of the synovial fluid samples. His-tagged PAD-2 and PAD-4 were used to confirm specific reactivity of the anti-PAD antibodies. Note that the recombinant His–tagged PAD-2 and PAD-4, used as positive controls, migrated at a slightly higher apparent mass because of the 6 histidine residues in the tag. Both antibodies reacted with higher and lower molecular weight polypeptides, indicating cross-reactivity with other unidentified proteins. Immunoblotting with anti-goat and anti-rabbit antibodies, without primary antibodies, confirmed that the synovial fluid proteins were not reacting directly with the secondary antibodies (results not shown).

DISCUSSION

In this study we have demonstrated that synovial fluid from patients with RA and from patients with SpA is characterized by an abundance of citrullinated proteins. The presence of abundant citrullinated proteins in synovial fluid from patients with SpA and patients with RA, compared with a lack of citrullinated proteins in OA synovial fluid, suggests that, similar to that in the synovial membrane, the presence of elevated levels of citrullinated proteins may be characteristic of inflammation, and not restricted to RA.

Our identification of α-enolase as a candidate citrullinated antigen in synovial fluid suggests that this previously neglected compartment of the rheumatoid joint is a site of expression of autoantigens. This observation is consistent with the findings from previous studies, which have demonstrated the presence of citrullinated fibrin in synovial fluid from patients with RA, but not in OA synovial fluid (15,16), and is consistent with a recent description of mutated and citrullinated vimentin in RA synovial fluid (17), although in that report, the material included lysed cells. We also found that the mean α-enolase levels in rheumatoid synovial fluid were increased at least 6-fold in the samples from RA patients, and increased at least 4-fold in the SpA synovial fluid. Because the levels in the OA samples fell below the lowest concentration of recombinant α-enolase on the standard curve, it may be that the ratio of concentrations in the 2 types of inflammatory arthritis tested herein may be even higher in comparison with OA.

The higher levels of anti–CEP-1 antibodies found in RA patients (60% of the RA synovial fluid samples) compared with the controls supports the concept that expression of citrullinated proteins is a product of inflammation, whereas the antibody response remains specific to RA. In this study we did not examine the relative concentration of anti–CEP-1 antibodies in synovial fluid compared with that in the serum. A highly significant enrichment has already been demonstrated for anti–CEP-1 and other ACPAs in more than 300 paired serum and synovial fluid samples in a separate study (Snir O, et al: unpublished observations).

We demonstrated reactivity with the anti–CEP-1 antibody in acidic fractions containing proteins with pH below the calculated pH of unmodified α-enolase. Reactivity with the AMC detection antibody was also evident in the more acidic fractions. This is consistent with some of the α-enolase being citrullinated in vivo, and therefore gaining acidity. However, we were unable to demonstrate specific citrullinated residues in synovial fluid α-enolase by tandem MS.

Similar findings were obtained for other abun-

dant proteins, mainly well-documented serum proteins, in which citrullination was demonstrated in comigrating polypeptides, as revealed by staining with the AMC antibody. The only sample in which deiminated arginines were demonstrable by MS was the in vitro citrullinated α-enolase, in which every arginine was deiminated, as observed both in this study and in our previous report (13). The failure to detect citrulline residues in the deiminated synovial fluid proteins is probably due to partial deimination, to the low stoichiometry of in vivo citrullination, and to the distribution of the modification over different arginines in different molecules. This is entirely consistent with the findings from another recent study, in which deimination was detected in synovial membrane proteins by both electrophoretic shift assay and staining with the AMC antibody, but not by MS, even in relatively abundant proteins such as citrullinated fibrinogen (12).

This is the first report demonstrating the presence of extracellular PAD-2 and PAD-4 in the synovial fluid, although it required removal of the abundant serum proteins, and subsequent concentration, to demonstrate the enzymes clearly. The anti–PAD-4 antibody reacted unequivocally at the appropriate molecular weight with all of the synovial fluid samples, with fainter bands visible in the OA samples, which again supports the hypothesis that synovial fluid is a site of active deimination in the presence of inflammation. In comparison, reactivity with the anti–PAD-2 antibody was only detectable in the RA and SpA samples. Therefore, it cannot be discounted that PAD-2 is also present in the OA samples, but was below the levels of detection used in this study. If both PAD-2 and citrullinated proteins are indeed absent in OA samples, the hypothesis can justly be made that it is the PAD-2 that is responsible for the deimination of the extracellular citrullination of synovial fluid proteins.

The presence of polypeptides of other molecular weights that was detected with both the anti–PAD-2 and anti–PAD-4 antibodies raises questions regarding the specificity of these reagents. Therefore, caution should be taken when using these reagents in assay systems such as immunohistochemistry, since there is no control for molecular weight, and therefore, there is the possibility of false-positive results.

In the present study we have provided evidence that the synovial fluid from patients with inflammatory arthritis is a site of expression of citrullinated proteins and up-regulation of α-enolase. The presence of PAD-2 and PAD-4, in which the extracellular environment would favor activation of the enzymes, supports the hypothesis that the synovial fluid may be an important site of expression of citrullinated proteins in inflammatory arthritis. The restriction of the immune response to patients with RA may help explain the chronic autoimmune response in the rheumatoid joint, characteristic of this disease.

ACKNOWLEDGMENTS

We thank Ms Alex Martin (Imperial College, London) for her expert technical support, and the European Union's Sixth Framework Programme project AutoCure for providing a forum for discussion and collaboration.

AUTHOR CONTRIBUTIONS

Dr. Venables had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study design. Kinloch, Lundberg, Lim, Venables.

Acquisition of data. Kinloch, Lundberg, Wait, Wegner, Lim, Zendman, Saxne, Malmström, Venables.

Analysis and interpretation of data. Kinloch, Lundberg, Wait, Wegner, Lim, Saxne, Malmström, Venables.

Manuscript preparation. Kinloch, Lundberg, Wait, Saxne, Malmström, Venables.

REFERENCES

1. Kinloch A, Lundberg K, Moyes D, Venables P. Pathogenic role of antibodies to citrullinated proteins in rheumatoid arthritis. Expert Rev Clin Immunol 2006;2:365–75.
2. Vossenaar ER, Zendman AJ, van Venrooij WJ, Pruijn GJ. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. Bioessays 2003;25:1106–18.
3. Vossenaar ER, Radstake TR, van der Heijden A, van Mansum MA, Dieteren C, de Rooij DJ, et al. Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages. Ann Rheum Dis 2004;63:373–81.
4. De Rycke L, Nicholas AP, Cantaert T, Kruithof E, Echols JD, Vandekerckhove B, et al. Synovial intracellular citrullinated proteins colocalizing with peptidyl arginine deiminase as pathophysiologically relevant antigenic determinants of rheumatoid arthritis–specific humoral autoimmunity. Arthritis Rheum 2005; 52:2323–30.
5. Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, et al. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 2003;34:395–402.
6. Chapuy-Regaud S, Sebbag M, Baeten D, Clavel C, Foulquier C, De Keyser F, et al. Fibrin deimination in synovial tissue is not specific for rheumatoid arthritis but commonly occurs during synovitides. J Immunol 2005;174:5057–64.
7. Makrygiannakis D, af Klint E, Lundberg I, Lofberg R, Ulfgren AK, Klareskog L, et al. Citrullination is an inflammation-dependent process. Ann Rheum Dis 2006;65:1219–22.
8. Masson-Bessiere C, Sebbag M, Durieux J, Nogueira L, Vincent C, Girbal-Neuhauser E, et al. In the rheumatoid pannus, anti-filaggrin autoantibodies are produced by local plasma cells and constitute a higher proportion of IgG than in synovial fluid and serum. Clin Exp Immunol 2000;119:544–52.
9. Spadaro A, Riccieri V, Alessandri C, Scrivo R, Valesini G. Usefulness of anti-cyclic citrullinate peptide antibody determina-

tion in synovial fluid analysis of patients with rheumatoid arthritis. Reumatismo 2006;58:116–20. In Italian.
10. Masson-Bessiere C, Sebbag M, Girbal-Neuhauser E, Nogueira L, Vincent C, Senshu T, et al. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the α- and β-chains of fibrin. J Immunol 2001;166:4177–84.
11. Matsuo K, Xiang Y, Nakamura H, Masuko K, Yudoh K, Noyori K, et al. Identification of novel citrullinated autoantigens of synovium in rheumatoid arthritis using a proteomic approach. Arthritis Res Ther 2006;8:R175.
12. Tilleman K, Union A, Cantaert T, De Keyser S, Daniels A, Elewaut D, et al. In pursuit of B-cell synovial autoantigens in rheumatoid arthritis: confirmation of citrullinated fibrinogen, detection of vimentin, and introducing carbonic anhydrase as a possible new synovial autoantigen. Proteomics-Clin Apps 2007;1: 32–46.
13. Kinloch A, Tatzer V, Wait R, Peston D, Lundberg K, Donatien P, et al. Identification of citrullinated α-enolase as a candidate autoantigen in rheumatoid arthritis. Arthritis Res Ther 2005;7: R1421–9.
14. Lundberg K, Kinloch A, Allison H, Sriskandan S, Moyes D, Venables P. Mapping of citrullinated α-enolase antibodies to an immunodominant epitope with high sequence similarity to bacterial enolase. Ann Rheum Dis 2007;66 Suppl 1:A6–69.
15. Skriner K, Adolph K, Jungblut PR, Burmester GR. Association of citrullinated proteins with synovial exosomes. Arthritis Rheum 2006;54:3809–14.
16. Takizawa Y, Suzuki A, Sawada T, Ohsaka M, Inoue T, Yamada R, et al. Citrullinated fibrinogen detected as a soluble citrullinated autoantigen in rheumatoid arthritis synovial fluids. Ann Rheum Dis 2006;65:1013–20.
17. Bang H, Egerer K, Gauliard A, Luthke K, Rudolph PE, Fredenhagen G, et al. Mutation and citrullination modifies vimentin to a novel autoantigen for rheumatoid arthritis. Arthritis Rheum 2007; 56:2503–11.

ARTHRITIS & RHEUMATISM
Vol. 58, No. 10, October 2008, pp 3009–3019
DOI 10.1002/art.23936

Antibodies to Citrullinated α-Enolase Peptide 1 Are Specific for Rheumatoid Arthritis and Cross-React With Bacterial Enolase

Karin Lundberg,[1] Andrew Kinloch,[1] Benjamin A. Fisher,[1] Natalia Wegner,[1] Robin Wait,[1] Peter Charles,[1] Ted R. Mikuls,[2] and Patrick J. Venables[1]

Objective. **To map the antibody response to human citrullinated α-enolase, a candidate autoantigen in rheumatoid arthritis (RA), and to examine cross-reactivity with bacterial enolase.**

Methods. **Serum samples obtained from patients with RA, disease control subjects, and healthy control subjects were tested by enzyme-linked immunosorbent assay (ELISA) for reactivity with citrullinated α-enolase peptides. Antibodies specific for the immunodominant epitope were raised in rabbits or were purified from RA sera. Cross-reactivity with other citrullinated epitopes was investigated by inhibition ELISAs, and cross-reactivity with bacterial enolase was investigated by immunoblotting.**

Results. **An immunodominant peptide, citrullinated α-enolase peptide 1, was identified. Antibodies to this epitope were observed in 37–62% of sera obtained from patients with RA, 3% of sera obtained from disease control subjects, and 2% of sera obtained from healthy control subjects. Binding was inhibited with homologous peptide but not with the arginine-containing control peptide or with 4 citrullinated peptides from elsewhere on the molecule, indicating that antibody binding was dependent on both citrulline and flanking amino acids. The immunodominant peptide showed 82% homology with enolase from *Porphyromonas gingivalis*, and the levels of antibodies to citrullinated α-enolase peptide 1 correlated with the levels of antibodies to the bacterial peptide ($r^2 = 0.803$, $P < 0.0001$). Affinity-purified antibodies to the human peptide cross-reacted with citrullinated recombinant *P gingivalis* enolase.**

Conclusion. **We have identified an immunodominant epitope in citrullinated α-enolase, to which antibodies are specific for RA. Our data on sequence similarity and cross-reactivity with bacterial enolase may indicate a role for bacterial infection, particularly with *P gingivalis,* in priming autoimmunity in a subset of patients with RA.**

Rheumatoid arthritis (RA) is a chronic inflammatory joint disorder that is considered to be autoimmune, although the autoantigens that trigger and sustain the immune response remain unknown. Over the last 10 years, investigations have shown that an essential feature of many autoantigens in RA is the posttranslational conversion of peptidyl arginine to peptidyl citrulline. Anti–citrullinated protein antibodies (ACPAs) are highly specific (98%) and sensitive (up to 80%) for RA (1,2), making citrullinated proteins strong candidates for driving the autoimmune response in this disease (for review, see ref. 3). Kuhn et al, for example, recently demonstrated that administration of anti–citrullinated fibrinogen antibodies enhanced disease severity in an

Supported by grants from the Arthritis Research Campaign and the Medical Research Council UK. Dr. Mikuls' work was supported by the NIH (grant AR-054539 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases).

[1]Karin Lundberg, MSc, PhD, Andrew Kinloch, MSc, Benjamin A. Fisher, MRCP, Natalia Wegner, MSc, MRES, Robin Wait, MA, PhD, Peter Charles, CSci, FIBMS, Patrick J. Venables, MD, FRCP: Imperial College London, London, UK; [2]Ted R. Mikuls, MD, MSPH: University of Nebraska and Omaha Veterans Affairs Hospital, Omaha, Nebraska.

Dr. Mikuls has received consulting fees, speaking fees, and/or honoraria from Amgen, Genentech, and Bristol-Myers Squibb (less than $10,000 each). Patent applications (GB0701417.8 and PCT/GB08/000267) have been filed to protect Drs. Venables and Lundberg's and Mr. Kinloch's intellectual property as coinventors of the discovery of citrullinated α-enolase peptide 1 (CEP-1) and other sequences from citrullinated α-enolase and their use in the diagnosis and treatment of rheumatoid arthritis; the coinventors receive no royalties.

Address correspondence and reprint requests to Karin Lundberg, MSc, PhD, Kennedy Institute of Rheumatology, Imperial College, 1 Aspenlea Road, London W6 8LH, UK. E-mail: k.lundberg@imperial.ac.uk.

Submitted for publication October 17, 2007; accepted in revised form June 27, 2008.

experimental mouse model of arthritis (4). In the majority of patients with RA, the presence of ACPAs antedates disease onset (5,6), and ACPA-positive patients present with a more severe and erosive form of arthritis (7–16). A commercial enzyme-linked immunosorbent assay (ELISA) based on cyclic citrullinated peptides (CCPs) detects ACPAs and is now routinely used in the diagnosis of RA.

Citrullination has a physiologic role in the generation of structural tissue such as skin, hair follicles, and the myelin sheaths of nerve fibers. In addition, the accumulation of citrullinated proteins has been described at sites of inflammation, including the joints of patients with all forms of arthritis (17), the brains of patients with multiple sclerosis (18) or Alzheimer's disease (19), and in the muscle fibers of patients with myositis (20). Hence, citrullinated proteins are present in the setting of both health and disease, while tolerance to citrullinated proteins appears to be selectively lost in patients with RA. Thus, in RA, it is the antibody response rather than the expression of antigen that is specific to the disease.

Genes and the environment interact in the development of this complex and heterogeneous disorder. Recent studies have demonstrated that the HLA–DRB1 shared epitope (SE) alleles, the best known genetic risk factor for RA, are associated with only anti-CCP antibody–positive RA, not anti-CCP antibody–negative RA, indicating that HLA SE alleles may be a specific risk factor for the production of ACPAs rather than the for RA itself (21,22). Furthermore, a strong gene–environment interaction between HLA SE alleles and cigarette smoking is present in anti-CCP antibody–positive patients but not in anti-CCP antibody–negative patients (21,23). Infectious agents, both bacterial and viral, have also been proposed as potential environmental stimuli (24–27), although to date, no single organism has survived as a compelling candidate for the etiology of the disease. Given that citrullinated proteins are target autoantigens in RA, the pathogen *Porphyromonas gingivalis,* which expresses the citrullinating enzyme peptidyl arginine deiminase (PAD) (28), could be an environmental trigger of RA in a manner similar to that proposed for smoking (21).

It is not clear which proteins harbor the epitopes targeted by ACPAs, although several candidates, including citrullinated fibrinogen (29), vimentin (30), and type II collagen (31), have been suggested, and we recently identified citrullinated α-enolase as another potential autoantigen (32). Alpha-enolase is abundantly expressed in the rheumatoid joint, antibodies targeting only the citrullinated form of the protein are specific for the disease, and our group recently demonstrated the presence in vivo of citrullinated α-enolase in synovial fluid from patients with RA (33). The molecule is also highly conserved throughout eukaryotes and prokaryotes and could therefore be a candidate for molecular mimicry between bacterial and host proteins (34). In the present study, we mapped the epitope of this anti–citrullinated α-enolase antibody response, using several citrullinated α-enolase peptides (CEPs), with the aim of examining cross-reactivity with bacterial enolase, which could prime autoimmunity in a subset of patients with RA.

PATIENTS AND METHODS

Serum samples from patients and control subjects. Serum samples were obtained, with informed consent and ethics approval from the Regional Research Ethics committee, from 102 consecutive patients with RA who were attending the Rheumatology Clinic at Charing Cross Hospital, London. The disease control group comprised 110 patients with other rheumatic diseases, including systemic lupus erythematosus (n = 32), Sjögren's syndrome (n = 31), Behçet's syndrome (n = 18), psoriatic arthritis (n = 5), and miscellaneous other rheumatic diseases (n = 24). Ninety-two control serum samples were obtained from healthy volunteers. Serum samples from 20 patients with spondylarthritides who were attending the Rheumatology Clinic at the Karolinska University Hospital in Stockholm, Sweden, were also collected, with informed consent and local ethics approval. In addition, serum samples were collected from an independent US cohort of 81 patients with RA obtained at baseline in clinical trials conducted by the Rheumatoid Arthritis Investigational Network (RAIN; coordinating center, University of Nebraska, Omaha) and from 82 age- and sex-matched healthy volunteers. All US samples were obtained with ethics approval from the local institutional review board. All patients with RA met the American College of Rheumatology (formerly, the American Rheumatism Association) 1987 revised criteria for the classification of RA (35).

Peptides. Fifteen cyclic 15–23-mer peptides were synthesized at Cambridge Research Biochemicals (Billingham, Cleveland, UK). The peptide sequences corresponded to amino acid sequences in human α-enolase (Swiss-Prot accession no. P06733) or *P gingivalis* enolase (Swiss-Prot accession no. AAQ66821), with the addition of cysteine residues at the amino and carboxy termini and the exchange of arginine for citrulline residues at certain positions (Table 1).

ELISAs. Ninety-six–well plates (MaxiSorp; Nunc, Roskilde, Denmark) were coated with α-enolase peptides at 10 μg/ml (diluted in a 50-m*M* carbonate buffer, pH 9.6) or with carbonate buffer alone and incubated overnight at 4°C. Wells were washed with phosphate buffered saline (PBS)–Tween (0.05%) and blocked with 2% bovine serum albumin (BSA) diluted in PBS for 3 hours at room temperature. Sera were diluted 1:50 in radioimmunoassay (RIA) buffer (10 m*M* Tris, 1% BSA, 350 m*M* NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate) supplemented with 10% fetal calf serum (FCS), added in duplicates, and incubated for 1.5 hours at room temperature. Plates were washed as described above and incubated with peroxidase-

Table 1. Alpha-enolase peptide sequences and the antibody response to these peptides in patients with RA and controls*

Peptide name	Sequence	RA (n = 102)	Disease controls (n = 110)	Healthy controls (n = 92)
1A (CEP-1)	ckiha-X-eifds-X-gnptvec	37	3†	2
1A *P gingivalis*	ckiig-X-eilds-X-gnptvec	34	1†	ND
1B	ckiha-R-eifds-X-gnptvec	40	ND	ND
1C	ckiha-X-eifds-R-gnptvec	20	ND	ND
1D	ckiha-R-eifds-R-gnptvec	5	2	3
2	cvdlftskglf-X-aavpsgc	10	0	2
3	cyealel-X-dndkt-X-ymgkgvskc	15	2	0
4	cekgvply-X-hiadlagnsec	16	2	2
5	cnf-X-eam-X-igaevyhnlknc	11	4	3
6	cf-X-sgkydldfkspddps-X-yc	12	5	0
7	cangwgvmvsh-X-sgetedtc	4	2	2
8	cakynqll-X-ieeelgskac	7	4	0
9	ckfag-X-nf-X-nplakc	15	1	1
10	ciqvvgddltvtnpk-X-iakac	6	ND	1
11	cnpk-X-iakavnekscnclc	18	ND	0

* Values are the percent positive, using an optical density at 450 nm cutoff value of 0.1. RA = rheumatoid arthritis; X = citrulline; R = arginine; ND = not done.
† A total of 130 disease controls were screened for reactivity with human and *Porphyromonas gingivalis* citrullinated α-enolase peptide 1 (CEP-1).

conjugated mouse anti-human IgG (Hybridoma Reagent Laboratory, Baltimore, MD) (diluted 1:1,000 in RIA buffer, 10% FCS) for 1 hour at room temperature. After a final wash (PBS–Tween, 0.05%), bound antibodies were detected with tetramethylbenzidine substrate (KPL, Gaithersburg, MD). The reaction was stopped by the addition of 1*M* H_2SO_4, and absorbance was measured at 450 nm in a Multiscan Ascent microplate reader (Thermolab Systems, Franklin, MA). A control serum was included on all plates to correct for plate-to-plate variation. The value for background optical density at 450 nm (OD_{450}) (wells coated with carbonate buffer alone) was subtracted from the peptide OD_{450} value. OD values above 0.1 were considered to be positive. ELISA of the serum samples obtained from the US cohort was performed in a similar manner but with 2% BSA in carbonate buffer as background and with an OD_{450} cutoff of 0.2 for positive samples. Anti-CCP antibody status was analyzed using the CCP2 kit (Eurodiagnostica, Malmo, Sweden), according to the manufacturer's instructions.

Inhibition assays. Inhibition experiments were performed in liquid phase using serum samples obtained from 6 anti–CEP-1–positive patients with RA or in solid phase using serum samples obtained from 9 double-positive (anti–CEP-1 positive and CCP positive) patients with RA. Sera were preincubated for 2 hours in RIA buffer containing increasing concentrations (0, 1, 10, and 100 μg/ml) of CEP-1, control peptide 1D, peptide 11, 4, or 9 (representing the second, third, and fourth most reactive peptides), or peptide 7 (representing the peptide with the lowest degree of reactivity). Alternatively, serum specimens were preabsorbed on plates coated with CEP-1 or carbonate buffer alone or were preabsorbed on commercial CCP2 plates. The liquid phase mixtures were centrifuged at 16,200*g* for 15 minutes before the supernatants were transferred to CEP-1–coated plates and assessed as described above for ELISAs. Inhibition of binding to CEP-1, for serum samples preabsorbed to CCP, was calculated in relation to the maximum inhibition, which by definition was set to 100% for samples preincubated with CEP-1.

Generation of anti–CEP-1 antibodies. A rabbit polyclonal anti–CEP-1 antibody was generated at Cambridge Research Biochemicals. Briefly, 2 rabbits were immunized subcutaneously every 2 weeks, for 10 weeks, with 200 μg keyhole limpet hemocyanin–conjugated peptide 1A in Freund's incomplete adjuvant per boost. Blood was collected 7 days after each injection, and sera were analyzed for the presence of anti–CEP-1 antibodies. When antibody titers reached significant levels, the animals were killed and their blood was harvested. The crude antisera were depleted of cross-reactive antibodies by chromatography on a thiopropyl–Sepharose column conjugated to peptide 1D. The unbound fraction was affinity-purified on a second thiopropyl–Sepharose column conjugated to peptide 1A. Anti–CEP-1–specific antibodies were eluted and further depleted of nonspecific antibodies by 3 subsequent passages through the depleting column. Human anti–CEP-1 antibodies from a patient with RA were purified with high titers of anti–CEP-1 antibodies. The serum was passed through a thiopropyl–Sepharose column conjugated to peptide 1A. Bound antibodies were eluted, after repeated PBS washes, using 3*M* GuHCl, and dialyzed against PBS. Purified rabbit anti–CEP-1 antibodies (0.39 mg/ml) and human anti–CEP-1 antibodies (0.6 μg/ml) were stored at −20°C until used further.

ELISA to assess CEP-1 specificity of affinity-purified anti–CEP-1 antibodies. Affinity-purified anti–CEP-1 antibodies (rabbit and human) were tested for their anti–CEP-1 specificity in serial dilutions (starting at 1:50 for the rabbit anti–CEP-1 antibody and at 1:2 for the human anti–CEP-1 antibody) on plates coated with peptide 1A or peptide 1D and assayed as described above for ELISAs.

Cloning and expression of human and bacterial enolase. Full-length human enolase and *P gingivalis* enolase were amplified by polymerase chain reaction (PCR), from vitamin D3–differentiated HL-60 cells and *P gingivalis* strain W83 (no.

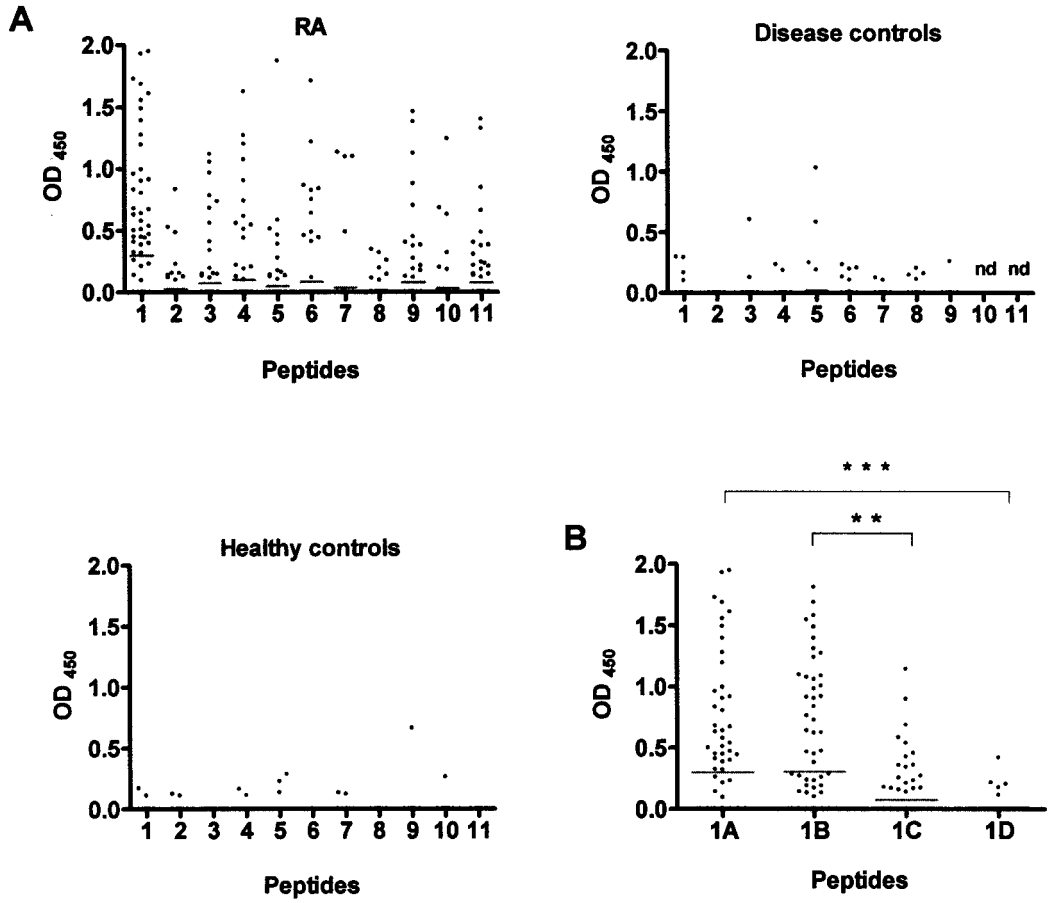


Figure 1. A, IgG response to citrullinated α-enolase peptides in patients with rheumatoid arthritis (RA; n = 102), disease controls (n = 110), and healthy controls (n = 92). **B,** Antibody response to the immunodominant peptides 1A, 1B, 1C, and 1D in patients with RA. The response to peptide 1A is dependent on the presence of citrulline, as demonstrated by the low antibody levels to the arginine-containing control peptide 1D. The second citrulline residue, rather than the first, is more important for antibody binding (compare peptides 1B and 1C). IgG was measured by enzyme-linked immunosorbent assay, and data are presented as the peptide optical density at 450 nm (OD_{450}) value with the background (carbonate buffer) OD_{450} value subtracted. Each dot represents the reactivity of 1 serum sample with the indicated peptide. Bars indicate the mean. ** = $P < 0.01$; *** = $P < 0.0001$. ND = not done.

BAA-308D-5; American Type Culture Collection, Rockville, MD), respectively. PCR products were ligated between *Bam* HI and *Xho* I restriction sites in a pGEX 6P3 expression vector (GE Healthcare, Bucks, UK), 3′ to the glutathione transferase (GST) coding site. GST–enolase protein expression was induced by isopropyl thiogalactose in the protease-deficient BL21 strain of *Escherichia coli*. Protein was purified using glutathione–Sepharose 4B (GE Healthcare), and the GST moiety was cleaved using PreScission Protease (GE Healthcare). Purified enolase, as determined by Coomassie staining and tandem mass spectometry analysis, was dialyzed against PBS and stored at −20°C until used further.

In vitro deimination. Recombinant human enolase and bacterial enolase were diluted to a concentration of 0.3 mg/ml in PAD buffer (0.1*M* Tris HCl, pH 7.6, 10 m*M* $CaCl_2$, 5 m*M* dithiothreitol) and incubated with rabbit skeletal PAD (Sigma, St. Louis, MO) at a concentration of 7 units/mg protein, for 3 hours at 50°C. Citrullination was terminated by the addition of 20 m*M* EDTA. Control proteins were treated similarly, apart from the addition of PAD. All samples were stored at −20°C until used further.

Silver staining and immunoblotting. Recombinant human enolase and *P gingivalis* enolase were electrophoresed on 4–12% NuPAGE Bis-Tris gels (Invitrogen, Paisley, UK) before silver staining, using a standard protocol, or before transfer to nitrocellulose membranes for immunoblotting. Briefly, membranes were blocked with 5% nonfat milk and incubated with rabbit anti–CEP-1 antibody (diluted 1:25) or human anti–CEP-1 antibody, diluted 1:2. Proteins were detected using peroxidase-conjugated secondary antibody (goat

anti-rabbit IgG, diluted 1:2,000; Dako, Glostrup, Denmark) or mouse anti-human IgG, diluted 1:500 (Hybridoma Reagent Laboratory, Baltimore, MD). Membranes were developed using the enhanced chemiluminescence technique (Amersham Biosciences, Little Chalfont, UK). Citrullinated proteins were detected using the Anti-Citrulline (Modified) Detection Kit (Upstate Biotechnology, Lake Placid, NY), in accordance with the manufacturer's instructions.

Statistical analysis. All statistical analyses were performed using the Mann-Whitney U test for independent groups.

RESULTS

Detection of a disease-specific antibody response to CEPs in patients with RA. Eleven citrullinated peptides (peptides 1–11) (Table 1), covering 15 of the 17 arginine residues within α-enolase, were selected on the basis of containing >1 potential citrulline residue within a 20–amino acid sequence, or by the demonstration (by mass spectrometry) in our previous study (32) that arginines had been deiminated to citrulline in vitro (32). Each peptide synthesized was tested for reactivity in 102 patients with RA, 110 disease control subjects, and 92 healthy control subjects. The results showed that 64% of patients with RA and 15% of the control subjects had IgG antibodies to 1 or several CEPs. The pattern of reactivity in patients with RA varied, with the majority of patients having an antibody response to multiple CEPs. In contrast, the antibody response in the control subjects was mainly restricted to one of the peptides, and the IgG antibody levels were significantly lower than those observed in patients with RA (Figure 1A).

Identification of an immunodominant B cell epitope within citrullinated α-enolase. Peptide 1 (1A) was the immunodominant peptide, which reacted with 37% of the RA serum samples compared with 2% of the healthy control samples and 3% of the disease control samples (Table 1). To map the antibody epitope further and to investigate the citrulline dependence of this antibody response, another 3 peptides (peptides 1B, 1C, and 1D) were synthesized (Table 1). The low reactivity of control peptide 1D, which does not contain any citrulline residues, confirms the importance of citrulline in this epitope (Figure 1B). The proportion of patients with RA whose sera reacted with peptide 1B (40%) was similar to the proportion whose sera reacted with peptide 1A but was higher than the proportion whose sera reacted with peptide 1C (20%). This, together with the higher antibody levels to peptides 1A and 1B compared with peptide 1C (Figure 1B), indicates that it is the second citrulline residue, rather than the first, that is more important for antibody recognition. Peptide 1A, containing both citrullinated residues, was chosen for further study and is referred to as CEP-1.

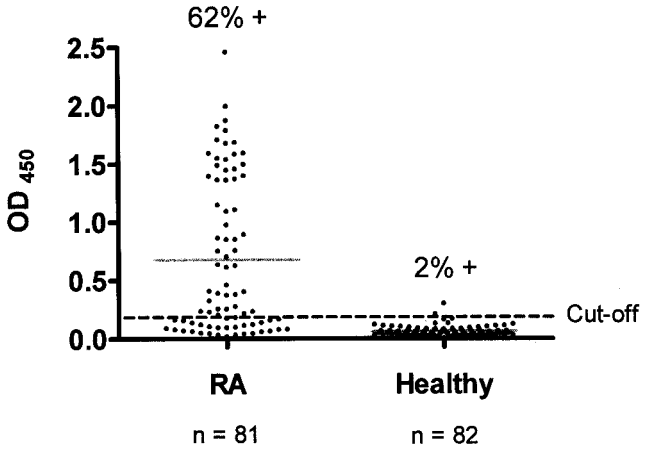


Figure 2. Anti–citrullinated α-enolase peptide 1 antibodies in the US cohort. Data were generated by enzyme-linked immunosorbent assay and are presented as the peptide optical density at 450 nm (OD_{450}) value with the background OD_{450} value subtracted. Broken line indicates the cutoff value for positive samples, based on the OD value for the ninety-eighth percentile of control subjects. Bar indicates the mean. RA = rheumatoid arthritis.

Anti-CCP antibodies in the same serum samples gave a diagnostic sensitivity of 71%, with a specificity of 98%. Seven (23%) of the anti-CCP antibody–negative patients with RA had positive results on the anti–CEP-1 ELISA (data not shown). Hence, combining the anti-CCP with the anti–CEP-1 ELISA results increased the overall sensitivity of ACPAs in this cohort to 78%.

Confirmation of the diagnostic sensitivity and specificity of the anti–CEP-1 ELISA in an independent cohort of patients with RA. To confirm that the high levels of antibodies to CEP-1 were not a peculiarity of the patients in our study, we used the anti–CEP-1 antibody ELISA to test 81 patients with RA and 82 healthy control subjects from a US cohort. To compare the data with the UK cohort, the OD value for the ninety-eighth percentile of control subjects was used as the cutoff point for positive samples. The sensitivity was increased to 62% in the US RA population, with a specificity of 98% (Figure 2). Disease controls were not examined in this part of the study.

Anti–CEP-1–specific as well as CCP cross-reactive antibodies in patients with RA. The epitope specificity of the anti–CEP-1 antibody response was tested in 2 separate inhibition assays. In the first experiment, inhibition of binding to CEP-1 was evaluated in sera from 6 anti–CEP-1–positive patients with RA, using

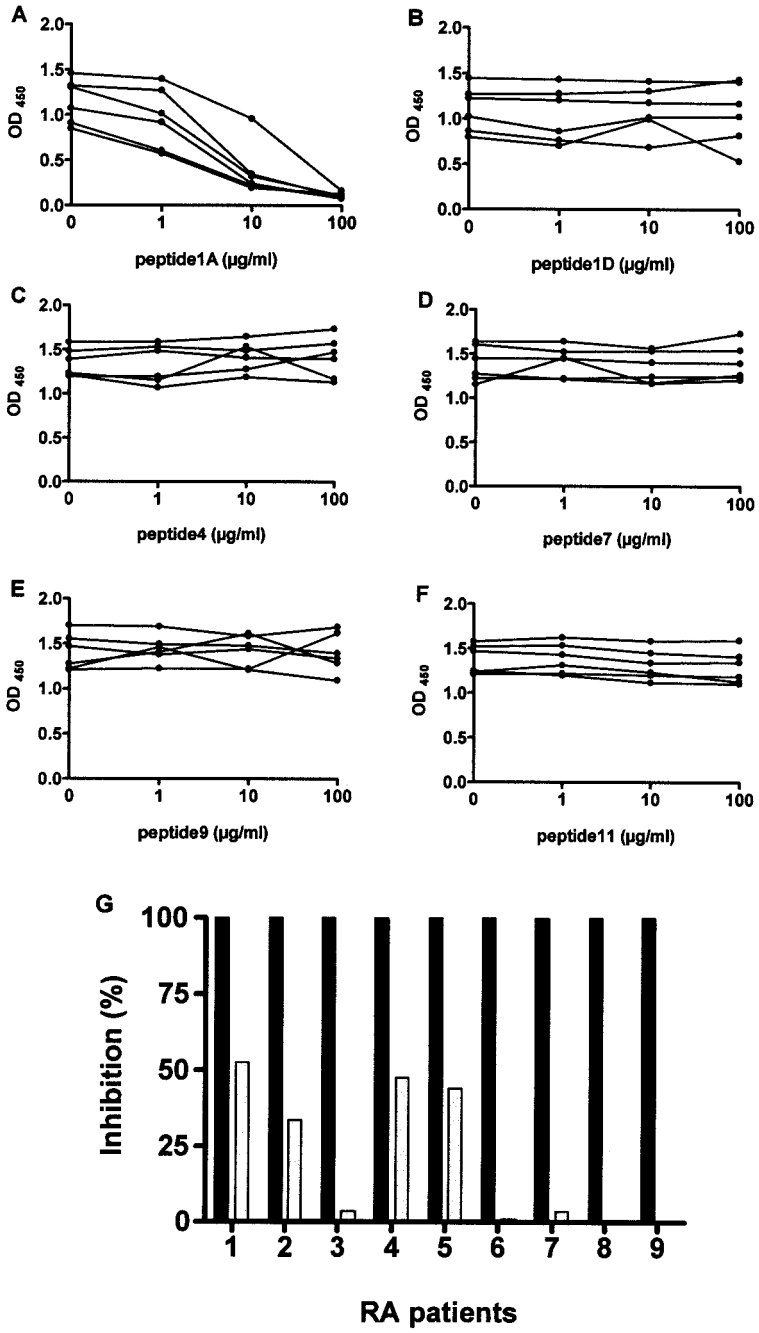


Figure 3. A–F, No cross-reaction of anti–citrullinated α-enolase peptide 1 (anti–CEP-1) antibodies with other citrullinated α-enolase epitopes. A dose-dependent decrease in binding to CEP-1–coated plates was seen in 6 anti–CEP-1–positive serum samples from patients with rheumatoid arthritis (RA), preincubated with increasing concentrations (0, 1, 10, and 100 μg/ml) of CEP-1 (**A**), while preincubation with the control peptide 1D (**B**), peptide 4 (**C**), peptide 7 (**D**), peptide 9 (**E**), or peptide 11 (**F**) did not result in decreased antibody binding. **G,** Differing degrees of cross-reactivity with cyclic citrullinated peptide (CCP) in individual serum specimens from 9 double-positive (anti–CEP-1 positive and anti-CCP positive) patients with RA following preincubation with CCP (shaded bars), as demonstrated by a reduction in CEP-1 reactivity that ranged from 0% to 53%. Results are expressed as the percentage of maximum inhibition, which by definition was set to 100% for preincubation with CEP-1 (solid bars). OD_{450} = optical density at 450 nm.

liquid phase. There was a dose-dependent inhibition by the homologous peptide, while there was no inhibition by the arginine-containing control peptide 1D or by the nonreactive citrullinated α-enolase peptide 7. Also, there was no inhibition when using citrullinated α-enolase peptides representing other reactive epitopes on the molecule, i.e., peptides 4, 9, and 11 (Figures 3A–F). In the second experiment, in which we examined cross-reactivity to CCP, it was necessary to use the commercially available CCP2 plates in a solid phase assay, because the sequences of the peptides used in the assay have not been published. In 9 double-positive (anti–CEP-1 positive and anti–CCP positive) RA samples, preabsorption on CCP plates showed inhibition of anti–CEP-1 binding, varying from 0% in 2 of the sera and increasing to a maximum of 53% in 1 sample, suggesting variable cross-reactivity of the antibodies (Figure 3G).

High sequence similarity between human and bacterial CEP-1. Human enolase and *P gingivalis* (Swiss-Prot accession no. AAQ66821) enolase were 51% identical at the amino acid level, across the whole protein. However, when comparing the sequence for peptide 1 (amino acids 5–21), the sequence identity increased to 82%, and the 9 amino acids spanning the immunodominant epitope on peptide 1 (amino acids 13–21) were 100% identical.

Antibodies to *P gingivalis* CEP-1 in patients with RA, and cross-reaction of purified anti–CEP-1 antibodies with bacterial enolase. We tested the 102 RA patients from the UK cohort for reactivity with the *P gingivalis* version of CEP-1, and 34% had positive results (Table 1); IgG antibody levels were similar to those observed for the human version of this peptide (data not shown). Sera from the 110 disease control subjects and from 20 patients with spondylarthritides were also tested for reactivity with *P gingivalis* CEP-1, and only 1 was positive (Table 1). In addition, serum samples from the US cohort were analyzed for anti–*P gingivalis* CEP-1 reactivity, and 54% of the patients with RA and 2% of the healthy control subjects had positive results. This antibody response correlated strongly ($r^2 = 0.803$, $P < 0.0001$) with the antibody response directed to the human version of CEP-1 (Figure 4).

To investigate whether antibodies directed to the immunodominant epitope of human α-enolase also recognized epitopes in *P gingivalis* enolase, recombinant *P gingivalis* enolase was in vitro treated with rabbit skeletal PAD or was left untreated, before immunoblotting with affinity-purified anti–CEP-1 antibodies. These antibodies were generated in rabbits or purified from RA sera,

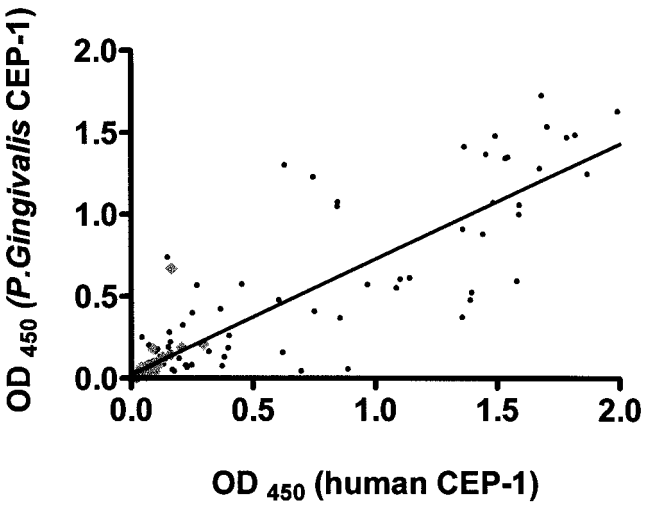


Figure 4. IgG antibodies to the *Porphyromonas gingivalis* version of CEP-1 in patients with RA (n = 81) and healthy control subjects (n = 82), as measured by enzyme-linked immunosorbent assay. Fifty-four percent of patients with RA were positive for anti–*P gingivalis* CEP-1 antibodies (circles), compared with 2% of healthy controls (diamonds). The antibody response to *P gingivalis* CEP-1 (y-axis) correlates strongly with the antibody response to the human version of this peptide (x-axis) ($r^2 = 0.803$, $P < 0.0001$). See Figure 3 for definitions.

and their specificity for CEP-1 was analyzed by ELISA. Serial dilutions of the rabbit polyclonal antibody (Figure 5A) showed a higher CEP-1 specificity than affinity-purified human immunoglobulins (Figure 5B), which also showed reactivity with the arginine-containing control peptide.

Silver staining (Figure 5C) demonstrated that *P gingivalis* enolase migrated at a slightly lower molecular weight than human α-enolase, consistent with the theoretical molecular weights of 45.8 kd for *P gingivalis* enolase and 47.2 kd for human α-enolase.

Western blots, using the Anti-Citrulline (Modified) Detection Kit, also confirmed that *P gingivalis* enolase had been successfully citrullinated in vitro by eukaryotic PAD (Figure 5D). The purified anti–CEP-1 antibodies demonstrated cross-reactivity with citrullinated *P gingivalis* enolase. Both the rabbit (Figure 5E) and the human (Figure 5F) anti–CEP-1 antibody preferentially bound the citrullinated form of the protein but also showed cross-reactivity to the noncitrullinated form. Control blots, using goat anti-rabbit IgG, mouse anti-human IgG, the anti–modifed citrulline antibody on partially modified membranes, or the goat anti-rabbit IgG secondary on modified membranes were all negative (data not shown).

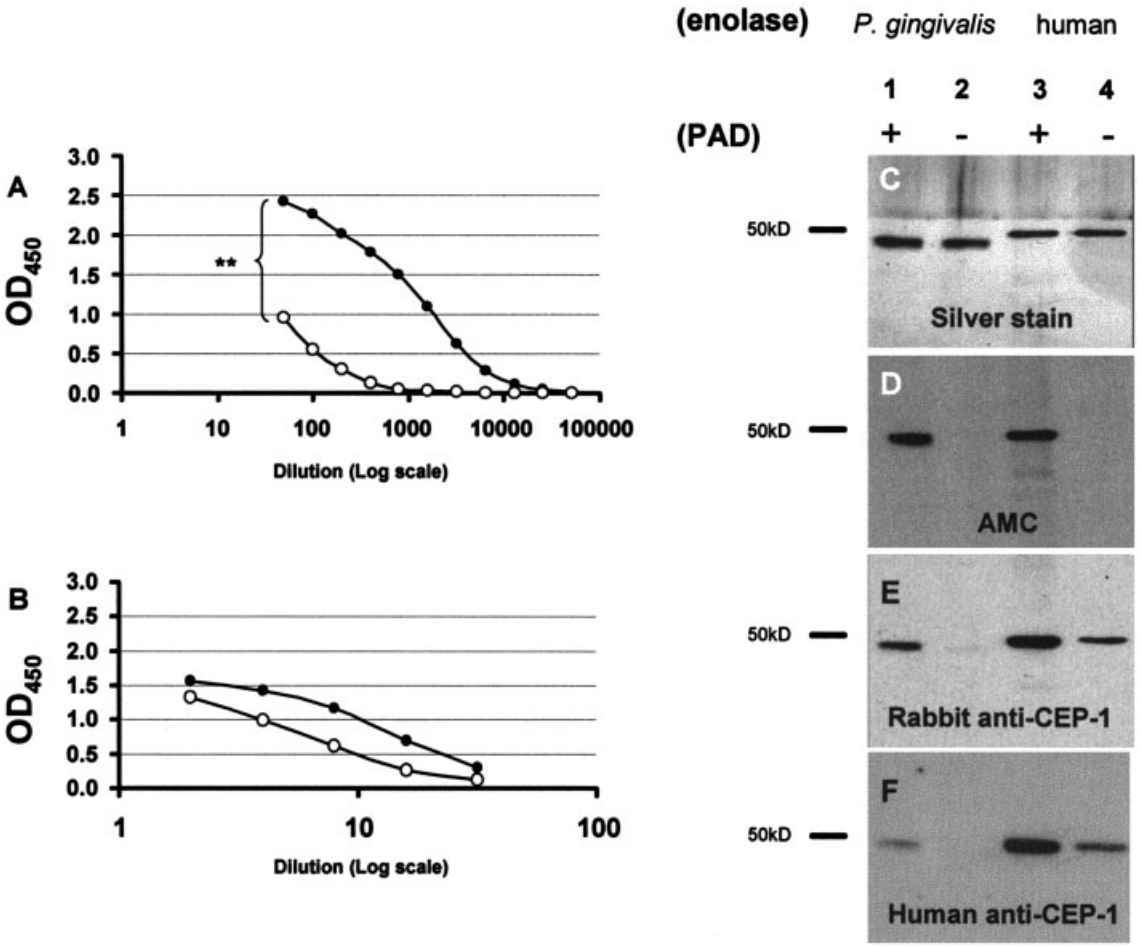


Figure 5. **A** and **B,** Enzyme-linked immunosorbent assays. **A,** CEP-1 specificity for the rabbit polyclonal antibody and **B,** cross-reactivity with the control peptide for the affinity-purified human antibody were demonstrated. IgG responses to CEP-1 (solid circles) and to the control peptide (open circles) were measured in serial dilutions. **C–F,** Silver staining and immunoblotting. **C,** Silver staining showed recombinant *Porphyromonas gingivalis* enolase (lanes 1 and 2) migrating at a slightly lower molecular weight than recombinant human α-enolase (lanes 3 and 4). **D,** Successful in vitro citrullination by rabbit skeletal peptidyl arginine deiminase (PAD) was demonstrated for *P gingivalis* enolase (lane 1) and human α-enolase (lane 3). **E** and **F,** Noncitrullinated *P gingivalis* enolase (lane 2) and human α-enolase (lane 4) were run in parallel as negative controls. Affinity-purified rabbit and human anti–CEP-1 antibodies cross-reacted with citrullinated *P gingivalis* enolase (lane 1), while a lower degree of reactivity was seen with the noncitrullinated form of the protein (lane 2). Citrullinated (lane 3) and noncitrullinated (lane 4) human α-enolase served as controls. Proteins were subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis. ** = $P < 0.01$. AMC = anti–modified citrulline (see Figure 3 for other definitions).

DISCUSSION

In this report, we describe the identification of a dominant B cell epitope in citrullinated α-enolase, CEP-1, which is reactive with 37–62% of RA sera and 2% of healthy control sera. This epitope shows high sequence identity to bacterial enolase, and affinity-purified antibodies, specific for this epitope, react with both human and bacterial forms of the molecule. These data provide a new model for bacterial infection priming the autoimmune response in RA.

RA is diagnosed according to a set of clinical criteria and probably comprises different disease entities, with different pathogenic pathways leading to a similar clinical outcome. For example, 2 subgroups can clearly be distinguished based on the presence or absence of ACPAs (36). It is also possible that there is a

range of disease-driving antigens within the ACPA-positive group of patients with RA (37). In a previous study, we characterized citrullinated α-enolase as one such candidate (32). Using immunoblotting of whole protein, we observed that serum samples from 46% of patients with RA reacted with citrullinated α-enolase, of which 7 samples (13%) also recognized the noncitrullinated protein. Sera from 15% of healthy control subjects reacted with both forms of the molecule, and we speculated that this was attributable to reactivity with noncitrullinated epitopes. In the present study, we attempted to identify RA-specific antibodies by testing cyclic peptides consisting of α-enolase sequences surrounding citrulline residues. In doing so, we have identified a dominant epitope, CEP-1, with a diagnostic sensitivity of 37% and a specificity of 98%.

To confirm that this sensitivity and high specificity were not a peculiarity of the sera used at our institution, we also tested 81 RA samples and 82 healthy control samples obtained from a separate unit and found that the sensitivity was higher (62%), with a specificity of 98%. This difference may be attributable to the fact that the US cohort was derived from patients participating in clinical trials; therefore, patients in the US cohort may have had more active disease than did patients in the UK cohort, which was derived from patients attending a regular clinic, many of whom were receiving treatment. Antibodies to native α-enolase have also been demonstrated in patients with SLE and those with Behçet's disease (38,39), although in our study, patients with these diagnoses were uniformly negative for anti–CEP-1 antibodies, again supporting the concept that this epitope has specificity for patients with RA.

The importance of citrulline in the anti–CEP-1 antibody response was demonstrated by the low degree of reactivity to the arginine-containing control peptide. However, it was also clear that neighboring amino acids constitute important antigenic determinants, because the response to the other citrullinated α-enolase peptides was much weaker. The 2 amino acids, serine and glycine, flanking the second citrulline residue in CEP-1 have previously been reported to enhance ACPA recognition and binding (37,40). Thus, it was not surprising to observe that peptide 1B, containing the Ser-Cit-Gly motif, had a higher sensitivity than peptide 1C, which lacks this motif. Our inhibition experiments demonstrate that anti–CEP-1 antibodies do not cross-react with other immunoreactive citrullinated α-enolase peptides. This suggests that antibodies to citrullinated α-enolase react independently with multiple epitopes, which characterizes an antigen-driven immune response. This in turn supports the concept that citrullinated α-enolase is a true autoantigen in RA. The concordance between anti–CEP-1 and anti-CCP antibodies, as well as the results from our inhibition experiments showing a variable degree of cross-reactivity between the 2 antigens, suggests that citrullinated α-enolase may be one of a family of citrullinated autoantigens, for which antibodies are screened, at least in part, by the anti-CCP assay.

At this stage, we do not claim that the anti–CEP-1 ELISA is a new diagnostic assay for RA. Modification of coating conditions, the use of different configurations of peptides, and a standard curve to ensure accurate quantification and reproducibility may increase the sensitivity of the assay. These refinements, together with further analysis of large cohorts, may show that antibodies to CEP-1 define a specific clinical or immunogenetic subset of RA. Thus far, our data using 2 independent cohorts of RA patients and appropriate normal controls and disease controls support the concept that citrullinated α-enolase is a true autoantigen, and that the sequence of the immunodominant peptide representing CEP-1 may be important in the pathogenesis of the disease in the proportion of patients who have the antibody.

In light of the hypothesis that RA may be precipitated by infection, it was intriguing to observe that the sequence of 9 amino acids spanning the immunodominant epitope on CEP-1 was 100% identical to that encoded by *P gingivalis*. *P gingivalis* is a gram-negative bacterium that causes adult periodontitis. Like RA, adult periodontitis is a chronic inflammatory disorder in which the accumulation of immune cells leads to local production of proinflammatory cytokines such as tumor necrosis factor α and interleukin-1β, metalloproteinases, and prostaglandins, which results in tissue swelling and degradation (for review, see refs. 41 and 42). RA is 4 times more common in patients with periodontitis than in the normal population (43), and higher levels of antibodies to *P gingivalis* (44), as well as a higher prevalence of advanced forms of periodontal destruction, have also been reported in patients with RA compared with control subjects. Finally, both rheumatoid factor production and anti-CCP antibody production have been associated with periodontal disease (41,45).

These data could imply a common underlying dysregulation of the host immune response in adult periodontitis and RA. However, there is also a striking genetic similarity between the 2 diseases in that both are associated with the HLA SE alleles (46,47). This, together with the fact that *P gingivalis* is the only bacterium

known to synthesize its own PAD enzyme (28), lends support to another hypothesis: that *P gingivalis* may be the "septic stimulus" in RA, as proposed by Rosenstein et al (42). Therefore, individuals with periodontal infection may already be exposed to citrullinated antigens, including citrullinated bacterial enolase, generated by host PAD during the inflammatory response or by bacterial PAD produced as a virulence factor to evade host defense (28). In the genetic context of the HLA SE alleles, and in the presence of danger signals, this could result in a pathologic immune response, with the formation of ACPAs. A subclinical arthritis could develop at a later time point, perhaps due to trauma or a viral infection, resulting in citrullination of synovial proteins. Under normal circumstances, this arthritis would be self-limiting. However, the presence of anti–CEP-1 antibodies could lead to cross-reactivity with citrullinated proteins in the joint and amplification of the inflammatory process, with progression to chronic RA.

We were, in fact, able to demonstrate such antibody cross-reactivity between human and bacterial enolase. Not only did patients with RA have antibodies reactive with the *P gingivalis* version of CEP-1, but anti–CEP-1 antibodies (raised in rabbits or purified from a patient with RA) bound to the immunodominant epitope in *P gingivalis* enolase. Hence, based on our data, we hypothesize that autoimmunity in the subset of RA patients with a humoral immune response to CEP-1 could be primed by bacterial infection, and that tolerance could be broken by citrullinated bacterial enolase. Cross-reactivity with citrullinated human α-enolase within the joint could then lead to the chronic destructive inflammation that characterizes RA.

ACKNOWLEDGMENTS

We wish to acknowledge Professor Dorian Haskard, Imperial College London, for providing serum samples from patients with Behçet's syndrome, and Dr. Vivianne Malmstrom and Professor Lars Klareskog, Karolinska Institutet, Stockholm, Sweden, for providing serum samples from patients with spondylarthritides. We would also like to thank James R. O'Dell, MD, as well as investigators and patients from the RAIN Network, for their contributions.

AUTHOR CONTRIBUTIONS

Dr. Lundberg had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study design. Lundberg Kinloch, Venables.

Acquisition of data. Lundberg, Kinloch, Fisher, Wegner, Wait, Charles, Mikuls, Venables.

Analysis and interpretation of data. Lundberg, Kinloch, Fisher, Wait, Venables.

Manuscript preparation. Lundberg, Kinloch, Mikuls, Venables.

Statistical analysis. Lundberg.

REFERENCES

1. Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000;43:155–63.
2. Van Venrooij WJ, Hazes JM, Visser H. Anticitrullinated protein/peptide antibody and its role in the diagnosis and prognosis of early rheumatoid arthritis. Neth J Med 2002;60:383–8.
3. Kinloch A, Lundberg K, Venables P. The pathogenic role of antibodies to citrullinated proteins in rheumatoid arthritis. Exp Rev Clin Immunol 2006;2:365–75.
4. Kuhn KA, Kulik L, Tomooka B, Braschler KJ, Arend WP, Robinson WH, et al. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. J Clin Invest 2006;116:961–73.
5. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;50:380–6.
6. Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.
7. Forslind K, Ahlmen M, Eberhardt K, Hafstrom I, Svensson B. Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: role of antibodies to citrullinated peptides (anti-CCP). Ann Rheum Dis 2004;63:1090–5.
8. Jansen LM, van Schaardenburg D, van der Horst-Bruinsma I, van der Stadt RJ, de Koning MH, Dijkmans BA. The predictive value of anti-cyclic citrullinated peptide antibodies in early arthritis. J Rheumatol 2003;30:1691–5.
9. Kastbom A, Strandberg G, Lindroos A, Skogh T. Anti-CCP antibody test predicts the disease course during 3 years in early rheumatoid arthritis (the Swedish TIRA project). Ann Rheum Dis 2004;63:1085–9.
10. Saraux A, Berthelot JM, Devauchelle V, Bendaoud B, Chales G, Le Henaff C, et al. Value of antibodies to citrulline-containing peptides for diagnosing early rheumatoid arthritis. J Rheumatol 2003;30:2535–9.
11. Vencovsky J, Machacek S, Sedova L, Kafkova J, Gatterova J, Pesakova V, et al. Autoantibodies can be prognostic markers of an erosive disease in early rheumatoid arthritis. Ann Rheum Dis 2003;62:427–30.
12. Kroot EJ, de Jong BA, van Leeuwen MA, Swinkels H, van den Hoogen FH, van 't Hof M, et al. The prognostic value of anti–cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. Arthritis Rheum 2000;43:1831–5.
13. Meyer O, Labarre C, Dougados M, Goupille P, Cantagrel A, Dubois A, et al. Anticitrullinated protein/peptide antibody assays in early rheumatoid arthritis for predicting five year radiographic damage. Ann Rheum Dis 2003;62:120–6.
14. Tamai M, Kawakami A, Uetani M, Takao S, Tanaka F, Nakamura H, et al. The presence of anti-cyclic citrullinated peptide antibody is associated with magnetic resonance imaging detection of bone marrow oedema in early stage rheumatoid arthritis. Ann Rheum Dis 2006;65:133–4.
15. Vallbracht I, Rieber J, Oppermann M, Forger F, Siebert U, Helmke K. Diagnostic and clinical value of anti-cyclic citrullinated

peptide antibodies compared with rheumatoid factor isotypes in rheumatoid arthritis. Ann Rheum Dis 2004;63:1079–84.
16. Van Gaalen FA, van Aken J, Huizinga TW, Schreuder GM, Breedveld FC, Zanelli E, et al. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. Arthritis Rheum 2004;50:2113–21.
17. Vossenaar ER, Smeets TJ, Kraan MC, Raats JM, van Venrooij WJ, Tak PP. The presence of citrullinated proteins is not specific for rheumatoid synovial tissue. Arthritis Rheum 2004;50:3485–94.
18. Moscarello MA, Wood DD, Ackerley C, Boulias C. Myelin in multiple sclerosis is developmentally immature. J Clin Invest 1994;94:146–54.
19. Ishigami A, Ohsawa T, Hiratsuka M, Taguchi H, Kobayashi S, Saito Y, et al. Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. J Neurosci Res 2005;80: 120–8.
20. Makrygiannakis D, af Klint E, Lundberg IE, Lofberg R, Ulfgren AK, Klareskog L, et al. Citrullination is an inflammation dependent process. Ann Rheum Dis 2006;65:1219–22.
21. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al, and the Epidemiological Investigation of Rheumatoid Arthritis Study Group. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;54:38–46.
22. Van der Helm-van Mil AH, Verpoort KN, Breedveld FC, Huizinga TW, Toes RE, de Vries RR. The HLA–DRB1 shared epitope alleles are primarily a risk factor for anti–cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum 2006;54: 1117–21.
23. Pedersen M, Jacobsen S, Klarlund M, Pedersen BV, Wiik A, Wohlfahrt J, et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. Arthritis Res Ther 2006;8:R133.
24. Origuchi T, Eguchi K, Kawabe Y, Yamashita I, Mizokami A, Ida H, et al. Increased levels of serum IgM antibody to staphylococcal enterotoxin B in patients with rheumatoid arthritis. Ann Rheum Dis 1995;54:713–20.
25. Takahashi Y, Murai C, Shibata S, Munakata Y, Ishii T, Ishii K, et al. Human parvovirus B19 as a causative agent for rheumatoid arthritis. Proc Natl Acad Sci U S A 1998;95:8227–32.
26. Blaschke S, Schwarz G, Moneke D, Binder L, Muller G, Reuss-Borst M. Epstein-Barr virus infection in peripheral blood mononuclear cells, synovial fluid cells, and synovial membranes of patients with rheumatoid arthritis. J Rheumatol 2000;27:866–73.
27. Saal JG, Steidle M, Einsele H, Muller CA, Fritz P, Zacher J. Persistence of B19 parvovirus in synovial membranes of patients with rheumatoid arthritis. Rheumatol Int 1992;12:147–51.
28. McGraw WT, Potempa J, Farley D, Travis J. Purification, characterization, and sequence analysis of a potential virulence factor from Porphyromonas gingivalis, peptidylarginine deiminase. Infect Immun 1999;67:3248–56.
29. Masson-Bessiere C, Sebbag M, Girbal-Neuhauser E, Nogueira L, Vincent C, Senshu T, et al. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the α- and β-chains of fibrin. J Immunol 2001;166:4177–84.
30. Vossenaar ER, Despres N, Lapointe E, van der Heijden A, Lora M, Senshu T, et al. Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin. Arthritis Res Ther 2004;6: R142–50.
31. Burkhardt H, Sehnert B, Bockermann R, Engstrom A, Kalden JR, Holmdahl R. Humoral immune response to citrullinated collagen type II determinants in early rheumatoid arthritis. Eur J Immunol 2005;35:1643–52.
32. Kinloch A, Tatzer V, Wait R, Peston D, Lundberg K, Donatien P, et al. Identification of citrullinated α-enolase as a candidate autoantigen in rheumatoid arthritis. Arthritis Res Ther 2005;7: R1421–9.
33. Kinloch A, Lundberg K, Wait R, Wegner N, Lim NH, Zendman AJ, et al. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. Arthritis Rheum 2008;58:2287–95.
34. Albert LJ, Inman RD. Molecular mimicry and autoimmunity. N Engl J Med 1999;341:2068–74.
35. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
36. Huizinga TW, Amos CI, van der Helm-van Mil AH, Chen W, van Gaalen FA, Jawaheer D, et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA–DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005;52:3433–8.
37. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101:273–81.
38. Pratesi F, Moscato S, Sabbatini A, Chimenti D, Bombardieri S, Migliorini P. Autoantibodies specific for α-enolase in systemic autoimmune disorders. J Rheumatol 2000;27:109–15.
39. Lee KH, Chung HS, Kim HS, Oh SH, Ha MK, Baik JH, et al. Human α-enolase from endothelial cells as a target antigen of anti–endothelial cell antibody in Behcet's disease. Arthritis Rheum 2003;48:2025–35.
40. Girbal-Neuhauser E, Durieux JJ, Arnaud M, Dalbon P, Sebbag M, Vincent C, et al. The epitopes targeted by the rheumatoid arthritis-associated antifilaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deimination of arginine residues. J Immunol 1999;162:585–94.
41. Mercado FB, Marshall RI, Bartold PM. Inter-relationships between rheumatoid arthritis and periodontal disease: a review. J Clin Periodontol 2003;30:761–72.
42. Rosenstein ED, Greenwald RA, Kushner LJ, Weissmann G. Hypothesis: the humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis. Inflammation 2004;28:311–8.
43. Mercado F, Marshall RI, Klestov AC, Bartold PM. Is there a relationship between rheumatoid arthritis and periodontal disease? J Clin Periodontol 2000;27:267–72.
44. Ogrendik M, Kokino S, Ozdemir F, Bird PS, Hamlet S. Serum antibodies to oral anaerobic bacteria in patients with rheumatoid arthritis. MedGenMed 2005;7:2.
45. Havemose-Poulsen A, Westergaard J, Stoltze K, Skjodt H, Danneskiold-Samsoe B, Locht H, et al. Periodontal and hematological characteristics associated with aggressive periodontitis, juvenile idiopathic arthritis, and rheumatoid arthritis. J Periodontol 2006; 77:280–8.
46. Katz J, Goultschin J, Benoliel R, Brautbar C. Human leukocyte antigen (HLA) DR4: positive association with rapidly progressing periodontitis. J Periodontol 1987;58:607–10.
47. Marotte H, Farge P, Gaudin P, Alexandre C, Mougin B, Miossec P. The association between periodontal disease and joint destruction in rheumatoid arthritis extends the link between the HLA-DR shared epitope and severity of bone destruction. Ann Rheum Dis 2006;65:905–9.

The International Journal of Biochemistry & Cell Biology 41 (2009) 390–397

Contents lists available at ScienceDirect

The International Journal of Biochemistry & Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Review

Evolutionarily conserved antigens in autoimmune disease: Implications for an infective aetiology

Natalia Wegner, Robin Wait, Patrick J. Venables*

Kennedy Institute of Rheumatology Division, Imperial College London, 65 Aspenlea Road, London W6 8LH, UK

ARTICLE INFO

Article history:
Available online 24 September 2008

Keywords:
Autoimmune disease
Conserved epitopes
Molecular mimicry
Infection
Rheumatoid
Evolution
Enolase

ABSTRACT

The immune system has evolved to eliminate or inactivate infectious organisms. An inappropriate response against self-components (autoantigens) can result in autoimmune disease. Here we examine the hypothesis that some evolutionarily conserved proteins, present in pathogenic and commensal organisms and their hosts, provide the stimulus that initiates autoimmune disease in susceptible individuals. We focus on seven autoantigens, of which at least four, glutamate decarboxylase, pyruvate dehydrogenase, histidyl-tRNA synthetase and alpha enolase, have orthologs in bacteria. Citrullinated alpha-enolase, a target for autoantibodies in 40% of patients with rheumatoid arthritis, is our main example. The major epitope is highly conserved, with over 90% identity to human in some bacteria. We propose that this reactivity of autoantibodies to shared sequences provides a model of autoimmunity in rheumatoid arthritis, which may well extend to other autoimmune disease in humans.

Contents

1. Introduction

The adaptive immune system provides an important line of defence against infectious agents and has evolved an array of tolerance mechanisms, which prevent inappropriate responses against self-(auto) antigens. Because the discrimination between self- and nonself-reactive lymphocytes is imperfect, some self-reactive lymphocytes may mature and, in the presence of danger signals from the innate immune system, become activated and cause autoimmune disease.

Autoimmune diseases are increasingly diagnosed by their associated autoantibodies. Traditionally, such antibodies are detected with immunological assays such as indirect immunofluorescence, which define antibodies targeting tissues including the thyroid, pancreas and gastric parietal cells. Using cell lines, antibodies to subcellular components, for example nuclei, nucleoli, ribosomes and mitochondria, can also be identified. The need for more disease-specific tests has driven the search for the actual molecules that are targeted by these antibodies, and has resulted in the dis-

Abbreviations: ANCA, anti-neutrophil cytoplasmic antigens; 2-OADC, 2-oxo-acid dehydrogenase complex; BLAST, basic local alignment search tool; CCP, cyclic citrullinated peptide; GABA, γ-aminobutyric acid; GAD, glutamate decarboxylase; HisRS, histidyl-tRNA synthetase; LPS, lipopolysaccharide; MS, multiple sclerosis; PAD, peptidyl arginine deiminase; PDC-E2, pyruvate dehydrogenase complex E2 component; PR3, proteinase 3; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.

* Corresponding author. Tel.: +44 20838334444.
E-mail address: p.venables@imperial.ac.uk (P.J. Venables).

1357-2725/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biocel.2008.09.012

Table 1
Organ-specific and non-organ-specific autoimmune diseases and their antigens.

	Disease	Generic autoantibody target	Specific autoantigen(s)	Bacterial orthologs	Reference
Organ specific	Hashimoto's thyroiditis	Thyrocyte	Thyroid peroxidase	?+	Daiyasu and Toh, 2000
	Grave's disease	Thyrocyte	Thyrotropin receptor		
	Type I diabetes mellitus	Pancreatic islet cell	Glutamic acid decarboxylase 65	+	García and López, 1995 Ueno, 2000
	Celiac disease	Endomyceal	Tissue transglutaminase		
	Autoimmune gastritis	Gastric parietal cell	Intrinsic factor		
	Myasthenia gravis	Acetylcholine receptor	Acetylcholine receptor	+	Bocquet et al., 2007[1] Hilf and Dutzler, 2008[2]
	Neuromyelitis optica (Devic's disease)	NMO IgG	Aquaporin-4	+	Calamita et al., 1995[3]
	Autoimmune haemolytic anaemia	Erythrocyte	various cell surface glycoproteins		
	Primary biliary cirrhosis	Mitochondria	Pyruvate dehydrogenase complex E2	+	Selmi et al, 2003
	Addison's diseae	Adrenal Cortex	21 hydroxylase		
	Goodpastures syndrome	Glomerular basement membrane	Type IV collagen		
	Sjogren's syndrome	Ro, La	Ro/SSA, La/SSB		
	Polymyositis	Jo-1 (and others)	Histidyl tRNA synthetase	+	Raben et al., 1994
	Rheumatoid Arthritis	Citrullinated (cit) proteins	(cit)fibrin, (cit)vimentin, (cit)collagen, (cit)α-enolase	+	Pancholi, 2001
	Systemic sclerosis	DNA binding proteins	Topoisomerase-1 Centromeric proteins A & B	+	Forterre et al., 2007[4]
	Wegener's granulomatosis	cANCA antigens	Proteinase-3	?+	
	Microscopic polyangiitis	pANCA antigens	Myeloperoxidase	?+	Daiyasu and Toh, 2000
Non-Organ Specific	Systemic lupus erythematosus	Nucleosomes Ribonucleoproteins Phospholipid	DNA, U1RNP, Sm, Ro, La B2-glycoprotein-1		

[+] Signifies existence of bacterial orthologs.
[a] Bocquet N, Prado de Carvalho L, Cartaud J, Neyton J, Le Poupon C, Taly A, Grutter T, Changeux J-P, Corringer P-J. A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. Nature 2007;445:116–9.
[b] Hilf RJC, Dutzler R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. Nature 2008;452:375–9.
[c] Calamita G, Bishai WR, Preston GM, Guggino WB, Agre P. Molecular cloning and characterization of AqpZ, a water channel from *Escherichia coli*. J Biol Chem 1995;270:29063–6.
[d] Forterre P, Gribaldo S, Gadelle D, Serre M-C. Origin and evolution of DNA topoisomerases. Biochimie 2007;89:427–46.

covery of autoantigens within the nucleus and other organelles, for which the autoantibody response is exquisitely antigen- and, in many cases, disease-specific. In parallel, there is accumulating evidence that some antigens are directly involved in pathogenesis.

Most autoimmune diseases have only a small number of associated autoantigens; some examples are shown in Table 1. While no doubt others will be discovered, the repertoire of human autoantibodies is relatively small (no more than a few hundred autoantigens) compared to around 21,000 genome-encoded proteins, implying considerable selectivity of the autoimmune response. Many autoantigens are members of evolutionarily conserved protein families which originated before the divergence of the prokaryotic and eukaryotic lineages. In this review we consider those autoantigens for which antibodies are diagnostically and/or

pathologically significant, rather than targets of autoantibodies that are present in healthy people. We have not addressed the issue of mechanisms, such as the evolution of the benign to a pathological autoimmune response, nor do we discuss T cell epitopes. We focus specifically on the stimulus that causes breakdown of tolerance and initiates or potentiates autoimmune disease. We propose that this stimulus is due to molecular similarities between prokaryotic and eukaryotic molecules, which in some cases may be due to evolutionary conservation.

2. Autoantigens and disease

For the purposes of this review, we use the accepted distinction between organ- and non-organ-specific autoimmunity, but have extended it to encompass a spectrum ranging from purely organ-specific to completely systemic (Table 1). In organ-specific autoimmunity, the antigens are usually hormones, enzymes or receptors largely restricted to a particular organ, often an endocrine or exocrine gland. The purest examples are Hashimoto's thyroiditis and Grave's disease, where both the autoantigens and the pathology are almost exclusively confined to the thyroid. Other important examples are type I diabetes, autoimmune gastritis and celiac disease. In non-organ-specific autoimmunity, the antigens are more widely distributed and may include nucleic acids, nucleoproteins, ribonucleoproteins, intracellular enzymes, membrane phospholipids, and plasma proteins. Most diseases in this group can loosely be defined as systemic autoimmune rheumatic disease. Systemic lupus erythematosus is the prototype, with antibodies directed to ubiquitous cellular antigens and involvement of multiple organs. Exceptions to this pattern are organ-specific diseases with ubiquitous autoantigens, such as polymyositis and primary biliary cirrhosis. Polymyositis is characterised by inflammation of muscle, but the main autoantigen (histidyl-tRNA synthetases) can be found in every human cell type. The same applies to primary biliary cirrhosis, an inflammatory liver disease, where the major autoantigen (pyruvate dehydrogenase complex E2 component) is present throughout the body.

Most of the specific antigens listed in Table 1 have only been described in the last decade and have enabled definitive classification of their respective diseases as truly autoimmune. A notable exception is multiple sclerosis (MS), where a number of candidate autoantibodies have been proposed, but none have been convincingly demonstrated to be disease-specific targets for the autoimmune response. It has even been argued that autoimmunity in MS may be entirely T-cell driven. The similarity of MS to rare diseases such as neuromyelitis optica (Devic's disease), for which the target antigen aquaporin-4 was reported as recently as 2005 (Lennon et al., 2005), may imply that there is a true B cell autoantigen in MS, which remains to be described. Other common chronic inflammatory diseases omitted from Table 1 include Crohn's disease and ulcerative colitis. Both exhibit many characteristics of autoimmune disease and may prove to have their own specific autoantigens.

3. Evolutionarily conserved proteins and autoimmune disease

3.1. Thyroid peroxidase and myeloperoxidase

Hashimoto's thyroiditis is probably the commonest autoimmune syndrome in humans. The disease is characterised by a T cell-mediated destruction of thyroid follicular cells, resulting in hypothyroidism. Antibodies to the main autoantigen, thyroid peroxidase, are found in up to 20% of adult females, though many subjects are asymptomatic (Prentice et al., 1990).

Thyroid peroxidase, a dimeric glycoprotein located on the surface of thyroid follicular cells, is responsible for the hydrogen peroxide-mediated oxidation of iodide to iodine for incorporation into tyrosine residues of thyroglobulin and subsequent production of the hormones thyroxine and triiodothyronine. It is a member of the myeloperoxidase superfamily of peroxidases which catalyse hydrogen peroxide-mediated oxidations and, among other functions, detoxify reactive oxygen species.

Some of the anti-neutrophil cytoplasmic antibodies (ANCA) associated with various types of systemic vasculitis, particularly microscopic polyangiitis (see below), turned out to target another family member, myeloperoxidase, a lysosomal enzyme abundant in the azurophilic granules of neutrophils, which catalyzes the production of bactericidal compounds during the respiratory burst.

A second peroxidase superfamily (Passardi et al., 2007) comprises proteins present in plants, fungi, many bacteria and some protists, which show very low levels of sequence similarity to myeloperoxidase family members. This similarity is too weak to establish an evolutionary relationship between the two families, but available X-ray structures indicate considerable similarity in the overall topology of the catalytic and heme binding regions and the disposition of helices (Picot et al., 1994), suggesting a conserved structure in spite of their divergent sequences.

The myeloperoxidase family was originally thought to be confined to mammals, but relatives were subsequently discovered in arthropods, *Caenorhabditis elegans* and bacteria, specifically an open reading frame from *Pseudomonas alcaligenes*, located in the Xcp-secretion gene cluster, which showed sequence similarity to the myeloperoxidase family, including conservation of the histidine residues involved in haem binding in myeloperoxidase (Daiyasu and Toh, 2000). We have detected a number of other potential myeloperoxidase family members in prokaryotes, including *Pseudomonas putida* and the marine organism *Planctomyces maris* (present authors, unpublished observations).

Recently, local sequence similarities have been reported between thyroid autoantigens (including the conserved haem binding region of thyroid peroxidase) and proteins from *Yersinia pseudotuberculosis* and *Borrelia burgdorferi* (Benvenga et al., 2006). Interestingly, infections with *B. burgdorferi* and *Yersinia* species have both been associated with autoimmune thyroid disease (Tomer and Davies, 1993; Vaccaro et al., 2002).

3.2. Proteinase 3

Anti-neutrophil cytoplasmic autoantibodies are associated with another form of vasculitis, Wegeners granulomatosis, but in this case the ANCA are primarily directed to proteinase 3 (PR3; also known as myeloblastin), a serine protease of the S1 family expressed in azurophilic granules of neutrophils. Related proteins present in these granules include neutrophil elastase, cathepsin G and the proteolytically inactive azurocidin. PR3 is able to degrade extracellular matrix components including elastin, fibronectin, laminin, vitronectin, and collagen types I, III and IV, thus facilitating migration of neutrophils through basement membranes. It also has microbicidal properties, which may be independent of its proteolytic activity, and may have a role in the release of surface-associated tumour necrosis factor alpha.

The major immunodominant linear epitopes are ATVQLPQ (residues 108–114) and RVGAHDP (residues 132–148) which, according to the crystal structure, are located within a relatively mobile, and thus possibly accessible, region towards the middle of the sequence (Fujinaga et al., 1996).

Orthologs of neutrophil serine proteases have been detected in bony fish (Wernersson et al., 2006), and tBLAST analysis of the collection of currently available microbial genomes retrieves a number of sequences with about 50% similarity and 30% identity, which appear to correspond to S1 proteases, including examples from *Myxococcus xanthus*, *Pseudoalteromonas tunicata*, and *Vibrio* species (present authors, unpublished observations).

There is some evidence for a bacterial role in Wegener's granulomatosis, as carriage of *Staphylococcus aureus* appears to increase the risk of relapse, which may be reduced by prophylactic antibiotic treatment (Kallenberg, 2008). However, there is no ortholog of PR3 in any of the 12 strains of *S. aureus* which have been completely sequenced, nor in *S. epidermidis* or *S. haemolyticus*. One hypothesis is that the effect is mediated via bacterial superantigens, which are potent antigen-independent T cell stimulators, though this would not explain the generation of specific anti-PR3 antibodies. An alternative theory is based on the observation that antibodies to peptides translated from the antisense strand of PR3, as well as anti-PR3 antibodies, are present in sera from some patients with Wegener's granulomatosis (Pendergraft et al., 2004). It was proposed that the antibody response to these complementary peptides could generate anti-idiotypic antibodies capable of cross-reactions to PR3 sense peptides, leading to autoimmunity. The complementary peptides could either be endogenous transcripts of the antisense PR3 DNA or microbial proteins containing regions of local similarity to complementary PR3 peptides. While this intriguing hypothesis remains unproven, it suggests a mechanism that could potentially increase the likelihood of microbial exposure leading to autoimmunity, as the repertoire of potentially mimetic exogenous sequences would be increased.

3.3. Pyruvate dehydrogenase complex

Primary biliary cirrhosis, a chronic liver disease leading to progressive destruction of intrahepatic bile ducts, and eventually to liver cirrhosis, is characterised by the presence of high-titre antimitochondrial antibodies. The major autoantigen, present in more than 90% of patients, has been identified as the E2 component of the pyruvate dehydogenase complex (PDC-E2). The pyruvate dehydrogenase complex is located within the matrix of the inner mitochondrial membrane, and in common with other 2-oxo-acid dehydrogenase complexes, consists of multiple copies of three enzymes. These are a thiamine pyrophosphate-dependent decarboxylase (E1), a coenzyme A-dependent acyltransferase (E2), and a dihydrolipoyl dehydrogenase (E3).

The E2 component is highly conserved between prokaryotes and humans, both in terms of amino acid sequence (typically about 40% identity and 55% similarity) and its overall fold. The immunodominant B and T cell epitopes on PDC-E2 are located around the inner lipoic acid domain (Shimoda et al., 1995; Van de Water et al., 1988). This region is exposed on the surface of the molecule, which may explain its dominant antigenicity (Howard et al., 1998).

Because of the similarity of prokaryotic and human PDC-E2, and because some studies have shown antibody and T cell cross-reactivity between them, an infectious aetiology for PBC has been suggested via molecular mimicry. Candidates have included mycobacterial and enterobacterial infections, but these associations have not been confirmed in all studies. For a review see Haydon and Neuberger (2000).

Recently, *Novosphingobium aromaticivorans* (formerly *Sphingomonas aromaticivorans*) has been suggested as a potential aetiological agent (Kaplan, 2004; Selmi et al., 2003), since the lipoyl domain of its PDC-E2 ortholog was claimed to be more similar to the human enzyme than that from other prokaryotes, and reactivity with sera from patients with primary biliary cirrhosis was much higher than observed with *E. coli*. However, comparison of both full length human PDC-E2 and its conserved lipoylated motif against 988 partial and complete microbial genomes using tBLAST found at least 70 prokaryotes in which the conserved lipoic acid attachment region was as, or more, similar to the human enzyme than to that from *N. aromaticivorans* (present authors; unpublished observations). This, together with the fact that Selmi et al. were unable to show any differences in carriage rates between PBC patients and controls might suggests that *N. aromaticivorans* is one among several possible candidates present in the human microbial flora. However, members of the Sphingomonodaceae, including *N. aromaticivorans*, are unusual among Gram negative bacteria, in that not only do they lack lipopolysaccharide (LPS), but their cell walls are rich in glycosphingolipids, which can bind to CD1d in the lysosomal compartement, leading to specific recognition by, and activation of, CD1d restricted NKT cells. An interesting recent study (Mattner et al., 2008) demonstrated that infection of mice with *N. aromaticivorans* induced an inflammatory liver pathology and anti-PDC-E2 IgG antibodies in a CD1 dependent manner. T cells were able to transfer the inflammation to naïve mice, in the absence of NKT cells, CD1d or bacterial infection.

3.4. Glutamic acid decarboxylase (GAD)

Glutamic acid decarboxylase is a pyridoxal 5′ phosphate dependent enzyme which catalyses the decarboxylation of glutamic acid to γ-aminobutyric acid, an inhibitory neurotransmitter in higher animals. Two GAD enzymes are present in mammals, GAD67 (or GAD1) and GAD65 (or GAD2), which in humans are respectively encoded by genes on chromosome 2 and chromosome 10. GAD67 is thought to be constitutively active and responsible for maintaining basal levels of GABA. This is reflected in the more severe phenotype of *Gad1* null mice, which have low GABA levels and die at birth, whereas *GAD2*$^{-/-}$ animals have near normal GABA and appear healthy when born, though subsequently are prone to anxiety and seizures.

GAD occurs predominantly in the brain, but both enzymes are also present in pancreatic beta cells, where their function is uncertain. GAD65, but not GAD 67, is an important autoantigen in type 1 diabetes and latent autoimmune diabetes of adults (Baekkeskov et al., 1990; Fenalti and Rowley, 2008), and autoantibodies can be detected long before the onset of overt symptoms. The sequences of both enzymes are very similar, except for the first 100 amino terminal residues, which are divergent. Interestingly, most of the epitopes which have been partially characterised do not seem to be associated with the divergent N-terminal region but rather cluster towards the C-terminus, though in areas of local dissimilarity, which, a recently published crystal structure (Fenalti et al., 2007) suggests, may be relatively mobile.

GAD is widely distributed (Ueno, 2000) and bacterial orthologs of about 30% similarity and 50% identity to human GAD65 are readily found, and include proteins from a number of pathogens, for example *Coxiella burnetti*, *Yersinia pestis* and *Vibrio* species. A previous report (Garcia and Lopez, 1995) suggested that the enzyme from *Streptococcus pneumoniae* was more similar to human GAD65 than examples from other prokaryotes, but this was probably an artefact of the limited number of bacterial sequences available at the time, since alignment of currently available sequences reveal many that are as, or more, closely related. The absence of a convincing association of diabetes mellitus (one of the most intensively studied of human autoimmune diseases) with any specific preceding bacterial infection tends to argue against epitope mimicry as a pathogenic mechanism.

Possible associations between type I diabetes and Coxsackie B virus infection have been frequently reported over the last 30 years.

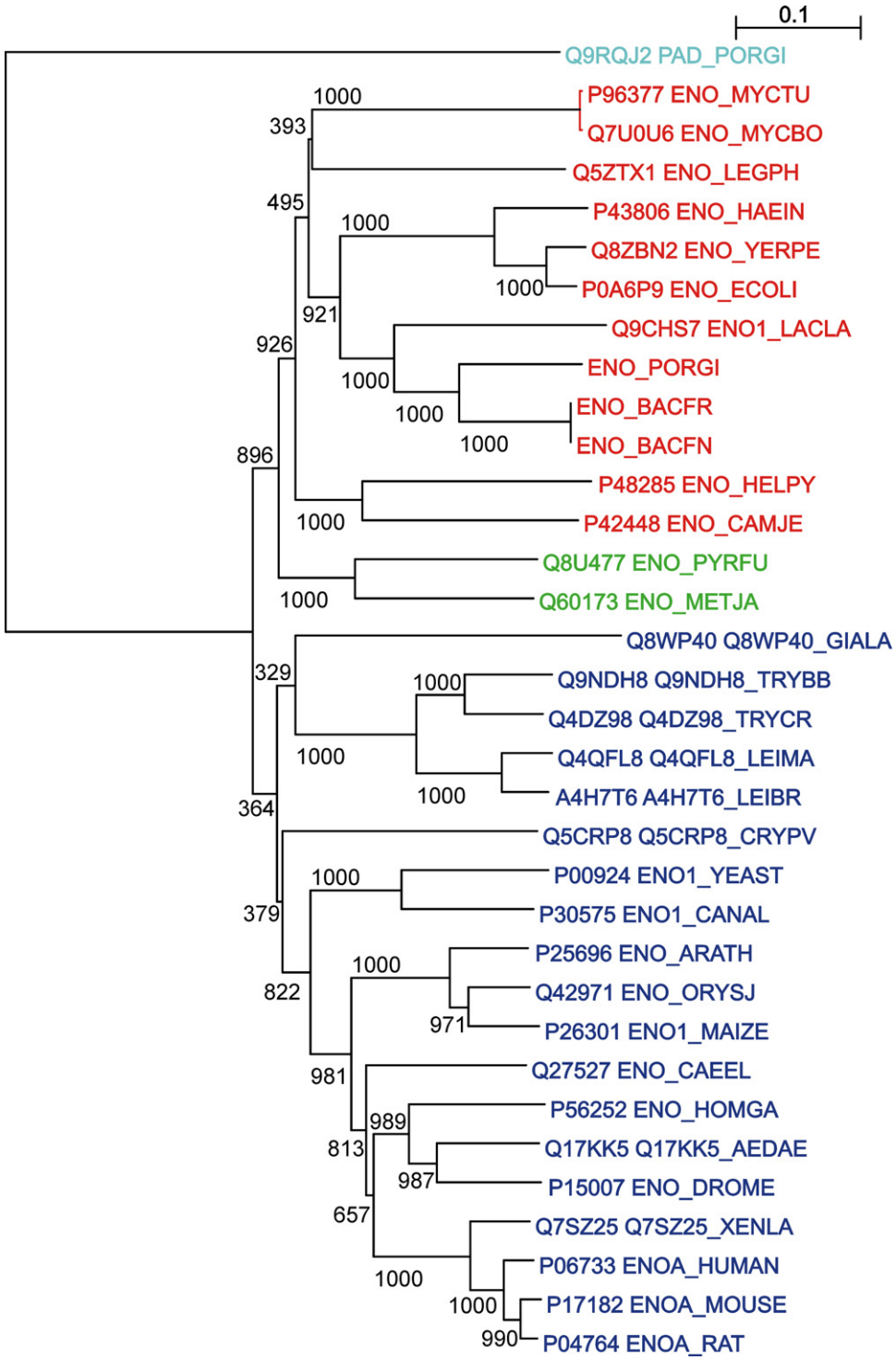


Fig. 1. Rooted phylogenetic tree for enolase from representative eukaryotic, archaeal and bacterial species. The tree was constructed using Clustal X version 2.0 (ftp://ftp.ebi.ac.uk/pub/software/clustalw2/) with the neighbour joining method. Peptidyl arginine deiminase from *Porphymonas gingivalis* was used as outgroup. Protein sequences in FASTA format were retrieved from UniProt. The tips are labelled with the corresponding accession numbers and SwissProt identifiers. (Key: red = bacteria; green = archaea; blue = eukaryotes; cyan = outgroup (*P. gingivalis* peptidyl arginine deiminase).) Branches are labelled with bootstrap values obtained from 1000 replicates. Trees with similar topology were also obtained using parsimony and maximum likelihood methods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

However, a recent systematic review of controlled studies of the relationship between Coxsackie virus serology and type 1 diabetes incidence concluded that existing epidemiological data did not provide convincing evidence for or against such an association (Green et al., 2004).

3.5. Histidyl-tRNA synthetase

Aminoacyl-tRNA synthetases catalyse the attachment of amino acids to their cognate tRNA molecules during protein translation. Being part of the basal machinery of protein synthesis, individual

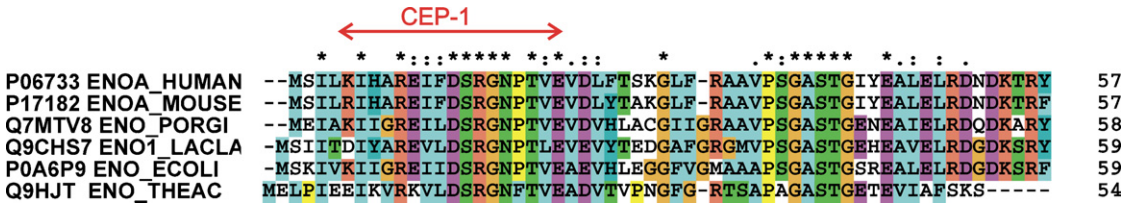


Fig. 2. Sequence alignment of the first 60 residues of mammalian and bacterial and archael enolases showing the high degree of conservation of the immunodominant epitope CEP-1 (KIHAREIFDSRGNPTVE; indicated by the red line). The alignment was created using Clustal X version 2.0. The sequences were obtained from UniProt. P06733 (*H. sapiens*); P17182 (*Mus musculus*); Q7MTV8 (*P. gingivalis*); Q9CHS7 (*Lactococcus lactis*); P0A6P9 (*E. coli*); Q9HJT (*Thermoplasma acidophilum*). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

aminoacyl-tRNA synthetases are highly conserved and are present in archaea, bacteria, and eucarya. Recently, four aminoacyl-tRNA synthetases were detected in *Acanthamoeba polyphaga* mimi virus (Abergel et al., 2005), though phylogenetic analysis suggests that they were possibly the result of complex lateral gene transfer from ancient hosts, rather than being survivors of a now lost viral protein translation apparatus (Abergel et al., 2007).

Aminoacyl-synthetases are structurally diverse, but can be classified into classes I and II, on the basis of conserved signature motifs and the differing folds of their active site regions (Arnez and Moras, 1997).

Autoantibodies to aminoacyl-tRNA synthetases occur in approximately 25% of patients with myositis, for a review see Mimori et al. (2007). Cytoplasmic histidyl-tRNA synthetase (HisRS) is the major target, but antibodies to glycyl-, alanyl-, isoleucyl-, threonyl-, and asparaginyl-tRNA synthase have also been reported.

Most of anti-HisRS antibodies in myositis patient sera bind to sites within the conserved N-terminal sequence of the enzyme (residues 1–47) near the catalytic domain, the predominant antigenic region being surface-exposed, and predicted to form a coiled-coil (Raben et al., 1994; Ramsden et al., 1989). This region, however, is absent from bacterial HisRS, which otherwise share about 25% identity and 40% similarity to human HisRS, which argues against a mechanism involving infection and epitope mimicry. Consistent with this, mysositis autoantibodies are reported to recognise only histidyl-tRNA synthetase from higher eukaryotes. Interestingly, the amino terminal domain contains a highly conserved 32 amino acid motif which is shared by other class I and class II tRNA aminoacyl synthetases from both eukaryotes and prokaryotes, including bacterial seryl-tRNA synthetases (Raben et al., 1994). Whether this has implications for the aetiology of the disease is unclear.

Encephalomyocarditis virus (EMCV) infection has been associated with myositis, as it is able to produce disease symptoms similar to myositis in mice. It has been reported that the VP1 component of the viral polyprotein has some regions of local similarity with HisRS, though the resemblance is presumably fortuitous and not evidence of a common ancestor (Walker and Jeffrey, 1988).

3.6. Enolase

Enolase catalyses the interconversion of phosphoenolpyruvate and 2-phosphoglycerate, the penultimate step in glycolysis. Glycolysis is an ancient metabolic pathway and its key enzymes are highly conserved (Pancholi, 2001).

Enolase from higher eukaryotes is a dimeric protein composed of three isoenzymes (α, β, γ), of which the most abundant is α-enolase. Unlike the other autoantigens discussed so far, the presence of antibodies to α-enolase is not disease-specific. Anti-α-enolase antibodies have been implicated in autoimmune diseases including post-streptococcal syndromes, Behçet's disease, systemic lupus erythematosus (SLE), ulcerative colitis, Crohn's disease, autoimmune hepatitis and lymphocytic hypophysitis, reviewed in Pratesi et al. (2000). Despite the high sequence similarity between enolases from different species, bacterial aetiology has only been studied for post-streptococcal autoimmune disease (Dale et al., 2006; Fontan et al., 2000). The fact that cell-surface exposed enolase is found not only in mammalian cells but also in prokaryotic and eukaryotic pathogens (e.g. *Streptococci*, *Pneumococci*, *S. aureus*, *Candida albicans*, *Pneumocystis carini*, and *Leishmania mexicana*) makes it a prime candidate for the generation of cross-reactive autoantibodies.

3.7. Citrullinated α-enolase

The family of peptidyl arginine deiminase (PAD) enzymes deiminate positively charged arginine residues in proteins to form neutral citrulline residues. It was first appreciated that citrullinated proteins were the dominant autoantigens in rheumatoid arthritis (RA) over a decade ago, when it was shown that serum samples from a panel of patients with RA preferentially bound to synthetic peptides containing citrulline, compared with serum from healthy controls (Schellekens et al., 1998). This observation led to the development of a new diagnostic test for RA, using randomly generated synthetic peptides of unpublished sequence, optimised for highest specificity (98%) and sensitivity (80%) for the diagnosis of RA. This test, known as anti-CCP2, predicts severe disease and may occur years before the onset of clinical symptoms. There is increasing evidence that these antibodies are directly involved in disease pathogenesis (Clavel et al., 2008; Kinloch et al., 2006) (Fig. 1).

Citrullinated α-enolase was first described by our group in 2005 as a candidate member of the family of citrullinated proteins targeted by autoantibodies in RA (Kinloch et al., 2005). Since then it has gained acceptance as a novel disease-specific autoantigen with distinctive immunogenetic association with HLA DRB1*0401 (Snir et al., 2008). We have identified a sequence (KIHAREIFDSRGNPTVE) near the amino terminus as the immunodominant epitope, and have shown that the citrullination of the second arginine makes the greatest contribution to its immunogenicity. This sequence, designated citrullinated enolase peptide-1 (CEP-1), is highly conserved between eukaryotes and prokaryotes as illustrated by the multiple sequence alignment in Fig. 2.

Western blot analysis using affinity-purified antibodies from RA sera showed cross-reactivity between citrullinated human α-enolase and citrullinated recombinant enolase from the periodontal pathogen *Porphyromonas gingivalis* (Lundberg et al., 2008). There is evidence that *P. gingivalis* expresses a PAD-like enzyme which can deiminate C-terminal arginine residues (McGraw et al., 1999), but, alternatively, human PAD enzymes, which are secreted in response to inflammation, could be responsible for citrullination of both human and bacterial enolase.

4. Conclusion

Why particular proteins elicit autoantibodies, while others do not, is unknown, but the example of GAD suggests that the reasons

are subtle, since GAD67 and GAD65 are closely related and present in the target tissue, but autoantibodies are mainly produced against GAD65. Certain intrinsic structural features seem to be necessary to trigger autoantibody formation, including surface charge density, and the presence of repetitive surface elements. Epitopes targeted by autoantibodies tend to be surface accessible, often in coiled-coil domains, and relatively mobile.

The evolutionary distances between humans and prokaryotes are enormous, so that even proteins subject to extreme functional constraints, such as those responsible for fundamental processes of protein synthesis and energy metabolism, may exhibit substantial sequence divergence. Thus, in most cases, the likelihood of linear epitopes in microbial proteins retaining sufficient similarity to trigger an autoimmune response is probably quite small. Moreover, the presence of microbial epitopes mimicking host proteins is probably a necessary rather than a sufficient condition for production of autoantibodies, which may require, in addition, an underlying genetic or other susceptibility in the host, and possibly some form of co-stimulation from the environment or the infecting microbe. The presence of glycosphingolipids able to bind CD1d in *N. aromaticivorans*, or the modification of arginine to citrulline, could be examples of the latter. The requirement for a conjunction of all these factors is likely to dramatically reduce the number of proteins able to induce an autoimmune response.

Of the seven major autoantigens examined in this review, four (glutamate decarboxylase, pyruvate dehydrogenase complex E2, histidyl-tRNA synthetase and alpha-enolase) have well conserved prokaryotic orthologs, most with significant regions of local identity. In the case of α-enolase, the immunodominant epitope is the most highly conserved and therefore potentially a true molecular mimic in genetic, functional and immunological terms. When Damian coined the term "molecular mimicry" (Damian, 1965), he focussed on the conservation of sequences as a consequence of the selective pressure exerted on the pathogen by immune surveillance. We conclude that molecular mimicry leading to autoimmunity may be another unwelcome effect of evolutionary conservation.

References

Abergel C, Chenivesse S, Byrne D, Suhre K, Arondel V, Claverie JM. Mimivirus TyrRS: preliminary structural and functional characterization of the first amino-acyl tRNA synthetase found in a virus. Acta Crystallogr Sect F Struct Biol Cryst Commun 2005;61:212–5.

Abergel C, Rudinger-Thirion J, Giege R, Claverie JM. Virus-encoded aminoacyl-tRNA synthetases: structural and functional characterization of mimivirus TyrRS and MetRS. J Virol 2007;81:12406–17.

Arnez JG, Moras D. Structural and functional considerations of the aminoacylation reaction. Trends Biochem Sci 1997;22:211–6.

Baekkeskov S, Aanstoot HJ, Christgau S, Reetz A, Solimena M, Cascalho M, et al. Identification of the 64 K autoantigen in insulin-dependent diabetes as the GABA-synthesizing enzyme glutamic acid decarboxylase. Nature 1990;347:151–6.

Benvenga S, Santarpia L, Trimarchi F, Guarneri F. Human thyroid autoantigens and proteins of Yersinia and Borrelia share amino acid sequence homology that includes binding motifs to HLA-DR molecules and T-cell receptor. Thyroid 2006;16:225–36.

Clavel C, Nogueira L, Laurent L, Iobagiu C, Vincent C, Sebbag M, et al. Induction of macrophage secretion of tumor necrosis factor alpha through Fcgamma receptor IIa engagement by rheumatoid arthritis-specific autoantibodies to citrullinated proteins complexed with fibrinogen. Arthritis Rheum 2008;58:678–88.

Daiyasu H, Toh H. Molecular evolution of the myeloperoxidase family. J Mol Evol 2000;51:433–45.

Dale RC, Candler PM, Church AJ, Wait R, Pocock JM, Giovannoni G. Neuronal surface glycolytic enzymes are autoantigen targets in post-streptococcal autoimmune CNS disease. J Neuroimmunol 2006;172:187–97.

Damian RT. Molecular mimicry in biological adaptation. Science 1965;147:824.

Fenalti G, Law RH, Buckle AM, Langendorf C, Tuck K, Rosado CJ, et al. GABA production by glutamic acid decarboxylase is regulated by a dynamic catalytic loop. Nat Struct Mol Biol 2007;14:280–6.

Fenalti G, Rowley MJ. GAD65 as a prototypic autoantigen. J Autoimmun, 2008. [Epub ahead of print].

Fontan PA, Pancholi V, Nociari MM, Fischetti VA. Antibodies to streptococcal surface enolase react with human alpha-enolase: implications in poststreptococcal sequelae. J Infect Dis 2000;182:1712–21.

Fujinaga M, Chernaia MM, Halenbeck R, Koths K, James MN. The crystal structure of PR3, a neutrophil serine proteinase antigen of Wegener's granulomatosis antibodies. J Mol Biol 1996;261:267–78.

Garcia E, Lopez R. *Streptococcus pneumoniae* type 3 encodes a protein highly similar to the human glutamate decarboxylase (GAD65). FEMS Microbiol Lett 1995;133:113–8.

Green J, Casabonne D, Newton R. Coxsackie B virus serology and Type 1 diabetes mellitus: a systematic review of published case-control studies. Diabet Med 2004;21:507–14.

Haydon GH, Neuberger J. PBC: an infectious disease? Gut 2000;47:586–8.

Howard MJ, Fuller C, Broadhurst RW, Perham RN, Tang JG, Quinn J, et al. Three-dimensional structure of the major autoantigen in primary biliary cirrhosis. Gastroenterology 1998;115:139–46.

Kallenberg CG. Pathogenesis of PR3-ANCA associated vasculitis. J Autoimmun 2008;30:29–36.

Kaplan MM. *Novosphingobium aromaticivorans*: a potential initiator of primary biliary cirrhosis. Am J Gastroenterol 2004;99:2147–9.

Kinloch A, Lundberg K, Moyes D, Venables PJ. The pathogenic role of antibodies to citrullinated proteins in rheumatoid arthritis. Expert Rev Clin Immunol 2006;2:365–75.

Kinloch A, Tatzer V, Wait R, Peston D, Lundberg K, Donatien P, et al. Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis. Arthritis Res Ther 2005;7:R1421–9.

Lennon VA, Kryzer TJ, Pittock SJ, Verkman AS, Hinson SR. IgG marker of optic-spinal multiple sclerosis binds to the aquaporin-4 water channel. J Exp Med 2005;202:473–7.

Lundberg K, Kinloch A, Fisher B, Wegner N, Wait R, Charles P, et al. Antibodies to citrullinated alpha-enolase peptide-1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. Arthritis Rheum 2008;58(10):3009–19.

Mattner J, Savage PB, Leung P, Oertelt SS, Wang V, Trivedi O, et al. Liver autoimmunity triggered by microbial activation of natural killer T cells. Cell Host Microbe 2008;3:304–15.

McGraw WT, Potempa J, Farley D, Travis J. Purification, characterization, and sequence analysis of a potential virulence factor from *Porphyromonas gingivalis*, peptidylarginine deiminase. Infect Immun 1999;67:3248–56.

Mimori T, Imura Y, Nakashima R, Yoshifuji H. Autoantibodies in idiopathic inflammatory myopathy: an update on clinical and pathophysiological significance. Curr Opin Rheumatol 2007;19:523–9.

Pancholi V. Multifunctional alpha-enolase: its role in diseases. Cell Mol Life Sci 2001;58:902–20.

Passardi F, Bakalovic N, Teixeira FK, Margis-Pinheiro M, Penel C, Dunand C. Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes. Genomics 2007;89:567–79.

Pendergraft 3rd WF, Preston GA, Shah RR, Tropsha A, Carter Jr CW, Jennette JC, et al. Autoimmunity is triggered by cPR-3(105-201), a protein complementary to human autoantigen proteinase-3. Nat Med 2004;10:72–9.

Picot D, Loll PJ, Garavito RM. The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1. Nature 1994;367:243–9.

Pratesi F, Moscato S, Sabbatini A, Chimenti D, Bombardieri S, Migliorini P. Autoantibodies specific for alpha-enolase in systemic autoimmune disorders. J Rheumatol 2000;27:109–15.

Prentice LM, Phillips DI, Sarsero D, Beever K, McLachlan SM, Smith BR. Geographical distribution of subclinical autoimmune thyroid disease in Britain: a study using highly sensitive direct assays for autoantibodies to thyroglobulin and thyroid peroxidase. Acta Endocrinol (Copenh) 1990;123:493–8.

Raben N, Nichols R, Dohlman J, McPhie P, Sridhar V, Hyde C, et al. A motif in human histidyl-tRNA synthetase which is shared among several aminoacyl-tRNA synthetases is a coiled-coil that is essential for enzymatic activity and contains the major autoantigenic epitope. J Biol Chem 1994;269:24277–83.

Ramsden DA, Chen J, Miller FW, Misener V, Bernstein RM, Siminovitch KA, et al. Epitope mapping of the cloned human autoantigen, histidyl-tRNA synthetase. Analysis of the myositis-associated anti-Jo-1 autoimmune response. J Immunol 1989;143:2267–72.

Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101:273–81.

Selmi C, Balkwill DL, Invernizzi P, Ansari AA, Coppel RL, Podda M, et al. Patients with primary biliary cirrhosis react against a ubiquitous xenobiotic-metabolizing bacterium. Hepatology 2003;38:1250–7.

Shimoda S, Nakamura M, Ishibashi H, Hayashida K, Niho Y. HLA DRB4 0101-restricted immunodominant T cell autoepitope of pyruvate dehydrogenase complex in primary biliary cirrhosis: evidence of molecular mimicry in human autoimmune diseases. J Exp Med 1995;181:1835–45.

Snir O, Widhe M, von Spee C, Lindberg J, Padyukov L, Lundberg K, et al. Multiple antibody reactivities to citrullinated antigens in sera from rheumatoid arthritis patients—association with HLA-DRB1 alleles. Ann Rheum Dis, 2008. [Epub ahead of print].

Tomer Y, Davies TF. Infection, thyroid disease, and autoimmunity. Endocr Rev 1993;14:107–20.

Ueno H. Enzymatic and structural aspects on glutamate dehydrogenase. J Mol Catal B: Enzym 2000;10:67–79.

Vaccaro M, Guarneri F, Borgia F, Cannavo SP, Benvenga S. Association of lichen sclerosus and autoimmune thyroiditis: possible role of *Borrelia burgdorferi*? Thyroid 2002;12:1147–8.

Van de Water J, Gershwin ME, Leung P, Ansari A, Coppel RL. The autoepitope of the 74-kDa mitochondrial autoantigen of primary biliary cirrhosis corresponds to the functional site of dihydrolipoamide acetyltransferase. J Exp Med 1988;167:1791–9.

Walker EJ, Jeffrey PD. Sequence homology between encephalomyocarditis virus protein VPI and histidyl-tRNA synthetase supports a hypothesis of molecular mimicry in polymyositis. Med Hypotheses 1988;25:21–5.

Wernersson S, Reimer JM, Poorafshar M, Karlson U, Wermenstam N, Bengten E, et al. Granzyme-like sequences in bony fish shed light on the emergence of hematopoietic serine proteases during vertebrate evolution. Dev Comp Immunol 2006;30:901–18.

Natalia Wegner
Karin Lundberg
Andrew Kinloch
Benjamin Fisher
Vivianne Malmström
Marc Feldmann
Patrick J. Venables

Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis

Authors' addresses
Natalia Wegner[1], *Karin Lundberg*[1], *Andrew Kinloch*[1], *Benjamin Fisher*[1], *Vivianne Malmström*[2], *Marc Feldmann*[1], *Patrick J. Venables*[1]
[1]The Kennedy Institute of Rheumatology, Imperial College London, London, UK.
[2]Department of Medicine, Karolinska Institutet, Stockholm, Sweden.

Correspondence to:
Patrick J. Venables
Imperial College London
Kennedy Institute Division
65 Aspenlea Road
W6 8LH London, UK
Tel.: +44 2083833444
Fax: +44 2083834499
e-mail: p.venables@imperial.ac.uk

Immunological Reviews 2010
Vol. 233: 34–54

Immunological Reviews
0105-2896

Summary: Rheumatoid arthritis (RA) is now clearly a true autoimmune disease with accumulating evidence of pathogenic disease-specific autoimmunity to citrullinated proteins. Citrullination, also termed deimination, is a modification of arginine side chains catalyzed by peptidylarginine deiminase (PAD) enzymes. This post-translational modification has the potential to alter the structure, antigenicity, and function of proteins. In RA, antibodies to cyclic citrullinated peptides are now well established for clinical diagnosis, though we argue that the identification of specific citrullinated antigens, as whole proteins, is necessary for exploring pathogenic mechanisms. Four citrullinated antigens, fibrinogen, vimentin, collagen type II, and α-enolase, are now well established, with others awaiting further characterization. All four proteins are expressed in the joint, and there is evidence that antibodies to citrullinated fibrinogen and collagen type II mediate inflammation by the formation of immune complexes, both in humans and animal models. Antibodies to citrullinated proteins are associated with HLA 'shared epitope' alleles, and autoimmunity to at least one antigenic sequence, the CEP-1 peptide from citrullinated α-enolase (KIHAcitEIFDScitGNPTVE), shows a specific association with HLA-DRB1*0401, *0404, 620W PTPN22, and smoking. Periodontitis, in which *Porphyromonas gingivalis* is a major pathogenic bacterium, has been linked to RA in epidemiological studies and also shares similar gene/environment associations. This is also the only bacterium identified that expresses endogenous citrullinated proteins and its own bacterial PAD enzyme, though the precise molecular mechanisms of bacterial citrullination have yet to be explored. Thus, both smoking and *Porphyromonas gingivalis* are attractive etiological agents for further investigation into the gene/environment/autoimmunity triad of RA.

Keywords: *rheumatoid, enolase, autoimmune, etiology, smoking, periodontitis*

Introduction

Rheumatoid arthritis (RA) is a systemic autoimmune disease of unknown etiology, affecting 0.5–1.0% of the adult population worldwide. It is a complex, multifactorial disease in which chronic inflammation of synovial joints and erosion of bone result in joint destruction, pain, disability, and a reduced life expectancy. Genetic influences are estimated to be responsible for around 50–60% of the risk of developing RA (1),

with environmental factors explaining the remainder. The recent characterization of citrullinated proteins as the main targets for the autoimmune attack in RA has enabled the investigation of the paradigm that the pathology of the disease is mediated by specific autoantibodies, which themselves are induced by the interaction of genetic risk factors and the environment (*Fig. 1*). This triad of factors – autoantibodies, genes, and environment – is of course a gross simplification of the diverse and complex pathways that underlie the clinical manifestations seen in RA. Nevertheless, it serves as a useful template for addressing a fundamental issue, namely, what actually causes the disease.

Autoimmunity in RA

In contrast to many other autoimmune diseases, where the major autoantigens and their involvement in disease pathology have been characterized at a molecular level several decades ago, the autoantigens in RA are only just beginning to be evaluated. Historically, rheumatoid factor (RF), an antibody reactive with the Fc portion of IgG, has been the main serological marker for the diagnosis of RA and is still used as one of the criteria for the classification of the disease (2). Therefore, it has been suggested that IgG is the autoantigen in RA. However, RFs also occur in other autoimmune diseases, infections, and in 5% of the healthy population, and thus may simply be a consequence of polyclonal B-cell activation. Given the lack of disease specificity of RFs, it is difficult to substantiate the concept that IgG is the specific antigen driving the destructive inflammatory response that characterizes RA.

Interest has focused recently on citrullinated proteins as the true autoantigens in RA. A prototype assay was described over 40 years ago as the anti-perinuclear factor test; though at the time, it was not known that it was detecting anti-citrullinated protein antibodies. Human buccal mucosa cells were used as the substrate, and antibodies reactive with granular structures around the nucleus were detected by indirect immunofluorescence (3). The assay had a high diagnostic specificity for RA but with variable sensitivity (reviewed in 4). This drawback, combined with the cumbersome nature of the assay, precluded its use for routine diagnosis. An antibody reactive with the keratinized layer of rat esophagus was subsequently reported as the so-called 'anti-keratin antibody' (5) and correlated strongly with the anti-perinuclear factor. More recently, it was shown that both anti-perinuclear and anti-keratin antibodies reacted with mature filaggrin (6–8), and a crucial discovery was that this reactivity was citrulline-specific (9, 10). Nowadays, enzymatic assays using cyclic citrullinated peptides (CCP) as surrogate target antigens are routinely used in RA diagnosis. The first generation anti-CCP assay (anti-CCP1) was based on citrulline-containing peptides from the filaggrin sequence (10, 11). This test had a diagnostic sensitivity of approximately 70% with a disease-specificity of 96% (11). Because filaggrin, a protein which is involved in epidermal differentiation and hydration, is found only in epithelial cells but not in the joint (12), it was concluded that citrullinated filaggrin acted as a surrogate antigen and that other citrullinated proteins were more likely to be driving the autoimmune response in RA (reviewed in 13). To improve the diagnostic sensitivity and specificity of the anti-CCP1 test, the anti-CCP2 assay was developed and is now widely used in diagnostic laboratories. The substrates for this test were selected from a large panel of randomly generated citrulline-containing peptides that were tested against RA and control serum, with the sequences giving the best discrimination in diagnostic sensitivity and specificity being adopted for a commercial assay. Using an enzyme-linked immunosorbent assay (ELISA), anti-CCP2 antibodies have a diagnostic sensitivity that is at least equivalent to that of RF (up to 82%) and a specificity that is much higher (98%) (reviewed in 14). Anti-CCP antibodies have been detected prior to the development of clinically apparent RA and are associated with more severe and erosive disease (15–19). The high diagnostic specificity and predictive value of these antibodies has stimulated interest in their involvement in the pathophysiology of RA.

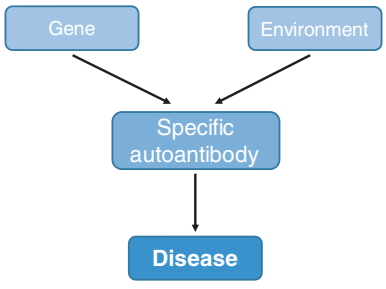


***Fig. 1.* Schematic representation of the hypothesis for the gene/environment/autoimmunity triad in the etiology of rheumatoid arthritis and other chronic autoimmune diseases.**

Physiological and pathological consequences of protein citrullination

Citrulline, in the context of a peptide backbone, is a non-standard amino acid that results from post-translational modification of arginine residues. This conversion, termed citrullination or deimination, reduces the net charge of the protein by the loss of a positive charge per citrulline residue

(*Fig.* 2). Citrullination is catalyzed by a family of calcium-dependent peptidylarginine deiminase (PAD) enzymes, which are activated by higher Ca^{2+} concentrations [$>10^{-6}$ M, varies with the source of PAD enzyme and *in vitro* substrate (20, 21)] than that present in intact cells [estimates range from 10^{-9} to 10^{-7} (22)]. Five members of the PAD family (PAD1, PAD2, PAD3, PAD4, and PAD6) with different tissue distribution, sub-cellular localization and substrates have been found in humans (23). Of particular relevance to RA are PAD2 and PAD4, as their expression has been demonstrated in rheumatoid synovial membrane (21, 24–26), synovial fluid cells (27), and extracellularly in synovial fluid (28).

Arginine residues within polypeptides often play a central role in the structural integrity of a protein, due to their ability to participate in ionic interactions with negatively charged amino acid side chains, substrates, and cofactors, and to form multiple hydrogen bonds to both the peptide backbone and other amino acid side chains (29). Arginine has also got the most polar of all the common amino acid side chains and is therefore the amino acid that is most likely to be found on the surface of proteins in an aqueous environment (29). Citrullination would be expected to destroy the ionic interactions, interfere with hydrogen bonds, and create new interactions. Hence, the conversion of arginine into citrulline may result in an altered three-dimensional structure and function of a protein.

In healthy physiology, citrullinated proteins are present in a variety of cells and tissues. Citrullination of keratin and filaggrin in the skin enables proteolytic cleavage and cross-linking during the final stages of keratinocyte differentiation, which plays an important role in the cornification of the epidermis (30, 31). Trichohyalin, the major structural protein of the hair follicle, and S100A3, a protein involved in the maturation of hair cuticle cells, are also PAD substrates and citrullination of these proteins is important in the formation of rigid structures (32). In the nervous system, physiological citrullination of myelin basic protein (MBP) is important in ensuring electrical insulation of the myelin sheaths. Pathological citrullination of MBP results in a more open conformation (reviewed in 33), increasing its susceptibility to cleavage by cathepsin D (34). Thus, citrullination of MBP is important in the investigation of demyelinating diseases such as multiple sclerosis.

At a cellular level, citrullination affects cytoskeletal stability. Inagaki *et al.* (20) have demonstrated that *in vitro* citrullination of murine vimentin filaments results in the disassembly of the filaments and explained this finding as a mechanism for non-reversible filament modulation taking place under pathological conditions, where the cell membrane disintegrates and allows for elevated intracellular Ca^{2+} concentrations. Similar results were also obtained for the other intermediate filament proteins desmin and glial fibrillary acidic protein (GFAP) (20, 35). Several studies suggest that citrullination of histones may be involved in transcriptional regulation (36–44) and in the decondensation of chromatin when forming neutrophil extracellular traps, as an innate response to infectious and inflammatory stimuli (45, 46).

Citrullination, as part of the inflammatory process, is only just beginning to be explored. Stimulation of peripheral blood mononuclear cells with IFN-γ and a synthetic dsRNA causes citrullination of the chemokines CXCL8 (47) and CXCL10 (48), with fundamental effects on their receptor usage, proteolytic processing, and biological activities. Furthermore, thrombin-catalyzed fibrinogen polymerization is defective when fibrinogen is citrullinated (49), due to citrullination at the thrombin cleavage site (arginine-16 in the α chain and arginine-14 in β chain) (50). The physiological role of this is unclear, but it might be an anti-inflammatory response to tissue inflammation and a negative regulator of fibrin deposition. Excessive fibrin deposition, with a potential pro-inflammatory role, is a common feature within RA joints (51), suggesting inefficient fibrinolysis. Hence, the fibrinolytic process, where plasmin degrades fibrin clots, may also be

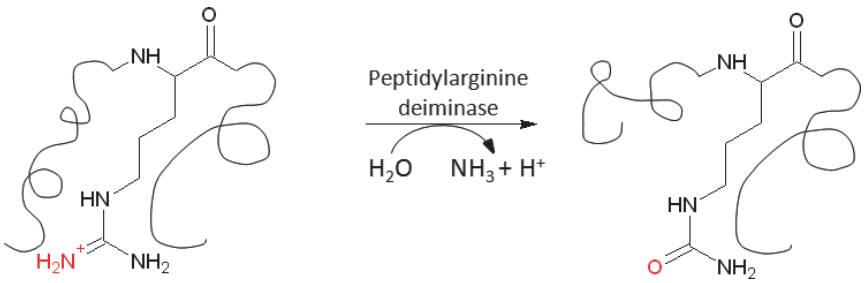


***Fig.* 2. Enzymatic conversion of peptidylarginine into peptidylcitrulline, catalyzed by the family of peptidylarginine deiminase enzymes.** The positively charged arginine guanidino group is converted into the neutral citrulline ureido group in the presence of water, yielding ammonia and a hydrogen ion as by-products. The possible consequences on the protein structure are shown schematically.

defective as a result of fibrin citrullination, as arginine residues are part of the cleavage sites.

In RA, the effects of citrullination are of particular importance with regard to the potential change in antigenicity of the proteins. The unfolding of certain proteins upon citrullination was reported to be similar to that observed in the presence of high concentrations of urea (>4.5 M) (30) and could clearly expose new epitopes to the immune system, either directly as a result of the changed three-dimensional structure or indirectly due to altered proteolytic processing and MHC class II presentation. Citrullination of rat serum albumin (RSA), for example, has been shown to break immune tolerance in rats in the presence of adjuvant. Animals immunized with citrullinated RSA developed an antibody response to both citrullinated and native RSA, while control animals immunized with native RSA did not (52). As we discuss below, the mere substitution of a single arginine by citrulline in a small peptide can fundamentally alter the binding of the peptide to both antibodies and to MHC class II molecules. Thus, the presence of citrullinated proteins as autoantigens and their distribution in inflamed tissues are of critical importance in our understanding of the pathogenesis of RA.

Citrullinated proteins in disease

Pathological protein citrullination is associated with inflammation, not just in RA but also in other forms of inflammatory arthritis (53–55) and unrelated inflammatory diseases, such as multiple sclerosis (56), glaucoma (57), myositis (54), and Alzheimer's disease (58). In psoriasis, citrullination is downregulated in keratinocytes, possibly reflecting their failure to undergo terminal differentiation in the formation of the cornified layer of the skin (59). In inflammatory arthritis, including RA, it is assumed that inflammation and the resulting release of pro-inflammatory stimuli and increased cell death, allowing PAD activation in a calcium-rich environment, might account for the accumulation of citrullinated proteins. We have shown that the synovial fluid is an extracellular compartment within the joint with particularly abundant expression of citrullinated proteins (28). In a study of 60 cell-free synovial fluids (20 rheumatoid, 20 spondylarthropathies, and 20 osteoarthropathies), we found citrullinated proteins in 34 out of the 40 inflammatory fluids (i.e. RA and spondyloarthropathies) but none or markedly reduced levels of citrullinated proteins in the osteoarthritis controls. The deiminating enzyme PAD4 was detected in all three disease groups, while detection of PAD2 was mainly restricted to the inflammatory synovial fluids. This provides some evidence that the extracellular (and therefore Ca^{2+}-rich) environment is able to support the activity of PADs and that PAD2 may be inducible by inflammation. However, intracellular citrullinated proteins in inflammatory arthritides were also observed by a number of investigators (12, 24, 60–63) and might result from a controlled increase of intracellular Ca^{2+} concentrations, for example in response to specific ion pump activation.

Autoantibodies to specific citrullinated proteins in RA

As accumulation of citrullinated proteins occurs in a variety of inflammatory conditions, it is now accepted that it is the antibody response rather than the presence of citrullinated proteins in general that characterizes RA. The question is, therefore, which specific antigens drive the immune response in RA? The cyclic citrullinated peptides used in the CCP2 test do not correspond to *in vivo* generated citrullinated proteins and are therefore of limited use for understanding the disease etiology and pathogenesis. However, four citrullinated candidate antigens, present in the joint, are now well established: citrullinated fibrinogen/fibrin (63), vimentin (64), collagen type II (65), and α-enolase (66).

Citrullinated fibrinogen

Fibrinogen is the precursor of fibrin, and citrullinated fibrin(ogen) is one of the best-established specific autoantigens in RA. The antigen, mainly as citrullinated fibrin, is abundantly expressed in inflammatory joints of humans (53, 55, 63, 67, 68) and experimental animals (69). It was originally identified by its cross-reaction with anti-citrullinated filaggrin antibodies, which were affinity-purified from RA serum. By immunoblotting, these antibodies reacted mainly with the α-and β-chains of fibrin (63). Later studies utilized the more convenient, soluble (citrullinated) fibrinogen for further characterization of the reactivity.

One large study of established RA suggested that antibodies to whole citrullinated fibrinogen are equal, in terms of diagnostic sensitivity and specificity, to the anti-CCP2 assay (70), though a more recent study of early RA found that the sensitivity of antibodies to citrullinated fibrinogen was slightly lower (66%) compared with the CCP2 test (72%) (71). In both studies, there was a close correlation but by no means exclusive relationship between anti-citrullinated fibrinogen and anti-CCP2 antibodies.

The most extensive study aimed at identifying the major immunodominant B-cell epitopes in human fibrin(ogen) was published by Sebbag *et al.* (72). Using 71 linear peptides derived from the α- and β-chains of fibrinogen (residues

36–644 and 45–491, respectively), five peptides containing immunodominant epitopes, out of a total of 18 reactive peptides, were identified by ELISA: four from the α-chain (^{36}GPcitVVEcitHQSACKDS^{50}; ^{171}VDIDIKIcitSCcitGSCS^{185}; ^{501}SGIGTLDGFcitHcitHPD^{515}; ^{621}citGHAKScitPVcitGIHTS^{635}) and one from the β-chain (^{60}citPAPPPISGGGYcitAcit^{74}). Although each of the 20 RA serum samples showed reactivity to a limited number of peptides (between two and seven), an overall inter-individual variability in reactivity profiles was observed. As the authors stated, conformation-dependent epitopes which are only formed in the context of the whole molecule might not have been detected using the applied technique, and it is not clear whether the identified epitopes exist *in vivo*. Assuming citrullination in the joint occurs after fibrinogen processing into fibrin, some peptides might be absent from mature fibrin. In another study, Nakayama-Hamada *et al.* (21) investigated *in vitro* citrullination of fibrinogen after incubation with recombinant human PAD2 or PAD4. They observed citrullination in protein regions corresponding to three of the peptide epitopes identified by Sebbag *et al.* (α-chain peptides 36–50 and 501–515 and β-chain peptide 60–74) but not to the other two. However, Nakayama-Hamada *et al.* used whole fibrinogen, which may well expose different protein regions to PAD than those corresponding to the peptides used by Sebbag *et al.*

The expression of antigen in the joint and the detection of disease-specific antibodies in the serum make citrullinated fibrinogen an attractive antigen for mediating the immunopathology of RA. This possibility was recently explored in an *in vitro* model system (73). Clavel *et al.* (73) found that human monocyte-derived macrophages produced TNF-α when incubated with immobilized immune complexes containing citrullinated fibrinogen and that the effect was mediated by Fcγ receptor IIa. Fibrinogen-containing immune complexes were demonstrated in serum samples from about 50% of anti-CCP-positive RA patients, and there was indirect evidence of complement component 3 (C3)-binding immune complexes colocalizing with fibrin(ogen) in the pannus tissue (74). These findings go a long way towards explaining the mechanisms of how immune complexes, containing citrullinated fibrinogen, might be important in the pathogenesis of RA by driving inflammation via Fcγ and complement receptors, resulting in inflammatory cytokine release (75).

An arthritogenic role for citrullinated fibrinogen was recently reported by Hill *et al.* (60), using C57B6 mice transgenic for human HLA-DRB1*0401. Transgenic and wildtype mice were immunized with unmodified or citrullinated human fibrinogen. Approximately 35% of transgenic mice immunized with citrullinated fibrinogen but none of those immunized with unmodified fibrinogen developed marked swelling of the hind paws, which was followed by ankylosis. The histological changes were not entirely typical of human RA. There was synovial hyperplasia, but only a mild lymphocytic infiltrate and a notable absence of neutrophil infiltration. Importantly, wildtype mice immunized with either form of fibrinogen did not develop arthritis, nor did mice, including the transgenic immunized with murine unmodified or citrullinated fibrinogen. These data indicate the importance of both the human MHC II genotype, in this case DRB1*0401, and citrullination of the antigen, in this case the human antigen, in mediating pathology.

We would argue that this model, although not perfect, is the first prototype for true RA as opposed to other models, such as collagen-induced arthritis, which are more representative of inflammatory arthritis as a whole. Other studies have shown that immunization of several wildtype mice with citrullinated human fibrinogen together with adjuvant breaks tolerance, with an antibody response to both citrullinated and unmodified fibrinogen, but this procedure does not induce arthritis (76, 77). However, this observation does not exclude the possibility that antibodies to citrullinated fibrinogen are important in downstream pathogenic events in other mouse models. For example, co-administration of a murine monoclonal antibody binding to citrullinated human fibrinogen with a sub-optimal dose of anti-collagen II antibodies exacerbated arthritis, compared with administration of anti-collagen II antibodies alone (78).

Citrullinated vimentin

Vimentin is an abundant intermediate filament protein, involved in the dynamic organization of the cytoskeleton, with a vital function in organelle transport, cell migration, and proliferation. Citrullinated vimentin was first described as the Sa antigen, detected by immunoblotting using placenta and spleen extracts (79). Anti-Sa antibodies were detected in approximately 40% of rheumatoid serum samples and were predictive of a more severe disease (80). The identity of the Sa antigen was subsequently suggested (81) and later confirmed (64) to be citrullinated vimentin. Citrullinated vimentin is expressed *in vitro* in macrophage-like cells after ionophore-induced Ca^{2+} influx (27, 82). As these ionophores and the concomitant Ca^{2+} influx cause apoptosis, it was suggested that citrullination of vimentin is associated with apoptosis. Later, two independent groups (83, 84) demonstrated the presence

of citrullinated vimentin in the inflamed joint. In one of these studies, citrullinated isoforms lacking the amino-terminal region were detected in synovial tissue of patients with various arthritides (83). These particular isoforms could be the result of caspase-3 cleavage, as it is known that vimentin is cleaved by various caspases during apoptosis (85, 86) and that the amino-terminal cleavage products promote apoptosis (86). In the other study, Bang *et al.* (84) reported various citrullinated and mutated isoforms of vimentin in cellular synovial fluid from RA patients and their reactivity with serum antibodies by immunoblotting. Mass-spectrometry analysis demonstrated that the vimentin isoforms contained certain amino acid mutations and modifications, in particular mutations of glycine residues into arginine residues at positions 16, 59, 145, and 147, and citrullination of a number of arginine residues. Based on these findings, recombinant human vimentin was expressed, mutated at positions glycine-16 and glycine-59 into arginine and at arginine-50 into histidine, and citrullinated *in vitro*. The commercial test using this antigen is known as the anti-modified citrullinated vimentin (anti-MCV) assay and was found to be of superior diagnostic sensitivity (82%) compared with the anti-CCP2 assay (72%), without loss of specificity (84). A similar difference in sensitivity was also found in patients with early RA (anti-MCV: 70.7% versus anti-CCP2: 57.9%) (87), while another study reported that anti-CCP2 was superior (88). In a systematic review of 14 studies comparing the two assays (89), the authors concluded that the anti-MCV assay could be used as an alternative to the anti-CCP2 assay for diagnosis, though importantly, all studies found a handful of samples which were positive for anti-MCV but negative for anti-CCP2 or *vice versa*. This means that it is best to adopt one of the assays as a 'gold standard' for the purposes of epidemiological studies, and, for most, this choice has tended to be anti-CCP2.

The high diagnostic sensitivity and specificity of anti-MCV as well as its predictive value for erosive disease provide indirect evidence for citrullinated vimentin as an important molecule for driving the pathological immune response in RA. A vimentin peptide (^{65}SAVRAcitSSVPGVR^{77}), derived from human vimentin with leucine-69 replaced by alanine, was shown to bind HLA-DRB1*0401 with higher affinity in its citrullinated (at position 70) than in its native form. Immunization of HLA-DRB1*0401 transgenic mice with the citrullinated but not the native peptide induced a CD4+ T-cell response (90). The substitution of a leucine for alanine was designed to lock the adjacent arginine/citrulline in a stable position in the P4 pocket of the HLA-DRB1*0401 groove, though unpublished results showed that the leucine-containing peptide produced very similar results (David Bell, personal communication). These experiments demonstrate that at least one citrullinated peptide, but not its arginine-containing counterpart, can drive a T-cell response in the context of the HLA-DRB1*0401 molecule. However, unlike citrullinated fibrinogen, citrullinated vimentin has yet to be demonstrated to be arthritogenic in an *in vivo* animal model.

Given the fact that human and murine vimentin are highly conserved (97.4% sequence identity), animal models using antibodies to human citrullinated vimentin could provide important insights into the pathogenicity of these autoantibodies. Furthermore, a potential role for vimentin in the pathology of RA is beginning to emerge from studies of extracellular vimentin in inflammation. A recent study showed that extracellular vimentin mediates TGF-β activation (91), and Mor-Vaknin *et al.* (92) reported that during the inflammatory response to bacteria, cell-surface expression and secretion of vimentin in monocyte-derived macrophages was enhanced by TNF-α stimulation and mediated killing of internalized *E. coli*. Other studies have shown extracellular vimentin fragments on endothelial cells (91, 93) and on inflammatory cells, including apoptotic T cells (94), neutrophils (95), activated platelets (96), and macrophages (92, 93). Interestingly, various bacteria use cell-surface vimentin as an attachment receptor (97–99). Taken together, the extracellular location and the data demonstrating *in vivo* citrullination of vimentin support a role for citrullinated vimentin in pathogenesis and, in the presence of additional risk factors such as bacterial infection, possibly priming of an autoantibody response.

Citrullinated collagen type II

Collagen type II is a fibrillar protein and the major component of articular cartilage. Antibodies reactive with citrullinated collagen type II were originally reported in patients with early RA (65). Using an immunodominant citrullinated peptide from collagen type II, C1^{III} (^{359}AcitGLTGcitPGDA^{369}), the diagnostic sensitivity for RA was 40% (65, 71), and it remains an attractive antigen because it is likely to be present in degraded cartilage. Antibodies to unmodified collagen type II can also be found in RA patient serum, though at a lower frequency (around 15–25%), and are also present in other inflammatory and autoimmune diseases (65, 100–103). In a rat model of collagen-induced arthritis, Lundberg *et al.* (52) found that citrullinated rat collagen type II was more arthritogenic than the native form of the molecule. A more recent study demonstrated that monoclonal antibodies to the citrullinated C1^{III} peptide bound cartilage and inflamed

synovial tissue and induced or enhanced arthritis in mice (104). Importantly, the same study demonstrated *in vivo* citrullinated collagen peptides in synovial fluid from patients with RA. These data lend powerful support for citrullinated collagen type II being important in mediating pathology in both human RA and murine arthritis, though whether citrullinated collagen or a related molecule can break tolerance and induce autoimmunity remains unknown.

Citrullinated α-enolase

α-enolase was first described as a candidate autoantigen reactive with 25% of early RA sera, using human placenta and epithelial cell line extracts, by Saulot *et al.* (105). However, using recombinant α-enolase as the substrate, only eight of the 36 previously positive samples showed reactivity, strongly suggesting the presence of crucial post-translational modifications of α-enolase from the placental/cell line tissue extracts. We independently described the citrullinated form of α-enolase as a polypeptide in deiminated lysates of HL-60 cells reactive with serum from patients with RA (66). α-enolase was abundantly expressed in the synovial membrane, and antibodies reacting only with the citrullinated form of the molecule were specific for RA (sensitivity approximately 40%), compared with healthy serum. More recently, using cyclic citrullinated α-enolase peptides covering 15 out of 17 arginine residues present in human α-enolase, we identified a peptide containing immunodominant B-cell epitope(s) (106). This region comprised amino acids 5–21 of α-enolase, with arginine-9 and arginine-15 replaced by citrulline (^{5}KIHAcitEIFDScitGNPTVE^{21}). Carboxy- and amino-terminal cysteine residues were added to enable the peptide to adopt a cyclic conformation, and the resultant peptide is referred to as citrullinated α-enolase peptide-1 (CEP-1). Using mass spectrometry, we have confirmed that both arginine-9 and arginine-15 can be citrullinated *in vitro* (66). A three-dimensional representation of the human α-enolase protein shows that the CEP-1 peptide is present near the amino-terminus on a pole of the molecule, comprising the first and part of the second β strand, and is surface-accessible (*Fig. 3A-C*). Arginine-15, which is the crucial amino acid in CEP-1, is located in the loop region between the first and the second β strand. Taken together, these observations could explain why CEP-1 might be a dominant epitope and why a cyclic version of the peptide may be required for the detection of antibodies. In our studies of over 300 serum samples, antibodies to CEP-1 were found in 37% of RA patients, 3% of disease controls, and 2% of healthy controls (106). Similar frequencies of anti-CEP-1 antibodies (41%) were found in a large cohort of Swedish patients with RA (71). A much lower frequency of antibodies (13.9%) to a linear peptide from the same region of α-enolase (amino acids 5–20), shortened by one amino acid and without added amino- and carboxy-terminal cysteine residues, was found in a separate study (107). We suggest that it is the linear rather than the cyclic version of the peptide that explains these differences, as only a proportion of linear peptides will adopt a conformation similar to the original epitope (11). Other factors may explain differences in autoantibody frequencies. For example, in our original description of anti-CEP-1, we were surprised to find a much higher frequency of anti-CEP-1 antibodies (62%) in a cohort of sera derived from the US (106). We have no explanation for this discrepancy, though it may have been related to different genetic backgrounds or different smoking habits (see below), details which were not available in this cohort. In addition, as the patients in the US cohort were participating in clinical trials, they may have had more active RA, resulting in a higher frequency of antibodies.

The prevalence of anti-CEP-1 antibodies may not necessarily represent the true frequency of anti-citrullinated α-enolase antibodies as a whole. Using recombinant whole human α-enolase protein citrullinated *in vitro*, we found that detection of antibodies by immunoblotting increased to 70% in RA compared with 5% in healthy controls (A. Kinloch, K. Lundberg, P.J. Venables, unpublished data) (*Fig. 5*). In this study, we used highly purified recombinant human α-enolase, which may explain the higher frequency than that found in our first report (40%) (66), in which a semi-purified, commercially available rabbit α-enolase was used. Using purified recombinant human α-enolase, the healthy controls reacted to non-citrullinated and citrullinated alpha-enolase with an equal frequency (5%) (*Fig. 5*), suggesting that antibodies to non-citrullinated epitopes on the molecule could occur in a substantial proportion of the population. This finding is in accordance with studies reporting anti-α-enolase antibodies in cancer and a whole range of other inflammatory diseases, among which Behçet's syndrome and autoimmune hypophysitis are particularly prominent (103–112).

Similar to other established antigens in human autoimmune diseases, enolase is a highly conserved and ubiquitous protein (reviewed in 113). Mammals express three enolase isotypes (α, β, and γ), which form homodimers (*Fig. 3*) or heterodimers. α-enolase is the most ubiquitous isotype and is expressed in a variety of tissues, while β-enolase is exclusively found in muscle tissues and γ-enolase in neuronal and neuroendocrine tissues. Enolase is a multifunctional protein: it catalyses the penultimate step in glycolysis but also plays a role in

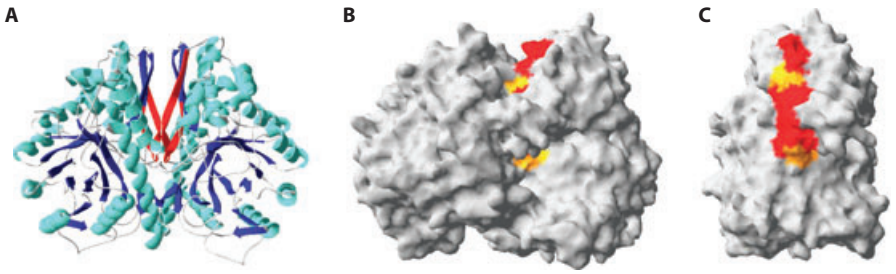


***Fig. 3.* Molecular representation of the human α-enolase dimer (A, B) and monomer (C).** (A) Ribbon diagram. The immunodominant B-cell epitope (CEP-1) is highlighted in red. Secondary structure elements (helices, light blue; strands, dark blue; loops, grey) are labeled. Bound magnesium and sulphate ions are not depicted. (B) Molecular surface model. CEP-1 is highlighted in red, with the two arginines in orange. (C) Molecular surface model of the monomer, looking onto the dimer interface. Color coding as in (B). This figure was generated using the software DeepView v4.0 (http://spdbv.vital-it.ch/) and is based on the X-ray crystal structure published by Kang *et al.* (132).

various processes, such as hypoxia tolerance, growth control, and in fibrinolysis, extracellular matrix remodeling and cell migration, due its ability to bind plasminogen on the cell surface of various eukaryotic cells (110, 112, 114–117) and a number of bacteria (118–125), fungi (126), and parasites (127–129). Using human cell lines and a monoclonal antibody directed towards the plasminogen-binding region of α-enolase, Lopez-Alemany *et al.* (117) showed that α-enolase, despite being responsible for only around 15–20% of the total plasminogen binding, accounts for up to 90% of plasminogen activation on the cell surface. It was also reported that plasminogen upregulates α-enolase gene expression in fibroblasts and peripheral blood mononuclear cells (130), supporting an important physiological role of α-enolase in the plasminogen system. A recent study by Wygrecka *et al.* (116) demonstrated a role for cell-surface α-enolase in inflammatory cell recruitment. They showed that cell-surface expression of α-enolase is upregulated in monocytes upon lipopolysaccharide stimulation, and enhances plasmin generation, monocyte cell migration, and matrix invasion in inflamed lungs (116).

The multifunctional nature of the molecule, involving interactions with both cells of the immune system and infectious agents or their products, provides the basis for a hypothesis of its involvement as an autoantigen in the etiology and pathogenesis of RA. We have already demonstrated increased expression of α-enolase in synovial membrane from patients with RA compared with OA (*Fig.* 4) and more recently shown that α-enolase is upregulated in cell-free synovial fluid in inflammatory arthritis (28). Using synovial fluid sample from a patient with RA, a shift towards a more acidic isoelectric point (pI) in the two-dimensional electrophoretic migration of alpha-enolase was observed (28). This relative 'acidic shift' towards the anode is a characteristic of protein citrullination due to the resulting decreased net positive charge. Taken together, these data provide some evidence that α-enolase is involved in the pathogenesis of RA, though animal model data, such as that already available for citrullinated fibrinogen and collagen, are awaited.

Given the multitude of physiological roles in which α-enolase is involved, an obvious question is as to what the possible consequences of citrullination on the structure and function of this protein are. A number of arginine residues on α-enolase, including the ones located within the CEP-1 epitope, can be citrullinated *in vitro* using rabbit PAD2 and α-enolase from HL-60 cell lysates (66). Citrullination is likely to change the conformation of the enolase monomer and could influence the non-covalent interactions involved in the formation of the enolase dimer. This might result in an altered glycolytic activity and plasminogen-binding potential. As the plasminogen-binding sites are located on the surface of the molecule [residues 250–256 and carboxy-terminal lysine residues (131)], autoantibodies could interfere with the coagulation pathway. Given the fact that excessive fibrin deposition is a common feature of RA joints (51) and that mice defective of fibrinolysis exhibit an exacerbated arthritis (132), we could speculate that citrullination of cell-surface α-enolase abrogates its plasminogen-binding and -activating function and contributes to the decreased fibrinolysis seen in RA joints.

Other potential citrullinated antigens

It is apparent that the targets of anti-citrullinated protein antibodies extend beyond the four listed above. For example, in the study that led us to identify citrullinated α-enolase as a candidate autoantigen in RA, we also identified a number of other citrullinated proteins in *in vitro* deiminated HL-60 cell lysates. Some of these also reacted specifically with rheumatoid serum, including elongation factor 1α (EF1α) and adenyl cyclase-associated protein 1 (CAP1) (66). In a more recent study, Goeb *et al.* (133) performed similar experiments, using

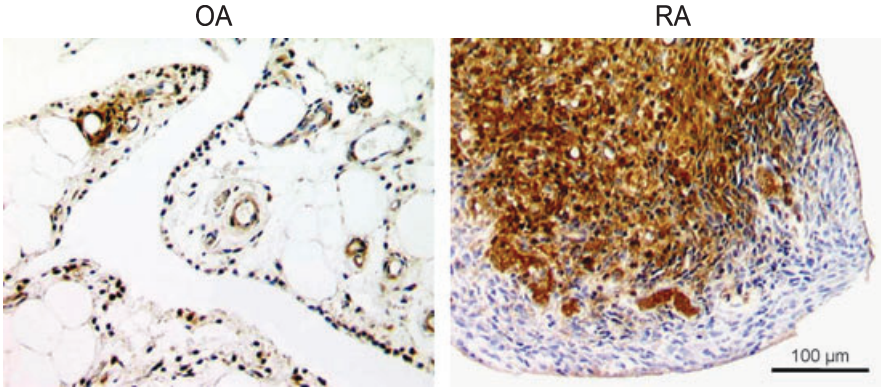


***Fig. 4.* Expression of α-enolase (stained brown) in synovial membranes from a patient with osteoarthritis (OA) and a patient with rheumatoid arthritis (RA).** Strongest staining was observed in the subsynovial layer in RA sections and in vascular endothelial cells in OA sections. Cell nuclei are counterstained in blue.

HL-60 cell lysates with no additional *in vitro* deimination step, where they identified a number of citrullinated glycolytic enzymes and molecular chaperons as potential targets. Citrullinated α-enolase was confirmed as a specific autoantigen in RA, but also novel citrullinated candidates such as citrullinated aldolase, phosphoglycerate kinase 1 (PGK1), calreticulin, heat shock protein 60 (HSP 60) and far upstream element-binding proteins 1 and 2 (FUSE-BPs), were described.

To determine which citrullinated proteins are generated and targeted in the rheumatoid synovium, Matsuo *et al.* (67) blotted synovial tissue extract from an RA patient with a pool of RA serum and in parallel with an anti-modified citrulline antibody. Eight citrullinated candidate antigens were reported: asporin, cathepsin D, β-actin, F-actin capping protein α-1 subunit (CapZα-1), albumin, histamine receptor, protein disulfide-isomerase ER60 precursor, glucose-regulated protein, and mitochondrial aldehyde dehydrogenase (ALDH2). Importantly, (citrullinated) fibrinogen was also detected, which could be seen as a proof of concept result for the method. However, the study did not provide convincing evidence that the peptides identified by MALDI-TOF MS contained citrulline residues. The presence of autoantibodies to recombinant citrullinated and non-citrullinated CapZα-1 was then investigated using an ELISA with serum samples from 30 RA patients, 28 OA patients, 19 with SLE, and 31 healthy controls. The results showed reactivity to citrullinated CapZα-1, which was increased in RA patients (53%) versus OA (7%), SLE (5%), and healthy controls (6%), but also to the non-citrullinated protein (RA 36%; OA 10%; SLE 5%; healthy controls 6%). More detailed analyses of individual samples revealed that citrulline-specific reactivity (i.e. antibodies to the citrullinated but not the unmodified protein) was observed in 16% of RA patients, in one OA patient (3%), and in none of the other controls.

Using a different approach, Okazaki *et al.* (134) performed immunoscreening of human chondrocyte cDNA expression libraries modified by PAD4. A citrullinated fragment of eukaryotic translation initiation factor 4G1 (eIF4G1) was identified. With ELISA and Western blot, they demonstrated that 48 out of 100 RA patients had autoantibodies to the citrullinated eIF4G1 fragment, nine out of 100 had autoantibodies to the native form, while no reactivity was detected in 44 healthy controls.

Confirmatory data in large-scale cohorts as well as studies where (i) the citrulline specificity, (ii) the disease specificity, and (iii) the *in vivo* demonstration of these new candidates are investigated will determine whether they represent 'true' autoantigens in RA or not. Screening strategies using anti-citrullinated protein antibodies purified from RA serum and synovium-derived citrullinated proteins in their native conformation should help to establish the repertoire of true citrullinated autoantigens in RA.

Antibodies to citrullinated proteins in RA: citrulline specific, peptide specific, or antigen specific?

Citrulline specific?

The identification of at least four good candidate citrullinated antigens in RA, with potential others awaiting further characterization, could suggest that it is merely the presence of citrulline that confers antigenicity and that the molecule on which the residue is found is irrelevant. This view is widely held, which has led to terms such as 'anti-citrulline autoimmunity'. In one study (107), 'anti-Sa' antibodies in an ELISA were detected not using citrullinated vimentin but

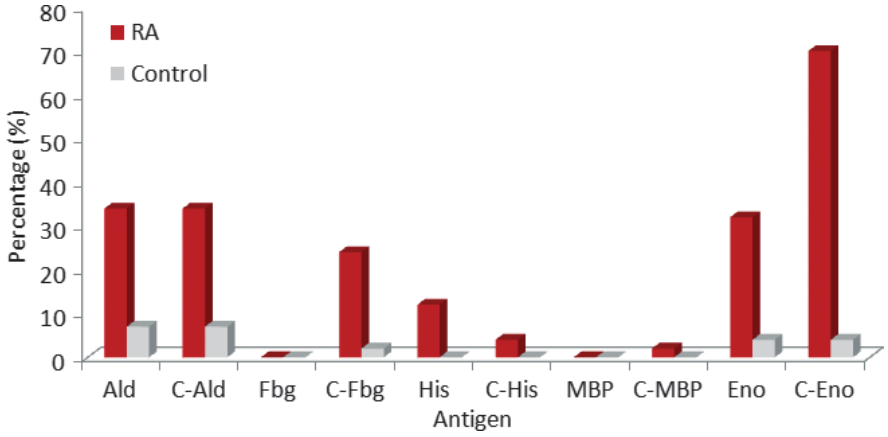


***Fig. 5.* Percentage of serum samples from RA patients (red bars) and healthy controls (grey bars) that are seropositive for the respective native or citrullinated antigens.** Serum samples from 50 patients with RA and 45 healthy controls were screened by immunoblotting for IgG reactivity to aldolase-A (Ald), fibrinogen (Fbg), total histones (His), myelin basic protein (MBP), and α-enolase (Eno), either in their native or *in vitro* citrullinated form (C-). Significant reactivity in RA serum was observed with aldolase-A, fibrinogen, histones, and α-enolase, though increased reactivity with the citrullinated form of the proteins was observed only for fibrinogen and α-enolase.

in vitro citrullinated MBP. This was justified on the basis that commercially available MBP was less expensive than vimentin and that it was a valid substitute based on the fact that it had a similar proportion of arginines. It was suggested that antibodies detected with citrullinated MBP by ELISA closely correlated with positive results obtained in the original anti-Sa test, detected using immunoblotting on human placenta extracts (135), although no data in support of this were shown. In contrast, our recent preliminary study of several purified antigens showed that only two out of 50 RA serum samples reacted weakly with citrullinated MBP (*Fig.* 5). It could be argued that the low frequency of this reactivity, as detected by Western blotting, was due to technical factors or to the presence of largely conformational epitopes detected by ELISA in the study described above. We would counter this by proposing that most true autoantibodies react with multiple epitopes on the same molecule, both linear and conformational. For example, in the same cohort, antibodies to citrullinated fibrinogen and α-enolase were readily detected by this method in a significant proportion of the serum samples (*Fig.* 5). Further evidence for the importance of more than just a citrulline residue comes from studies showing that not all proteins become targets for anti-citrulline antibodies after *in vitro* citrullination [for example, citrullinated bovine albumin (9) and citrullinated fibrinogen γ chain (63)].

Peptide specific?

Another concept, which is widely held and which has more evidence in its support, is that the autoantibody response in RA is determined by the presence of citrulline residues in association with certain flanking amino acids that facilitate antibody binding. In the first study of sequences derived from filaggrin (10), it was shown that peptides containing citrulline flanked by neutral and relatively flexible residues, such as glycine and serine, showed greater sensitivity for the diagnosis of RA than those which contained strongly charged residues or imposed a rigid structural conformation, such as proline. In the case of CEP-1 from α-enolase, there are two citrulline residues, the second of which at position 15 is flanked by a serine and glycine. Girbal-Neuhauser *et al.* (9) demonstrated that two citrulline-containing peptides from filaggrin harbor epitopes recognized by RA serum. Both peptides contained a citrulline residue flanked by a serine and a histidine residue (9). Evidence for the importance of structural homology of the surrounding amino acid residues on immunodominant epitopes from different antigens comes from a subsequent study, where Sebbag *et al.* (72) used five fillagrin-derived peptides, including the two described above, in inhibition studies, showing that these peptides can inhibit RA serum reactivity towards citrullinated fibrinogen. As the identified immunodominant B-cell epitopes on fibrinogen do not share sequence homology with the filaggrin peptides, it was concluded that structural homology of surrounding amino acids is the crucial factor. Thus, it could be argued that the antibody response in RA is entirely dependent on the presence of citrulline residues with appropriate surrounding amino acids with structural homology and that the candidate antigens are no more than molecules which happen to possess these sequences.

Antigen specific?

More recent data support the concept that it is the whole molecule that drives the autoimmune response rather than specific citrulline-containing epitopes. Using absorption experiments, we showed that a given serum sample reacting with citrullinated α-enolase peptides reacted independently with different epitopes on the molecule (106). Similar results were also found for citrullinated fibrinogen by Sebbag *et al.* (72) and in a study of different citrullinated antigens by Snir *et al.* (71). Anti-CEP-1 reactivity was not inhibited by absorption with a citrullinated collagen type II peptide, though complete inhibition occurred with the CEP-1 peptide. While this result might be due to sufficient dissimilar structural homology between the CEP-1 and collagen type II peptides, the occurrence of independent antibody reactivities to different (non-cross-reactive) epitopes on a handful of antigens strongly suggests that the ACPA response is antigen-specific. Absorption of serum with whole citrullinated fibrinogen, which potentially contains a large number of citrullinated peptide epitopes, had little effect on binding to peptides from the other antigens (α-enolase and collagen type II) tested. This study provided evidence that at least three citrullinated molecules, fibrinogen, collagen type II, and α-enolase, could justifiably be termed 'specific citrullinated antigens'. A heat map was constructed to visually compare autoantibody reactivities to CCP2, citrullinated vimentin peptide (residues 60–75), citrullinated fibrinogen (whole protein), CEP-1, and citrullinated collagen type II peptide (residues 359–369) within individual rheumatoid serum samples (Fig. 6). The majority of the serum samples reacted with more than one antigen, though each serum had its own pattern of reactivity and most fell within the anti-CCP2 positive population. These findings are in accordance with the concept that disease-specific autoimmunity is characterized by reactivity to a 'family' of anatomically or functionally related antigens (113).

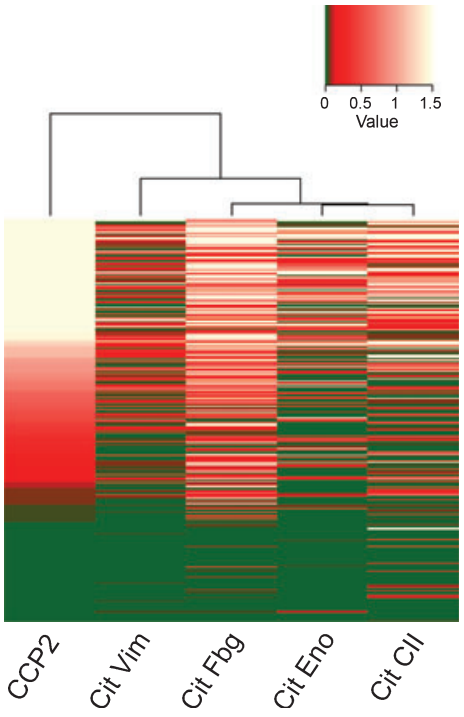


***Fig. 6.* Clustering of IgG RA-associated antibodies.** IgG reactivity to CCP2, a citrullinated vimentin peptide (aa 60–75) (Cit Vim), citrullinated full-length fibrinogen (Cit Fbg), the immunodominant epitope (CEP-1) of citrullinated α-enolase (Cit Eno), and the triple helical peptide of the CII C1-epitope (aa 359-369) (CitCII) was measured in serum from 291 RA patients. Each row represents one patient. The color scale shows the relative degree of antibody reactivity, from low reactivity (green) to high reactivity (light red). The order of the columns is according to their relationship with each other and anti-CCP.

Genes and autoimmunity to specific citrullinated antigens

Autoimmune diseases are polygenic, with each gene conferring a relatively small risk but when added together producing an interacting profile leading to an enhanced probability for a common phenotype (reviewed in 136). In most autoimmune diseases, the MHC class II provides the strongest risk, with individual alleles often showing more association with specific autoantibodies than with the disease itself. For example, in a cohort of 157 patients with systemic lupus erythematosus, the well-established association with HLA-DR3-containing haplotypes was confirmed, with an odds ratio (OR) of 2.5, though in the small subset of patients positive for anti-La antibodies (13%) the OR rose to an astonishing 71 (137). Analogous results were also published in relationship to polymyositis (138). Again, DR3 showed an association with the disease as a whole, but a marked increased risk was seen in patients with antibodies to PM/Scl. In the same cohort, a group with antibodies to U1 RNP all possessed at least one copy of DR4.

It is now beginning to emerge that a similar pattern of susceptibility, namely that MHC class II alleles are more associated with autoantibody-specificities than with disease, also applies to RA. The major susceptibility alleles for RA are present in HLA-DR4, -DR1, and -DR10, principally DRB1*0101, *0102, *0401, *0404, *0405, *0408, *1001, and *1402. All share variants of the EQKRAA motif, present in the third hypervariable region of the DRB chain, which constitutes part of the peptide binding cleft of the class II molecule. These alleles are collectively known as the 'shared epitope' (SE), implying that any of these alleles confer risk for RA (139). The universality of the SE hypothesis has been challenged by the findings that DR1 alleles (DRB1*0101 and DRB1*0102)

tend to associate with milder disease (reviewed in 140). Thus, when analyzing DRB1*SE susceptibility to individual autoantibodies, it is important to consider the DR1 and DR4 haplotypes separately.

In RA, it is now well established that the SE primarily determines the presence of anti-CCP antibodies rather than the disease itself (141, 142). Not surprisingly, given that anti-CCP antibodies predict more severe disease, this association is the strongest for DR4 alleles. The link with DR4 has also been confirmed with antibodies to specific citrullinated antigens, including antibodies to a citrulline-containing peptide from vimentin (^{60}VYATcitSSAVcitLcitSSVP^{75}) (143), CEP-1 from α-enolase, and, to a lesser extent, the collagen type II peptide already described above (71). Discrepant results have been obtained in relationship to fibrinogen. Using a peptide (^{6}NEEGFFSAcitGHcitPLDKK^{22}) from the fibrinogen β chain as antigen, there was no association with the SE (143), whereas using the whole molecule citrullinated *in vitro*, an association was shown both with SE in general and DR4 in particular (71). These results could indicate that the initial autoimmune response, determined by the MHC class II, was directed against a fibrinogen epitope on a different part of the molecule or that antibodies to citrullinated fibrinogen may be more of a downstream event in pathogenesis, resulting from epitope spreading. However, it should be emphasized that in the animal model of citrullinated fibrinogen-induced arthritis (60), only mice transgenic for human HLA-DRB1*0401 developed arthritis, whereas wildtype controls did not. This study provides some evidence that HLA-DRB1*0401 is required for an immune response to citrullinated fibrinogen.

Citrullinated α-enolase and vimentin appear to be the most promising candidates for priming an MHC class II-determined autoimmune response, though a direct interaction between peptide and HLA-DRB1*SE alleles has yet to be demonstrated for α-enolase. Nevertheless, there is indirect evidence for epitopes within citrullinated α-enolase being important in relation to the SE. We have recently studied three large cohorts, of approximately 1000 RA cases, and demonstrated that the HLA-DRB1*SE association is considerably stronger for anti-CCP positive patients that also have antibodies to CEP-1, compared with anti-CCP positive patients that do not have these antibodies (144).

Genome-wide association studies have contributed significantly to the identification of other gene loci in polygenic diseases (reviewed in 137), and the following loci have now been confirmed to associate with RA in several large cohorts: *TRAF1-C5*, *STAT4*, *REL*, *TNFAIP3*, *CTLA*, and *CD40*. Of relevance to citrullination, *TRAF1-C5* has been associated with anti-CCP positive disease (145), as have polymorphisms of the PAD4 gene in Japanese and Korean populations but notably not in those derived from Western Europeans (reviewed in 137). In one large case–control study, at least three loci, the MHC, the *PTPN22*, and variants around *IL2RA* (chromosome 10p15), were associated with RA, with several other SNPs being identified as candidates for further investigation (146).

In spite of the discovery of increasing numbers of susceptibility genes, it is MHC that consistently provides the strongest genetic risk, which, as discussed above, has already enjoyed more that 30 years of investigation. Polymorphisms of *PTPN22* have also already been described as a risk factor, not only for RA (147) but also for other autoimmune diseases (137), and may well be a risk factor for autoimmunity in general. The gene encodes a tyrosine phosphatase involved in T and B-cell signaling, and the susceptibility allele 620W of *PTPN22*, which is present in approximately 17% of healthy controls and in 28% of individuals with RA, disrupts the proline-rich motif of the protein that is important for its normal function as a negative regulator of T and B-cell activation (147). Like the SE, the risk allele of *PTPN22* is more related to the anti-CCP positive subset of RA than to the disease as a whole, and when combined in an additive analysis as a 'gene/gene interaction' with the SE, produces an even higher OR than either gene alone (148). In our recent study, we dissected this effect further by showing that the combined effect of the two genes conferred a higher risk for the anti-CEP-1 positive/anti-CCP positive population (OR = 17.8) compared with the anti-CEP-1 negative/anti-CCP positive (OR = 4.6) (143). In this study, we did not investigate the possibility that other antibodies to other citrullinated antigens would show the same powerful associations with interacting risk genes. Given the data published on HLA-DRB1*SE interactions (144), it might be predicted that antibodies to citrullinated vimentin peptides may show a similar effect. Preliminary studies suggest that this is the case (Renne Toes, personal communication), though this kind of analysis in large cohorts of patients has yet to be published.

In our study discussed above, we also found an association between CEP-1-positive disease and polymorphism in the bromodomain-containing 2 (*Brd2*) gene. *Brd2* has not previously been linked to RA or any other autoimmune condition. A substantial literature exists on genetic polymorphism in *Brd2* in association with idiopathic generalized epilepsy (reviewed in 149). *Brd2*, situated within the MHC region, encodes a ubiquitously expressed protein kinase, associated with transcription complexes and acetylated chromatin during mitosis. It has been implicated in fundamental cellular processes, such

as cell-cycle regulation and transcriptional regulation (150, 151). Mice deficient in Brd2 do not complete embryogenesis and neurogenesis (152), while constitutive expression of Brd2 was shown to cause malignancy in mice, similar to human diffuse large B-cell lymphoma (151). The association between *Brd2* and CEP-1-positive RA needs to be confirmed in other cohorts, and further studies will determine the functional consequences of variants of the gene and their relationship with autoantibody specificities in RA.

Environmental factors in the development of autoimmunity to specific citrullinated proteins

Smoking

Cigarette smoking is today the best known environmental risk factor for the development of RA. The link to smoking was first recognized in 1987 (153), as an unexpected finding in a study investigating the association between RA and the use of oral contraceptive, and later confirmed in a number of case-control and cohort studies. Perhaps, the most striking results were presented from the Arthritis and Rheumatism Council Twin Study, where 13 pairs of monozygotic twins, discordant for RA and smoking, were identified, and in 12 out of 13 cases the RA patient was also the smoker (154).

Cigarette smoking constitutes more of a specific risk for RF-positive (155–160) and most notably anti-CCP-positive RA (141, 161–163). Furthermore, this risk is strongly dependent on the presence of HLA-DRB1*SE alleles, and evidence exists for a biological gene–environment interaction (142, 161–164). More recently, we have found that smoking and HLA-DRB1*SE is primarily associated with the anti-CEP-1-positive subset of anti-CCP-positive RA (144). In a large case–control study, comprising 1000 RA cases and 872 matched healthy controls, we compared the anti-CEP-1-positive/-CCP-positive population with the anti-CEP-1 negative/-CCP-positive population. In the subset with antibodies to CEP-1, we could demonstrate an impressive OR of 37 for smokers with the SE and the susceptibility allele of *PTPN22* compared with an OR of six in the corresponding subset negative for anti-CEP-1 (*Fig. 7*). These data strongly suggest that CEP-1 is at least one peptide from a specific autoantigen which links smoking to HLA–DRB1*SE in the development of RA.

It is currently not known how smoking may contribute to the development of CEP-1-positive/CCP-positive RA. Cigarette smoke contains thousands of potential toxic compounds that may cause tissue damage and inflammation (165). Recruitment and activation of immune cells, levels of CRP, matrix metalloproteinases, IL-6, IL-1β, TNF-α, and serum levels of fibrinogen, as well as the expression of Fas on T and B cells, have been reported to be elevated in response to cigarette smoke (reviewed in 166). In support of the hypothesis linking smoking-induced inflammation and citrullination are data demonstrating higher levels of both PAD2 and citrullinated proteins in the lungs of smokers as compared with non-smokers (142, 167). The presence of IgA anti-CCP in approximately 60% of the IgG anti-CCP-positive patients (168) also points to the involvement of mucosal surfaces, such as the lung, in generating or sustaining autoantibodies to citrullinated proteins. Furthermore, the occurrence of IgA anti-CCP has been reported to be associated with smoking (169). However,

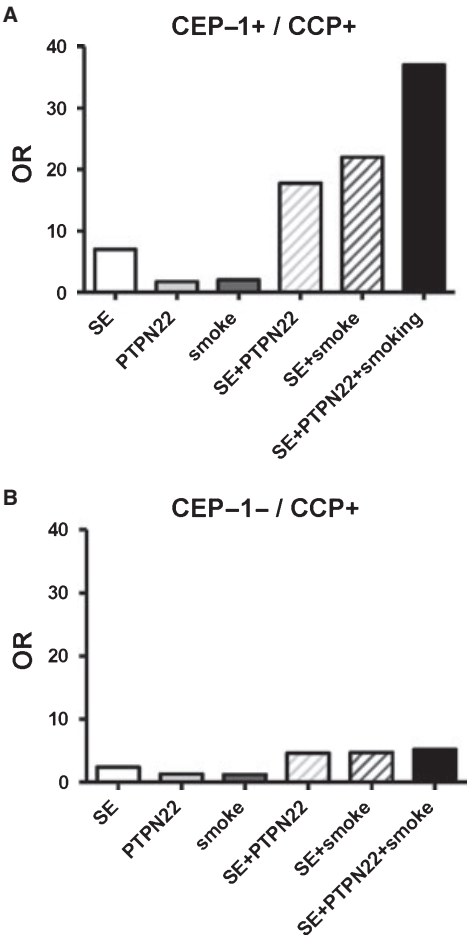


***Fig. 7.* Single and combined effects of the three risk factors HLA-DRB1*SE (SE), the W620 allele of PTPN22 (PTPN22) and cigarette smoking (smoke), for the two anti-CCP positive RA subsets.**
(A) CEP-1+/CCP+ and (B) CEP-1−/CCP+. Odds ratios (OR) were calculated by means of logistic regression in a case-control analysis (1000 RA cases and 872 healthy controls), where adjustments for age, residential area and gender were made. Both the single and the combined effects of the three risk factors were strongest for the CEP-1+/CCP+ subset.

not all anti-CEP-1-positive/anti-CCP-positive and HLA-DRB1*SE-positive patients are past or current smokers, and the interaction between smoking and HLA-DRB1*SE could not be confirmed in three North American cohorts (170). Other factors, such as air pollution, silica dust exposure, or interstitial lung disease could have masked the effect of smoking. Alternatively, the generation of citrullinated antigens and an antibody response could occur, in a similar fashion as hypothesized for smoking, at another site in response to other inflammatory stimuli, such as infection.

Infection: *Porphyromonas gingivalis*

Infection as a contributory or even sole etiological factor in RA has been suggested and discussed for decades (171, 172). Here, we focus on a possible role of infection with *Porphyromonas* (P.) *gingivalis* in the development of autoimmunity to citrullinated proteins, based on its expression of a bacterial PAD enzyme (173) with the potential to generate citrullinated epitopes both on bacterial antigens and human autoantigens. This bacterium plays a significant role in the progression of chronic periodontitis (174) and has been proposed as an environmental risk factor for the development of RA (106, 175–182).

The oral cavity is a diverse ecological reservoir for a large number of bacterial species, with an estimated 500 colonizing species (183). P. *gingivalis* is a Gram-negative bacterium and is the most extensively studied of all major periodontopathogenic organisms (reviewed in 184). It is commonly present in periodontitis patients as a biofilm in the gingival crevice and intracellularly in oral epithelial cells but can also be found in periodontally healthy subjects, although at lower numbers (185–187). P. *gingivalis* produces an array of virulence factors, most notably extracellular cysteine proteases called gingipains, to support its adherence, growth, tissue invasion and degradation, and to evade and subvert the immune system (184, 188–192).

Analogous to RA, periodontitis is a chronic inflammatory disease characterized by erosion of bone. The prevalence of periodontitis in the American population was reported to be 4.2% (193), although this varies depending on the definition criteria of periodontitis and the applied examination protocol, among other factors. Periodontitis usually begins with a self-sustaining inflammation as a result of plaque accumulation (i.e. gingivitis) and then progresses into a chronic, erosive disease mediated by TNF-α, IL-1β, prostaglandins E_2, and metalloproteinases. With progressing inflammation, lymphocytes and monocytes accumulate, and more degradative and pro-inflammatory molecules are released, leading to erosion of adjacent bone.

Studies in our laboratory have recently shown that P. *gingivalis* contains a range of endogenous bacterial citrullinated proteins (N. Wegner, K. Lundberg, P.J Venables, unpublished data). Citrullination of a human protein, fibrinogen, by P. *gingivalis* has been stated (177) and recently repeatedly cited (175, 178, 194, 195), despite the fact that experimental proof supporting these claims has never been published, including the referenced article which is cited in these publications.

The citrullination of bacterial and/or host proteins by an external PAD enzyme in the infectious context of chronic gingival inflammation could prime autoimmunity to citrullinated proteins in individuals with additional susceptibility factors. Two small-scale studies have found slightly higher than expected rates of anti-CCP antibodies in patients with periodontitis. Two out of 27 (196) and two out of 49 (174) periodontitis patients, who had no history or current manifestations of RA, were found to be anti-CCP positive. We have previously shown by Western blotting that autoantibodies to CEP-1, affinity-purified from an RA patient, cross-react with *in vitro* citrullinated recombinant P. *gingivalis* enolase (106). Whether P. *gingivalis* enolase is an *in vivo* target of bacterial PAD has yet to be demonstrated. Both bacterial enolase (Fig. 8) and PAD (173) are present on the cell surface of P. *gingivalis*. In addition, exposure of neutrophils, epithelial cells, and platelets to P. *gingivalis* has been shown to increase intracellular

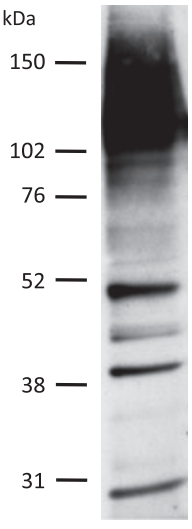


***Fig. 8.* Western blot showing the presence of extracellular enolase in the cell membrane fraction of *Porphyromonas gingivalis*.** Enolase was detected using an antibody raised against a conserved peptide from the amino-terminus of human α-enolase.

Ca^{2+} concentrations in these cells via cleavage and activation of human proteinase-activated receptor-2 (PAR-2) (197–202), which might lead to intracellular activation of human PAD enzymes. Hence, human PAD enzymes could also contribute to the citrullination of proteins at the site of gingival inflammation. Very little is known about the enzymatic and functional properties of bacterial PAD, which might generate different epitopes than the human PAD enzymes, and hence needs to be addressed in future studies.

There is also epidemiological evidence linking periodontitis and RA (reviewed in 178) and some evidence of a common pattern of genetic susceptibility and environmental risk factors. Genetic influences account for around 50% of the total risk for developing chronic periodontitis (203). Gene polymorphisms associated with periodontitis were reported for a number of genes encoding cytokines, Fc receptors, and, notably, MHC class II (reviewed in 204, 205). Susceptibility to periodontitis has been linked to the same HLA-DRB1 alleles which predispose to RA, specifically the *0401 allele. In a study of 48 patients with severe periodontitis, Bonfil *et al.* (206) found that HLA-DRB1 subtypes *0401, *0404, *0405, and *0408 are more frequent in patients with severe periodontitis compared with healthy controls. Other studies (207–211) also demonstrated an increased frequency of HLA-DR4 in patients with various forms of periodontitis.

Smoking, now established as an environmental risk factor for RA, has also been linked to periodontitis with an OR in the order of 2–6 (reviewed in 212), and there is a dose–response relationship between the number of cigarettes smoked per day and the odds of developing disease (213). Smokers have more severe periodontitis, as measured by the degree of alveolar bone erosion and/or loss. In a large study of 1361 subjects, smokers were at greater risk for severe bone loss than non-smokers, with odds ratios of 3.25 and 7.28 for light and heavy smokers, respectively (214). Levels of cotinine, the principle metabolite of nicotine, have been shown to correlate directly with periodontal destruction (215, 216). The direct role of smoking in the pathogenesis of periodontitis is uncertain, but studies have shown increased levels of TNF-α in gingival crevicular fluid of smoking as compared with non-smoking periodontitis patients (217–219). As discussed previously, cigarette smoke has been demonstrated to have several pro-inflammatory properties (reviewed in 166). Using various *in vitro* systems and compounds to mimic smoking, a trend towards functionally compromised phenotypes in phagocytic cells, B, and T lymphocytes has been observed (reviewed in 212, 220). However, it is important to bear in mind that *in vitro* smoking studies use different compounds, cells, and experimental setups to mimic smoking and therefore afford only a limited reproducibility of the true short- and long-term physiological events accompanying tobacco smoking.

The direct association between smoking and both periodontitis and RA also points out an important confounding factor for epidemiological studies on the association between periodontitis and RA. Some studies (195, 221) included statistical adjustment for smoking in their analysis; however, these were based on questionnaires, often containing inadequate documentation of smoking variables, with the potential to inflate the incidence of RA in individuals with periodontitis (222). In spite of these limitations, it can be concluded that periodontitis and RA have similar genetic and environmental risk factors leading to analogous patterns of pathology. A definitive causative role for *P. gingivalis* requires much larger and appropriately adjusted epidemiological studies and more molecular work on the targets for its PAD enzyme and the mechanisms for the generation of potential autoantigens.

Conclusion

We have summarized evidence that antibodies to specific citrullinated antigens can be traced to the immune response genes, HLA-DRB1*0401 and HLA-DRB1*0404, with the *PTPN22* gene and the novel gene *BRD2* conferring additional risk to subsets of the autoimmune response in RA. Smoking and *P. gingivalis* are attractive environmental risk factors, because both are likely to be associated with citrullination of autoantigens *in vivo*, though the mechanisms involved require further elucidation. Whether, in the presence of susceptibility genes, citrullination itself, through exposure of new self-epitopes, or citrullination by the bacterial PAD in an infectious context are enough to break tolerance has yet to be investigated. An alternative hypothesis, suggested by the identification of citrullinated enolase as at least one disease-specific autoantigen that is conserved in prokaryotes, raises the possibility that an immune response to the bacterial form of the protein could trigger autoimmunity through molecular mimicry. Whatever the mechanisms involved, both epidemiology and molecular immunology point to specific citrullinated protein antigens as being key factors that link genes and environment to the etiology of RA. We propose that further investigation of these mechanisms could lead us towards the cause of the disease.

References

1. MacGregor AJ, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;**43**:30–37.
2. Arnett FC, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;**31**:315–324.
3. Nienhuis RL, Mandema E. A new serum factor in patients with rheumatoid arthritis; the antiperinuclear factor. Ann Rheum Dis 1964;**23**:302–305.
4. Nijenhuis S, Zendman AJ, Vossenaar ER, Pruijn GJ, vanVenrooij WJ. Autoantibodies to citrullinated proteins in rheumatoid arthritis: clinical performance and biochemical aspects of an RA-specific marker. Clin Chim Acta 2004;**350**:17–34.
5. Young BJ, Mallya RK, Leslie RD, Clark CJ, Hamblin TJ. Anti-keratin antibodies in rheumatoid arthritis. Br Med J 1979;**2**:97–99.
6. Simon M, et al. The cytokeratin filament-aggregating protein filaggrin is the target of the so-called ''antikeratin antibodies,'' autoantibodies specific for rheumatoid arthritis. J Clin Invest 1993;**92**: 1387–1393.
7. Sebbag M, et al. The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies. J Clin Invest 1995;**95**: 2672–2679.
8. Hoet RM, Boerbooms AM, Arends M, Ruiter DJ, van Venrooij WJ. Antiperinuclear factor, a marker autoantibody for rheumatoid arthritis: colocalisation of the perinuclear factor and profilaggrin. Ann Rheum Dis 1991;**50**:611–618.
9. Girbal-Neuhauser E, et al. The epitopes targeted by the rheumatoid arthritis-associated antifilaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deimination of arginine residues. J Immunol 1999;**162**: 585–594.
10. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;**101**:273–281.
11. Schellekens GA, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000;**43**:155–163.
12. Baeten D, et al. Specific presence of intracellular citrullinated proteins in rheumatoid arthritis synovium: relevance to antifilaggrin autoantibodies. Arthritis Rheum 2001;**44**: 2255–2262.
13. van Boekel MA, Vossenaar ER, van den Hoogen FH, van Venrooij WJ. Autoantibody systems in rheumatoid arthritis: specificity, sensitivity and diagnostic value. Arthritis Res 2002;**4**:87–93.
14. van Venrooij WJ, Zendman AJ. Anti-CCP2 antibodies: an overview and perspective of the diagnostic abilities of this serological marker for early rheumatoid arthritis. Clin Rev Allergy Immunol 2008;**34**: 36–39.
15. Nielen MM, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;**50**:380–386.
16. Kroot EJ, et al. The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. Arthritis Rheum 2000;**43**:1831–1835.
17. Rantapaa-Dahlqvist S, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;**48**:2741–2749.
18. Vencovsky J, et al. Autoantibodies can be prognostic markers of an erosive disease in early rheumatoid arthritis. Ann Rheum Dis 2003;**62**:427–430.
19. van Gaalen FA, et al. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. Arthritis Rheum 2004;**50**: 709–715.
20. Inagaki M, Takahara H, Nishi Y, Sugawara K, Sato C. $Ca2^{+}$-dependent deimination-induced disassembly of intermediate filaments involves specific modification of the amino-terminal head domain. J Biol Chem 1989;**264**:18119–18127.
21. Nakayama-Hamada M, et al. Comparison of enzymatic properties between hPADI2 and hPADI4. Biochem Biophys Res Commun 2005;**327**:192–200.
22. Orlov SN, Grygorczyk R, Kotelevtsev SV. Do we know the absolute values of intracellular free calcium concentration? Cell Calcium 2003;**34**:511–515.
23. Vossenaar ER, Zendman AJ, van Venrooij WJ, Pruijn GJ. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. BioEssays 2003;**25**: 1106–1118.
24. De Rycke L, et al. Synovial intracellular citrullinated proteins colocalizing with peptidyl arginine deiminase as pathophysiologically relevant antigenic determinants of rheumatoid arthritis-specific humoral autoimmunity. Arthritis Rheum 2005;**52**: 2323–2330.
25. Chang X, et al. Localization of peptidylarginine deiminase 4 (PADI4) and citrullinated protein in synovial tissue of rheumatoid arthritis. Rheumatology (Oxford) 2005;**44**: 40–50.
26. Foulquier C, et al. Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation. Arthritis Rheum 2007;**56**: 3541–3553.
27. Vossenaar ER, et al. Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages. Ann Rheum Dis 2004;**63**:373–381.
28. Kinloch A, et al. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. Arthritis Rheum 2008;**58**: 2287–2295.
29. Borders CL Jr, et al. A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens. Protein Sci 1994;**3**:541–548.
30. Tarcsa E, Marekov LN, Mei G, Melino G, Lee SC, Steinert PM. Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. J Biol Chem 1996;**271**:30709–30716.
31. Gyorgy B, Toth E, Tarcsa E, Falus A, Buzas EI. Citrullination: a posttranslational modification in health and disease. Int J Biochem Cell Biol 2006;**38**:1662–1677.
32. Kizawa K, Takahara H, Troxler H, Kleinert P, Mochida U, Heizmann CW. Specific citrullination causes assembly of a globular S100A3 homotetramer: a putative Ca^{2+} modulator matures human hair cuticle. J Biol Chem 2008;**283**:5004–5013.
33. Harauz G, Musse AA. A tale of two citrullines–structural and functional aspects of myelin basic protein deimination in health and disease. Neurochem Res 2007;**32**:137–158.
34. Pritzker LB, Joshi S, Gowan JJ, Harauz G, Moscarello MA. Deimination of myelin basic protein. 1. Effect of deimination of arginyl residues of myelin basic protein on its structure and susceptibility to digestion by cathepsin D. Biochemistry 2000;**39**: 5374–5381.
35. Steinert PM, Idler WW. Postsynthetic modifications of mammalian epidermal alpha-keratin. Biochemistry 1979;**18**:5664–5669.
36. Hagiwara T, Nakashima K, Hirano H, Senshu T, Yamada M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. Biochem Biophys Res Commun 2002;**290**:979–983.

37. Nakashima K, Hagiwara T, Yamada M. Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. J Biol Chem 2002;**277**: 49562–49568.
38. Wang Y, et al. Human PAD4 regulates histone arginine methylation levels via demethylimination. Science 2004;**306**: 279–283.
39. Cuthbert GL, et al. Histone deimination antagonizes arginine methylation. Cell 2004;**118**:545–553.
40. Hidaka Y, Hagiwara T, Yamada M. Methylation of the guanidino group of arginine residues prevents citrullination by peptidylarginine deiminase IV. FEBS Lett 2005;**579**:4088–4092.
41. Raijmakers R, et al. Methylation of arginine residues interferes with citrullination by peptidylarginine deiminases in vitro. J Mol Biol 2007;**367**:1118–1129.
42. Thompson PR, Fast W. Histone citrullination by protein arginine deiminase: is arginine methylation a green light or a roadblock? ACS Chem Biol 2006;**1**:433–441.
43. Yao H, et al. Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis. J Biol Chem 2008; **283**:20060–20068.
44. Li P, et al. Regulation of p53 target gene expression by peptidylarginine deiminase 4. Mol Cell Biol 2008;**28**:4745–4758.
45. Neeli I, Khan SN, Radic M. Histone deimination as a response to inflammatory stimuli in neutrophils. J Immunol 2008;**180**:1895–1902.
46. Wang Y, et al. Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. J Cell Biol 2009;**184**:205–213.
47. Proost P, et al. Citrullination of CXCL8 by peptidylarginine deiminase alters receptor usage, prevents proteolysis, and dampens tissue inflammation. J Exp Med 2008;**205**: 2085–2097.
48. Loos T, et al. Citrullination of CXCL10 and CXCL11 by peptidylarginine deiminase: a naturally occurring posttranslational modification of chemokines and new dimension of immunoregulation. Blood 2008;**112**: 2648–2656.
49. Okumura N, Haneishi A, Terasawa F. Citrullinated fibrinogen shows defects in FPA and FPB release and fibrin polymerization catalyzed by thrombin. Clin Chim Acta 2009;**401**:119–123.
50. Nakayama-Hamada M, Suzuki A, Furukawa H, Yamada R, Yamamoto K. Citrullinated fibrinogen inhibits thrombin-catalysed fibrin polymerization. J Biochem 2008;**144**: 393–398.
51. Weinberg JB, Pippen AM, Greenberg CS. Extravascular fibrin formation and dissolution in synovial tissue of patients with osteoarthritis and rheumatoid arthritis. Arthritis Rheum 1991;**34**:996–1005.
52. Lundberg K, et al. Citrullinated proteins have increased immunogenicity and arthritogenicity and their presence in arthritic joints correlates with disease severity. Arthritis Res Ther 2005;**7**:R458–R467.
53. Vossenaar ER, Smeets TJ, Kraan MC, Raats JM, van Venrooij WJ, Tak PP. The presence of citrullinated proteins is not specific for rheumatoid synovial tissue. Arthritis Rheum 2004;**50**:3485–3494.
54. Makrygiannakis D, et al. Citrullination is an inflammation-dependent process. Ann Rheum Dis 2006;**65**:1219–1222.
55. Chapuy-Regaud S, et al. Fibrin deimination in synovial tissue is not specific for rheumatoid arthritis but commonly occurs during synovitides. J Immunol 2005;**174**: 5057–5064.
56. Moscarello MA, Wood DD, Ackerley C, Boulias C. Myelin in multiple sclerosis is developmentally immature. J Clin Invest 1994;**94**:146–154.
57. Bhattacharya SK, Crabb JS, Bonilha VL, Gu X, Takahara H, Crabb JW. Proteomics implicates peptidyl arginine deiminase 2 and optic nerve citrullination in glaucoma pathogenesis. Invest Ophthalmol Vis Sci 2006;**47**:2508–2514.
58. Ishigami A, et al. Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. J Neurosci Res 2005;**80**: 120–128.
59. Ishida-Yamamoto A, Senshu T, Takahashi H, Akiyama K, Nomura K, Iizuka H. Decreased deiminated keratin K1 in psoriatic hyperproliferative epidermis. J Invest Dermatol 2000;**114**:701–705.
60. Hill JA, et al. Arthritis induced by post-translationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. J Exp Med 2008;**205**:967–979.
61. Cantaert T, Coucke P, De Rycke L, Veys EM, De Keyser F, Baeten D. Functional haplotypes of PADI4: relevance for rheumatoid arthritis specific synovial intracellular citrullinated proteins and anticitrullinated protein antibodies. Ann Rheum Dis 2005;**64**:1316–1320.
62. Smeets TJ, Vossenaar ER, van Venrooij WJ, Tak PP. Is expression of intracellular citrullinated proteins in synovial tissue specific for rheumatoid arthritis? Comment on the article by Baeten et al Arthritis Rheum 2002;**46**:2824–2826. author reply 2826-2827.
63. Masson-Bessiere C, et al. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. J Immunol 2001;**166**:4177–4184.
64. Vossenaar ER, et al. Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin. Arthritis Res Ther 2004;**6**: R142–R150.
65. Burkhardt H, Sehnert B, Bockermann R, Engstrom A, Kalden JR, Holmdahl R. Humoral immune response to citrullinated collagen type II determinants in early rheumatoid arthritis. Eur J Immunol 2005;**35**:1643–1652.
66. Kinloch A, et al. Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis. Arthritis Res Ther 2005;**7**:R1421–R1429.
67. Matsuo K, et al. Identification of novel citrullinated autoantigens of synovium in rheumatoid arthritis using a proteomic approach. Arthritis Res Ther 2006;**8**:R175.
68. Tilleman K, et al. Chronically inflamed synovium from spondyloarthropathy and rheumatoid arthritis investigated by protein expression profiling followed by tandem mass spectrometry. Proteomics 2005;**5**: 2247–2257.
69. Vossenaar ER, et al. Citrullination of synovial proteins in murine models of rheumatoid arthritis. Arthritis Rheum 2003;**48**:2489–2500.
70. Vander Cruyssen B, et al. Diagnostic value of anti-human citrullinated fibrinogen ELISA and comparison with four other anti-citrullinated protein assays. Arthritis Res Ther 2006;**8**:R122.
71. Snir O, et al. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. Ann Rheum Dis 2009;**68**:736–743.
72. Sebbag M, et al. Epitopes of human fibrin recognized by the rheumatoid arthritis-specific autoantibodies to citrullinated proteins. Eur J Immunol 2006;**36**:2250–2263.
73. Clavel C, et al. Induction of macrophage secretion of tumor necrosis factor alpha through Fcgamma receptor IIa engagement by rheumatoid arthritis-specific autoantibodies to citrullinated proteins complexed with fibrinogen. Arthritis Rheum 2008; **58**:678–688.
74. Zhao X, et al. Circulating immune complexes contain citrullinated fibrinogen in rheumatoid arthritis. Arthritis Res Ther 2008;**10**:R94.
75. van Venrooij WJ, Pruijn GJ. An important step towards completing the rheumatoid arthritis cycle. Arthritis Res Ther 2008; **10**:117.

76. Hida S, Miura NN, Adachi Y, Ohno N. Influence of arginine deimination on antigenicity of fibrinogen. J Autoimmun 2004;**23**:141–150.
77. Rubin B, Sonderstrup G. Citrullination of self-proteins and autoimmunity. Scand J Immunol 2004;**60**:112–120.
78. Kuhn KA, et al. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. J Clin Invest 2006;**116**:961–973.
79. Despres N, Boire G, Lopez-Longo FJ, Menard HA. The Sa system: a novel antigen-antibody system specific for rheumatoid arthritis. J Rheumatol 1994;**21**:1027–1033.
80. Goldbach-Mansky R, et al. Rheumatoid arthritis associated autoantibodies in patients with synovitis of recent onset. Arthritis Res 2000;**2**:236–243.
81. Menard HA, Lapointe E, Rochdi MD, Zhou ZJ. Insights into rheumatoid arthritis derived from the Sa immune system. Arthritis Res 2000;**2**:429–432.
82. Asaga H, Yamada M, Senshu T. Selective deimination of vimentin in calcium ionophore-induced apoptosis of mouse peritoneal macrophages. Biochem Biophys Res Commun 1998;**243**:641–646.
83. Tilleman K, Van Steendam K, Cantaert T, De Keyser F, Elewaut D, Deforce D. Synovial detection and autoantibody reactivity of processed citrullinated isoforms of vimentin in inflammatory arthritides. Rheumatology (Oxford) 2008;**47**:597–604.
84. Bang H, et al. Mutation and citrullination modifies vimentin to a novel autoantigen for rheumatoid arthritis. Arthritis Rheum 2007;**56**:2503–2511.
85. Morishima N. Changes in nuclear morphology during apoptosis correlate with vimentin cleavage by different caspases located either upstream or downstream of Bcl-2 action. Genes Cells 1999;**4**:401–414.
86. Byun Y, Chen F, Chang R, Trivedi M, Green KJ, Cryns VL. Caspase cleavage of vimentin disrupts intermediate filaments and promotes apoptosis. Cell Death Differ 2001;**8**:443–450.
87. Mathsson L, et al. Antibodies against citrullinated vimentin in rheumatoid arthritis: higher sensitivity and extended prognostic value concerning future radiographic progression as compared with antibodies against cyclic citrullinated peptides. Arthritis Rheum 2008;**58**:36–45.
88. Ursum J, et al. Antibodies to mutated citrullinated vimentin and disease activity score in early arthritis: a cohort study. Arthritis Res Ther 2008;**10**:R12.
89. Luime JJ, Colin EM, Hazes JM, Lubberts E. Does anti-MCV has additional value as serological marker in the diagnostic and prognostic work-up of patients with rheumatoid arthritis? A systematic review. Ann Rheum Dis 2009;in press.
90. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. J Immunol 2003;**171**:538–541.
91. Nishida Y, Shibata K, Yamasaki M, Sato Y, Abe M. A possible role of vimentin on the cell surface for the activation of latent transforming growth factor-beta. FEBS Lett 2009;**583**:308–312.
92. Mor-Vaknin N, Punturieri A, Sitwala K, Markovitz DM. Vimentin is secreted by activated macrophages. Nat Cell Biol 2003;**5**:59–63.
93. Xu B, deWaal RM, Mor-Vaknin N, Hibbard C, Markovitz DM, Kahn ML. The endothelial cell-specific antibody PAL-E identifies a secreted form of vimentin in the blood vasculature. Mol Cell Biol 2004;**24**:9198–9206.
94. Boilard E, Bourgoin SG, Bernatchez C, Surette ME. Identification of an autoantigen on the surface of apoptotic human T cells as a new protein interacting with inflammatory group IIA phospholipase A2. Blood 2003;**102**:2901–2909.
95. Moisan E, Girard D. Cell surface expression of intermediate filament proteins vimentin and lamin B1 in human neutrophil spontaneous apoptosis. J Leukoc Biol 2006;**79**:489–498.
96. Podor TJ, et al. Vimentin exposed on activated platelets and platelet microparticles localizes vitronectin and plasminogen activator inhibitor complexes on their surface. J Biol Chem 2002;**277**:7529–7539.
97. Bryant AE, Bayer CR, Huntington JD, Stevens DL. Group A streptococcal myonecrosis: increased vimentin expression after skeletal-muscle injury mediates the binding of *Streptococcus pyogenes*. J Infect Dis 2006;**193**:1685–1692.
98. Murli S, Watson RO, Galan JE. Role of tyrosine kinases and the tyrosine phosphatase SptP in the interaction of Salmonella with host cells. Cell Microbiol 2001;**3**:795–810.
99. Zou Y, He L, Huang SH. Identification of a surface protein on human brain microvascular endothelial cells as vimentin interacting with *Escherichia coli* invasion protein IbeA. Biochem Biophys Res Commun 2006;**351**:625–630.
100. Choi EK, Gatenby PA, McGill NW, Bateman JF, Cole WG, York JR. Autoantibodies to type II collagen: occurrence in rheumatoid arthritis, other arthritides, autoimmune connective tissue diseases, and chronic inflammatory syndromes. Ann Rheum Dis 1988;**47**:313–322.
101. Cook AD, Rowley MJ, Mackay IR, Gough A, Emery P. Antibodies to type II collagen in early rheumatoid arthritis. Correlation with disease progression. Arthritis Rheum 1996;**39**:1720–1727.
102. Cook AD, et al. Antibodies to type II collagen and HLA disease susceptibility markers in rheumatoid arthritis. Arthritis Rheum 1999;**42**:2569–2576.
103. Cook AD, Mackay IR, Cicuttini FM, Rowley MJ. IgG subclasses of antibodies to type II collagen in rheumatoid arthritis differ from those in systemic lupus erythematosus and other connective tissue diseases. J Rheumatol 1997;**24**:2090–2096.
104. Uysal H, et al. Structure and pathogenicity of antibodies specific for citrullinated collagen type II in experimental arthritis. J Exp Med 2009;**206**:449–462.
105. Saulot V, et al. Presence of autoantibodies to the glycolytic enzyme alpha-enolase in sera from patients with early rheumatoid arthritis. Arthritis Rheum 2002;**46**:1196–1201.
106. Lundberg K, et al. Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. Arthritis Rheum 2008;**58**:3009–3019.
107. Ioan-Facsinay A, et al. Marked differences in fine specificity and isotype usage of the anti-citrullinated protein antibody in health and disease. Arthritis Rheum 2008;**58**:3000–3008.
108. Terrier B, et al. Anti-fibroblast antibodies from systemic sclerosis patients bind to α-enolase and are associated with interstitial lung disease. Ann Rheum Dis 2009;in press.
109. Cappello P, et al. An integrated humoral and cellular response is elicited in pancreatic cancer by alpha-enolase, a novel pancreatic ductal adenocarcinoma-associated antigen. Int J Cancer 2009;**125**:639–648.
110. Pancholi V. Multifunctional alpha-enolase: its role in diseases. Cell Mol Life Sci 2001;**58**:902–920.
111. Pratesi F, Moscato S, Sabbatini A, Chimenti D, Bombardieri S, Migliorini P. Autoantibodies specific for alpha-enolase in systemic autoimmune disorders. J Rheumatol 2000;**27**:109–115.
112. Terrier B, Degand N, Guilpain P, Servettaz A, Guillevin L, Mouthon L. Alpha-enolase: a target of antibodies in infectious and autoimmune diseases. Autoimmun Rev 2007;**6**:176–182.
113. Wegner N, Wait R, Venables PJ. Evolutionarily conserved antigens in autoimmune disease: implications for an infective aetiology. Int J Biochem Cell Biol 2009;**41**:390–397.

114. Miles LA, Dahlberg CM, Plescia J, Felez J, Kato K, Plow EF. Role of cell-surface lysines in plasminogen binding to cells: identification of alpha-enolase as a candidate plasminogen receptor. Biochemistry 1991;**30**:1682–1691.
115. Redlitz A, Fowler BJ, Plow EF, Miles LA. The role of an enolase-related molecule in plasminogen binding to cells. Eur J Biochem 1995;**227**:407–415.
116. Wygrecka M, et al. Enolase-1 promotes plasminogen-mediated recruitment of monocytes to the acutely inflamed lung. Blood 2009;**113**:5588–5598.
117. Lopez-Alemany R, et al. Inhibition of cell surface mediated plasminogen activation by a monoclonal antibody against alpha-Enolase. Am J Hematol 2003;**72**:234–242.
118. Whiting GC, Evans JT, Patel S, Gillespie SH. Purification of native {alpha}-enolase from *Streptococcus pneumoniae* that binds plasminogen and is immunogenic. J Med Microbiol 2002;**51**:837–843.
119. Pancholi V, Fischetti VA. alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic Streptococci. J Biol Chem 1998;**273**:14503–14515.
120. Bergmann S, Rohde M, Chhatwal GS, Hammerschmidt S. alpha-Enolase of *Streptococcus pneumoniae* is a plasmin(ogen)-binding protein displayed on the bacterial cell surface. Mol Microbiol 2001;**40**:1273–1287.
121. Bergmann S, et al. Identification of a novel plasmin(ogen)-binding motif in surface displayed alpha-enolase of *Streptococcus pneumoniae*. Mol Microbiol 2003;**49**:411–423.
122. Ge J, Catt DM, Gregory RL. Streptococcus mutans surface alpha-enolase binds salivary mucin MG2 and human plasminogen. Infect Immun 2004;**72**:6748–6752.
123. Kolberg J, et al. *Streptococcus pneumoniae* enolase is important for plasminogen binding despite low abundance of enolase protein on the bacterial cell surface. Microbiology 2006;**152**:1307–1317.
124. Yavlovich A, Rechnitzer H, Rottem S. Alpha-enolase resides on the cell surface of *Mycoplasma fermentans* and binds plasminogen. Infect Immun 2007;**75**:5716–5719.
125. Agarwal S, Kulshreshtha P, Bambah Mukku D, Bhatnagar R. alpha-Enolase binds to human plasminogen on the surface of *Bacillus anthracis*. Biochim Biophys Acta 2008;**1784**:986–994.
126. Fox D, Smulian AG. Plasminogen-binding activity of enolase in the opportunistic pathogen *Pneumocystis carinii*. Med Mycol 2001;**39**:495–507.
127. Mundodi V, Kucknoor AS, Alderete JF. Immunogenic and plasminogen-binding surface-associated alpha-enolase of *Trichomonas vaginalis*. Infect Immun 2008;**76**:523–531.
128. Vanegas G, Quiñones W, Carrasco-López C, Concepción J, Albericio F, Avilán L. Enolase as a plasminogen binding protein in *Leishmania mexicana*. Parasitology Res 2007;**101**:1511–1516.
129. Jolodar A, Fischer P, Bergmann S, Büttner DW, Hammerschmidt S, Brattig NW. Molecular cloning of an [alpha]-enolase from the human filarial parasite *Onchocerca volvulus* that binds human plasminogen. Biochim Biophys Acta 2003;**1627**:111–120.
130. Sousa LP, et al. Plasminogen/plasmin regulates alpha-enolase expression through the MEK/ERK pathway. Biochem Biophys Res Commun 2005;**337**:1065–1071.
131. Kang HJ, Jung SK, Kim SJ, Chung SJ. Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme. Acta Crystallogr D Biol Crystallogr 2008;**64**:651–657.
132. Yang YH, Carmeliet P, Hamilton JA. Tissue-type plasminogen activator deficiency exacerbates arthritis. J Immunol 2001;**167**:1047–1052.
133. Goeb V, et al. Candidate autoantigens identified by mass spectrometry in early rheumatoid arthritis are chaperones and citrullinated glycolytic enzymes. Arthritis Res Ther 2009;**11**:R38.
134. Okazaki Y, et al. Identification of citrullinated eukaryotic translation initiation factor 4G1 as novel autoantigen in rheumatoid arthritis. Biochem Biophys Res Commun 2006;**341**:94–100.
135. Boire G, et al. Anti-Sa antibodies and antibodies against cyclic citrullinated peptide are not equivalent as predictors of severe outcomes in patients with recent-onset polyarthritis. Arthritis Res Ther 2005;**7**:R592–R603.
136. Gregersen PK, Olsson LM. Recent advances in the genetics of autoimmune disease. Annu Rev Immunol 2009;**27**:363–391.
137. McHugh NJ, Owen P, Cox B, Dunphy J, Welsh K. MHC class II, tumour necrosis factor alpha, and lymphotoxin alpha gene haplotype associations with serological subsets of systemic lupus erythematosus. Ann Rheum Dis 2006;**65**:488–494.
138. Wedderburn LR, et al. HLA class II haplotype and autoantibody associations in children with juvenile dermatomyositis and juvenile dermatomyositis-scleroderma overlap. Rheumatology (Oxford) 2007;**46**:1786–1791.
139. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;**30**:1205–1213.
140. Zanelli E, Breedveld FC, de Vries RR. HLA class II association with rheumatoid arthritis: facts and interpretations. Hum Immunol 2000;**61**:1254–1261.
141. Klareskog L, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;**54**:38–46.
142. van der Helm-van Mil AH, Verpoort KN, Breedveld FC, Huizinga TW, Toes RE, de Vries RR. The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum 2006;**54**:1117–1121.
143. Verpoort KN, et al. Fine specificity of the anti-citrullinated protein antibody response is influenced by the shared epitope alleles. Arthritis Rheum 2007;**56**:3949–3952.
144. Mahdi H, et al. Specific interaction between genotype, smoking and autoimmunity to citrullinated a-enolase in the etiology of rheumatoid arthritis. Nat Genet 2009; in press.
145. Plenge RM, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis–a genomewide study. N Engl J Med 2007;**357**:1199–1209.
146. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;**447**:661–678.
147. Begovich AB, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;**75**:330–337.
148. Kallberg H, et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 2007;**80**:867–875.
149. Lucarini N, Verrotti A, Napolioni V, Bosco G, Curatolo P. Genetic polymorphisms and idiopathic generalized epilepsies. Pediatr Neurol 2007;**37**:157–164.
150. Peng J, Dong W, Chen L, Zou T, Qi Y, Liu Y. Brd2 is a TBP-associated protein and recruits TBP into E2F-1 transcriptional complex in response to serum stimulation. Mol Cell Biochem 2007;**294**:45–54.
151. Lenburg ME, Sinha A, Faller DV, Denis GV. Tumor-specific and proliferation-specific gene expression typifies murine transgenic B cell lymphomagenesis. J Biol Chem 2007;**282**:4803–4811.
152. Gyuris A, et al. The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis. Biochim Biophys Acta 2009;**1789**:413–421.

153. Vessey MP, Villard-Mackintosh L, Yeates D. Oral contraceptives, cigarette smoking and other factors in relation to arthritis. Contraception 1987;**35**:457–464.
154. Silman AJ, Newman J, MacGregor AJ. Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins. Arthritis Rheum 1996;**39**:732–735.
155. Heliovaara M, Aho K, Aromaa A, Knekt P, Reunanen A. Smoking and risk of rheumatoid arthritis. J Rheumatol 1993;**20**:1830–1835.
156. Uhlig T, Hagen KB, Kvien TK. Current tobacco smoking, formal education, and the risk of rheumatoid arthritis. J Rheumatol 1999;**26**:47–54.
157. Masdottir B, Jonsson T, Manfredsdottir V, Vikingsson A, Brekkan A, Valdimarsson H. Smoking, rheumatoid factor isotypes and severity of rheumatoid arthritis. Rheumatology (Oxford) 2000;**39**:1202–1205.
158. Criswell LA, et al. Cigarette smoking and the risk of rheumatoid arthritis among postmenopausal women: results from the Iowa Women's Health Study. Am J Med 2002;**112**:465–471.
159. Mattey DL, et al. Relationship among the HLA-DRB1 shared epitope, smoking, and rheumatoid factor production in rheumatoid arthritis. Arthritis Rheum 2002;**47**:403–407.
160. Stolt P, et al. Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann Rheum Dis 2003;**62**:835–841.
161. Pedersen M, et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. Arthritis Res Ther 2006;**8**:R133.
162. van der Helm-van Mil AH, Verpoort KN, le Cessie S, Huizinga TW, de Vries RR, Toes RE. The HLA-DRB1 shared epitope alleles differ in the interaction with smoking and predisposition to antibodies to cyclic citrullinated peptide. Arthritis Rheum 2007;**56**:425–432.
163. Michou L, et al. Associations between genetic factors, tobacco smoking and autoantibodies in familial and sporadic rheumatoid arthritis. Ann Rheum Dis 2008;**67**:466–470.
164. Linn-Rasker SP, et al. Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. Ann Rheum Dis 2006;**65**:366–371.
165. Pryor WA, Stone K. Oxidants in cigarette smoke. Radicals, hydrogen peroxide, peroxynitrate, and peroxynitrite. Ann NY Acad Sci 1993;**686**:12–27. discussion-28.
166. Costenbader KH, Karlson EW. Cigarette smoking and autoimmune disease: what can we learn from epidemiology? Lupus 2006;**15**:737–745.
167. Makrygiannakis D, et al. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. Ann Rheum Dis 2008;**67**:1488–1492.
168. Verpoort KN, et al. Isotype distribution of anti-cyclic citrullinated peptide antibodies in undifferentiated arthritis and rheumatoid arthritis reflects an ongoing immune response. Arthritis Rheum 2006;**54**:3799–3808.
169. Svard A, Kastbom A, Reckner-Olsson A, Skogh T. Presence and utility of IgA-class antibodies to cyclic citrullinated peptides in early rheumatoid arthritis: the Swedish TIRA project. Arthritis Res Ther 2008;**10**:R75.
170. Lee HS, et al. Interaction between smoking, the shared epitope, and anti-cyclic citrullinated peptide: a mixed picture in three large North American rheumatoid arthritis cohorts. Arthritis Rheum 2007;**56**:1745–1753.
171. Carty SM, Snowden N, Silman AJ. Should infection still be considered as the most likely triggering factor for rheumatoid arthritis? Ann Rheum Dis 2004;**63** (Suppl):ii46–ii49.
172. Zhang X, Pacheco-Tena C, Inman RD. Microbe hunting in the joints. Arthritis Rheum 2003;**49**:479–482.
173. McGraw WT, Potempa J, Farley D, Travis J. Purification, characterization, and sequence analysis of a potential virulence factor from *Porphyromonas gingivalis*, peptidylarginine deiminase. Infect Immun 1999;**67**:3248–3256.
174. Dzink JL, Socransky SS, Haffajee AD. The predominant cultivable microbiota of active and inactive lesions of destructive periodontal diseases. J Clin Periodontol 1988;**15**:316–323.
175. Mikuls TR, et al. Antibody responses to *Porphyromonas gingivalis* (*P. gingivalis*) in subjects with rheumatoid arthritis and periodontitis. Int Immunopharmacol 2009;**9**:38–42.
176. Mercado FB, Marshall RI, Bartold PM. Inter-relationships between rheumatoid arthritis and periodontal disease. A review. J Clin Periodontol 2003;**30**:761–772.
177. Rosenstein ED, Greenwald RA, Kushner LJ, Weissmann G. Hypothesis: the humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis. Inflammation 2004;**28**:311–318.
178. de Pablo P, Chapple ILC, Buckley CD, Dietrich T. Periodontitis in systemic rheumatic diseases. Nat Rev Rheumatol 2009;**5**:218–224.
179. Ogrendik M, Kokino S, Ozdemir F, Bird PS, Hamlet S. Serum antibodies to oral anaerobic bacteria in patients with rheumatoid arthritis. MedGenMed 2005;**7**:2.
180. Mercado F, Marshall RI, Klestov AC, Bartold PM. Is there a relationship between rheumatoid arthritis and periodontal disease? J Clin Periodontol 2000;**27**:267–272.
181. Ortiz P, et al. Periodontal therapy reduces the severity of active rheumatoid arthritis in patients treated with or without tumor necrosis factor inhibitors. J Periodontol 2009;**80**:535–540.
182. Klareskog L, Ronnelid J, Lundberg K, Padyukov L, Alfredsson L. Immunity to citrullinated proteins in rheumatoid arthritis. Annu Rev Immunol 2008;**26**:651–675.
183. Paster BJ, et al. Bacterial diversity in human subgingival plaque. J Bacteriol 2001;**183**:3770–3783.
184. Lamont RJ, Jenkinson HF. Life below the gum line: pathogenic mechanisms of *Porphyromonas gingivalis*. Microbiol Mol Biol Rev 1998;**62**:1244–1263.
185. Colombo AV, da Silva CM, Haffajee A, Colombo AP. Identification of intracellular oral species within human crevicular epithelial cells from subjects with chronic periodontitis by fluorescence in situ hybridization. J Periodontal Res 2007;**42**:236–243.
186. Colombo AV, Silva CM, Haffajee A, Colombo AP. Identification of oral bacteria associated with crevicular epithelial cells from chronic periodontitis lesions. J Med Microbiol 2006;**55**:609–615.
187. Rudney JD, Chen R. The vital status of human buccal epithelial cells and the bacteria associated with them. Arch Oral Biol 2006;**51**:291–298.
188. Chen Z, Potempa J, Polanowski A, Wikstrom M, Travis J. Purification and characterization of a 50-kDa cysteine proteinase (gingipain) from *Porphyromonas gingivalis*. J Biol Chem 1992;**267**:18896–18901.
189. Travis J, Pike R, Imamura T, Potempa J. *Porphyromonas gingivalis* proteinases as virulence factors in the development of periodontitis. J Periodontal Res 1997;**32**:120–125.
190. Potempa J, Sroka A, Imamura T, Travis J. Gingipains, the major cysteine proteinases and virulence factors of *Porphyromonas gingivalis*: structure, function and assembly of multidomain protein complexes. Curr Protein Pept Sci 2003;**4**:397–407.
191. Imamura T, Travis J, Potempa J. The biphasic virulence activities of gingipains: activation and inactivation of host proteins. Curr Protein Pept Sci 2003;**4**:443–450.

192. Holt SC, Ebersole JL. *Porphyromonas gingivalis*, *Treponema denticola*, and *Tannerella forsythia*: the ''red complex'', a prototype polybacterial pathogenic consortium in periodontitis. Periodontol 2000 2005;**38**:72–122.
193. Borrell LN, Burt BA, Taylor GW. Prevalence and Trends in Periodontitis in the USA: from the NHANES III to the NHANES, 1988 to 2000. J Dental Res 2005;**84**:924–930.
194. Liao F, Li Z, Wang Y, Shi B, Gong Z, Cheng X. *Porphyromonas gingivalis* may play an important role in the pathogenesis of periodontitis-associated rheumatoid arthritis. Med Hypotheses 2009;**72**:732–735.
195. de Pablo P, Dietrich T, McAlindon TE. Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population. J Rheumatol 2008;**35**:70–76.
196. Havemose-Poulsen A, et al. Periodontal and hematological characteristics associated with aggressive periodontitis, juvenile idiopathic arthritis, and rheumatoid arthritis. J Periodontol 2006;**77**:280–288.
197. Izutsu KT, et al. Involvement of calcium in interactions between gingival epithelial cells and *Porphyromonas gingivalis*. FEMS Microbiol Lett 1996;**144**:145–150.
198. Lourbakos A, et al. Cleavage and activation of proteinase-activated receptor-2 on human neutrophils by gingipain-R from *Porphyromonas gingivalis*. FEBS Lett 1998;**435**: 45–48.
199. Belton CM, Goodwin PC, Fatherazi S, Schubert MM, Lamont RJ, Izutsu KT. Calcium oscillations in gingival epithelial cells infected with *Porphyromonas gingivalis*. Microbes Infect 2004;**6**:440–447.
200. Lourbakos A, et al. Arginine-specific protease from *Porphyromonas gingivalis* activates protease-activated receptors on human oral epithelial cells and induces interleukin-6 secretion. Infect Immun 2001;**69**:5121–5130.
201. Lourbakos A, et al. Activation of protease-activated receptors by gingipains from *Porphyromonas gingivalis* leads to platelet aggregation: a new trait in microbial pathogenicity. Blood 2001;**97**:3790–3797.
202. Holzhausen M, et al. Protease-activated receptor-2 activation: a major role in the pathogenesis of *Porphyromonas gingivalis* infection. Am J Pathol 2006;**168**:1189–1199.
203. Michalowicz BS, et al. Evidence of a substantial genetic basis for risk of adult periodontitis. J Periodontol 2000;**71**:1699–1707.
204. Kinane DF, Shiba H, Hart TC. The genetic basis of periodontitis. Periodontol 2000 2005;**39**:91–117.
205. Yoshie H, Kobayashi T, Tai H, Galicia JC. The role of genetic polymorphisms in periodontitis. Periodontol 2000 2007;**43**: 102–132.
206. Bonfil JJ, et al. A ''case control'' study on the role of HLA DR4 in severe periodontitis and rapidly progressive periodontitis. J Clin Periodontol 1999;**26**:77–84.
207. Katz J, Goultschin J, Benoliel R, Brautbar C. Human leukocyte antigen (HLA) DR4. Positive association with rapidly progressing periodontitis. J Periodontol 1987;**58**:607–610.
208. Alley CS, et al. HLA-D and T lymphocyte reactivity to specific periodontal pathogens in type 1 diabetic periodontitis. J Periodontol 1993;**64**:974–979.
209. Firatli E, et al. Association between HLA antigens and early onset periodontitis. J Clin Periodontol 1996;**23**:563–566.
210. Dyer JK, et al. HLA-D types and serum IgG responses to Capnocytophaga in diabetes and periodontitis. J Dent Res 1997;**76**: 1825–1832.
211. Stein J, Reichert S, Gautsch A, Machulla HKG. Are there HLA combinations typical supporting for or making resistant against aggressive and/or chronic periodontitis? J Periodontal Res 2003;**38**:508–517.
212. Kinane DF, Chestnutt IG. Smoking and periodontal disease. Crit Rev Oral Biol Med 2000;**11**:356–365.
213. Tomar SL, Asma S. Smoking-attributable periodontitis in the United States: findings from NHANES III. National Health and Nutrition Examination Survey. J Periodontol 2000;**71**:743–751.
214. Grossi SG, et al. Assessment of risk for periodontal disease. II. Risk indicators for alveolar bone loss. J Periodontol 1995;**66**: 23–29.
215. Machtei EE, et al. Longitudinal study of prognostic factors in established periodontitis patients. J Clin Periodontol 1997;**24**: 102–109.
216. Yamamoto Y, et al. Association between passive and active smoking evaluated by salivary cotinine and periodontitis. J Clin Periodontol 2005;**32**:1041–1046.
217. Bostrom L, Linder LE, Bergstrom J. Smoking and cervicular fluid levels of IL-6 and TNF-alpha in periodontal disease. J Clin Periodontol 1999;**26**:352–357.
218. Bostrom L, Linder LE, Bergstrom J. Clinical expression of TNF-alpha in smoking-associated periodontal disease. J Clin Periodontol 1998;**25**:767–773.
219. Bostrom L, Linder LE, Bergstrom J. Influence of smoking on the outcome of periodontal surgery. A 5-year follow-up. J Clin Periodontol 1998;**25**:194–201.
220. Barbour SE, et al. Tobacco and smoking: environmental factors that modify the host response (immune system) and have an impact on periodontal health. Crit Rev Oral Biol Med 1997;**8**: 437–460.
221. Pischon N, et al. Association among rheumatoid arthritis, oral hygiene, and periodontitis. J Periodontol 2008;**79**:979–986.
222. Hujoel PP, Drangsholt M, Spiekerman C, DeRouen TA. Periodontitis-systemic disease associations in the presence of smoking–causal or coincidental? Periodontol 2000 2002;**30**:51–60.

ARTHRITIS & RHEUMATISM
Vol. 62, No. 9, September 2010, pp 2662–2672
DOI 10.1002/art.27552

Peptidylarginine Deiminase From *Porphyromonas gingivalis* Citrullinates Human Fibrinogen and α-Enolase

Implications for Autoimmunity in Rheumatoid Arthritis

Natalia Wegner,[1] Robin Wait,[1] Aneta Sroka,[2] Sigrun Eick,[3] Ky-Anh Nguyen,[4] Karin Lundberg,[1] Andrew Kinloch,[1] Shauna Culshaw,[5] Jan Potempa,[6] and Patrick J. Venables[1]

Objective. **To investigate protein citrullination by the periodontal pathogen *Porphyromonas gingivalis* as a potential mechanism for breaking tolerance to citrullinated proteins in rheumatoid arthritis (RA).**

Methods. **The expression of endogenous citrullinated proteins was analyzed by immunoblotting of cell extracts from *P gingivalis* and 10 other oral bacteria. *P gingivalis*–knockout strains lacking the bacterial peptidylarginine deiminases (PADs) or gingipains were created to assess the role of these enzymes in citrullination. Citrullination of human fibrinogen and α-enolase by *P gingivalis* was studied by incubating live wild-type and knockout strains with the proteins and analyzing the products by immunoblotting and mass spectrometry.**

Results. **Endogenous protein citrullination was abundant in *P gingivalis* but lacking in the other oral bacteria. Deletion of the bacterial PAD gene resulted in complete abrogation of protein citrullination. Inactivation of arginine gingipains, but not lysine gingipains, led to decreased citrullination. Incubation of wild-type *P gingivalis* with fibrinogen or α-enolase caused degradation of the proteins and citrullination of the resulting peptides at carboxy-terminal arginine residues, which were identified by mass spectrometry.**

Conclusion. **Our findings demonstrate that among the oral bacterial pathogens tested, *P gingivalis* is unique in its ability to citrullinate proteins. We further show that *P gingivalis* rapidly generates citrullinated host peptides by proteolytic cleavage at Arg-X peptide bonds by arginine gingipains, followed by citrullination of carboxy-terminal arginines by bacterial PAD. Our results suggest a novel model where *P gingivalis*–mediated citrullination of bacterial and host proteins provides a molecular mechanism for generating antigens that drive the autoimmune response in RA.**

Rheumatoid arthritis (RA) is characterized by disease-specific autoimmunity to citrullinated proteins. Citrullination is a posttranslational modification of arginine residues that is mediated by the family of peptidylarginine deiminases (PADs). Citrullinated fibrin(ogen) and α-enolase are 2 of the physiologic proteins that are targeted by anti–citrullinated protein antibodies in RA (1–5). Fibrinogen is the precursor of fibrin, and autoan-

Ms. Wegner and Drs. Wait, Lundberg, and Venables' work was supported by a grant from the Arthritis Research Campaign, UK (DKCR F33018). Drs. Lundberg and Venables also received funding from the AutoCure (Curing Autoimmune Rheumatic Disease) Consortium. Dr. Culshaw's work was supported by Medical Research Scotland (168RFG). Dr. Potempa's work was supported by grants from the NIH (DE-09761) and from the Department of Scientific Research, Polish Ministry of Science and Education (1642/B/P01/2008/35). The Faculty of Biochemistry, Biophysics, and Biotechnology of the Jagiellonian University Krakow receives grants from the European Union (POIG.02.01.00-12-064/08; Molecular Biotechnology for Health).

[1]Natalia Wegner, MRes, Robin Wait, PhD, Karin Lundberg, PhD, Andrew Kinloch, PhD, Patrick J. Venables, MD: Imperial College London, London, UK; [2]Aneta Sroka, PhD: Jagiellonian University, Krakow, Poland; [3]Sigrun Eick, MD: University Hospital Jena, Jena, Germany; [4]Ky-Anh Nguyen, PhD: Westmead Centre for Oral Health and University of Sydney, Sydney, Australia; [5]Shauna Culshaw, PhD: University of Glasgow, Glasgow, UK; [6]Jan Potempa, PhD: Jagiellonian University, Krakow, Poland, and University of Louisville, Louisville, Kentucky.

Drs. Potempa and Venables contributed equally to this work.

Dr. Venables has received consulting fees for online and telephone consultations with Guidepoint Global and Medefield (less than $10,000 each), and Drs. Venables, Lundberg, and Kinloch have submitted patent applications (GB0701417.8 and PCT/GB08/000267) for citrullinated α-enolase peptide 1 for use in the diagnosis and treatment of rheumatoid arthritis, for which no royalties are received.

Address correspondence and reprint requests to Patrick J. Venables, MD, Imperial College London, Kennedy Institute of Rheumatology Division, 65 Aspenlea Road, London W6 8LH, UK. E-mail: p.venables@imperial.ac.uk.

Submitted for publication October 5, 2009; accepted in revised form April 29, 2010.

tibodies to citrullinated fibrin(ogen) are found in up to 66% of patients with RA (6). Alpha-enolase is an evolutionarily conserved, multifunctional protein (7) that is best known for its role in glucose metabolism and, more recently, as a plasminogen-binding protein on the surface of various mammalian and prokaryotic cell types (8,9). Autoantibodies to citrullinated α-enolase can be found in 40–60% of patients with RA (4–6). The pathogenicity of these autoantibodies may be mediated by the formation of immune complexes with citrullinated host proteins in the joint and by activation of downstream inflammatory pathways via complement fixation and Fcγ receptor activation (10–13).

It is not yet known which factors trigger the breakdown of tolerance to citrullinated proteins. Protein citrullination is part of healthy physiology, with citrullinated filaggrin having been identified in healthy skin (14), and is part of the inflammatory response in general (15), whereas the formation of autoantibodies to citrullinated proteins is largely restricted to RA (16). Deposits of citrullinated fibrin have been found in a variety of inflammatory joint conditions without an accompanying autoantibody response (17,18). Additional environmental and genetic risk factors are therefore likely to be required.

To date, tobacco exposure and the presence of certain alleles in the HLA–DRB1 locus with a common peptide-binding motif, collectively known as the shared epitope, have been identified as susceptibility factors for the development of autoantibodies to citrullinated proteins (19,20), α-enolase and vimentin in particular (21), but these do not explain the total risk. Additional etiologic pathways require consideration, with the periodontal pathogen *Porphyromonas gingivalis* being a prime candidate for investigation.

Periodontitis, in which *P gingivalis* is a major causative agent, is a chronic inflammatory disease of the supporting tissues of the teeth, with an estimated prevalence of 4.2% in the US population (22). *P gingivalis* can be detected in 80–90% of periodontitis patients and in 10–30% of healthy subjects (23,24). The bacterium has recently attracted interest based on epidemiologic links between RA and periodontitis (25) and the description of a novel bacterial PAD (26) (hereinafter called PPAD), suggesting a potential etiologic role of *P gingivalis* in RA through the generation of citrullinated antigens.

The pathophysiologic mechanisms of periodontitis are similar to those of RA. The condition is characterized by the resorption of the supporting bony structure around the teeth and is mediated by a variety of proinflammatory molecules, including tumor necrosis factor α, interleukin-1β, prostaglandin E_2, and matrix metalloproteinases (27). A number of studies have indicated a positive association between the prevalence of periodontitis and RA (25,28), even when adjusted for smoking, which is a major risk factor for both diseases. We have shown that RA-specific autoantibodies to citrullinated α-enolase peptide 1, the immunodominant B cell epitope of human α-enolase, cross-react with in vitro–citrullinated enolase from *P gingivalis* (5), raising the possibility of molecular mimicry between epitopes from citrullinated bacterial and human enolases. *P gingivalis* is the only prokaryote described to date that expresses a functional bacterial PAD, though its physiologic substrates are unknown, as are the molecular mechanisms of citrullination.

PPAD displays no amino acid sequence similarity to the human PAD enzymes, and a previous study indicated that it might preferentially target carboxy-terminal arginine residues (26). This is in contrast to the human enzymes, which efficiently deiminate internal arginine residues (29). Citrullination of bacterial and host proteins and peptides by *P gingivalis* PAD could therefore create new epitopes and, given the infectious context providing endogenous and exogenous danger signals, trigger a latent antibody response to citrullinated bacterial and host proteins in susceptible individuals.

In the present study, we aimed to elucidate the molecular requirements for bacterial and human protein citrullination by *P gingivalis* PAD and thus advance our understanding of the potential underlying mechanisms for the generation of citrullinated antigens and the induction of autoimmunity in RA.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *Porphyromonas gingivalis* wild-type strain (W83), *P gingivalis* clinical isolates obtained from patients with severe periodontitis (MaRL, D243, JH16, and J430), and *P gingivalis* mutants (Δ*ppad*, *ppad*+, Δ*rgp*, Δ*kgp*, and Δ*rgp*+*kgp*) were grown in Schaedler anaerobe broth (Oxoid), supplemented with 5% sheep blood, at 37°C in an anaerobic chamber (90% N_2, 5% CO_2, and 5% H_2). Erythromycin or tetracycline was used at 5 μg/ml or 1 μg/ml, respectively, on solid media. The concentrations were doubled for selective growth in liquid culture.

Other anaerobic oral bacteria (*Prevotella intermedia* H13 [clinical isolate], *Prevotella oralis* [ATCC 33269], *Capnocytophaga gingivalis* [ATCC 33624], and *Capnocytophaga ochracea* [ATCC 27872]) were grown in Schaedler anaerobe broth, supplemented with 2.5 μg/liter of vitamin K, at 37°C in an anaerobic chamber (90% N_2, 5% CO_2, and 5% H_2). *Fusobacterium nucleatum* (ATCC 10953) was grown in Schaedler anaerobe broth at 37°C in an anaerobic chamber (80% N_2,

10% CO_2, and 10% H_2). *Aggregatibacter actinomycetemcomitans* (ATCC 43718) was grown in tryptic soy broth (Sigma), supplemented with 6% yeast extract and 8% glucose, at 37°C in an atmosphere consisting of 5% CO_2. Aerobic bacteria (*Streptococcus constellatus* [ATCC 27823], *Streptococcus gordonii* [ATCC 10558], *Streptococcus sanguinis* [ATCC 10556], and *Streptococcus salivarius* [ATCC 7073]) were grown on Columbia agar plates, supplemented with 8% defibrinated sheep blood or brain/heart infusion broth.

Construction of *P gingivalis* mutant strain *ppad*. A 1-kb region 3′ to the *P gingivalis ppad* gene (GenBank accession no. 2552184; locus tag PG1424) was amplified by polymerase chain reaction (PCR; primers 5′-GCTCTAGATGGAATCCGTGAGACAATG and 5′-TAAGCATGCGATATTTGTCGGAAGGACTC) for insertion into the *Xba* I and *Sph* I sites of the pUC19 plasmid (New England BioLabs). An erythromycin resistance cassette *ermF/ermAM* from plasmid pVA2198 was amplified and inserted into the *Sma* I and *Xba* I sites of the modified pUC19 plasmid. The resultant plasmid was modified further by incorporating an amplified 1-kb region 5′ to the *ppad* gene (primers 5′-AAGAGCTCAAGCACGTAATAAGGACAATGA and 5′-TTATCCCGGGTGTTCCTGAACATATGATAAGATCT) into the *Sac* I and *Sma* I sites to create the deletional inactivation plasmid construct (pΔ*ppad*) or by incorporating the entire *ppad* gene and a 1-kb region 5′ to the gene (primers 5′-AAGAGCTCAAGCACGTAATAAGGACAATGA and 5′-TTATCCCGGGTGTCTACCTGAGGAGTATTCT) into the *Sac* I and *Sma* I sites to create the control mutant construct (p*ppad*+) to control for possible polar effects.

The correct placement and orientation of the DNA segments were confirmed by sequencing. The modified plasmid constructs were integrated into the *P gingivalis* W83 genome by a double-crossover recombination event by electroporation using standard protocols (30). Erythromycin-resistant clones were subcultured on selective plates, and genomic integration was confirmed by PCR using primers from outside of the cloned regions surrounding the *ppad* gene.

Construction of *P gingivalis* mutant strains *rgp* and *kgp*. The general procedure for construction of the Δ*rgp* and Δ*kgp* mutants has been described elsewhere (30). Homologous recombination of the $\triangle$*kgp* plasmid into the *P gingivalis* $\triangle$*rgp* mutant genome resulted in a *kgp-rgp*–deficient mutant (Δ*rgp*+*kgp*). The respective phenotypes were confirmed by enzymatic assays and Western blot analysis.

Preparation of bacterial whole-cell extracts. Bacterial cultures were grown in liquid media until the early stationary phase. Twenty milliliters of the culture was centrifuged at 10,000*g* for 15 minutes at 4°C, and the resulting bacterial pellet was resuspended in phosphate buffered saline (PBS). The optical density (OD) at 600 nm was measured and adjusted to 1.0 with PBS, and the suspension was sonicated on ice. Sodium azide (final concentration 0.02% volume/volume) was added to all samples as a preservative.

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). Protein samples were mixed with reducing 4× lithium dodecyl sulfate (LDS) sample buffer (Invitrogen), heated for 10 minutes at 70°C, and resolved on 12% NuPAGE Bis-Tris gels (Invitrogen) using MOPS running buffer. After electrophoresis, proteins were stained with the Coomassie-based stain InstantBlue (Triple Red) or transferred to nitrocellulose membranes for immunoblotting. For analysis of fibrinogen/enolase-derived peptides, 10–20% Tricine gels, 2× Tricine sample buffer, and Tricine SDS running buffer (all from Invitrogen) were used, and protein bands were visualized using a standard silver staining protocol.

Detection of citrullinated proteins by immunoblotting and dot-blotting. Citrullinated proteins were detected using an anticitrulline (modified) detection kit (Upstate/Millipore) in accordance with the manufacturer's instructions. For dot-blotting, 10 μl of sample was spotted onto an equilibrated nitrocellulose membrane (0.1-μm Protran membrane; Whatman) and allowed to dry before proceeding with the standard Western blotting protocol. Controls were performed in which the modification step or the primary antibody was omitted to control for nonspecific binding by the primary antibody to structures other than modified citrulline side chains or for binding of the secondary antibody to proteins other than the primary antibody, respectively.

Analysis of fibrinogen and α-enolase citrullination by live *P gingivalis*. *P gingivalis* was cultured as described above, and the $OD_{600\ nm}$ was measured. Bacterial cells were pelleted, washed in ice-cold PBS, and resuspended in assay buffer (10 m*M* HEPES, 150 m*M* NaCl, 1 m*M* $CaCl_2$, pH 7.5, and 10 m*M* L-cysteine) to a final $OD_{600\ nm}$ of 1.0. Purified fibrinogen was purchased from Sigma-Aldrich (catalog no. F3879), and recombinant human α-enolase was expressed in *Escherichia coli* as previously described (31). The proteins were diluted in HEPES buffer at a concentration of 0.5 mg/ml. Equal volumes of protein solution and bacterial cell suspension were mixed, and an aliquot was immediately withdrawn (corresponds to time point of 1 minute). The cultures were then incubated at 37°C on a shaking platform, and further aliquots were withdrawn after 1.5 hours, 3 hours, and 6 hours. Bacterial cells were immediately removed from all aliquots by centrifugation. The resulting supernatant was used for analysis on SDS-PAGE gels and immunoblotting as described above and for protein precipitation with 15% *meta*-phosphoric acid, leaving small peptides in solution, and subsequent analysis by high-performance liquid chromatography (HPLC). HPLC peak fractions were collected and subjected to peptide analysis by mass spectrometry.

HPLC analysis. HPLC was performed with a Shimadzu VP series chromatograph equipped with a Supelcosil LC-318 reverse-phase column measuring 25 cm × 4.6 mm (Sigma-Aldrich). Samples (100 μl) were injected and eluted with a gradient of H_2O/0.1% trifluoroacetic acid (TFA) (solution A) and 80% acetonitrile/0.08% TFA (solution B) and were monitored at 215 nm with a Shimadzu SPD-10A UV-Vis detector. Peaks were collected manually and freeze-dried prior to analysis by mass spectrometry.

In-gel digestion of proteins for mass spectrometry. Protein bands were excised with a scalpel, and in-gel digestion was performed using a robotic system (Investigator ProGest; Genomic Solutions). The bands were washed in 100 m*M* ammonium bicarbonate buffer and dehydrated in 100% acetonitrile. Cysteine residues were reduced with 10 m*M* DTT, then carboamidomethylated with 55 m*M* iodoacetamide. Digestion was performed for 6 hours at 37°C by the addition of modified porcine trypsin (10 μl at 6.5 ng/μl in 25 m*M* ammonium bicarbonate), and peptides were recovered by sequential extraction with 25 m*M* ammonium bicarbonate buffer, 5% formic acid, and acetonitrile. Extracts were pooled,

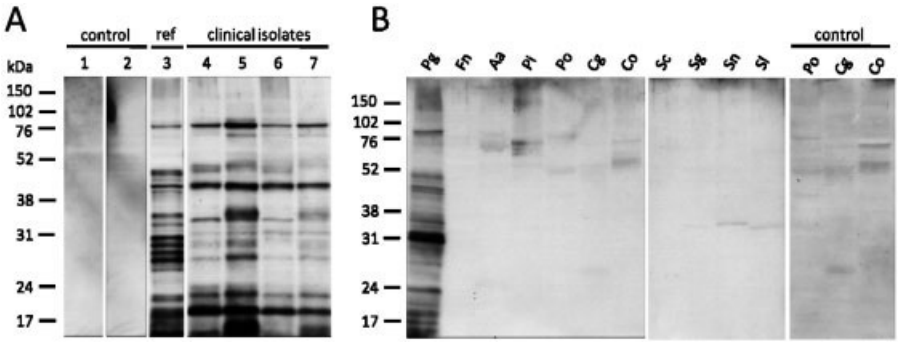


Figure 1. Expression of endogenous citrullinated proteins is ubiquitous in *Porphyromonas gingivalis*, but not in 10 other oral bacteria. **A,** Protein citrullination in total cell extracts of the *P gingivalis* wild-type reference strain W83 (lane 3) and in 4 clinical isolates (lanes 4–7, corresponding to strains MaRL, D243, JH16, and J430) was analyzed by immunoblotting with anti–modified citrulline (AMC) antibody. Controls in which the modification step (lane 1) or the secondary antibody (lane 2) had been omitted were run in parallel. Molecular mass markers are shown on the left. **B,** Total cell extracts of 10 other prominent oral bacteria were tested for endogenous protein citrullination using the AMC antibody. Background signals were confirmed to stem from nonspecific binding of the primary antibody (control). Pg = *Porphyromonas gingivalis*; Fn = *Fusobacterium nucleatum*; Aa = *Aggregatibacter actinomycetemcomitans*; Pi = *Prevotella intermedia*; Po = *Prevotella oralis*; Cg = *Capnocytophaga gingivalis*; Co = *Capnocytophaga ochracea*; Sc = *Streptococcus constellatus*; Sg = *Streptococcus gordonii*; Sn = *Streptococcus sanguinis*; Sl = *Streptococcus salivarius*.

lyophilized, and redissolved in 0.1% formic acid prior to performing mass spectrometry.

Analysis by mass spectrometry. Tandem electrospray mass spectra were recorded with a Q-Tof hybrid quadrupole/orthogonal acceleration time-of-flight spectrometer (Waters) interfaced to a CapLC chromatograph. Freeze-dried peptide samples were redissolved in 0.1% formic acid, and 6 μl was injected onto a PepMap C18 column (300 μm $\times$ 0.5 cm; LC Packings) and eluted with an acetonitrile/0.1% formic acid gradient at a flow rate of 1 μl/minute. The capillary voltage was set to 3,500V, and data-dependent product ion scans were performed on precursor ions with charge states of 2, 3, or 4 over a survey mass range of m/z 400 to 1,400. The raw spectra were smoothed, deisotoped, transformed onto a singly charged mass/charge (m/z) axis using a maximum entropy method as implemented in the peptide auto module of MassLynx (Waters), and then saved in the peaklist (pkl) format prior to database searching.

Proteins were identified by correlation of uninterpreted spectra to entries in the Swiss-Prot/TrEMBL database using the ProteinLynx Global Server (version 1.1; Waters) and a local installation of Mascot, version 2.2 (www.matrixscience-.com). The database used was a FASTA format composite constructed in-house by merging Swiss-Prot, TrEMBL, and associated splice variants (release date May 26, 2009; 8,413,758 sequences). Searches were run in error-tolerant mode, and no mass or taxomic constraints were applied. The initial enzyme specificity was set to trypsin, but subsequent searches of the *P gingivalis* digestions of fibrinogen and enolase were repeated with no enzyme specificity in order to match peptides resulting from the combination of gingipain activity with other enzymes, such as aminopeptidases and carboxypeptidases. All spectra matching citrullinated peptides were reviewed manually by interpretation of sequence-specific fragment ions to confirm the presence and location of the citrulline residue and to exclude other modifications, such as deamidation of aspartic acid, which also result in a mass increase of 1 dalton.

RESULTS

Expression of endogenous citrullinated proteins is unique for *P gingivalis*. In order to test whether *P gingivalis* citrullinates its own endogenous proteins, cultures of *P gingivalis*, comprising reference strain W83 and 4 clinical isolates from patients with periodontal disease, were grown to stationary phase, and whole-cell lysates were analyzed by immunoblotting using anti–modified citrulline (AMC) antibody. We observed strong, distinct bands, with a similar pattern of citrullinated proteins in all strains tested (Figure 1A). The gene encoding PPAD was detected in all *P gingivalis* strains tested, including the clinical isolates, as determined by PCR (data not shown). Subcellular fractionation of the wild-type strain further showed that the majority of citrullinated proteins were associated with the periplasm and the outer and inner membrane fractions (data not shown), which is a typical feature of bacterial virulence factors and antigens.

To examine whether endogenous citrullination is a unique ability of *P gingivalis* within the community of oral pathogens, we tested whole cell lysates of 10 other

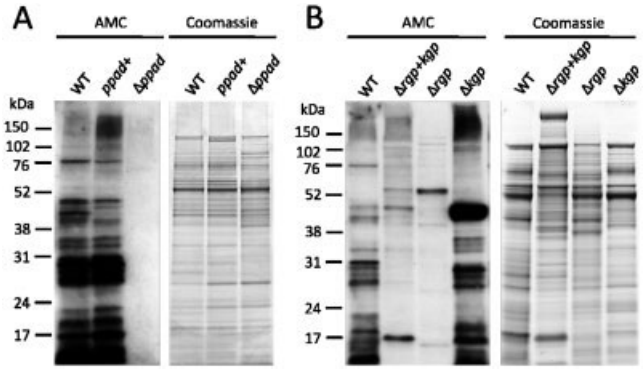


Figure 2. Citrullination in *Porphyromonas gingivalis* depends on the bacterial peptidylarginine deiminase (PPAD) and is influenced by arginine gingipain–mediated proteolytic cleavage of substrate proteins. **A,** A *P gingivalis* mutant strain lacking the PPAD gene (Δ*ppad*) was constructed, and total cell extracts were analyzed for the presence of citrullinated proteins by immunoblotting with anti–modified citrulline (AMC) antibody or by staining with the Coomassie-based stain InstantBlue. A control mutant containing the entire *ppad* gene and the antibiotic cassette (*ppad*+) was created to control for possible polar effects. The *P gingivalis* wild-type (WT) strain W83 was used as positive control. Molecular mass markers are shown on the left. **B,** *P gingivalis* mutant strains lacking arginine gingipain (Δ*rgp*), lysine gingipain (Δ*kgp*), or both (Δ*rgp*+*kgp*) proteolytic activities were analyzed for the presence of citrullinated proteins by immunoblotting with AMC antibody or by staining with the Coomassie-based stain InstantBlue.

oral organisms for the presence of citrullinated proteins. None were detected except in *P gingivalis* (Figure 1B), suggesting that functional PAD enzymes are absent from the other strains tested. Weak bands were noticeable in a number of strains, but were the result of nonspecific antibody binding (see controls in Figure 1B).

Endogenous protein citrullination is dependent on the bacterial PAD enzyme in cooperation with protein cleavage by arginine gingipains. To confirm that the observed endogenous protein citrullination in *P gingivalis* is due to the enzymatic activity of PPAD, and to rule out the possibility of a second, uncharacterized, bacterial PAD enzyme, we created a *P gingivalis* W83–knockout strain (Δ*ppad*) by replacement of the entire *ppad*-encoding DNA sequence with an antibiotic cassette. A strain in which the antibiotic cassette was inserted behind the *ppad* gene was used as a control against polar effects from genetic manipulations (*ppad*+). Immunoblotting for citrullinated proteins showed that Δ*ppad* entirely lacked endogenous citrullinated proteins, while the wild-type strain and the control strain (*ppad*+) showed a similar pattern and intensity of citrullinated proteins (Figure 2A). These data demonstrated that PPAD is essential for citrullination of endogenous proteins in *P gingivalis* and that it possesses only 1 peptidylarginine deiminase.

We then examined how citrullination depends on the activity of the major virulence factors in *P gingivalis*, which are called gingipains (32,33). Gingipains are potent proteases and cleave various proteins/peptides after either arginine residues (i.e., arginine gingipain [Rgp]) or lysine residues (i.e., lysine gingipain [Kgp]), resulting in peptides with carboxy-terminal arginine or lysine residues. The reported preference of native PPAD for carboxy-terminal arginine residues in vitro (26) could be mediated through the activity of arginine gingipains, and as such, this would have important consequences for the type of citrullinated peptides that can be generated by *P gingivalis*. Hence, we studied *P gingivalis* mutants lacking functional arginine (Δ*rgp*), lysine (Δ*kgp*), or both (Δ*rgp*+*kgp*) types of gingipains for endogenous citrullination. Immunoblotting of whole-cell lysates showed a significantly decreased level, but not complete abrogation, of citrullinated proteins in the Δ*rgp* and Δ*rgp*+*kgp* strains, but not in the Δ*kgp* strain (Figure 2B), confirming that arginine gingipains play a role in protein citrullination, probably by generating proteins with carboxy-terminal arginine residues that are subsequently citrullinated by PPAD. The residual citrullinated proteins seen in the Δ*rgp* and Δ*rgp*+*kgp* strains might be due to the presence of proteins that naturally contain a carboxy-terminal arginine residue, which had not been proteolytically processed, and therefore appeared at different molecular weights as compared with the wild-type.

***P gingivalis* rapidly generates citrullinated fibrinogen and α-enolase peptides by proteolytic cleavage at Arg-X peptide bonds, followed by citrullination of carboxy-terminal arginines.** Having found that citrullination of endogenous *P gingivalis* proteins depends on the presence of PPAD and is influenced by arginine gingipains, the question arose whether the same principles apply to human proteins. We initially chose human fibrinogen for this study, since it is a major RA autoantigen in its citrullinated form (2,3,34) and a major part of the inflammatory response in general because of its function in the coagulation and platelet aggregation cascade. Fibrinogen is also involved in the pathogenesis of periodontitis, where it is abundantly found in the periodontal lesion, being an established target protein of gingipains (35). We investigated the potential of *P gingivalis* to citrullinate human fibrinogen using intact, live wild-type, Δ*ppad*-knockout, and Δ*rgp*-knockout strains.

Fibrinogen was rapidly cleaved by wild-type *P gingivalis* (Figure 3A), which is consistent with previous

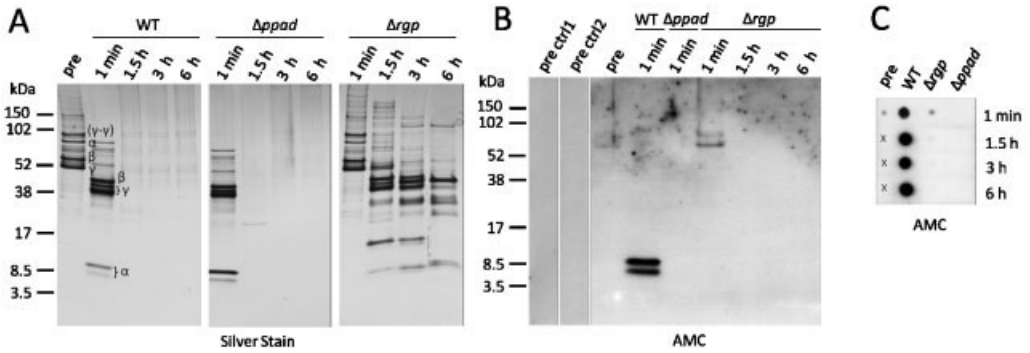


Figure 3. *Porphyromonas gingivalis* rapidly cleaves human fibrinogen through arginine gingipain activity, and the resulting peptides are citrullinated at the carboxy-terminus by bacterial peptidylarginine deiminase (PAD). **A,** Fibrinogen fragments were incubated for 1 minute, 1.5 hours, 3 hours, or 6 hours with *P gingivalis* wild-type (WT) strain or with mutant strains lacking bacterial PAD (Δ*ppad*) or arginine gingipain (Δ*rgp*), resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis, and visualized using silver staining. A fibrinogen sample prior to incubation with *P gingivalis* (pre) served as a control. Protein bands were analyzed by mass spectrometry. Bands labeled α, β, and γ indicate the α, β, and γ chains of fibrinogen. **B,** All samples with visible protein bands from the experiments shown in **A** were tested for citrullination by immunoblotting with anti–modified citrulline (AMC). Controls for the preincubation sample were performed in which the modification step (pre ctrl1) or the secondary antibody (pre ctrl2) was omitted. **C,** All samples from the experiments shown in **A** were analyzed for the presence of citrullinated peptides by dot-blotting. Areas on the membrane where samples were not applied are marked as ×.

reports (36). A similar degradation pattern was observed with the Δ*ppad* strain, indicating that citrullination of substrate proteins, including gingipains themselves, is not essential to the proteolytic potency of gingipains. As expected, fibrinogen samples incubated with the Δ*rgp* strain showed considerably decreased proteolytic cleavage and a different pattern of cleaved peptides, with the residual proteolytic activity being mainly due to cleavage by lysine gingipains and other proteinases and peptidases from *P gingivalis* (37–40). Our analysis of the identity of the protein bands by mass spectrometry confirmed that the majority were derived from any of the 3 fibrinogen chains (Figure 3A). To exclude cleavage of protein by plasma-derived proteases, which may contaminate fibrinogen, we further performed control reactions in which protein alone was incubated in assay buffer, and no such cleavage was detected.

Next, we examined whether the cleaved fibrinogen fragments had been citrullinated by *P gingivalis*. Immunoblotting of all samples that contained protein bands (see Figure 3A) detected 2 citrullinated peptide bands at ~8.5 kd, mapping to the amino-terminal region of the fibrinogen α-chain, in samples incubated with *P gingivalis* wild-type, but not Δ*ppad* or Δ*rgp* (Figure 3B), confirming that Arg-X proteolytic cleavage of fibrinogen is a prerequisite for subsequent citrullination by PPAD. We further observed a weak positive signal in fibrinogen samples taken before incubation with *P gingivalis* and in the Δ*rgp* sample taken at 1 minute, while the corresponding controls (the modification and the conjugate controls) were negative (Figure 3B), suggesting that purified human fibrinogen, as used in these experiments, is already endogenously citrullinated by human PADs.

The fact that the majority of fibrinogen had been degraded within minutes and that only 2 citrullinated proteins/peptides could be detected by immunoblotting suggested that the majority of generated peptides was smaller than the size limit of the peptide gels (~3 kd) used for this analysis. Thus, we applied a dot-blot technique, which confirmed that the wild-type *P gingivalis* cells rapidly degraded and citrullinated fibrinogen into small citrullinated peptides (Figure 3C) and that both arginine gingipains and PPAD were required, since no positive signals were observed with the Δ*ppad* and Δ*rgp* strains.

We then aimed to identify the amino acid sequence of the citrullinated fibrinogen peptides and determine the position of the citrulline residue. To this end, we fractionated the peptides derived from the wild-type and Δ*ppad* strains by HPLC and analyzed the eluted peak fractions by liquid chromatography tandem mass spectrometry. We identified a total of 30 peptides derived from fibrinogen (results not shown). The majority of the identified peptides were the product of proteolytic cleavage after either an arginine residue or a lysine residue, which is consistent with the results de-

Figure 4. Sequences of citrullinated peptides from human fibrinogen generated after incubation with *Porphyromonas gingivalis*. Amino acid sequences of human fibrinogen α-chain (Fib A; Swiss-Prot entry P02671) and fibrinogen β-chain (Fib B; Swiss-Prot entry P02675) are shown. Peptides were detected using liquid chromatography tandem mass spectrometry. Citrullinated peptides detected after incubation of fibrinogen with *P gingivalis* wild-type strain are underlined. Fibrinopeptides A and B and the thrombin cleavage sites are indicated. Cit = citrulline.

scribed above, but further proteolytic processing by other peptidases, particularly at the amino-terminus after glycine, alanine, and serine residues, was also evident. In samples incubated with *P gingivalis* wild-type, but not Δ*ppad*, we found 4 peptides that contained a carboxy-terminal citrulline residue (Figure 4): [1]ADS-

Table 1. Mass spectrometry of citrullinated peptides generated after incubation of human fibrinogen or α-enolase with *Porphyromonas gingivalis* wild-type strain and their respective arginated peptides generated with the Δ*ppad* strain*

Protein, sequence	Amino acid positions	*P gingivalis* strain	m/z ratio	Mascot score
Fibrinogen α-chain				
−.ADSGEGDFLAEGGGVCit.G	1–16	Wild type	769.29 (+2)	56
−.ADSGEGDFLAEGGGVR.G	1–16	Δ*ppad*	768.77 (+2)	99
Fibrinogen α-chain				
R.GGSTSYGTGSETESPCit.N	253–268	Wild type	546.91 (+3)	51
R.GGSTSYGTGSETESPR.N	253–268	Δ*ppad*	786.82 (+2)	78
Fibrinogen α-chain				
K.ESSSHHPGIAEFPSCit.G	540–554	Wild type	819.83 (+2)	98
S.SHHPGIAEFPSR.G	543–554	Δ*ppad*	445.54 (+3)	39
Fibrinogen β-chain				
R.PAPPPISGGGYCit.A	31–42	Wild type	585.30 (+2)	26
R.PAPPPISGGGYR.A	31–42	Δ*ppad*	584.77 (+2)	65
α-enolase				
S.TGIYEALELCit.D	41–50	Wild type	583.30 (+2)	24
S.TGIYEALELR.D	41–50	Δ*ppad*	582.79 (+2)	56

* Amino acids in the uncleaved proteins located carboxy–terminal and amino–terminal to the identified peptides are indicated to demonstrate the sites of proteolytic cleavage. Citrulline (Cit) residues that were identified are underlined. For the mass/charge (m/z) ratio, the numbers in parentheses are the peptide ion charge state.

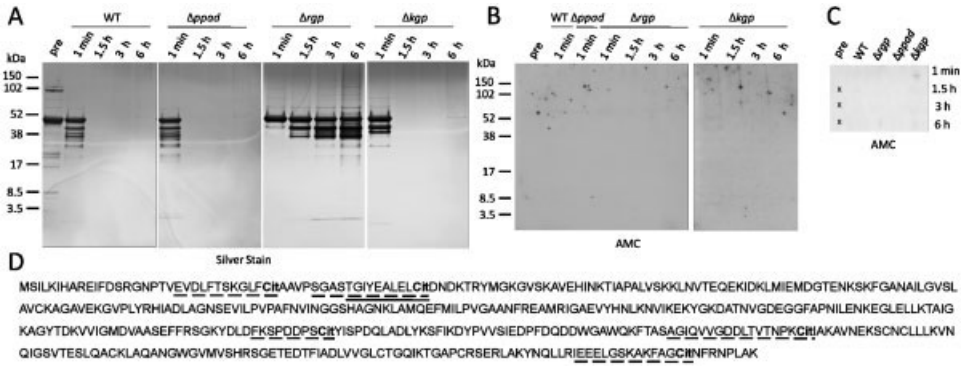


Figure 5. Human α-enolase is rapidly cleaved by *Porphyromonas gingivalis* gingipains, and citrullinated peptides are detectable by mass spectrometry. **A,** Analogous to the experiments with fibrinogen shown in Figure 3, human α-enolase was incubated for 1 minute, 1.5 hours, 3 hours, or 6 hours with *P gingivalis* wild-type (WT) strain or with mutant strains lacking bacterial peptidylarginine deiminase (Δ*ppad*), arginine gingipain (Δ*rgp*), or lysine gingipain (Δ*kgp*), resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis, and visualized using silver staining. An α-enolase sample prior to incubation with *P gingivalis* (pre) served as a control. **B,** All samples with visible protein bands from the experiments shown in **A** were tested for citrullination by immunoblotting with anti–modified citrulline (AMC). **C,** All samples from the experiments shown in **A** were analyzed for the presence of citrullinated peptides by dot-blotting. Areas on the membrane where samples were not applied are marked as ×. **D,** Citrullinated peptides detected by mass spectrometry after incubation of enolase with *P gingivalis* wild-type and Δ*kgp* are underlined with a continuous line and with a dashed line, respectively. The amino acid sequence of human α-enolase (Swiss-Prot entry P06733) is shown. Cit = citrulline.

GEGDFLAEGGGVCit^{16}, ^{31}PAPPPISGGGYCit^{42}, ^{253}GGSTSYGTGSETESPCit^{268}, and ^{540}ESSSHHPGIAEFPSCit^{554}. Carboxy-terminal arginine-containing peptides were detected only in the Δ*ppad*, but not the wild-type, strain (Table 1), suggesting that citrullination of fibrinogen is tightly linked to cleavage by arginine gingipains. The combined data support the concept that target proteins such as human fibrinogen are cleaved by arginine gingipains, generating suitable peptide substrates for subsequent citrullination at the exposed carboxy-terminal arginine residue by *P gingivalis* PAD.

To further test this concept, we performed analogous experiments using recombinant human α-enolase. Similar to the findings with fibrinogen, α-enolase was rapidly degraded by the wild-type and Δ*ppad* strains and less so by the Δ*rgp* strain (Figure 5A). Using immunoblotting on peptide SDS-PAGE gels (Figure 5B) as well as dot-blotting (Figure 5C), no citrullination could be detected. Analysis of the samples derived from the wild-type and Δ*ppad* strains by mass spectrometry revealed only 1 citrullinated peptide in the wild-type (^{41}TGIYEALELCit^{50}) (Figure 5D and Table 1), among a total of 17 peptides detected. The arginine-containing counterpart of this citrullinated peptide was detected in the samples incubated with Δ*ppad*. Analogous to fibrinogen, no peptides with carboxy-terminal arginine were detected in the wild-type samples. The proportion of peptides cleaved at residues other than arginine and lysine was higher than that found in fibrinogen, suggesting extensive cleavage by non–arginine/lysine-specific peptidases. Combined with the higher relative number of lysine residues in α-enolase (8.8% versus 6.9% in fibrinogen), this might result in the generation of short peptides, some of which might be citrullinated but would be too short to be detected using these methods. We therefore incubated *P gingivalis* Δ*kgp* with α-enolase (Figures 5A–C) and detected 5 citrullinated peptides by mass spectrometry (Figure 5D), confirming that PPAD is able to citrullinate α-enolase peptides. Using the AMC antibody dot-blot, which relies on long peptides that are hydrophobic enough to bind to the membrane, we observed weak positive signals with the Δ*kgp* strain, which decreased with time (Figure 5C), again suggesting extensive proteolytic degradation by other proteinases.

DISCUSSION

In the present study, we found evidence that the periodontal pathogen *Porphyromonas gingivalis* is an alternative source in the human host for generating citrullinated proteins and peptides. The underlying mechanism—proteolytic cleavage and subsequent citrul-

lination at carboxy-terminal arginine residues—differs from that of the human PAD enzymes, which citrullinate internal arginine residues in whole proteins the most efficiently. This finding suggests that protein citrullination by the bacterial PAD has the potential to generate epitopes to which immunologic tolerance does not exist, not only due to the presence of foreign citrullinated proteins from the bacterium, but also through a foreign mode of proteolytic processing and posttranslational modification of host antigens. It also indicates that citrullination of bacterial proteins at internal arginines, as a potential mechanism for triggering autoantibodies via molecular mimicry (5), is more likely to be due to the action of human PAD enzymes that are present at the site of inflammation.

Of the 11 oral bacterial species tested, endogenous citrullinated proteins were detected exclusively in *P gingivalis*, indicating that a bacterial PAD gene is expressed or is active only in this bacterium among those tested. To substantiate this finding, we performed similarity searches using BLAST and PSI-BLAST (41). This revealed numerous ortholog distantly related to PPAD among prokaryotes, including several of the oral organisms we tested. Most share the predicted conserved catalytic residues of PPAD and other members of the guanidino group–modifying enzyme superfamily, although they most likely possess agmatine iminohydrolase or arginine deiminase, rather than peptidylarginine deiminase activity (42).

Using fibrinogen as a model antigen, we showed that *P gingivalis* rapidly generated small fibrinogen peptides with carboxy-terminal citrulline residues. Fibrinopeptide A, which normally results from thrombin cleavage of the fibrinogen α-chain after arginine-16, was also detected in its citrullinated form (^{1}ADSGEGDFLAEGGGV$\underline{\text{cit}}^{16}$) in samples incubated with *P gingivalis* wild-type, but only in the native, arginine-containing form in the Δ*ppad* samples. It is known that *P gingivalis* gingipain–mediated degradation of human fibrinogen inhibits fibrinogen polymerization and results in the localized bleeding tendency that is typical of chronic periodontitis (35). A recent study showed that in intact fibrinogen, internal citrullination at arginine-16 by mammalian PAD impairs thrombin-catalyzed cleavage and fibrin polymerization (43), indicating at least 2 possible pathogenic roles of citrullinated fibrinogen in RA: serving as an autoantigen and disturbing the coagulation cascade and linked pathways.

The pathophysiologic role of fibrinogen peptides with carboxy-terminal citrulline residues, which are generated by the concerted action of gingipain and PPAD, is as yet unknown, thus opening up a novel area for future investigations. Similarly, it is known that *P gingivalis* arginine gingipains cleave a number of other human proteins, releasing biologically active peptides with important roles in immunity and inflammation, such as C5a (44) and bradykinin (45,46), and simultaneous citrullination of these peptides by *P gingivalis* PAD might have a previously unappreciated role in human disease.

The lower levels of detectable citrullination of α-enolase peptides with the *P gingivalis* wild-type strain are likely to be the result of a combination of physiologic and technical factors. Enolase has a lower percentage of arginine residues (3.9%) as compared with fibrinogen (5.2%) and a higher percentage of lysine residues (enolase 8.8% versus fibrinogen 6.9%). It also appears to be more extensively cleaved by non-Arg/Lys peptidases. In combination, this would result in fewer suitable PPAD substrates, and very small peptides overall, which would not be detected with the methods used in the present study. Thus, lower levels of detectable citrullination may simply be due to a relative paucity of the substrate and technical shortcomings with the detection of short peptides.

Herein, we have demonstrated that *P gingivalis* efficiently citrullinates its own proteins and peptides from host fibrinogen and, to a lesser extent, α-enolase. The 2 major findings of this study—that proteolytic processing is required for citrullination by *P gingivalis* and that host peptides with exclusively carboxy-terminal citrulline residues are generated—provides a strong basis for future in vivo studies aimed at identifying citrullinated peptides at the site of gingival inflammation and exploring their potency for triggering a T cell and/or B cell response. Citrullinated host peptides generated by *P gingivalis* are likely to expose epitopes previously hidden from immune surveillance, which in the context of bacterial infection in a genetically susceptible host, may trigger an immune response. The slightly increased prevalence of anticitrullinated protein antibodies reported in patients with periodontitis as compared with healthy controls (47,48) supports this concept, but the lower frequency and titer than are found in RA patients suggest that *P gingivalis* infection is not sufficient on its own for the mature autoimmune response. However, once tolerance is breached, we predict that exposure to host proteins in the inflamed joint, which have been citrullinated by human PADs (31), leads to intra- and intermolecular epitope spreading to additional peptides from the initiating proteins and other autoantigens.

We therefore propose a "two-hit" model of RA, based first on the breakdown of tolerance to specific

citrullinated peptides generated by *P gingivalis* at the site of gingival inflammation and followed by epitope-spreading to other host citrullinated proteins in the inflamed joint. This self-sustaining immune response would then result in the chronic and destructive inflammation that typifies RA. The unique nature of the bacterial deiminase, along with its location on the cell surface of the bacterium (26), provides a target for treatment designed to prevent this otherwise incurable disease.

ACKNOWLEDGMENTS

We are thankful to Drs. Andrzej Kozik, Maria Rapala-Kozik, and Anna Golda (Jagiellonian University, Krakow, Poland) for their help with the HPLC analysis and to Mr. Anto Jose (Glasgow Dental School, Glasgow, UK) for preparation of bacterial strains.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Venables had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Wegner, Wait, Sroka, Lundberg, Potempa, Venables.

Acquisition of data. Wegner, Wait, Sroka, Eick, Nguyen, Kinloch, Culshaw, Venables.

Analysis and interpretation of data. Wegner, Wait, Sroka, Potempa, Venables.

REFERENCES

1. Wegner N, Lundberg K, Kinloch A, Fisher B, Malmstrom V, Feldmann M, et al. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. Immunol Rev 2010;233:34–54.
2. Masson-Bessiere C, Sebbag M, Girbal-Neuhauser E, Nogueira L, Vincent C, Senshu T, et al. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the α- and β-chains of fibrin. J Immunol 2001;166:4177–84.
3. Sebbag M, Moinard N, Auger I, Clavel C, Arnaud J, Nogueira L, et al. Epitopes of human fibrin recognized by the rheumatoid arthritis-specific autoantibodies to citrullinated proteins. Eur J Immunol 2006;36:2250–63.
4. Kinloch A, Tatzer V, Wait R, Peston D, Lundberg K, Donatien P, et al. Identification of citrullinated α-enolase as a candidate autoantigen in rheumatoid arthritis. Arthritis Res Ther 2005;7: R1421–9.
5. Lundberg K, Kinloch A, Fisher BA, Wegner N, Wait R, Charles P, et al. Antibodies to citrullinated α-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. Arthritis Rheum 2008;58:3009–19.
6. Snir O, Widhe M, von Spee C, Lindberg J, Padyukov L, Lundberg K, et al. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. Ann Rheum Dis 2009;68:736–43.
7. Wegner N, Wait R, Venables PJ. Evolutionarily conserved antigens in autoimmune disease: implications for an infective aetiology. Int J Biochem Cell Biol 2009;41:390–7.
8. Wygrecka M, Marsh LM, Morty RE, Henneke I, Guenther A, Lohmeyer J, et al. Enolase-1 promotes plasminogen-mediated recruitment of monocytes to the acutely inflamed lung. Blood 2009;113:5588–98.
9. Pancholi V, Fischetti VA. Alpha-enolase, a novel strong plasmin-(ogen) binding protein on the surface of pathogenic streptococci. J Biol Chem 1998;273:14503–15.
10. Clavel C, Nogueira L, Laurent L, Iobagiu C, Vincent C, Sebbag M, et al. Induction of macrophage secretion of tumor necrosis factor α through Fcγ receptor IIa engagement by rheumatoid arthritis–specific autoantibodies to citrullinated proteins complexed with fibrinogen. Arthritis Rheum 2008;58:678–88.
11. Zhao X, Okeke NL, Sharpe O, Batliwalla FM, Lee AT, Ho PP, et al. Circulating immune complexes contain citrullinated fibrinogen in rheumatoid arthritis. Arthritis Res Ther 2008;10:R94.
12. Trouw LA, Haisma EM, Levarht EW, van der Woude D, Ioan-Facsinay A, Daha MR, et al. Anti–cyclic citrullinated peptide antibodies from rheumatoid arthritis patients activate complement via both the classical and alternative pathways. Arthritis Rheum 2009;60:1923–31.
13. Hill JA, Bell DA, Brintnell W, Yue D, Wehrli B, Jevnikar AM, et al. Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. J Exp Med 2008;205: 967–79.
14. Gyorgy B, Toth E, Tarcsa E, Falus A, Buzas EI. Citrullination: a posttranslational modification in health and disease. Int J Biochem Cell Biol 2006;38:1662–77.
15. Makrygiannakis D, af Klint E, Lundberg IE, Lofberg R, Ulfgren AK, Klareskog L, et al. Citrullination is an inflammation-dependent process. Ann Rheum Dis 2006;65:1219–22.
16. Van Venrooij WJ, Zendman AJ. Anti-CCP2 antibodies: an overview and perspective of the diagnostic abilities of this serological marker for early rheumatoid arthritis. Clin Rev Allergy Immunol 2008;34:36–9.
17. Vossenaar ER, Smeets TJ, Kraan MC, Raats JM, van Venrooij WJ, Tak PP. The presence of citrullinated proteins is not specific for rheumatoid synovial tissue. Arthritis Rheum 2004;50:3485–94.
18. Chapuy-Regaud S, Sebbag M, Baeten D, Clavel C, Foulquier C, De Keyser F, et al. Fibrin deimination in synovial tissue is not specific for rheumatoid arthritis but commonly occurs during synovitides. J Immunol 2005;174:5057–64.
19. Van der Helm-van Mil AH, Verpoort KN, Breedveld FC, Huizinga TW, Toes RE, de Vries RR. The HLA–DRB1 shared epitope alleles are primarily a risk factor for anti–cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum 2006;54: 1117–21.
20. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;54:38–46.
21. Mahdi H, Fisher BA, Kallberg H, Plant D, Malmstrom V, Ronnelid J, et al. Specific interaction between genotype, smoking and autoimmunity to citrullinated α-enolase in the etiology of rheumatoid arthritis. Nat Genet 2009;41:1319–24.
22. Borrell LN, Burt BA, Taylor GW. Prevalence and trends in periodontitis in the USA: from the NHANES III to the NHANES, 1988 to 2000. J Dent Res 2005;84:924–30.
23. Takeuchi Y, Umeda M, Sakamoto M, Benno Y, Huang Y, Ishikawa I. Treponema socranskii, Treponema denticola, and Porphyromonas gingivalis are associated with severity of periodontal tissue destruction. J Periodontol 2001;72:1354–63.
24. Griffen AL, Lyons SR, Becker MR, Moeschberger ML, Leys EJ.

Porphyromonas gingivalis strain variability and periodontitis. J Clin Microbiol 1999;37:4028–33.
25. De Pablo P, Chapple IL, Buckley CD, Dietrich T. Periodontitis in systemic rheumatic diseases. Nat Rev Rheumatol 2009;5:218–24.
26. McGraw WT, Potempa J, Farley D, Travis J. Purification, characterization, and sequence analysis of a potential virulence factor from Porphyromonas gingivalis, peptidylarginine deiminase. Infect Immun 1999;67:3248–56.
27. Graves D. Cytokines that promote periodontal tissue destruction. J Periodontol 2008;79:1585–91.
28. De Pablo P, Dietrich T, McAlindon TE. Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population. J Rheumatol 2008;35:70–6.
29. Sugawara K, Oikawa Y, Ouchi T. Identification and properties of peptidylarginine deiminase from rabbit skeletal muscle. J Biochem 1982;91:1065–71.
30. Nguyen KA, Travis J, Potempa J. Does the importance of the C-terminal residues in the maturation of RgpB from Porphyromonas gingivalis reveal a novel mechanism for protein export in a subgroup of gram-negative bacteria? J Bacteriol 2007;189:833–43.
31. Kinloch A, Lundberg K, Wait R, Wegner N, Lim NH, Zendman AJ, et al. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. Arthritis Rheum 2008;58:2287–95.
32. Travis J, Pike R, Imamura T, Potempa J. Porphyromonas gingivalis proteinases as virulence factors in the development of periodontitis. J Periodontal Res 1997;32:120–5.
33. Potempa J, Sroka A, Imamura T, Travis J. Gingipains, the major cysteine proteinases and virulence factors of Porphyromonas gingivalis: structure, function and assembly of multidomain protein complexes. Curr Protein Pept Sci 2003;4:397–407.
34. Klareskog L, Ronnelid J, Lundberg K, Padyukov L, Alfredsson L. Immunity to citrullinated proteins in rheumatoid arthritis. Annu Rev Immunol 2008;26:651–75.
35. Imamura T, Potempa J, Pike RN, Moore JN, Barton MH, Travis J. Effect of free and vesicle-bound cysteine proteinases of Porphyromonas gingivalis on plasma clot formation: implications for bleeding tendency at periodontitis sites. Infect Immun 1995;63:4877–82.
36. Ally N, Whisstock JC, Sieprawska-Lupa M, Potempa J, Le Bonniec BF, Travis J, et al. Characterization of the specificity of arginine-specific gingipains from Porphyromonas gingivalis reveals active site differences between different forms of the enzymes. Biochemistry 2003;42:11693–700.
37. Nelson D, Potempa J, Kordula T, Travis J. Purification and characterization of a novel cysteine proteinase (periodontain) from Porphyromonas gingivalis: evidence for a role in the inactivation of human α_1-proteinase inhibitor. J Biol Chem 1999;274:12245–51.
38. Jagels MA, Travis J, Potempa J, Pike R, Hugli TE. Proteolytic inactivation of the leukocyte C5a receptor by proteinases derived from Porphyromonas gingivalis. Infect Immun 1996;64:1984–91.
39. Masuda K, Yoshioka M, Hinode D, Nakamura R. Purification and characterization of arginine carboxypeptidase produced by Porphyromonas gingivalis. Infect Immun 2002;70:1807–15.
40. Banbula A, Bugno M, Goldstein J, Yen J, Nelson D, Travis J, et al. Emerging family of proline-specific peptidases of Porphyromonas gingivalis: purification and characterization of serine dipeptidyl peptidase, a structural and functional homologue of mammalian prolyl dipeptidyl peptidase IV. Infect Immun 2000;68:1176–82.
41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10.
42. Shirai H, Mokrab Y, Mizuguchi K. The guanidino-group modifying enzymes: structural basis for their diversity and commonality. Proteins 2006;64:1010–23.
43. Nakayama-Hamada M, Suzuki A, Furukawa H, Yamada R, Yamamoto K. Citrullinated fibrinogen inhibits thrombin-catalysed fibrin polymerization. J Biochem 2008;144:393–8.
44. Popadiak K, Potempa J, Riesbeck K, Blom AM. Biphasic effect of gingipains from Porphyromonas gingivalis on the human complement system. J Immunol 2007;178:7242–50.
45. Monteiro AC, Scovino A, Raposo S, Gaze VM, Cruz C, Svensjo E, et al. Kinin danger signals proteolytically released by gingipain induce fimbriae-specific IFN-γ- and IL-17-producing T cells in mice infected intramucosally with Porphyromonas gingivalis. J Immunol 2009;183:3700–11.
46. Imamura T, Pike RN, Potempa J, Travis J. Pathogenesis of periodontitis: a major arginine-specific cysteine proteinase from Porphyromonas gingivalis induces vascular permeability enhancement through activation of the kallikrein/kinin pathway. J Clin Invest 1994;94:361–7.
47. Mikuls TR, Payne JB, Reinhardt RA, Thiele GM, Maziarz E, Cannella AC, et al. Antibody responses to Porphyromonas gingivalis (P. gingivalis) in subjects with rheumatoid arthritis and periodontitis. Int Immunopharmacol 2009;9:38–42.
48. Havemose-Poulsen A, Westergaard J, Stoltze K, Skjodt H, Danneskiold-Samsoe B, Locht H, et al. Periodontal and hematological characteristics associated with aggressive periodontitis, juvenile idiopathic arthritis, and rheumatoid arthritis. J Periodontol 2006;77:280–8.

arthritis research&therapy

REVIEW

Bacterial and human peptidylarginine deiminases: targets for inhibiting the autoimmune response in rheumatoid arthritis?

Pamela Mangat[1], Natalia Wegner[1], Patrick J Venables*[1] and Jan Potempa[2,3]

Abstract

Peptidylarginine deiminases (PADs) convert arginine within a peptide (peptidylarginine) into peptidylcitrulline. Citrullination by human PADs is important in normal physiology and inflammation. *Porphyromonas gingivalis*, a major pathogen in periodontitis, is the only prokaryote described to possess PAD. *P. gingivalis* infection may generate citrullinated peptides, which trigger anti-citrullinated peptide antibodies. In susceptible individuals, host protein citrullination by human PADs in the joint probably perpetuates antibody formation, paving the way for the development of chronic arthritis. Blockades of bacterial and human PADs may act as powerful novel therapies by inhibiting the generation of the antigens that trigger and sustain autoimmunity in rheumatoid arthritis.

Introduction

Arginine is a positively charged, hydrophilic amino acid that is often found on the surface of proteins, where it participates in ionic interactions with other amino acid side chains and forms stabilizing hydrogen bonds with both the peptide backbone and amino acid side chains. These characteristics make it a key amino acid in the three-dimensional organization of proteins and in the interaction with other biological molecules. Hence, post-translational modification of arginine can alter the three-dimensional protein structure and function and potentially expose previously hidden epitopes to the immune system. Deimination (citrullination) of arginine side chains (peptidylarginine) to form peptidylcitrulline is one of many recognized post-translational modifications of this amino acid. This post-translational conversion is catalyzed by the family of peptidylarginine deiminase (PAD) enzymes. The process of protein citrullination plays a vital role in normal physiology, in which it is involved in the formation of rigid structures such as hair, skin, and myelin sheaths [1]. Aberrant citrullination has been observed in diseases of the skin and nervous system and in inflammatory arthritides, of which rheumatoid arthritis (RA) is one example [1]. Despite the ubiquity of citrullinated proteins, the autoantibody response to citrullinated proteins is largely restricted to RA [2]. The switch that leads to the generation of antibodies to citrullinated peptides and thus loss of immune tolerance to citrullinated proteins is likely to involve a complex interplay of individual genetic and environmental factors.

Citrullination by human and bacterial peptidylarginine deiminases

In humans, a family of five PAD enzymes (PAD1 to 4 and PAD6), encoded by five genes clustered on chromosome 1p35-36, has been described [3]. Apart from PAD4, which can translocate to the nucleus, PAD enzymes are typically found in the cytoplasm of various cell types and show a characteristic tissue distribution. The localization and functions of each of the human PAD enzymes are summarized in Table 1. Homologous amino acid sequences for some or all of these PADs exist in other eukaryotic species, such as the mouse, chicken, frog, and bony fish. Among prokaryotic species, PAD activity has, to date, been described in *Porphyromonas gingivalis* only [4]. *P. gingivalis* is a major pathogen in periodontitis, a disease that (akin to RA) is a chronic inflammatory disorder characterized by pro-inflammatory cytokine production and erosion of bone.

As protein citrullination in the joint is not specific to RA [5] and autoantibodies to citrullinated proteins precede the clinical signs of RA [6], it has been proposed that oral citrullination of human and bacterial proteins by *P. gingivalis* PAD (PPAD) in an infectious context prior to the onset of RA could break tolerance and trigger a latent antibody response against citrullinated protein [7].

*Correspondence: p.venables@imperial.ac.uk
[1]The Kennedy Institute of Rheumatology Division, Imperial College, 65 Aspenlea Road, London, W6 8LH, UK
Full list of author information is available at the end of the article

Table 1. Localization and function of human peptidylarginine deiminase enzymes

	Localization	Function	Reference
PAD1	Epidermis, hair follicles, arrector pili muscles, and sweat glands	Citrullination of filaggrin and keratin, facilitating proteolysis and crosslinking of the proteins and contributing to skin cornification. Maintains hydration of stratum corneum and epidermis barrier function. Differentiation of hair follicles.	[66-68]
PAD2	Brain astrocytes, sweat glands, arrector pili muscles, skeletal muscle, spleen, macrophages, monocytes, epidermis, synovial tissue, and synovial fluid	Citrullination of myelin basic protein in the brain and spinal cord, promoting electrical insulation of myelin sheaths. Citrullination of vimentin in apoptotic monocytes and macrophages.	[45,46,66,67,69-73]
PAD3	Upper layers of epidermis and hair follicles	Citrullination of trichohyalin, contributing to directional hair growth.	[66-68]
PAD4	Hematopoietic cells and inflamed rheumatoid synovium	Citrullination of transcriptional coactivator p300 and histones H2A, H3, and H4, regulating gene expression by chromatin remodelling. Citrullination of fibrin, contributing to chronic inflammation in rheumatoid arthritis. P53-dependent citrullination of proteins following DNA damage, translocation of histone chaperone nucleophosmin, and p53-mediated inhibition of tumor cell growth.	[35,44,45,74]
PAD6	Ovary and testis tissue and peripheral blood leukocytes	Amino acids known to be conserved in PAD enzymatic activity are not conserved in PAD6. Function and enzymatic activity remain unclear.	[3,73]

PAD, peptidylarginine deiminase.

Once tolerance is breached, citrullination of host proteins by human PADs perpetuates the immune response through epitope spreading and cross-reactivity, resulting in chronic inflammatory disease (Figure 1). Citrullination by both human and bacterial PAD enzymes may thus provide a target for inhibiting the immune response at an early stage in the inflammatory pathway of RA.

The best-established autoantigens in RA include α-enolase, fibrinogen, vimentin, and type II collagen (reviewed in [1]) and all are efficiently deiminated by mammalian PADs. In theory, citrullinated peptides from these antigens could also be generated by PPAD, although this has yet to be demonstrated experimentally. Alpha-enolase is of particular interest in this respect because it is highly conserved among eukaryotes and prokaryotes. A sequence of nine amino acids (Asp-Ser-Arg-Gly-Asn-Pro-Thr-Val-Glu) spanning the immunodominant epitope on the peptide known as citrullinated enolase peptide-1 (CEP-1) is 100% identical to the corresponding region in *P. gingivalis* enolase, and affinity-purified antibodies to CEP-1 react with recombinant enolase citrullinated *in vitro* from both humans and *P. gingivalis* [8], providing an attractive target for molecular mimicry between human and bacterial species.

Etiological association between periodontitis and rheumatoid arthritis

The rationale for considering both human and *P. gingivalis* PADs in the etiology and pathology of RA is also based on epidemiological data suggesting an association between the two diseases (reviewed in [9]). Periodontitis and RA are chronic inflammatory disorders characterized by erosion of bone and production of pro-inflammatory cytokines. The reported prevalence of periodontitis is highly variable; in one large study of the American population, the prevalence was 4.2% [10]. Epidemiological studies have shown that RA is more prevalent among patients with periodontal disease (3.95%) than in the general population (1%) [11]. In addition, patients with RA have a higher frequency of advanced periodontal disease than the general population [12]. *P. gingivalis*, *Treponema denticola*, and *Tannerella forsythia* are some of the major Gram-negative bacteria that exist as part of a complex bacterial biofilm in the gingival crevice and are linked to the development and progression of periodontitis but can also be found in lower numbers in periodontally healthy subjects [13]. Long-term plaque accumulation and an interplay of host and bacterial factors result in chronic inflammation and tissue damage. Destruction of the adjacent bone and periodontal ligament attachment may eventually lead to tooth loss [14]. *P. gingivalis* antibody levels have been shown to correlate with anti-CCP (anti-cyclic citrullinated peptide) antibody titres [15], making this periodontopathic oral bacterium an attractive candidate environmental trigger in the development of RA.

Several research groups have reported an increased variety and number of oral bacterial DNA and antibodies targeting these bacteria in serum and synovial fluid of patients with RA and other inflammatory joint diseases compared with controls (non-inflammatory arthritides or healthy donors) [15-19]. Oral bacterial DNA could

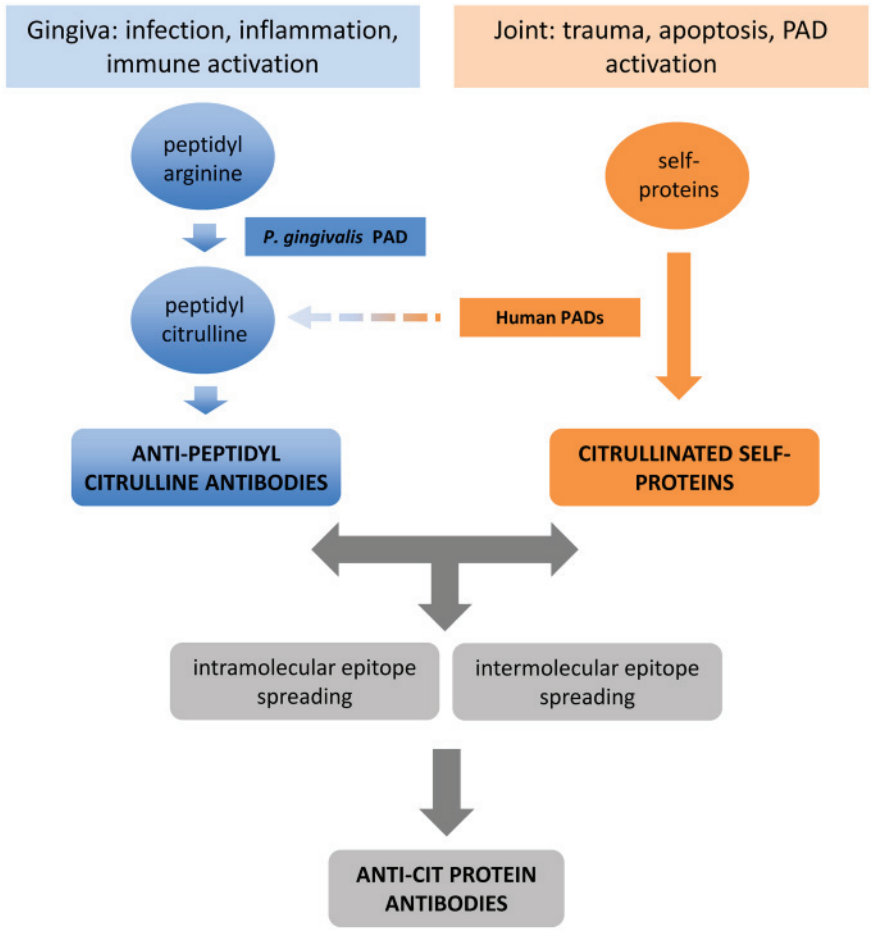


Figure 1. Simplified model illustrating the hypothesis that *Porphyromonas gingivalis*-mediated citrullination triggers anti-citrulline autoimmunity in rheumatoid arthritis. Citrullination by *P. gingivalis* peptidylarginine deiminase (PAD) in the inflammatory context of periodontitis produces bacterial and host-derived citrullinated peptides to which the immune system mounts a humoral immune response with the production of peptidylcitrulline antibodies. Inflammation-induced citrullination by human PAD enzymes in the gingiva is also possible (dashed arrow). Tissue injury and inflammation in the joint lead to activation of human PAD enzymes and citrullination of host proteins, such as α-enolase, vimentin, fibrin(ogen), and collagen type II. Peptidylcitrulline antibodies bind citrullinated host and bacterial peptides, which may show molecular mimicry, and in genetically susceptible individuals (presence of the certain HLA alleles), intra- and intermolecular epitope spreading leads to a sustained immune response with the formation of high-affinity antibodies to host citrullinated proteins.

reach the joint as free DNA or intracellularly in immune cells. Owing to the stringent growth requirements of live oral bacteria, their presence in the joint is unlikely, and no viable organisms have been obtained from synovial fluid [19]. However, these observations need to be interpreted with caution since many bacterial antibody assays using whole-bacterium lysates are of questionable specificity, and the same applies to polymerase chain reaction-based detection and DNA-DNA hybridization using a complex nucleic acid mixture containing an excess of human DNA. A number of antibiotics used in the treatment of periodontitis, such as tetracyclines and clarithromycin, are efficacious in the treatment of RA [20-24], although to date there has been no direct evidence that this therapeutic effect is due to their anti-bacterial activity. For example, minocycline has anti-inflammatory and anti-apoptotic effects that are separate from its anti-bacterial role and that are mediated by inhibition of nitric oxide synthase [25], matrix metalloproteinases [26], and caspases [27]. As will be discussed below, minocycline and other tetracycline derivatives may also be direct inhibitors of human PAD4 [28] and *P. gingivalis* arginine-gingipains [29], which are potent proteinases and major virulence factors in periodontal disease.

Human peptidylarginine deiminases in disease

In normal physiology, PAD enzymes are involved in regulatory processes such as epidermal differentiation, maturation of hair follicles, insulation of nerve fibers, and epigenetic regulation. Aberrant citrullination contributes

to skin diseases such as psoriasis and neurological disorders such as multiple sclerosis, Alzheimer disease, and prion disease [30-32]. Citrullination of histones and other nuclear proteins by PAD4 is involved in transcriptional regulation and response to cellular stresses and contributes to the innate immune response through the formation of neutrophil extracellular traps [33-36]. Recently, citrullination of various chemokines has been shown to have functional roles in receptor binding and signalling, proteolytic cleavage, and extravasation of neutrophils [37,38]. Furthermore, citrullination appears to play a role in the coagulation system and associated pathways, and this is supported by the findings that *in vitro* citrullinated fibrinogen shows impaired thrombin-catalyzed fibrin polymerization [39] and *in vitro* citrullination of antithrombin with PAD4 abolishes its thrombin-inhibitory activity [40]. Both citrullinated fibrin(ogen) and citrullinated antithrombin were detected in patients with inflammatory arthritis [40,41].

Citrullination is thus a widespread phenomenon in normal physiology and inflammation, although targeting citrullinated proteins for an autoimmune response is relatively restricted to RA as shown by the high specificity of anti-citrullinated peptide antibodies for RA [2]. Therefore, it is important to consider which of the deiminases are used for generating the antigens that drive this autoimmunity. On the transcriptional level, various single-nucleotide polymorphisms in the *PADI4* gene have been associated with RA in Asian but not in Caucasian populations (reviewed in [42]). Suzuki and colleagues [43] showed that the presence of the disease-associated *PADI4* haplotype led to a more stable mRNA, which they suggested increased PAD4 expression and thus levels of citrullinated proteins. However, as PAD inhibitors would work on the post-transcriptional level, we will focus on the expression of PAD enzymes. PAD2 and PAD4 expression has been demonstrated in rheumatoid synovium [44] and synovial fluid cells [45] and extracellularly in synovial fluid [46]. PAD4 differs from other PAD isotypes in its capacity to undergo nuclear translocation due to the presence of a nuclear localization sequence and this translocation has been shown to be induced by tumor necrosis factor-alpha in murine and human oligodendroglial cell lines [47]. PAD expression in the synovial tissue is not specific to RA. It occurs in a variety of inflammatory synovitides [41] and diseases such as inflammatory bowel disease, polymyositis, and interstitial pneumonia [48]. While PAD2 is expressed in the synovia of both patients with inflammatory arthritis and osteoarthritis (OA), PAD4 is predominantly expressed in the synovia of patients with inflammatory arthritides rather than OA [44]. The converse was observed in the extracellular compartment, where Kinloch and colleagues [46] showed the presence of PAD4 in the synovial fluid of patients with RA, spondyloathropathies, and OA, while PAD2 expression was found in both groups of patients with inflammatory arthritis but was notably absent in those with OA. PAD2 and PAD4 expression in the synovium correlates with inflammatory cell infiltration, synovial lining thickness, and vascularity of the deep synovium [44]. Foulquier and colleagues [44] demonstrated PAD2 and PAD4 in close proximity to citrullinated fibrin deposits, although simultaneous detection of the two enzymes in the same area was rare.

Bacterial peptidylarginine deiminase

P. gingivalis, considered a primary pathogen in chronic periodontitis, is a Gram-negative, non-motile anaerobic bacterium that is the only prokaryote described to date to express a functional endogenous PAD enzyme [4]. To date, investigations of bacterial deiminases have focused mainly on enzymes that use free, non-peptidyl arginine or arginine derivatives such as arginine deiminase (ADI). ADIs are enzymes that catalyze the deimination of free arginine to citrulline, releasing ammonia. They are key enzymes in the widespread anaerobic pathway of arginine degradation and many pathogenic microorganisms use this pathway for energy production. Since ADIs are missing in higher eukaryotes, the enzyme constitutes a potential anti-parasitic and anti-bacterial drug target [49]. The other group of structurally and functionally related enzymes produced by most bacterial species consists of agmatine deiminases (agmatine iminohydrolases, or AIHs). AIHs deiminate agmatine (a decarboxylation product of arginine) to *N*-carbamoylputrescine and ammonia.

On the amino acid sequence level, PPAD shows no relation to eukaryotic PAD; instead, position-specific iterative-basic local alignment search tool (PSI-BLAST) search connects PPAD to the AIH family (Figure 2). Although the molecular structure of PPAD is unknown, its sequence similarity to AIHs with conservation of key catalytic and guanidino-binding residues indicates that the catalytic domain shares the common α/β-propeller fold of the guanidine-group modifying enzyme (GME) superfamily, which includes human PADs, microbial ADI, aminotransferases, dimethylarginine dimethylaminohydrolases, and AIH [50]. Of note, the database annotation of AIH is confusing since these enzymes are often referred to as 'Porphyromonas-type peptidyl-arginine deiminases' although they most likely do not possess PPAD activity. The three-dimensional structure of PPAD was predicted to consist of the amino-terminal catalytic α/β-propeller domain, followed by an immunoglobulin-like β sandwich. In comparison, the published structure of human PAD4 is composed of two amino-terminal immunoglobulin-like β sandwich domains, followed by the catalytic α/β-propeller domain [51].

Figure 2. Alignment of amino acid sequence of catalytic domains of *Porphyromonas gingivalis* PAD (PPAD) (residues 86 to 363), AIH from *Dyadobacter fermentans* DSM 18053 (locus Dfer_2630, residues 60 to 352), and human PAD4 (residues 306 to 556). Residues identical in PPAD and AIH and/or PAD4 are highlighted. Guanidino-binding (#) and catalytic residues (*) that are conserved in all families of guanidino-group modifying enzyme superfamily are indicated. The amino-terminal sequence of each enzyme is unique. In PAD4, the amino-terminal portion is folded into two consecutive immunoglobulin-like β-sandwich domains preceding the catalytic domain harboring the α/β-propeller fold [51]. A long 200-residue carboxy-terminal extension of PPAD is predicted to adopt an immunoglobulin-like β-sandwich structure [50]. AIH, agmatine iminohydrolase; PAD, peptidylarginine deiminase.

Unlike mammalian enzymes, PPAD is able to deiminate both free arginine and peptidylarginine ([4] and our own unpublished observations) and preferentially targets carboxy-terminal arginine, although internal citrullination cannot be excluded. Furthermore, deimination by human PAD is calcium-dependent in contrast to that by PPAD, which does not appear to require any specific cofactors ([4,52] and our own unpublished observations). *P. gingivalis* has been shown, however, to increase intracellular calcium concentrations by cleavage of proteinase-activated receptor 2 (PAR 2), a G protein-coupled receptor found on the neutrophil surface, which may in turn promote human PAD activation [53].

The physiological role of PPAD is unclear. It was suggested that production of ammonia during deimination enhances the survival of *P. gingivalis* within the periodontal pocket [4]. Indeed, ADI- and AIH-catalyzed ammonium production among bacterial species is known to act as a virulence factor, promoting the survival of microbial pathogens in the host environment. Ammonia neutralizes acidic environments and thereby optimizes gingipain and PPAD function, inactivates hemagglutinins, promotes ATP production, and has negative effects on neutrophil function [4,54]. Furthermore, it can be speculated that PPAD acts as a virulence factor by generating citrullinated peptides, which may assist the bacterium in spreading and circumventing the humoral immune response. However, the requirements for citrullination by PPAD have not been well investigated to date and it is unknown whether the citrullinated peptides are immunogenic.

Thus, we conclude that PPAD may be more relevant to the initiation of autoimmunity at a site distant from the joint, such as the gingiva, and that PAD2 and PAD4 are important in generating autoantigens that perpetuate autoimmunity in RA once tolerance is breached. Further work is required to identify the regulation and substrate specificity of each enzyme in order to establish a more precise role in the autoimmune response.

Therapeutic peptidylarginine deiminase blockade in rheumatoid arthritis

Although PAD4 has been most extensively studied as a potential therapeutic target in RA (mainly based on the availability of a crystal structure [51]), PAD2 may also be important. It is proposed that selective inhibition of PAD would reduce the levels of citrullinated proteins and consequently suppress the humoral immune response directed to citrullinated antigens in RA. Because PAD4 has an important physiological role in regulating gene expression and PAD4 translocates into the nucleus from the cytosol, potential inhibitors may need to be selective for the extracellular compartment or other PAD isotypes to avoid unwanted effects on gene transcription. It is, however, not known whether intracellular or extracellular PAD is important in the pathophysiology of RA.

Paclitaxel is a chemotherapeutic agent that was initially derived from the bark of the Pacific yew tree. It inhibits angiogenesis by interfering with microtubule function in cell mitosis, migration, chemotaxis, and intracellular transport [55]. In addition, in the millimolar range (half-maximal inhibitory concentration [IC_{50}] = approximately 5 mM), paclitaxel inhibits PAD isolated from bovine brain [56]. It has been shown to prevent the induction of collagen-induced arthritis (CIA) and cause significant regression of existing CIA [57]. An open-label multicenter phase II study of paclitaxel in patients with RA was completed in July 2008, although results of this are still pending [58].

Other PAD inhibitors include F-amidine [*N*-α-benzoyl-N^5-(2-fluoro-1-iminoethyl)-L-ornithine amide], Cl-amidine [*N*-α-benzoyl-N^5-(2-chloro-1-iminoethyl)-L-ornithine amide], and 2-chloroacetamidine, of which Cl-amidine was reported to be the most potent (IC_{50} = 5.9 μM) [59]. *Ex vivo* studies with F-amidine and Cl-amidine, using a cell line and an assay measuring PAD4-mediated citrullination of a nuclear protein and the resulting enhancement in binding to another protein, indicated that these inhibitors are bioavailable [59,60]. F-amidine irreversibly inhibits PAD4 via the specific modification of Cys 645, an active-site residue that is critical for enzyme catalysis. Cys 645 acts as a nucleophile to form a thiouronium intermediate that is hydrolyzed to form citrulline. Cl-amidine and 2-chloroacetamidine are thought to act via a similar mechanism [59,61]. Inactivation by F-amidine and Cl-amidine is calcium-dependent [60]. *In vitro* studies with PAD4 have shown that calcium binding leads to a conformational change that moves Cys 645 and His 471 into positions that are competent for catalysis [51] and presumably reactive with F-amidine and Cl-amidine. This is of therapeutic importance as these compounds would therefore be expected to inhibit PAD4 in its activated state only at sites of inflammatory activity such as the synovium and not the inactive enzyme at other sites in the body, limiting toxicity [59]. Willis and colleagues [62] recently showed that Cl-amidine treatment in CIA is able to inhibit clinical disease activity scores by 55%, 53%, and 42% in the 50, 10, and 1 mg/kg per day groups, respectively. Histological severity scores and complement C3 deposition scores paralleled the decreases in disease activity. In addition, mice receiving Cl-amidine showed reduced epitope spreading by peptide microarray, especially to citrullinated joint antigens. Interestingly, there were no changes in the percentages of T-cell, B-cell, or monocyte populations in treated mice compared with controls [62]. These results suggest that Cl-amidine may represent a novel class of RA therapeutics that specifically target citrullination.

Bhattacharya and colleagues [63] demonstrated that human astrocytes subject to pressure showed elevated PAD2 levels, increased intracellular calcium concentrations, and increased citrullination. Treatment with the cell-permeable calcium chelating agent BAPTA-AM (1,2-bis-(o-Aminophenoxy)-ethane-N,N,N',N'-tetraacetic acid tetraacetoxymethyl ester) resulted in decreased intracellular calcium concentration and PAD2 levels. These results suggest that calcium modulation may be an alternative therapeutic strategy in modulating PAD activity and citrullination, although we would argue that this mechanism is too broad to be applicable in practice.

On the basis of the therapeutic use of tetracyclines in RA [23], Knuckley and colleagues [28] screened tetracycline derivatives (minocycline, doxycycline, tetracycline, and chlortetracycline) for their potential to inhibit PAD4 activity. Chlortetracycline was identified as the most potent inhibitor (IC_{50} = 100 μM) and was suggested to bind to a region distal from the active site [28]. Streptomycin, an aminoglycoside antibiotic, was also tested because of its two guanidinium groups that could act as inhibitors of PAD4. Streptomycin was found to inhibit PAD4, though with a lower potency (IC_{50} = approximately 1.8 mM), and was suggested to bind within or in close proximity to the active site. The data suggest that these compounds could provide a valuable scaffold for engineering inhibitors with greater potency and selectivity.

Porphyromonas gingivalis peptidylarginine deiminase as a target for treatment in rheumatoid arthritis

The unique nature of PPAD in terms of its different amino acid sequence, cofactor requirement, and domain organization compared with human PADs (Figure 2), along with its location on the bacterial cell surface [4], would make this enzyme a potential target in the treatment of RA provided that its possible involvement in disease etiology or pathology is substantiated in future studies. Development of therapeutics targeting PPAD is

further encouraged by advances in design and synthesis of inhibitors against parasite-derived ADI with potentials to be used as anti-parasitic agents [64]. Since ADI, PADs, and PPAD are likely to use the same catalytic machinery to deiminate (peptidyl)arginine (Figure 2), a similar chemistry may be applied to develop PPAD inhibitors. The calcium-independent deimination of carboxy-terminal arginine residues specific to PPAD can be explored to develop highly selective compounds with little or no cross-reactivity with host enzymes.

McGraw and colleagues [4] reported that native PPAD, purified from the bacterial culture supernatant, was missing the N-terminus inferred from the DNA sequence because of proteolysis at the Arg43-Ala44 peptide bond. This might have been an artifact caused by the potent proteases, arginine-gingipains, which co-purified with PAD at the initial stages of protein purification, or might have true biological significance (for example, arising during export of the enzyme from the cell to form the mature protein). A recent paper on PPAD reported that the full-length, uncleaved form was unstable and had only 40% activity when compared with the truncated form of the enzyme [52]. Future studies aimed at identifying the mature, *in vivo* form of PPAD and its enzymology are required in order to pin down the biologically relevant form of the enzyme and as such the more appropriate target for therapeutic blockade.

As the PPAD enzyme is not well studied, there are no published studies on possible therapeutic inhibitors. To gain insight into the catalytic mode of PPAD, McGraw and colleagues [4] tested various compounds that might interfere with the catalytic cysteine residue (Cys 351) or substrate binding. They reported that the serine- and cysteine-protease inhibitor leupeptin is able to completely inhibit PPAD at millimolar levels (5 mM), with other inhibitors such as thiourea, thio-L-citrulline, and the serine- and cysteine-protease inhibitor TLCK (*N*-alpha-p-tosyl-L-lysine chloromethyl ketone) being inhibitory at higher concentrations (12.5 to 50 mM) [4]. Apart from the relatively low inhibitory potency, these compounds are either toxic (thiourea) or unselective (thio-L-citrulline is a potent inhibitor of nitric oxide synthase) [65] but nonetheless provide a basis for the development of more potent, specific inhibitors.

Conclusions

We have summarized a possible role for PPAD in breaking tolerance to citrullinated proteins, with human PAD2 or PAD4 or both maintaining the generation of citrullinated antigens in the joint. However, the evidence remains speculative and clearly requires further investigation of the mechanisms of activity of the enzymes involved and how the apparently unique PAD encoded by *P. gingivalis* could generate immunogenic peptides. If these hypotheses are further substantiated, PAD blockade has the potential to switch off autoimmunity at the point of initiation and inhibit the maintenance of the pathology in RA. Thus, inhibition of bacterial and human PADs could become the first treatment targeting the generation of the actual antigens that drive the disease.

Abbreviations

ADI, arginine deiminase; AIH, agmatine iminohydrolase; CEP-1, citrullinated enolase peptide-1; CIA, collagen-induced arthritis; Cl-amidine, *N*-α-benzoyl-N^5-(2-chloro-1-iminoethyl)-L-ornithine amide; F-amidine, *N*-α-benzoyl-N^5-(2-fluoro-1-iminoethyl)-L-ornithine amide; IC_{50}, half-maximal inhibitory concentration (concentration of inhibitor that yields 50% inhibition); OA, osteoarthritis; PAD, peptidylarginine deiminase; PPAD, *Porphyromonas gingivalis* peptidylarginine deiminase; RA, rheumatoid arthritis.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the Arthritis Research Campaign (NW and PJV) and in part by grants DE 09761 and 1642/B/P01/2008/35 from the National Institutes of Health (Bethesda, MD, USA) and the Department of Scientific Research of the Polish Ministry of Science and Education, respectively (JP). The Faculty of Biochemistry, Biophysics, and Biotechnology of the Jagiellonian University Krakow is a recipient of structural funds from the European Union (grant number POIG.02.01.00-12-064/08 – 'Molecular biotechnology for health').

Author details

[1]The Kennedy Institute of Rheumatology Division, Imperial College, 65 Aspenlea Road, London, W6 8LH, UK. [2]Department of Microbiology, Faculty of Biochemistry, Biophysics, and Biotechnology, Jagiellonian University, ul. Gronostajowa 7, 30-387 Krakow, Poland. [3]School of Dentistry, Oral Health and Systemic Disease, University of Louisville, 501 South Preston Street, Louisville, KY 40202, USA.

Published: 2 June 2010

References

1. Wegner N, Lundberg K, Kinloch A, Fisher BA, Malmstrom V, Feldmann M, Venables PJ: **Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis.** *Immunol Rev* 2009, **233:**1-21.
2. Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC, van Venrooij WJ: **The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide.** *Arthritis Rheum* 2000, **43:**155-163.
3. Chavanas S, Mechin MC, Takahara H, Kawada A, Nachat R, Serre G, Simon M: **Comparative analysis of the mouse and human peptidylarginine deiminase gene clusters reveals highly conserved non-coding segments and a new human gene, PADI6.** *Gene* 2004, **330:**19-27.
4. McGraw WT, Potempa J, Farley D, Travis J: **Purification, characterization, and sequence analysis of a potential virulence factor from *Porphyromonas gingivalis*, peptidylarginine deiminase.** *Infect Immun* 1999, **67:**3248-3256.
5. Vossenaar ER, Smeets TJ, Kraan MC, Raats JM, van Venrooij WJ, Tak PP: **The presence of citrullinated proteins is not specific for rheumatoid synovial tissue.** *Arthritis Rheum* 2004, **50:**3485-3494.
6. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, Habibuw MR, Vandenbroucke JP, Dijkmans BA: **Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors.** *Arthritis Rheum* 2004, **50:**380-386.
7. Rosenstein ED, Greenwald RA, Kushner LJ, Weissmann G: **Hypothesis: the humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis.** *Inflammation* 2004, **28:**311-318.
8. Lundberg K, Kinloch A, Fisher BA, Wegner N, Wait R, Charles P, Mikuls TR, Venables PJ: **Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase.** *Arthritis Rheum* 2008, **58:**3009-3019.

9. de Pablo P, Chapple IL, Buckley CD, Dietrich T: **Periodontitis in systemic rheumatic diseases.** *Nat Rev Rheumatol* 2009, **5:**218-224.
10. Borrell LN, Burt BA, Taylor GW: **Prevalence and trends in periodontitis in the USA: the [corrected] NHANES, 1988 to 2000.** *J Dent Res* 2005, **84:**924-930.
11. de Pablo P, Dietrich T, McAlindon TE: **Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population.** *J Rheumatol* 2008, **35:**70-76.
12. Mercado F, Marshall RI, Klestov AC, Bartold PM: **Is there a relationship between rheumatoid arthritis and periodontal disease?** *J Clin Periodontol* 2000, **27:**267-272.
13. Socransky SS, Haffajee AD: **Periodontal microbial ecology.** *Periodontol 2000* 2005, **38:**135-187.
14. Tatakis DN, Kumar PS: **Etiology and pathogenesis of periodontal diseases.** *Dent Clin North Am* 2005, **49:**491-516.
15. Mikuls TR, Payne JB, Reinhardt RA, Thiele GM, Maziarz E, Cannella AC, Holers VM, Kuhn KA, O'Dell JR: **Antibody responses to *Porphyromonas gingivalis* (*P. gingivalis*) in subjects with rheumatoid arthritis and periodontitis.** *Int Immunopharmacol* 2009, **9:**38-42.
16. Moen K, Brun JG, Madland TM, Tynning T, Jonsson R: **Immunoglobulin G and A antibody responses to *Bacteroides forsythus* and *Prevotella intermedia* in sera and synovial fluids of arthritis patients.** *Clin Diagn Lab Immunol* 2003, **10:**1043-1050.
17. Moen K, Brun JG, Valen M, Skartveit L, Eribe EK, Olsen I, Jonsson R: **Synovial inflammation in active rheumatoid arthritis and psoriatic arthritis facilitates trapping of a variety of oral bacterial DNAs.** *Clin Exp Rheumatol* 2006, **24:**656-663.
18. Ogrendik M, Kokino S, Ozdemir F, Bird PS, Hamlet S: **Serum antibodies to oral anaerobic bacteria in patients with rheumatoid arthritis.** *MedGenMed* 2005, **7:**2.
19. Martinez-Martinez RE, Abud-Mendoza C, Patino-Marin N, Rizo-Rodriguez JC, Little JW, Loyola-Rodriguez JP: **Detection of periodontal bacterial DNA in serum and synovial fluid in refractory rheumatoid arthritis patients.** *J Clin Periodontol* 2009, **36:**1004-1010.
20. Ogrendik M: **Treatment of rheumatoid arthritis with ornidazole: a randomized, double-blind, placebo-controlled study.** *Rheumatol Int* 2006, **26:**1132-1137.
21. Ogrendik M: **Levofloxacin treatment in patients with rheumatoid arthritis receiving methotrexate.** *South Med J* 2007, **100:**135-139.
22. Ogrendik M: **Effects of clarithromycin in patients with active rheumatoid arthritis.** *Curr Med Res Opin* 2007, **23:**515-522.
23. O'Dell JR, Haire CE, Palmer W, Drymalski W, Wees S, Blakely K, Churchill M, Eckhoff PJ, Weaver A, Doud D, Erikson N, Dietz F, Olson R, Maloley P, Klassen LW, Moore GF: **Treatment of early rheumatoid arthritis with minocycline or placebo: results of a randomized, double-blind, placebo-controlled trial.** *Arthritis Rheum* 1997, **40:**842-848.
24. Kloppenburg M, Breedveld FC, Terwiel JP, Mallee C, Dijkmans BA: **Minocycline in active rheumatoid arthritis. A double-blind, placebo-controlled trial.** *Arthritis Rheum* 1994, **37:**629-636.
25. Amin AR, Attur MG, Thakker GD, Patel PD, Vyas PR, Patel RN, Patel IR, Abramson SB: **A novel mechanism of action of tetracyclines: effects on nitric oxide synthases.** *Proc Natl Acad Sci U S A* 1996, **93:**14014-14019.
26. Greenwald RA, Golub LM, Lavietes B, Ramamurthy NS, Gruber B, Laskin RS, McNamara TF: **Tetracyclines inhibit human synovial collagenase *in vivo* and *in vitro*.** *J Rheumatol* 1987, **14:**28-32.
27. Chen M, Ona VO, Li M, Ferrante RJ, Fink KB, Zhu S, Bian J, Guo L, Farrell LA, Hersch SM, Hobbs W, Vonsattel JP, Cha JH, Friedlander RM: **Minocycline inhibits caspase-1 and caspase-3 expression and delays mortality in a transgenic mouse model of Huntington disease.** *Nat Med* 2000, **6:**797-801.
28. Knuckley B, Luo Y, Thompson PR: **Profiling Protein Arginine Deiminase 4 (PAD4): a novel screen to identify PAD4 inhibitors.** *Bioorg Med Chem* 2008, **16:**739-745.
29. Imamura T, Matsushita K, Travis J, Potempa J: **Inhibition of trypsin-like cysteine proteinases (gingipains) from Porphyromonas gingivalis by tetracycline and its analogues.** *Antimicrob Agents Chemother* 2001, **45:**2871-2876.
30. Jang B, Kim E, Choi JK, Jin JK, Kim JI, Ishigami A, Maruyama N, Carp RI, Kim YS, Choi EK: **Accumulation of citrullinated proteins by up-regulated peptidylarginine deiminase 2 in brains of scrapie-infected mice: a possible role in pathogenesis.** *Am J Pathol* 2008, **173:**1129-1142.
31. Gyorgy B, Toth E, Tarcsa E, Falus A, Buzas EI: **Citrullination: a posttranslational modification in health and disease.** *Int J Biochem Cell Biol* 2006, **38:**1662-1677.
32. Young DS, Meersman F, Oxley D, Webster J, Gill AC, Bronstein I, Lowe CR, Dear DV: **Effect of enzymatic deimination on the conformation of recombinant prion protein.** *Biochim Biophys Acta* 2009, **1794:**1123-1133.
33. Neeli I, Dwivedi N, Khan S, Radic M: **Regulation of extracellular chromatin release from neutrophils.** *J Innate Immun* 2009, **1:**194-201.
34. Wang Y, Li M, Stadler S, Correll S, Li P, Wang D, Hayama R, Leonelli L, Han H, Grigoryev SA, Allis CD, Coonrod SA: **Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation.** *J Cell Biol* 2009, **184:**205-213.
35. Tanikawa C, Ueda K, Nakagawa H, Yoshida N, Nakamura Y, Matsuda K: **Regulation of protein Citrullination through p53/PADI4 network in DNA damage response.** *Cancer Res* 2009, **69:**8761-8769.
36. Neeli I, Khan SN, Radic M: **Histone deimination as a response to inflammatory stimuli in neutrophils.** *J Immunol* 2008, **180:**1895-1902.
37. Proost P, Loos T, Mortier A, Schutyser E, Gouwy M, Noppen S, Dillen C, Ronsse I, Conings R, Struyf S, Opdenakker G, Maudgal PC, Van Damme J: **Citrullination of CXCL8 by peptidylarginine deiminase alters receptor usage, prevents proteolysis, and dampens tissue inflammation.** *J Exp Med* 2008, **205:**2085-2097.
38. Loos T, Mortier A, Gouwy M, Ronsse I, Put W, Lenaerts JP, Van Damme J, Proost P: **Citrullination of CXCL10 and CXCL11 by peptidylarginine deiminase: a naturally occurring posttranslational modification of chemokines and new dimension of immunoregulation.** *Blood* 2008, **112:**2648-2656.
39. Okumura N, Haneishi A, Terasawa F: **Citrullinated fibrinogen shows defects in FPA and FPB release and fibrin polymerization catalyzed by thrombin.** *Clin Chim Acta* 2009, **401:**119-123.
40. Chang X, Yamada R, Sawada T, Suzuki A, Kochi Y, Yamamoto K: **The inhibition of antithrombin by peptidylarginine deiminase 4 may contribute to pathogenesis of rheumatoid arthritis.** *Rheumatology (Oxford)* 2005, **44:**293-298.
41. Chapuy-Regaud S, Sebbag M, Baeten D, Clavel C, Foulquier C, De Keyser F, Serre G: **Fibrin deimination in synovial tissue is not specific for rheumatoid arthritis but commonly occurs during synovitides.** *J Immunol* 2005, **174:**5057-5064.
42. Suzuki A, Yamada R, Yamamoto K: **Citrullination by peptidylarginine deiminase in rheumatoid arthritis.** *Ann N Y Acad Sci* 2007, **1108:**323-339.
43. Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K: **Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis.** *Nat Genet* 2003, **34:**395-402.
44. Foulquier C, Sebbag M, Clavel C, Chapuy-Regaud S, Al Badine R, Méchin MC, Vincent C, Nachat R, Yamada M, Takahara H, Simon M, Guerrin M, Serre G: **Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation.** *Arthritis Rheum* 2007, **56:**3541-3553.
45. Vossenaar ER, Radstake TR, van der Heijden A, van Mansum MA, Dieteren C, de Rooij DJ, Barrera P, Zendman AJ, van Venrooij WJ: **Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages.** *Ann Rheum Dis* 2004, **63:**373-381.
46. Kinloch A, Lundberg K, Wait R, Wegner N, Lim NH, Zendman AJ, Saxne T, Malmstrom V, Venables PJ: **Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis.** *Arthritis Rheum* 2008, **58:**2287-2295.
47. Mastronardi FG, Wood DD, Mei J, Raijmakers R, Tseveleki V, Dosch HM, Probert L, Casaccia-Bonnefil P, Moscarello MA: **Increased citrullination of histone H3 in multiple sclerosis brain and animal models of demyelination: a role for tumor necrosis factor-induced peptidylarginine deiminase 4 translocation.** *J Neurosci* 2006, **26:**11387-11396.
48. Makrygiannakis D, af Klint E, Lundberg IE, Lofberg R, Ulfgren AK, Klareskog L, Catrina AI: **Citrullination is an inflammation-dependent process.** *Ann Rheum Dis* 2006, **65:**1219-1222.
49. Galkin A, Kulakova L, Sarikaya E, Lim K, Howard A, Herzberg O: **Structural insight into arginine degradation by arginine deiminase, an antibacterial and parasite drug target.** *J Biol Chem* 2004, **279:**14001-14008.
50. Shirai H, Mokrab Y, Mizuguchi K: **The guanidino-group modifying enzymes: structural basis for their diversity and commonality.** *Proteins* 2006, **64:**1010-1023.
51. Arita K, Hashimoto H, Shimizu T, Nakashima K, Yamada M, Sato M: **Structural**

basis for Ca(2+)-induced activation of human PAD4. *Nat Struct Mol Biol* 2004, **11:**777-783.
52. Rodriguez SB, Stitt BL, Ash DE: Expression of peptidylarginine deiminase from Porphyromonas gingivalis in *Escherichia coli*: enzyme purification and characterization. *Arch Biochem Biophys* 2009, **488:**14-22.
53. Lourbakos A, Chinni C, Thompson P, Potempa J, Travis J, Mackie EJ, Pike RN: Cleavage and activation of proteinase-activated receptor-2 on human neutrophils by gingipain-R from *Porphyromonas gingivalis*. *FEBS Lett* 1998, **435:**45-48.
54. Niederman R, Brunkhorst B, Smith S, Weinreb RN, Ryder MI: Ammonia as a potential mediator of adult human periodontal infection: inhibition of neutrophil function. *Arch Oral Biol* 1990, **35 Suppl:**205S-209S.
55. Belotti D, Rieppi M, Nicoletti MI, Casazza AM, Fojo T, Taraboletti G, Giavazzi R: Paclitaxel (Taxol(R)) inhibits motility of paclitaxel-resistant human ovarian carcinoma cells. *Clin Cancer Res* 1996, **2:**1725-1730.
56. Pritzker LB, Moscarello MA: A novel microtubule independent effect of paclitaxel: the inhibition of peptidylarginine deiminase from bovine brain. *Biochim Biophys Acta* 1998, **1388:**154-160.
57. Brahn E, Tang C, Banquerigo ML: Regression of collagen-induced arthritis with taxol, a microtubule stabilizer. *Arthritis Rheum* 1994, **37:**839-845.
58. ClinicalTrials.gov: search of: paclitaxel AND rheumatoid: a study using intravenous paxceed to treat patients with rheumatoid arthritis [http://clinicaltrials.gov/ct2/results?term=paclitaxel+AND+rheumatoid].
59. Luo Y, Arita K, Bhatia M, Knuckley B, Lee YH, Stallcup MR, Sato M, Thompson PR: Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization. *Biochemistry* 2006, **45:**11727-11736.
60. Luo Y, Knuckley B, Lee YH, Stallcup MR, Thompson PR: A fluoroacetamidine-based inactivator of protein arginine deiminase 4: design, synthesis, and in vitro and in vivo evaluation. *J Am Chem Soc* 2006, **128:**1092-1093.
61. Stone EM, Schaller TH, Bianchi H, Person MD, Fast W: Inactivation of two diverse enzymes in the amidinotransferase superfamily by 2-chloroacetamidine: dimethylargininase and peptidylarginine deiminase. *Biochemistry* 2005, **44:**13744-13752.
62. Willis V, Gizinski A, Knuckley B, Banda N, Cordova K, Luo Y, Causey C: Treatment with Cl-amidine, a peptidyl arginine deiminase (PAD) inhibitor significantly reduces collagen-induced arthritis (CIA) [abstract]. *Arthritis Rheum* 2009, **60(suppl 10):**665.
63. Bhattacharya SK, Bhat MB, Takahara H: Modulation of peptidyl arginine deiminase 2 and implication for neurodegeneration. *Curr Eye Res* 2006, **31:**1063-1071.
64. Li L, Li Z, Chen D, Lu X, Feng X, Wright EC, Solberg NO, Dunaway-Mariano D, Mariano PS, Galkin A, Kulakova L, Herzberg O, Green-Church KB, Zhang L: Inactivation of microbial arginine deiminases by L-canavanine. *J Am Chem Soc* 2008, **130:**1918-1931.
65. Frey C, Narayanan K, McMillan K, Spack L, Gross SS, Masters BS, Griffith OW: L-thiocitrulline. A stereospecific, heme-binding inhibitor of nitric-oxide synthases. *J Biol Chem* 1994, **269:**26083-26091.
66. Nachat R, Mechin MC, Takahara H, Chavanas S, Charveron M, Serre G, Simon M: Peptidylarginine deiminase isoforms 1-3 are expressed in the epidermis and involved in the deimination of K1 and filaggrin. *J Invest Dermatol* 2005, **124:**384-393.
67. Nachat R, Mechin MC, Charveron M, Serre G, Constans J, Simon M: Peptidylarginine deiminase isoforms are differentially expressed in the anagen hair follicles and other human skin appendages. *J Invest Dermatol* 2005, **125:**34-41.
68. Mechin MC, Enji M, Nachat R, Chavanas S, Charveron M, Ishida-Yamamoto A, Serre G, Takahara H, Simon M: The peptidylarginine deiminases expressed in human epidermis differ in their substrate specificities and subcellular locations. *Cell Mol Life Sci* 2005, **62:**1984-1995.
69. Sambandam T, Belousova M, Accaviti-Loper MA, Blanquicett C, Guercello V, Raijmakers R, Nicholas AP: Increased peptidylarginine deiminase type II in hypoxic astrocytes. *Biochem Biophys Res Commun* 2004, **325:**1324-1329.
70. Ishigami A, Ohsawa T, Hiratsuka M, Taguchi H, Kobayashi S, Saito Y, Murayama S, Asaga H, Toda T, Kimura N, Maruyama N: Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. *J Neurosci Res* 2005, **80:**120-128.
71. Urano Y, Watanabe K, Sakaki A, Arase S, Watanabe Y, Shigemi F, Takeda K, Akiyama K, Senshu T: Immunohistochemical demonstration of peptidylarginine deiminase in human sweat glands. *Am J Dermatopathol* 1990, **12:**249-255.
72. Ishigami A, Ohsawa T, Asaga H, Akiyama K, Kuramoto M, Maruyama N: Human peptidylarginine deiminase type II: molecular cloning, gene organization, and expression in human skin. *Arch Biochem Biophys* 2002, **407:**25-31.
73. Vossenaar ER, Zendman AJ, van Venrooij WJ, Pruijn GJ: PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* 2003, **25:**1106-1118.
74. Hagiwara T, Nakashima K, Hirano H, Senshu T, Yamada M: Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem Biophys Res Commun* 2002, **290:**979-983.

doi:10.1186/ar3000
Cite this article as: Mangat P, *et al.*: Bacterial and human peptidylarginine deiminases: targets for inhibiting the autoimmune response in rheumatoid arthritis? *Arthritis Research & Therapy* 2010, **12:**209.

OPINION

Periodontitis in RA—the citrullinated enolase connection

Karin Lundberg, Natalia Wegner, Tülay Yucel-Lindberg and Patrick J. Venables

Abstract | Autoimmunity in rheumatoid arthritis (RA) is characterized by an antibody response to citrullinated proteins. Two of the risk factors for RA—*HLA-DRB1* shared epitope alleles and smoking—are also associated with periodontitis, which is largely, but not exclusively, caused by *Porphyromonas gingivalis* infection. Furthermore, RA and periodontitis have a similar pathophysiology, characterized by destructive inflammation. The citrullination of proteins by *P. gingivalis* and the subsequent generation of autoantigens that drive autoimmunity in RA represents a possible causative link between these two diseases. Antibodies directed towards the immunodominant epitope of human citrullinated α-enolase cross-react with a conserved sequence on citrullinated *P. gingivalis* enolase. On the basis of this cross-reactivity, in this Perspectives article we explore the hypothesis of molecular mimicry in the etiology of RA, with citrullinated enolase as the specific antigen involved.

Lundberg, K. *et al. Nat. Rev. Rheumatol.* **6**, 727–730 (2010); published online 7 September 2010; doi:10.1038/nrrheum.2010.139

Introduction

The etiological paradigm of rheumatoid arthritis (RA) is that an environmental agent triggers an autoimmune reaction in a genetically predisposed individual. The discovery of citrulline-specific autoimmunity could be the key to our understanding of these gene–environment interactions and hence to the cause(s) of RA.[1,2] Citrullination is the enzymatic conversion of peptidylarginine to peptidylcitrulline, a post-translational modification catalyzed by a family of calcium-dependent peptidylarginine deiminases (PADs). Although physiological citrullination of specific substrates occurs in a variety of cells and tissues in healthy individuals, and pathological citrullination seems to be a general phenomenon associated with inflammation, tolerance to citrullinated proteins is selectively lost in patients with RA. Hence, the role of citrulline-specific autoimmunity in RA has led to the suggestion (first proposed by Rosenstein *et al.*[3]), that the bacterium *Porphyromonas gingivalis*, which expresses an enzyme with PAD-activity (denoted PPAD) and is a major cause of periodontitis,[4] is also involved in the etiology of RA. In this Perspectives article, we propose that the association between RA and periodontitis is causative, based on the hypothesis that infection with *P. gingivalis* reduces tolerance to citrullinated antigens, and that one antigen in particular, citrullinated α-enolase, is central to the initiation of the pathogenic pathway leading to RA (Figure 1).

Competing interests
The authors declare no competing interests.

Links between RA and periodontitis

A number of studies have demonstrated an association between RA and periodontitis and have been reviewed elsewhere.[3,5–7] RA is more prevalent in individuals with periodontitis and vice versa in comparison to the general population, and one study showed that patients with RA were more likely than individuals without RA to be edentulous and to have periodontitis, independent of age, sex, ethnicity and smoking habits.[8] In addition to an epidemiological link between RA and periodontitis, the two diseases have a similar pathobiology.[5] The chronic inflammation associated with both diseases is characterized by a dominance of proinflammatory cytokines, notably tumor necrosis factor (TNF), leading to the upregulation of matrix metalloproteinases that cause tissue destruction and erosion of periarticular bone and periodontal bone, respectively.[2,5,6] This similarity between RA and periodontitis could arise from shared susceptibility genes and environmental factors leading to a comparable dysregulation of the host immune response.

Notwithstanding the similarities in their pathophysiology, the two diseases probably differ in their inflammatory effector mechanisms. Periodontitis can be thought of as an infectious disease caused by a network of interacting oral pathogens, dominated by *P. gingivalis*. However, since periodontitis has a strong genetic background,[9] it can also be considered as an inflammatory disease, resulting from an abnormal host defense in response to infection. By contrast, evidence is accumulating that RA is a true autoimmune disease where, in the majority of patients, inflammation is driven by pathogenic autoantibodies to citrullinated proteins.[1] Tissue damage is thought to be mediated by immune complexes binding to Fcγ receptors, leading to TNF secretion.[2]

Risk factors for RA and periodontitis

Genetic factors are thought to contribute considerably (~50%) to the etiology of both RA and periodontitis.[9,11] Well-established risk factors for RA, such as *HLA-DRB1* shared epitope (SE) alleles, the 620W polymorphism of the protein tyrosine phosphatase, non-receptor type 22 gene (*PTPN22*), and cigarette smoking, are found predominantly in the subset of patients with antibodies to citrullinated proteins.[1] These antibodies are commonly detected using a commercial enzyme-linked immunosorbent assay based on synthetic cyclic citrullinated peptides (CCPs), with high diagnostic specificity (98%) for RA.[1] However, CCPs are not physiological autoantigens, and the peptides used in the CCP assay do not show sequence homology with any known human protein. This assay is simply a generic test for the presence of antibodies to a number of possible citrullinated proteins present *in vivo*. At least four citrullinated autoantigens in RA have now been defined—citrullinated fibrinogen, vimentin, collagen type II, and α-enolase, with many others awaiting full characterization.[2]

We have shown that approximately half of the anti-CCP-positive patients with RA have antibodies to citrullinated α-enolase peptide 1 (CEP-1).[10] In a case–control analysis of 1,000

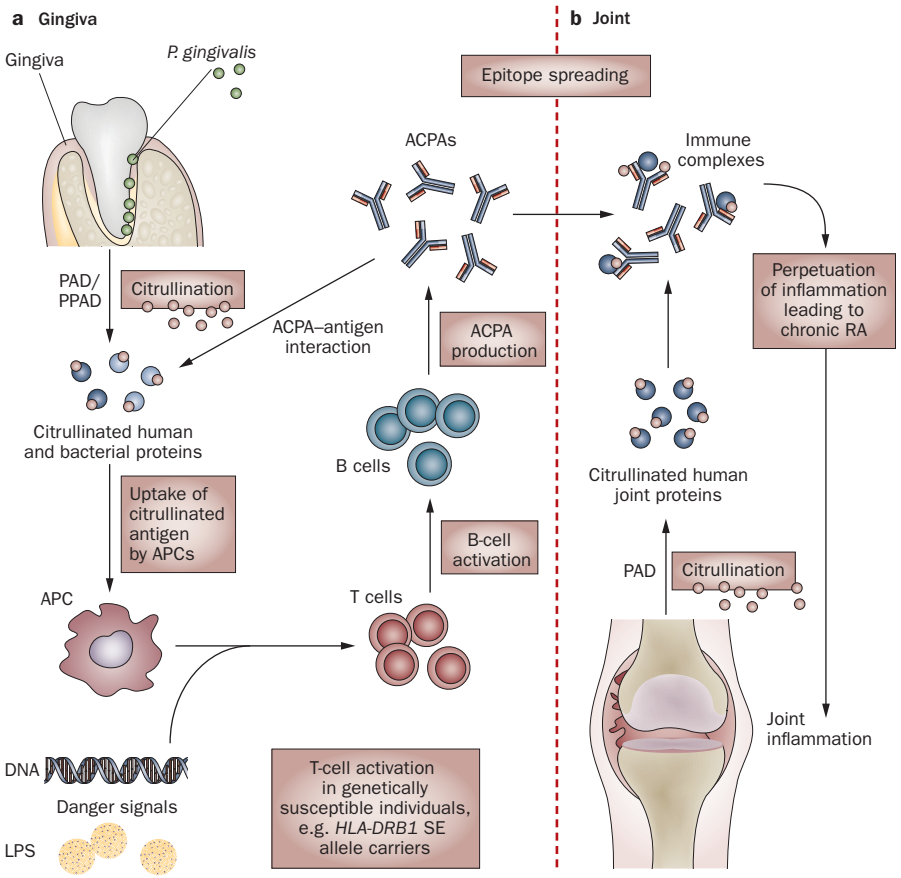


Figure 1 | Schematic illustration of the etiological hypothesis for *P. gingivalis* and citrullinated α-enolase involvement in RA. **a** | Infection by *P. gingivalis* leads to citrullination in the gingiva of bacterial and/or human proteins (by either PPAD, human PAD, or both). In the presence of danger signals, such as LPS and DNA, pathogenic T cells are activated by APCs (citrullinated antigen in the context of *HLA-DRB1* SE). T-cell-mediated activation of pathogenic B cells results in the production of antibodies specific for citrullinated proteins. **b** | A second inflammatory event occurs in the joint, leading to citrullination of joint proteins and the formation of immune complexes in the joint (through epitope spreading, cross-reactivity with citrullinated joint proteins, or both). The resulting perpetuation of the inflammatory process eventually causes chronic RA. Abbreviations: ACPA, anti-citrullinated protein/peptide antibody; APC, antigen presenting cell; LPS, lipopolysaccharide; *P. gingivalis*, *Porphyromonas gingivalis*; PAD, peptidylarginine deiminases; PPAD, *P. gingivalis* PAD; RA, rheumatoid arthritis; SE, shared epitope.

patients with RA and 872 healthy controls, we compared the CEP-1$^+$/CCP$^+$ group with the CEP-1$^-$/CCP$^+$ subset, and demonstrated that *HLA-DRB1* SE alleles, 620W *PTPN22*, and cigarette smoking, each constitute a risk for the development of antibodies to CEP-1, rather than to CCP.[12] The combined effect of these three risk factors showed an odds ratio of 39.0 (95% CI 15.6–98.0) for the development of CEP-1$^+$/CCP$^+$ disease, compared with an odds ratio of 4.3 (95% CI 2.2–8.6) for CEP-1$^-$/CCP$^+$ disease.[12]

By contrast, less-definitive evidence has come from genetic association studies of patients with periodontitis. Several cytokine and Fcγ receptor gene variants have been identified, each of which confers a relatively small risk of periodontitis, although no consensus on the risk of these variants has been reached to date.[13] The *HLA-DRB1* SE alleles have been associated with periodontitis;[14–16] however, the relevant studies were small in scale and larger cohorts, using uniform disease classification criteria, are needed before this genetic link can be confirmed. Cigarette smoking is another risk factor common to both RA and periodontitis. Smokers have a greater risk of developing severe periodontitis and bone loss compared with nonsmokers.[17] Importantly, evidence exists that the association between RA and periodontitis is independent of smoking,[8] although few studies have adjusted for this confounding factor.

P. gingivalis and citrullination

The Gram-negative bacterium *P. gingivalis* is commonly present, together with other oral bacteria, as a biofilm in the gingival crevice and intracellularly in oral epithelial cells. Notably, periodontitis is not always caused by *P. gingivalis*, and conversely, *P. gingivalis* infection does not always result in periodontitis. However, higher numbers of bacteria are present in the gingival crevice of patients with periodontitis than in individuals without periodontitis.[18] *P. gingivalis* produces an array of virulence factors, particularly extracellular cysteine proteases called gingipains, to support its adherence, growth, tissue invasion and tissue degradation, and to evade and subvert the host immune system.[19]

The most striking feature of this bacterium, in the context of autoimmunity in RA, is that *P. gingivalis* expresses its own unique citrullinating enzyme, PPAD.[4] PPAD has no sequence homology with human PADs, and seems to differ in substrate-specificity by showing a marked preference for the citrullination of carboxy-terminal arginines.[4] Studies in our laboratory have shown that *P. gingivalis* contains a range of endogenous citrullinated proteins that are not present in other common oral pathogens.[20] We have also, for the first time to our knowledge, demonstrated citrullination of human fibrinogen and human α-enolase by *P. gingivalis*.[20] Citrullination of these proteins was dependent on PPAD, but also on arginine-gingipains, which proteolytically cleave proteins at arginine residues, thus generating short peptides containing carboxy-terminal arginine residues that are subsequently citrullinated by PPAD.[20]

The enolase connection

We first identified citrullinated α-enolase as a candidate autoantigen by immunoblotting HL60 cell lysates with serum samples from patients with RA, and sequencing a reactive 47 kDa band by mass spectrometry.[21] In a subsequent study,[10] we screened serum from 102 patients with RA, 110 patients with other rheumatic diseases and 92 healthy controls for reactivity with 11 cyclic CEPs, covering 15 of the 17 arginine residues within α-enolase. We went on to map the immunodominant B-cell epitope(s) to the CEP-1 peptide, which corresponds to amino acids 5–21 in the full-length protein. CEP-1 contains two arginine residues replaced by citrulline, and we demonstrated that the second of these (citrulline 15) was most important for antibody recognition.[10] Reactivity against the arginine-containing control peptide was similarly low in patients and controls. A quantitative enzyme-linked immunosorbent

assay using CEP-1 has now become the basis of several published studies of over 3,000 serum samples, which have demonstrated an anti-CEP-1 antibody frequency of approximately 40% in patients with RA, and a disease-specificity of 97%.[10,12,22]

Although CEP-1 could possibly act as a surrogate marker for similar (linear) epitopes in other candidate autoantigens already proposed as being involved in the pathogenesis of RA, no significant sequence similarity was found with any other human protein outside the enolase family (α, β, and γ isoenzymes). However, several hits mapped to bacterial enolases, specifically *P. gingivalis* enolase.[10,23] Although full-length human α-enolase and *P. gingivalis* enolase are 51% identical, they have 82% sequence identity at the region corresponding to CEP-1 (Figure 2a). In addition, the nine amino acids at positions 13–21, which surround the second citrulline residue at position 15, are 100% identical. Furthermore, in patients with RA, the antibody response to the *P. gingivalis* version of CEP-1 correlated strongly with that to human CEP-1 ($r^2 = 0.8$, $P < 0.0001$).[10] More intriguing was the observation that anti-CEP-1 antibodies, affinity-purified from a patient with RA, cross-reacted with *in vitro* citrullinated recombinant *P. gingivalis* enolase (Figure 2b),[10] thus suggesting an etiological hypothesis based on molecular mimicry.

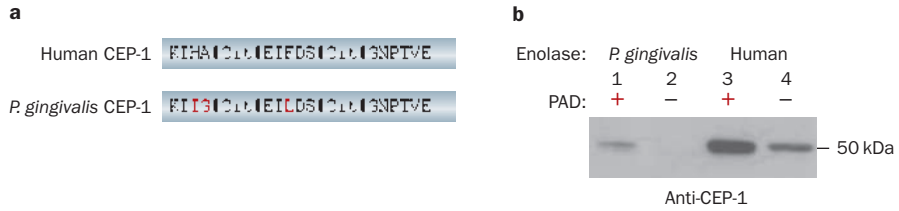


Figure 2 | *P. gingivalis* and citrullination. **a** | Human and *P. gingivalis* CEP-1 (amino acids 5–21 of the full length proteins) are 82% identical at the amino acid level. Residues that differ are highlighted in red. **b** | Human anti-CEP-1 antibodies, affinity-purified from a patient with RA, cross-react with *in vitro* citrullinated recombinant *P. gingivalis* enolase. *In vitro* citrullinated recombinant human α-enolase serves as a positive control (cross-reactivity of the anti-CEP-1 antibody with the native form of human α-enolase has previously been demonstrated[10]). Abbreviations: CEP-1, citrullinated α-enolase peptide 1; Cit, citrulline; *P. gingivalis*, *Porphyromonas gingivalis*; PAD, (rabbit) peptidylarginine deiminase. Panel 2b is reproduced with permission from John Wiley and Sons © Lundberg, K. *et al. Arthritis Rheum.* **58**, 3009–3019 (2008).

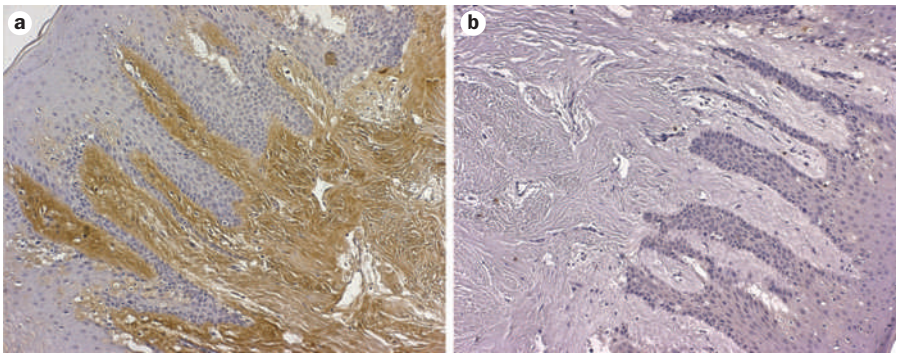


Figure 3 | Citrullinated proteins are present in the gingiva of a patient with periodontitis. Immunohistochemical staining of extracellular and intracellular proteins in the connective tissue of the gingiva using **a** | F95 IgM mouse monoclonal anticitrulline antibody and **b** | mouse IgM (as a control). Brown coloration indicates the presence of citrullinated proteins. Images are shown at ×100 magnification. Information regarding the development, the characterization and the use of the F95 antibody for immunohistochemistry is published in Nicholas *et al.*[30]

Etiological hypothesis

Expression of anti-CCP antibodies precedes the onset of clinical RA by several years.[24] In a study of 15 'pre-RA' serum samples selected on the basis of *HLA-DRB1* SE alleles and anti-CCP antibodies being present, we detected anti-CEP antibodies in nine samples, five of which were directed towards CEP-1.[25] This finding suggests that the initial breakdown of immunological tolerance to citrullinated proteins, including α-enolase, might occur as a consequence of an inflammatory event outside the joint. Furthermore, the presence of IgA CCP and IgA CEP-1[22] might indicate the involvement of mucosal surfaces, such as the gingiva. Hence, we hypothesize that anti-CEP-1 antibodies can be generated, as a consequence of *P. gingivalis*-induced citrullination of enolase, in the gingiva of genetically susceptible individuals.

We have demonstrated the presence of citrullinated proteins in gingival biopsies from patients with periodontitis (Figure 3).[26] However, we have not yet determined whether these proteins are of bacterial or human origin, whether they are citrullinated by bacterial or human PAD, or if this phenomenon is specific to periodontitis. Despite these limitations, it is intriguing to speculate that the presentation of peptides derived from these citrullinated proteins (in the context of *HLA-DRB1* SE molecules) during chronic exposure to danger signals, such as bacterial lipopolysacharides and DNA, might facilitate the activation of pathogenic T cells. *P. gingivalis* has been shown to preferentially stimulate T cells to produce interleukin-17, a cytokine implicated in the pathogenesis of RA.[7] Following B-cell activation with the subsequent production of anti-citrullinated protein antibodies, and a second inflammatory event in the synovium, mechanisms of epitope spreading or molecular mimicry, or both, could result in cross-reactivity with citrullinated joint proteins, such as citrullinated human α-enolase, perpetuating the inflammatory process into chronic RA (Figure 1). Notably, the initial T-cell-mediated and B-cell-mediated antibody response might be directed towards *P. gingivalis*-generated carboxy-terminal citrulline-containing epitopes, and this reactivity might be lost by the time RA develops, as a consequence of epitope spreading to epitopes containing internal citrulline, generated by human PADs.

An association between antibody titres to *P. gingivalis* and CCP has previously been demonstrated in a population of patients with RA.[27] The anti-*P. gingivalis* assay, which used a crude lysate of bacterial cells, could be criticized for a lack of specificity and the likely presence of citrullinated proteins in the extract. Nevertheless, the test can differentiate between individuals infected with *P. gingivalis* and those who are not.[28] By reanalyzing the data from two published studies on RA serum samples,[10,27] we have preliminary evidence for an association between the presence of antibodies to *P. gingivalis* and expression of anti-CEP-1 antibodies in RA (Lundberg, K. *et al.* unpublished data), indicating an association between the immune response to *P. gingivalis* and citrullinated α-enolase. However, these

findings could potentially be explained by cross-reactivity of CEP-1-specific antibodies with citrullinated epitopes in the *P. gingivalis* lysate used in the assay.

A study published in 2006 reported a somewhat higher than expected frequency of anti-CCP antibodies in patients with generalized aggressive periodontitis when compared with controls, though the difference was not statistically significant.[29] This potentially interesting finding needs to be confirmed and further investigated in large-scale periodontitis cohorts. On the basis of our data, we hypothesize that it would be beneficial to also analyze anti-CEP-1 and *HLA-DRB1* status in these cohorts. Likewise, the presence of *P. gingivalis* infection should be investigated, using more-specific antibacterial antibody assays, in well-characterized RA cohorts. We also propose that a similar analytical strategy, as outlined in our study on the specific interaction between genotype, smoking, and autoimmunity to citrullinated α-enolase in the etiology of RA,[12] should be used to attain a comprehensive picture of the genetic epidemiology of periodontitis-associated RA. Implementation of this strategy would address the possibility that *P. gingivalis* might be driving immune reactions that contribute to the development of RA in a large proportion of patients.

Conclusions

No direct evidence exists of a role for *P. gingivalis* in the development of RA. However, *P. gingivalis* is known to citrullinate both endogenous and human proteins, and RA-specific anti-CEP-1 antibodies cross-react with citrullinated *P. gingivalis* enolase. Furthermore, *HLA-DRB1* SE alleles and smoking, risk factors for both RA and periodontitis, are associated with the production of anti-CEP-1 antibodies. We also have preliminary evidence that citrullinated proteins are present in the gingiva of patients with periodontitis and that there is a link between the presence and the levels of anti-CEP-1 and anti-*P. gingivalis* antibodies in patients with RA. Therefore, we believe that further studies into the relationship between RA and periodontitis, with a focus on the molecular mechanisms linking *P. gingivalis* and citrulline immunity, are warranted and might reveal that the old hypothesis—that RA is precipitated by infection—is true, at least for a subgroup of patients.

*The Kennedy Institute of Rheumatology, Imperial College London, 65 Aspenlea Road, London W6 8LH, UK (**K. Lundberg, N. Wegner, P. J. Venables**). Department of Dental Medicine, Karolinska Institutet, SE-141 04, Huddinge, Sweden (**T. Yucel-Lindberg**).*

Correspondence to: P. J. Venables
p.venables@imperial.ac.uk

1. Klareskog, L., Rönnelid, J., Lundberg, K., Padyukov, L. & Alfredsson, L. Immunity to citrullinated proteins in rheumatoid arthritis. *Annu. Rev. Immunol.* **26**, 651–675 (2008).
2. Wegner, N. *et al.* Autoimmunity to citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. *Immunol. Rev.* **233**, 34–54 (2010).
3. Rosenstein, E. D., Greenwald, R. A., Kushner, L. J. & Weissmann, G. Hypothesis: the humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis. *Inflammation* **28**, 311–318 (2004).
4. McGraw, W. T., Potempa, J., Farley, D. & Travis, J. Purification, characterization, and sequence analysis of a potential virulence factor from *Porphyromonas gingivalis*, peptidylarginine deiminase. *Infect. Immun.* **67**, 3248–3256 (1999).
5. Mercado, F. B., Marshall, R. I. & Bartold, P. M. Inter-relationships between rheumatoid arthritis and periodontal disease. A review. *J. Clin. Periodontol.* **30**, 761–772 (2003).
6. de Pablo, P., Chapple, I. L., Buckley, C. D. & Dietrich, T. Periodontitis in systemic rheumatic diseases. *Nat. Rev. Rheumatol.* **5**, 218–224 (2009).
7. Ogrendik, M. Rheumatoid arthritis is linked to oral bacteria: etiological association. *Mod. Rheumatol.* **19**, 453–456 (2009).
8. de Pablo, P., Dietrich, T. & McAlindon, T. E. Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population. *J. Rheumatol.* **35**, 70–76 (2008).
9. Michalowicz, B. S. *et al.* Periodontal findings in adult twins. *J. Periodontol.* **62**, 293–299 (1991).
10. Lundberg, K. *et al.* Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. *Arthritis Rheum.* **58**, 3009–3019 (2008).
11. MacGregor, A. J. *et al.* Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum.* **43**, 30–37 (2000).
12. Mahdi, H. *et al.* Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis. *Nat. Genet.* **41**, 1319–1324 (2009).
13. Kinane, D. F., Shiba, H. & Hart, T. C. The genetic basis of periodontitis. *Periodontol. 2000* **39**, 91–117 (2005).
14. Katz, J., Goultschin, J., Benoliel, R. & Brautbar, C. Human leukocyte antigen (HLA) DR4. Positive association with rapidly progressing periodontitis. *J. Periodontol.* **58**, 607–610 (1987).
15. Marotte, H. *et al.* The association between periodontal disease and joint destruction in rheumatoid arthritis extends the link between the HLA-DR shared epitope and severity of bone destruction. *Ann. Rheum. Dis.* **65**, 905–909 (2006).
16. Bonfil, J. J. *et al.* A "case control" study on the role of HLA DR4 in severe periodontitis and rapidly progressive periodontitis. Identification of types and subtypes using molecular biology (PCR.SSO). *J. Clin. Periodontol.* **26**, 77–84 (1999).
17. Kinane, D. F. & Chestnutt, I. G. Smoking and periodontal disease. *Crit. Rev. Oral Biol. Med.* **11**, 356–365 (2000).
18. Colombo, A. V., da Silva, C. M., Haffajee, A. & Colombo, A. P. Identification of intracellular oral species within human crevicular epithelial cells from subjects with chronic periodontitis by fluorescence *in situ* hybridization. *J. Periodontal Res.* **42**, 236–243 (2007).
19. Potempa, J., Sroka, A., Imamura, T. & Travis, J. Gingipains, the major cysteine proteinases and virulence factors of *Porphyromonas gingivalis*: structure, function and assembly of multidomain protein complexes. *Curr. Protein Pept. Sci.* **4**, 397–407 (2003).
20. Wegner, N. *et al.* Peptidylarginine deiminase from *Porphyromonas gingivalis* citrullinates human fibrinogen and alpha-enolase: implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum.* doi: 10.1002/art.27552.
21. Kinloch, A. *et al.* Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis. *Arthritis Res. Ther.* **7**, R1421–R1429 (2005).
22. Snir, O. *et al.* Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. *Ann. Rheum. Dis.* **68**, 736–743 (2009).
23. Wegner, N., Wait, R. & Venables, P. J. Evolutionarily conserved antigens in autoimmune disease: implications for an infective aetiology. *Int. J. Biochem. Cell Biol.* **41**, 390–397 (2009).
24. Rantapää-Dahlqvist, S. *et al.* Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum.* **48**, 2741–2749 (2003).
25. Lundberg, K. The pathobiology underpinning autoantibodies to citrullinated protein antigens [abstract SP0064]. *Ann. Rheum. Dis.* **68** (Suppl. 3), 21 (2009).
26. Yucel-Lindberg, M. *et al.* Detection of citrullinated proteins and PAD-2 in periodontal tissue [abstract THU0155]. *Ann. Rheum. Dis.* **69** (Suppl. 3), 69 (2010).
27. Mikuls, T. R. *et al.* Antibody responses to *Porphyromonas gingivalis* (*P. gingivalis*) in subjects with rheumatoid arthritis and periodontitis. *Int. Immunopharmacol.* **9**, 38–42 (2009).
28. Ebersole, J. L. *et al.* Human serum antibody responses to oral microorganisms. IV. Correlation with homologous infection. *Oral Microbiol. Immunol.* **2**, 53–59 (1987).
29. Havemose-Poulsen, A. *et al.* Periodontal and hematological characteristics associated with aggressive periodontitis, juvenile idiopathic arthritis, and rheumatoid arthritis. *J. Periodontol.* **77**, 280–288 (2006).
30. Nicholas, A. P. & Whitaker, J. N. Preparation of a monoclonal antibody to citrullinated epitopes: its characterization and some applications to immunohistochemistry in human brain. *Glia* **37**, 328–336 (2002).

Acknowledgments
We thank Professor A. P. Nicholas of the University of Alabama at Birmingham, USA, for the kind donation of the F95 IgM mouse monoclonal anticitrulline antibody.

Author contributions
All authors contributed equally to researching data for the article, providing a substantial contribution to discussions of the content, writing the article, and to review and/or editing of the manuscript before submission.