The role of Human Endogenous Retrovirus K10 (HERV-K10) in the pathogenesis of rheumatoid arthritis via molecular mimicry

Malgorzata Marianna Trela BSc, MSc

A thesis submitted in partial fulfilment of the requirements of the University of Wolverhampton for the degree of Doctor of Philosophy

March 2019

This work or any part thereof has not previously been presented in any form to the University or to any other body whether for the purposes of assessment, publication or for any other purpose (unless otherwise indicated). Save for any express acknowledgments, references and/or bibliographies cited in the work, I confirm that the intellectual content of the work is the result of my own efforts and of no other person.

ABSTRACT

Rheumatoid arthritis (RA) is a chronic joint disease of unknown aetiology. The autoimmune nature of RA is underlined by abundant generation of rheumatoid factor (RF) autoantibodies to IgG1 Fc, and anti-citrullinated protein antibodies (ACPA) to citrullinated autoantigens such as fibrinogen. Although RA pathogenesis has not been elucidated, genetic predisposition, environmental insults and viral pathogens are considered contributory factors. Human endogenous retrovirus K10 (HERV-K10) is one such virus as it retained the capacity to produce viral particles in RA synovium. This study set out to explore how HERV-K10 Gag matrix region could contribute to RA pathogenesis and perpetuation, with particular emphasis on its ability to mimic host autoantigens. We showed that Gag region exhibits high levels of sequence and structural homology to IgG1 Fc and it could provide a key epitope important for auto-reactivity in RA.

Analysis of how HERV-K10 may evoke immune responses in RA was broadened by investigation of serological cross-reactivity of novel anti-K10 polyclonal antibody (PAbMAG) with IgG1 Fc. We showed that PAbMAG cross-reacted with linear and conformational epitopes on IgG1 Fc. In a further development, we showed a significantly elevated mean IgG response to HERV-K10 epitopes in serum samples from RA patients when compared to other arthritides. These data suggest that molecular mimicry between viral and host proteins has the potential to lead to antigen-driven high-affinity RF IgG immunological reactivity in RA.

Finally, we broadened our study of mimicry in RA by the investigation of citrullinated autoantigens. Structural studies demonstrated high levels of homology between citrullinated fibrinogen, IgG1 Fc and HERV. We further explored how protein citrullination affects the cross-reactivity of autoantibody responses in RA. These experiments revealed that generation of neoepitopes through citrullination of HERV-K10 and autoantigens IgG1 Fc and fibrinogen enhanced the reactivity of RA sera to these targets. Moreover, we showed that RF autoantibodies could mediate responses to a classical ACPA target fibrinogen, only when it is citrullinated, in the absence of ACPAs. These data provide a new insight into the initiation and propagation of immunological responses in RA and how viral/host molecular mimics and citrullination could modulate serum cross-reactivity profiles in RA.

CONTENTS

I would like to dedicate this work to my late Father, Dr Ryszard Trela whose kind heart, wisdom, passion for life, science and another human being continue to inspire me throughout my life.

ACKNOWLEDGEMENTS

First and foremost I want to express my profound gratitude to my mentor and the mastermind of this project, Professor Paul Nelson (*in memoriam*), who saw my potential before I could see it in myself and provided me the opportunity to work in his research group. He introduced me not only to the world of autoimmunity research, but also to the fascinating world of philosophical thinking in everyday life. His excellent guidance, knowledge, unwavering integrity and encouragement enabled me to grow both as a person and as a scientist and complete this work successfully. I feel humbled and honoured to have contributed to his professional legacy.

I would like to express my special appreciation and thanks to Dr Kesley Attridge, who did not hesitate to step in and become my supervisor in the 2nd year of my doctorate. You have been a tremendous mentor to me, your advice on both research as well as on my career have been invaluable and the joy and enthusiasm that you have for immunology is contagious and motivational, even during tough times in the PhD pursuit.

I am also very thankful to Dr Shantha Perera and Prof Paul Rylance, my other supervisors, for their undivided support and encouragement as well as for their expert guidance and for the many times they trusted me to represent our group.

I am extremely grateful to our collaborators at The Royal Wolverhampton NHS Trust, consultant rheumatologists, particularly Dr Thomas Sheeran, and research nurses, whose admirable efforts in patient recruitment and sample collection enabled me to complete my work.

I would also like to thank the South Staffordshire Medical Foundation, Rotha Abraham Bequest, the Royal Wolverhampton Hospital Charity, the New Cross Kidney Patients Association, and the James Beattie Charitable Trust for supporting this work by providing research grants.

I gratefully thank Dr Margaret Goodall from the University of Birmingham who has been a true inspiration to me and whose expertise and enthusiastic attitude helped me to arrive at my goal. She has taught me, both consciously and unconsciously, how good experimental immunology is done and I am very thankful for the excellent example she provided as a successful woman immunologist.

I want to express my gratitude to all the technical staff at the University of Wolverhampton and to my colleagues in the lab, in particular Denise Roden for their excellent technical assistance.

Warm thanks are expressed to Dan Blakeway who have made these years enjoyable both in the lab and also at leisure, we have shared the ups and downs during the PhD studies and are almost at the end of this road!

A special thank you to Hoda, you have been one of my greatest supporters since the moment we met. Your words of encouragement, understanding and advice on both professional and personal level have been invaluable and you are truly my soul-sister.

To my dearest friends, Ania, Claire, Kasia, Shawna, Hamida, Zaynab, Nelly, the Marras and Yip families, your are deeply thanked for your wonderful friendship along the years. Long discussions with good food and even better friends have reminded me about the important things in life.

I render my deepest thanks to Dr Ewa, Andrzej and Tomasz Banyś who I consider my family and who I look up to. Your constant support, encouragement and unlimited belief in my abilities from the very beginning have made this thesis possible.

A special thanks to my parents, Drodzy Rodzice, jestem Wam głęboko wdzięczna za Wasze wsparcie, wyrzecznia oraz za to że od początku mojego życia wierzyliście we mnie i byliście moimi najlepszymi nauczycielami. To dzięki Wam ta praca ziściła się.

Most of all, I want to thank my partner Rafael, whose support, understanding and patience has helped me complete this thesis.

LIST OF FIGURES

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 5:

Chapter 6:

Chapter 7:

LIST OF TABLES

Chapter 1:

Chapter 2:

Chapter 3:

LIST OF ABBREVIATIONS

Ab……...antibody

ACR…….……..American College of Rheumatology

Ag……..……..antigen

ACPA……..anti-citrullinated protein antibody

AID……..……activation-induced cytidine deaminase

AIRE……..……...autoimmune regulator

ALS……..……..amyotrophic lateral sclerosis

APC……..……..antigen presenting cells

BAFF……..……..….B cell activating factor

BCR……..……..…B cell receptor

Bregs……..……...regulatory B cells

CCR……..……..CC chemokine receptor

CD……..……cluster of differentiation

CDR……complementarity determining regions

CH……..……..…heavy chain constant domain

CIA……..……collagen-induced arthritis

CAIA…….collagen antibody-induced arthritis

CL……..……...light chain constant domain

CMV……..……..…cytomegalovirus

CRP……..……...C-reactive protein

CSF……..……....cerebrospinal fluid

CSR……..……...class-switch recombination

cTEC……..……...cortical thymic epithelial cells

CXCL……..……..……...CXC chemokine ligand

DAMP……...damage-associated molecular pattern

DC……..……..…..dendritic cell

DMARD……..……..disease-modifying antirheumatic drug

DN……..……..……double negative

DP……...double positive

EBV……..…..Epstein-Barr virus

ECM……..…extracellular matrix

ELF……..…ectopic lymphoid follicle

Env……..envelope

ERV……..……endogenous retrovirus

ESR……..…..…erythrocyte sedimentation rate

EULAR……...European League Against Rheumatism

Fab……..…….fragment antigen binding

Fc……..……fragment crystallisable region

FLS……..…….…fibroblast-like synoviocytes

Foxp3……..…….…..forkhead box p3

Gag……..……..group specific antigen

GCT……..……....germ cell tumours

HCV……..……...hepatitis C virus

HERV……..…....human endogenous retrovirus

HHV……..…….……human herpes virus

HIV……..……...human immunodeficiency virus

HLA……..……human leukocyte antigen

HSC……hematopoietic precursor stem cells

HTLV……..……..human T cell leukaemia virus

IC……..……..…immune complex

IDDM……..……insulin dependent diabetes mellitus

IFNγ……..interferon γ

Ig……..……immunoglobulin

IL……..……...interleukin

iTreg……..……..induced regulatory T cells

JAK…….……....janus kinase

LINE…….long interspersed nuclear element

LTR……..…….long-terminal repeat

MHC……..….major histocompatibility complex

MLS……...macrophage-like synoviocytes

MMP……membrane metalloproteinase

MS……..……..……..multiple sclerosis

MSRV……..multiple sclerosis associated retrovirus

mTEC……..….medullary thymic epithelial cells

NET……..……neutrophil extracellular trap

NSAID……..non-steroidal anti-inflammatory drug

nTreg……..…...…….natural regulatory T cell

OA……..……osteoarthritis

PBMC…….peripheral blood mononuclear cells

Pol……..……...polyprotein

PsA……..……....psoriatic arthritis

PAD……..……peptidylarginine deiminase

PAMP...……..pathogen-associated molecular pattern

PD……..……....periodontitis

PD-1……..……programmed death 1

PRR..……..….pattern recognition receptor

PTM……..……...post-translational modification

RA……..……...rheumatoid arthritis

RANKL…….receptor activator of nuclear factor kappa B ligand

RF……..……....rheumatoid factor

RN……..……...rheumatoid nodules

ROS……..…….reactive oxygen species

RT-PCR……...reverse transcription-polymerase chain reaction

SAg……..……..superantigen

SE……..…….shared epitope

SHM……..……...somatic hypermutation

SLE……..……...systemic lupus erythematosus

SP……..……..single positive

SS……..……..Sjögren syndrome

TCR……..……...T cell receptor

Tfh……..……....T follicular helper

TGFβ……..……transforming growth factor β

Th……..…….T helper

TLR……...toll-like receptor

TNFα……..……...tumour necrosis factor α

TRA……..……tissue-restricted antigens

Treg...……..regulatory T cell

V, D, J……..……variable, diversity, joining

VEGF……vascular endothelial growth factor

VH...……..…..…….heavy chain variable domain

VL……..……....light chain variable domain

1. INTRODUCTION

1.1. Overview of the immune system

The immune system evolved as a rapid, specific, and multilayered defence against harmful pathogens or their biologic products. It functions as an interactive network of lymphoid organs, cells, humoral factors and cytokines across a broad spectrum of challenges preserving the integrity of the host. It can be broadly divided into two subsystems, the innate and adaptive arms, determined by the speed and specificity of the responses. Although the two elements of the immune system exhibit distinct effector mechanisms, they are intricately intertwined and mutually supportive.

1.1.1. Innate immunity

The innate immune system provides the first line of defence against infections or injury via diverse components and mechanisms which are typically present before the onset of a pathogenic attack. Anatomical barriers in the form of the skin and mucosal membranes prevent breach of the host and allow an equilibrium between local harmful environmental insults (toxins, pathogens) and favourable components such as commensal microbiota (Catrina *et al.*, 2016). Innate immunity employs two basic mechanisms of immune recognition: recognition of foreign substances (non-self) and recognition of missing self. Recognition of microbial non-self becomes activated by ‘danger signals’, known as pathogen-associated molecular patterns (PAMPs), which constitute molecular motifs common to various pathogenic microorganisms. These molecules become intercepted by germline-encoded pattern recognition receptors (PRRs) present on neutrophils, macrophages, NK cells and dendritic cells (DCs) which

signal the presence of infection and trigger phagocytosis (Medina, 2016). Toll-like receptors (TLRs), which constitute a major family of PRRs, are responsible for the recognition of motifs broadly shared by pathogenic organisms and act as critical mediators of innate immunity (Skwarczynski and Tóth, 2016). The TLR family comprises two groups based on subcellular localization with cell-surface receptors (TLR1, TLR2, TLR4, TLR5 and TLR6) recognising primarily bacterial and fungal PAMPs and intracellular receptors in endosomal compartments (TLR3, TLR7, TLR8, TLR9 and murine TLR11) sensing viral PAMPs as well as self and foreign nucleic acid structures (Kay, Scotland and Whiteford, 2014). TLRs can also recognise damage-associated molecular patterns (DAMPs) released in the body from damaged cells which are thereby tagged for elimination. Recognition of missing self relies on the detection of functional molecules typically expressed on normal cells unique to the host, which cease to be presented upon infection or cellular transformation (Medzhitov and Janeway, 2002). This strategy is also employed by the complement system. One of the complement components, C-type lectin mannose receptors recognise repeated mannose units on the surface of infectious agents and trigger an enzymatic cascade that ultimately leads to cell lysis. The complement system can also be initiated via an alternative pathway through the binding of C3b to bacterial cells or viral particles. Complement, together with TLRs, act as mediators between the innate and adaptive systems. Indeed, the classical complement pathway is dependent on the formation of soluble antigen-antibody complexes which are characteristic of adaptive immune responses (Hajishengallis and Lambris, 2010).

1.1.2. Adaptive immunity

A clear dialogue exists between innate and adaptive immunity where dendritic cells form the chief interface between these two systems. This interaction is initiated upon endocytosis of a pathogen by DCs. For example, following TLR ligation, bacteria are internalised into the phagosome, wherein bacterial peptide antigens are processed and loaded onto major histocompatibility complex II. MHC-peptide is then trafficked to the cell surface for presentation to cells of the adaptive immune system. DC directly infected by an intracellular pathogen, such as a virus, recognise cytosolic PAMPs with the use of RIG-1 like receptors. Viral products are then processed, coupled with MHC class I and displayed at the plasma membrane where they can trigger adaptive immunity. Once activated, dendritic cells migrate to secondary lymphoid organs, such as lymph nodes, where they present antigens to T lymphocytes to initiate adaptive immune responses to the vast majority of potential pathogens, be they extra- or intracellular (Iwasaki and Medzhitov, 2010).

Two major cell populations, T and B lymphocytes, occupy the centre stage of adaptive immunity by dictating the specificity of immune responses and orchestrating the effector arms of these processes. T lymphocytes recognise antigenic peptides complexed with an MHC molecule. Recognition of a particular antigen is facilitated by highly specific T cell receptors (TCRs) which, unlike germline-encoded PRRs derived from innate responses, are assembled via a series of genomic rearrangements known as somatic V(D)J recombination. TCR specificity is determined by each assembled V-D-J cassette which represents a vast number of possible permutations of recombinations of the gene segments of each type (V, D or J). This repertoire of pathogen detection regions encoded by the TCR greatly exceeds the number of known pathogens by many orders of magnitude (Bonilla and Oettgen, 2010). Activation of native T lymphocytes is initiated upon antigen recognition and is followed by their

proliferation and migration to infection sites. Small populations of these cells will differentiate into long-lived memory T lymphocytes which can mount protective immune responses upon pathogen re-encounter (Farber, Yudanin and Restifo, 2014).

Pre-existing diversity in the antigen receptor repertoire is also used by B lymphocytes which constitute second major cell population of adaptive immunity. These cells identify their cognate antigen in its native form displayed by dendritic cells, or engage with soluble antigen in the extracellular space (Heesters *et al*., 2016). In contrast to TCR, the antigen receptors on B lymphocytes, known as B cell receptors (BCRs) are surface-bound immunoglobulins which can be secreted as soluble receptors. The process of BCRs assembly from V(D)J gene segments is entirely analogous to the mechanism for TCR. Following antigen engagement, B cells receive stimulatory signals from cognate T cells which encourage them to proliferate into short-lived plasma cells (plasmablasts) or form specialised germinal centres where they differentiate into long-lived memory B cells (Harwood and Batista, 2010). Upon antigen re-exposure, these cells will exhibit rapid and robust responses resulting in proliferation and plasma cell formation (Kurosaki, Kometani and Ise, 2015). The ability of T and B lymphocytes to communicate in this highly sophisticated and specific manner greatly contributes to host defence against infection.

1.1.3. B lymphocytes

B lymphocytes develop from hematopoietic precursor stem cells (HSCs) in the bone marrow where they undergo distinct developmental stages to attain their antigen specificity via V(D)J rearrangement of the BCR, which may be further enhanced by random addition of ‘non-templated’ N nucleotides onto DNA strands via terminal dideoxyl transferase (TdT). Before entering the periphery, B cell progenitors are

subject to rigorous checkpoints for highly self-reactive BCRs which undergo deletion. This process constitutes the antigen-independent phase of B cell development which transpires in the absence of any interaction with exogenous antigen (Pieper, Grimbacher and Eibel, 2013, Motea and Berdis, 2010). Subsequently, the antigen-dependent phase of B lymphocyte differentiation occurs. Immature B cells leave the bone marrow and mature in the spleen into B2-type cells termed naïve follicular or marginal zone cells. A third category of mature B lymphocytes is the B1-type cells which unlike the other two types derive from foetal liver and neonatal bone marrow and are subdivided into the B1a ($CD5^+$) and B1b ($CD5^-$). These cells reside primarily in the pleural and peritoneal cavities (Montecino-Rodriguez and Dorshkind, 2012). B1-type together with marginal zone B cells exhibit heightened propensity for differentiation into plasma cells which are the source of natural serum immunoglobulin M (IgM). They have been defined as innate-like cells due to their rapid and shortlived T-cell independent antibody responses of poor specificity. These two cell types express BCRs in conjunction with pathogen-recognising TLRs which are essential for their immune responses and contribute to the interaction between the innate and adaptive systems. Moreover, B1-type and marginal zone B cells have been implicated in perpetuating autoimmune responses as a result of their polyspecific BCR repertoire (Rawlings *et al*., 2012). T-dependent follicular B cell responses on the other hand are mediators of highly specific adaptive immune responses. Antigen recognised by the surface IgM of the B lymphocyte is processed and presented in complex with MHC class II to follicular helper CD4+ T cells at the border of the B cell follicle and T cell zone. The T lymphocytes in turn generate accessory signals which enable B cell activation and further differentiation. Assistance is predominantly provided via CD40 signalling and cytokines IL-4 and IL-21 which induce isotype switching from IgM to IgG, IgA, or IgE (Iwasaki and Medzhitov, 2015). Upon T cell priming, antigen-activated B

cells can either exit the follicle to become short-lived plasmablasts, or migrate to the follicles to establish germinal centres. At this site B lymphocytes diversify their antibody variable regions through somatic hypermutation (SHM) which leads to the production of high-affinity antibodies for the antigen involved. This process constitutes affinity maturation and results in the production of long-lived plasma cells, many of which travel to the bone marrow and continue releasing antibody for several weeks. Memory B cells are also generated and can differentiate into plasma cells upon recall antigen exposure. In addition to the immunoglobulin gene rearrangement and hypermutations in the variable regions which provide high antigen specificity, class-switch recombination (CSR) is catalysed by components of the SHM machinery, including activation-induced cytidine deaminase (AID). During CSR, the constant regions of IgM and IgD (Cμ and Cδ, respectively) become replaced by Cα, Cε or Cγ, giving rise to IgA, IgE or IgG isotypes which determine the type of the immune response. Therefore, CSR provides functionally distinct antibodies while preserving antigen specificity (Bürgler, 2017).

1.1.4. Antigen and antibody interactions

One of the main events in the neutralisation of pathogenic microorganisms or foreign particles by the adaptive immune system is the specific binding between immunoglobulins and their target antigens. The connection between an antibody (Ab) and an antigen (Ag) requires numerous non-covalent interactions between the antigen-binding site and the epitope (antigenic determinant) (Kringelum *et al*., 2013). Since epitopes have the ability to bind both soluble Ab molecules and membrane-bound BCRs, they are often termed B cell epitopes to distinguish them from the proteolytically cleaved T cell epitopes and will be referred to as epitopes in this section. Epitopes of

proteins are generally classified as either continuous or structural (conformational). The latter consist of key amino acid residues which are brought together by protein folding, whereas continuous epitopes are composed of a single, linear fragment of the antigen's amino acid residues which come into direct physical contact with antibody residues (Regenmortel, 2009). Antibody molecules are large (150kDa) plasma glycoproteins composed of an immunoglobulin monomer which constitute their basic functional unit. Abs consist of four polypeptide chains: two identical heavy (H) and two identical light (L) chains linked by disulphide bonds (Cuesta *et al*., 2010). They each create the arms of a Y-shaped structure know as fragment antigen binding (Fab), giving rise to hypervariable Ag binding sites or paratopes (Sela-Culang, Kunik and Ofran, 2013). The latter include six flexible complementarity determining regions (CDRs) which harbour hotspot residues and mediate Ag recognition by adopting a three-dimensional confirmation (Nikoloudis, Pitts and Saldanha, 2014; Wei *et al*., 2015) (Fig. 1.01). Each immunoglobulin monomer is bivalent and can bind concurrently two identical antigenic molecules resulting in an avidity-driven increase in Ab affinity and its long-term retention on the antigen (Deyev and Lebedenko, 2009). The Fab consist of two variable domains (VH and VL) and two constant domains (CH1 and CL). The base of the Y, known as the fragment crystallisable (Fc) region is composed of two heavy chain domains, CH2 and CH3 and connected with two Fabs via a hinge region. Fc fragment confers the isotype of the immunoglobulin and mediates the effector functions of the Ab such as complement binding, half-life, avidity and Fc receptor binding (Janda *et al*., 2016). Immunoglobulin G constitutes the major serum isotype and comprises four subtypes, IgG1, IgG2, IgG3 and IgG4 (Panda and Ding, 2015). Antibody molecules can be produced by different clones of plasma B cell and therefore exhibit polyclonal responses to different epitopes in the same antigen whereas

monoclonal antibodies originate from the same clone and bind to a unique epitope (Saper, 2009).

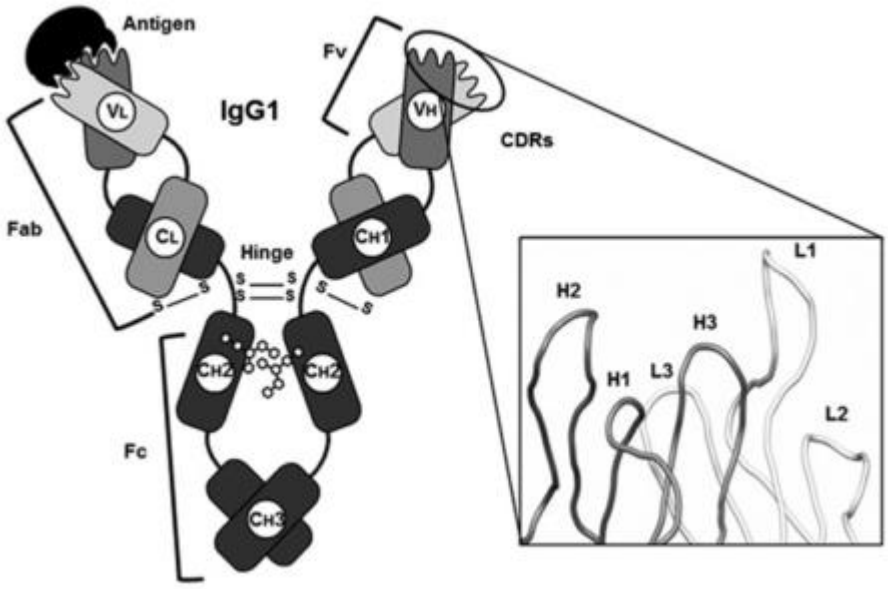


Figure 1.01. Structure of immunoglobulin G (IgG). Depicted are the heavy (dark gray) and light (lighter gray) chains with fragment antigen binding (Fab) and constant (Fc) domains. The insert represents the six complementarity determining regions (CDRs) which form the antigen-binding site. Reproduced from: (Haji-Ghassemi *et al*., 2015).

1.1.5. T lymphocytes

As with B cells, T lymphocytes originate from bone marrow resident multipotential HSCs which upon differentiation become either common myeloid progenitors or common lymphoid progenitors with the latter being B and T cell precursors. In order to acquire their identity, T lymphocytes migrate to the thymus via the blood (Krueger, Ziętara and Łyszkiewicz, 2017). Upon interaction with thymic cells and chemokines, T cell progenitors migrate to the outer cortex where they differentiate based on the cell-surface expression of CD4 and CD8, progressing from the double-negative (DN) into double-positive (DP) cells. These cells subsequently mature into single-positive (SP) cells which express either CD4 or CD8 exclusively. At this stage, developing T cells express low levels of the membrane TCR and are attracted to the thymic medulla where they mature and subsequently exit into the circulation to patrol peripheral tissues (Hosokawa and Rothenberg, 2017). Mature T lymphocyte populations, CD4+ and CD8+ T cells have very distinct effector functions. CD8+ cells act to remove cells infected with intracellular pathogens, including viruses and dysfunctional cells. They induce cytolytic molecules such as granzymes/perforin and Fas/FasL pathways to kill infected cells (Müllbacher *et al*., 2002). In contrast, CD4+ T cells serve a helper function by producing a range of cytokines to augment the phagocytic activity of macrophages and assist B cells in generation of high affinity antibodies. Mature T lymphocytes are activated upon engagement of their TCRs with antigenic peptide displayed on MHC molecules. CD8+ T cells recognise antigen (typically 9 amino acids in length) presented by MHC class I, whilst CD4+ T cells require antigen presentation (generally 15 amino acid peptide) by MHC class II (Kiecker *et al*., 2004; Luckheeram *et al*., 2012). The type of effector response generated depends upon the nature of the pathogen. All nucleated cells display MHC class I molecules complexed with endogenous antigens, therefore cells infected with intracellular pathogens will be

cleared by CD8+ T cells. In contrast, MHC class II molecules are present on professional antigen presenting cells (APCs), including macrophages, B cells and dendritic cells (Parkin and Cohen, 2001), and initiate CD4+ T cell responses to extracellular pathogens.

1.2. Immunological tolerance and autoimmunity

The ability to mediate self-defence while avoiding excessive immune reactions that would be deleterious to the host is a fundamental attribute of the immune system. Continuous generation of an enormous diversity of antigen recognition receptors inherent to random V(D)J gene rearrangement inevitably leads to the development of self-reactive B and T cells. The immune system has therefore evolved robust mechanisms ensuring the generation and preservation of lymphocyte self-tolerance. Classically, these processes are mediated within the bone marrow and the thymus where they constitute central tolerance, or in the periphery forming peripheral tolerance.

1.2.1. Central tolerance

Central B cell tolerance comprises the regulatory mechanisms that negatively select immature B lymphocytes which react with self-antigens in the bone marrow. Ligation of the BCR by self-peptides results from the random nature of V(D)J rearrangement. It has been reported that up to 75% of BCRs are self-reactive at this stage, therefore stringent checkpoints exist to prevent B cells from migrating into peripheral compartments (Nemazee, 2017). Autoreactive immature clones become eliminated through the mechanisms of clonal deletion and anergy (partial or total unresponsiveness). They can also undergo receptor editing which enables reprogramming of the Ig V gene fragment to form a non-autoreactive BCR. If this process proves unsuccessful in removing self-reactive receptors, deletion occurs. High avidity interactions of immature B cells with self-antigen leads to cell apoptosis whilst low avidity interactions are thought to permit receptor editing (Pelanda and Torres, 2012). Some weakly self-reactive clones emerge into the periphery without being

eliminated or inactivated but do not differentiate and are effectively ignored (clonal ignorance) (Tsubata, 2017).

In order to maintain immune homeostasis, developing T lymphocytes undergo verification at several immune checkpoints prior to their maturation and exit from the thymus. The first stage of the selection process involves DN progenitors (CD4-CD8-) which will not be allowed to progress to DP thymocytes (CD4+CD8+) without rearranging their TCR β chain locus. The TCR affinity for MHC-self-peptide presented by cortical thymic epithelial cells (cTEC) determines the fate of the remaining T lymphocytes. DP cells which fail to bind (Gururajan, Sindhava and Bondada, 2014) with a minimum level of affinity and avidity die by neglect, whereas T cells that interact with self-MHC are positively selected and migrate to the medullary area to differentiate into CD4 or CD8 SP thymocytes (Xing and Hogquist, 2012). Here, medullary thymic epithelial cells (mTEC) present a broad array of tissue-specific self-antigens. The autoimmune regulator (AIRE), together with other costimulatory molecules, plays an important role in the expression of a broad array of these antigens (Metzger and Anderson, 2011). Strongly self-reactive TCRs represent a potential danger of autoimmunity and they can elicit a number of outcomes. Thymocytes exhibiting autoreactivity are either eliminated through negative selection (clonal deletion), anergised or are skewed to the regulatory lineage. Central tolerance mechanisms are so efficient that out of approximately fifty million T lymphocytes produced daily in the thymus, under 10% of these precursors will be released into the circulation (Klein *et al*., 2014).

1.2.2. Peripheral tolerance

Although 90% of autoreactive B cell clones become eliminated through deletion, receptor editing or anergy, a small number of self-reactive cells will survive and migrate to the periphery as transitional B cells. Here peripheral tolerance mechanisms, including the interaction between tonic BCR-mediated signals and B lymphocyte stimulator (BAFF/BLyS) signalling, determine whether transitional B lymphocytes undergo positive or negative selection (Gururajan, Sindhava and Bondada, 2014). Moreover, if a naïve mature B cell recognises its antigen in peripheral tissues in the absence of cognate T cell help (CD40 engagement or cytokine stimulation) it will be functionally inactivated or undergo apoptosis (Tobón, Izquierdo and Cañas, 2013). Alternatively, mature B lymphocytes become partially activated which results in their exclusion from lymphoid follicles and premature death (Romagnani, 2006).

Should self-reactive T lymphocytes avoid deletion through the process of thymic negative selection, several mechanisms operate on autoreactive mature clones in the periphery. One level of peripheral control comprises physical separation of potentially self-reactive T lymphocytes from tissue-restricted antigens (TRA), such as that facilitated by the blood-brain barrier. Naïve T cells are directed towards the lymph node by the expression of CCR7 and CD62L which limit their migration to the nonlymphoid peripheral tissues with a high density of TRAs (Mueller, 2010). Immunologically privileged sites are protected from the entry of naïve clones through low expression of selectin and integrin receptors, which are necessary for transendothelial migration. Occasional T lymphocyte infiltration of immune-privileged organs such as the ocular system is associated with high expression of immunomodulatory mediators such as IL-10/TGFβ which suppress autoreactive clones (Shechter, London and Schwartz, 2013). However, due to the migration of antigen-presenting mature dendritic cells into the lymph nodes, the majority of naïve T cells do not need to enter the non-lymphoid

tissues to encounter their cognate antigen. Tolerogenic dendritic cells are a specialised subset which present TRA under non-inflammatory conditions in the absence of adequate costimulatory signals. This results in Fas- and Bim-mediated apoptotic cell death of autoreactive T lymphocytes or anergy induction (Luckashenak *et al*., 2008; Raker, Domogalla and Steinbrink, 2015). Differentiation of highly self-reactive T lymphocytes into the FOXP3+ natural regulatory (nTreg) cell lineage constitutes a vital process for peripheral tolerance. These cells enter the periphery and have the ability to suppress autoreactive T cell clones. Here, nTregs convert to FOXP3+ induced T cells (iTreg) upon TGFβ or retinoic acid stimulation and suppress the activation and expansion of autoreactive cells (Bilate and Lafaille, 2012; Lourenço and La Cava, 2011). Therefore, deletion, anergy, clonal ignorance and active regulation by nTreg and iTreg result in the maintenance of peripheral tolerance.

1.2.3. Autoimmunity

Despite the broad array of mechanisms utilised by the immune system to maintain self-tolerance, a dysregulation between effector and regulatory responses can lead to autoimmune reactions. Whilst a low degree of autoimmunity may be beneficial to the host, such as in anti-tumour responses, persistent high levels of self-reactivity result in tissue damage and disease (Cristaldi *et al*., 2011). This may arise through a breakdown in tolerance, together with a genetic predisposition, gender, and environmental agents. In general, both innate and acquired immune systems may play a role in the spectrum of organ specific and non-organ specific (systemic) autoimmune diseases. Despite numerous differences in their clinical manifestations, autoimmune disorders appear to follow sequential stages of initiation, propagation and resolution (Rosenblum, Remedios and Abbas, 2015).

Evidence for the role of predisposing genetic polymorphisms in autoimmunity derives from the identification of disease clustering in families and higher disease concordance rates in monozygotic (25-57%) and dizygotic twins (2-9%) (Hewagama and Richardson, 2009). The MHC locus, also known as the human leukocyte antigen (HLA), has been strongly associated with susceptibility to autoimmune disease but the mechanisms involved are not fully elucidated (Matzaraki *et al*., 2017). Moreover, since approximately 80% of individuals suffering from autoimmunity are women, sex hormones and epigenetic silencing of one X chromosome in females is believed to contribute to their predisposition to autoimmune diseases (Amur, Parekh and Mummaneni, 2012). However, the apparent gender bias appears to be disease-specific since SLE, MS, Hashimoto's thyroiditis and RA have been shown to be more prevalent in women (Voskuhl, 2011) whilst autoimmune diabetes prevalence is higher in men (Brahmkshatriya *et al*., 2012). The nature of the environmental factor in autoimmunity is also unclear, however recent evidence suggests that environmentally-induced epigenetic alterations (without genome modification) could mediate certain forms of autoreactivity (Brooks and Renaudineau, 2015). Other potential environmental triggers of autoimmunity include exposure to toxins, microbial infections, nutrition, dysbiosis and UV irradiation. Infectious agents have long been suspected to induce or exacerbate autoimmune disorders via different possible mechanisms such as epitope spreading, molecular mimicry, bystander activation and excessive pattern recognition receptor activation (Trela, Nelson and Rylance, 2016; Vojdani, 2014). These mechanisms together with an inflammatory environment encourage the self-perpetuating nature of autoimmune conditions.

1.3. Rheumatoid Arthritis

1.3.1. General Introduction

The term ‘arthritis’ derives from the Greek arthron meaning ‘joint’ and is used for over 200 forms of musculoskeletal, connective tissue and non-articular disorders which affect both young and old alike. These arthritides can have different clinical manifestations and be of either an inflammatory or non-inflammatory origin. The most common form of chronic inflammatory arthritis is Rheumatoid Arthritis (RA) with a prevalence of approximately 1% of the population worldwide (Gerlag, Norris and Tak, 2016). Other inflammatory arthropathies include ankylosing spondylitis, gout, psoriatic arthritis (PsA), spondyloarthritis, and systemic lupus erythematosus (SLE) (Chang *et al*., 2016; Valesini *et al*., 2015) (Fig. 1.02). Although these rheumatic disorders are typically associated with certain elements of joint inflammation, their presentation can be systemic or organ-specific.

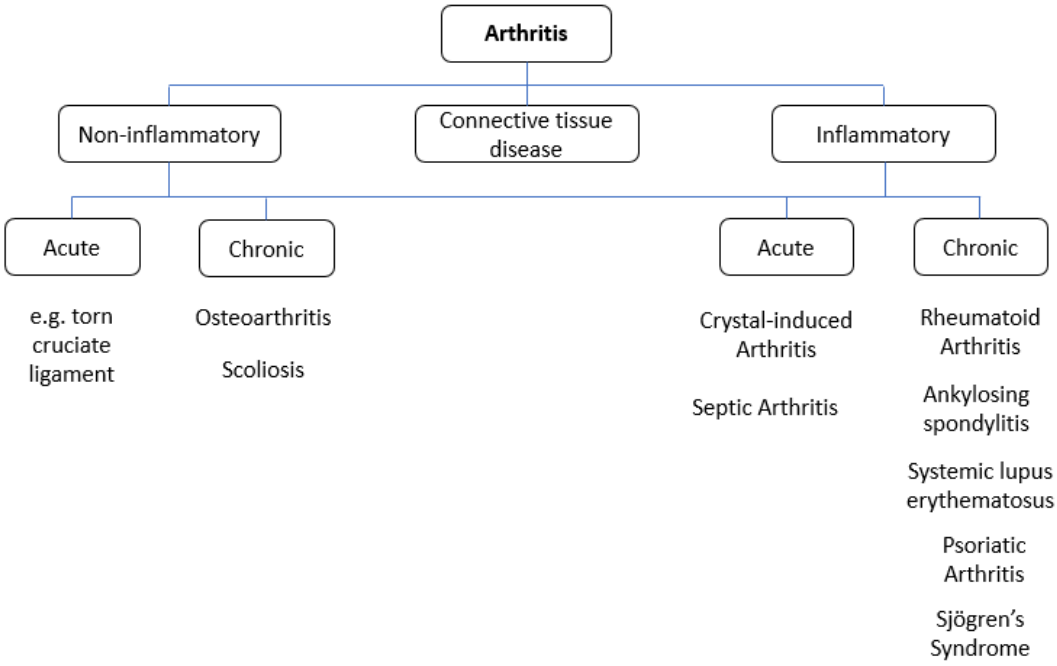


Figure 1.02. Classification of the arthritides. Adapted from: (Freimanis, 2008)

1.3.2. Inflammation

The mechanism of inflammation constitutes a crucial protective strategy in response to pathogenic intruders and tissue injury. It enables restoration of cellular homeostasis by the removal of harmful pathogens as well as the healing of damaged cells. Therefore, inflammatory responses are considered as an element of innate immunity and first line of host defence (Foley, 2013). They can be classified as either systemic or localised with an acute or chronic response.

The entire course of inflammation, its duration and intensity need to be closely controlled to prevent collateral damage and pathology. In several autoimmune conditions, this process extends toward systemic acute phase mechanisms via the increased generation of inflammatory cytokines and acute phase plasma proteins such as C-reactive protein (Ahmed, 2011). The inflammatory signals which trigger a chain of non-specific protective events, limit inadvertent targeting of host tissues. Local inflammatory response is initiated and symptoms such as redness, pain, swelling, heat and functional impairment ensue (Ashley, Weil and Nelson, 2012). Within minutes after the injury, soluble factors such as bradykinin and prostaglandins facilitate vasodilation and increased vascular permeability resulting in oedema at the site of injury (Serra *et al*., 2017). Within hours of onset, neutrophils mediated by pro-inflammatory cytokines (i.e. interleukin-1, interleukin-6, tumour necrosis factor-alpha) and adhesion molecules, such as selectins or integrins, interact with vascular endothelium near the perturbed tissue.

In a sequence of adhesive steps, circulating neutrophils are tethered to the endothelial surface and transmigrate into the damaged tissue, following chemotactic gradients in the area (Muller, 2013). Upon entering the area of disturbance, neutrophils promote the chemotaxis of other immune cells via upregulation of receptors for chemokines and

expression of immunoregulatory cytokines such as interferon-γ. Activated neutrophils express high-affinity Fc receptors which contribute to the recognition of Ig-opsonised microorganisms (Futosi, Fodor and Mócsai, 2013). Moreover, neutrophils release reactive oxygen species (ROS) and lytic proteases such as defensins or cathepsins to destroy pathogens. Phagocytosis is facilitated by neutrophil extracellular traps (NETs) which immobilise harmful microorganisms. Nevertheless, neutrophils may also mediate the tissue damage through generation of membrane metalloproteinases (MMPs) which cleave different forms of collagen and degrade extracellular matrix (ECM) (Kolaczkowska and Kubes, 2013).

Concomitantly with the arrival of neutrophils, free-flowing monocytes are recruited to the site and differentiate into mature macrophages. These cells promote further phagocytosis and inflammatory responses through the production of cytokines including IL-1, IL-12 and TNF-α as well as inducible nitric oxide synthase (Koh and DiPietro, 2011). The release of these cytokines causes further recruitment and infiltration of monocytes and neutrophils at the site of injury. Consequently, an amplification of acute inflammation, in the presence of persistent antigen, may result in a sustained (chronic) inflammatory response leading to overt tissue destruction and/or autoimmunity (Selders *et al*., 2017).

1.3.3. Clinical manifestations of RA

RA, like the majority of rheumatologic disorders, is a multifactorial disease characterised by a wide variation of clinical and pathologic features with unpredictable prognosis. This heterogeneity, which leads to considerable challenges in the development of diagnostics and therapeutic strategies, results in periods of remission and relapse (Chang *et al*., 2016). The disease typically presents as symmetrical

polyarticular arthritis with a gradual onset of additive joint tenderness, stiffness and swelling. RA primarily affects small diarthrodial joints such as metacarpophalangeal or proximal interphalangeal joints, followed by larger joints including ankles, shoulders, and knees.

In approximately 60% of RA patients the onset is acute with palindromic (episodic) joint involvement or with monoarticular presentation which lead to the joint destruction. Up to one-third of patients experience systemic symptoms with prominent myalgia, fever, decreased appetite, weight loss and fatigue (Jeffery, 2010). In order to make an early diagnosis of RA, an individual is assessed according to the most recent classification criteria for RA developed in 2010 by the American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) (Table 1.01). Such criteria enable stratification of patient subgroups, aid diagnosis and prevent individuals from developing the chronic destructive RA (Aletaha *et al*., 2010).

In addition to joint inflammation patients with RA, in particular with early polyarticular onset and aggressive course of the disease, may also develop extra-articular complications. Fibrinoid subcutaneous or internal rheumatoid nodules (RN) occur in around 30% of seropositive patients. They may be co-presented with vasculitis which affects small and mid-sized vessels resulting in cardiac abnormalities, peripheral gangrene and inflammatory eye disease. Pulmonary involvement is also not uncommon and includes pleurisy, pulmonary RN and interstitial lung disease (Romanowska-Prochnicka *et al*., 2013). Overall, RA typically progresses from the periphery to more proximal joints and leads to substantial locomotor impairment within 10 to 20 years without effective treatment strategies (Venables and Maini, 2017). Therefore, clinical, laboratory and radiologic findings play a crucial role in early diagnosis and effective therapy to improve clinical outcomes in RA patients (Sudoł-Szopińska, Jans and Teh, 2017).

2010 ACR/EULAR CLASSIFICATION CRITERIA FOR RA	
Requirements for using the criteria: Patients should have at least 1 swollen joint (synovitis) which is not better explained by another condition. To classify for RA: Add all the applicable scores from categories A-D. A score of ≥ 6 is required for classification as definite RA.	
	SCORE
A. Joint involvement	
1 large joint	0
2-10 large joints	1
1-3 small joints	2
4-10 small joints	3
>10 joints (at least 1 small joint)	5
B. Serology	
RF- and ACPA-	0
Low RF+ and low ACPA+	2
High RF+ and high ACPA+	3
C. Acute-phase reactants	
Normal CRP and normal ESR	0
Abnormal CRP or abnormal ESR	1
D. Duration of symptoms	
< 6 weeks	0
≥ 6 weeks	1

Table 1.01. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis. Adapted from (Aletaha *et al*., 2010). ACPA, anti-citrullinated protein antibody; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate.

1.3.4. Rheumatoid Arthritis (RA)

RA is a chronic autoimmune inflammatory disease associated with progressive degradation of articular cartilage and underlying bone. It primarily affects the joints but extra-articular features such as rheumatoid nodules, pulmonary manifestations or vasculitis, and systemic effects are common. In industrialised countries, RA has an incidence rate of 0.5% to 1% (Choy, 2012), although some populations such as North American Natives may have rates as high as 5% (McDougall, Hurd and Barnabe, 2017). Being the most common inflammatory systemic autoimmune condition, rheumatoid arthritis incurs considerable socioeconomic burden with an estimated annual cost in the UK of £4 billion (NICE, 2009). RA is characterised by the production of rheumatoid factor (RF) autoantibodies reactive with the constant region (Fc) of IgG and anti-citrullinated peptide antibodies (ACPAs) (Brink *et al*., 2016). There is a strong association with the HLA system, responsible for displaying processed peptide to T lymphocytes, which accounts for 30-50% of the genetic susceptibility to RA (Boissier *et al*., 2012). Although RA is present in all ethnic groups and at all ages, its prevalence increases together with age (Innala *et al*., 2014). Moreover, hormonal factors appear to contribute to RA development with women being at three times greater risk than men (Quintero *et al*., 2012).

1.3.5. Pathogenesis of RA

The normal joint lining (synovium) is composed of synoviocytes (FLS and MLS), which provides structural support, lubrication and nutrients to the cartilage. The sublining layer consists of blood and lymphatic vessels, lipocytes, fibroblasts, sparse lymphocytes and macrophages in a bed of collagenous extracellular matrix (Smith, 2011).

Rheumatoid arthritis typically affects synovium-lined joints and is usually found in the small joints of the hands and feet, although larger joint involvement is also frequent. The arthritis is characteristically symmetrical and often leads, if untreated, to cartilage and bone erosion with other systemic complications. Cell-mediated immune responses contribute to initiation and propagation of tissue-destructive processes (Fig. 1.03). Although the exact cause of RA has not been fully elucidated, a large body of evidence supports a role of *HLA-DR1* and *HLA-DR4* alleles (*DRB1*0101, *0401, *0404, *0405, *0901*) which reside on chromosome 6 and encode MHC II on APCs. Common amino acid sequences (QRRAA, KRRAA and RRRAA) at positions 70-74 in the MHC II peptide-binding groove comprise the shared epitope (SE) which confers susceptibility to RA and is associated with the presence of RFs or ACPAs (Boissier *et al*., 2012). Moreover, *HLA-DRB1*1001* SE allele can accommodate arginine-replacing citrulline residues in its antigen binding pockets, which elicits citrullinated protein-specific CD4+ T cell and ACPA responses. On conversely, *HLA-DR3* alleles appear to be associated with ACPA negative RA and a less severe disease course. Non-HLA loci such as *PTPN22*, *PADI4*, *STAT4*, *TRAF1* and *CD40* account for only around 5% of the genetic burden of RA (Kurkó *et al*., 2013).

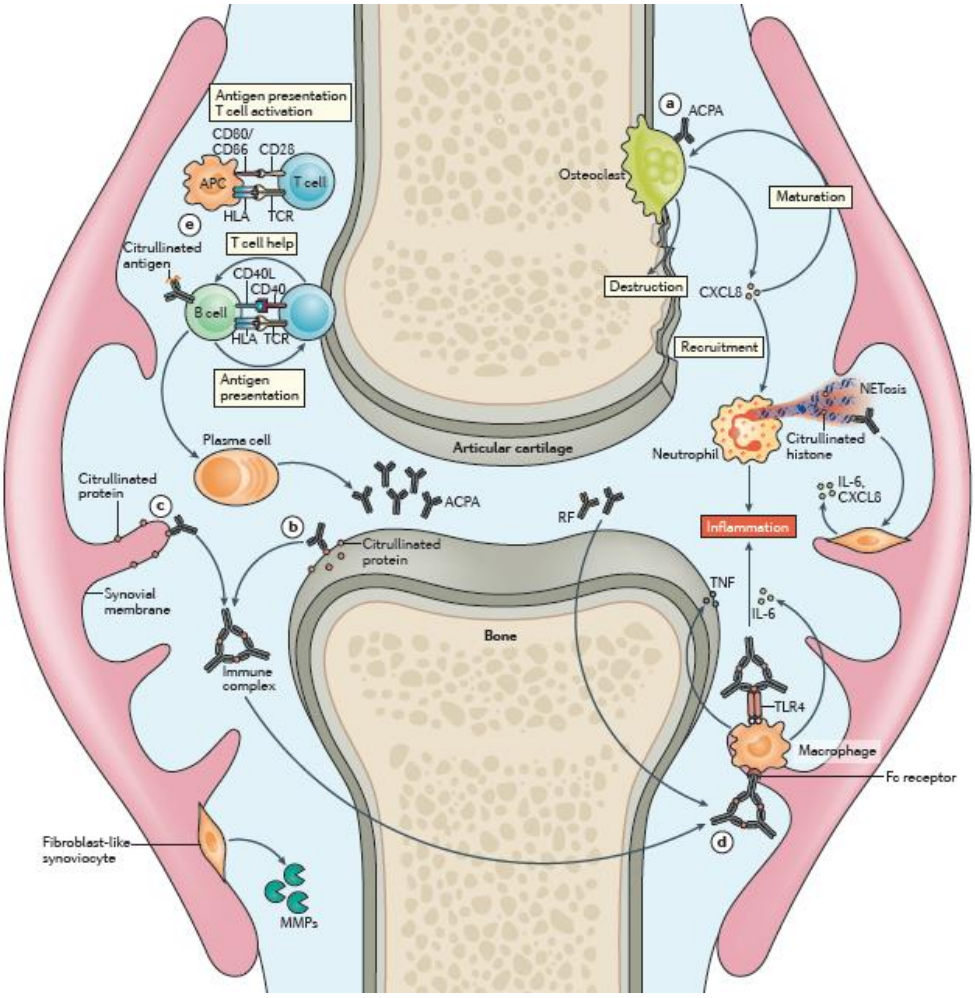


Figure. 1.03. The inflammatory processes in active and established rheumatoid arthritis (RA). RA pathology involves the infiltration of numerous inflammation-inducing cells into synovial fluid and synovium. Articular manifestations result from the presentation of MHC II-bound antigen to T cells which release chemokines and pro-inflammatory cytokines that further perpetuate the inflammatory reactions. TNF-α is the predominant cytokine in RA joint which indirectly leads to cartilage and bone injury via immune cell recruitment and induction of osteoclast differentiation. Reproduced from: (Malmström, Catrina and Klareskog, 2017).

The earliest process in the disease pathogenesis is activation of dendritic cells by autologous antigen which together with other APCs, including macrophages, and activated B lymphocytes, display RA-related antigens to T lymphocytes (Choy, 2012). Autoantigenic peptide can also bind to antibody resulting in the formation of IgG-containing immune complexes (IC), such as those containing RF and ACPA autoantibodies found abundantly in serum and synovial fluid from RA patients. These ICs together with CD4+ T cell-derived IFN-γ activate the monocyte/macrophage cell lineage which promotes inflammatory responses (Cooper *et al*., 2012).

Macrophages constitute central effectors of synovitis and are present in large quantities in the rheumatoid joint. Through the release of pro-inflammatory cytokines such as TNF-α, IL-1, IL-6 and matrix metalloproteinases (MMPs) they activate endothelial cells leading to acute phase reactions and cartilage degradation. Macrophages encourage the recruitment of additional leukocytes to the inflamed joints via expression of chemokines and adhesion molecules (Roberts, Dickinson and Taams, 2015). Recruited neutrophils, which constitute the most abundant cells in synovial fluid, release degradative enzymes and reactive oxygen species (ROS) upon immune complex encounter. These oxygen radicals, inappropriately generated in RA pathology, lead to genetic damage which can result in immunoglobulin mutations and thus RF production. Moreover, the enhanced release of neutrophil extracellular traps during cellular death (NETosis) exposes a range of citrullinated antigens which are targets of ACPAs (Rosas, Correa and Henriques, 2017).

Conditions within the rheumatoid joint such as hypoxia and production of antiapoptotic cytokines promote neutrophil survival from 4 hours (physiological half-life) to several days contributing to the chronic nature of synovitis. Further generation of cytokines such as interferons leads to harmful positive-feedback loops for additional T cell, neutrophil, macrophage and B cell interactions. Moreover, within the synovial

membrane there is a dramatic increase in activated fibroblast-like (FLS) and macrophage-like synoviocytes (MLS) which contribute to overproduction of pro-inflammatory cytokines. FLS, which express various ECM proteins including vimentin and collagens, display an aggressive phenotype within the synovium. Therefore, the rheumatoid joint transforms from a quiescent relatively acellular milieu to an invasive, hyperplastic membrane, known as rheumatoid pannus, which increases in FLS and MLS density from 2-3 layers to up to 15 layers. This structure becomes densely infiltrated with innate and adaptive immune cells such as macrophages, plasma cells and neutrophils and erodes the adjacent articular cartilage and subchondral bone (Rosas, Correa and Henriques, 2017).

Moreover, synovial neo-angiogenesis mediated by vascular endothelial growth factor (VEGF) is a key process for the development of proliferative synovitis (Bartok and Firestein, 2010; Araki and Mimura, 2016). Pannus invasion, displaying a 'tumour-like' phenotype, causes radiologically detected joint space narrowing which can also be induced by immune complexes deposited on the articular cartilage surface (Sudoł-Szopińska, Jans and Teh, 2017). Structural deformities are initially triggered by the activation of bone-resorbing osteoclasts by immune complexes and subsequently aggravated by pro-inflammatory cytokines which lead to osteoclast differentiation via RANKL (receptor activator of nuclear factor kappa B ligand) (McInnes and Schett, 2017). Eventually, the joint space is destroyed leading to the fibrous or bony fusion of the bone ends termed ankylosis (Sophia and Ramesha, 2017).

Strong environmental factors operate on a background of genetic pre-disposition to trigger the development of RA. Cigarette smoking is the dominant environmental pro-disease factor which affects the prevalence and severity of RA in a dose-dependent manner. Smoking doubles the risk of developing the disease. It also augments the risk of ACPA production in individuals with *HLA-DRB1* SE alleles (Chang *et al*., 2014).

Other bronchial irritants such as silica dust have also been implicated in RA as a risk factor (Burska *et al*., 2014).

Moreover, infectious agents have been proposed to perpetuate the disease. Among them, *Porphyromonas gingivalis,* recognised as causative pathogenic organism linked with smoking-related periodontitis (PD), which promotes aberrant tissue citrullination. PD is associated with *HLA-DR4* expression and is at least two-fold more prevalent in RA patients (Koziel, Mydel and Potempa, 2014; Lee *et al*., 2015). In addition, other infectious pathogens such as Epstein-Barr virus (EBV), *Escherichia coli* or *Proteus mirabilis* have been proposed to trigger RA through the mechanisms of molecular mimicry (sequence or structural homology) (Smolen, Aletaha and McInnes, 2016).

Epigenetic regulation of gene expression through DNA methylation status in chromatin appear to play an important role in RA development. This process integrates environmental and genetic risk factors. Altered methylation signatures (global hypomethylation), which were found in T cells and FLS derived from RA patients, can affect gene expression levels and properties of these cell types (Viatte, Plant and Raychaudhuri, 2013). Epigenetic modifications of FLS also contribute to their intrinsic activation and increased expression of demethylated repetitive DNA elements such as the retrotransposon LINE-1 (Klein and Gay, 2015; Hewagama and Richardson, 2009). Moreover, environmental triggers such as smoking could lead to substantial genome-wide alterations in DNA methylation (Zeilinger *et al*., 2013).

1.3.6. Post-Translational Modifications (PTMs) of proteins in RA

Human cells are capable of sustaining their high efficiency as a result of several finely coordinated mechanisms such as DNA transcription or protein synthesis which are directly dependent upon chemical reactions known as posttranslational modifications. PTMs encompass any alteration that affects a protein's chemical structure or property. They take place at the time of or after protein translation and are generally catalysed by enzymes. Currently more than 300 PTMs are recognised with over 200 of them being enzyme-mediated. It is widely accepted that PTMs are physiological processes essential for the development and evolution of all living organisms (Mastrangelo *et al*., 2015). PTMs and their dysregulation have attracted increasing attention over the last few decades as potential contributing factors in numerous pathological conditions such as rheumatoid arthritis. There are several PTMs including citrullination, carbamylation and glycosylation that seem to play an important role in RA pathogenesis and the impairment of tolerance.

1.3.6.1. Citrullination

Citrullination (deimination) entails the switch of the iminic nitrogen of protein-bound arginine to oxygen, in the presence of Ca^{2+}-dependent peptidylarginine deiminases (PAD) enzymes, resulting in the formation of a non-standard amino acid citrulline. Each residue conversion leads to 0.984 Da mass increase and the loss of one positive charge, with the latter affecting the acidity and iso-electric point of the protein's side chain. In humans, PADs occur in five isoforms i.e. PAD1-4 and PAD6 which are tissue restricted and closely controlled (Witalison, Thompson and Hofseth, 2015). Proteins undergo citrullination in the physiological development of tissues and organs as well as in inflammation (Pratesi *et al*., 2014).

In RA synovium, intrinsic PADs are activated in neutrophils, which due to their excessive mobilisation to the articular space constitute 90% of all RA joint cells. Neutrophils arrive in synovium already primed for the production of reactive oxygen species which together with calcium-induced PADs lead to hypercitrullination of proteins (Rosas, Correa and Henriques, 2017; Damgaard *et al*., 2017). The conformational alterations resulting from citrullination can lead to neoepitope production and thus generation of new autoantigens, different to those apprehended by self-tolerance mechanisms (Valesini *et al*., 2015). In this case, elevated number of citrullinated autoantigens stimulate the generation of ACPA antibodies which contribute to IC formation and complement fixation (Tan and Smolen, 2016). Citrullination is also associated with NET production and has several different implications for protein conformation such as change in protein folding and tertiary peptide structure. Moreover, new citrullinated epitopes can be generated by PAD derived from *P.gingivalis* bacterium or via nanoparticles inhaled with tobacco smoke (Mastrangelo *et al*., 2015). Therefore, it has been implied that RA initiation could take place at mucosal compartments such as the lungs or gums (Catrina *et al*., 2017). A total of 53 citrullinated proteins have been detected in RA sera and synovium to date, which have been collectively termed the ‘citrullinome’ (van Beers *et al*., 2013).

1.3.6.2. Carbamylation

Carbamylation is a physiological, non-enzymatic PTM which involves binding of a cyanate group on self-proteins and leads to the formation of a carbamyl group upon urea degradation. Excessive carbamylation can lead to the loss of native protein conformation and a breakdown of tolerance (Shi *et al*., 2014). Lysine carbamylation leads to the formation of homocitrulline in RA patients (Pruijn, 2015) and has been

suggested as a link between smoking, uremia and inflammation which may enhance cyanate levels resulting in the formation of neoepitopes (Turunen *et al*., 2015; Firestein and Mcinnes, 2017). Moreover, anti-carbamylated protein antibodies, with low cross-reactivity with ACPAs, can be detected in some RA patients however their pathogenic role is unknown (Spinelli *et al*., 2016).

1.3.6.3. Glycosylation

Glycosylation involves the attachment of sugar moieties on oxygen or nitrogen atoms of the protein's side chain and is essential for protein folding or cellular interactions. Alterations in the Fc-connected carbohydrate backbone through the addition of sialic acid, galactose, fucose and GlcNAc residues affects humoral immune responses (Scherer *et al*., 2010). In RA, it can lead to pathological antigen presentation to the immune system as a result of aberrant glycosylation of synovial and plasma proteins (Mastrangelo *et al*., 2015). Abnormalities in *N*-glycosylation of the Fc region of IgG have also been reported in RA with a decrease in galactosylated and sialylated glycans in proinflammatory phenotypes (Huang *et al*., 2017). Moreover, different glycosylation patterns may affect immunological properties of RF and ACPA autoantibodies in RA (Bondt *et al*., 2017).

1.3.7. The immunology of RA

1.3.7.1. The role of B Lymphocytes in RA

B lymphocytes are likely to be responsible for RA pathogenesis via several mechanisms including autoantibody production, antigen presentation, T lymphocyte activation, cytokine secretion, bone homeostasis and ectopic lymphoneogenesis (Marston, Palanichamy and Anolik, 2010). A high number of mature naïve autoreactive B cells accumulate in the B cell compartment of RA patients due to defective central and peripheral tolerance checkpoints (Menard *et al*., 2011). The relative contribution of B cells to the disease pathogenesis of RA seems variable, as inferred by the presence of seropositive and seronegative RA (Ajeganova and Huizinga, 2014). However, the importance of B cells in seropositive RA progression has been demonstrated by the success of neutralising anti-CD20 (rituximab) treatment (Buch *et al*., 2011).

Initially suggested by the presence of rheumatoid factor and more recently by the identification of anti-citrullinated peptide autoantibodies, Ab-dependent and Ab-independent pathogenic roles of B cells are corroborated by a growing body of human studies and experimental work. Autoantibodies lead to the formation of immune complexes which are thought to play a major pathogenic role through costimulation of TLR, B cell receptor engagement and activation of the complement cascade (Elshahaly *et al*., 2012). Arthritogenic autoantibodies may be generated from joint-derived synovial plasma cells which originate from locally activated B lymphocytes. Molecular evidence indicates that lymphoid infiltrates organise within the inflamed RA synovial joint as B/T cell discrete areas with germinal centre-like structure and activity. This tertiary lymphoid tissue (ectopic lymphoid follicle, ELF) containing follicular

dendritic cells is present in around 40% of RA patients and is promoted by synovial fibroblasts via secretion of chemokines such as CXCL13, CXCL21 and CXCL12. Moreover, effector B cells may contribute to ELF formation and osteoclastogenesis by secreting lymphotoxin-β and RANKL, respectively (Hamze *et al*., 2013).

Synoviocytes, by producing B-cell activating factor and a proliferation-inducing ligand, also propagate local antigen-specific autoreactive responses through AID-mediated somatic hypermutation and class switch recombination in B cells (Jones and Jones, 2016). B cell clones within ELFs exhibit highly mutated variable V_H and V_L genes which have been acquired locally as confirmed by lineage-tree analysis suggesting a critical role for ELFs in the initiation or/and amplification of RA pathology. In some patients, over 30% of the general synovial B cell response has a bias toward citrullinated antigens. Lymphoid aggregates can also form in extra-articular sites such as the lung and subchondral bone marrow. ELF formation is closely correlated with T follicular helper (Tfh) cells which perpetuate chronic autoantibody production through CD40L and IL-21 signalling (Bombardieri, Lewis and Pitzalis, 2017).

B cells can also serve as efficient APCs to activate T cells and enable optimal memory development in the CD4+ T cells. In RA, their antigen uptake was observed to be highly selective, in particular for RF+ B cells, as compared to nonspecific uptake by other professional APCs. Moreover, B cells could activate pathogenic T cell reactions via cytokines such as IL-6 which is produced in chronic autoimmune responses. This pro-inflammatory cytokine regulates the balance between Th17 cells and Tregs in the presence of TGFβ. Dysregulation of IL-6 leads to Th17-medited B cell proliferation, differentiation, and antibody production suggesting a positive feedback loop between these two cell types (Bugatti *et al*., 2014). In addition, effector B cells can affect other pathologically relevant populations through the release of a wide spectrum of inflammatory cytokines including TNF-α, IL-12, IL-23 and IL-1α (Hamze *et al*., 2013).

The role of B cells in regulating bone homeostasis is believed to be developmental and stage-dependent. Inflammatory cytokines (IL-6, IL-1, IL-17, TNF-α) together with RANKL, promote bone loss in rheumatoid joint.

Moreover, ACPA-mediated osteoclastogenesis, which likely promotes bone resorption and disease perpetuation, has also been observed in RA patients (Meednu *et al*., 2016; Kurowska *et al*., 2017). Recent findings also indicate the coexistence of a different B cell subset termed B regulatory cells (Bregs). In active RA, the immunosuppressive function of these cells on Th1 and Th17 responses is partly impaired (Bankó *et al*., 2017). In addition, Breg levels are reduced in the systemic circulation and enriched in the synovial fluid of RA patients, when compared to healthy individuals, which might indicate their active migration to the inflamed synovium (Flores-Borja *et al*., 2013).

1.3.7.2. The role of T Lymphocytes in RA

For many decades RA has been considered a T-cell driven disease. CD4+ T cells are the predominant subtype of lymphocytes in RA synovial infiltrate and appear to be central in the disease pathogenesis (Aterido *et al*., 2014). They contribute to tissue damage, both in the synovium and in extra-articular sites, via diverse mechanisms including pro-inflammatory cytokine release and cooperation with B cells for autoantibody production. Breakdown of immunological tolerance and enhanced self-antigen recognition by CD4+ T cells is associated with certain MHC II alleles (particularly HLA-DRB1) and dysregulation of signalling molecules such as PD-1 and LYP. In RA synovium, infiltrating CD4+ T cells interact with DCs and B cells which are crucial for their co-stimulation and thus for the progression of the disease. Through the production of TNF-α, IFN-γ, IL-1α and IL-15, CD4+ T cells reduce collagen production by FLS and induce the release of pro-inflammatory cytokines by FLS and

macrophages, which triggers cartilage damage (Castro-Sánchez and Roda-Navarro, 2017). CD4+ T cells are also a major source of RANKL which induces osteoclast generation and therefore contributes to bone resorption (Elshahaly *et al*., 2012). Hyper-activation of this cell population in RA has been reported not only for synovium-infiltrating but also for peripheral blood T cells (López-Santalla *et al*., 2011). In addition to distortions of the CD4+ T cell repertoire and phenotype in RA patients, premature cellular aging with telomere shortening in this cell population appears to contribute to their autoreactivity (Ponchel *et al*., 2012).

Upon activation, CD4+ T cells differentiate into distinct effector subsets, Th1, Th2, Th17, Tfh and Treg cells, which may become dysregulated in RA and drive synovitis. Disease activity appears to correlate with Th1/Th2 balance biased towards the former subset with increased IFN-γ and TNF-α expression and reduced expression of IL-6 and IL-10 by Th2 subset in RA plasma. Conversely, this balance is also disturbed by RA microparticles via TNF-α reduction and IL-6 and IL-10 increase (He *et al*., 2017).

Th17 cells play a predominant role in RA pathology through the release of IL-17 in the synovium. This cytokine promotes inflammation by stimulating the production of pro-inflammatory mediators including TNF-α, IL-6, IL-1β, chemokines and metalloproteinases which induce osteoclastogenesis and cartilage damage. Moreover, increased IL-17 levels correlate with more severe synovial lesions (Noack and Miossec, 2014). The pathogenic role of IL-17 is corroborated by the beneficial effects of the Th17/IL-17 inhibitor (secukinumab) in the treatment of RA (Li *et al*., 2012; Blanco *et al*., 2017). Tregs are present at increased levels in rheumatoid synovium compared with peripheral blood indicating an active recruitment of this cell subset to the affected joint. In RA, they exhibit impaired suppressive functions associated with polymorphisms in their suppressive molecules and are unable to control the production of pro-inflammatory cytokines such as TNF-α and IL-6 (Morita *et al*., 2016). However,

due to developmental and functional discrepancies between Treg subsets among RA populations, their specific identification and role in RA is still elusive (Fessler *et al*., 2017).

Furthermore, there is a growing evidence base which implies that CD8+ T cell homeostasis is dysregulated in RA and could play a role in disease pathogenesis. The contribution of this cell population to RA is demonstrated by the association of MHC I polymorphisms with disease risk. CD8+ T cells comprise around 40% of all infiltrating T cells in rheumatoid synovium and are detected in the preclinical stages of RA development. The levels of memory CD8+ T cells appear to be increased in the peripheral blood and synovial fluid from RA patients when compared to healthy individuals suggesting their active generation. The frequency of these cells correlates with IgM-RF levels in the serum. Through the production of cytolytic proteins, CD8+ T cells are also vital for the formation of ectopic germinal centres in RA synovium (Carvalheiro, da Silva and Souto-Carneiro, 2013; Carvalheiro *et al*., 2015; Ramwadhdoebe *et al*., 2016). In arthritic joints, CD8+ T cells constitute a heterogenous population with both pro-inflammatory function via IL-6/TNF-α release and anti-inflammatory effects through IL-4/IL-10. The regulation and balance of both of these cell subtypes was reported to be impaired and resistant to suppression in RA. Moreover, the TCR repertoire of synovial CD8+ T cells in RA has been observed to be specific for several types of viruses including EBV, cytomegalovirus and influenza virus (Petrelli and van Wijk, 2016).

1.3.7.3. The role of Rheumatoid Factor (RF) autoantibody in RA

Rheumatoid factors, first reported by Waaler in 1940 as antibodies reacting with gamma globulins, are autoantibodies targeting the C-terminal part of the constant

region (Fc) of IgG heavy chain in the γ2-γ3 cleft. They are detected in 60-80% of RA patients and together with ACPAs, represent the major serological marker for disease diagnosis in the current classification criteria. The presence of these autoantibodies precedes the onset of RA symptoms by several years. RFs have also been detected in other autoimmune diseases such as systemic lupus erythematosus and Sjögren syndrome as well as in non-autoimmune conditions. Moreover, RF positivity has been observed in normal individuals upon secondary immune responses to infections and immunisations. However, in healthy subjects these antibodies are typically polyreactive IgMs of low affinity and act in host defence through the generation and clearance of immune complexes (IC). Their global variability in distribution appears to be affected by genetic and environmental elements, with the highest prevalence of up to 30% reported in North American Indians (Ingegnoli, Castelli and Gualtierotti, 2013; Anquetil *et al*., 2015). They resemble natural antibodies generated by CD5+ B1 cells which unlike B2 cells do not undergo somatic hypermutation. In RA, RF+ B cells are activated more efficiently by immune-complexed IgGs when compared to their monomeric variant, implying that effective BCR cross-linking is necessary for the propagation of B cell activation signal (Song and Kang, 2010a). Moreover, RF may be produced as a result of B cell stimulation by EBV which acts as polyclonal activator (Adtani and Malathi, 2015). EBV has been reported at elevated frequencies in rheumatoid synovium (Tan and Smolen, 2016), particularly in RF positive RA patients (Westergaard *et al*., 2015). Local formation of RF has been observed in the inflamed synovium where ectopic lymphoid follicles are found (Song and Kang, 2010b). In RA, RFs undergo isotype switching from IgM to IgA and IgG and affinity maturation which might be T cell dependent. Due to their local production, levels of RFs in the synovium are likely higher than those in the circulation (Jones *et al*., 2013). The arthritogenic role of RF remains unclear, however high IgA RF titres have been associated with bone

erosions, extra-articular symptoms, and poor prognosis. High-affinity RFs in synovium may potentiate inflammation and antigen trapping through enhanced IgG-binding and complement fixation (Song *et al*., 2010b). Moreover, IgG RFs can form large ICs and stimulate inflammatory cells leading to RA perpetuation. High RA activity, independent of ACPA presence, is also exhibited by patients with RFs (Tan and Smolen, 2016).

Pepscan and *in-silico* epitope mapping studies have contributed to identification of potential RF epitopes within CH_2 and CH_3 domains of IgG1 Fc in the pathological mechanism of RA. It is apparent that a series of 7 to 10 CH_3 regions constitute potential major conformational IgG sites targeted by RF (Westwood *et al*., 2008; Nelson *et al*., 2003). Although RFs are not RA-specific, they constitute important diagnostic and prognostic markers and may be the first autoantibodies to appear at the site of disease (Brink *et al*., 2016).

1.3.7.4. The role of Anti-Citrullinated Protein Antibody (ACPA) in RA

The other major group of autoantibodies present in around 60 to 70% of RA patients' sera constitute anti-citrullinated proteins/peptide antibodies (ACPAs) (Fernandes-Cerqueira *et al*., 2015). They also occur in other conditions such as multiple sclerosis (MS), psoriatic arthritis and SLE but at a reduced frequency (Abdel Fattah *et al*., 2009; Singh *et al*., 2011). ACPAs have been observed to exhibit a high specificity for RA although the existence of citrullinated synovial proteins is not disease-specific (Jilani and Mackworth-Young, 2015). ACPA positivity has been incorporated in the 2010 ACR/EULAR classification criteria for RA and they have been shown to predate the onset of RA by several years (Valesini *et al*., 2015). The targets of ACPAs identified thus far in RA constitute numerous citrullinated proteins including fibrinogen, fillagrin, vimentin, type II collagen, α-enolase, aggrecan and several viral antigens (Glaant *et*

al., 2016). However, recognition of abundant citrullinated autoantigens might suggest that no traditional antigenic determinant is targeted by ACPAs and that the culpable autoantigens can indeed change over time instigating new reactivities in the ACPA responses (Trier *et al*., 2015). Therefore, antibodies generated against citrullinated epitopes could sustain the immune response via epitope spreading and cross-reactivity with a broad array of citrullinated proteins (Pratesi *et al*., 2013; Quirke *et al*., 2014).

1.3.8. Treatment strategies in RA

The treatment of rheumatoid arthritis has undergone significant changes over the last 20 years and focused on rapid diagnosis and treat-to-target approach. There is currently no reliable cure for RA thus the existing treatment strategy has evolved towards early, aggressive therapy to increase the likelihood of clinical remission, to prevent joint damage and systemic manifestations (Losina and Katz, 2017). An optimum therapeutic window of opportunity of up to 6 months at the onset of RA symptoms appears to have lasting benefits in terms of disease progression (Raza and Filer, 2015).

Treatment of RA is a multifaceted process which incorporates both pharmacological and non-pharmacological approaches such as physical therapy, lifestyle changes and surgical intervention (Sophia and Ramesha, 2017). Conventional RA therapies include non-steroidal anti-inflammatory drugs (NSAIDs) such as aspirin, diclofenac, ibuprofen, and sulphonamides, which inhibit inflammation by blocking prostaglandin production. They are effective analgesic drugs, which improve physical function but do not prevent joint degradation and may cause many adverse effects such as bleeding and gastrointestinal toxicity (Crofford, 2013). Glucocorticoids such as prednisone are a class of steroid hormones which have rapid disease-modifying and anti-inflammatory

effects. However, they are associated with long-term adverse events such as hypertension or osteoporosis and are typically used as adjunct therapy (Cantley, Smith and Haynes, 2009).

One of the most widely used approaches for the effective treatment of RA are disease-modifying antirheumatic drugs (DMARDs) which have anti-inflammatory properties and inhibit structural damage progression. They are of synthetic and biological origin and are used for long-term disease control. Synthetic DMARDs are further defined as conventional synthetic (e.g. methotrexate, leflunomide, hydroxychloroquine, gold salts) and targeted synthetic such as janus kinase (JAK)-inhibitors (tofacitinib, baricitinib), with the latter modulating specific inflammatory targets (Smolen, Aletaha and McInnes, 2016).

Biological DMARDs achieve their therapeutic effect through TNF-α, IL-1 and IL-6 receptor inhibition, T cell co-stimulation blockade and B cell depletion. Currently, five TNF-α inhibitors including monoclonal antibodies (infliximab, adalimumab, certolizumab pegol, and golimumab) and a TNF-receptor construct (etanercept) are approved for RA therapy. Whereas monoclonal antibodies tocilizumab, abatacept and rituximab target IL-6 receptor, T cell co-stimulation and B cells, respectively (Rein and Mueller, 2017). Anakinra together with TNF inhibitors constitute 'first-generation' biologicals and acts as an IL-1 receptor antagonist. However, the clinical potency of this approach is inferior to that of anti-TNF-α therapy. The clinical efficacy of biological DMARDs is comparable across all types and DMARDs can be used as monotherapy or in combination with other pharmaceutical agents.

Although a significant proportion of RA patients respond favourably to DMARD treatment, there have been reports of toxicity and numerous adverse events such as gastrointestinal complications, renal damage and serious infection risk (Umićević

Mirkov and Coenen, 2013). Due to their substantial cost, the use of DMARDs has been restricted. However, the advent of biosimilars, which are clinically equivalent to existing biological agents such as infliximab, rituximab or etanercept, provides an appreciably cheaper alternative, which makes them widely available (Smolen *et al*., 2017). Nevertheless, despite recent advances in therapeutic approaches for RA, up to 40% of patients do not respond to treatment and only 20-25% of patients achieve complete remission (Zampeli, Vlachoyiannopoulos and Tzioufas, 2015). Therefore, the need for new RA therapies and strategies, which could achieve cessation of disease progression and disability, still remains.

1.3.9. Animal models of RA

Animal models of arthritis can function as informative tools to analyse the pathophysiological pathways of the developing immune response, establish drug targets and test potential therapeutic agents (Mcnamee, Williams and Seed, 2015). The most commonly utilised animal species, which recapitulate many features of human RA, are mice and rats.

RA animal models are categorised into induced models (via arthritogenic stimuli or active/passive immunisation) and spontaneous models (Fischer *et al*., 2017). Collagen-induced arthritis (CIA) is the archetypical experimental model of inflammatory arthritis induced in genetically susceptible animals (MHC haplotype q) by immunisation with collagen II, the major component of articular cartilage. Mice or rats progress on to acute erosive polyarthritis within three weeks of onset with extensive cellular infiltrates and joint damage. CIA species exhibit autoreactive B/T cell responses and produce both RF and ACPA autoantibodies. On the other hand, in antigen-induced arthritis animals develop monoarticular pannus and lymphoid follicles upon injection of

exogenous protein antigen (e.g. methylated bovine serum albumin) (Alves *et al*., 2016). Whereas in streptococcal cell-wall induced model animals exhibit systemic symptoms with chronic erosive polyarthritis which is thought to be caused by microbial superantigens that stimulate T lymphocytes with specific *Vβ* genes in an antigen-independent fashion.

Arthritis can also be induced in susceptible strains of rats following administration of adjuvants such as complete or incomplete Freund's adjuvant, pristine and squalene, possibly due to increased response to self-antigens. Animals in this model exhibit chronic relapsing disease course or go into remission depending on the adjuvant used (Mcnamee, Williams and Seed, 2015). Autoreactive sera from RA patients or CIA can be used to generate collagen antibody-induced arthritis (CAIA) model. In this experimental model, animals develop invasive pannus which leads to cartilage and bone damage within 8 days, however the disease induction is B and T lymphocyte independent (Caplazi *et al*., 2015). In K/BxN antibody-transfer arthritis model, mice spontaneously develop a progressive arthritis with rapid onset. This model is mediated by neutrophils, immune complex accumulation and TNF-α and therefore focuses on immune recruitment and synovial damage (Rohrbach *et al*., 2012).

A number of inbred mouse strains are susceptible to development of spontaneous arthritis and several models have been established including SKG, GP130, IL-1 $RA^{-/-}$, and TNF transgenic models. However, these models do not exhibit full penetrance and reproducible arthritis progression (Kollias *et al*., 2011; Fischer *et al*., 2017). Studies involving animal models in autoimmunity are rendered difficult due to the complexity of disease susceptibility loci which may or may not be present in different species. Moreover, certain translated targets such as endogenous retroviruses, which are of primary interest in this thesis, may be specific to humans therefore no suitable animal model might be available for their assessment *in vivo* (Curtin *et al*., 2015).

Nevertheless, this research has been instrumental to the development of novel biological agents and can serve as a useful tool in considering the heterogeneity of human RA (Germolec *et al*., 2012).

1.4. Human Endogenous Retroviruses (HERVs)

1.4.1. Introduction

Retroviruses, found in all mammals, abundant among vertebrates and other species, constitute a diverse family of small ssRNA viruses capable of reverse transcribing their genomes into dsDNA molecule and integrating it into the chromosomal DNA of the host. Retroviruses exist in exogenous or endogenous forms. Exogenous retroviruses such as human immunodeficiency virus (HIV) and human T cell leukaemia virus (HTLV) target primarily somatic cells and remain infectious through the production of viral particles, or virions. Whereas endogenous retroviruses (ERVs) have become permanently embedded within germline DNA and are inherited in a Mendelian fashion (Escalera-Zamudio and Greenwood, 2016). ERVs are therefore relics of ancestral, and possibly extinct, exogenous retroviruses. Upon insertion into the host's germline, most ERVs have become defective for replication through frameshift mutations, insertions and deletions. They are therefore unable to assemble their virions and thus transmit infections between individuals. However, certain ERVs such as HERV-K are replication competent and can express their virus-like particles (Weiss, 2016). Cross-species viral transmission has been reported between baboons and cats (Shimode, Nakagawa and Miyazawa, 2015) as well as between gibbons and koalas (Tarlinton, Meers and Young, 2006). Moreover, the process of ongoing viral endogenisation has been reported in different extant species such as mouse or koala where both forms of the virus co-exist and can be transmitted horizontally as well as vertically. However, the vast majority of endogenised viruses constitute non-infectious 'retroviral fossils' including human ERVs (Vargiu *et al*., 2016). Although HERVs can be found in all nucleated cells, their mRNA and protein expression varies among individuals. They are broadly distributed

throughout the genome and exist as single or multicopy retroelements on chromosomes as a result of their numerous amplifications via reinfection/retrotransposition (Trela, Nelson and Rylance, 2016; Nelson *et al*., 2014b).

HERVs have similar although basic genomic structure common to exogenous retroviruses and are composed of three highly conserved genetic regions, *Gag*, *Pol* and *Env*, flanked by two long-terminal repeats (LTRs) which regulate viral gene insertion and expression. Group specific antigen (*Gag*) codes for viral matrix, capsid and nucleocapsid, *Pol* encodes enzymes reverse transcriptase, protease and integrase, and *Env* encodes the surface and transmembrane envelope domains (Fig. 1.04) (Ryan, 2016).

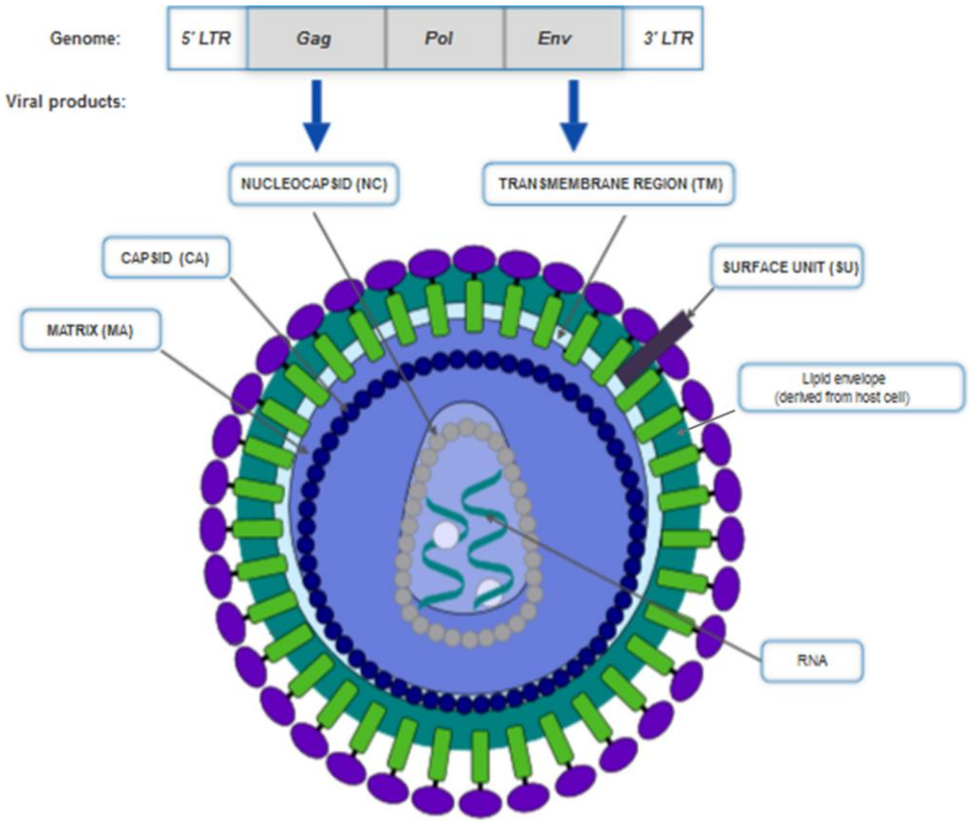


Figure 1.04. Schematic diagram of structural organisation of human endogenous retrovirus with the basic genomic structure and encoded proteins.

Due to the lack of standardised nomenclature, HERVs have been classified within three groups based on sequence and structural similarity in the *Pol* region to present day exogenous retroviruses. Thus, Class I, Class II and Class III HERVs are related to Gammaretroviruses, Betaretroviruses and Spumaviruses, respectively (Table 1.02) (Tugnet *et al*., 2013). Phylogenetic studies allowed identification of around 30 distinct HERV groups ('families'), with the majority expressed in old-world monkeys and some recently integrated elements such as HERV-K (HML-2) which are human-specific and sporadically insertionally polymorphic (Fig. 1.05) (Young, Stoye and Kassiotis, 2013).

Class I	Class II	Class III
Group 1: HERV-HF HERV-H (RTVL-H, RGH) HERV-F (HERV-FXA, HERV-Fb) HERV-F/H (HERV-Fc1, HERV-Fc2 **Group 2: HERV- RW** HERV-W HERV-R (ERV-9) HERV-P (HuERS-P, HuRRS-P) **Group 3: HERV- ER1** HERV-E (4-1, ERVA, NP-2) 51-1 HERV-R (ERV-3) RRHERV-1 **Group 4: HERV-T** HERV-T (CRTK1, CRTK6, S71) **Group 5: HERV-IP** HERV-I (RTVL-I) HERV-IP-T47D (ERV-FTD) **Group 6: HERV-FRD** ERV-FRD **Other: HRES-1**	Group 1: **HML-1** HERV-K (HML-1.1) Group 2: **HML-2** HERV-K10 HERV-K18 HERV-K103 HERV-K105 HERV-K111 HERV-K113 HERV-K115 HERV-K-HTDV **Group 3: HERV-K** (HML-3.1) **Group 4: HML-4** HERV-K-T47D **Group 5: HML-5** HERV-K- NMWV2 **Group 6: HML-6** HERV-K (HML-6p) **Group 7: HML-7** HERV-K- NMWV7 **Group 8: HML-8** HERV-K- NMWV3 **Group 9: HML-9** HERV-K- NMWV9 **Group 10: HML-10** HERV-KC4	**HERV-L** **HERV-L** (HERV18) **HERV-U2** **HERV-U3** **HERV-S**

Table 1.02. Classification of human endogenous retroviruses (HERVs). Class I HERVs cluster with Gammaretroviruses, Class II with Betaretroviruses, Class III with Spumaviruses. Adapted from Trela *et al.*, 2016.

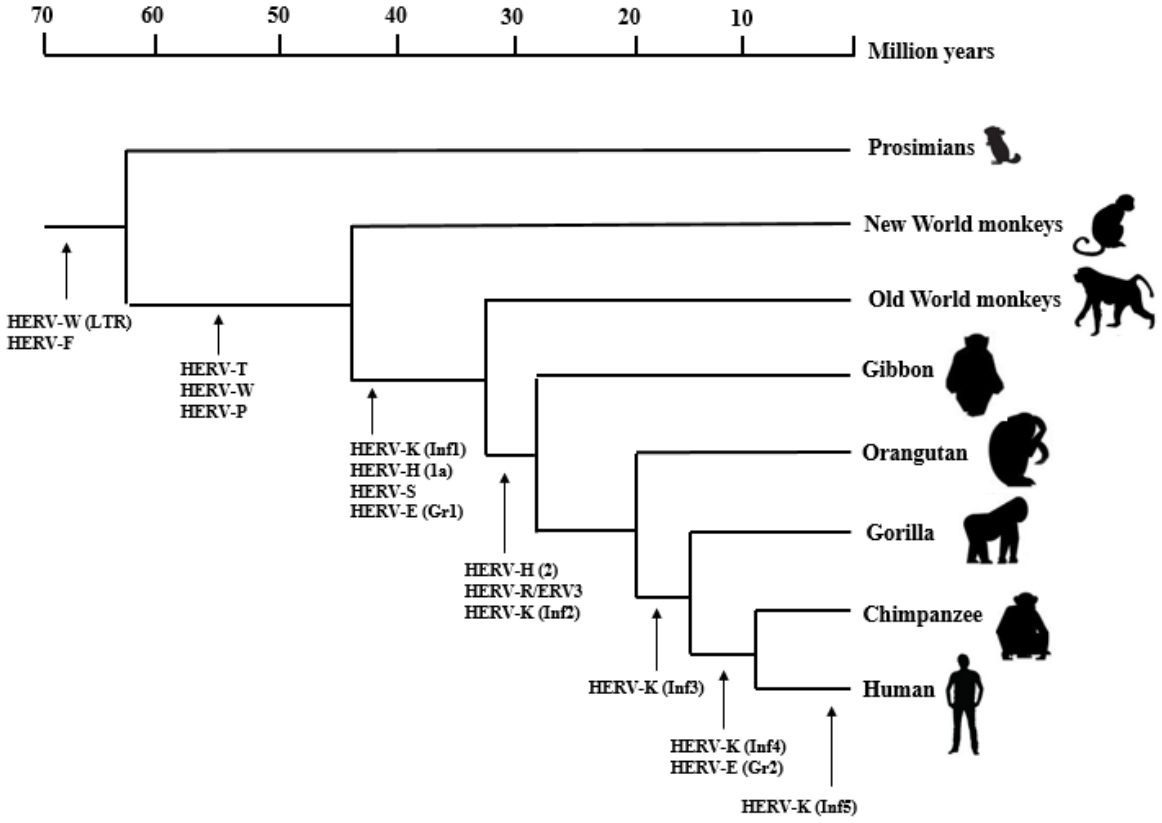


Figure 1.05. Cladogram showing events resulting in the introduction of HERV groups within the primate lineage. Adapted from (Freimanis, 2008).

1.4.2. HERV expression in humans

The haploid human genome occupies a total of 3 billion DNA base pairs of which only about 3% code for proteins. In contrast, long terminal repeat (LTR) retrotransposons namely HERVs and their solitary LTRs, represent approximately 8% of human DNA sequences (Fig. 6) which contribute to over 500,000 retroelements present in the human genome (circa 45%) (Yu, Zhao and Zhu, 2013).

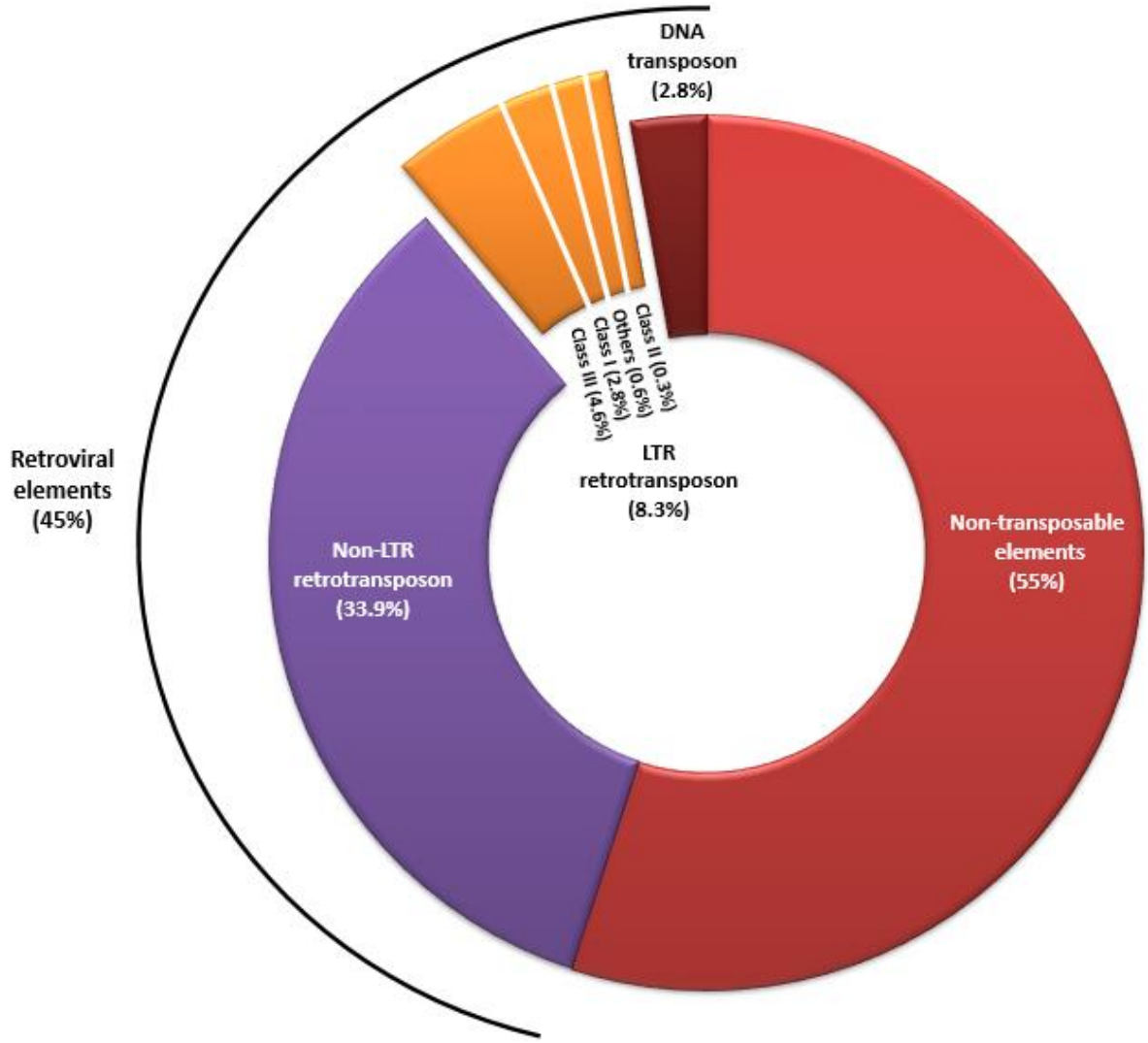


Figure 1.06. Schematic diagram representing approximate DNA breakdown of the human genome composition highlighting percentage of HERV classes. Reproduced from (Trela, Nelson and Rylance, 2016)**.**

Multiple HERVs integrated into primate genomes over 30 million years ago and have been passed onto offspring through vertical inheritance (Krzysztalowska-Wawrzyniak *et al*., 2011). Even though the majority of retrotransposons are present in the intergenic/intronic regions, some of them harbour promoters and enhancers and can exhibit relevant host function in promoting mRNA synthesis of nearby host genes (Pi *et al*., 2010). HERV LTRs naturally possess transcriptional regulatory signals thus their involvement in regulation of genes is very probable (Cohen, Lock and Mager, 2009). Therefore, depending on their location relative to host genes, HERVs can serve as transcriptional promoters or repressors of adjacent genes.

Although the acquisition of these transposable elements is usually considered to be detrimental, parasitic or redundant, depending on the HERV, their existence may be neutral or they can impart a beneficial effect on mammalian chromosomes by facilitating genomic and regulatory evolution. This advantage for the population outweighs possible negative effects that are likely to be removed through selection. HERVs can also serve as restriction factors and therefore interfere with other infecting exogenous retroviruses by blocking viral receptors on the host cell (Weiss, 2016).

In addition to providing protection against exogenous viral infection, several endogenous viral gene products have evolved to contribute to other important physiological functions. The best-characterized of these proteins are known as syncytins which constitute Env fragments of HERV-W and HERV-FRD known as syncytin-1 and syncytin-2, respectively. They serve in human placental development through syncytiotrophoblast formation and cell fusion promotion. Syncytin-2, HERV-H Env and HERV-K proteins exhibit immunosuppressive properties which may serve in protection of the developing foetus from attack by the maternal immune system. ERV-3 (HERV-R) enhanced expression has been observed in different foetal tissues and may play a role in placental development and cellular differentiation (Kaer and Speek,

2013; Jern and Coffin, 2008). The enhanced expression of numerous HERV families in human placentas has likely resulted from repressing LTR hypomethylation within these tissues (Schroeder *et al*., 2015). Syncytin-1 may also exert its fusogenic activity in striated muscles and osteoclasts (Søe *et al*., 2011). Moreover, increased activity of HERV-H and HERV-K has been observed in early human embryogenesis and embryonic stem cells. Therefore, it has been suggested that these families may contribute to embryonic development and pluripotency (Meyer *et al*., 2017).

HERVs, through the insertion of their regulatory sequences, can also generate novel patterns of genomic expression. This has been observed in ERV-9 which has an important role in the maintenance of male germ line integrity (Liu and Eiden, 2011) and in HERV-E which controls the expression of salivary amylase and may have had an impact on high-starch diet in humans (Gonzalez-Hernandez *et al*., 2012). More examples of possible functional effects of HERV retroelements can be found in Table 1.03.

Retroelement	Role	References
Alu	Autoimmune lymphoproliferative syndrome (ALPS); Hyper-immunoglobulin M (HIGM) Syndrome Affects molecular organisation of the human MHC class II region	(Tighe *et al*., 2002; Apoil *et al*., 2006; Andersson *et al*., 1998)
ERV-1 (MER21A)	Involved in endometriosis and breast cancer; Protein glycosylation in humoral immunity; tumor progression; Oestrogen synthesis	(Conley and Hinshelwood, 2001; Oduwole *et al*., 2004; Cohen, Lock and Mager, 2009; Tsuchiya *et al*., 2005; Harduin-Lepers *et al*., 2001)
HERV-E (LTR-2B)	Neurodegenerative diseases; Heparin-binding growth factor in neural development	(Kadomatsu and Muramatsu, 2004; Deuel *et al*., 2002; Marchionini *et al*., 2007)
HERV-K (DQ-LTR3)	Enhanced susceptibility to RA and IDDM (LTR present in the gene coding for MHC class II receptor HLA-DQ)	(Larek-Rapala *et al*., 2011; Seidl *et al*., 1999)
HERV-K (DRB7)	MHC class II-dependent superantigen in IDDM	(Andersson *et al*., 1998)
HERV-P (LTR9)	Inhibition of apoptosis and innate immunity	(Liston, Fong and Korneluk, 2003; Chamaillard *et al*., 2003)
LINE (L1)	Chronic granulomatous disease	(Meischl *et al*., 2000)
SVA	X-linked agammaglobulinemia	(Rohrer *et al*., 1999)
HERV-P	Neuronal apoptosis inhibitory protein	(Romanish *et al*., 2007)
MER39B	Immune and cardiovascular systems	(Robson, Sévigny and Zimmermann, 2006)
ERV-1 (MER41E)	Pregnancy (pre-eclampsia)	(Overgaard *et al*., 2001; Winn *et al*., 2009)

ERV-L (MER54B)	Cancer; Cell-surface glycoprotein with secreted cytokine component	(Chang and Pastan, 1996; Hassan and Ho, 2008)
HERV-9 (LTR12C)	Immune response to intracellular pathogens; Axon guidance in brain and cell-cell interactions in immune system	(Oduwole *et al*., 2004; Tsuchiya *et al*., 2005; Kumanogoh and Kikutani, 2004)
ERV-9 LTR	Regulation of basal and tissue-specific retroviral transcription in MHC class II region (HLA-DR)	(Andersson *et al*., 1998)
MaLR (THE1D)	T-cell proliferation	(Minami *et al*., 1993)
HERV-E (4-1)	Regulation of the human salivary amylase	(Samuelson, Phillips and Swanberg, 1996; Ting *et al*., 1992)
HML-5 (LTR22B)	Insulin-like growth factor involved in placental morphogenesis	(Bièche *et al*., 2003)
HERV-9 (LTR12D)	Signalling regulation (oxidoreductases)	(Gabrielli *et al*., 1995; Shafqat *et al*., 2006)
HERV-E	Apolipoprotein CI	(Medstrand, Landry and Mager, 2001)
HERV-H	Phospholipase A2-like	(Feuchter-Murthy, Freeman and Mager, 1993)
HERV-L (MLT2B3)	Carbohydrate synthesis	(Isshiki *et al*., 1999)
HERV-E (LTR2)	Lipid metabolism	(Shachter, 2001)

Table 1.03. Association of retroviral elements with human physiological functions and disorders as reported previously within the literature.

1.4.3. HERVs and human disease

HERVs have been repeatedly discussed as etiological factors in a large number of clinical human pathologies. HERVs association with multiple human malignancies have been well demonstrated and they appear to exert their tumourigenic effects through oncogene transduction and retroviral insertion near protooncogenes. HERV-K mRNA and proteins are overexpressed in ovarian cancers, melanoma tissue and cell lines as well as in germ cell tumours (GCT) such as teratocarcinoma, where retrovirus-like particles are released (Hohn, Hanke and Bannert, 2013). Increased levels of HERV-K *Env* RNA and immunoglobulins against viral proteins have also been found in breast and prostate cancer patients whilst *Gag* RNA was also detected in these cancers and in chronic lymphocytic leukaemia patients (Wang-Johanning *et al*., 2014; Wallace *et al*., 2014). HERV-W-derived syncytins have been detected in human malignancies and are suggested to play a role in cancer cell modification or metastasis (Kassiotis, 2014). Both HERV-W and HERV-K Env domains expressed in tumours may exert immunosuppressive properties and therefore prevent its rejection (Morozov, Dao Thi and Denner, 2013). In addition, HERV-H products have been detected in colon cancer tissues (Alves *et al*., 2008) and in Hodgkin's lymphoma patients (Woo *et al*., 2014).

HERVs have been demonstrated to be associated with viral infections through mechanisms which may involve cooperation or recombination with exogenous viruses. HIV-1 infection appears to increase the expression levels of HERV-K (HML-2) RNA in peripheral blood mononuclear cells (PBMCs) or serum from affected patients. Moreover, enhanced expression of HERV-K Gag products has been reported in CD4+ and CD8+ T lymphocytes from HIV-1 infected individuals (van der Kuyl, 2012). It has been reported that HERV-K may also contribute to the perpetuation of long-term

infection by Epstein-Barr virus and human herpes virus 6 (HHV-6) leading to the development of associated clinical conditions (Vincendeau *et al*., 2015).

Enhanced expression of HERV products has also been reported in a variety of neurological diseases. Increased levels of HERV-W Env and Gag proteins have been observed in muscle biopsies from patients with amyotrophic lateral sclerosis (ALS), whereas HERV- K activity was found in the brain tissue from those patients. Both HERV-W and HERV-K are also upregulated in CSF, blood and brain biopsies from patients with schizophrenia and bipolar disorder (Douville and Nath, 2014). In multiple sclerosis (MS), which constitutes an autoimmune neurodegenerative disease, HERV-W Env proteins were found in PBMCs and brain lesions of affected patients. The corresponding HERV-W mRNA levels have been frequently detected in cerebrospinal fluid (CSF), plasma and brain samples from patients with MS (do Olival *et al*., 2013). Moreover, enhanced expression of HERV-H/F family HERV-Fc1 RNA has been reported in T cells and plasma from individuals with MS (Laska *et al*., 2012). Various HERVs have also been implicated in other autoimmune disorders (Table 1.04), which will be discussed in greater depth in the following section.

Autoimmune disorder	Human Endogenous Retroviruses
Rheumatoid Arthritis (RA)	HERV-K10 HTLV-1 (p19 antigen) HERV-W (MSRV) ERV-9 HERV-K113 HERV-L ERV-3/λ4-1
Systemic Lupus Erythematosus (SLE)	HERV-E (Clone 4-1) HERV-H HERV-K HRES-1 HIAP-1 ERV-3
Multiple Sclerosis (MS)	HERV-W (MSRV) HERV-H/F HERV-Fc1 HERV-K18
Insulin Dependent (Type 1) Diabetes Mellitus (IDDM)	HERV-K18 LTR13
Sjögren's Syndrome (SS)	HERV-E *env* protein (4-1) HRES-1 HIAP-1 HERV-K113
Juvenile Idiopathic Arthritis (JIA)	HERV-K18
Psoriasis	HERV-E HERV-K HERV-W ERV9
Opitz Syndrome	HERV-K HERV-E
Congenital heart block	ERV-3
Autoimmune liver disease (Primary biliary cirrhosis)	HIAP
Addisons disease	LTR13
Alopecia areata	HIAP
Essential thrombocytopenia	HERV-K10
Mixed connective tissue disease	HERV-E (4-1)

Table 1.04. HERVs associated with autoimmune diseases. Adapted from (Trela, Nelson and Rylance, 2016; Freimanis, 2008).

1.4.4. HERVs and their association with RA

Several exogenous viruses including EBV, parvovirus B19, cytomegalovirus and hepatitis C have been detected within the synovial tissue, although there is no clear causative association between these pathogens and RA (Tugnet *et al*., 2013). The potential contribution of endogenous retroviruses in the aetiology and pathogenesis of rheumatoid arthritis partly stems from the ineffective attempts to detect known human infectious retrovirus, such as HIV-1 or HTLV-1 in the rheumatoid synovium (di Giovine *et al*., 1994). Although antibodies to both of these viruses have been reported in RA sera, molecular investigations failed to identify viral sequences, which would indicate active infection (Nelson *et al*., 1994). Therefore, cross reactions with HERVs or autoantigens were suggested (Brookes *et al*., 1992). Moreover, HTLV-I-derived (p19, p24) and HIV-derived (p17, p24) antigens were detected in RA synovium, however no antibody reactivity was demonstrated to those targets (Ziegler *et al*., 1989; Ziegler and Thomas, 1991). In addition, viral particles resembling those of type C retrovirus have been reported in RA synovial fluid (Kalden and Gay, 1994).

Human retrovirus 5 (HRV-5) has also been implicated in the disease upon detection of increased levels of proviral DNA in the inflamed synovium and blood samples from rheumatoid patients (Griffiths *et al*., 1999), however this finding was not further corroborated (Piper *et al*., 2006). With the advent of more sensitive assays, such as reverse transcription-polymerase chain reaction (RT-PCR), enhanced levels of several HERVs including HERV-K, HERV-L and ERV-9 were detected in synovial fluid from patients with RA (Nakagawa *et al*., 1997). ERV3 mRNA and ERV λ4-1 mRNA species were found in PBMCs and synovial tissues from RA patients, however their expression was also detected in osteoarthritis and healthy subjects (Takeuchi *et al*., 1995). HERV-E CD5 has been reported in B cell subsets (B-1 type) which are responsible for the generation of rheumatoid factor autoantibodies in rheumatoid patients. This virus is

believed to be involved in the formation of truncated CD5 protein end products, which in turn lead to reduced BCR transduction and autoantibody production (Renaudineau *et al*., 2005).

Moreover, enhanced HERV-K18 superantigen transcripts levels have been reported in PBMCs and synovial fluid mononuclear cells from patients with juvenile rheumatoid arthritis (Sicat, Sutkowski and Huber, 2005). Whilst the prevalence of a relatively young virus HERV-K113 was significantly increased among RA patients (Krzysztalowska-Wawrzyniak *et al*., 2011). HERV-K (HML-2) Rec products have been detected in arthritis and normal synovia, however the authors concluded that it is unlikely to be causal of the disease mechanism. They also observed that HERV-K (HML-2) Env-derived transcripts expression was decreased in rheumatoid synovial cells when compared to normal synovium (Ehlhardt *et al*., 2006). On the other hand, Mameli and colleagues reported an increased reactivity of RA serum to HERV-K-6 Env region in Sardinian population (Mameli *et al*., 2017).

The viral loads of HERV-K (HML-2) type 1 and type 2 were found to be increased in RA plasma and synovial fluid samples from individuals with RA and correlated with the disease activity (Reynier *et al*., 2009). Nelson and colleagues reported a significant increase in the frequency of HERV-K10 mRNA in RA synovial fluid, when compared to healthy controls (Nelson *et al*., 1999), which was further corroborated in PBMCs obtained from rheumatoid patients (Ejtehadi *et al*., 2006). Significant up-regulation of HERV-K gag1 gene transcript levels in PBMCs and of serum antibody response to HERV-K10 Gag region (KPR) has been reported in rheumatoid patients by Freimanis and colleagues (Freimanis *et al*., 2010). Moreover, HERV-K10 solitary LTR element (DQ-LTR3) has been reported at increased levels within HLA-DQB1 locus in RA patients compared to healthy subjects. It has been suggested as a potential disease

marker for RA and could play a role in enhanced susceptibility to the disease (Seidl *et al*., 1999; Pascual *et al*., 2001).

1.4.5. Mechanisms by which HERVs may induce autoimmunity

HERVs may be associated with diseases in humans as causative agents through several mechanisms including superantigen motifs, neoantigens gene modulation and molecular mimicry.

1.4.5.1. Superantigenicity

Superantigens (SAgs) are proteins derived from bacteria and viruses which are potent mitogenic activators of B and T cells independent of the specificity of their antigen receptor. Once secreted, SAgs do not need to be processed for presentation as classical antigens and interact with antigen presenting cells and MHC II outside of the peptide binding groove. This unorthodox binding results in Vβ-restricted activation of up to 20% of total resting T cells and extensive cytokine secretion which leads to immune dysregulation (Solanki, Srivastava and Singh, 2008). Conrad and colleagues reported the first HERV SAg derived from IDDMK$_{1,2}$22 Env protein in type 1 diabetes (Conrad *et al*., 1997), which was subsequently observed to correspond to the HERV-K18 *Env* allele with superantigenic activity in IDDM (Stauffer *et al*., 2001). Moreover, HERV-K18 expression levels are increased in subjects with juvenile idiopathic arthritis suggesting possible SAg involvement in overstimulation of auto-reactive T lymphocytes (Sicat, Sutkowski and Huber, 2005). Emmer and colleagues have reported a polyclonal expansion of Vβ T cells in the brain samples from MS patients which appears to be stimulated by MSRV SAg. They also implied that B cell SAg may

contribute to MS pathogenesis via oligoclonal bands (immunoglobulin) stimulation. Moreover, SAg-activated T lymphocytes can also lead to bystander cell damage (Emmer, Staege and Kornhuber, 2014). Elevated frequencies of Vβ T cells have also been reported in rheumatoid synovium which may suggest SAg involvement in RA aetiology (Li *et al*., 2013). Moreover, SAg could stimulate B lymphocytes to generate rheumatoid factor autoantibodies and activate cytokine-producing synovial cells in RA (Schiffenbauer, 1999).

1.4.5.2. Neoantigenicity

Being an integral part of the human genome, HERV-derived proteins are expected to be perceived as auto-antigens and tolerated by lymphocytes. However, if they are not presented during thymic selection they can become immunogenic neo-antigens (Gröger and Cynis, 2018). Viral protein combinations of endogenous and/or exogenous viruses could lead to the formation of neo-antigens (Balada, Ordi-Ros and Vilardell-Tarrés, 2009). Moreover, post-translational modifications such as citrullination converts native epitopes into neo-antigens which may differ sufficiently to drive loss of tolerance reactions in RA (Lerner, Aminov and Matthias, 2016).

1.4.5.3. Trascriptional activation

HERV expression in autoimmunity can be modulated by infectious agents such as exogenous viruses. It has been observed that HERV-K18 Env protein expression can be transactivated by EBV in resting B cells, which leads to SAg formation (Hsiao *et al*., 2006). HHV-6 has also been reported to induce transcriptional activation of HERV-K18 Env (Turcanova, Bundgaard and Höllsberg, 2009) whereas human simplex virus type

1 (HSV-1) induced HERV-W Gag and Env proteins in MS (Ruprecht *et al*., 2006). Assinger and colleagues have recently reported that human cytomegalovirus (CMV) upregulates HERV-K (HML-2), HERV-W, HERV-F, HERV-L and HERV-T activity (Assinger *et al*., 2013). Moreover, HIV-1 Tat protein has been shown to transactivate HERV-K (HML-2) LTR whilst HTLV-1 Tax protein upregulated HERV-W and HERV-H LTRs. HERV-K Gag is also upregulated in HIV-1-infected CD8+ T lymphocytes (van der Kuyl, 2012).

1.4.5.4. Host immune response modulation

HERV sequences can become integrated into any fragment of the human genome and may lead to activation, inhibition and differential splicing of regulatory genes. Transposable element-mediated defects in the *Fas* gene can lead to apoptosis disruption in autoreactive lymphocytes in SLE. Moreover, HERV-K10 has been reported to have an integration site inside the C2 complement gene (Perl, 1999) whilst HERV-K(C4) contains the entire sequence within C4 complement gene and therefore could affect the complement cascade. HERV Env proteins constitute possible modulators of cytokine expression which in turn can act on lymphocyte responses. In SLE, HERV clone 4-1 is capable of inducing T cells, pro-inflammatory cytokines and anergy in PBMCs. HERV integration within MHC I and MHC II genes is also believed to contribute to self-reactivity through defective MHC expression and polymorphisms in autoimmune-prone DR haplotypes (Balada, Ordi-Ros and Vilardell-Tarrés, 2009). Moreover, LTR-mediated expression of ectopic host genes and generation of HERV gene products could cause or exacerbate disease. HERVs have also been reported to activate both the innate and adaptive immune systems (Lock *et al*., 2014; Young, Stoye and Kassiotis, 2013). HERVs have the ability to induce chronic inflammation through

the activation of PRRs of the innate immune system. HERV-W Env protein elicits pro-inflammatory cytokine release through interaction with CD14 and TLR4 which could elicit autoinflammatory and autoimmune conditions. HERVs could also stimulate the production of autoantibodies by PRRs as shown with TLR-activating self nucleic acids in SLE. Alternatively, HERV proteins could lead to immune suppression as demonstrated for HERV-FRD Env in pregnancy protection and for HERV-K Env in immunosuppressive cytokine (IL-10) production (Magiorkinis and Hurst, 2015).

1.4.5.5. Molecular mimicry

The notion of molecular mimicry relies on the observation that microbial agents may share structural motifs or sequence homology with that of a host component. As a result, an immune response to a foreign epitope could also induce cross-reactive antibodies and T cell immunity directed against host proteins. Ultimately, this mechanism may lead to a loss of clonal anergy by B and T lymphocytes, consecutive reactivity to a range of autoantigens via epitope spreading and result in autoimmune injury (Nelson *et al*., 2014b).

A classic example of post-infectious autoimmune responses evoked by molecular mimicry events is observed in rheumatic fever and glomerulonephritis following infection with the gram-positive bacterium *Streptococcus pyogenes*. Host antibodies generated to bacterial epitopes, virulence factors and exotoxins cross-react with cardiac and renal components leading to organ damage (Cusick, Libbey and Fujinami, 2012).

Exogenous viral agents have been implicated in the pathogenesis of multiple human autoimmune diseases (Table 1.05) including parvovirus B19 infection in RA. Here, viral capsid protein VP1 is homologous with human cytokeratin and persistent infection with

this virus can perpetuate the generation of antiviral antibodies that show cross-reactivity with collagen II, a major autoantigen in RA (Lunardi *et al*., 2008). Epstein-Barr virus has also been reported to share sequence similarities with RA-specific proteins such as synovial proteins or shared epitope QKRAA motif which is homologous with EBV-encoded gp110 protein (Costenbader and Karlson, 2006). Much interest has also surrounded the finding of molecular mimicry between HERVs and host proteins in autoimmune diseases.

Immune-mediated disease (antigen mimicked)	Exogenous viral agent	References
Acute disseminating encephalomyelitis (Myelin basic protein)	Measles virus, rabies vaccine, herpesvirus 6 (HHV-6), coronavirus, influenza virus hemagglutinin, Epstein-Barr virus (EBV), Semliki Forest virus	(Tenembaum *et al*., 2007)
Myocarditis (Cardiac myosin)	Coxsackie virus	(Cunningham, 2004)
Coeliac disease (Transglutaminase)	Adenovirus 12, hepatitis C virus (HCV)	(Dieterich *et al*., 1997; Plot and Amital, 2009)
Cogan's syndrome (SSA/Ro; connexin 26)	Reovirus III major core protein lambda 1	(Lunardi *et al*., 2002)
Essential mixed cryoglobulinemia (IgG-Fc)	HCV	(De Re *et al*., 2006)
Graft vs host disease (HLA-DR, aminopeptidase N)	Human cytomegalovirus (hCMV)	(Nauclér, Larsson and Möller, 1996)
Insulin dependent diabetes (type I) (Islet antigens- GAD 65, proinsulin carboxypeptidase H, tyrosine phosphatase IA-2)	Coxsackie B4 virus, rubella, rotavirus, herpes, rhinovirus, hantavirus, flavivirus and retrovirus	(van der Werf *et al*., 2007; Baum *et al*., 1995; Jones and Armstrong, 1995; Endl *et al*., 1997; Steed and Stappenbeck, 2014)
Multiple sclerosis (Myelin basic protein)	Corona, measles, mumps, EBV, herpes virus	(Oldstone, 1998; Libbey, McCoy and Fujinami, 2007; Wucherpfennig and Strominger, 1995; Fujinami *et al*., 2006)
Myasthenia gravis (Acetylcholine receptor, neurofilaments)	Herpes simplex virus type 1 gpD	(Oldstone, 1998; Schwimmbeck *et al*., 1989)
Primary biliary cirrhosis (HLA-DR)	hCMV	(Shimoda *et al*., 2003; Selmi *et al*., 2011; Kita *et al*., 2002)
Rheumatoid arthritis (Collagen type II)	Parvovirus B19	(Sfriso *et al*., 2010)
Scleroderma (Tetraspan novel antigen-2)	hCMV UL94 protein	(Lunardi *et al*., 2000)

Stiff-person syndrome (GAD65)	hCMV (pUL57)	(Ali *et al*., 2011; Hiemstra *et al*., 2001)
Systemic lupus erythematosus (Ro60 kD, SmB')	EBV (nuclear antigen-1)	(McClain *et al*., 2005)
Sjögren's syndrome (Ro60 kD)	Coxsackie virus	(Stathopoulou *et al*., 2005)

Table 1.05. Association of exogenous viral agents with human autoimmune diseases as reported previously within the literature.

Multiple sequence homologies between HERV-W (MS associated retrovirus element, MSRV) Env regions and myelin proteins have been demonstrated in multiple sclerosis. These viral fragments are found in MS active brain lesions and are believed to contribute to the disease pathogenesis. Increased anti-HERV-W Env reactivity has also been observed in MS serum samples compared to healthy controls (Ramasamy, Joseph and Whittall, 2017; do Olival *et al*., 2013). MSRV-Env pathogenic effects have been successfully antagonised by the GNbAC1 monoclonal antibody which has recently undergone a Phase 2b clinical trial and exhibits anti-inflammatory properties and permits remyelination in MS patients (Derfuss *et al*., 2015). This biological agent is currently also in a Phase 2 clinical trial for type 1 diabetes (IDDM) as increased expression of HERV-W Env has been detected in serum, PBMCs and pancreata from IDDM patients and has been suggested as a causative factor in disease pathogenesis (Levet *et al*., 2017).

Moreover, our group has reported multiple homologous regions between HERV proteins and autoantigens in systemic lupus erythematosus. One such protein is HRES-1 p30 gag protein which shares amino acid sequence similarity with several SLE autoantigens including SmD1, Ribosome P, SmN, Ro/SSA-60, La/SSB, anti-U1 RNP and histone H2B. We also demonstrated several regions of homology between HERV-K10 and IgG1 Fc in RA (Trela, Nelson and Rylance, 2016; Tugnet *et al*., 2013).

1.5. Aims and Objectives

A major aim for this research project was to analyse in more detail the role of molecular mimicry in Rheumatoid Arthritis via two major mechanisms: HERV-K10 involvement and post-translational autoprotein citrullination.

We wished to ascertain the processes by which HERVs may play a role in, or trigger RA immunopathology. We also wished to establish whether the citrullination of RA autoantigens and HERV-K10 viral targets could perpetuate the disease. Several approaches were used to establish the foregoing aims. An overview of the approach taken is set out below.

- *In-silico* investigation of the homologous amino acid regions between HERV-K10, IgG1 Fc and other autoantigens in rheumatoid arthritis (Chapter 3)
- Although HERVs are of great interest in the pathogenesis of diseases, there is a current lack of reagents to fully investigate the link. Therefore, an anti-HERV-K10 polyclonal antibody was developed and characterised as a potential diagnostic reagent in RA and in detection of HERV-K10 expression (Chapter 4)
- Serological analysis of RA patient serum samples for the detection of HERV-K10 viral titres and their association with the disease (Chapter 5)
- Investigation of the effects of endogenous protein citrullination on RA autoantibody reactivity (Chapter 6)

Hypothesis:

We propose that immunodominant peptides of HERV-K10 matrix region could facilitate a secondary, antigen-driven immune response in rheumatoid arthritis via mechanisms of molecular mimicry with the key RA autoantigen, IgG1 Fc.
Further, we hypothesise that this phenomenon could be enhanced by global hypercitrullination within RA synovial joints.

2. MATERIALS AND METHODS

2.1. Ethics approval

Ethics approval for working with NHS patient samples during the project was provided by the Black Country Research Ethics (REC reference number 08/985). Ethical approval for the project was also sought from the University Life Sciences Ethics Committee (LSEC/201516/SP/148). The ethical approvals were in line with the revised Declaration of Helsinki, 2013 (https://www.wma.net/policy/current-policies/, 2013).

2.2. Collection and processing of patient samples

The Black Country Research Ethics Committee granted ethical approval for the collection of patient samples at New Cross Hospital, Wolverhampton and at Cannock Chase Hospital. All patients with RA fulfilled the diagnostic criteria of the American College of Rheumatology (ACR). Blood samples were also obtained from patients with systemic lupus erythematosus, psoriatic arthritis, osteoarthritis, and from healthy individuals. All blood samples were collected in sterile BD Vacutainer Serum Separator Tubes and the serum fraction was transferred into cryovials after centrifugation for storage at -80°C. Samples were collected from patients after obtaining written consent. In order to maximize accuracy of the investigation all blood samples were processed within 1 hour. All steps were performed under aseptic conditions.

2.3. *In silico* bioinformatic analysis

In-silico (bioinformatic) tools were used to acquire accession numbers and sequences of HERV-K10 Gag and the most prevalent autoantigens in rheumatoid arthritis identified from the literature including aggrecan, alpha 1B-glycoprotein, alpha 1 antitrypsin, alpha 2HS-glycoprotein, alpha-enolase, BRAF, calpastatin, collagen type II, fibrinogen beta and gamma, filaggrin, gelsolin, glial fibrillary acidic protein, human cartilage glycoprotein 39, IgG1 Fc, keratin type II, lumican, peptidyl arginine deiminase 4, tubulin beta-chain and vimentin. Their protein sequences were extracted from the National Center for Biotechnology Information/Genbank online database (www.ncbi.nlm.nih). Bioinformatic analysis of the protein residues was performed using established algorithms (Westwood *et al*., 2008), based on physicochemical parameters, to ascertain the antigenic profiles and immunodominant regions of each protein. Algorithms included hydrophilicity, residue polarity, surface accessibility, flexibility and β-turn propensity.

2.3.1. B Lymphocyte epitope mapping

B cell epitope predictions in this study were based on four assumptions made in previously published investigations.

I. B cell epitopes can comprise continuous (linear/sequential) or discontinuous (conformational) amino acid fragments (Potocnakova, Bhide and Pulzova, 2016).
II. B cell epitopes are surface exposed and solvent accessible polar amino acids depleted of hydrophobic residues (Kringelum *et al*., 2013).
III. B cells epitopes are typically situated within regions of high flexibility and exhibit site mobility (Mukonyora, 2015).

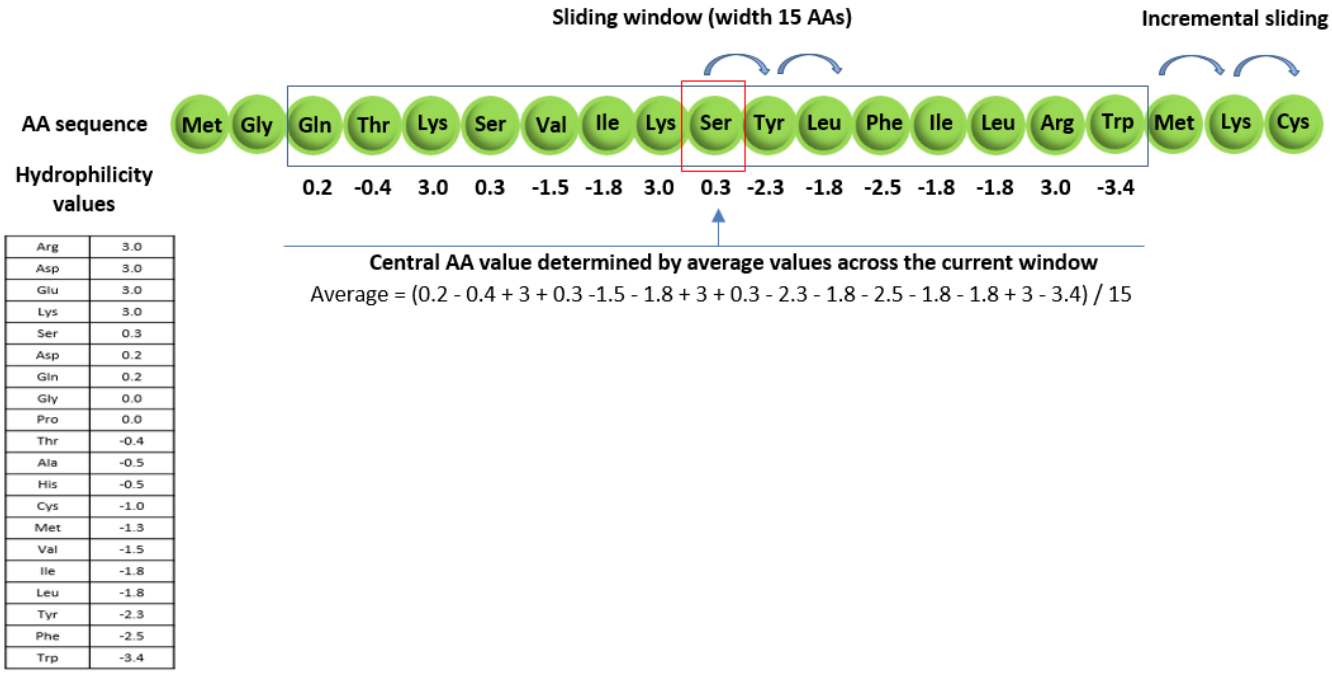
IV. Complex proteins contain numerous overlapping B cell epitopes with some showing immunodominant properties (Annadurai, 2008).

Linear B cell epitope predictions were performed using two online software programmes -ExPASy ProtScale (http://www.expasy.ch/cgi-bin/protscale.pl), and BcePred software (http://crdd.osdd.net/raghava/bcepred/), with the latter used as an additional programme to compare and contrast the performance of ExPASy and thus enhance prediction accuracy. BcePred programme offers predictions based on automated combinations of all five physicochemical algorithms as opposed to propensity measures used by ExPASy that depend on individual properties and require manual formatting (El-Manzalawy and Honavar, 2010).

These primary structure properties have been established as the most common criteria for identification of linear regions of antigenicity (Sanchez-Trincado, Gomez-Perosanz and Reche, 2017; Mukonyora, 2015). Thus, hydrophilicity criterion for B cell epitope mapping was employed on the assumption that hydrophilic residues, which confer antigenicity, are typically exposed on the protein surface (Hopp and Woods, 1981; Hopp, 1993). The accuracy of predictions is further enhanced when combined with other amino acid propensity scales such as accessibility (Lins, Thomas and Brasseur, 2003) and flexibility (Parker, Guo and Hodges, 1986; Karplus and Schulz, 1985), which confer suitable residue positioning and mobility, respectively (Saha and Raghava, 2004; Vihinen, Torkkila and Riikonen, 1994). Two additional criteria that improve B cell epitope predictions constitute residue polarity (with more polar residues situated on the surface of proteins) (Lee and Richards, 1971; Shrake and Rupley, 1973) and secondary structure criterion termed β-turns (Pellequer, Westhof and Van Regenmortel, 1993), which correlates with antigenicity, acts as ligand binding sites, and is responsible for protein globular structure (Yu-Bin Dong *et al*., 2003).

2.3.2. The Sliding window approach

To improve the robustness of algorithm-based linear B cell epitope predictions, a ‘sliding window’ principle was incorporated within a query protein sequence by ExPASy ProtScale and BcePred servers. Upon establishing a window size, which in this case was 15 amino acids wide, the window moved along the length of the sequence incrementally one amino acid (AA) at a time. The residue in the middle of each window had its value estimated by assigning it the mean value of the hydrophilicity parameter of all amino acids within the window. Then the window was shifted by a single position to the next residue (maintaining a consistent level of overlap) and the process continued until the end of the query sequence was reached (Figure 2.01). The propensity scores, which depend on correlations between amino acid properties, were utilised to predict the antigenic residues and their location within the query sequence (Claverie and Notredame, 2007; El-Manzalawy and Honavar, 2010).



Arg	3.0
Asp	3.0
Glu	3.0
Lys	3.0
Ser	0.3
Asp	0.2
Gln	0.2
Gly	0.0
Pro	0.0
Thr	-0.4
Ala	-0.5
His	-0.5
Cys	-1.0
Met	-1.3
Val	-1.5
Ile	-1.8
Leu	-1.8
Tyr	-2.3
Phe	-2.5
Trp	-3.4

Figure 2.01. Sliding window principle in action. Hydrophilicity values established by Hopp and Woods (1981) (box on the left) are allocated to individual amino acids within a default window of 15 residues. An average of all 15 hydrophilicity values is established and assigned to the amino acid in the centre of the window (Serine; red box). The window then shifts by a single position along the full length of the protein sequence and the process is repeated for every residue (Sivalingam and Shepherd, 2012). AA – amino acid.

2.3.3. T Lymphocyte epitope mapping

Additional epitope predictions were performed to identify potential T cell epitopes which may coincide with B cell epitopes. Peptides which are capable of binding to HLA-DRB1*01:01, *01:04, *04:01, *04:04, *04:05, *09:01 (RA alleles) were analysed using the Immune Epitope Data Base (www.iedb.org) server. The default peptide length of 15 amino acid residues, which was applied in the analysis, encompasses the core nonamer peptides that are expected to bind to the HLA DR molecule and form the major portion of the T cell epitope. The prediction method selected was the Stabilised Matrix Method (SMM) where the peptides are graded according to their predicted binding affinities IC_{50} value (indicating the concentration of peptide (nM) needed to achieve 50% saturation of the HLA molecule). Thus, a lower number implies higher affinity and peptides with IC_{50} <50 nM are considered to bind with high affinity, <500 nM with intermediate affinity and <5000 nM with low affinity.

2.3.4. Pairwise epitope alignment

The level of homology (similarity) between the extracted protein sequences of HERV-K10 Gag and RA autoantigens was evaluated using global pairwise Basic Local Alignment Search Tool (BLAST) analysis carried out online (https://www.ncbi.nlm.nih.gov/BLAST/). They were subsequently scanned for regions of homology using ExPASy SIM (http://web.expasy.org/sim/) and LALIGN (http://www.ebi.ac.uk/Tools/psa/lalign/) online local alignment tools with their default parameters. Global and local alignments depend on algorithms which assess an optimal alignment of two sequences with the former creating an end-to-end sequence comparison and the latter matching the most similar regions within the two query sequences (Brown, 2000). BLOck SUbstitution Matrix 62 (BLOSUM62) was applied

within the local similarity programmes to score pairs of aligned residues. Produced alignments were visually inspected for regions of homology.

2.3.5. Replacement score determination

Replacement scores were obtained from a substitution matrix developed by Tűdős *et al.* (1990) for all amino acid pairs to allow the identification of homology between specific amino acid pairs. Therefore, the higher the correlation between residues (higher positive value), the greater the similarity between them. Individual replacement scores were extracted for matching amino acid pairs and applied manually to predicted alignments between two query sequences.

2.4. Molecular modelling

Visualisation of epitopes and potential structural homologies between protein fragments was undertaken using PyMOL (Molecular Graphics System, Version 1.2r3pre, Schrodinger, LLC: https://www.pymol.org/) and UCSF Chimera (http://www.cgl.ucsf.edu/chimera/) molecular visualisation software. Molecular models of the homologous regions between amino acid pairs were generated and inspected visually for structural mimicry.

The crystal structure of IgG1 Fc (entry: 4DZ8) was extracted from the Protein Data Bank (www.rcsb.org) and used as a template for a structural model of HERV-K10 Gag generated by the SWISS-MODEL server (https://swissmodel.expasy.org/).

In-silico simulation of post-translational modification of arginine residues to citrulline was achieved using a PyMOL PyTMs citrullination plugin (Warnecke *et al*., 2014). Three-dimensional models of the human fibrinogen (entry: 3GHG) and IgG1 Fc (entry:

5JII) were obtained from the Protein Data Bank. Their ribbon models were processed with PyTMs plugin and enhanced using UCSF Chimera.

2.5. Peptide synthesis and bleed evaluation

Following bioinformatic analysis, the HERV-K10 Gag immunodominant fragment denoted by our group as MAG1: (Matrix Gag peptide 1) was synthesised commercially (Severn Biotech, Kidderminster, UK) as N-terminal biotinylated and non-biotinylated peptides of different lengths: 15-mer (RIGKELKQAGRKGNI;), 12-mer (RIGKELKQAGRK), 9-mer (RIGKELKQA), 7-mer (IGKELKQ), 5-mer (GKELK), 4-mer (GKEL), and 3-mer (GKE). Polyclonal antibodies were generated commercially (Severn Biotech) to the 15-mer MAG1, which was used to immunize New Zealand White rabbits. The IgG fraction was subsequently purified from immunised rabbit sera (final bleed), acquired at 42 days following the first antigen injection, using NAb Protein G affinity columns (Thermo Fisher Scientific, Rugby, UK). Total protein concentration of the purified rabbit IgG was assessed using the BCA Protein Assay Kit (Thermo Fisher Scientific).

2.6. Serological analysis and antibody optimisation

2.6.1. Antibodies for immunotesting

Primary antibodies				
Specificity	**Application**	**Valency**	**Supplier**	**Dilution/Concentration**
MAG1	ELISA/ICC/IF/WB	PAb	Severn Biotech	2 µg/ml
MAG3	ELISA	PAb	Severn Biotech	8 µg/ml
MAG4	ELISA	PAb	Severn Biotech	2 µg/ml
Citrulline	ELISA	PAb	Abcam	8 µg/ml
Citrulline	ELISA	PAb	Millipore	8 µg/ml
Citrulline	ELISA	PAb	Abcam	8 µg/ml
Fibrinogen	ELISA, WB	PAb	Abcam	1 µg/ml
Pan-IgG (A57H)	ELISA, WB	MAb	Birmingham University	1 µg/ml
HLA-ABC	ICC/IF	MAb	DAKO	2 µg/ml
Secondary antibodies				
Specificity	**Conjugate**		**Supplier**	**Dilution/Concentration**
Human IgG F(ab')2	HRP		Bio-rad	1/100
IgG (isotype control)	FITC		GeneTex	1/500
Goat IgG	HRP		Abcam	1/5000
Mouse IgG	HRP		Sigma-Aldrich	1/4000
Rabbit IgG	HRP		DAKO	1/5000
Rabbit IgM	HRP		DAKO	1/5000
Rabbit IgA	HRP		DAKO	1/5000
MAG1	FITC		In-house	25 µg/ml

Table 2.01. Antibodies for immunotesting in ELISA, immunocytochemistry (ICC), immunofluorescence (IF), flow cytometry (FACS) and western blotting (WB). PAb-polyclonal antibody, MAb-monoclonal antibody.

2.6.2. Indirect Enzyme-Linked Immunosorbent Assay (ELISA)

Wells of a 96-well plate (2 HB Immulon, Dynex Technologies, Worthing, UK for non-biotinylated peptides or NeutrAvidin Protein plates, Thermo Fisher Scientific, for biotinylated peptides) were coated with the indicated peptide at the indicated concentration and incubated overnight at 4°C. Plates were then washed three times with phosphate-buffered saline (PBS, pH 7.2) using an automatic Wellwash Microplate washer (Thermo Fisher Scientific), and blocked with 2% bovine serum albumin (BSA) in PBS (for purified primary antibody preparations) or with 1% Hammarsten grade casein in PBS (for crude serum preparations) at room temperature for 1h. After three washes with PBS/Tween20 (PBS-T) (0.1%), 50 µl of a purified primary antibody at the indicated concentration (Table 2.01) diluted in PBS-T (0.05%)/ BSA (2%) or human patient sera diluted in 1% Hammarsten casein/PBS at 1/200 was added to the wells and incubated at room temperature for 1 hour. After washing three times with PBS-T (0.1%), horseradish peroxidase (HRP)-conjugated secondary antibody was diluted in PBS-T (0.1%)/ BSA (2%) to the indicated dilution (Table 2.01) and incubated at room temperature for 1 hour. Plates were washed with PBS-T (0.1%) before 50 µl of 3, 3', 5, 5'-tetramethyl benzidine (TMB; Sigma-Aldrich, Cambridge, UK) liquid substrate system was added to each well. The colourimetric reaction was allowed to develop for 4min in the dark at room temperature and was subsequently stopped by addition of 50µl of 2N hydrochloric acid. The optical density (OD) was measured at 450 nm using a GloMax Multi Microplate Reader (Promega, Madison, WI, USA).

2.6.3. Inhibition ELISA

An inhibition assay was performed as described previously (Freimanis *et al*., 2010) on two 96-well plates i.e. target and inhibition plate. Standard curves were obtained for each antibody of interest using the indirect ELISA protocol (Section 2.6.1.). Primary antibody dilutions that gave 90% of maximal binding with a corresponding antigen were selected. The antibody was subsequently diluted in PBS-T (0.05%)/ BSA (2%) and pre-incubated with a 2-fold serial dilution of the appropriate or control peptide on the inhibitor plate at room temperature for 1 hour. Incubated preparations (100 µl) were then transferred to the target plate pre-coated with the corresponding peptide and the ELISA assay was performed as outlined in the indirect ELISA protocol above.

2.6.4. Quantitative rheumatoid factor assay

The presence of rheumatoid factor (RF) autoantibody titres in patient serum samples was assessed using a commercially available sandwich ELISA kit (Abnova, Taiwan). The assays were performed according to the manufacturer's instructions, with positive and negative controls supplied by the kit included on each plate. Patient RF titres were established by interpolation from a calibration curve produced for calibrator RF sera provided in the kit.

2.6.5. Quantitative anti-citrullinated protein antibody assay

A quantitative commercial ELISA test (Axis-Shield, Dundee, UK) was used to detect anti-CCP (ACPA) autoantibodies in 46 RA serum samples. The assays were performed according to the manufacturer's instructions, with positive and negative controls supplied by the kit included on each plate. Patient ACPA titres were

established by interpolation from a calibration curve produced for calibrator ACPA sera provided in the kit.

2.6.6. Immunocytochemistry

All stages of this protocol were performed in a humidity chamber. Cell sections were rehydrated in PBS and fixed with 20 µl of acetone at 4°C for 5 min. Cell sections were blocked with PBS/ BSA (3%) for 5 min at room temperature and incubated for 40 min with a primary antibody diluted to the indicated concentration (Table 2.01) using PBS/ BSA (3%) at room temperature. Slides were washed twice with PBS for 5 mins under agitation, before a secondary anti-species HRP-conjugated antibody (Agilent) was added at the indicated dilution (Table 2.01) for 30 min at room temperature. Following two washes with PBS, cell sections were developed for 8 min in 0.5 mg/ml diaminobenzidine hydrochloride (DAB; Sigma-Aldrich). Cells were then counter-stained with Mayer's hematoxylin (Sigma-Aldrich) before a single drop of Fluoromount (Sigma-Aldrich) was added to the slide. Sterile glass coverslips were placed over the stained cell monolayers and sealed with nail varnish before drying for 5 min. Images were acquired using an inverted phase contrast microscope (Olympus UK, UK) at x20 magnification and processed using Image J (NIH, Maryland, USA).

2.6.7. Immunofluorescence and confocal laser microscopy

All stages of this protocol were performed in a humidity chamber. Cell sections were rehydrated in PBS and fixed with 20 µl of acetone/methanol (50/50) at 4°C for 5 min. Sections were then blocked with PBS/ BSA (3%) for 1 hour. Blocking buffer was removed and cells were incubated for 1 hour with a primary antibody diluted to the

indicated concentration (Table 2.01) using PBS/ BSA (3%). Cells were washed twice with PBS for 5 mins under agitation and treated with a secondary fluorescein isothiocyanate (FITC)-conjugated antibody (Agilent) at the indicated dilution (Table 2.01) for 30 min in the dark. Following three washes with PBS (5 min each), cell sections were incubated for 5 min with 4', 6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich) at 1/1000 dilution in PBS. Cells were then washed twice in PBS before a single drop of Fluoromount (Sigma-Aldrich) was added to the slide. Sterile glass coverslips were placed over the stained cell monolayers and sealed with nail varnish before drying for 5 min in the dark. Images were acquired using a Zeiss LSM 510 inverted laser scanning confocal microscope with a 40x objective. Images were processed using Image J (NIH, Maryland, USA).

2.6.8. Western blotting

Protein samples were resuspended in Laemmli sample buffer (Sigma-Aldrich) and heated at 96°C for 30min to denature proteins. Samples were stacked on a 4% (v/v) stacking gel and separated out on a 6% (v/v) sodium dodecyl-sulfate polacrylamide gel (SDS-PAGE) at 200 V, 400 mA for 35 min using a BIO-RAD Mini-PROTEAN tetra cell system (BIO-RAD Ltd., Hertfordshire, UK). Proteins were subsequently transferred onto a nitrocellulose membrane by electroblotting at 100 V, 400 mA for 1 hour. The membranes were stained with Ponceau S staining solution (Sigma-Aldrich) to ascertain successful protein transfer. After destaining in PBS, the membrane was blocked with PBS-T (0.05%)/ non-fat skimmed milk (Marvel) (5%) for 1 hour at room temperature. Primary antibody at the indicated concentration (Table 2.01) or human patient sera at 1/200 diluted in PBS-T (0.05%)/ non-fat skimmed milk (Marvel) were incubated with the membrane overnight at 4°C, followed by three 15 min washes in

PBS-T. Any bound antibodies were detected with anti-species HRP-labelled conjugate diluted in PBS-T (0.05%)/ non-fat skimmed milk (Marvel) at the indicated dilution (Table 2.01) upon 1 hour incubation at room temperature. Following washing, membrane was exposed to enhanced chemiluminescence detection system (ECL+, Amersham, Buckinghamshire, UK) for 5 min. Bands were visualized using a STORM 860 Phosphorimager (GMI, MI, USA) at 450 nm.

2.6.9. Flow cytometry

All cells were cultured in triplicate in 6 well flat bottom plates at $1x10^6$ cells per well at 37°C, 5% CO_2. Single cell suspensions were obtained by trypsinisation of adherent cells. For intracellular staining, cells were fixed and permeabilised with 2 ml of ice-cold methanol for 15 min at room temperature. After washing with PBS, cells were stained with FITC-conjugated primary antibody at the indicated concentration (Table 2.01) diluted in PBS/ BSA (1%) / Triton (0.06%) and incubated for 30 min at room temperature. Cells were then washed with PBS and acquired on an Accuri™ C6 Plus (BD Biosciences, San Jose, CA, USA) flow cytometer. Analysis was performed using FlowJo (TreeStar, Ashland, OR, USA).

2.6.10. Antibody labelling with fluorescein isothiocyanate (FITC)

In-house FITC conjugation of anti-MAG1 antibody was achieved using a commercial conjugation kit (Abcam, Cambridge, UK) according to the manufacturer's instructions.

2.7. *In vitro* cell culture

All cell culture work was performed in a laminar flow hood under appropriate aseptic conditions. Cell lines were maintained in DMEM or RPMI-1640 cell culture medium (Sigma-Aldrich) supplemented with 10% (v/v) fetal bovine serum (FBS) (Sigma-Aldrich) and 1% penicillin 10000 units/ml - streptomycin 10000 µg/ml solution (Sigma-Aldrich). The culture medium is henceforth referred to as “complete medium”. Four adherent cancer cell lines were grown for work within the project: GH teratocarcinoma (DSMZ no. ACC 811), KM-H2 Hodgkin lymphoma (DSMZ no. ACC 8), MCF-7 (ATCC HTB-22) and T-47D (ATCC HTB-133) breast cancer cell lines. Cells were grown in freshly made up medium supplemented according to their requirements (DMEM for MCF-7 and GH cell lines and RPMI-1640 for T-47D and KM-H2 cell lines).

2.7.1. Propagation of adherent cells

Cells were cultured in complete medium and incubated at 37°C with 5% carbon dioxide (CO_2) in a humidified air incubator prior to use. Cultures were inspected using an inverted microscope (Olympus, UK) to assess cell confluency and the presence of bacterial or fungal contamination. Spent medium was discarded using disposable sterile Pasteur glass pipettes and the monolayer was washed twice with a volume equivalent to half the culture medium of sterile phosphate buffered saline (PBS) to remove any FBS particles. After washing, 0.25% (v/v) trypsin-ethylene diamine tetra acetic acid (EDTA) solution (Sigma-Aldrich) was added to the cell monolayer using 1 ml per 25cm^2 of surface area to dislodge cells. The flask was tilted to allow the whole monolayer to be exposed to trypsin before incubating at 37°C for 5 min or until cell detachment was evident. Trypsin was deactivated by adding 9ml of complete medium. Cells were transferred to a sterile universal and centrifuged at 250 *gx* for 5 min.

Supernatant was removed and the cell pellet was re-suspended in 10 ml of pre-warmed complete medium. Cells were then counted using a bright-line haemocytometer (Sigma-Aldrich). The average of four different counts was used to establish cell numbers in a 10 ml suspension. Following cell counting, cell suspension was either split in a ratio of 1 to 5 into a new 75 cm^2 cell culture flask or seeded into multi-well cell culture plates to be used for subsequent procedures. Cells were incubated at 37°C and 5% CO_2 in humidified air until cultures reached 70-90% confluency at approximately 4-5 days at which point further cell passage or splitting was performed.

2.7.2. Cytospin preparation for staining from single cell suspension

Single cell suspensions of adherent cells were generated through trypsinisation as described in 2.7.1., and cells were subsequently counted using a Coulter counter (Z1; Beckman Coulter Inc., High Wycombe, UK) to obtain approximately $1x10^5$ cells per cytospin. Cells were centrifuged in culture medium at 2000rpm for 4 min and washed three times in PBS. The pellet was resuspended in 200 µl of PBS and cytospun in Shandon™ EZ Single Cytofunnel™ (Thermo Fisher Scientific) using a Cytospin™ 4 cytocentrifuge (Thermo Fisher Scientific) for 5 min at 2000rpm onto poly-L-lysine coated glass slides (Sigma-Aldrich). Mounted slides were air dried and used immediately or stored at -80°C.

2.7.3. Cryopreservation of cells

Cells were counted and diluted to 2-5 x 10^6 cells per ml of 10% (v/v) sterile dimethyl sulfoxide (DMSO) (Sigma-Aldrich) in pre-chilled neat FBS. The suspensions were stored in cryovials overnight in a -80°C freezer for slow cooling after which they were

transferred into a cryogenic freezer containing liquid nitrogen at a temperature of -180°C.

2.7.4. Resuscitation of frozen cell lines

Cryovials containing cell aliquots were retrieved from liquid nitrogen, moved into a water bath and heated to 37°C for about 1 min until about 80% defrosted. The contents of the vial were transferred into 10 ml of pre-warmed complete medium and centrifuged at 250 *gx* for 5 min. The supernatant was discarded and cells were re-suspended in 10 ml of fresh growth medium. Cells were transferred into a 25 cm^2 cell culture flask and an appropriate volume of growth medium was added to allow cell growth. Medium in each flask was replaced every 48-72 hours and cells were passaged until confluency.

2.8. *In vitro* citrullination of protein targets

2.8.1. Protein and peptide citrullination

Arginine residues from fibrinogen, IgG1 Fc, MAG1 and MAG3 peptides were converted into citrulline residues in vitro using an active recombinant peptidylarginine deaminase (PAD) enzyme cocktail (SignalChem, Richmond, British Columbia, Canada). Proteins/peptides were incubated with PAD enzyme at a ratio of 10:1 (protein/peptide:PAD enzyme) in 0.1M Tris-HCl (pH 7.6)/ 10mM $CaCl_2$/ 5mM dithiothreitol (DTT, Sigma-Aldrich) buffer for 3 hours at 55°C. Citrullinated samples were stored at -80°C.

2.8.2. Citrulline-specific reporter assay

A rhodamine-phenylglyoxal (Rh-PG, Cayman Chemical, Ann Arbor, Michigan, USA) citrulline specific chemical probe, was used to assess in-house protein citrullination. Citrulline labelling was performed at 37°C for 30 mins with 100μM Rh-PG in 20% (w/v) trichloroactic acid (TCA) at 1:1 ratio with citrullinated protein. Commercially citrullinated fibrinogen (Cayman Chemical) was used as a positive control for citrullination. For detection in western blotting, Rh-PG labelled protein samples were resuspended in loading buffer (2 μg protein/lane), separated via SDS-PAGE and visualised as described in 2.6.8.

2.8.3. Antibody digestion with pepsin

Serum antibodies from ACPA+RF- RA cohort were reduced via pepsin digestion to obtain intact $F(ab')_2$ fragments. 0.4% pepsin solution (Sigma-Aldrich) was prepared in cold 10mM HCl to activate the enzyme. Sera from 9 RA patients with ACPA+RF- profile were pooled and incubated with pepsin solution at a 20 : 1 (serum : enzyme) ratio for 1h at 37°C. Pepsin was then inactivated by raising the pH to 6 to stop further digestion and the reaction mixture was affinity purified using Protein G spin column (Thermo Fisher Scientific) which immobilised the Fc fragments allowing $F(ab')_2$ separation. Intact $F(ab')_2$ fragments were resuspended in PBS to the initial sample volume and the pH was adjusted to neutral (pH 7.2). Low molecular weight fragments were removed by diafiltration with a 30-kDa nominal molecular weight cut-off membrane (Merck Millipore, Watford, UK).

2.9. Statistics

Statistical analysis of differences between 2 groups was performed using unpaired T-tests in GraphPad Prism (GraphPad Prism version 7.03, San Diego, CA, USA). All analyses were two-tailed with a 95% confidence interval. *, p = <0.005, **, p = <0.01, ***, p = <0.001, ****, p=<0.0001, ns = not significant. For a comparison of more than two group means, one-way analysis of variance (ANOVA) was performed.

3. MOLECULAR MIMICRY BETWEEN RA AUTOANTIGENS AND HERV-K10 PEPTIDES

3.1. Introduction

Autoimmune diseases may arise through a breakdown in tolerance, together with a genetic predisposition, hormonal factors, and environmental agents. Potentially autoreactive B and T cells are deleted during ontogeny through apoptosis although a small proportion escape this mechanism but are subject to peripheral tolerance (Tobón, Izquierdo and Cañas, 2013). A breakdown in tolerance at this stage may be facilitated by a number of mechanisms including molecular mimicry, epitope spreading, superantigen motifs and bystander activation of cells through microbial agents (Vojdani, 2014), leading to autoimmunity. A number of autoimmune diseases, including rheumatoid arthritis (RA) are primarily antibody (B cell) mediated aided by T cell help that facilitates isotype switching and affinity maturation. Several clinically relevant autoantibodies are present in patients including 'rheumatoid factor' and antibodies to citrullinated products (ACPAs) such as vimentin, fibrinogen and collagen type II (Ajeganova *et al*., 2016; Glaant *et al*., 2016). In particular IgM and IgG rheumatoid factor autoantibodies bind to immunoglobulin IgG1 Fc and are present in 75-80% of RA patients. The presence of RF is associated with a poorer prognosis in terms of disease outcome (Song and Kang, 2010b).

Human endogenous retroviruses (HERVs), which account for up to 8% of the human genome, have been recently implicated as potential aetiological agents within RA. Whilst many HERVs have become defective over time through accumulated genetic and frame-shift mutations a few, including the human endogenous virus HERV-K10, have retained intact open reading frames within *gag*, *pol* and *env* genes with the

capacity to produce viral products (Löwer *et al*., 1993) and particles in RA synovial tissue (Stransky *et al*., 1993; Takeuchi *et al*., 1995). Accumulated evidence from our research group and others has suggested that HERV-K10 could play a pivotal role as a viral trigger of RA or augment the immunopathogenesis of disease potentially through the mechanism of molecular mimicry (Trela, Nelson and Rylance, 2016; Balada, Ordi-Ros and Vilardell-Tarrés, 2009).

The premise of molecular mimicry is that microbial agents may share homology in amino acid sequence, or a structural motif, with human proteins (Tugnet *et al*., 2013). As a consequence, an immune response to a given pathogen could also induce undesired B and T cell reactivity to a host protein. In effect, a pathogenic organism which constitutes a molecular mimic of the host protein could lead to the generation of autoantibodies in RA and thus autoimmune reactions. Therefore, we have sought to identify through *in silico* analysis key epitopes of HERV-K10 Gag, the matrix region which is a pre-requisite for viral particle formation (Löwer *et al*., 1993), that may share identical or similar epitopes on a key RA autoantigen, IgG1 Fc. Moreover, we also investigated the possibility of molecular mimicry within other RA autoantigens, including fibrinogen, vimentin and collagen type II.

3.2. Results

3.2.1. Epitope mapping and comparison of BcePred versus ExPASy Protscale B lymphocyte epitope prediction software

In silico B cell epitope mapping and non-intersecting pairwise alignments were performed on HERV-K10 Gag and RA autoantigens using established algorithms within the software (Westwood *et al*., 2008). Figure 3.01 summarises the process of selecting potential antigenic determinants from a database of both nucleotide and protein sequences registered in the NCBI database.

The retrospective prediction of established epitopes of autoantigens associated with RA pathogenesis was incorporated within this investigation as a measure of independent validation of the method utilised here. A set of 19 autoantigens implicated in RA was selected and review of the literature was performed to identify all experimentally determined linear epitopes published previously for these antigens. They were subsequently employed to verify the performance of both ExPASy and BcePred local alignment platforms. Table 3.01 demonstrates details for each RA autoantigen extracted from the literature that contributed to this dataset.

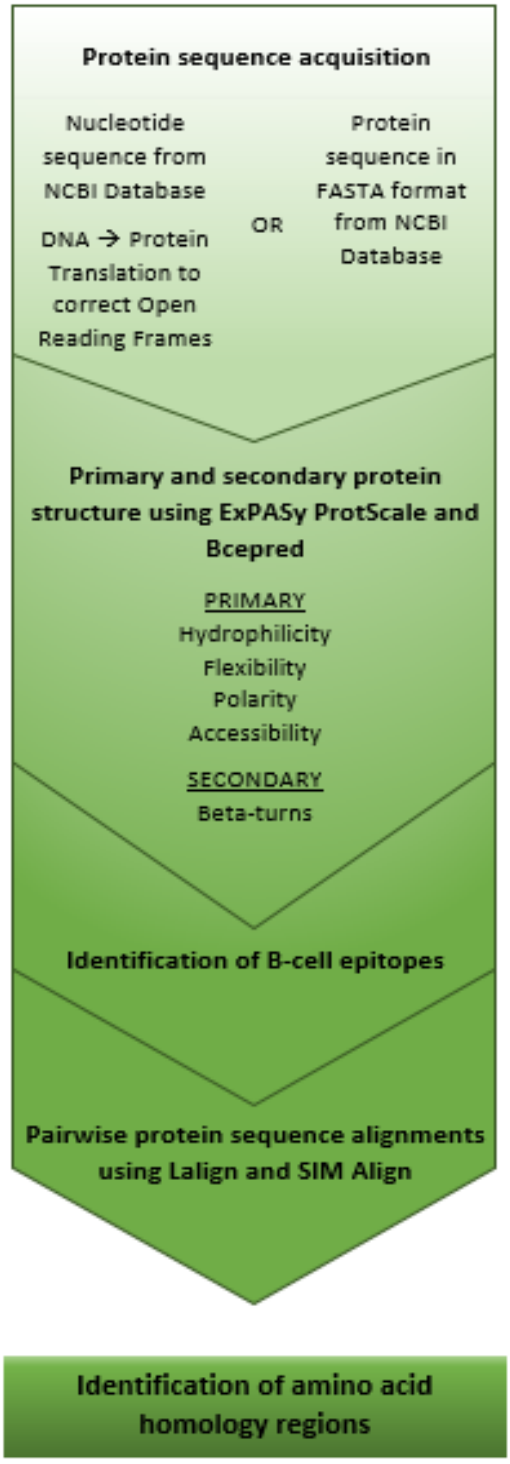


Figure 3.01. Workflow presenting the main components of the *in silico* epitope mapping. This procedure was applied to viral and host proteins for pairwise alignments of those sequences to establish the degree of homology between them.

Autoantigen		Accession number	Details	Reference
1	Aggrecan	NM_001135	Integral part of the extracellular matrix in cartilaginous tissue	(de Jong *et al*., 2010; Falconer *et al*., 2016)
2	Alpha 1B-glycoprotein (A1BG)	P04217	Possible role in antigen presentation	(Udby, Johnsen and Borregaard, 2010; Biswas *et al*., 2013)
3	Alpha 1 antitrypsin (A1AT)	P01009	Protection from inflammatory cells enzymes	(Janciauskiene *et al*., 2011; Blaurock *et al*., 2016)
4	Alpha 2 HS glycoprotein (AHSG)	P02765	Regulator of osteogenesis	(Yi *et al*., 2009; Saroha *et al*., 2012)
5	Alpha-enolase	P06733.2	Abundantly expressed in the synovial membrane microenvironment	(Terrier *et al*., 2007; Fan *et al*., 2015)
6	BRAF	P15056	Role in pro-inflammatory cytokine generation	(Charpin *et al*., 2010; Arnoux *et al*., 2016)
7	Calpastatin	BC013579	Modulator of articular cartilage proteolysis	(Auger *et al*., 2009; Fukuta *et al*., 2011)
8	Collagen type II	NM_001844	Major component of articular and hyaline cartilage	(Wegner *et al*., 2010; Lindh *et al*., 2014)
9	Fibrinogen Beta and Gamma	P02675 P02679	Regulation of cell adhesion and spreading	(van Beers *et al*., 2010; Alemany *et al*., 1996)
10	Filaggrin	NM_002016	Filament aggregating protein within synovial epithelium	(Union *et al*., 2002; van Venrooij and Pruijn, 2014)
11	Gelsolin	P06396	Actin filaments remodelling inside cells	(Mazumdar *et al*., 2012)
12	Glial fibrillary acidic protein (GFAP)	P14136	Abundantly contained in chondrocytes	(García, Weigum and Koke, 2003; Kanazawa *et al*., 2017)
13	Human cartilage glycoprotein 39	NM_001276	Modulator of tissue remodelling	(Boots *et al*., 2007)
14	IgG1 Fc	AF150959	Effector functions of the antibody	(Janda *et al*., 2016)
15	Keratin type II	Q9NSB2.2	Role in maintaining cell structural integrity	(Wang *et al*., 2015; Bhattacharjee *et al*., 2016)
16	Lumican	AAA85268	Regulator of collagen fibril organisation	(Chang *et al*., 2013; Gesteira *et al*., 2017)
17	Peptidyl arginine deiminase 4 (PAD4)	NP_036519	Role in the generation of citrullinated epitopes	(Auger *et al*., 2012)
18	Tubulin beta-chain (TUBB)	NP_821133	Major component of cytoskeleton and microtubules	(Libusová *et al*., 2005; Badillo-Soto *et al*., 2016)
19	Vimentin	NM_003380	Abundant intermediate filament protein involved in collagen stabilisation	(Challa and Stefanovic, 2011)

Table 3.01. Most prevalent autoantigens in rheumatoid arthritis selected as gold standards. RA autoantigens obtained from the literature used for the validation of B cell epitope prediction methodologies outlined within this chapter.

Using this dataset, both servers for linear epitope prediction were evaluated for accuracy and success of predicting established RA autoantigenic epitopes. Analysis was performed using five physico-chemical residue properties i.e. hydrophilicity, accessibility, flexibility, polarity and β-turns (Table 3.02) and each of the 19 autoantigens was used to grade the prediction success rate for each platform. As demonstrated in Table 3.03 and Figure 3.02 (A) BcePred server achieved 21% increase in prediction accuracy over ExPASy Protscale, predicting 17/19 (89%) established epitopes whilst ExPASy recognised 13/19 (68%) regions. Discrepancies between the two epitope prediction platforms were observed with BcePred predicting some epitopes that ExPASy considered insignificant and vice versa (Figure 3.02 B). Following this analysis, it was decided to incorporate both methods as complementary tools for increased sensitivity and specificity in B cell epitope predictions, an approach previously employed by others (Freimanis *et al*., 2010). All 19 autoantigen sequences were subsequently aligned with one another in order to identify regions of sequence homology between antigen pairs (Table 3.04) which could potentially contribute to RA perpetuation through mimicry mechanisms and autoantibody generation. Homologies of 5 or more consecutive residues between antigenic pairs were selected as significant since antibodies typically recognise linear epitopes of between 5 to 8 amino acids. Moreover, identical peptide segments of 5 or more residues are considered to be rare in nature (Trela, Nelson and Rylance, 2016) and were therefore of interest in our study.

Scale	
ExPASy Protscale	**BcePred**
Hydrophilicity (Hopp and Woods, 1981)	Hydrophilicity (Parker, Guo and Hodges, 1986)
Flexibility (Bhaskaran and Ponnuswamy, 2009)	Flexibility (Karplus and Schulz, 1985)
Accessibility (Janin, 1979)	Accessibility (Emini *et al*., 1985)
Polarity (Grantham, 1974)	Polarity (Ponnuswamy, Prabhakaran and Manavalan, 1980)
β-turns (Chou and Fasman, 1974)	β-turns (Pellequer, Westhof and Van Regenmortel, 1993)

Table 3.02. Parameters used by ExPASy Protscale and BcePred servers to predict linear B cell epitopes. Algorithm profiles were evaluated and optimised by the indicated research groups.

Autoantigen	Published sequence	ExPASy Protscale	BcePred
1	92-GRVRVNSAY-100	M	M
2	22-AIFYETQPSLWAESE-36	M	M
3	377-LEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK-418	0	M
4	104-QLKEHAVEGDCDFQLLK-120	M	M
5	56-RYMGKGVS-63	0	M
6	566-YLHAKSIIHRDLKSNNIFLH-585	M	M
7	305-DKDGKPLLPEPEEKP-319	M	M
8	324-GPRGLPGERGRTGPAGAA-341	M	M
9	43-ARGHRPLDKKREEAPSLRPA-62	M	M
10	831-HSASQDGQDTIRGHPGSS-848	0	M
11	35-SQAGAPQGRVPEARPN-50	0	M
12	60-AGFKETRAS-68	M	M
13	263-RSFTLASSETGVG-275	M	M
14	256-TPEVTCVVVDVSHED-270	M	M
15	157-NAQRVKKDEKEQI-169	0	M
16	326-YECLRVANEVTLN-338	0	0
17	211-VRVFQATRGKLSSKCSVVLG-230	M	M
18	81-FGQIFRPDNFVFGQS-95	M	M
19	428-LDSLPLVDTH-437	M	0
	Total successfully predicted (out of 19)	13	17
	Combined % prediction success (successful predictions / total number of published epitopes)	94.7% (18/19)	

Table 3.03. Summary of the levels of success of two in silico methods of B cell epitope prediction, ExPASy Protscale and BcePred, in assessing 19 previously published linear epitopes. M = match between published epitopes and epitopes predicted in this analysis. 0 = no match between predicted epitopes and epitopes published in the literature.

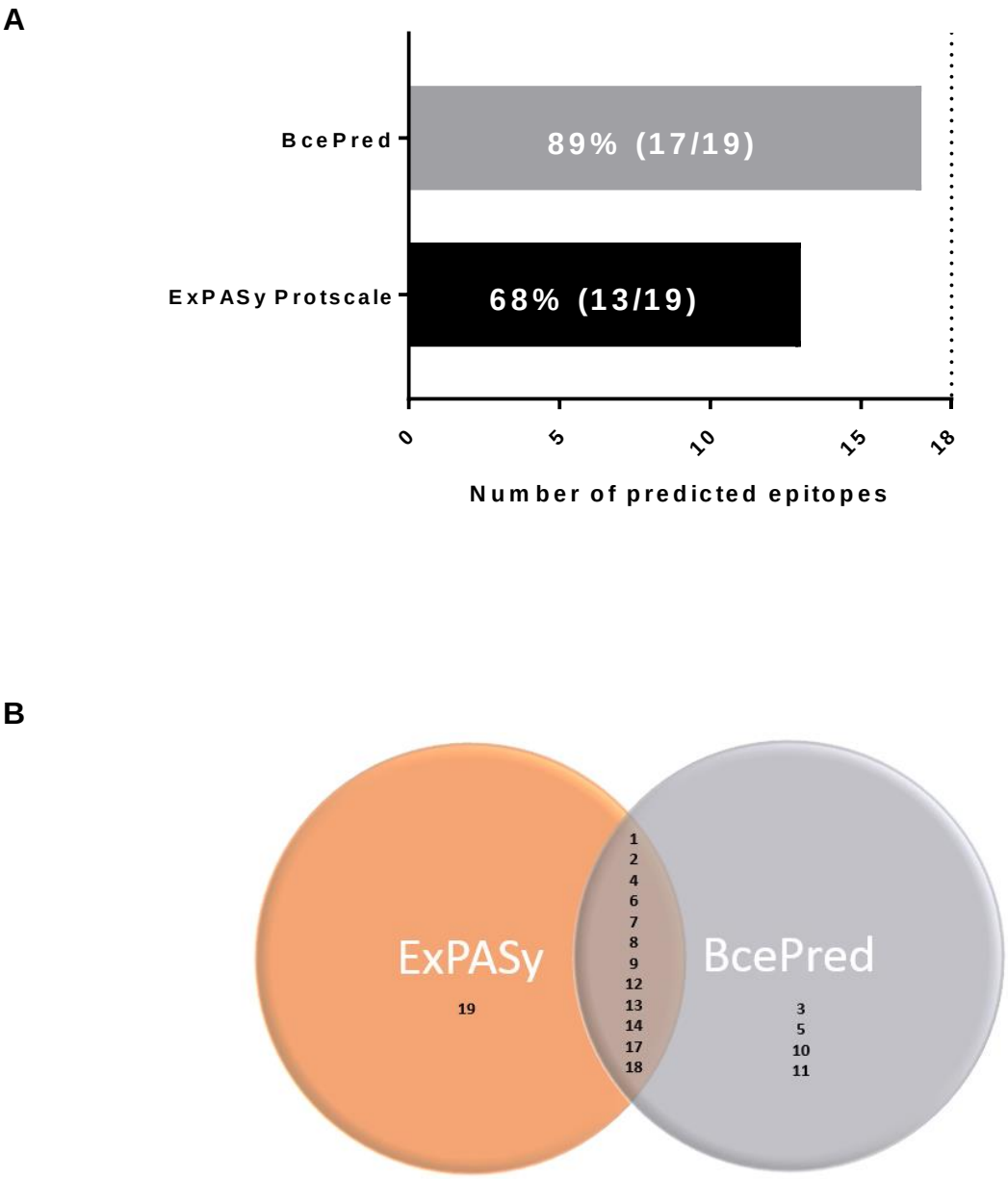


Figure 3.02. B cell epitope prediction success of ExPASy and BcePred platforms in identifying 19 previously published linear epitopes. (A) % prediction success presented as successful predictions / total number of published epitopes. **(B)** Overlap between the two platforms in predicting previously published epitopes. Numbers represent the particular epitope sequences listed in Table 3.03.

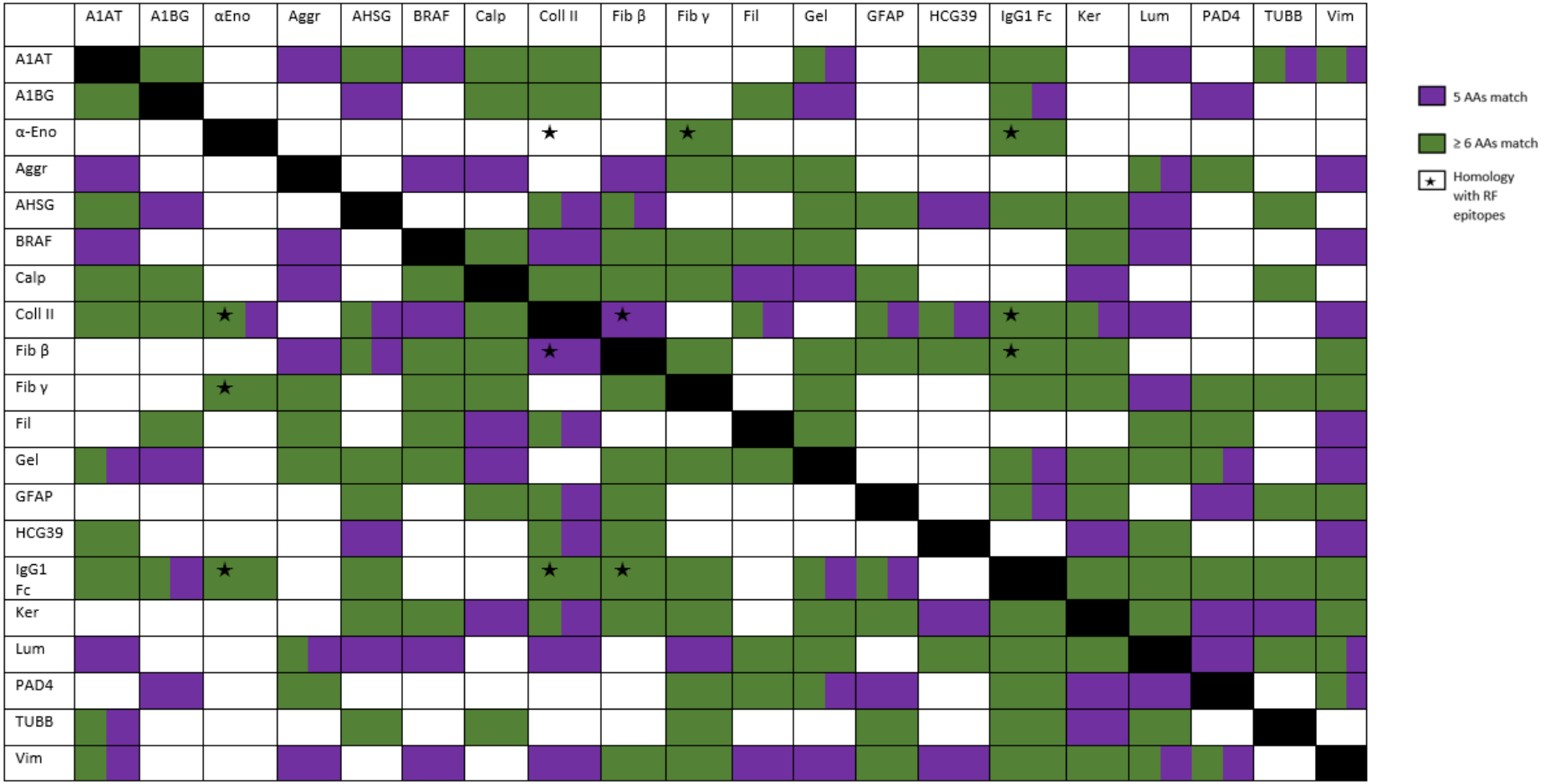


Table 3.04. Degree of homology between 19 most prevalent autoantigens in rheumatoid arthritis that were included in the epitope mapping in this investigation. All sequences used for putative autoantigens were published reference sequences on the NCBI public database. Whole protein sequences of all autoantigens in the table were aligned against each other using publicly available bioinformatics software ExPASy SIM and LALIGN. Autoantigens: A1AT- Alpha-1-antitrypsin, A1BG- α1B-glycoprotein, α-Eno- α-enolase, Aggr- Aggrecan, AHSG- Alpha 2 HS glycoprotein, BRAF- B-raf murine sarcoma viral oncogene homolog B1, Calp- Calpastatin, Coll II- Collagen Type II, Fib β- Fibrinogen beta, Fib γ- Fibrinogen gamma, Fil- Filaggrin, Gel- Gelsolin, GFAP- Glial fibrillary acidic protein, HCG39- Human Cartilage glycoprotein-39, IgG1 Fc- Fc fragment of immunoglobulin G1, Ker- Keratin type II cuticular Hb4, Lum- Lumican, PAD4- Peptidyl arginine deiminase 4, TUBB- Tubulin beta-chain, Vim- Vimentin. AAs- amino acids, RF- rheumatoid factor.

3.2.2. Identification of HERV-K10 epitope candidates for testing their potential role in molecular mimicry

When the sensitivity and specificity of prediction methods was evaluated, the technique was applied to predict potential epitopes present within HERV-K10 Gag region. A total of six antigenic regions were predicted for both viral matrix and capsid/nucleocapsid (Table 3.05). Resulting epitopes were aligned with IgG1 Fc sequence in order to identify regions of sequence homology. Table 3.06 demonstrates the alignments between epitopes of both HERV-K10 Gag and IgG1 Fc derived regions which exhibited high percentage of homology (over 70%). These data highlighted two peptide sequences, each composed of 4 linear residues (NGQP, PPSR) with 100% identity in terms of alignment scores with HERV-K10, and three 5 residue IgG1 Fc sequences (GKEYK, PELLG, VDKSR) with 80% identity to HERV-K10. Overall, the respective substitutions (tyrosine/leucine, leucine/methionine and valine/isoleucine) of these particular peptides scored highly in terms of amino acid replacement values (+16 to +24), meaning that they are structurally similar. The alignment score of one IgG1 Fc 7 residue sequence KVSNKAL, exhibited 71% similarity to HERV-K10. However, whilst the single amino acid substitution asparagine / threonine, provided a reasonable replacement score (+12), the second substitution alanine/asparagine fared poorly (-8).

	B-cell epitope prediction (using ExPASy Protscale and BcePred)
HERV-K10	**Matrix**
	25 39 RGGV**<u>KVSTKNL</u>**IKLF
	60 72 WKRI**<u>GKELK</u>**QAGR
	164 176 EGKG**<u>PELMG</u>**PSES
	Capsid and Nucleocapsid
	20 31 NPMA**<u>PPSR</u>**QGSE
	32 44 LHEI**<u>IDKSR</u>**KEGD
	366 377 KFDK**<u>NGQP</u>**LSGN

Table 3.05. Potential B cell antigenic regions mapped out on the surface of human endogenous retrovirus K10 in matrix, capsid and nucleocapsid regions. Immunodominant epitopes (underlined, in bold) were predicted using ExPASy Protscale and BcePred *in silico* bioinformatics platforms. Numerals indicate amino acid position with reference to protein accession numbers: HERV-K10 Gag1 (Matrix) (AAA88030.1) and Gag2 (Capsid and Nucleocapsid) (AAA88031.1). Amino acid single letter code: G=Glycine, A=Alanine, S=Serine, T=Threonine, C=Cysteine V=Valine, L=Leucine, I=Isoleucine, M=Methionine, F=Phenylalanine D=Aspartate, E=Glutamate N=Asparagine, Q=Glutamine, P=Proline, K=Lysine, R=Arginine, H=Histidine, W=Tryptophan, Y=Tyrosine.

HERV-K10 / autoantigen	Bioinformatic Pairwise Alignments					
	Matrix			Capsid and Nucleocapsid		
HERV-K10	29-KVS**T**K**N**L	64-GKE**L**K	168-PEL**M**G	24-PPSR	36-**I**DKSR	370-NGQP
IgG1 Fc	322-KVS**N**K**A**L	316-GKE**Y**K	232-PEL**L**G	352-PPSR	412-**V**DKSR	384-NGQP
Align. Score (%)	5/7 (71)	4/5 (80)	4/5 (80)	4/4 (100)	4/5 (80)	4/4 (100)
Repl. Score	+12 -8	+16	+24	N/A	+46	N/A

Table 3.06. Amino acid sequence alignments for HERV-K10 and IgG1 Fc. Matching homologous amino acid pairs were established using publicly available bioinformatics software ExPASy SIM and LALIGN. Residues highlighted in bold indicate amino acid substitutions. Align. Score**-**Alignment score shown as number of identical residues / number in peptide segment. Repl. Score-Replacement score for amino acid substitution according to Tűdős *et al.,* 1990. Total % homology of HERV-K10 Gag1 and Gag2 with IgG1 Fc (Accession number AF150959.1) circa 5%.

We next sought to determine whether the predicted B cell epitopes coincide with potential T cell epitope predictions as T cell help is essential in B cell mediated mechanisms in RA. Both HERV-K10 Gag and IgG1 Fc regions were assessed by the IEDB server for binding to the HLA DRB1 molecule, which appears to confer susceptibility to RA. As demonstrated in Table 3.07 there were 4 epitopes in both HERV-K10 Gag and IgG1 Fc with an IC_{50} < 500 nM predicted to be able to bind to the HLA DRB1 molecule with intermediate affinity whilst the remaining two epitopes bound with either high (IC_{50} < 50 nM) or low (IC_{50} < 5000 nM) affinity. From this analysis it was concluded that HERV-K10 Gag epitopes are capable of being presented by MHCs associated with RA. Overall, the predicted T cell epitopes from HERV-K10 Gag were located within regions homologous to the IgG1 Fc protein and therefore coincided with linear B cell epitope predictions.

HERV-K10 Gag	Predicted binding			Homologous IgG1 Fc epitope	Predicted binding		
	Allele (HLA)	IC_{50}	Strength		Allele (HLA)	IC_{50}	Strength
29-KVSTKNL	DRB1*04:05	171	Inter	322-KVSNKAL	DRB1*01:01	153	Inter
64-GKELK	DRB1*01:01	46	High	316-GKEYK	DRB1*01:01	153	Inter
168-PELMG	DRB1*01:01	382	Inter	232-PELLG	DRB1*01:01	38	High
24-PPSR	DRB1*01:01	107	Inter	352-PPSR	DRB1*01:01	119	Inter
36-IDKSR	DRB1*01:01	1795	Low	412-VDKSR	DRB1*01:01	101	Inter
370-NGQP	DRB1*01:01	356	Inter	384-NGQP	DRB1*04:04	2726	Low

Table 3.07. T-cell epitope predictions using IEDB MHC-II binding predictions for alleles HLA-DRB1*01:01, *01:04, *04:01, *04:04, *04:05, *09:01. Potential T-cell epitopes were predicted using the publicly available bioinformatics methodology Stabilised Matrix Method (SMM). Numerals indicate IC_{50} values representative of binding affinity: < 50 nM - high affinity, < 500 nM intermediate affinity, < 5000 nM low affinity. Inter – predicted intermediate affinity of binding.

3.2.3. Incorporation of negative controls into the study

Frequently small datasets include a significant bias for negative and positive data, as data was determined thorough a rigorous set of parameters (Gibb *et al*., 2015; Soria-Guerra *et al*., 2015). In order to reduce potential bias and its effect on a set of results we incorporated non-specific epitopes which were expected to show insignificant or no binding. Therefore, potential epitopes determined on HERV-K10 Gag and IgG1 Fc proteins were aligned with two presumed negative control sequences which were predicted by ExPASy and BcePred from the two control sequences i.e. tartrate resistant acid phosphatase (TRAP) protein and from retroviral plant protein of *Arabidopsis thaliana* (Table 3.08). No sequence homology was observed for any of the investigated proteins hence a minimal bias was presumed within the study structure.

Peptide	Organism	Accession number	Sequence
Gypsy/Ty3	Arabidopsis thaliana	AAG51464.1	^{413} VAVADGRKLNVDGQIK ^{428}
TRAP	Homo sapiens	P13686	^{311} SGKSLFKTRLPRRARP ^{325}
MAG1	HERV-K10	AAA88030.1	^{60} RIGKELKQAGRKGNI ^{72}
IgG1 Fc	Homo sapiens	AF150959.1	^{90} LNGKEYKCKVSNKAL ^{104}

Table 3.08. The potential antigenic determinants predicted for tartrate resistant acid phosphatase (TRAP) and plant retroelement of *Arabidopsis thaliana* negative controls. Epitopes were determined using the ExPASy Protscale and BcePred platforms. The two underlined epitopes indicate those used as controls for the pairwise alignment with HERV-K10 and IgG1 Fc epitopes to reduce bias in the investigation and were employed as negative control peptides in subsequent serological studies.

3.2.4. Shortlisting of candidate epitopes for synthesis as peptides

Those epitopes predicted for HERV-K10 Gag that demonstrated the highest degree of homology with IgG1 Fc autoantigen were compared with IgG1 Fc sequence to determine the best candidates for peptide synthesis and screening of patient serum. This was important, as it determined whether the predicted viral epitopes, which constitute potential molecular mimics involved in triggering RA pathogenesis, correspond to antigenic regions on IgG1 Fc that are recognised by and accessible to rheumatoid factor autoantibodies in RA.

Table 3.09 shows that 232-PELLG-236 was present on a lower hinge region residues and part of the fx1 β–pleated face of the CH2 domain which were not solvent accessible. Sequences 316-GKEYK-320 and 322-KVSNKAL-328 were present on exposed segments of the CH2 domain and located largely on the fy2 β–pleated face and on part of the b5 and b6 loop (or bend) of the domain, respectively. Both sequences were solvent accessible. Sequence 352-PPSR-355 was located on the fx1 β–pleated face of the CH3 domain whilst sequence 384-NGQP-387 was present on the b3 loop of the domain. Sequence 412-VDKSR-416 was located largely on the fx4 β–pleated face and a small segment of the b5 loop of the CH3 domain. Since B cell epitopes are typically located in the vicinity or within structural loops (Zheng *et al*., 2015), it was established that all regions except PELLG were capable of interaction with antibodies/rheumatoid factor, therefore PELLG was excluded from further investigation. Sequences were subsequently analysed for antigenicity and physico-chemical properties. As presented in Table 3.10. the remaining five regions were found to be within the protein antigenic regions. In addition, all sequences, with the exception of KVSNKAL, were also recognised by RFs. Moreover, properties of individual residues comprising the predicted epitopes of HERV-K10 Gag and their neighbouring amino acids were evaluated to assess their potential for interaction with

autoantibodies. Table 3.12 shows that KVSNKAL, GKELK, PPSR and IDKSR epitopes contained polar residues with positive charge (highlighted in green) which are crucial for antibody reactivity (Dam *et al*., 2014), whereas NGQP epitope was predominantly composed of non-charged (cyan) and aliphatic (magenta) residues, and was therefore excluded from further analysis.

AA no. 231 241 251 261 271

IgGFc1 APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD

---e1---/----- fx1 (CH2 domain) -----/---b1-----/---------- fx2----------/--- b2 ---/---------fy1-------/

AA no. 281 291 301 311 321

IgGFc1 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA

---b3-----/--------fx3----------/-b4---/----------fx4----------/------b5----/---------fy2--------/----b6-----/

AA no. 331 341 351 361 371

IgGFc1 PIEKTISKAK GQPREPQVYT LPPSRDELTK NQVSLTCLVK GFYPSDIAVE

---------fy3----------/-e2-/----- fx1 (CH3 domain) ----/-b1---/----------fx2----------/-b2-/--------

AA no. 381 391 401 411 421

IgGFc1 WESNGQPENN YKTTPPVLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE

-fy1/---b3------/--------fx3----------/-b4--/----------fx4----------/----b5--------/----------fy2----------

AA no. 431 441

IgGFc1 ALHNHYTQKS LSLSPGK

--b6---/--------fy3----------/

XX – RF accessible K10 mimic YY – RF non-accessible K10 mimic

Table 3.09. Four rheumatoid factor-accessible epitopes on IgG1 Fc with their location within the protein secondary structure. Regions of homology with HERV-K10 Gag epitopes accessible XX (boxed, green) and non-accessible YY (boxed, red) to RF antibodies identified through bioinformatics. Sequences aligned with the protein structure taken from Nelson *et al.,* 2008. IgG constant domain segments: face X (fx1-4), face Y (fy1-3), bends/loops (b1-6), non-β-structure (e1, e2).

AA no.	231
IgGFc1	APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD
AA no.	281
IgGFc1	GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA
AA no.	331
IgGFc1	PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE
AA no.	381
IgGFc1	WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE
AA no.	431
IgGFc1	ALHNHYTQKSLSLSPGK

XX – K10 mimic YY – antigenic regions ZZ - RF epitopes

Table 3.10. Identification of five HERV-K10 homologues within antigenic regions of IgG1 Fc. Homologues XX (boxed), antigenic regions YY (shaded, non-boxed) and published epitopes recognised by rheumatoid factor antibodies identified through bioinformatics and indicated in red.

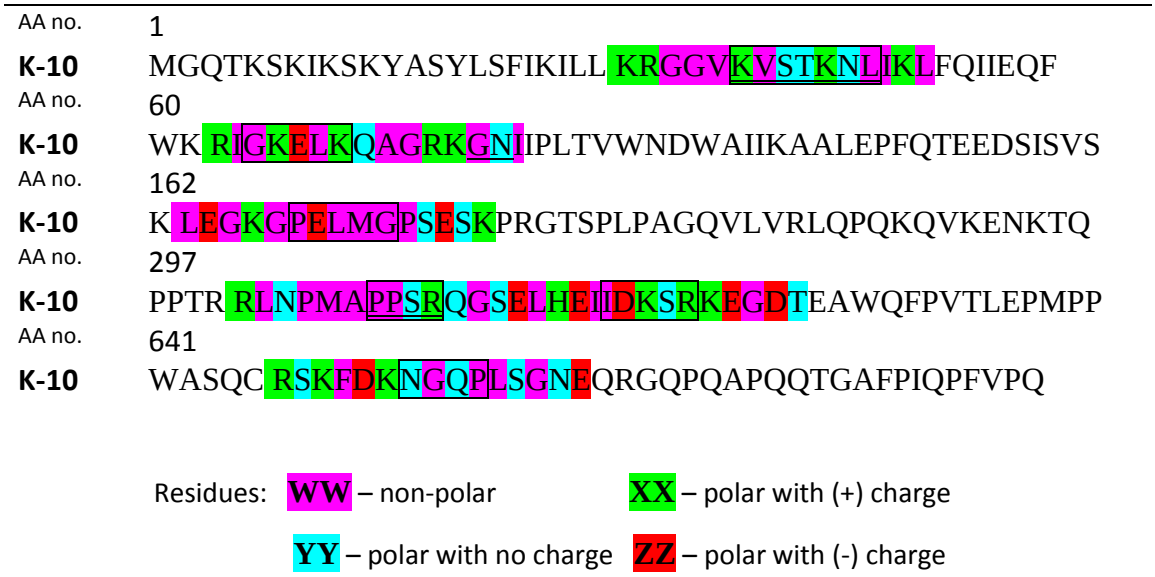


Table 3.11. Suggested HERV-K10 matrix epitopes have physico-chemical features of immunodominant epitope. Amino acid sequence of Gag epitopes (boxed), which constitute IgG1 Fc mimics, were identified through bioinformatics. Physico-chemical properties were specified for the epitopes and neighbouring residues constituting 15-mer peptides.

Four homologous sequences, KVSNKAL, GKELK, PPSR and IDKSR, which were selected as potential candidate epitopes were subsequently converted into 3D molecular models in order to investigate their structural mimicry. In broad terms, HERV-K10 and IgG1 Fc peptides showed very little difference in overall mesh structure. This enabled us to demonstrate the conformational similarity between epitopes within the RA autoantigen and viral sequences which are likely to be targeted by rheumatoid factor autoantibodies in RA (Table 3.12, columns 1 and 2). Additional molecular modelling using a PDB file of IgG1 Fc positioned the candidate homologues within the tertiary protein structure (Table 3.12, column 3) and showed that all four potential peptide mimics were surface exposed and accessible epitopes Visualisation of the viral epitopes within HERV-K10 tertiary structure is currently not achievable since the protein has not been crystallised and therefore K10 molecular model is not available.

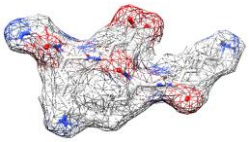
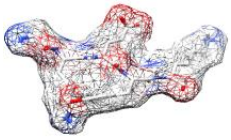
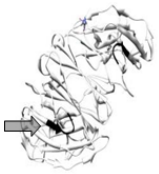
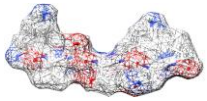
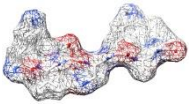
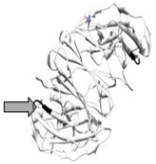
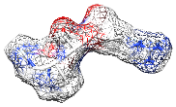
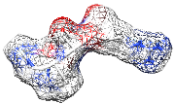
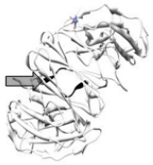
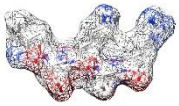
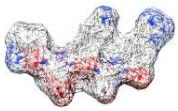
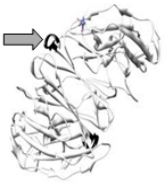
HERV-K10 epitope	IgG1 Fc mimic	Location on IgG1 Fc
Matrix		
GKE**L**K 	GKE**Y**K 	IgG1 Fc **CH2** 
KVS**T**K**N**L 	KVS**N**K**A**L 	IgG1 Fc **CH2** 
Capsid and Nucleocapsid		
PPSR 	PPSR 	IgG1 Fc **CH3** 
IDKSR 	**V**DKSR 	IgG1 Fc **CH3** 

Table 3.12. Molecular models of homologous epitopes located on HERV-K10 Gag1 / Gag2 and IgG1 Fc presented as 'mesh structure'. Three dimensional structures of homologous regions between HERV-K10 and IgG1 Fc were modelled in PyMOL and UCSF Chimera molecular graphics systems and presented in mesh mode together with their amino acid sequences. Location of Fc1 mimics is presented on IgG Fc1 model (arrowed).

Table 3.13 demonstrates those epitopes that are most likely to contribute to triggering and/or perpetuation of RA via mechanisms of molecular mimicry. Novel peptides of significant interest are all immunodominant 'MAG' (MAtrix Gag) peptides that exhibited homology to epitopes on IgG1 Fc and coincide with rheumatoid factor binding regions. MAG1 is a predicted immunodominant and antigenic B cell epitope which appears to resembles the shared epitope motif Q/R-K/R-x-x-A that confers susceptibility to RA and is associated with the presence of rheumatoid factors (de Almeida, Ling and Holoshitz, 2011). It is also enriched in polar residues with positive charge, features characteristic for B cell epitopes (Kringelum *et al*., 2013), which are structurally homologous with those on the corresponding fragment of IgG1 Fc. MAG1 is predicted to bind strongly (IC_{50} = 46) to the RA-associated HLA DRB1 molecule and can be potentially presented to Th cells with high affinity. Therefore, the peptide within the MAG1 region was selected as the primary focus for subsequent serological investigation.

Peptide name	**HERV-K10 Gag1 (MAG1)**	**HERV-K10 Gag2 (MAG2)**	**HERV-K10 Gag2 (MAG3)**	**HERV-K10 Gag1 (MAG4)**
Peptide sequence	RI**GKELK**QAGRKGNI	PMA**PPSR**QGSE	I**IDKSR**KEGDT	KRGGV**KVSTKNL**
Homology with IgG1 Fc	5 amino acid mimic with high (+16) replacement score	4 amino acid mimicry (no replacement residues)	5 amino acid mimicry with high (+46) replacement score	7 amino acid mimicry with moderate (+12, -8) replacement score
Epitope within HERV-K10 immuno-dominant region (structure)?	Yes (matrix)	Yes (capsid/ nucleocapsid)	Yes (capsid/ nucleocapsid)	Yes (matrix)
Epitope within RF binding region?	Yes	Yes	Yes	No
Comments	- 1st most antigenic peptide on K10 Gag1 - high affinity binding (IC_{50} = 46) by HLA DRB1 - resembles shared epitope motif Q/R-K/R-x-x-A	N/A	N/A	- 2nd most antigenic peptide on K10 Gag1

Table 3.13. Peptides selected as best homologue candidates for implicating HERV-K10 in the pathogenesis of rheumatoid arthritis. Peptides were selected through bioinformatics for their homology to antigenic regions on the main target of rheumatoid factor autoantibody, IgG1 Fc. Immunodominant epitopes were underlined and indicated in bold.

3.3. Discussion

Activation of autoreactive B cells, which upon T cell stimulation produce high levels of autoantibodies that bind to antigens at antigenic determinant regions, appears to constitute a key element in the pathogenesis of RA. Therefore, knowledge of B cell epitopes of proteins is of particular interest in RA as they can contribute to disease diagnosis and development of immune-therapeutics. Unlike time-consuming and labour-intensive experimental epitope mapping, bioinformatic screening provides a faster and low-cost alternative for preliminary epitope mapping. Owing to advances in bioinformatic tools and their use in the investigation of HERV involvement in autoimmunity, including *in silico* assessment of HERV-W in MS (Ramasamy, Joseph and Whittall, 2017) and HRES-1 in SLE (Nelson *et al*., 2014b), the role of HERV-K10, one of the most biologically active forms of endogenous virus, has gained increasing interest in autoimmune arthritis (Nelson *et al*., 2004; Ehlhardt *et al*., 2006). Due to the recent amplification of the majority of HERV-K elements (less than 3 million years ago), they preserved some of their retroviral open reading frames and thus the potential to form viral particles (Lemaître *et al*., 2014).

In this chapter a method for the prediction of potential viral B cell epitopes, which may play a role in the disease pathogenesis, was developed and optimised. An *in silico* approach was employed that permitted epitope mapping of autoantigens and HERVs previously associated with RA. For the purpose of this investigation, we compiled 19 previously published autoantigens which have been implicated in the disease process and that represent linear epitopes. These 19 epitopes were used as cross-validation controls to assess the performance of B cell epitope prediction platforms. This epitope list could be further enhanced by other antigens that are targeted by antibodies in RA such as citrullinated histones (Pratesi *et al*., 2014). The existing procedure employed in-house involves predictions based on five properties of individual amino acids i.e.

hydrophilicity, flexibility, polarity, solvent accessibility, and β-turn propensity, which when considered together have been reported to result in successful epitope predictions (Soria-Guerra *et al*., 2015). Moreover, these computer-aided strategies of B cell epitope predictions have been successfully pursued by several groups in design and analysis of anti-viral candidate vaccines. Cheung *et al*. (2010) used *in silico* approach, with similar level of successful predictions to the one established in our study, to predict epitopes on H5N1 influenza A, which he later combined with *in vitro* studies to confirm the peptide immunogenic potential, whilst Zheng *et al*. (2017) and Alonso-Padilla, Lafuente and Reche (2017) identified promising linear B cell epitopes for HBV and EBV targeting.

Therefore all 5 algorithms were used to perform initial studies using default values established by authors of original experimental antigenic propensity scales (Table 3.02). Two parallel epitope prediction strategies, which rely on protein composition algorithms, were employed to ensure prediction reliability (Table 3.03). It became evident that both prediction tools exhibited varying levels of accuracy. Overall, the BcePred server, which unlike ExPASy allows users to select two or more residue properties simultaneously and does not require manual prediction calculation, achieved better performance. When evaluated using the pre-published regions, BcePred predicted 89% (17/19) of epitopes compared to the 68% (13/19) established by ExPASy. However, when applied together, the prediction platforms exhibited a 94.7% (18/19) combined success rate. The importance of comparative analysis of different epitope prediction tools for enhanced accuracy has been recognised by others (Guedes *et al*., 2018; Moult, 2005) and has therefore been employed in this study. Although it appears that predictions obtained through BcePred outperform those made by ExPASy, it should be borne in mind that this evaluation relied on a small dataset of 19 epitopes. Studies with low sample size and small datasets have been reported to

yield biased performance estimates (Button *et al*., 2013; El-Manzalawy and Honavar, 2010). Therefore, any future experiments should incorporate large and non-redundant datasets. Another limitation is related to underperformance of linear epitope predictors which can label some amino acids incorrectly as non-epitope or epitope residues and therefore lead to false negative or false positive results, respectively. This is complicated by processes such as epitope spreading, which involves diversification of epitope specificity from dominant to subdominant/cryptic epitopes recognised by antibodies, and molecular mimicry which is considered a strong contributor to B cell epitope spreading (Cornaby *et al*., 2015). Nevertheless, sequence-based epitope mapping is advantageous for predictions where the structure of the target antigen is not known (El-Manzalawy and Honavar, 2010), which is the case with HERV proteins, and this prediction method has been shown to achieve an accuracy of up to 72.5% (Potocnakova, Bhide and Pulzova, 2016).

Once developed and optimised, this methodology was used to identify potential epitopes of HERV-K10 Gag (Table 3.05). Retroviral Gag polyproteins are crucial for the generation of viral particles and contribute to the assembly of key structural regions including matrix, capsid and nucleocapsid. These polyproteins assemble at the plasma membrane and, despite their loss of infectivity due to inactivating mutations, facilitate the release of the viral particles out of the cell (Kraus *et al*., 2011). Nevertheless, constituting part of the human chromosome, HERV-encoded fragments should be perceived as autoantigens unless they fail to undergo thymic presentation and are regarded as neo-antigens (Gröger and Cynis, 2018). They could also share immunologic epitopes with host proteins leading to immunoreactivity (Fujinami *et al*., 2006). This phenomenon between human endogenous retroviruses and autoantigens, known as molecular mimicry, has been considered a mechanism in the aetiology and pathogenesis of several autoimmune diseases. In this light, *in silico* studies contributed

to the identification of homologous epitope pairs between HERV-W and myelin sheaths in multiple sclerosis. HERV-W proteins are believed to induce autoreactivity and ineffective myelin repair due to its pro-inflammatory properties and oligodendrocyte toxicity (do Olival *et al*., 2013). Novel biological therapy (GNbAC1 monoclonal antibody) targeting and blocking these viral regions in MS is currently in clinical phase IIb trials (Curtin *et al*., 2015). In addition, this biological agent is also in clinical development for type I diabetes where it has been shown to neutralise the cytotoxic effects of HERV-W on glucose and insulin metabolism (Curtin *et al*., 2018). Moreover, computationally predicted homologous fragments have been detected between HRES-1 virus and several autoantigens in systemic lupus erythematosus (Nelson *et al*., 2014b). Molecular mimicry has also been considered a pathogenic mechanism in RA traditionally between *Proteus mirabilis* bacterium and sequences found in HLA-DR4/1 haplotypes (Rashid and Ebringer, 2012) and, more recently, HERV-K10 (Trela, Nelson and Rylance, 2016). We therefore assessed potential homology between HERV-K10 Gag and IgG1 Fc, the key known autoantigen protein in RA, using a local alignment programme (Table 3.06).

In total, six regions were found that were either identical (NGQP, PPSR) or possessed one or two amino acid substitutions (GKEY/LK, PELL/MG, V/IDKSR, KVSN/TKA/NL). Moreover, varying degrees of homology were observed between the 19 most prevalent autoantigens in RA with 157 matches of 6 consecutive residues and 93 matches of 5 consecutive residues some of which, including fibrinogen and α-enolase, coincide with regions on IgG1 Fc protein recognised by rheumatoid factor autoantibody (Table 3.04).

Overall, the sequence homology observed between different proteins with strings of 4 residues or more declined with increasing peptide length. However, the relative significance of the latter was important when discerned from the binomial expansion of 20^n (20, total number of amino acids; n, number of residues). Here the probability of a

unique set of residues in a peptide chain of 3, 4, 5, 6, and 7 amino acids was estimated as 1 in 8000, 160,000, 3 million, 64 million, 1,128 million and greater, respectively. Consequently it might be perceived that a peptide containing a unique set of residues of ≥ 5 (4 consecutive residues considered as borderline match) that was identical to a second peptide, could be significant when assessing potential molecular mimics and constitute a 'rare event'. In addition, a high replacement score for a given amino acid substitution (e.g. +24 for valine / isoleucine) that would share similar physio-chemical properties was seen as pivotal for retaining peptide / protein structure and function (Tűdős, Cserző and Simon, 1990). Indeed, our mesh configurations of molecular models (Table 3.11) highlighted little perturbation in shape with acceptable amino acid substitutions.

Both molecular modelling and sequence analysis (Table 3.09 and Table 3.10) of IgG1 Fc revealed that 4 HERV-K10 / IgG1 Fc peptide mimics were exposed on the surface of the protein, accessible, and thus likely to be antigenic sites. This observation was consistent with epitope mapping studies in that all 4 regions coincided with linear antigenic determinants identified by rheumatoid factor antibodies. In general, relatively small linear epitopes (of 5-8 amino acids) can be recognised by rheumatoid factors and specifically generated antibodies (Westwood *et al*., 2008). In addition, conformational epitopes from amino acids brought into juxtaposition through molecular folding may depend on one or two key residues within a given epitope (Nelson *et al*., 1990).

In light of the above, both linear and conformational epitopes might contribute towards molecular mimicry in RA and possible epitope spreading should additional B cells become activated (Cornaby *et al*., 2015). These processes might also be facilitated by T cell mimics that would promote T cells to facilitate isotype switching. Thus, it is likely that short viral peptide sequences of HERV-K10 could mimic regions of IgG1 Fc in RA.

In turn, this might help drive autoantibody production particularly if the IgG macromolecule was no longer recognised as body's own protein by the mechanism of self-tolerance (Benson *et al*., 2010) due to the physio-chemical degradation or structural alteration of IgG e.g. mediated by defects in glycosylation.

Ordinally, HERV-K integrated into the human genome as closely linked viral peptide sequences about 40 million years ago, to subsequently become separated out during evolution. We observed that the motif KVSTKNL preceded GKELK in HERV-K10 Gag by 28 residues. In the case of IgG1 Fc however, we noticed that the motif order was reversed and separated by a single residue. Since HERVs are widely distributed within the human genetic material as a result of multiple amplification and transposition events (Gröger and Cynis, 2018), and have the potential to generate mutations (Löwer *et al*., 1993), we hypothesise that this viral region became conserved within IgG1 Fc giving rise to the predicted regions of homology. It is interesting to note that other HERV families have been integrated into the human genome as conserved regions to carry out physiological roles such as HERV-W functional protein syncytin-1, which is essential for placenta formation, or HERV-H involved in embryonic development (Hurst and Magiorkinis, 2017). Moreover, as a result of their relatively recent amplification, HERV-K elements exhibit strong polymorphism at viral insertion sites and at specific loci, therefore it is possible that certain individuals will have alleles with a particular viral sequence whilst others may not have a functional allele in their chromosomes. This may affect HERV-K expression and lead to differences in immune responses (Lemaître *et al*., 2014).

Those epitopes of HERV-K10 Gag which shared sequence homology with IgG1 Fc were short-listed as candidates for generation of short polypeptides and for subsequent anti-HERV-K10 antibody characterisation as well as serological screening of patients' blood for reactivity to viral epitopes (Table 3.13). Of the short-listed epitopes

MAG1 was the first peptide selected for synthesis. This peptide containing 15 residues, which was found to be the optimum number of residues in an average B cell epitope (Kringelum *et al*., 2013), showed the best alignment with IgG1 Fc, the primary target of rheumatoid factor autoantibodies. This epitope was in the top four most antigenic regions on HERV-K10 Gag, in addition to being situated within the K10 immunodominant region and rheumatoid factor binding region. It is also located within the flexible β-turn structure which is highly characteristic of an epitope. MAG1 peptide contains multiple polar and charged residues with two hydrophobic amino acids (glycine, leucine) situated within the epitope fragment which are typically located closest to the antibody (Rahman *et al*., 2016). Moreover, T cell epitope prediction analysis identified that MAG1 can be potentially presented to Th cells with high affinity (IC_{50} < 50 nM) by HLA DRB1, the Class II molecule that has been associated with RA (Boissier *et al*., 2012). Lastly, this peptide resembles the shared epitope motif Q/R-K/R-x-x-A which resides within HLA DRβ chain and constitutes a major risk factor for severe arthritis (Ling *et al*., 2010). Thus percentage homology, peptide length, replacement residue value and molecular modelling may all contribute to the successful identification of peptide mimics that could be otherwise overlooked when using gross homology searches between proteins. Once identified *in silico* peptides can be developed into research tools, diagnostic applications and therapeutic agents such as blocking peptides used in renal idiopathic membranous neuropathy to prevent autoantibody interactions induced by molecular mimicry mechanisms (Havasi *et al*., 2017).

Currently there appears to be a limited understanding of the properties that characterise a functional B cell epitope, therefore the performance of prediction methods is not optimal and requires standardisation. This is further complicated by the stochasticity of antibody binding, discontinuous nature of numerous B cell epitopes

with closely spaced residues (Rahman *et al*., 2016) and high genetic variation in the MHC molecule for T cell predictions (de Jong *et al*., 2010). Therefore, further development of standard benchmark datasets used to assess the prediction approaches and meta-servers merging the predictions of several prediction methods is required. Hence, it is necessary to incorporate the *in silico* predictions with experimental data in order to allocate function (Soria-Guerra *et al*., 2015), an approach which we used in subsequent chapters.

4. GENERATION AND CHARACTERISATION OF A POLYCLONAL ANTIBODY TO HERV-K10 GAG

4.1. Introduction

The generation of reliable diagnostic immunoassays or immunotherapeutic strategies requires, among others, scrupulous evaluation of potential antibody reagents. Polyclonal antibodies enable detection of a variety of epitopes and therefore identify antigen from distinct orientations. This may constitute a salient feature in some systems which, with the use of a single epitope or certain immobilisation procedures, could otherwise exhibit assay restriction phenomenon i.e. an antibody performing well in one assay system but reacting poorly in another. Therefore, it is necessary to verify the reactivity profiles of novel reagents in different experimental techniques.

We have previously described (chapter 3) the use of bioinformatic algorithms to predict antigenic peptides within host and viral proteins that can be tested in immunoassay techniques. In this study, we identified a highly antigenic peptide on the matrix protein of HERV-K10 (denoted as MAG1: RIGKELKQAGRKGNI) with sequence and structural similarity to IgG1 Fc, which has also been identified as a key antigenic target for RF autoantibodies in RA (Nelson *et al*., 2014a). This finding suggested a potential region for molecular mimicry that exhibited little variation in overall structure and sequence. In order to further investigate the role of HERV-K10 mimicry with IgG1 Fc in RA, in this chapter we developed a polyclonal antibody (PAbMAG1) to the MAG1 peptide sequence and tested the reactivity of this antibody to IgG1 Fc mimics. At present, there is a paucity of antibodies to HERV-K10. Hence a panel of standardised reagents would provide specific laboratory probes to determine molecular mimicry, epitope spreading and antibody reactivity to native and modified proteins. Therefore, in this chapter, we

set out to generate and characterise the anti-MAG1 (PAbMAG1) antibody in a range of immunoassay systems.

4.2. Results

4.2.1. Peptide/antigen selection and antiserum development

As demonstrated in Table 4.01, we have determined four synthetic peptides (MAG1 - 4) of 11-15 residues using *in silico* investigation of the putative HERV-K10 Gag protein sequence, which were then synthesised commercially. These peptides exhibit sequence similarity to IgG1 Fc autoantigen and may constitute potential regions for molecular mimicry in RA. Additionally, MAG1 peptides of different lengths (3-, 4-, 5-, 7-, 9-, 12- and 15-mer) were developed to analyse the effect of epitope size on antibody binding. Two further peptides termed NC1GTY3 and 8023 incorporating epitopes derived from a retroelement of *Arabidopsis thaliana* and tartrate resistant acid phosphatase, respectively, were used as negative control peptides. In our bioinformatic analysis we predicted the 15-mer peptide sequence MAG1, which contains the immunodominant 'GKELK' region homologous to that on IgG1 Fc 'Key' peptide (LNG**KEY**KCKVSNKAL) derived from IgG CH2 domain, as highly antigenic and likely to provoke an immune response. A polyclonal antibody was therefore generated for investigation of HERV-K10 mimics. New Zealand white rabbits were immunised with PPD-conjugated MAG1 peptide by Severn Biotech Ltd. and antisera PAbMAG1 were harvested upon subsequent fortnightly antigen boosters.

Peptide name	Length	Sequence	Organism derived	Accession no.	Justification
MAG1	15-mer 12-mer 9-mer 7-mer 5-mer 4-mer 3-mer	RIGKELKQAGRKGNI RIGKELKQAGRK RIGKELKQA IGKELKQ GKELK GKEL GKE	HERV-K10 Gag	MI14123	Most antigenic K10 peptide Homology to published 'GKEYK' epitope on IgG1 Fc (synthesised as IgG1 Fc 'Key' peptide) (Westwood *et al*., 2008)
MAG2	11-mer	PMAPPSRQGSE	HERV-K10 Gag	MI14123	Homology to published epitope on IgG1 Fc. Control sequence
MAG3	11-mer	IIDKSRKEGDT	HERV-K10 Gag	MI14123	Homology to epitope on IgG1 Fc Control sequence
MAG4	12-mer	KRGGVKVSTKNL	HERV-K10 Gag	MI14123	Homology to epitope on IgG1 Fc Control sequence
IgG1 Fc Key	15-mer	LNGKEYKCKVSNKAL	Homo sapiens	AF150959.1	Published IgG1 Fc epitope homologous with MAG1 'GKELK' antigenic epitope
NC1GTY3 (Gypsy/Ty 3)	16-mer	VAVADGRKLNVDGQIK	*Arabidopsis thaliana* Gypsy/Ty3 fragment	AAG51464.1	Plant retroelement unrelated to HERVs
8023 (TRAP)	15-mer	GKSLFKTRLPRRARP	Homo sapiens	P13686	Tartrate resistant acid phosphatase protein unrelated to HERVs

Table 4.01. Peptides selected as candidates for synthesis and testing as coating antigens against PAbMAG1 and patient serum samples in ELISA. Underlined residues represent immunodominant epitopes identified through bioinformatics.

4.2.2. Bleed evaluation

Commercially synthesised 15-mer MAG1 peptide and its respective bleeds were evaluated in a peptide-based ELISA system to ascertain their functionality in addition to establishing the optimal concentration of purified serum for use within the assay. Rabbit tail bleeds were acquired pre-immunisation and at 14-day intervals including 14, 28 and 42 days following consecutive antigen booster injections. Serial dilutions of the crude PAbMAG1 pre-, 1st, 2nd, and terminal bleeds ranging from 1/125, down a two-fold dilution series to 1/8000 were analysed against a single concentration of MAG1 peptide (8µg/ml), which was previously optimised by our group (Freimanis *et al*., 2010), pre-coated on a polystyrene ELISA plate by passive adsorption. Results, demonstrated in Figure 4.01, indicated an increase in the strength of PAbMAG1 reactivity to the immunogen, upon administration of the consecutive boosters. This was recorded as increasing absorbance levels between all four successive bleeds with terminal bleed showing the highest reactivity across all antigen dilutions. As anticipated the pre-bleed exhibited no reactivity to the immunogen and therefore provides baseline serum antibody levels.

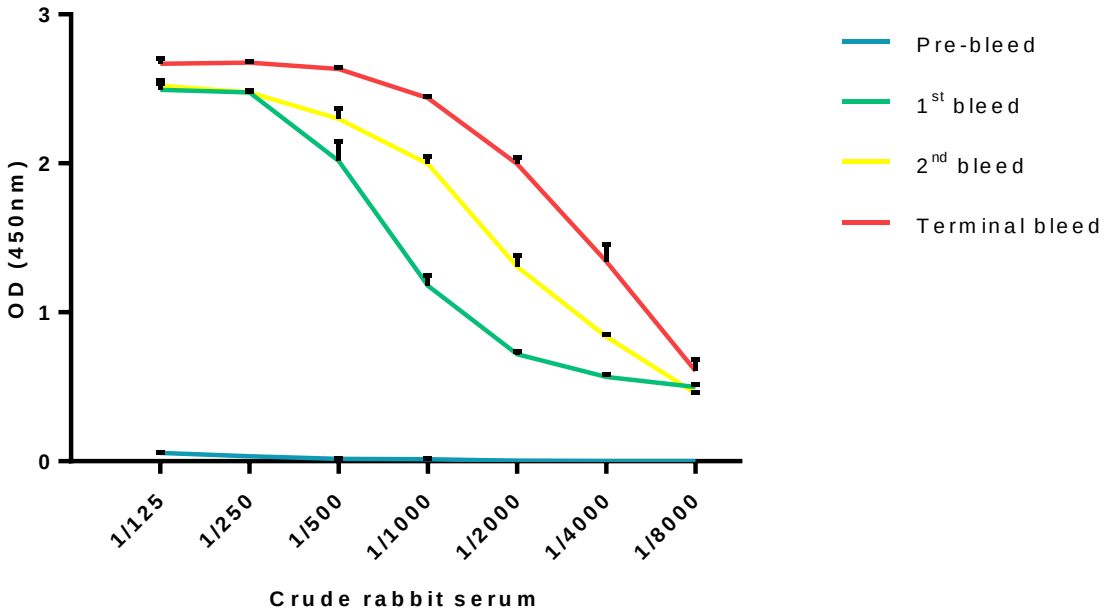


Figure 4.01. The reactivity of pre-bleed, 1st bleed, 2nd bleed and terminal bleed from the immunised rabbit with MAG1 peptide. Crude rabbit sera at indicated dilutions were incubated with MAG1 coating antigen at 8 µg/ml in ELISA system. Each bleed raised against MAG1 in the above figure represents a single sample taken from the rabbit at different time points over two months post immunisation. Pre-bleed represents levels of anti-MAG1 antibodies prior to immunisation with the peptide. 1st bleed shows levels of anti-MAG1 antibodies 14 days post-immunisation. 2nd bleed shows antibody titres after 28 days and the terminal bleed shows levels of anti-MAG1 antibodies 42 days post-immunisation. Error bars show standard error of the mean (SEM) for 3 experiments.

4.2.3. Optimisation of anti-MAG1 polyclonal antibody

In order to optimise the PAbMAG1 for subsequent immunoassays, antiserum from the terminal bleed was purified for the IgG fraction using affinity chromatography (Protein G column) and the protein concentration was determined. Purified antibody was subsequently verified for IgG isotype by anti-isotype ELISA (Figure 4.02). We next decided to titrate the PAbMAG1 against MAG1 peptides of different lengths, to determine the optimal antigen-antibody binding, and against two negative control peptides (NC1GTY3 and 8023) to measure antibody reactivity at different concentrations. The assays were performed in both non-biotinylated and biotinylated systems to determine the effect of peptide orientation within an ELISA plate on antibody accessibility. Data presented in Figure 4.03 and Figure 4.04 shows the antibody titres to MAG1 peptides with different number of residues, each peptide encompassing the entire or part of the immunodominant epitope sequence (GKELK). PAbMAG1 showed low levels of reactivity to short non-biotinylated peptides (3-7 mer) which improved with increasing peptide length (Figure 4.03). The efficiency of epitope presentation was drastically enhanced when biotin-conjugated peptides were used as antigens (Figure 4.04).

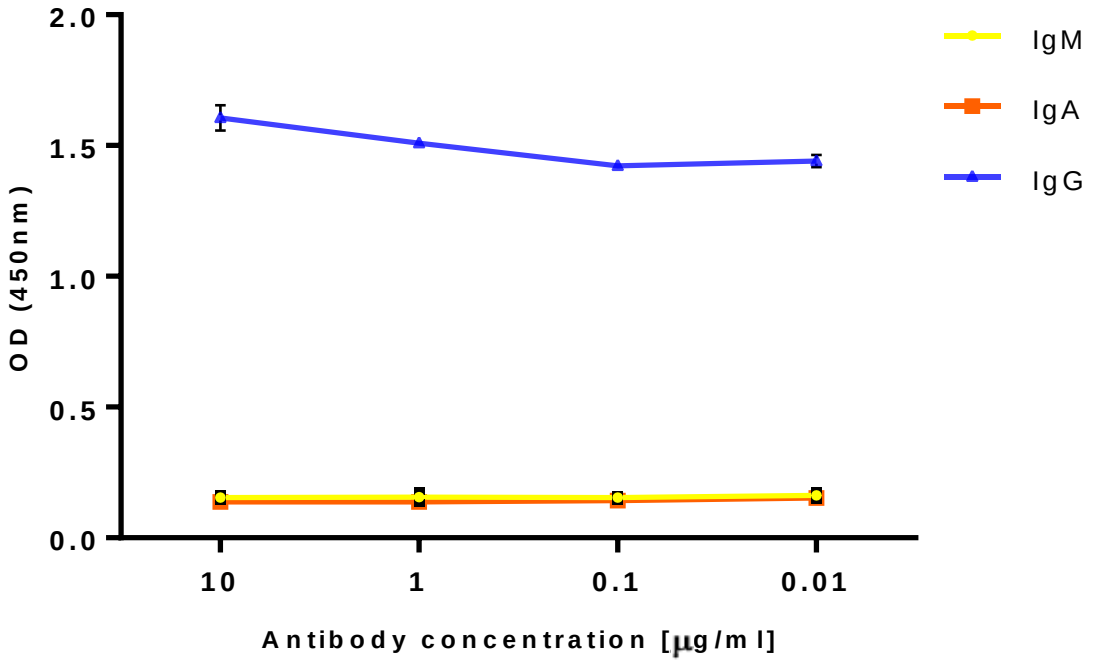


Figure 4.02. Verification of the isotype of PAbMAG1 purified terminal bleed with class-specific secondary antibodies. Purified IgG fraction from rabbit final serum sample collected 42 days post-immunisation with MAG1 peptide was titrated against MAG1 coating antigen at 8 µg/ml and probed with anti-isotype secondary Ab (anti-IgM, IgA, IgG) at 1/5000 in ELISA system. SEM is shown for 3 experiments**.**

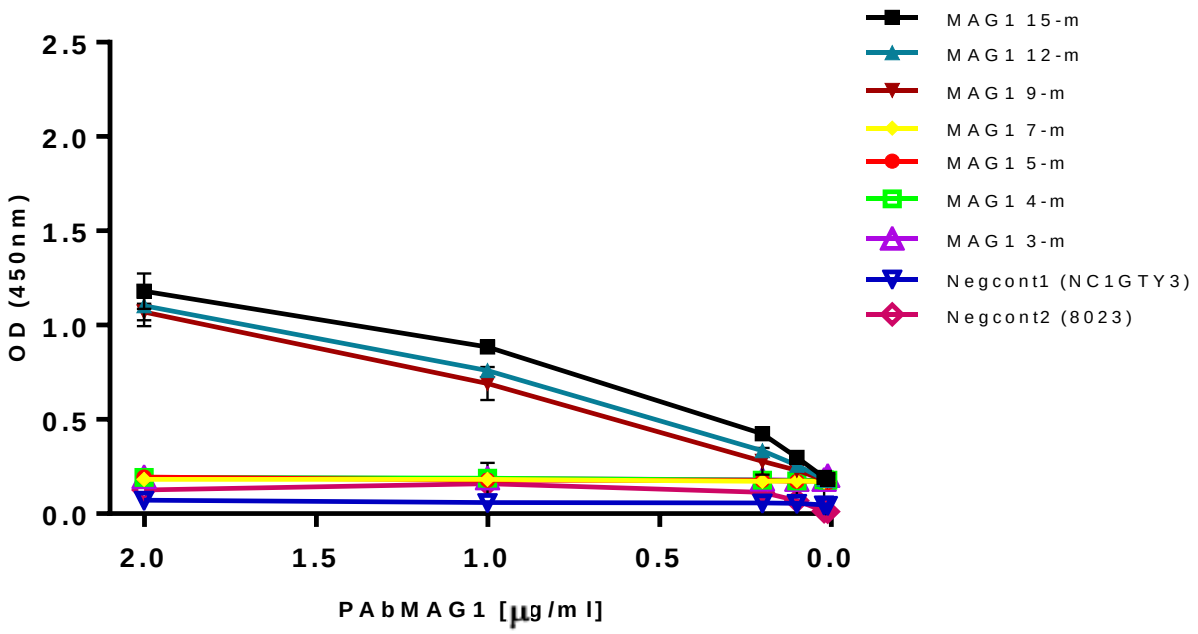


Figure 4.03. The reactivity of PAbMAG1 with MAG1 peptides of different lengths tested in non-biotinylated ELISA. Purified IgG fraction from MAG1-immunised rabbit serum was titrated against non-biotinylated MAG1 coating antigen consisting of between 3 and 15 residues and negative control peptides at 8 µg/ml. SEM is shown for 3 experiments.

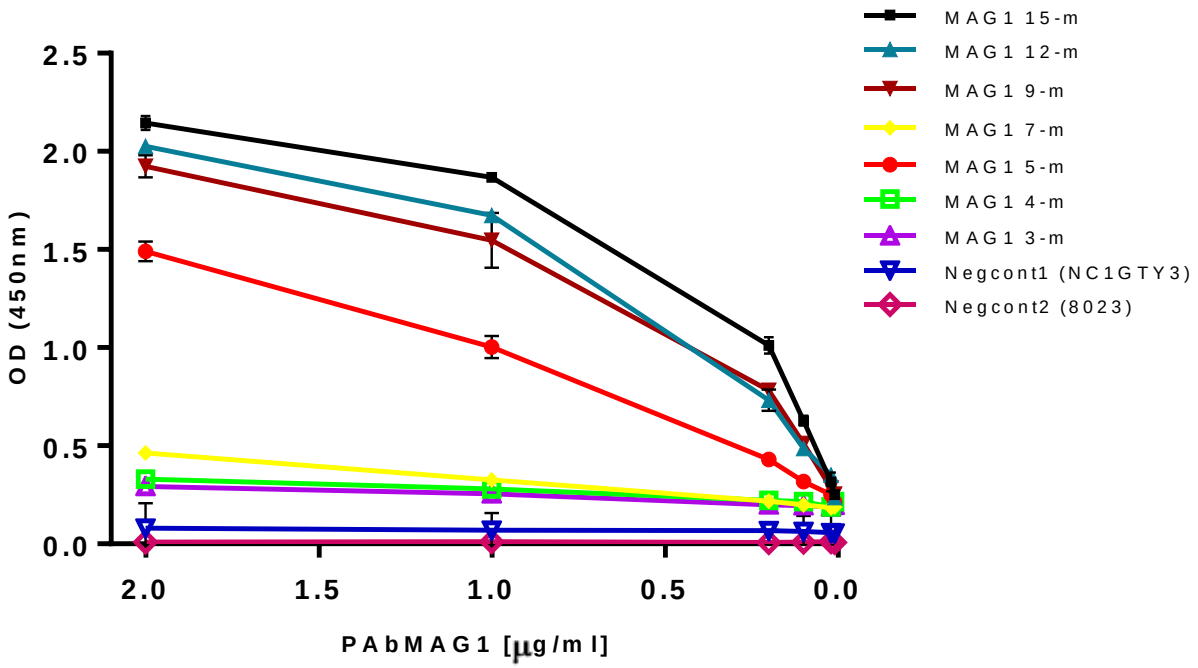


Figure 4.04. The reactivity of purified IgG PAbMAG1 with MAG1 peptides of different lengths tested in biotinylated ELISA. Purified IgG fraction from MAG1-immunised rabbit serum was titrated against biotinylated MAG1 coating antigen consisting of between 3 and 15 residues and negative control peptides at 8 µg/ml. SEM is shown for 3 experiments.

This was particularly noticeable for short biotinylated peptides where the improvement in epitope orientation is reflected in increased OD values compared to their non-biotinylated counterparts (Figure 4.05). As a result, the best PAbMAG1 reactivity was achieved at 2 µg/ml of the MAG1 15-mer biotinylated peptide, which was therefore used in subsequent immunoassays. PAbMAG1 showed negligible reactivity with two negative control peptides (OD < 0.1) in both systems. Additionally, no cross-reactivity was observed with three other MAG peptides in ELISA (Figure 4.06) therefore validating the specificity of PAbMAG1 to its target antigen (MAG1).

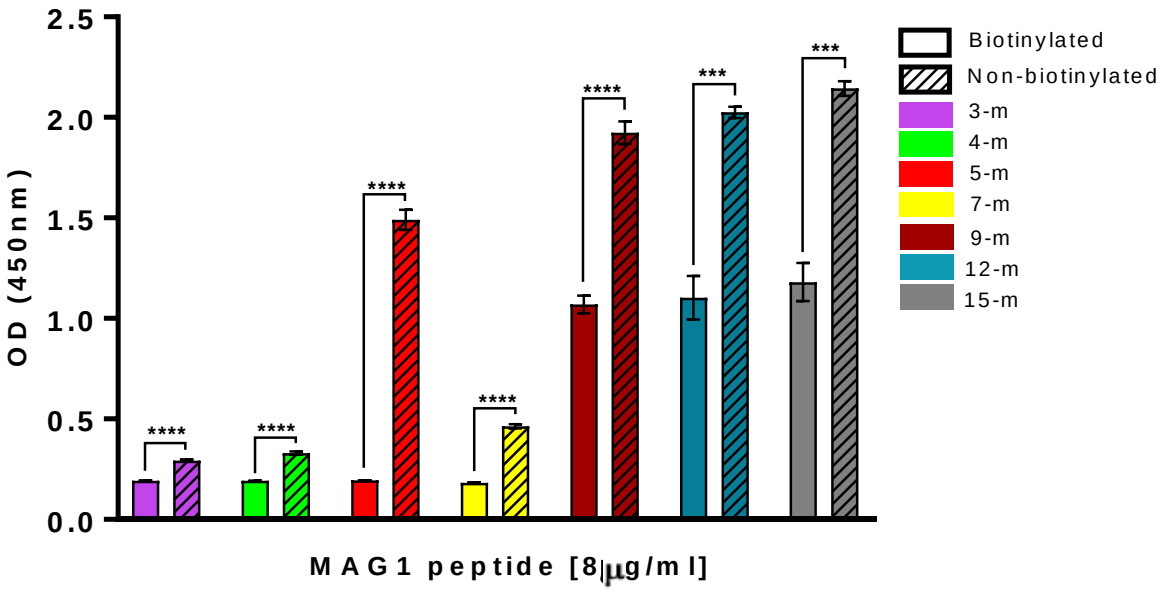


Figure 4.05. The effect of biotinylation of MAG1 peptides of different lengths on PAbMAG1 reactivity in ELISA. Purified IgG fraction from MAG1-immunised rabbit serum was incubated with non-biotinylated and biotinylated MAG1 coating antigen consisting of between 3 and 15 residues at 8 µg/ml. Bars represent means and SEM is shown for 3 experiments. ***, $p \leq 0.001$, ****, $p \leq 0.0001$.

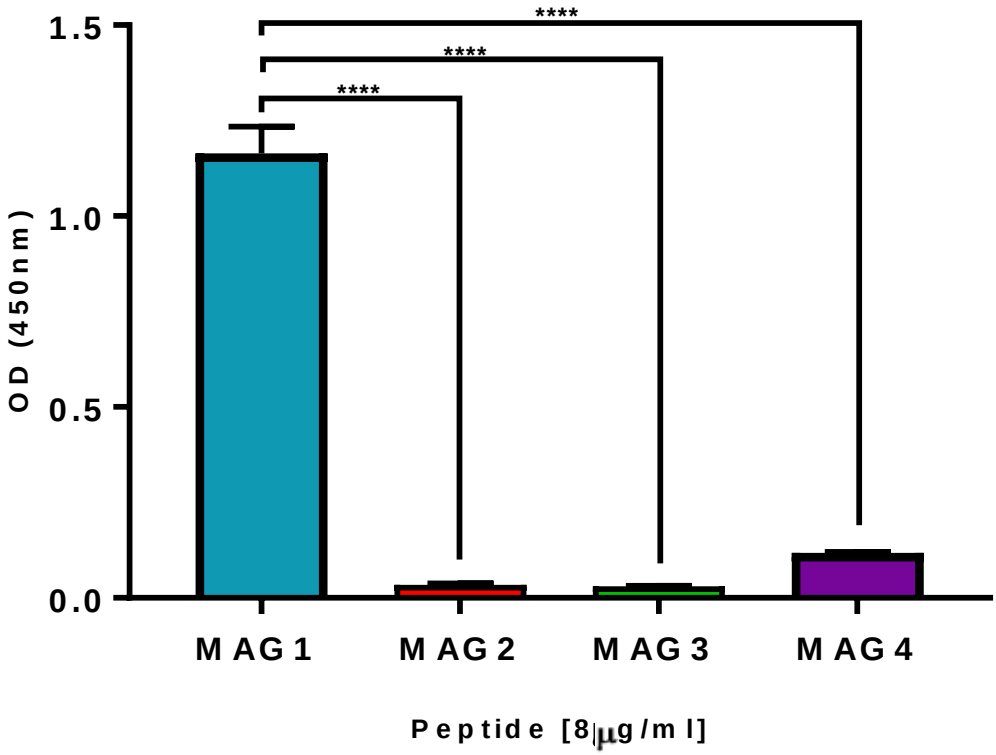
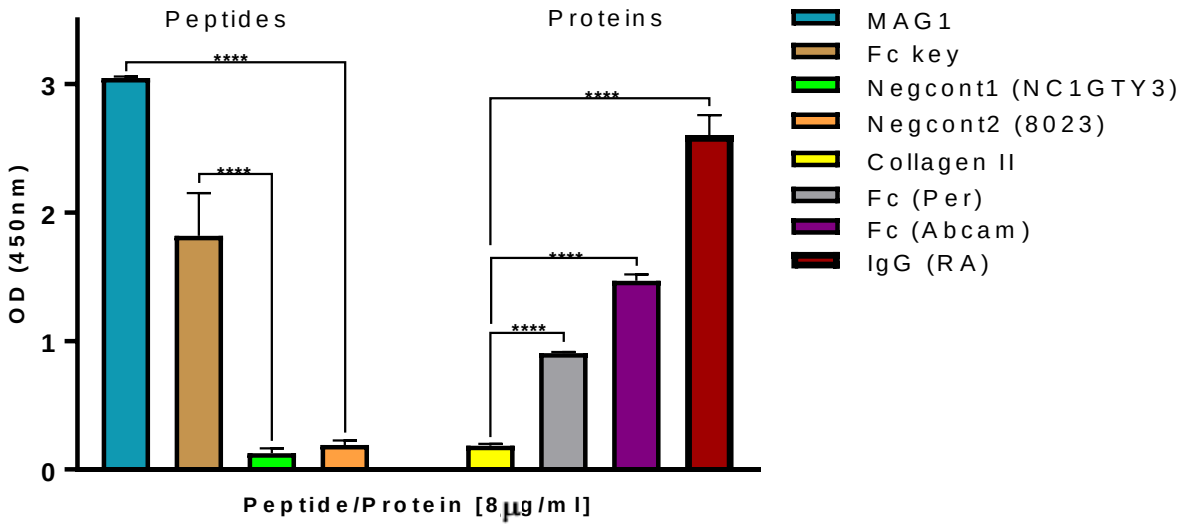


Figure 4.06. Cross-reactivity profile of PAbMAG1 with a panel of four MAG biotinylated peptides. Purified IgG fraction from MAG1-immunised rabbit serum was incubated at 2 µg/ml with biotinylated MAG1, MAG2, MAG3 and MAG4 coating antigens at 8 µg/ml in ELISA system. SEM is shown for 3 experiments. ****, $p \leq 0.0001$.

4.2.4. Assessment of PAbMAG1 antibody reactivity to IgG1 Fc

In chapter 3 our bioinformatic analysis highlighted six potential homologous amino acid sequence regions between HERV-K10 and IgG1 Fc autoantigen with the GKELK (MAG1) Gag matrix segment constituting a potentially strong molecular mimic that might be targeted in RA. Therefore, this finding prompted us to investigate cross-reactivity profiles of newly developed PAbMAG1 to IgG1 Fc preparations. As demonstrated in Figure 4.07 A, PAbMAG1 antibody reacted to both linear (Key peptide; mean OD value = 1.8) and conformational epitopes of IgG1 Fc retrieved from a patient with heavy chain disease (Per; OD = 0.9), where paraproteins with only heavy chain components and no light chains are generated (Malati, 2001), and commercially-produced Fc (Abcam; OD = 1.45). It also exhibited strong reactivity (OD = 2.6) to a purified IgG preparation obtained from RA serum. No cross-reactivity was observed with negative control peptides and collagen type II. A control pan-IgG monoclonal antibody A57H, which recognises the 383-SNGQPENN-390 peptide sequence on the IgG CH3 domain (Nelson *et al*., 1997) was used to validate reactivity against IgG1 Fc proteins and served as negative control against peptide preparations (Figure 4.07 B) as its target epitope does not share sequence homology with MAG1 or IgG1 Fc key peptides.

A



B

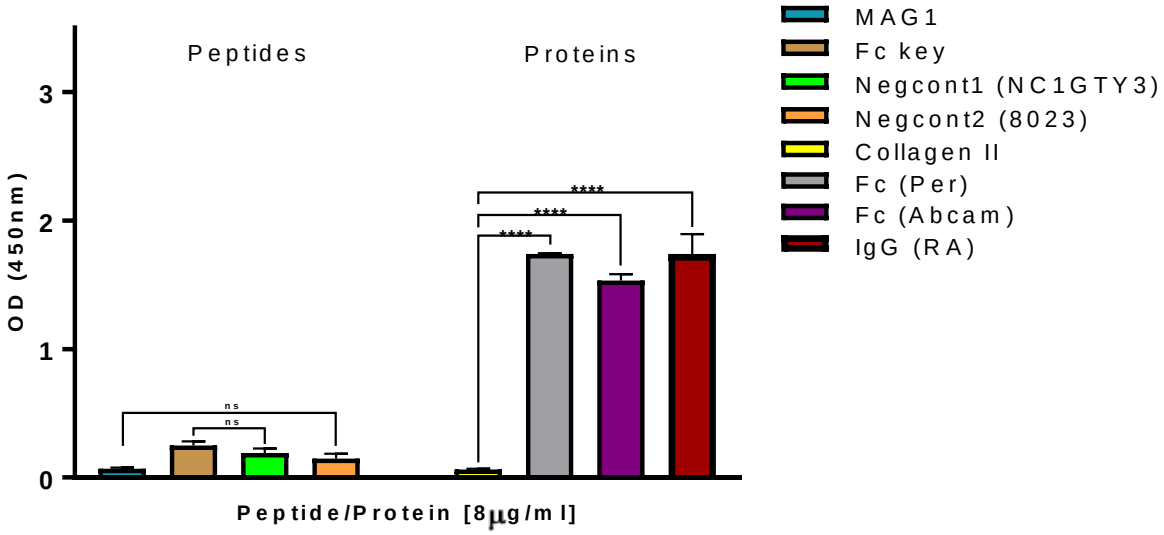


Figure 4.07. Cross-reactivity of PAbMAG1 with linear/peptide and conformational/protein epitopes of IgG1 Fc and HERV-K10 MAG1 mimics. Anti-K10 antibody PAbMAG1 at 2 µg/ml (A) and pan-IgG A57H control monoclonal antibody at 1 µg/ml (B) were incubated with protein and peptide coating antigens (8 µg/ml) in ELISA system. IgG1 Fc preparations were retrieved from a patient with heavy chain disease (Fc Per) and commercially-produced (Fc Abcam, Fc key). Purified IgG control from an RA patient and negative controls (NC1GTY3, 8023, collagen II) were also incorporated in this assay. SEM is shown for 3 experiments. ****, $p \leq 0.0001$, *ns* = not significant.

4.2.5. Inhibition assay confirming specificity of PAbMAG1 for MAG1 peptide

We next decided to perform an inhibition ELISA, to ascertain the specificity of PAbMAG1 for its respective peptide MAG1 (RIGKELKQAGRKGNI). This analysis also validated the *in silico* investigation confirming that the immunodominant epitope was exposed and able to react with specific antisera. Purified PAbMAG1 at its optimum concentration of 2 µg/ml was incubated with ten-fold dilutions of MAG1 15-mer peptide and negative control peptides before the preparation was transferred to a microtitre plate pre-coated with biotinylated MAG1 15-mer peptide. Results demonstrated in Figure 4.08 confirm that PAbMAG1 was specific for its respective antigen as an increase in the antibody binding to plate-bound MAG1, and thus increasing OD values, correlated with decreasing concentrations of pre-incubated MAG1 peptide. This confirms successful inhibition of the PAbMAG1 by its target peptide. High antibody specificity was reflected in the steep gradient of the curve. Conversely, the absorbance values of PAbMAG1 preincubated with non-related negative control peptides (NC1GTY3 and 8023) remained high and unchanged suggesting that they did not inhibit PAbMAG1 binding to its target peptide.

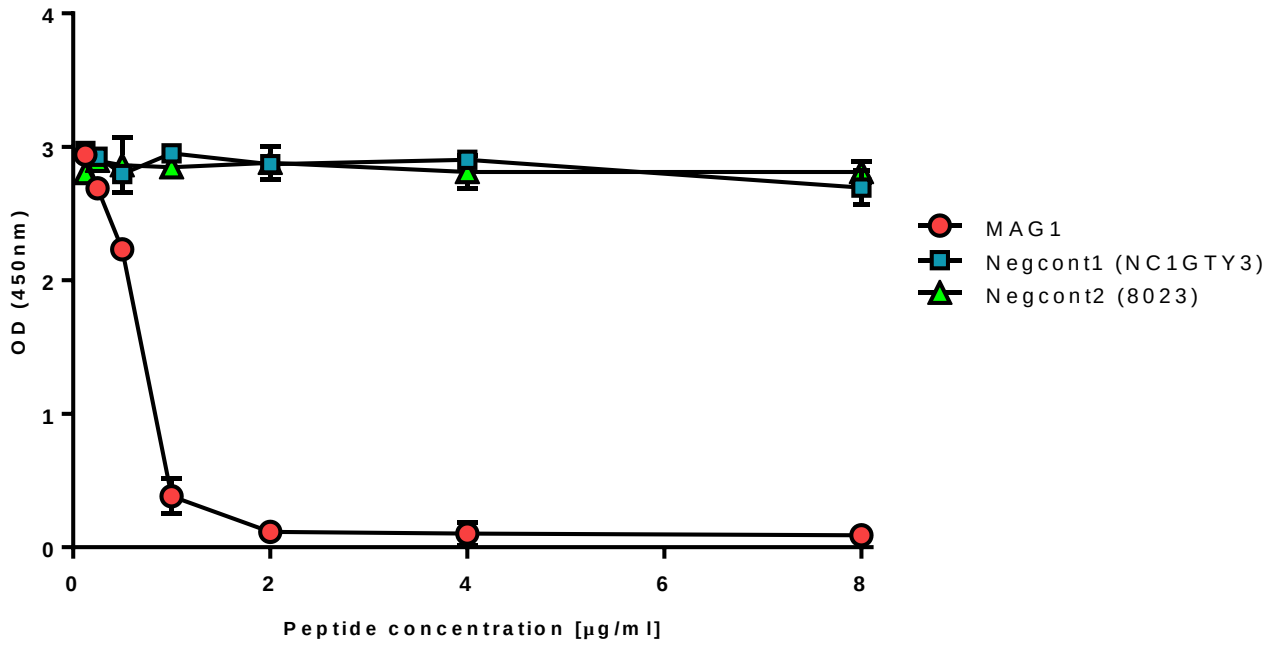


Figure 4.08. Inhibition of PAbMAG1 achieved with MAG1 peptide demonstrating the antibody specificity. PAbMAG1 at 2 µg/ml was pre-incubated with MAG1 peptide or NC1GTY3 and 8023 negative control peptide preparations (each at 2, 4, 6 and 8 µg/ml) for 1 hour prior to the antibody:peptide transfer to an ELISA plate pre-coated with MAG1 peptide at 8 µg/ml. SEM is shown for 3 experiments.

4.2.6. Assessment of HERV-K10 expression in cultured cells using immunocytochemistry and immunofluorescence *in vitro* assays

To determine its diagnostic potential, an *in vitro* analysis of the polyclonal antibody generated to HERV-K10 MAG1 peptide (PAbMAG1) was performed using immunocytochemistry (ICC) and immunofluorenscence (IF) on a number of cancer cell lines including GH, MCF-7, T-47D and KM-H2 (Figures 4.09 to 4.12), which were previously used for HERV-K expression studies (Simmons, 2016). GH teratocarcinoma cell line was reported to harbour HERV-K Gag particles (Boller *et al*., 2008) whilst MCF-7 and T-47D breast cancer cell lines were demonstrated to have elevated levels of HERV-K10 transcripts and viral particles (Ejthadi *et al*., 2005). On the other hand, KM-H2 Hodgkin's lymphoma cell line was shown to not express HERV-K Gag transcripts (Freimanis *et al*., 2010). We used anti-human HLA-ABC (MHC I) as a positive control antibody to ascertain that the staining procedure was successful. Rabbit IgG isotype control antibody and no primary antibody controls were incorporated to measure the level of non-specific background signal caused by primary and secondary antibodies, respectively. PAbMAG1 was applied as primary antibody at 2 µg/ml as used in the previous ELISA assays.

As expected, GH teratocarcinoma cell line exhibited strong cytoplasmic staining with PAbMAG1 in both ICC and IF (Figure 4.09 A and B, respectively) therefore serving as a positive control cell line. MCF-7 and T-47D breast cancer cell lines, exhibited moderate (MCF-7; Figure 4.10 A and B) and weak (T-47D; Figure 4.11 A and B) cytoplasmic immunostaining with PAbMAG1 in both systems. The staining in these malignant cells appears diffuse throughout the cytoplasm of the cell, which indicates cytoplasmic localisation of the retroviral Gag protein, that is necessary for viral particle formation before budding (Löwer *et al*., 1993). Upon staining with PAbMAG1, such patterns were not observed in KM-H2 Hodgkin's lymphoma cell line (Figure 4.12 A and

B), which showed negligible levels of staining, a finding supported by the lack of detectable HERV-K transcripts (Freimanis *et al*., 2010), and was therefore used as a negative control cell line. Cytoplasmic immunostaining was noted in all four cell lines tested with anti-HLA antibody, whilst no immunostaining was detected for isotype control and no primary antibody negative control. Mean fluorescence intensity (MFI) was assessed for all four cell lines (Figure 4.13) and validated the *in silico* investigation of accessible epitopes performed in chapter 3, confirming the binding of PAbMAG1 to HERV-K10 Gag epitopes.

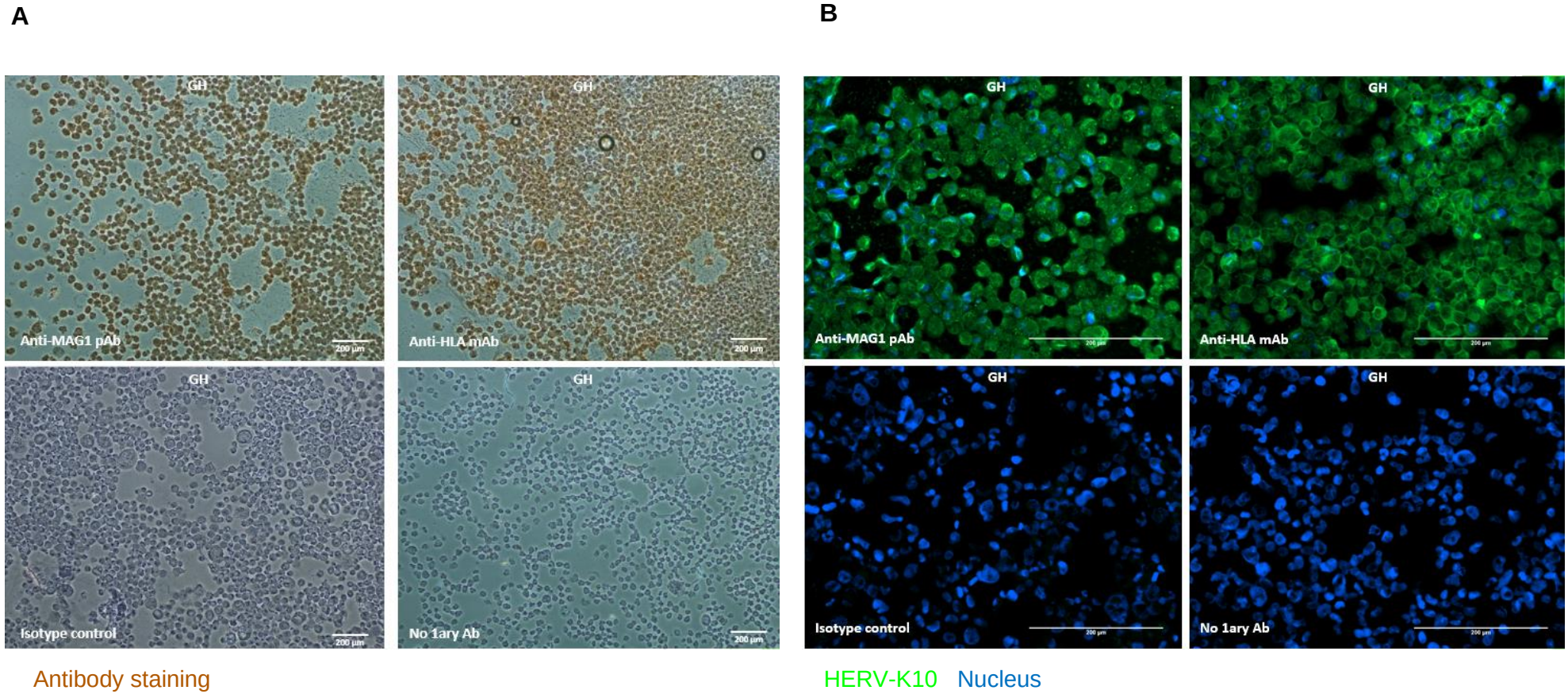


Figure 4.09. Examples of HERV-K10 Gag immunostaining in representative cytospins of GH teratocarcinoma positive control cell line. Cells at 10^5 were cytospun and stained with PAbMAG1, anti-HLA positive control antibody, isotype negative control antibody or secondary antibody only. Images show cytoplasmic labelling of GH teratocarcinoma cells in immunocytochemistry observed as brown staining (A; magnification x20) or as green staining (FITC) with blue DAPI-stained nuclei in immunofluorescence (B; magnification x40). Data are representative of 3 experiments.

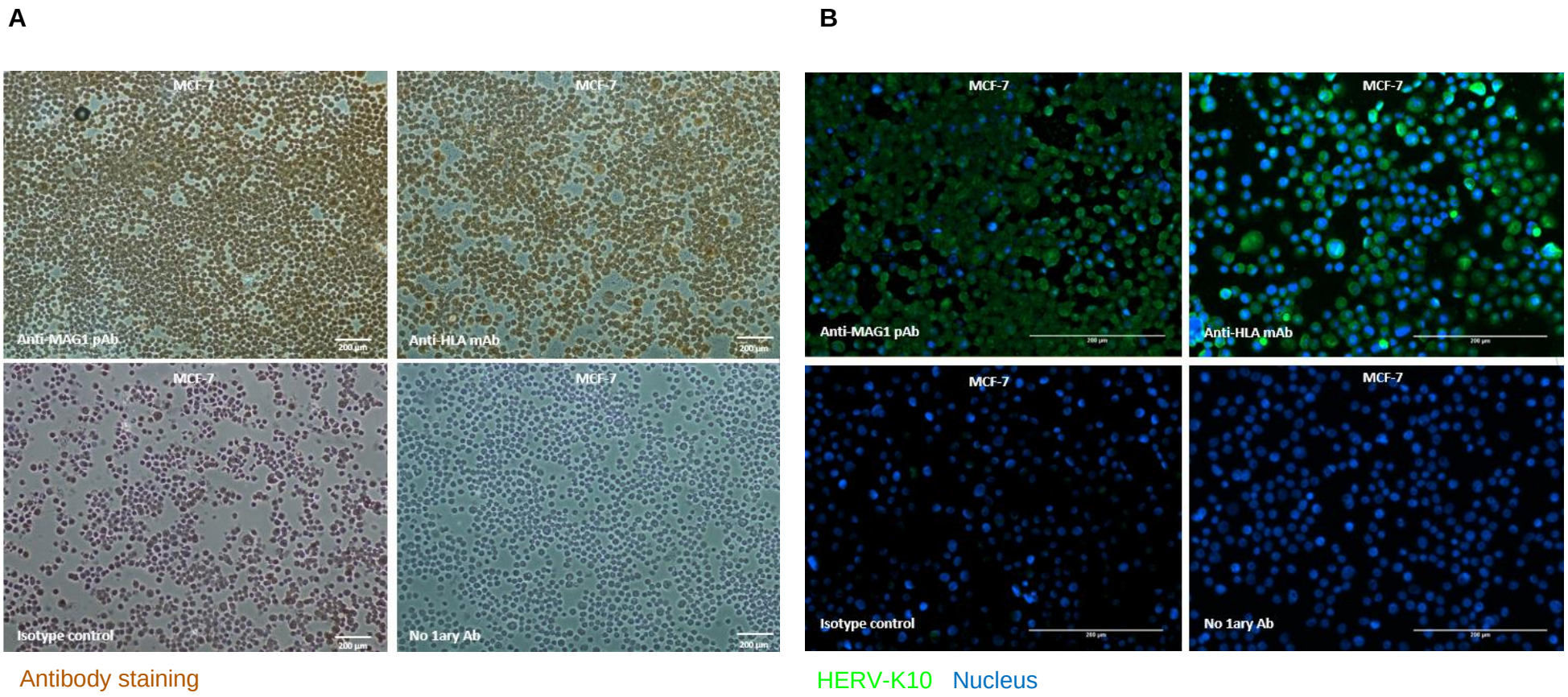


Figure 4.10. Examples of HERV-K10 Gag immunostaining in representative cytospins of MCF-7 breast cancer cell lines. Cells at 10^5 were cytospun and stained with PAbMAG1, anti-HLA positive control antibody, isotype negative control antibody or secondary antibody only. Images show cytoplasmic labelling of MCF-7 breast cancer cells in immunocytochemistry observed as brown staining (A; magnification x20) or as green staining (FITC) with blue DAPI-stained nuclei in immunofluorescence (B; magnification x40). Data are representative of 3 experiments.

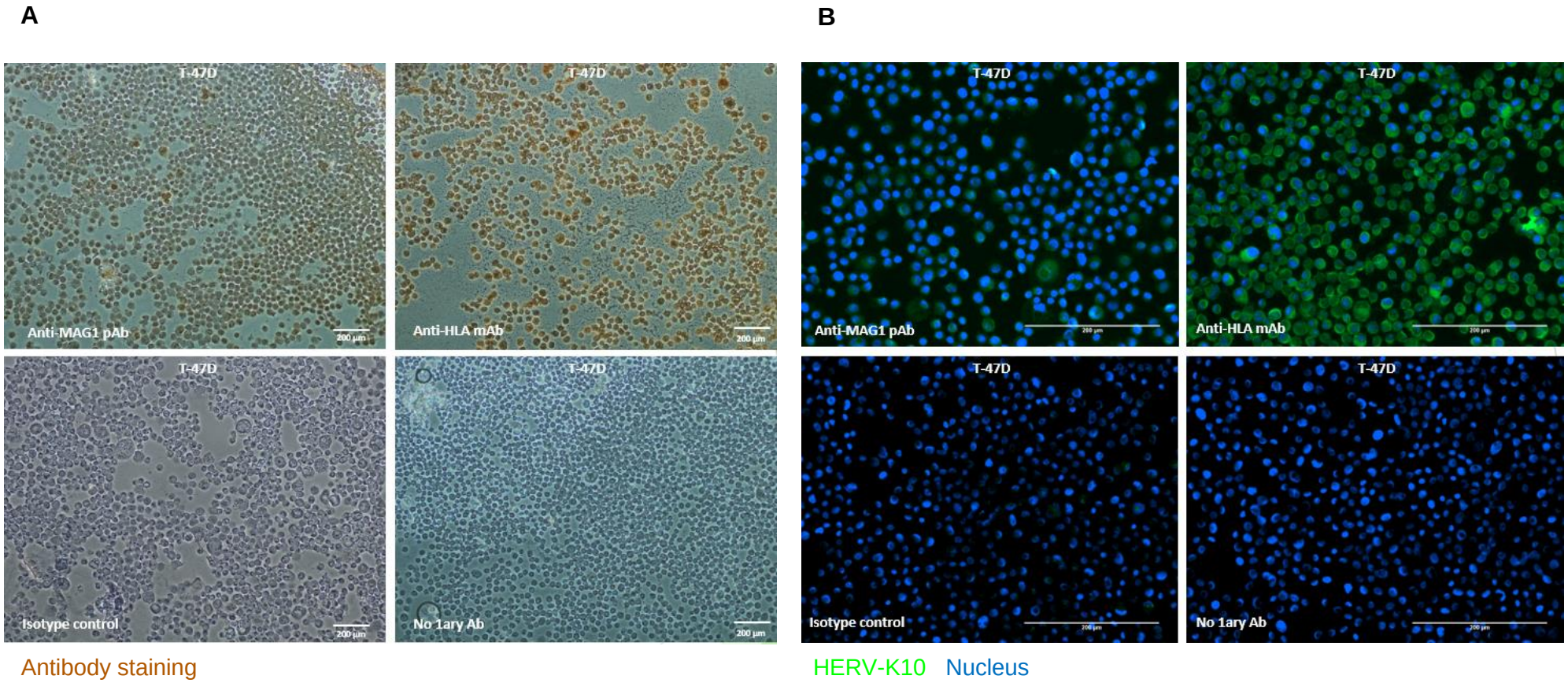


Figure 4.11. Examples of HERV-K10 Gag immunostaining in representative cytospins of T-47D breast cancer cell lines. Cells at 10^5 were cytospun and stained with PAbMAG1, anti-HLA positive control antibody, isotype negative control antibody or secondary antibody only. Images show cytoplasmic labelling of T-47D breast cancer cells in immunocytochemistry observed as brown staining (A; magnification x20) or as green staining (FITC) with blue DAPI-stained nuclei in immunofluorescence (B; magnification x40). Data are representative of 3 experiments.

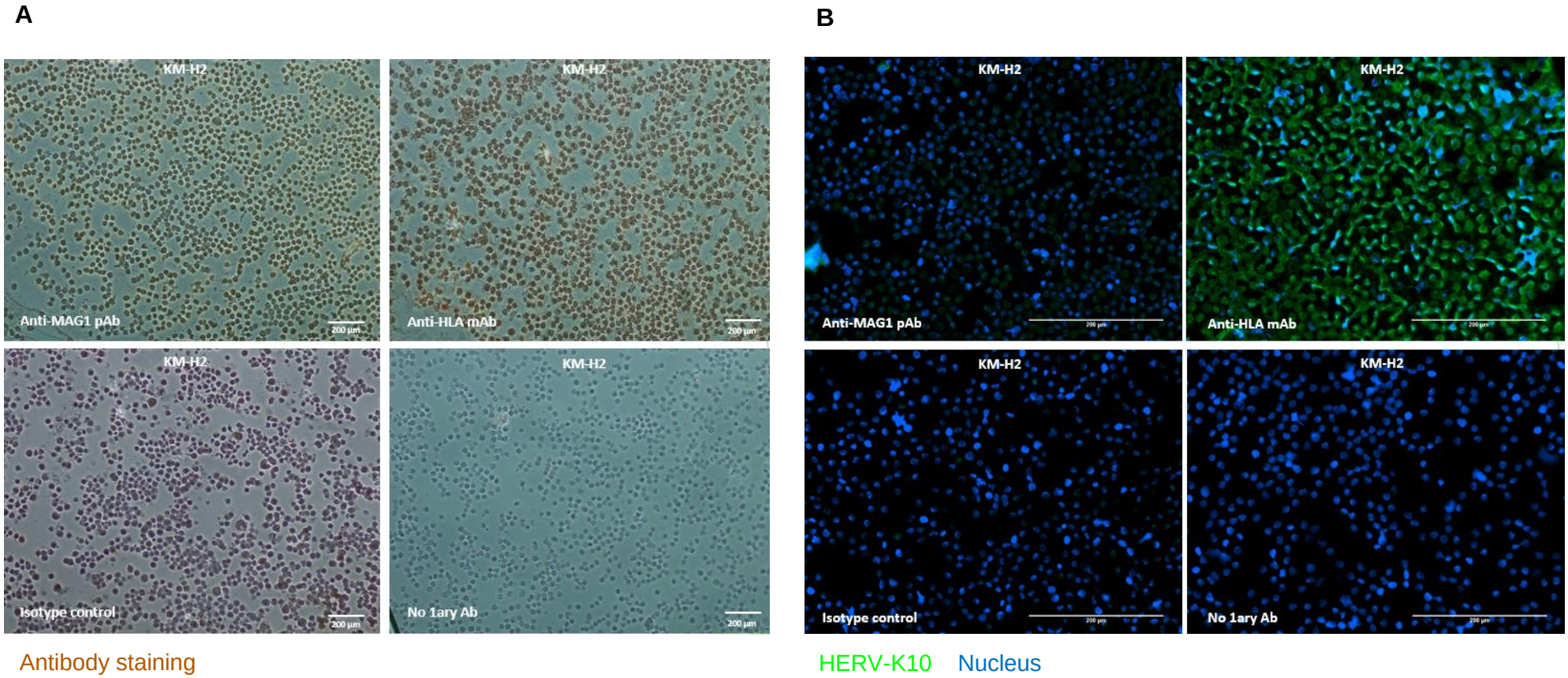


Figure 4.12. Examples of HERV-K10 Gag immunostaining in representative cytospins of KM-H2 Hodgkin lymphoma negative control cell line. Cells at 10^5 were cytospun and stained with PAbMAG1, anti-HLA positive control antibody, isotype negative control antibody or secondary antibody only. Images show cytoplasmic labelling of KM-H2 Hodgkin lymphoma cells in immunocytochemistry observed as brown staining (A; magnification x20) or as green staining (FITC) with blue DAPI-stained nuclei in immunofluorescence (B; magnification x40). Data are representative of 3 experiments.

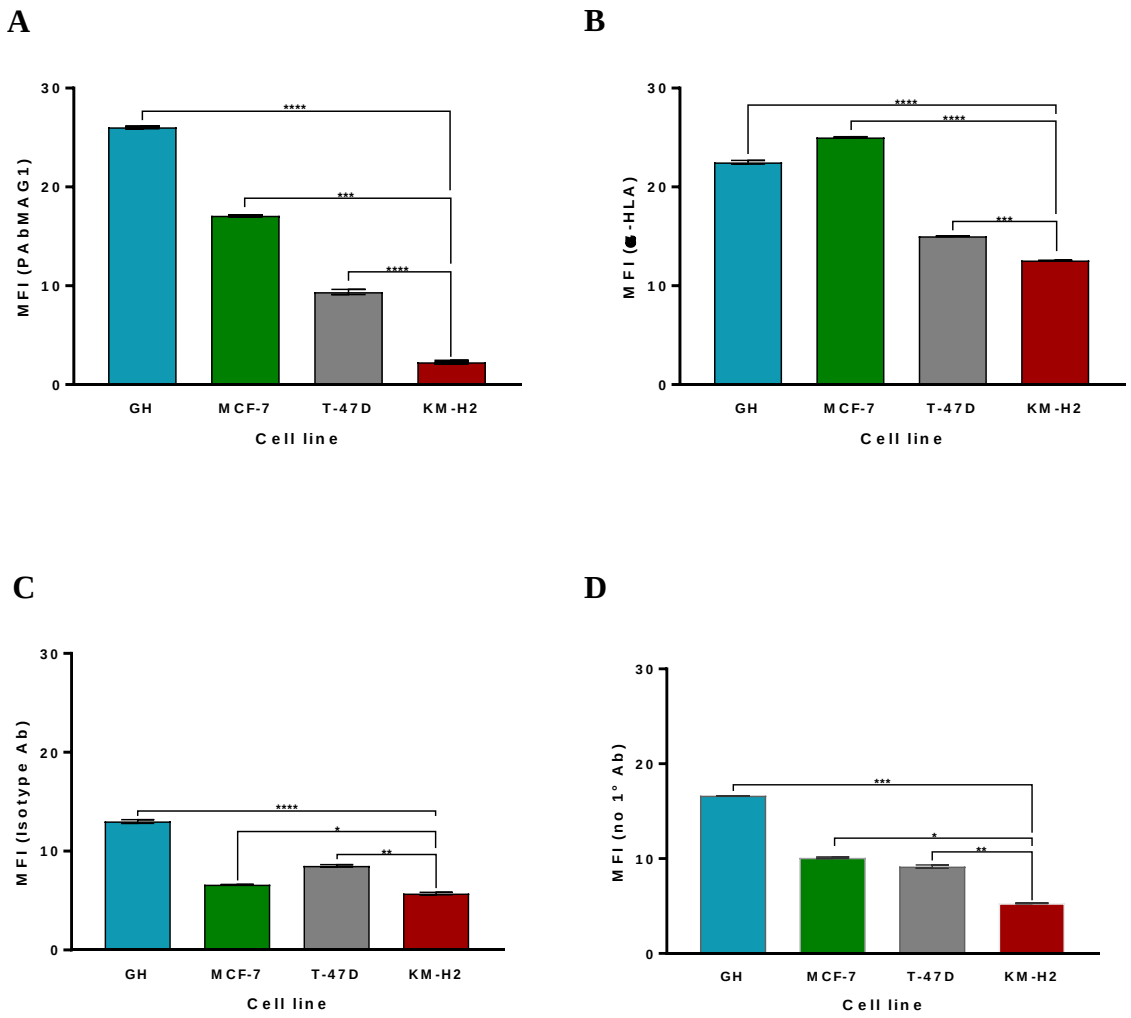


Figure 4.13. Mean fluorescence intensity (MFI) of intracellular staining observed in four malignant cell lines upon incubation with PAbMAG1 and control antibodies. Graph shows MFI values observed for PAbMAG1 (A), anti-HLA antibody (B), isotype control (C) and no primary antibody control (D) used in the immunofluorescence assays (Fig. 4.09-4.12 B). Cell lines used included HERV-K10-harbouring GH teratocarcinoma positive control, MCF-7 and T-47D breast cancer and KM-H2 Hodgkin lymphoma negative control cell line. Bars represent means and SEM is shown for 3 experiments. *, $p \leq 0.05$, **, $p \leq 0.01$, ***, $p \leq 0.001$, ****, $p \leq 0.0001$.

4.2.7. Assessment of PAbMAG1 antibody reactivity to IgG1Fc linear epitopes using western blotting

To ascertain the reactivity to potential linear epitopes within a protein structure, the PAbMAG1 antibody was evaluated by western blotting. Since peptides such as MAG1 are not suitable for electrophoresis due to their small molecular weight, PAbMAG1 was tested against two IgG1 Fc protein preparations, one from a patient (Per) and the other commercially-produced (Abcam). Bovine serum albumin (BSA) served as a negative control protein in this system. Protein electrophoretic transfer to nitrocellulose was validated by the control monoclonal pan-IgG A57H antibody, which bound to both IgG1 Fc proteins, in reduced (band 30-35 kDa) and non-reduced (band 60-65 kDa) fragments as well as some intact IgG (150kDa) (Figure 4.14). PAbMAG1 exhibited strong reactivity, in relation to A57H antibody, against both IgG1 Fc preparations (Figure 4.15) across different protein dilutions, and detected protein bands within non-reduced (60-65 kDa) Fc fragments, validating this antibody for the detection of linear IgG1 Fc epitopes. This antibody also bound to Fc epitopes within some intact IgG proteins from the patient. No reactivity was observed to BSA protein targets.

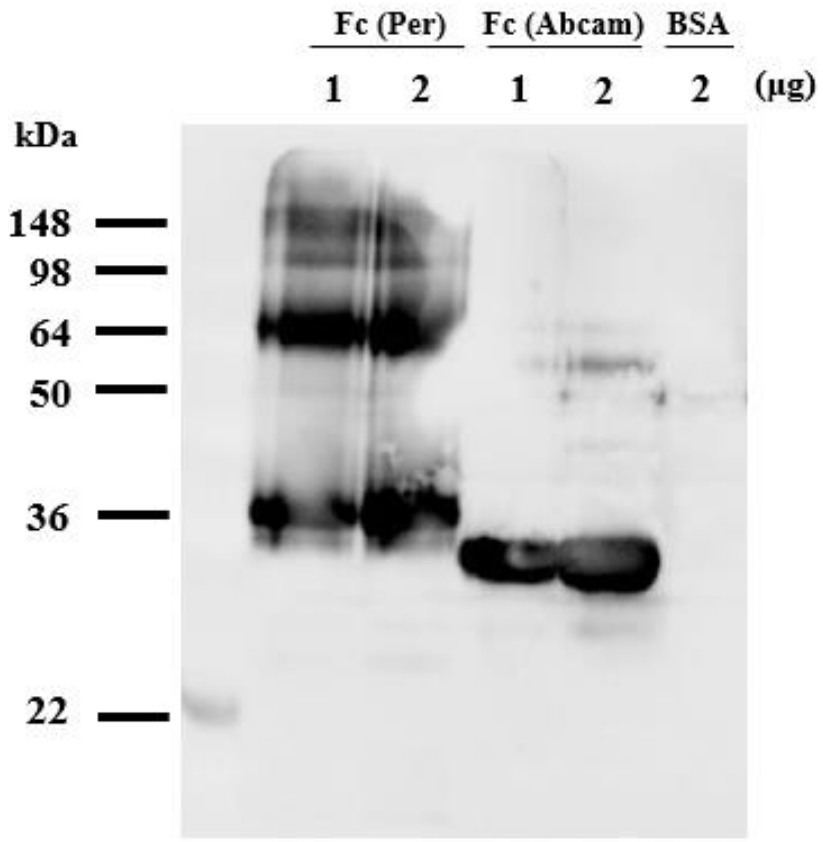


Figure 4.14. Positive reactivity of pan-IgG A57H monoclonal antibody to two IgG1 Fc preparations. IgG1 Fc from a patient (Per) and from a commercial source (Abcam) were analysed by SDS-PAGE and western blotting and probed with A57H antibody at 1 µg/ml. Immunocomplexes were revealed by enhanced chemiluminescence. Bovine serum albumin (BSA) was used as a negative control. Data are representative of 2 experiments.

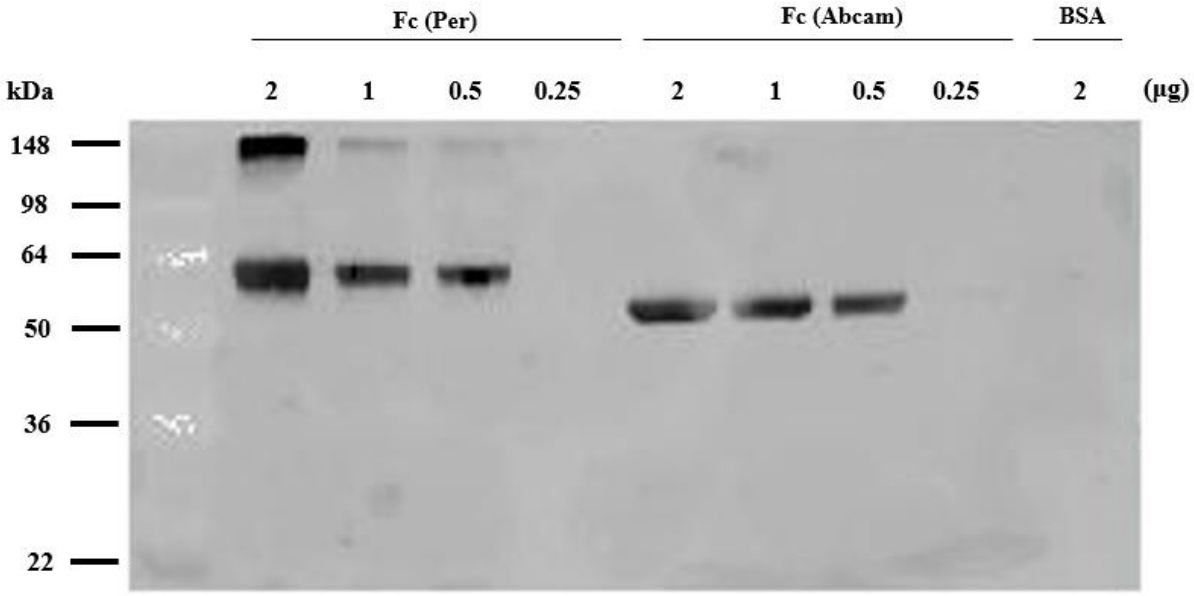


Figure 4.15. Positive reactivity of PAbMAG1 to linear epitopes within IgG1 Fc preparations. IgG1 Fc from a patient (Per) and from a commercial source (Abcam) were analysed by SDS-PAGE and western blotting and probed with PAbMAG1 at 2 µg/ml. Immunocomplexes were revealed by enhanced chemiluminescence. Bovine serum albumin (BSA) was used as a negative control. Data are representative of 2 experiments.

4.2.8. Assessment of HERV-K10 expression in cultured cells using a flow cytometry based *in vitro* assay

To further characterise the PAbMAG1 antibody, we wished to determine its performance in detection of intracellular HERV-K10 proteins by flow cytometry. To this end, we conjugated PAbMAG1 antibody with fluorescein isothiocyanate (FITC) to reduce the levels of potential non-specific binding from a secondary antibody and verified this procedure with an immunofluorescent staining protocol using four malignant cell lines as per Section 4.26 (Figure 4.16). We next assessed the expression of HERV-K10 intracellular proteins with PAbMAG1 at its optimum concentration of 2 µg/ml within the same cell lines by flow cytometry. We used anti-human HLA-ABC (MHC I) as a positive control antibody to ascertain that the staining procedure was successful. To account for background cell staining we incorporated a rabbit IgG FITC-conjugated isotype control and no primary antibody negative control to this analysis. As depicted in Figure 4.17, PAbMAG1 antibody performed poorly in the K10-harbouring GH cell line, showing virtually no staining and lower mean fluorescence intensity (MFI :1439.28) than the isotype (MFI : 16476.19) and no primary antibody (MFI : 3272.85) controls, when compared to the HLA positive control (MFI : 95719.28). This poor performance of PAbMAG1 antibody, resulting most probably from assay restriction phenomenon or cell fixation with acetone, was observed across all four cell lines (Figure 4.18). To ensure that the lack of staining by PAbMAG1 is not a consequence of suboptimal antibody concentration or a prozone effect, we performed a titration of PAbMAG1 and isotype control antibodies in the GH positive control cell line (Figure 4.19). From these data we concluded that unlike the isotype control antibody, PAbMAG1 titrated poorly across all dilutions and therefore does not appear to be suitable for use in intracellular flow cytometry applications.

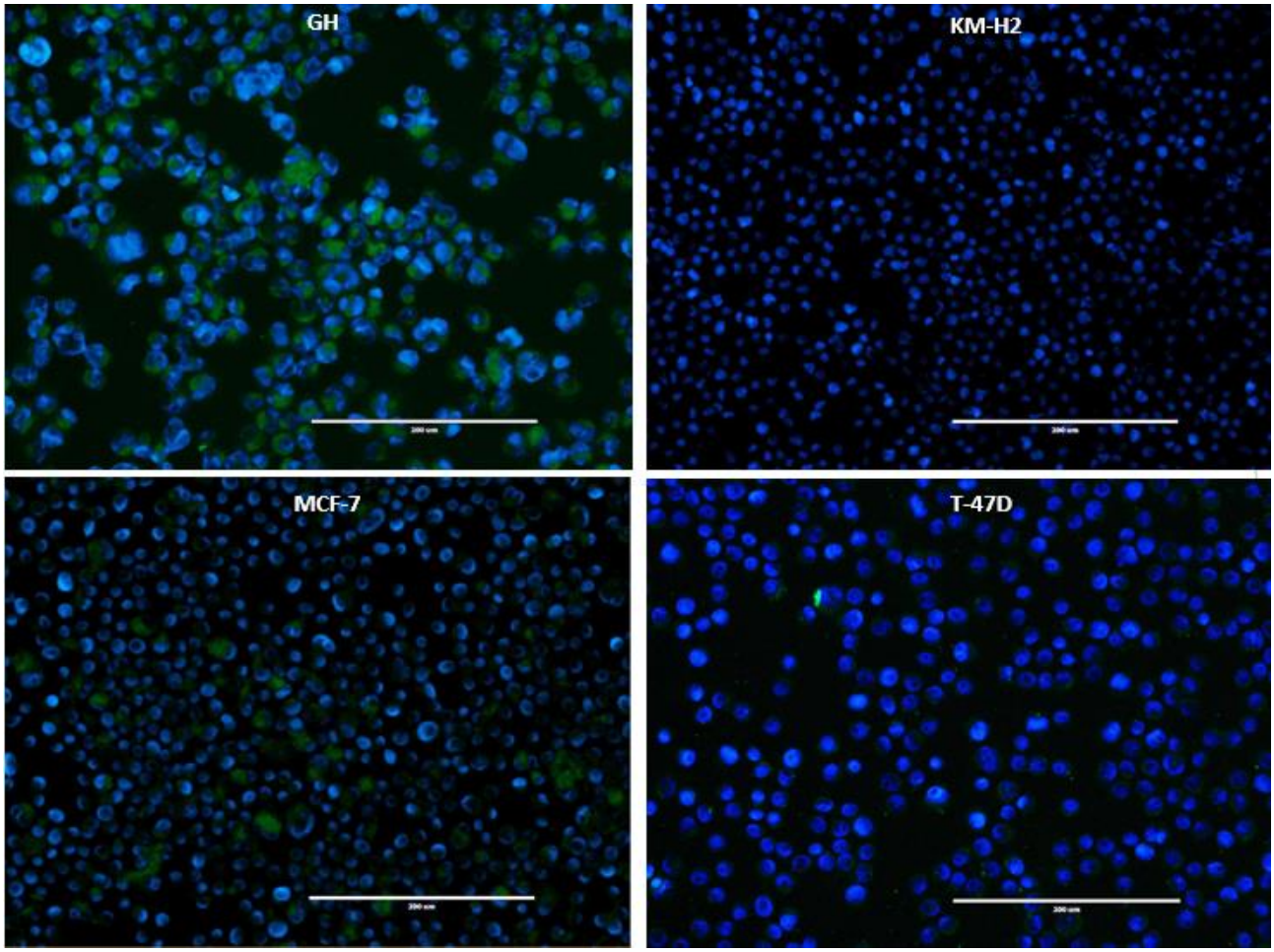


Figure 4.16. HERV-K10 Gag immunostaining of GH, MCF-7, T-47D and KM-H2 cancer cell lines with fluorochrome-tagged PAbMAG1 antibody. Cells at 10^5 were cytospun and stained with in-house FITC-conjugated PAbMAG1 at 2 µg/ml. Fluorescent microscopy images show cytoplasmic labelling of HERV-K10 (green) and blue DAPI-stained nuclei in four cell lines including K10-harbouring GH teratocarcinoma positive control, MCF-7 and T-47D breast cancer and KM-H2 Hodgkin lymphoma negative control cell line. Data are representative of 3 experiments.

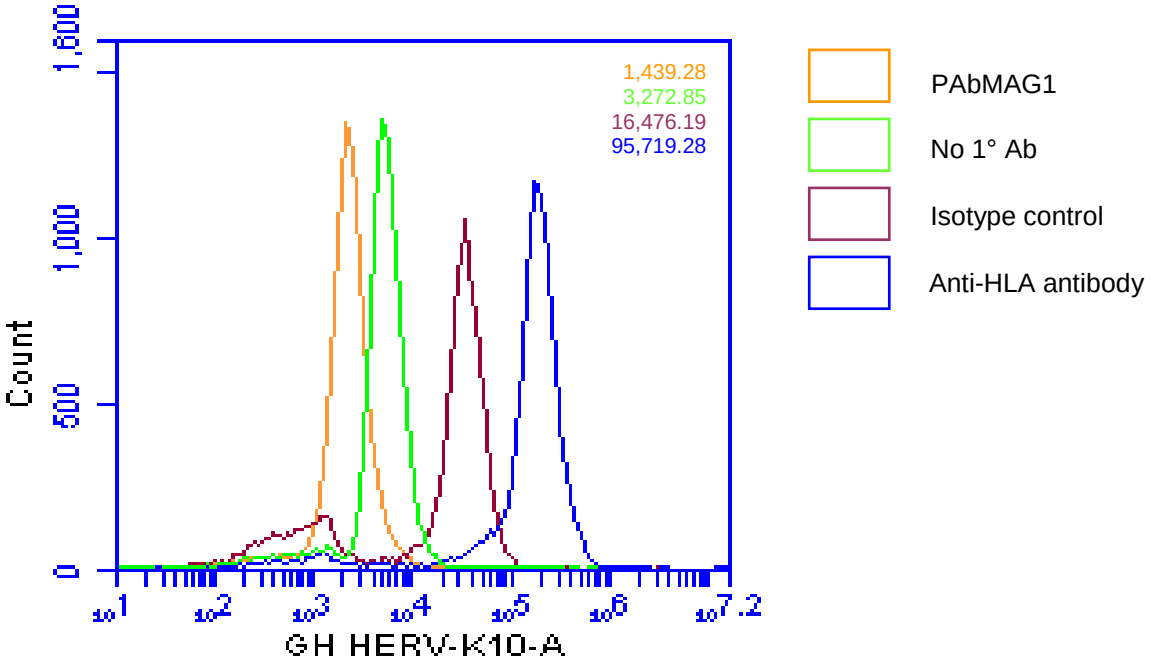


Figure 4.17. Assay restriction profile of PAbMAG1 antibody in K10-harbouring GH teratocarcinoma cell line determined by flow cytometry. 10^6 GH cells were permeabilised and stained with FITC-tagged PAbMAG1 at 25 µg/ml for intracellular HERV-K10. Representative histogram shows mean fluorescence intensity of GH staining with PAbMAG1 (orange, 1439.28) and controls including anti-HLA (blue, 95719.28), isotype Ab control (purple, 16476.19) and no primary Ab control (green, 3272.85). Data are representative of 3 experiments.

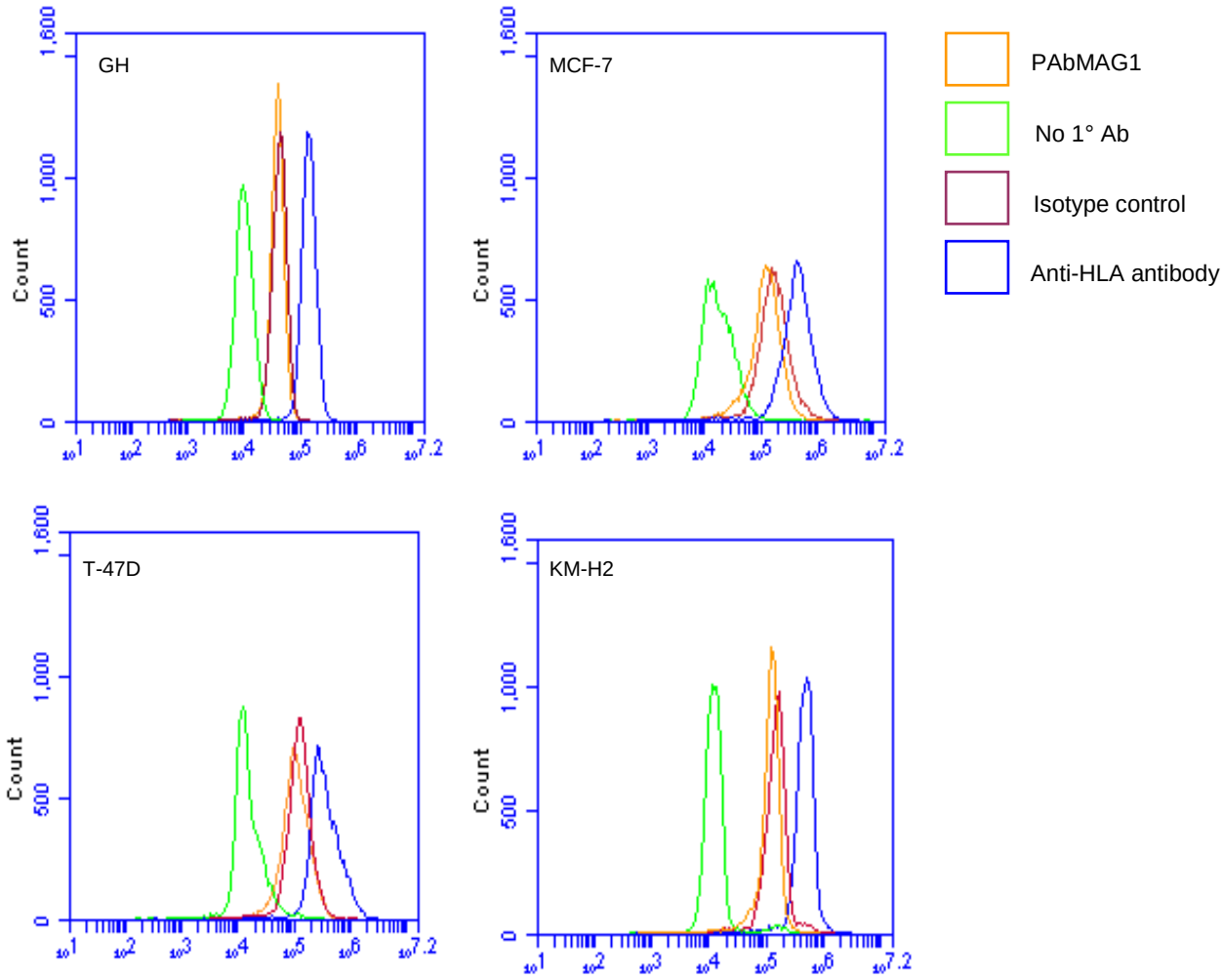


Figure 4.18. Assay restriction profile of PAbMAG1 in four cancer cell lines determined by flow cytometry. 10^6 GH, MCF-7, T-47D and KM-H2 cells were permeabilised and stained with PAbMAG1 at 25 µg/ml for intracellular HERV-K10 and with control antibodies. Representative histograms show PAbMAG1, isotype control and anti-HLA control staining for all four malignant cell lines. Data are representative of 3 experiments.

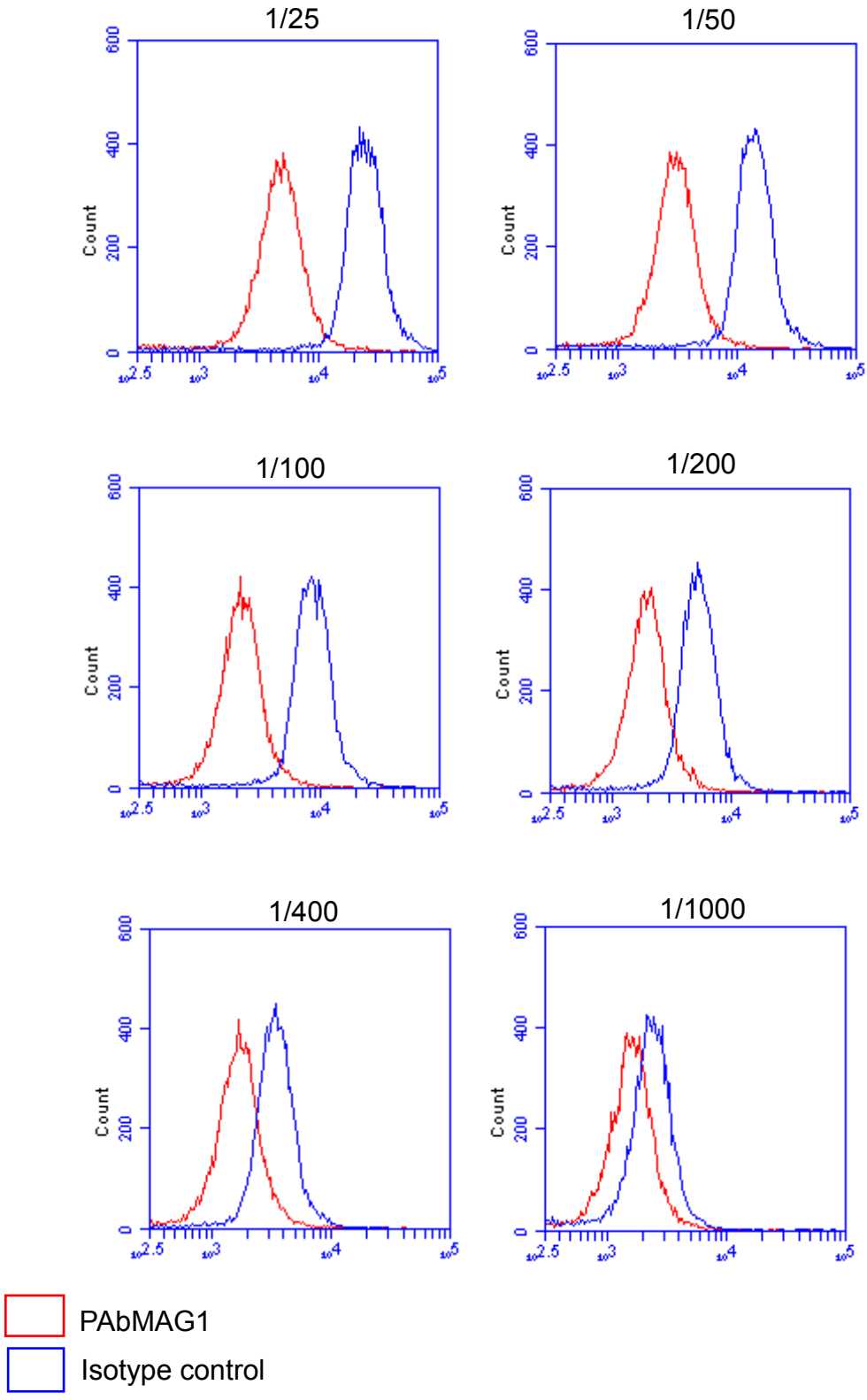


Figure 4.19. Poor antibody titration of PAbMAG1 in K10-harbouring GH teratocarcinoma cell line in flow cytometry. 10^6 GH cells were permeabilised and stained with PAbMAG1 and an IgG isotype control at six dilutions (1/25, 1/50, 1/100, 1/200, 1/400, 1/1000) for analysis by flow cytometry. Data are representative of 3 experiments.

4.3. Discussion

Polyclonal antibodies (PAbs) generated in vitro using synthetic peptides are widely used in biomedical research since they are relatively simple and inexpensive to produce. Moreover, they can recognise an antigen from different orientations, therefore the screening and characterisation of PAbs is vital for the development of sensitive immunoassay systems. There are currently no commercially available antibody reagents to HERV-K10 Gag, which has been implicated as a possible trigger in a number of diseases including rheumatoid arthritis (Nelson *et al*., 2014a). Evidently the availability of standardised antibody reagents to this viral region, with determined strength of reactivity and potential restriction profiles in certain immunoassays, would find useful applications in both a clinical and research context.

In this chapter, a polyclonal antibody was generated against a highly antigenic synthetic peptide epitope MAG1 within the viral matrix, to obtain a tool for HERV-K10 Gag protein detection in different assay applications. Epitopes are typically categorised into two groups, linear epitopes composed of sequential residues and conformational epitopes created by protein folding to somewhat resemble the native structure (Sivalingam and Shepherd, 2012). The latter are generally used for applications which rely on the protein natural structural context such as flow cytometry or therapeutic strategies, whereas the former are preferred in protein-denaturing assays such as western blotting or immunocytochemistry/ immunofluorescence (Forsström *et al*., 2015). Traditional protein immunization protocols involve synthetic peptides which act as immunogens (Brown *et al*., 2011). Therefore, we have used a 15-mer MAG1 linear peptide conjugated to KLH carrier, which improves antibody affinity by inducing isotype switching and CDR maturation (Patra *et al*., 2014), to produce a high affinity polyclonal antibody (PAbMAG1) to the HERV-K10 Gag region in a rabbit (Figure 4.01). High levels of reactivity were observed by ELISA for PAbMAG1 with its specific linear target

but not with the other MAG peptides, which confirmed antibody specificity. This finding is in agreement with the observations made by Forsstrom *et al*. (2015) who reported that antibodies generated to a unique peptide target exhibit less cross-reactivity than those developed using a full-length protein. Since the size of epitopes is highly diverse with linear stretches of 3-15 amino acids on average (Kringelum *et al*., 2013), we wished to determine how the number of epitope residues affects the PAbMAG1 binding in ELISA. To this end, we synthesised a panel of MAG1 peptides of different lengths (3-, 4-, 5-, 7-, 9-, 12-, and 15-mer), each containing some or all immunodominant GKELK residues. We observed that the levels of PAbMAG1 reactivity diminished with decreasing length of MAG1 peptide (Figure 4.03) with best binding reported for 15-, 12- and 9-mers and negligible binding for shorter peptides. This finding suggests that a minimum set of supporting residues is required for effective antigen-antibody binding. As reported by other groups (Stave and Lindpaintner, 2013), it appears that when interacting molecules possess complementary structures the strength of their interaction is more pronounced. This might be affected by the neighbouring residues as certain amino acids such as serine or glycine are more neutral and flexible than others (e.g. proline) which might be strongly charged or structurally rigid (Turunen *et al*., 2015; Juarez *et al*., 2016). Moreover, we observed that the efficiency of epitope presentation drastically improved when biotinylated peptides were used as coating antigens in complex with neutravidin in the ELISA system (Figure 4.04), when compared to their non-biotinylated counterparts. This was particularly noticeable for shorter peptides (7-mer and less; Figure 4.05) and suggested that the target epitopes were more favourably presented and more accessible on the microtitre plate. In this case, structured orientation of a peptide is facilitated by a neutravidin-biotin bridge which prevents the peptide from binding to the plastic plate in a random fashion (via hydrophobic interactions) (Crocker and Murray, 2003). This is in agreement with

observations made by Welch *et al*. (2017) who reported a two-fold improvement in antibody reactivity with the use of biotin linker. Interestingly, PAbMAG1 showed increased reactivity to 5-mer MAG1 peptide when compared to 7-mer peptide, which suggests that the epitope accessibility may have been sterically compromised in the 7-mer by non-polar isoleucine and neutral glutamine residues on either side of the epitope.

Finally, our interest in molecular mimicry between HERV-K10 Gag and IgG1 Fc led us to also examine immunological cross-reactivity of PAbMAG1 to IgG1 Fc. In using our ELISA system, we demonstrated the antibody specificity for MAG1 peptide and IgG1 Fc epitopes (Figure 4.07), which supports the premise of molecular mimicry between the viral and host proteins in RA. Variation in OD values between the protein and peptide preparations results most probably from the nature of epitope mimics with conformational and linear segments recognised by the antibody (Ponomarenko and Van Regenmortel, 2009). Since antibody efficacy in one assay method is not typically predictive of its success in another (Brown *et al*., 2011), PAbMAG1 was tested for reactivity in western blotting where it proved positive against accessible epitopes on IgG1 Fc. In this technique, proteins are generally rendered linear from conformational states (Westwood and Hay, 2001) and hence our antibody PAbMAG1, appeared to cross-react with continuous epitopes on IgG1 Fc preparations (Figure 4.16), with comparable strength of reactivity to that observed for pan-IgG A57H monoclonal antibody (Figure 4.15).

Immunoreactivity of PAbMAG1 to conformational HERV-K10 targets was also assessed in cell preparations using immunocytochemistry (ICC), immunofluorescence (IF) and flow cytometry in four cancer cell lines. In ICC and IF, PAbMAG1 exhibited reactivity to intracellular viral antigens within the cytoplasm of GH, MCF-7 and T-47D cell lines (Figure 4.10 to 4.13), which were previously shown to harbour HERV-K viral

products (Ejtehadi *et al*., 2005; Löwer *et al*., 1993). The observed cytoplasmic viral protein distribution is in agreement with published Gag localisation (Stake & Bann 2013, Chiang and Chang 2005). The lack of staining reported in KM-H2 negative control cell line (Figure 4.13), was also previously reported (Freimanis *et al*., 2010).

PAbMAG1 showed a clear assay restriction profile in flow cytometry and failed to recognise intracellular viral targets within all three HERV-K-harbouring cell lines (Figures 4.18 to 4.20). This phenomenon may have resulted from the manner in which the target antigenic determinants were presented, modified or masked. In flow cytometry, target epitopes are presented in their native state therefore antibodies generated to a short peptide may be unable to bind to such epitopes as a result of their unstructured features (Forsström *et al*., 2015). In addition, it has been reported that certain procedures within a flow cytometry protocol, such as acetone fixation as used in this investigation, can perturb the antigen conformation and mask or denature the target epitopes thus significantly impairing antibody reactivity (Brown *et al*., 2011; Bull *et al*., 1999). Consequently, the target epitopes in all four cell lines may not be accessible or might be structurally compromised. Moreover, since a 15-mer peptide was selected for immunisation, it is possible that the number of determinants on the native protein identified by PAbMAG1 was restricted. Evidently, this could have affected the antibody performance in this system and resulted in the assay restriction phenomenon revealed in our investigation. Interestingly, the polyclonal antibodies to MAG1 were effective in ELISA where epitopes are presented in a relatively native conformation, as well as in ICC, IF and western blotting against epitopes in various states of denaturation (Brown *et al*., 2011), suggesting that the target epitopes were favourably displayed and accessible, which was previously observed for other antibodies developed using synthetic peptides (Waldron *et al*., 2002). Accordingly, other groups have reported that antibodies generated to linear targets exhibit more

effective binding in immunoassays with either completely or partially denatured epitopes (Stave and Lindpaintner, 2013; Forsström *et al*., 2015).

Since there is a paucity of anti-HERV-K10 reagents, the novel PAbMAG1 generated against Gag protein sequence and assessed for reactivity in a range of immunoassays could serve as an effective tool for assessment of K10 Gag protein expression. For instance, HERV-K10 Gag proteins are synthesised in testicular tumours, in particular in seminoma cells, therefore PAbMAG1 antibodies may be suited for screening of tumour biopsies of those patients in ICC and IF (Sauter *et al*., 1995). Therefore, this approach could help to further explore the role of HERVs in this cancer. Moreover, it could be used for the detection of Gag protein levels in choriocarcinomas, teratomas and other germ cell tumours, which have been shown to possess increased K10 levels (Flockerzi *et al*., 2008). Analysis of tumour tissues for HERV-K10 expression with this antibody could also be performed by immunoblotting since Gag polyprotein renders an 80 kDa product (Sauter *et al*., 1995). This would considerably add to the repertoire of non-invasive screening methods for such cancers.

In addition, PAbMAG1 could be employed in research applications for the detection of cytoplasmic viral particles in K10-producing cell lines such as Tera 1 (Nelson *et al*., 2014a) and GH teratocarcinoma cell lines (Patience *et al*., 1998). Since K10 Gag protein has been previously detected in biological samples such as peripheral blood from patients with essential thrombocythemia (Morgan and Brodsky, 2004) and in synovial fluid from RA patients (Freimanis *et al*., 2010), PAbMAG1 could be used to capture the Gag proteins from such samples in ELISA for the assessment of viral expression. Finally, PAbMAG1 could serve to establish the principle for the development of a monoclonal anti-K10 Gag reagent of high specificity which could be considered as a therapeutic tool for conditions where HERV-K10 may be implicated in the disease pathogenesis. It could be used analogously to the bacterial exotoxin-

neutralising monoclonal antibodies raxibacumab, obiltoxaximab and bezlotoxumab (Wang-Lin *et al*., 2018) as well as the anti-HERV-W monoclonal antibody (GNbAC1) used as a potential treatment for MS (Madeira *et al*., 2016).

In this chapter, on developing a polyclonal antibody (PAbMAG1) to a synthetic peptide of HERV-K10 containing the epitope GKELK, we established antibody reactivity profiles in a range of immunoassay systems and demonstrated immunological cross-reactivity to IgG1 Fc. This observation clearly substantiates the premise of molecular mimicry between HERV-K10 and IgG1 Fc, and an association with RF binding epitopes in RA (Nelson *et al*., 2014a; Westwood *et al*., 2008). Consequently, this work prompted us to assess serological reactivity of RA patient sera to the immunodominant epitope in the Gag region of HERV-K10 and to investigate potential molecular mimicry with a key autoantigen in RA, IgG1 Fc.

5. SEROLOGICAL INVESTIGATION OF RA PATIENT SAMPLES FOR REACTIVITY TO HERV-K10

5.1. Introduction

The aetiology of RA remains elusive, although a genetic predisposition, together with other triggers including viral agents are considered to be contributory factors. More recently, human endogenous retrovirus HERV-K10 has been implicated as a potential aetiological agent within RA and has been linked to producing viral particles in RA synovial tissue (Nelson *et al*., 2014a; Takeuchi *et al*., 1995; Stransky *et al*., 1993). It has been suggested that the HERV-K10 Gag region could play a pivotal role as a viral trigger of RA or could augment disease pathogenesis through mechanisms such as molecular mimicry, epitope spreading and provision of superantigen motifs. The premise of molecular mimicry implies structural homology between the microbial agent and human proteins (Tugnet *et al*., 2013; Balada, Vilardell-Tarrés and Ordi-Ros, 2010). As a consequence, an immune response to a virus could also induce undesired B and T cell reactivity to a host protein. In effect, an inflammatory response with ensuing collateral damage would occur. In chapter 3 we used *in silico* analysis and molecular modelling to demonstrate key epitopes of HERV-K10 Gag that may share identical or similar epitopes on a key RA autoantigen, IgG1 Fc (Trela, Nelson and Rylance, 2016). Moreover, we used a polyclonal antibody to one potential HERV-K10 Gag / IgG1 Fc mimic (MAG1) to validate our finding (chapter 4) (Nelson *et al*., 2014a). A functional Gag region is a pre-requisite for viral particle formation. The complete translated Gag of HERV-K10 stems from two overlapping reading frames encoding matrix, and capsid plus nucleocapsid segments, to produce a precursor protein which is cleaved to generate individual protein products. These regions are essential in the assembly of

viral particles and exhibit increased levels of transcriptional activity in RA patients (Löwer *et al*., 1993; Ono *et al*., 1986; Freimanis *et al*., 2010). The matrix lies directly underneath the lipid envelope which mainly provides an outer viral protective coat and facilitates the attachment of envelope glycoproteins. Following degradation of the lipid envelope, the matrix would be readily exposed to the adaptive and innate immune systems. RA is considered to be antigen-driven and thus it is of interest to establish whether the class of generated antibody represents a secondary immune response. Previous studies from our group have demonstrated a general antibody response (IgM) to HERV-K in RA patients as compared to controls, which may suggest a continual supply of antigenic / immunogenic viral products to the immune system (Freimanis *et al*., 2010). More recently, we noted significantly higher antibody reactivity to HERV-K10 Gag in RA as compared to inflammatory, non-inflammatory and healthy controls (Nelson *et al*., 2014a). Consequently, a fragment of the Gag matrix might supply a reservoir of antigenic peptides which could initiate and perpetuate a secondary (IgG) immune response to this endogenous retrovirus. In this chapter, we therefore set out to investigate reactivity of IgG antibodies in RA patients to a highly antigenic matrix segment of HERV-K10, MAG1 region.

5.2. Results

5.2.1. Validation of reactivity to MAG1 in RA

We have previously demonstrated the reactivity of purified IgG anti-MAG1 (PAbMAG1) antibodies from an immunised rabbit to both HERV-K10 and IgG1 Fc in a number of immunoassays. We therefore set out to assess mimicry of viral and host proteins in RA by testing the reactivity of IgG antibodies obtained from RA patients, with specificity analogous to or equal to that of RFs, to MAG1 antigenic peptide. To this end, we purified two RA serum samples for IgG fraction via affinity chromatography and after protein quantification, we titrated the purified serum samples against biotinylated MAG1 peptide in ELISA. As presented in Figure 5.01, both RA patients exhibited strong anti-MAG1 reactivity which was comparable for both samples and suggested that RF or RF-like autoantibody in patient sera recognised the K10 viral mimic. Moreover, subsequent titration of purified PAbMAG1 against its target MAG1 peptide revealed a striking parallelism in titration profiles between serum antibodies from two RA patients and PAbMAG1 over a wide range of concentrations suggesting that an identical or highly similar epitope on MAG1 was recognised in all three cases. Therefore, our finding prompted us to investigate the anti-MAG1 IgG reactivity with a higher number of samples in a further six RA patients. We also set out to verify the effect of MAG1 biotinylation on the reactivity of antibodies from purified RA sera since peptide biotinylation imparted beneficial effect on PAbMAG1 reactivity (chapter 4). As demonstrated in Figure 5.02, all RA serum samples exhibited very low levels of activity to non-biotinylated MAG1 peptide, which was comparable to that for two negative control peptides. In this assay, PAbMAG1 exhibited high levels of reactivity to its non-biotinylated target peptide, as shown in chapter 4, and therefore acted as a positive control. We observed marked increase in the reactivity of patient autoantibodies to biotinylated MAG1 peptide (Figure 5.03), when compared to its non-biotinylated

counterpart, which suggested improved steric epitope orientation in this system and was therefore selected for subsequent serological analysis. Unsurprisingly, PAbMAG1 detected its target peptide in both systems (max. OD = 2), as demonstrated before (chapter 4, Figure 4.03 and 4.04) and served as a control antibody.

Since RA serum samples are clinically assayed for the disease markers in their crude form (without affinity purification), we therefore set out to analyse the level of reactivity to biotinylated MAG1 peptide from crude preparations of RA sera. The antibody reactivity was tested for six RA patients with different RF levels and three healthy controls, which were resuspended in our standard blocking buffer (2% BSA / PBS) at 1:200. We observed poor differentiation between patient serum samples and healthy controls (Figure 5.04 A) as well as considerable non-specific reactivity to the negative control (NC1GTY3) peptide (Figure 5.04 B). In order to improve the sensitivity of our assay, we mitigated the issue of serum matrix effect, arising from a series of matrix components in complex biological samples that cause interference, with the use of different blocking agent (1% Hammarsten-grade casein / PBS). This blocking buffer is devoid of carbohydrate and fatty acids and therefore reduces the chances of cross-reactivity with serum matrix components as previously shown by Sigdel *et al*. (2012). As demonstrated in Figure 5.05 A and Figure 5.05 B, we found a significant decrease in non-specific reactivity across all the serum samples to MAG1 peptide and NC1GTY3 control, respectively, which confirmed the effectiveness of this blocking agent and was therefore used in subsequent serological assays. This is in accordance with previous reports demonstrating that serum non-specific binding can be rectified with effective blocking agent (Rosenberg-Hasson *et al*., 2014; Güven *et al*., 2014).

Interestingly, upon establishing an effective protocol, we observed increased reactivity of RF positive sera from RA patients to MAG1 peptide when compared to RF negative sera and healthy controls (Figure 5.05 A). This finding prompted us to investigate in greater detail the possible correlation between RF autoantibody levels and anti-MAG1 reactivity in section 5.2.2.

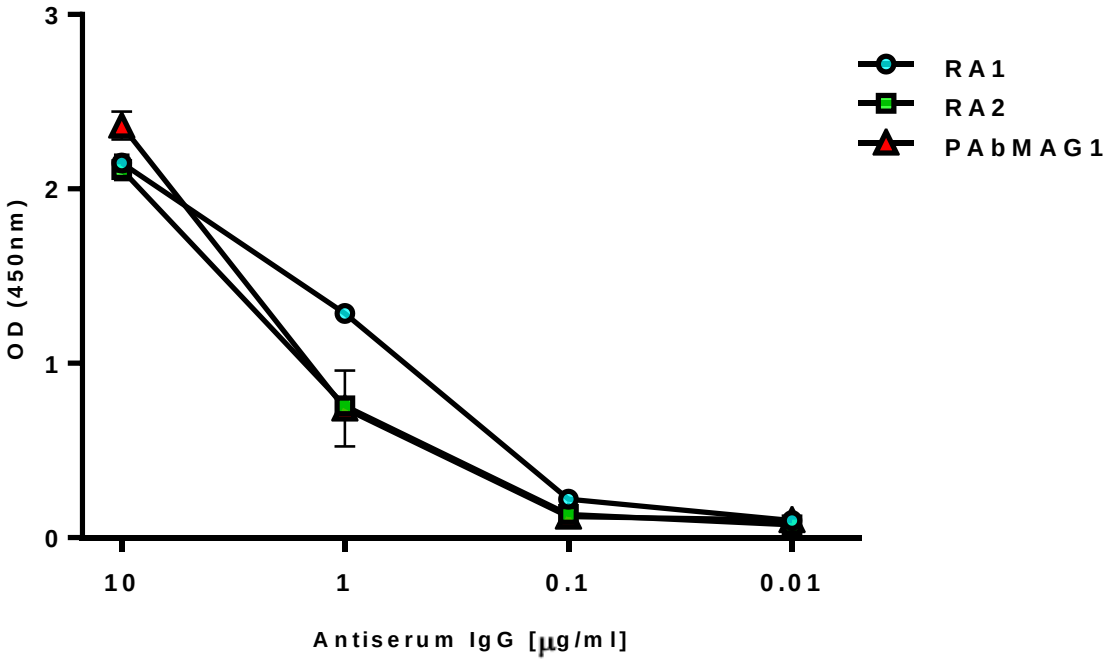


Figure 5.01. Parallelism of titration profiles of IgG from two RA patients and PAbMAG1 antibody to biotinylated MAG1 peptide. Purified IgG fraction from RA patient serum samples and PAbMAG1 from rabbit sera were titrated against MAG1 coating antigen (8 µg/ml) in biotinylated ELISA system. SEM is shown for 3 experiments.

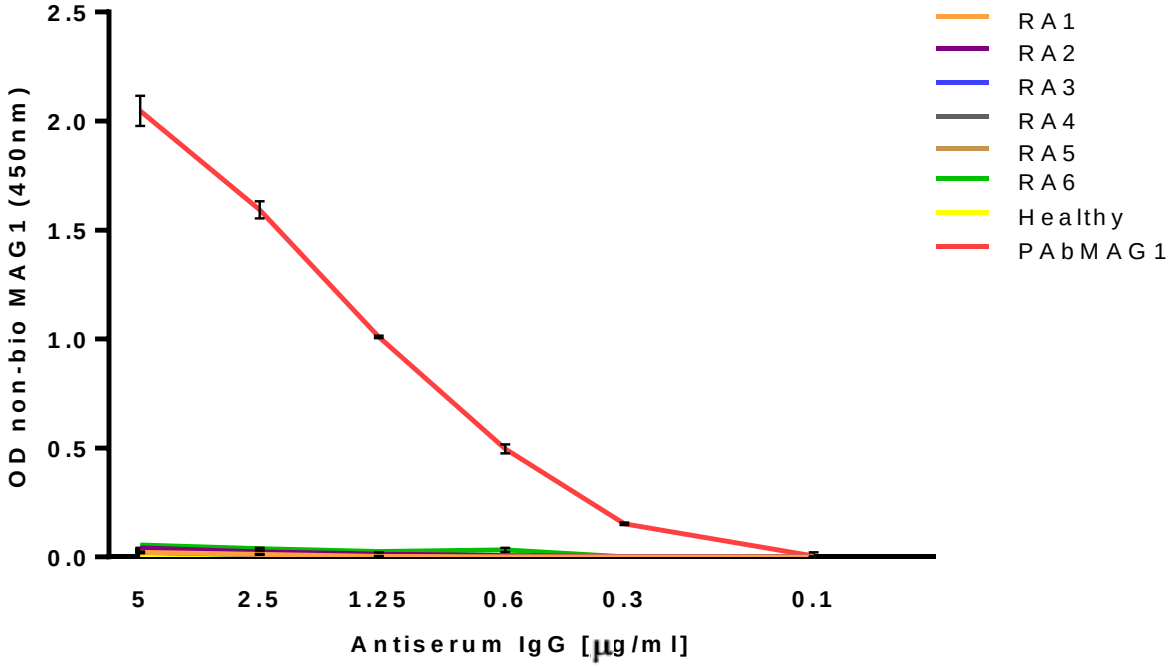


Figure 5.02. Poor titration profiles of IgG from six patients with RA to non-biotinylated MAG1 peptide. Purified IgG fraction from RA patient and control serum samples were titrated against MAG1 coating antigen (8 µg/ml) in non-biotinylated ELISA system. SEM is shown for 3 experiments.

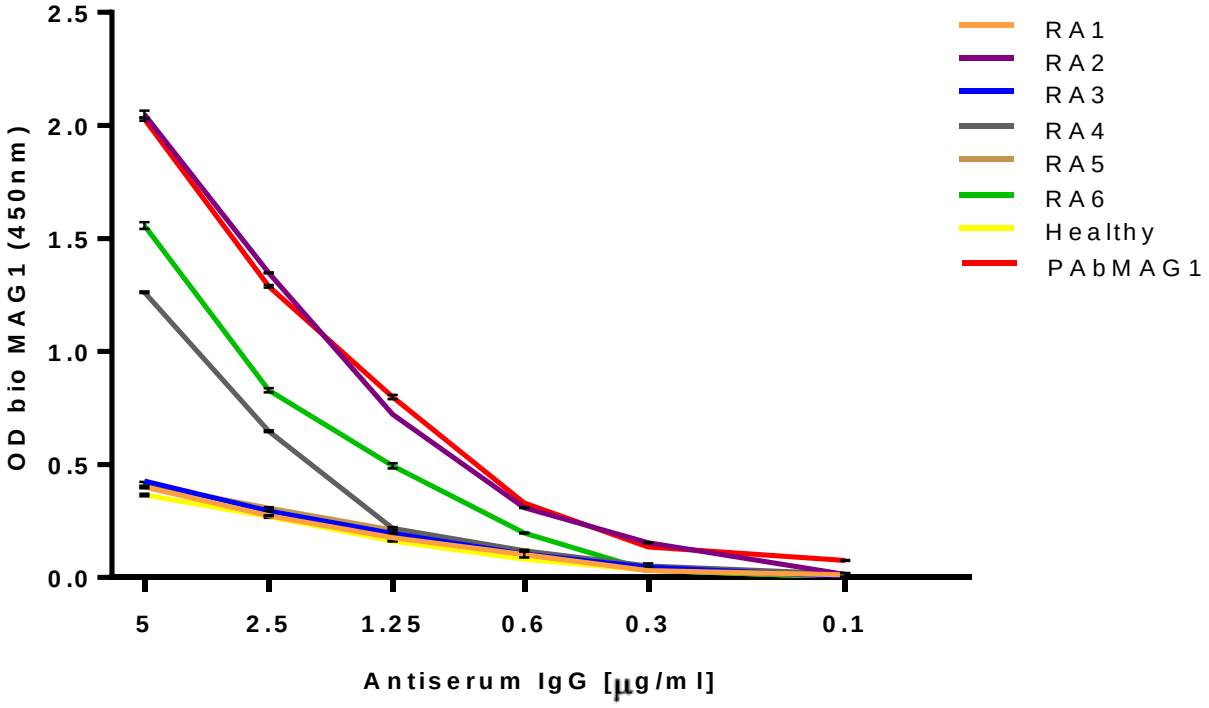


Figure 5.03. Enhanced titration profiles of IgG from six patients with RA to biotinylated MAG1 peptide. Purified IgG fraction from RA patient and control serum samples were titrated against MAG1 coating antigen (8 µg/ml) in biotinylated ELISA system. SEM is shown for 3 experiments.

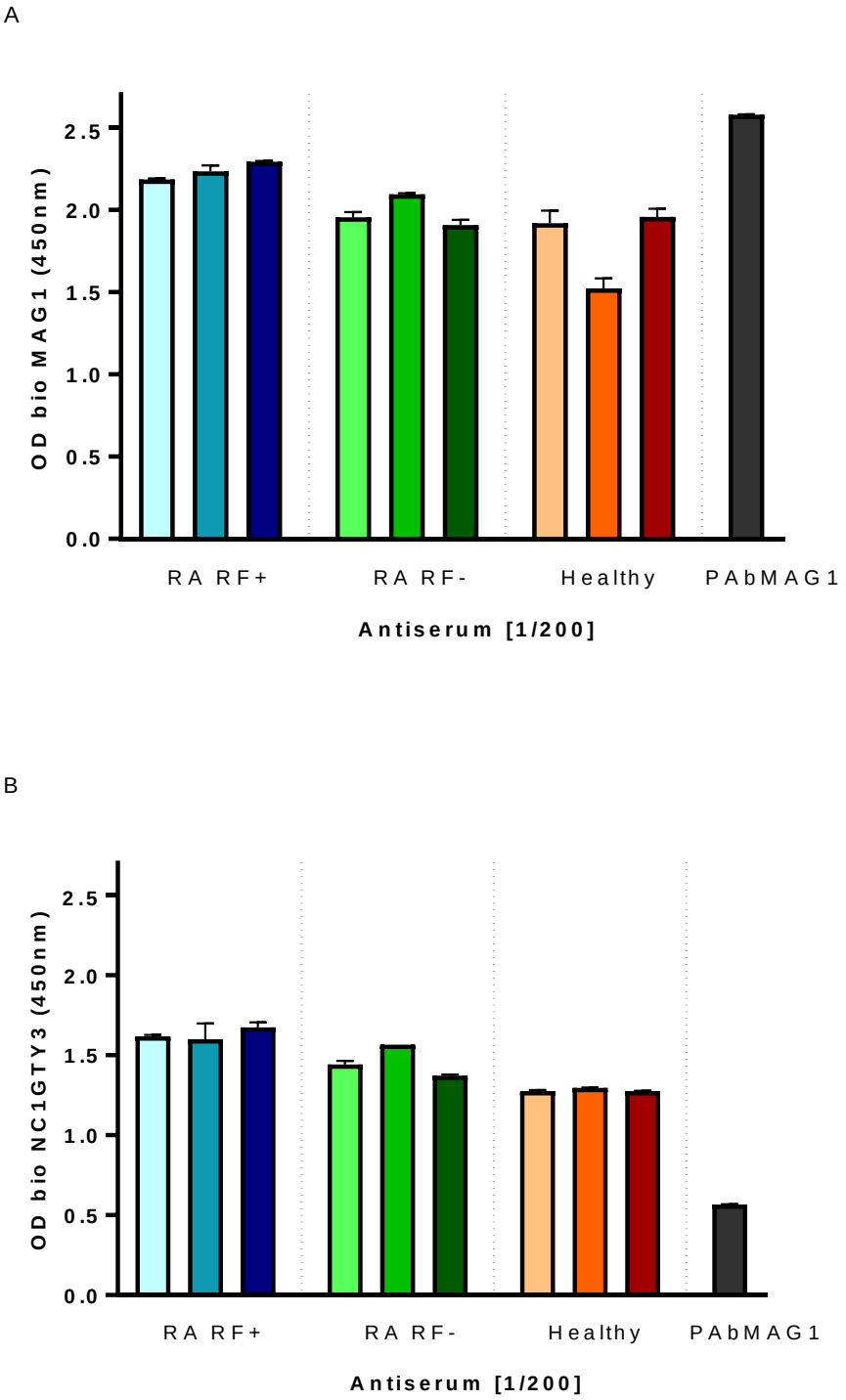
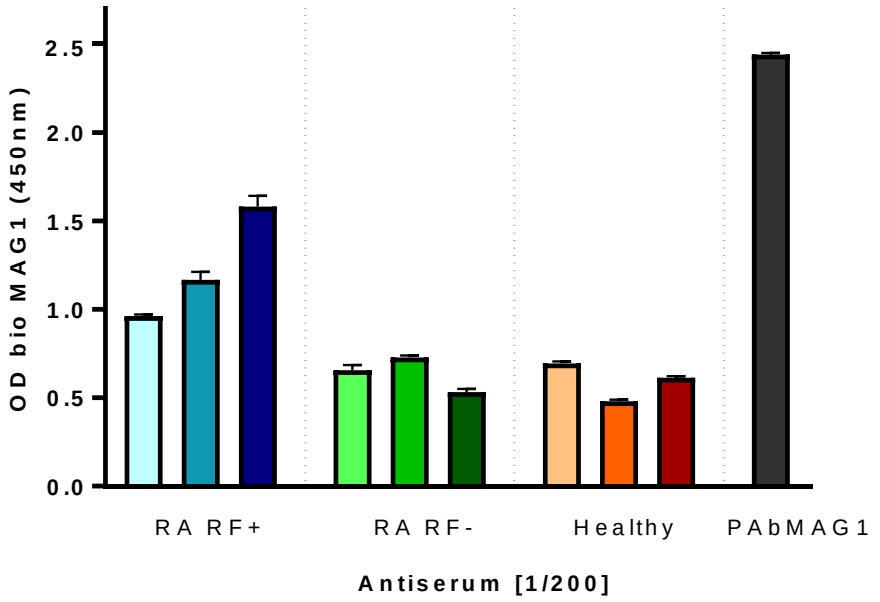


Figure 5.04. High levels of non-specific binding observed for patient serum diluted in BSA blocker. Purified IgG fractions from RF seropositive and RF seronegative RA patients and healthy control serum samples were probed against MAG1 (A) and NC1GTY3 coating antigens (B) (both at 8 µg/ml) in biotinylated ELISA system. Bars represent individual patients and SEM is shown for 3 experiments. RA - rheumatoid arthritis, RF - rheumatoid factor.

A



B

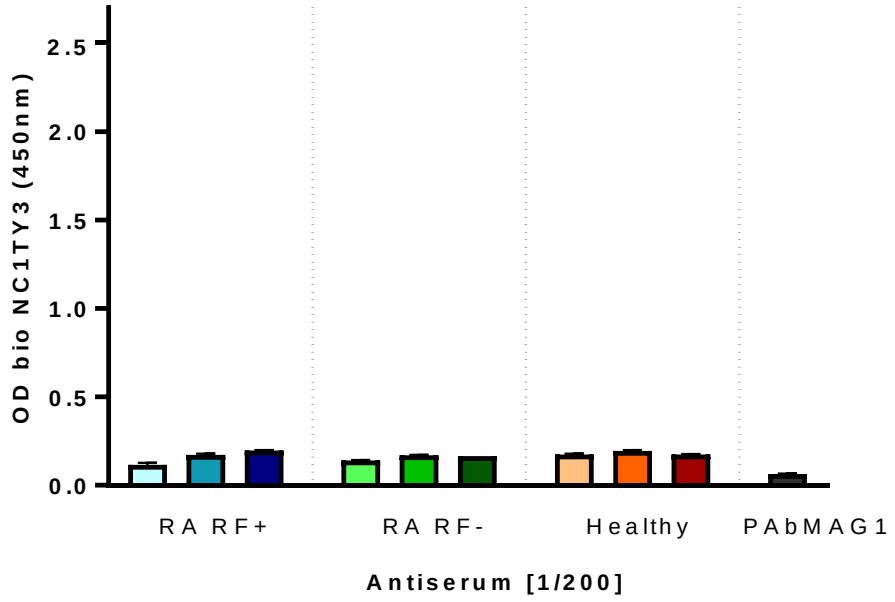


Figure 5.05. Low levels of non-specific binding observed for patient serum diluted in Hammarsten casein blocker. Purified IgG fractions from RF seropositive and RF seronegative RA patients and healthy control serum samples were probed against MAG1 (A) and NC1GTY3 coating antigens (B) (both at 8 µg/ml) in biotinylated ELISA system. Bars represent individual patients and SEM is shown for 3 experiments. RA - rheumatoid arthritis, RF - rheumatoid factor.

5.2.2. Serological analysis of patient cohorts for anti-MAG1 reactivity

Upon the observation that RA autoantibodies show increased reactivity with MAG1 viral peptide, it was important to more accurately determine whether IgG serum antibodies from RA react with HERV-K10 MAG1 and to compare this reactivity to the serological responses of patients with other arthritides and healthy controls. We therefore performed a study in a larger RA patient cohort and incorporated several control cohorts. Antiserum was obtained from whole venous blood of RA patients (n = 124) and the reactivity to biotinylated MAG1 15-mer peptide was evaluated and compared to control groups including non-inflammatory rheumatic disease (osteoarthritis, OA) (n = 10), inflammatory autoimmune arthritis (psoriatic arthritis, PsA) (n =13), autoimmune rheumatic disease (systemic lupus erythematosus, SLE) (n = 8) and healthy samples (H) (n = 30). These control cohorts, in which HERVs may play a causative role or which show similar inflammatory profiles to RA, were employed to examine the possibility that the increase in K10 expression was a by-product of inflammation.

Optical density (OD) values for samples tested in each cohort are demonstrated in Figure 5.06. We found a disease-specific increase in RA patient reactivity to MAG1 peptide identified as a significant difference between the mean IgG antibody response in those patients as compared to OA (*, $p \leq 0.05$), PsA (**, $p \leq 0.01$), and healthy controls (*, $p \leq 0.05$). No significant difference was observed between RA and SLE cohorts. All control cohorts exhibited noticeable clustering around the mean value. A marked overlap in anti-HERV-K10 antibody reactivity was observed between all cohorts, although data stratification (Table 5.01) revealed a large proportion of RA samples (45 / 124; 36%) with OD values statistically higher (OD ≥ 1) when compared to controls.

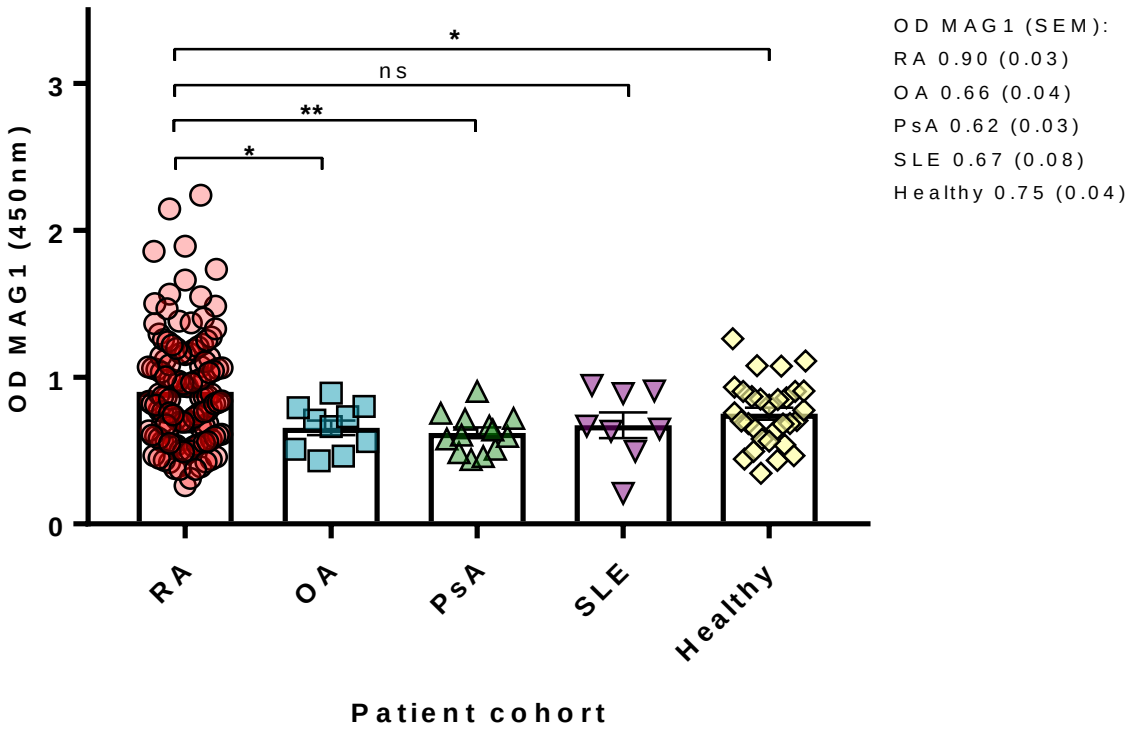


Figure 5.06. Increased mean serological response to HERV-K10 biotinylated matrix peptide MAG1 in patients with RA. Crude serum samples from RA patients and control cohorts (1/200) were incubated with MAG1 coating antigen (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means and SEM (in brackets) is shown for 3 experiments. *, $p \leq 0.05$, **, $p \leq 0.01$, *ns* = not significant. P value calculated by ANOVA was 0.0030. RA: rheumatoid arthritis, OA: osteoarthritis, PsA: psoriatic arthritis, SLE: systemic lupus erythematosus, OD: optical density.

OD MAG1	RA (n)	% RA	OA (n)	% OA	PsA (n)	% PsA	SLE (n)	% SLE	HEALTHY (n)	% HEALTHY
< 0.5	15	12	2	20	3	23	2	25	5	17
0.5 - 0.99	64	52	8	80	10	77	6	75	21	70
1 - 1.49	35	28	0	0	0	0	0	0	4	13
≥ 1.5	10	8	0	0	0	0	0	0	0	0
TOTAL	124	100%	10	100%	13	100%	8	100%	30	100%

Table 5.01. Stratification of serum reactivity to MAG1 peptide in RA patients and control cohorts. Patient stratification is based on the results acquired for their serum reactivity with MAG1 coating antigen (8 µg/ml) in biotinylated ELISA system. Distribution of patient samples (n) and percentile of total is reported per OD ranges for RA, OA, PsA, SLE and healthy controls.

To determine whether the increased reactivity observed for RA cohort in Figure 5.06 was linked to RF titre, I evaluated the reactivity of RF+ and RF- serum samples in comparison to controls. We noticed a significantly increased reactivity of RF+ patients as compared to RF- group (***, $p \leq 0.001$), OA (**, $p \leq 0.01$), PsA (***, $p \leq 0.001$), SLE (**, $p \leq 0.01$), and healthy controls (***, $p \leq 0.001$) (Figure 5.07). There was no significant difference between the RF- cohort and control groups which indicates that RF status might have a direct effect on anti-viral response. Moreover, our data showed a weak positive correlation between RF levels and anti-MAG1 reactivity ($R^2 = 0.02$) (Figure 5.08).

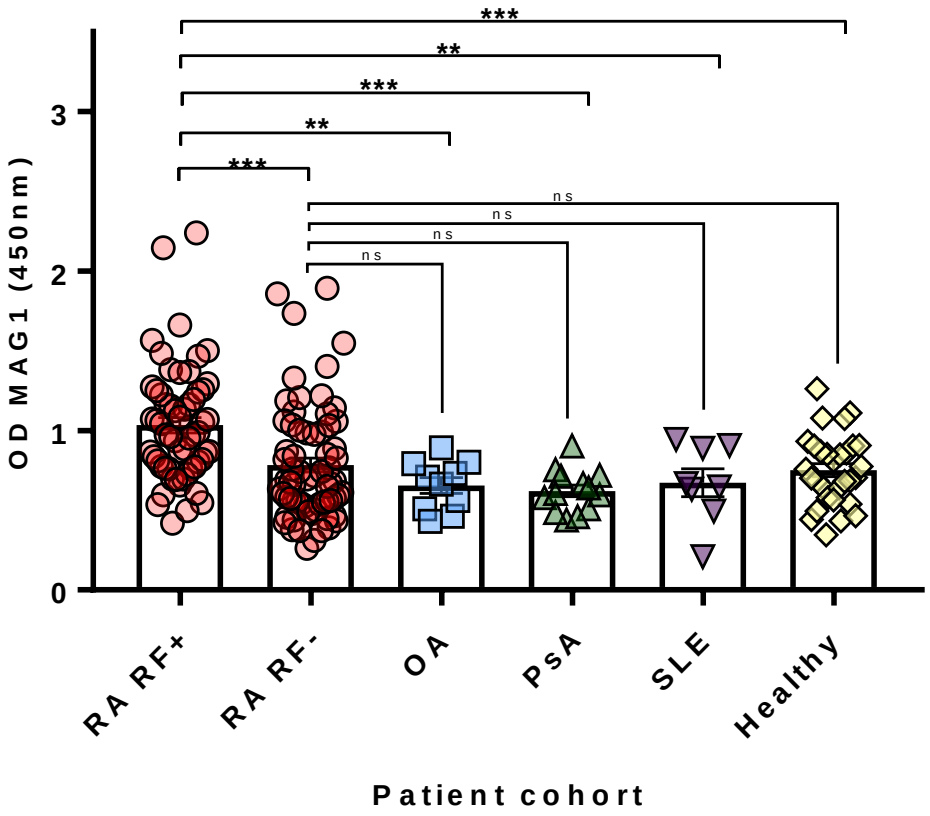


Figure 5.07. Increased mean serological response to HERV-K10 biotinylated matrix peptide MAG1 in RA patients with positive rheumatoid factor profiles. Crude serum samples from seropositive and seronegative RA patients and control cohorts (1/200) were incubated with MAG1 coating antigen (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means and SEM is shown for 3 experiments. **, $p \leq 0.01$, ***, $p \leq 0.001$, *ns* = not significant. P value calculated by ANOVA was < 0.0001. RF: rheumatoid factor.

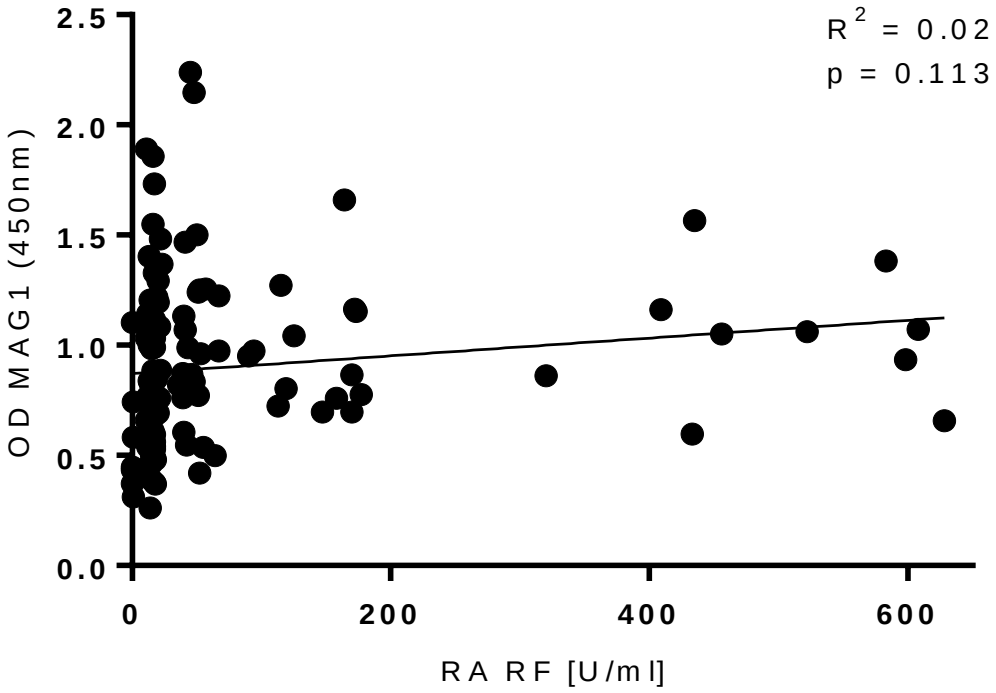


Figure 5.08. Correlation of RA serological activity to HERV-K10 Gag matrix peptide MAG1 with rheumatoid factor levels. RA crude serum reactivity to MAG1 coating antigen (8 µg/ml) assessed in biotinylated ELISA system for each RA patient was correlated with the levels of RF autoantibody detected in their serum using a commercial quantitative ELISA test.

A negative control peptide NC1GTY3 was also used to analyse patient serum samples for reactivity to a random non-human HERV epitope and to remove bias from the investigation. Patterns of reactivity reported for this peptide were intriguing as some patient cohorts, i.e. RA, PsA, SLE and healthy controls, showed lower OD values with comparative distribution ($\bar{x}$ OD = 0.09, 0.095, 0.094 and 0.09, respectively), whilst the OA cohort exhibited higher absorbance levels ($\bar{x}$ OD = 0.126) (Figure 5.09).

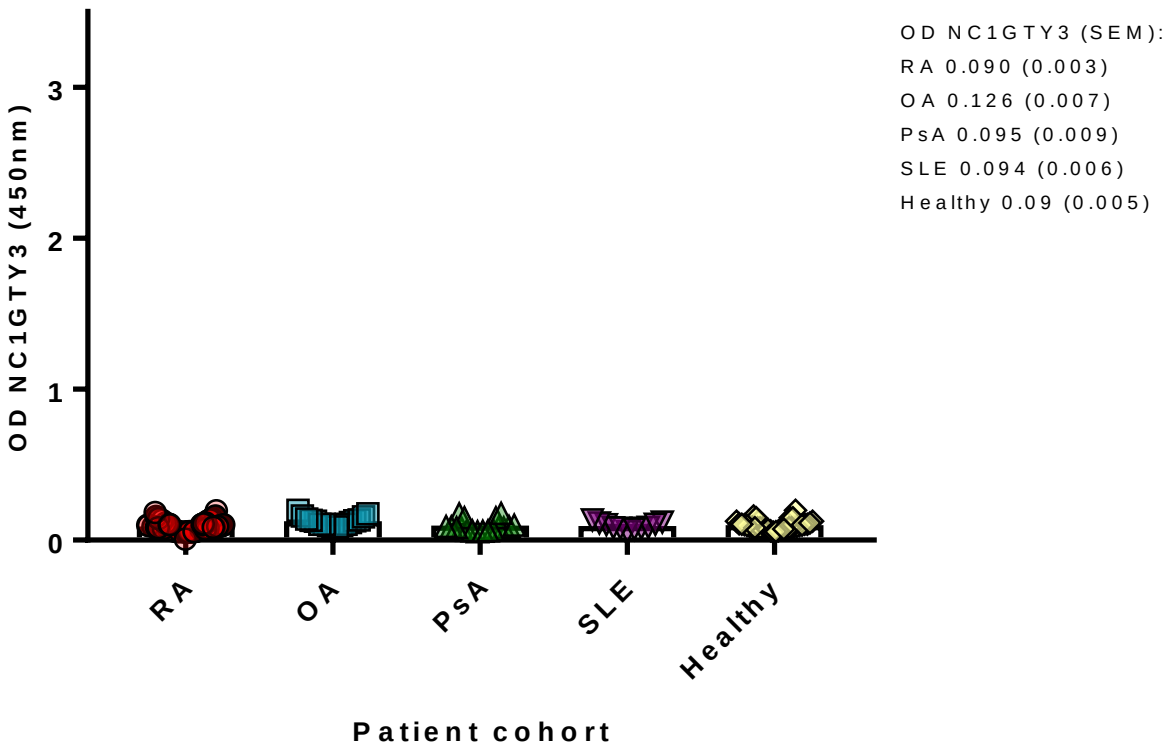


Figure 5.09. Negative serological response to control peptide NC1GTY3 in patients with RA and control cohorts. Crude serum samples from RA patients and control cohorts (1/200) were incubated with NC1GTY3 negative control coating antigen (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means and SEM (in brackets) is shown for 3 experiments.

5.2.3. Serological reactivity of patient serum antibodies to other HERV-K10 peptides

Because RA-derived antibodies in our previous experiments showed increased reactivity to MAG1 viral peptide, we wished to determine whether other homologous HERV-K10 peptide sequences which harbour RF epitopes would evoke similar responses. To achieve this, we assessed RA serum reactivity to MAG2, MAG3 and MAG4 biotinylated peptides and compared them with responses from the control cohorts.

We found comparably high disease specific increases in RA patient reactivity to all three MAG peptides as compared to control cohorts (Figures 5.10, 5.11 and 5.14). With peptides MAG2 and MAG3, there was a significant difference between the mean IgG antibody response in RA patients as compared to OA, SLE and healthy controls (*, $p \leq 0.05$ for all cohorts) and no significant difference with PsA cohort (Figure 5.10 and Figure 5.11, respectively). All control groups demonstrated marked sample clustering around the mean values. Furthermore, we observed a very weak positive correlation between RF levels and anti-MAG2 / anti-MAG3 reactivity ($R^2 = 0.001$) (Figure 5.12 and Figure 5.13, respectively).

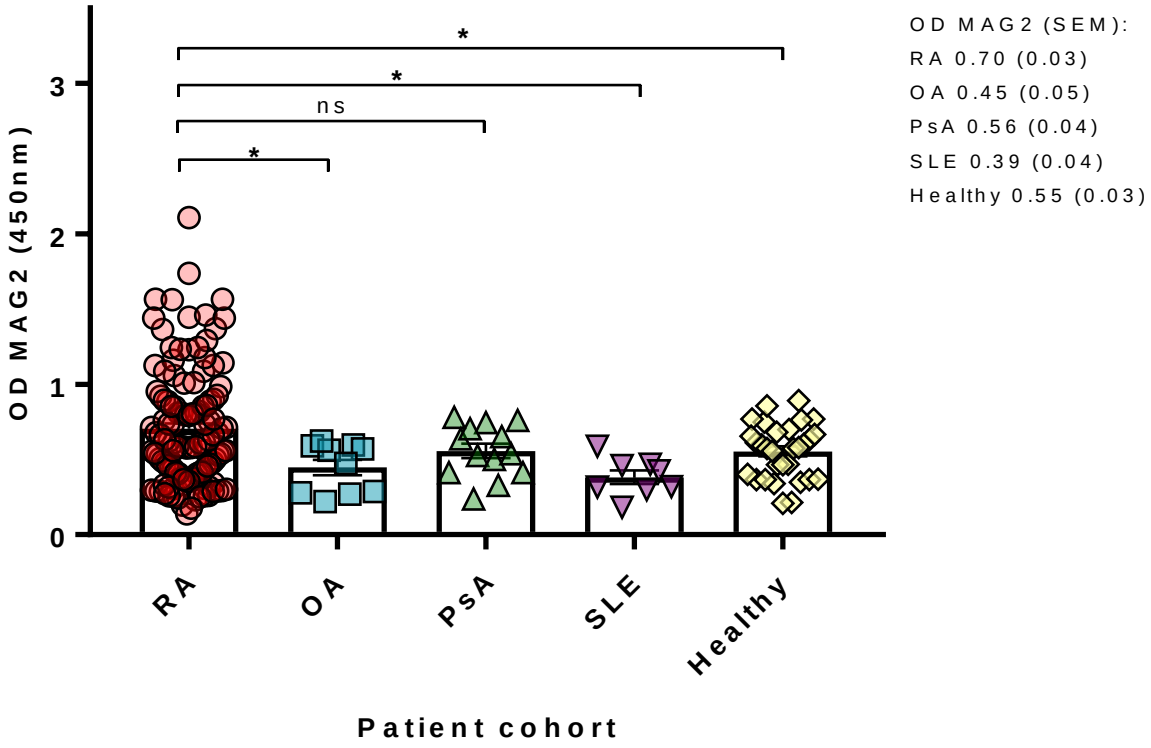


Figure 5.10. Increased mean serological response to HERV-K10 biotinylated matrix peptide MAG2 in patients with RA. Crude serum samples from RA patients and control cohorts (1/200) were incubated with MAG2 coating antigen (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means and SEM (in brackets) is shown for 3 experiments. *, $p \leq 0.05$, *ns* = not significant. P value calculated by ANOVA was 0.0048.

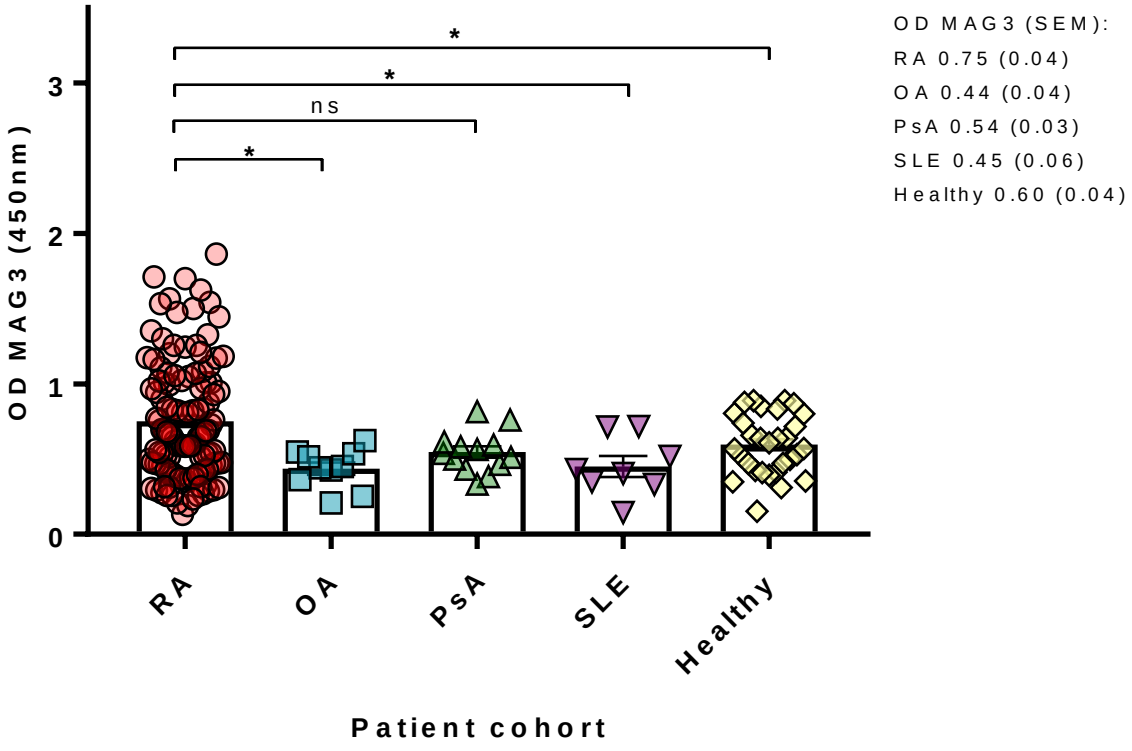


Figure 5.11. Increased mean serological response to HERV-K10 biotinylated matrix peptide MAG3 in patients with RA. Crude serum samples from RA patients and control cohorts (1/200) were incubated with MAG3 coating antigen (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means and SEM (in brackets) is shown for 3 experiments. *, $p \leq 0.05$, *ns* = not significant. P value calculated by ANOVA was 0.0019.

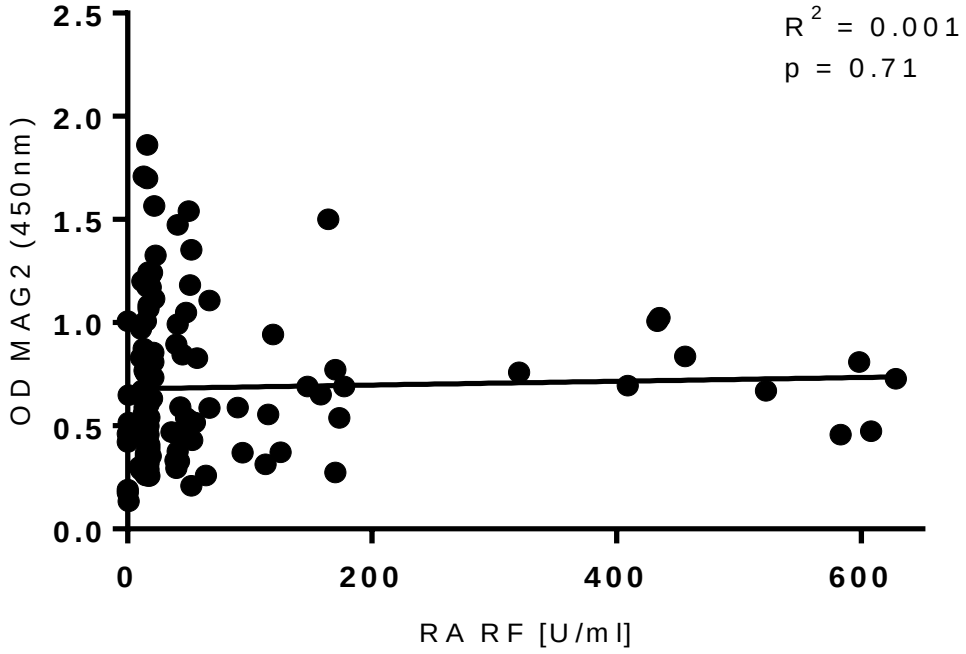


Figure 5.12. Correlation of RA serological activity to HERV-K10 Gag matrix peptide MAG2 with rheumatoid factor levels. RA crude serum reactivity to MAG2 coating antigen (8 µg/ml) assessed in biotinylated ELISA system for each RA patient was correlated with the levels of RF autoantibody detected in their serum using a commercial quantitative ELISA test.

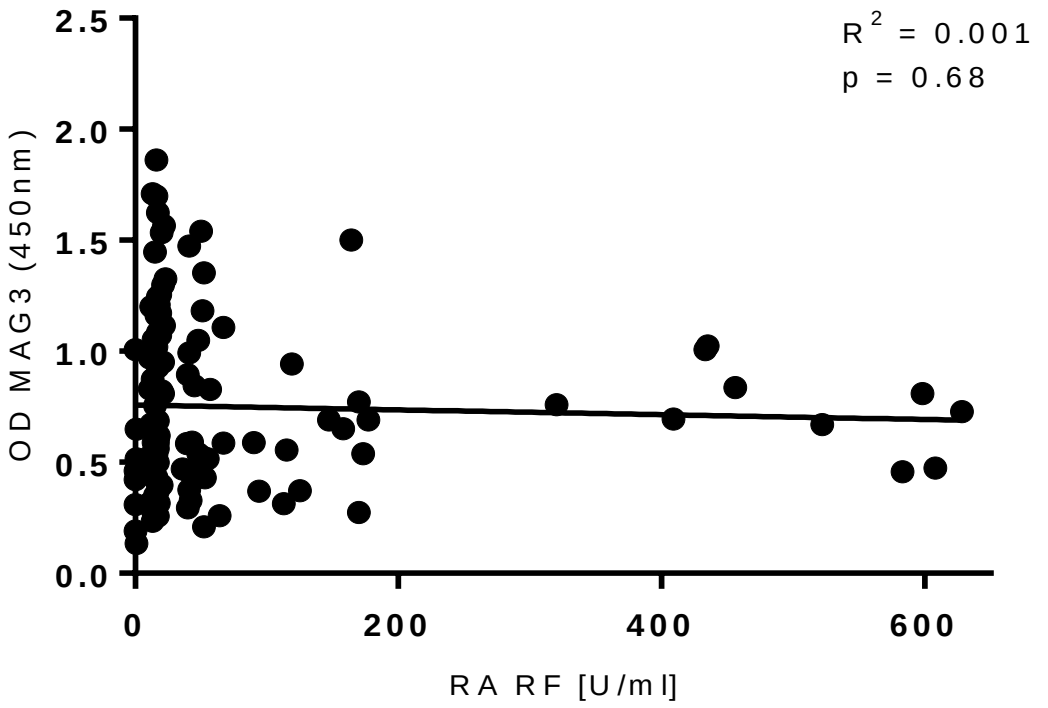


Figure 5.13. Correlation of RA serological activity to HERV-K10 Gag matrix peptide MAG3 with rheumatoid factor levels. RA crude serum reactivity to MAG3 coating antigen (8 µg/ml) assessed in biotinylated ELISA system for each RA patient was correlated with the levels of RF autoantibody detected in their serum using a commercial quantitative ELISA test.

We also found that when using the MAG4 peptide (Figure 5.14), RA sera showed the highest mean levels of reactivity (OD = 0.82) and a significantly increased response when compared to OA, PsA and healthy control groups (*, $p \leq 0.05$) but not to SLE cohort. A weak positive correlation, comparable to that with MAG1 peptide, was noticed between RF levels and the anti-MAG4 response ($R^2 = 0.02$) (Figure 5.15).

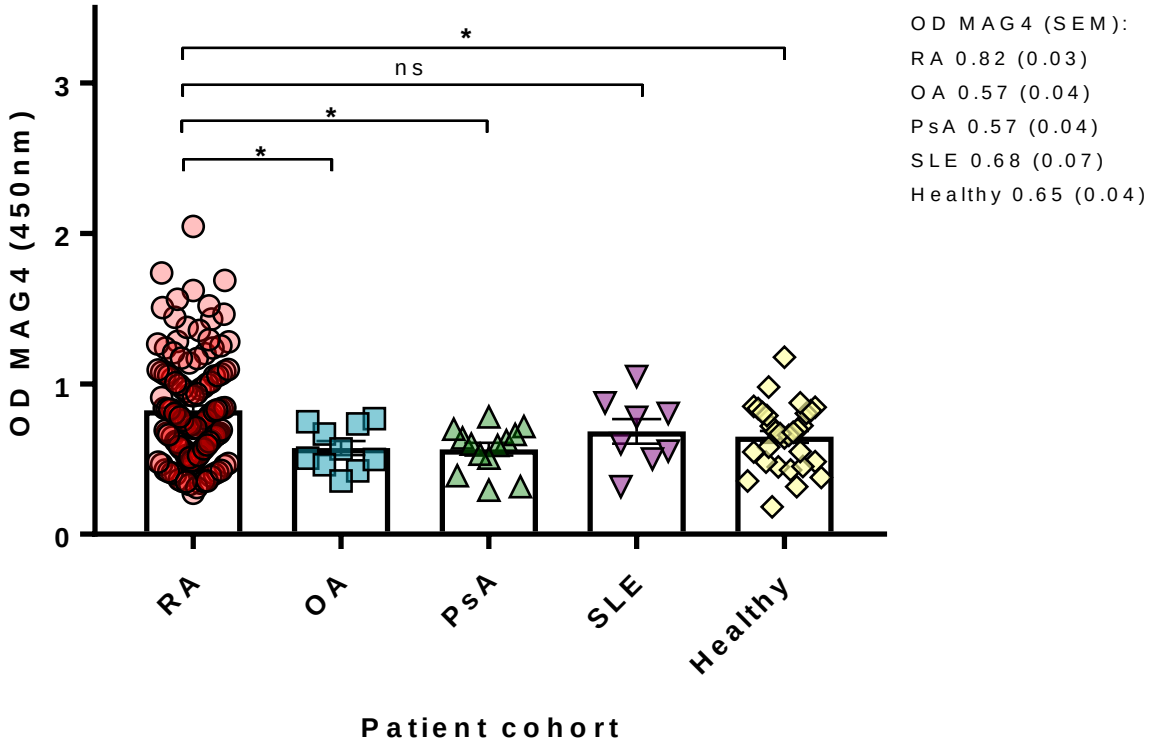


Figure 5.14. Increased mean serological response to HERV-K10 biotinylated matrix peptide MAG4 in patients with RA. Crude serum samples from RA patients and control cohorts (1/200) were incubated with MAG4 coating antigen (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means and SEM (in brackets) is shown for 3 experiments. *, $p \leq 0.05$, *ns* = not significant. P value calculated by ANOVA was 0.0018.

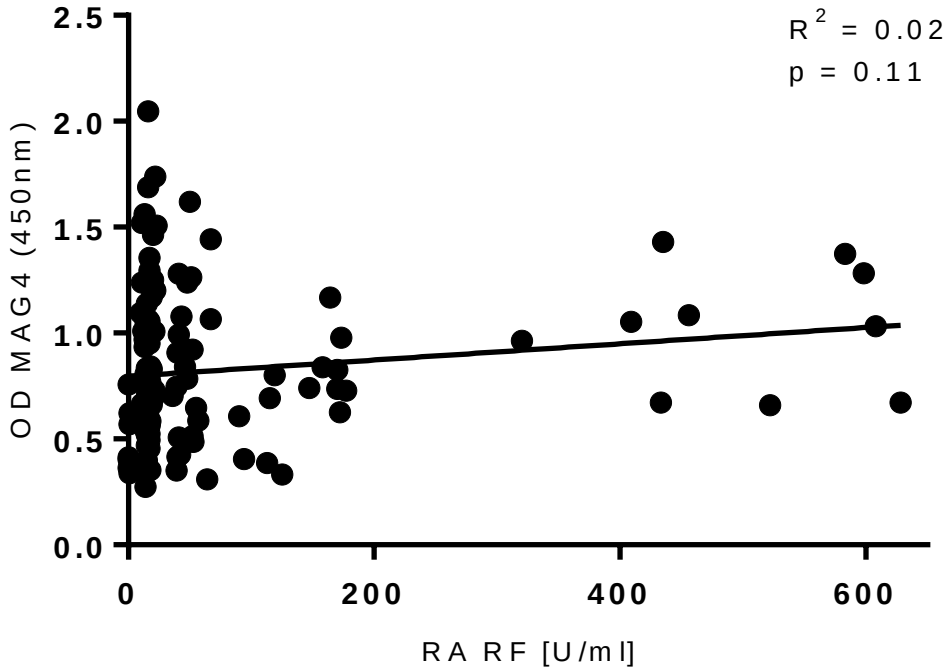


Figure 5.15. Correlation of RA serological activity to HERV-K10 Gag matrix peptide MAG4 with rheumatoid factor levels. RA crude serum reactivity to MAG4 coating antigen (8 µg/ml) assessed in biotinylated ELISA system for each RA patient was correlated with the levels of RF autoantibody detected in their serum using a commercial quantitative ELISA test.

Levels of RA disease markers may fluctuate in response to disease activity and treatment during the course of a clinical protocol (Barra *et al*., 2013). Since HERV-K10 reactivity could be affected by RA therapeutic agents and K10 may act as the triggering event in RA autoimmunity, we wished to assess anti-K10 reactivity in early arthritis as well as at different time intervals. Out of 124 RA patients recruited for this study, 20 patients were newly diagnosed RA cases (within up to 6 months before serum sample collection) exposed to limited treatment, typically consisting of two Disease Modifying Anti-Rheumatic Drugs. We therefore compared the mean serum antibody reactivity of the 20 RA patients diagnosed in 2017 and the remainder diagnosed pre-2017 to MAG1 – 4 peptides. In this scenario, we observed increased antibody reactivity of newly diagnosed RA patients to all four MAG peptides as compared to the RA patients who were diagnosed earlier (Figure 5.16A). Of all peptides tested, only MAG4 showed the greatest difference in reactivity between the two patient cohorts which was significant (**, $p \leq 0.01$) with mean OD values of 1.037 (2017 group) and 0.866 (pre-2017 group). No correlation was observed between RF levels of the newly diagnosed RA patients and their serum anti-MAG reactivity with any of the peptides (Figure 5.16B i-iv). A further 20 patients had their blood sample taken at more than one time point, i.e. at 3, 6 or / and 9 months, in addition to their initial sample (0 m). Of those 20 patients, 17 provided their sample 3 months after the first sample of whom 10 exhibited higher reactivity to MAG1 peptide when compared to 0 m sample, whilst the remaining 7 patients showed lower reactivity compared to the initial sample (Figure 5.17). Four patients provided their ‘6 m’ sample, half of which exhibited higher reactivity than the ‘0 m’ sample. A single ‘9 m’ sample which was assessed showed higher anti-MAG1 reactivity than the initial and the ‘3 m’ samples and lower reactivity when compared to the ‘6 m’ sample. Overall, there was no significant trend towards higher reactivity to

MAG1 in the months following diagnosis, but the longer-term loss of reactivity between the 2017 and pre-2017 cohorts was significant for MAG4 peptide.

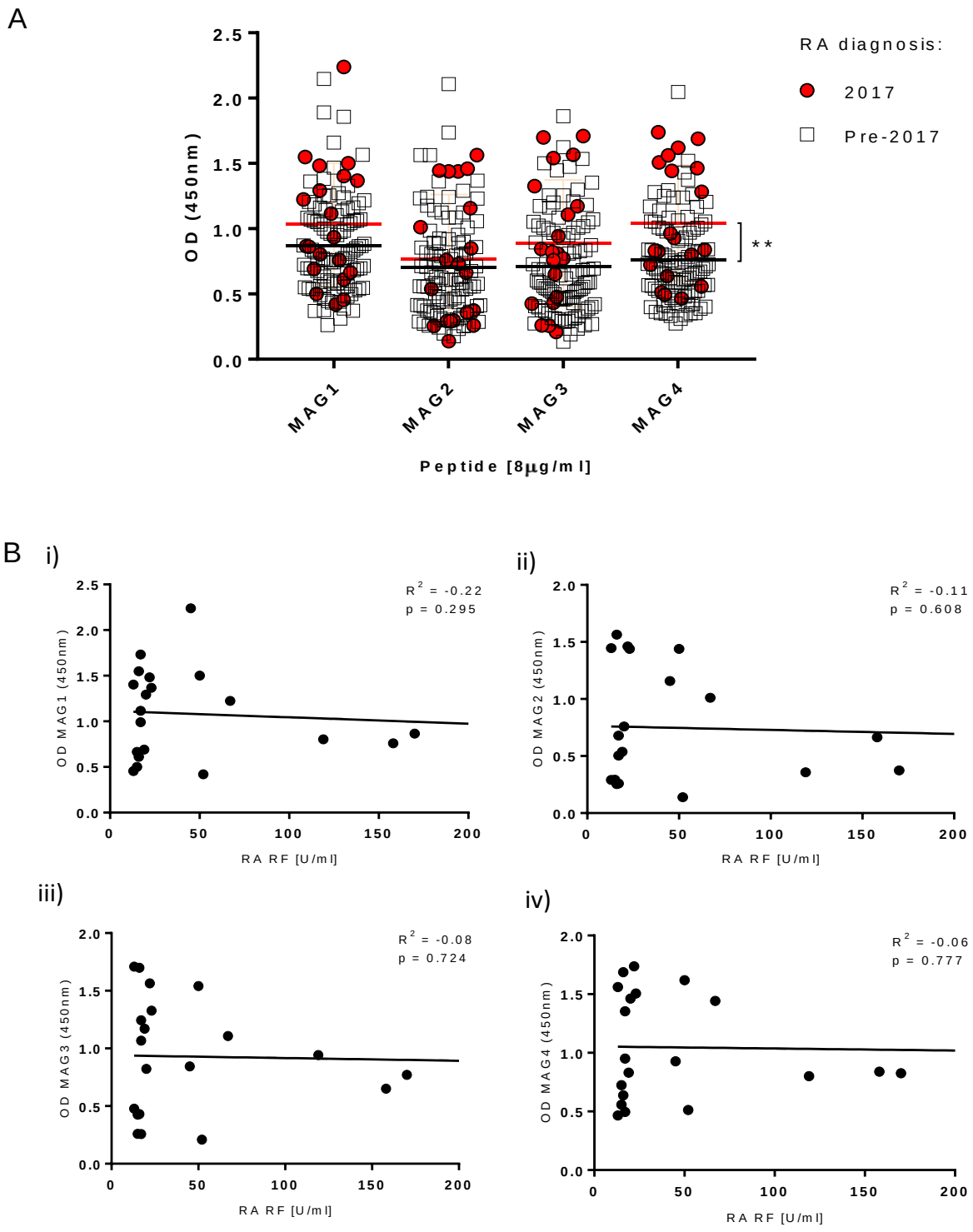


Figure 5.16. Patients with newly diagnosed (2017) RA demonstrate serological reactivity with HERV-K10 peptides MAG1 – 4 that does not correlate with rheumatoid factor levels. Crude serum samples collected from RA patients (1/200) were incubated with four separate MAG coating antigens (8 µg/ml) in biotinylated ELISA system (A). RA crude serum reactivity to MAG antigens assessed for each newly diagnosed patient was correlated with the levels of RF autoantibody detected in their serum using a commercial quantitative ELISA test (B i-iv). Symbols represent individual subjects; bars represent means. **, $p \leq 0.01$.

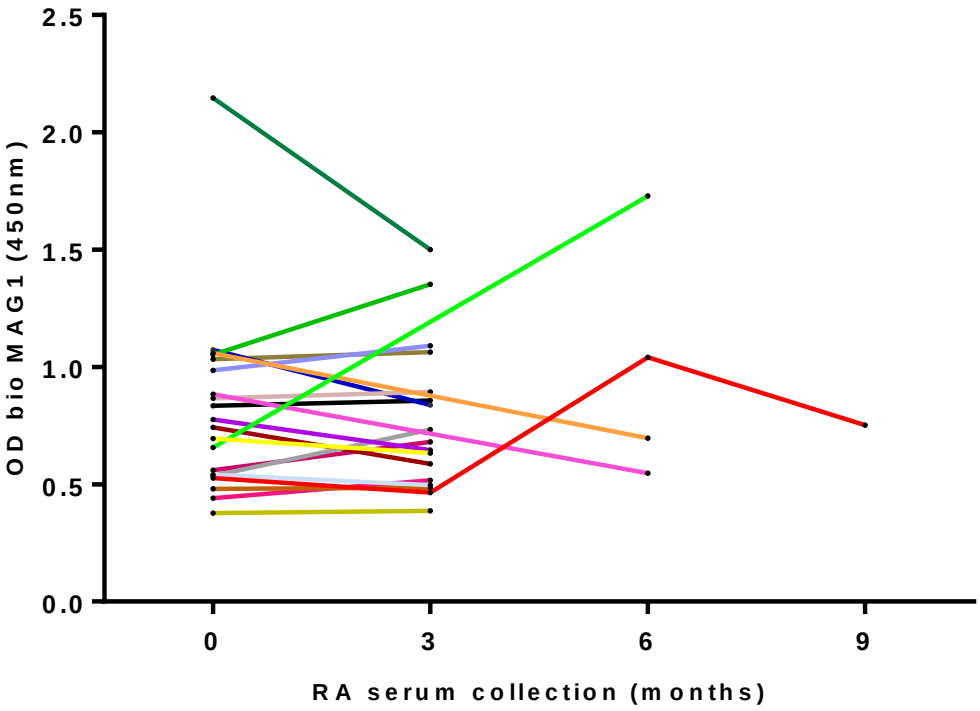


Figure 5.17. Serological response to HERV-K10 biotinylated matrix peptide MAG1 at different intervals reported for 20 patients with RA. Crude serum samples collected from RA patients (1/200) were incubated with MAG1 coating antigen (8 µg/ml) in biotinylated ELISA system. Initial sample (0), sample obtained at 3 months (3), 6 months (6) and 9 months (9) interval.

Since age has been reported to have influence on disease activity in RA (Pawłowska *et al*., 2011) and age-associated decline in IgG responses to HERV-K Env has been previously observed (Kim *et al*., 2016), we evaluated the extent of anti-MAG1 response in relation to patients' age. We found that there was a weak negative correlation with age (R^2 = -0.15) within the RA cohort with mean age of 61 years at the time of sample collection (Figure 5.18). In addition, as hormonal factors appear to contribute to autoimmunity (Quintero *et al*., 2012) and to differentially influence RA development (Alpízar-Rodríguez *et al*., 2016), HERV expression may be affected by those hormonal signals (Nelson *et al*., 2004). Therefore, we assessed the levels of anti-HERV-K10 reactivity with regard to patient sex and demonstrated weakly raised mean OD values in females to MAG1, MAG2 and MAG3 but not to MAG4 peptide, as compared to males, however no statistical significance was observed between these two cohorts (Figure 5.19).

In order to analyse the levels of anti-MAG1 reactivity in all 124 RA sera and to identify any correlations between the levels of serum reactivity to MAG1 and other MAG peptides, a collective heatmap of anti-K10 reactivity was created (Figure 5.20) and structured according to anti-MAG1 OD values from low (bright green) to high (red). We found that RA antibody reactivity to MAG1 peptide is predominantly concentrated around medium to high OD values (OD: 0.5 – 0.99 and 0.1 – 1.49, respectively) with significant overlap between peptides i.e. reactivity levels for MAG1 often reflect the levels observed for other MAG peptides. There was a small proportion of patient samples which exhibit either low reactivity profiles (OD: 0.1 – 0.49; bright green) or very high reactivity (OD: 1.5 – 2.5; red) across all MAG peptides.

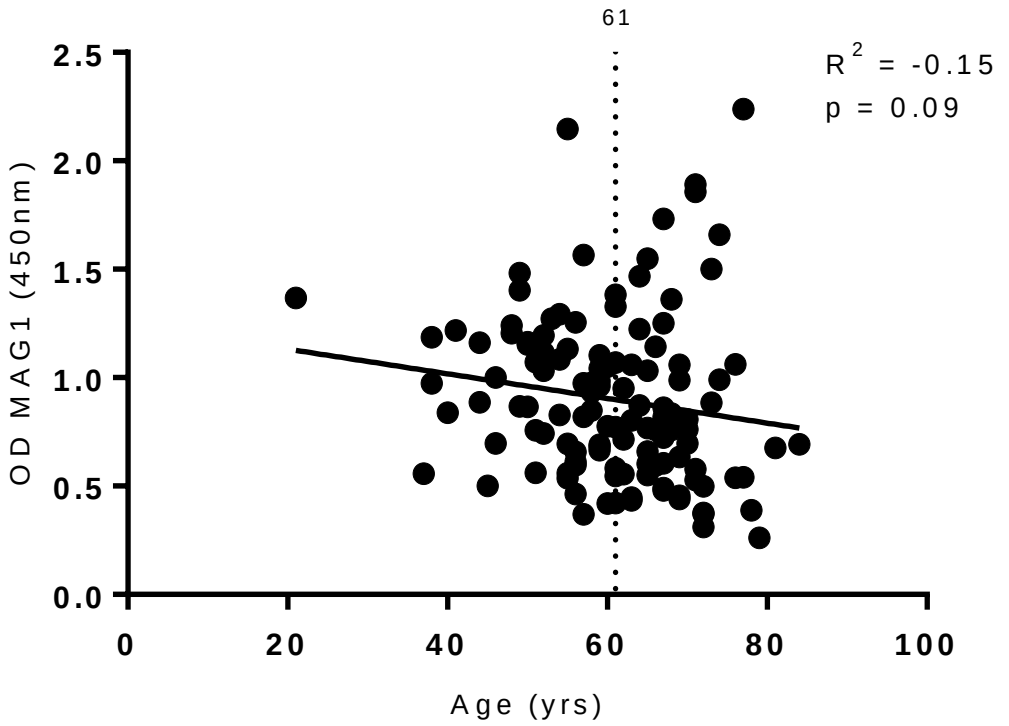


Figure 5.18. Negative correlation of RA serological activity to HERV-K10 peptide MAG1 with patient's age. RA crude serum reactivity to MAG1 coating antigen (8 µg/ml) assessed in biotinylated ELISA system for each RA patient was correlated with patient's age. Dotted line indicates patient's mean age (61 years).

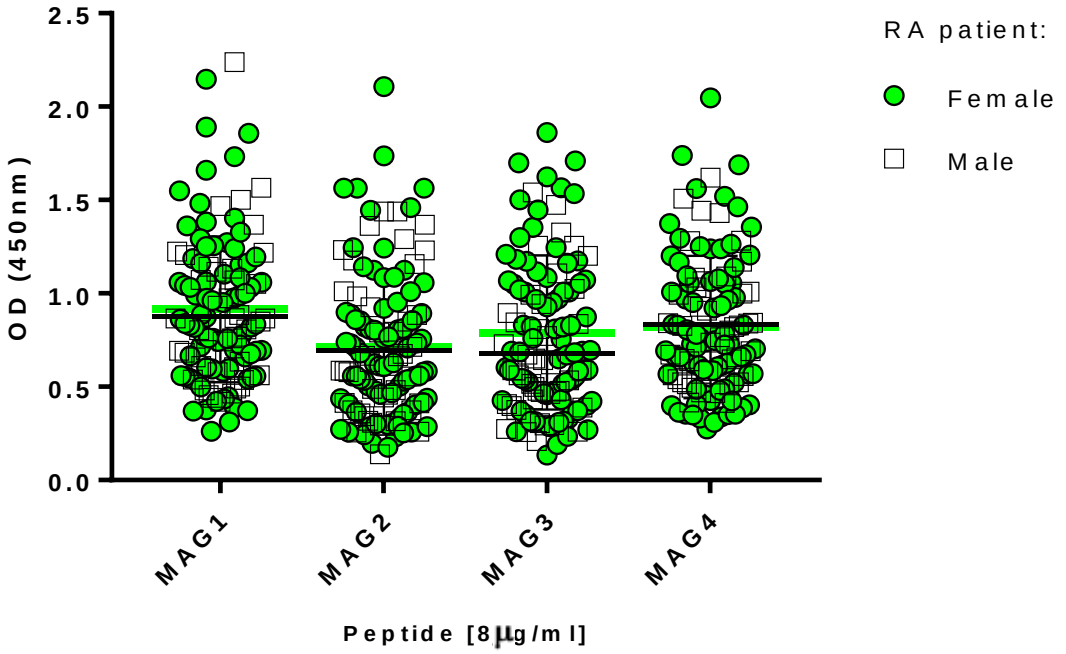


Figure 5.19. Marginally enhanced mean serological response to HERV-K10 peptides MAG1 - 4 in female versus male patients with RA. Crude serum samples collected from RA patients (1/200) were incubated with four different MAG coating antigens (8 µg/ml) in biotinylated ELISA system. Symbols represent individual subjects; bars represent means.

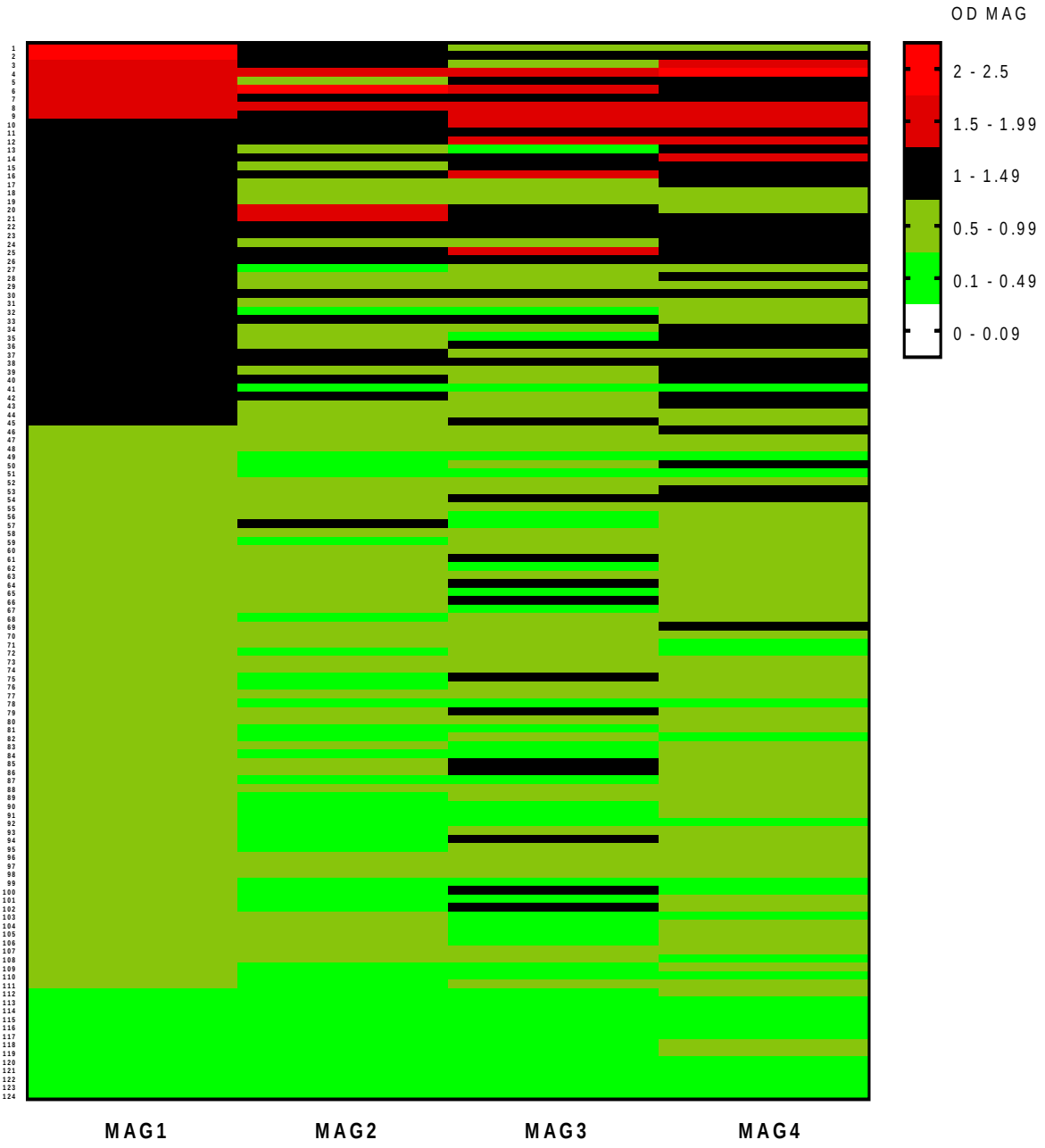


Figure 5.20. Heat map of antibody profiles showing the intensity of reactivity to four HERV-K10 MAG peptide mimics of IgG1 Fc in individual RA patient serum samples. Each column of the heat map represents an antigen-specific antibody response to MAG1-4 coating peptides (8 µg/ml) in biotinylated ELISA, while each row represents an individual RA patient (n = 124). Samples are sorted by decreasing reactivity (OD values) for MAG1. The colour scale shows the relative degree of antibody reactivity, from low reactivity (white) to high reactivity (light red). Red indicates OD: 1.99-2.5; black, OD: 1-1.49; green, OD: 0.1-0.99; white, OD: 0-0.09.

5.3. Discussion

The immunopathogenesis of rheumatoid arthritis remains elusive, however a viral agent, together with genetic background, could confer susceptibility to disease onset. The detection and role of human endogenous retroviruses has gained growing interest in RA as a result of several serological and molecular studies in clinically active individuals (Tugnet *et al*., 2013; Herve *et al*., 2002). One of the most biologically active forms of endogenous virus, HERV-K10, has proved particularly intriguing in RA patients (Ehlhardt *et al*., 2006; Nelson *et al*., 2004). We have previously shown that this agent, classified within the HML-2 family, may act as a candidate elicitor of autoimmune responses in RA through cross-reactivity with host proteins termed molecular mimicry. HERV-K10 bares structural homology with some autoantigens in RA, however the chief RA autoantigen, IgG1 Fc, has not been previously explored. Increased serological response of IgM class antibodies to HERV-K10 Gag matrix products, crucial for viral particle formation, was reported for RA patients as compared to inflammatory and healthy controls (Freimanis *et al*., 2010). This finding was substantiated by molecular investigation which revealed upregulation of K10 mRNA in PBMCs and synovial fibroblasts from RA patients with basal level in controls, consistent with the ubiquitous expression of HERV-K10 in the human genome (Freimanis *et al*., 2010; Ejtehadi *et al*., 2006; Nelson, Freimanis and Roden, 2008).

Our group has previously observed an increased secondary serological response to a single HERV-K10 matrix fragment in a small cohort of RA patients (Nelson *et al*., 2014a). Consequently, this finding prompted us to assess the IgG antibody response to HERV-K10 Gag in a large RA patient cohort and to investigate additional homologous peptide mimics (MAG1 – 4) which harbour RF epitopes, with MAG1 constituting the main region of interest.

Our optimised HERV-K10 ELISA system was used to examine patient serum samples, identifying different degrees of reactivity to K10 matrix epitopes in RA patients as compared to patients with OA, PsA, SLE, and healthy individuals. We have demonstrated significantly increased mean IgG antibody reactivity toward all four K10 peptides in RA patients as compared to control cohorts with the highest mean observed for MAG1 peptide (OD = 0.9) (Fig. 5.06). This finding is in agreement with our *in silico* investigation (chapter 3) which revealed that MAG1 peptide contained the most immunodominant epitope with homology to IgG1 Fc autoantigen and RA shared epitope motif. In specifically defining anti-IgG reactivity to HERV-K10 in RA, our observations support the notion of an antigen driven immune response, which does not simply constitute a by-product of inflammation. Clearly a continuous supply of endogenous antigen would undoubtedly promote a secondary immune response that would facilitate class switching from IgM to IgG. It may also suggest some specificity for K10 in RA, although there remain a number of other arthritides that have yet to be assessed (e.g. gout, ankylosing spondylitis). We also observed clear variations in reactivity within the RA cohort which is consistent with previous serological reports regarding HERV expression in other conditions such as SLE showing various response among patients (Perl, 2003). The heterogeneity of HERV activity exhibited by RA and, to some extent, the control groups with MAG1 – 4 peptides may reflect the distinct immune reactions among patients or the presence of specific polymorphisms (variants) within HERV-K10 (Moyes, Griffiths and Venables, 2007). Interestingly, HERV sequence polymorphism has been suggested as a contributing factor to allelic variation in the human genome and thus a possible association of certain haplotypes with autoimmunity such as that of HRES-1 (genotype I allele) with SLE (Magistrelli *et al*., 1999). Moreover, some insertionally polymorphic HERVs have been integrated into only a proportion of individuals e.g. HERV-K(C4) gene is present in approximately 70%

of the population (Dangel *et al*., 1994). This observation implies that rheumatoid arthritis does not comprise a single disease entity (Firestein and Mcinnes, 2017) which has also been recognised for other autoimmune conditions such as SLE due to the variety of clinical manifestations in both diseases (Agmon-Levin *et al*., 2012).

Interestingly, we also noticed a significantly increased ($p \leq 0.001$) anti-MAG1 reactivity of RF+ versus RF- RA patient cohort (Fig. 5.07), with the latter reported to be more heterogenous than the seropositive RA (Malmström, Catrina and Klareskog, 2017). This finding suggests that RF+ serum antibodies are more cross-reactive with MAG1 epitopes when compared to RF- sera. Given that HERV-K10 MAG peptides exhibit molecular mimicry with epitopes located on IgG1 Fc, which are within rheumatoid factor binding region, we next wished to determine whether there was a correlation between anti-MAG IgG reactivity and the levels of IgG RF in RA patient serum. We observed a weak positive correlation of the antibody reactivity to all four MAG peptides with RF levels (Figs. 5.08, 5.12, 5.13, 5.15), the latter reflecting disease reactivity for seropositive individuals (Aletaha, Alasti and Smolen, 2015). While the Disease Activity Score (DAS) was not available at the time of RA serum collection, it was fascinating to note that Reynier *et al*. (2009) reported a substantial association between HERV-K viral load in peripheral blood and the disease activity (as established by the DAS28 score) in RA during active disease period. Consequently, further studies employing DAS should be undertaken to verify this finding for HERV-K10.

Given the complexity of immunoreactivity, certain levels of RF autoantibodies could be generated even in patients diagnosed with seronegative RA by means of standardised kits with a cut-off value of antibody levels at < 20 U/ml for RF-negative patients (Ajeganova and Huizinga, 2014). Indeed RF has been shown to be false-negative around 31% of the time (Walker-Cillo, 2010). Therefore, our findings, which may suggest HERV-K10 involvement in a considerable number of predisposed RA patients,

such as those recognised within this study, could potentially contribute to the development of an additional diagnostic assay for RA.

In addition, we found that there was a small overlap in the levels of anti-K10 reactivity between RA and control cohorts, which may be indicative of the ubiquitous nature of HERV-K10 in the general population with 30-50 copies per haploid genome (Nelson *et al*., 2014a). To date there are no extensive population studies concerning the basal levels of expression of anti-HERV antibodies in healthy individuals and different diseases for direct comparison with our findings. Therefore, our data provides a level of base-line activity, particularly in healthy individuals that may be useful to future studies in RA and other medical conditions. Overall, despite the abundance of antibodies present in patient serum samples, we demonstrated negligible cross-reactivity of serum IgG to the negative control peptide NC1GTY3, which confirmed the specificity of anti-MAG response. Moreover, this validated our bioinformatics methodology and the performance of our assays.

This Ab isotype arises upon numerous somatic hypermutations and thus affinity maturation of antigen-binding sites (Doria-Rose and Joyce, 2015), and may reflect the tenacity of Ab binding in our anti-HERV-K10 ELISA system. Recalling our earlier data in this chapter and chapter 4, we found that non-biotinylated MAG peptides performed poorly when exposed to the patient serum samples due to potential epitope masking via steric hinderance. Complementary binding was therefore maximised through the use of a neutravidin-biotin peptide bridge which improved serum antibody binding to MAG peptides. Biotinylation has been shown to enhance analyte binding by more than 200-fold which is consistent with our observations (Trilling, Beekwilder and Zuilhof, 2013; Freimanis *et al*., 2010).

We also observed that the levels of anti-MAG reactivity were enhanced in the 20 patients with early RA, who were only exposed to DMARD therapy, as compared with the patients with established disease (Fig. 5.16). This notion was previously reported for anti-collagen autoantibodies in RA (Whittingham, Stockman and Rowley, 2017) together with the observation that RF autoantibody levels mirrored the disease severity. Moreover, substantial fluctuations of anti-K10 antibody activity were detected in further 20 patients for whom sequential serum samples of up to 9 months since the initial sample were available (Fig. 5.17). These results are consistent with changing RF autoantibody levels in patients responding to therapies (Aletaha, Alasti and Smolen, 2015). When correlating the difference in the mean IgG antibody levels to MAG1 with patient's age we demonstrated a weak negative correlation between this dataset (Fig. 5.18). In support of this, it has previously been shown that IgG responses to HERV-K Env decline with increasing age (Kim *et al*., 2016). In addition, we have previously reported minor fluctuations in anti-K10 antibody levels with age in control patient cohorts, which could serve as an index of nominal anti-HERV-K10 response in future autoimmunity research (Nelson *et al*., 2014a). It is interesting to note that HERV-K reactivation is consistent with human adult-onset diseases and / or senescence where HERVs, normally silenced by DNA methylation, undergo gradual global demethylation with age (Balestrieri *et al*., 2015). Global DNA hypomethylation has been reported in cells and tissues derived from patients with RA (Klein, Ospelt and Gay, 2012), where typical mean age of onset ranges from 30 to 50 years (Machado-Alba, Ruiz and Medina Morales, 2015). Moreover, HERV hypomethylation, which lead to increased viral expression, has also been demonstrated in other autoimmune conditions including SLE (Wu *et al*., 2015).

Further, we determined the impact of gender on the antibody reactivity to all four K10 epitopes as it is suggested that HERVs may contain hormone responsive element

which plays a role in viral expression (Trela, Nelson and Rylance, 2016). We observed a slight increase in the mean IgG levels to MAG1 and MAG3 peptides among female subjects as compared to male patients (Fig. 5.19), which is in agreement with the notion that hormonal factors may influence HERV expression (Nelson *et al*., 2004). In addition, RA prevalence is 2-fold to 3-fold greater among women compared to men, which indicates hormonal involvement providing further support for our findings (Smolen *et al*., 2018). Moreover, an increased HERV-K10 mRNA and protein expression has also been detected in oestrogen-stimulated cell model systems (Ejthadi *et al*., 2005), which may therefore suggest that oestrogen could increase HERV expression leading to RA perpetuation.

In this chapter, on screening patients' serum samples, we found significant anti-IgG reactivity to all four HERV-K10 MAG peptides in RA patients compared with disease control cohorts and healthy subjects. This finding supports the notion of molecular mimicry between endogenous retrovirus and IgG1 Fc that could augment high-affinity anti-HERV-K10 and autoantibody production in RA. In future experiments, it will be important to assess the significance of subtle structural variations within epitopes, possibly arising from post-translational modifications, which could contribute to variable reactivities of the RA-associated autoantibodies with HERV-K10 peptides.

6. IMPACT OF HERV-K10 AND AUTOANTIGEN CITRULLINATION ON RA SERUM ANTIBODY REACTIVITY

6.1. Introduction

The autoimmune nature of RA is accentuated by abundant generation of autoantibodies with diverse antigenic specificities and varying diagnostic value. One such autoantibody, rheumatoid factor (RF) constitutes a hallmark of seropositive RA and binds to the IgG1 Fc fragment forming large immune complexes in the synovium (Ingegnoli, Castelli and Gualtierotti, 2013). The other major group of autoantibodies present in around 60 to 70% of RA patients' sera constitute antibodies against citrullinated proteins / peptides (ACPA) (Fernandes-Cerqueira *et al*., 2015). Citrullination is a form of post-translational modification involving conversion of protein-bound arginine residue into the non-standard amino acid citrulline in the presence of peptidylarginine deiminase enzymes. This process, which is believed to contribute to RA pathogenesis, promotes neo-(auto) antigen production to which self-tolerance has never been established (Valesini *et al*., 2015; Badillo-Soto *et al*., 2016). The targets of ACPAs identified thus far in RA synovium constitute a number of structural citrullinated proteins such as fibrinogen, which is one of the most abundant autoantigens in RA, as well as vimentin, collagen type II, and several viral antigens (Wegner *et al*., 2010; Trouw, Huizinga and Toes, 2013). However, recognition of numerous citrullinated autoantigens might suggest that no traditional antigenic determinant is targeted by ACPAs and that the culpable autoantigens can indeed change over time instigating new reactivities in the ACPA response (Trier *et al*., 2015). Moreover, Tsuda *et al*. (2015) demonstrated that ACPAs might be induced by molecular mimicry due to exposure to multiple citrullinated viral proteins. Therefore, antibodies generated

against citrullinated epitopes could sustain the immune response via epitope spreading and cross-reactivity with a broad array of citrullinated proteins. Although RF and ACPA are autoantibodies to vastly different target proteins in terms of their origin and function, it has been suggested that RF may also bind to immune complexes containing an antigen composed of citrullinated sequences of arginine-rich cellular proteins (Tan and Smolen, 2016).

We have previously demonstrated immunological cross-reactivity of anti-HERV-K10 antibodies to the IgG1 Fc region and an increased IgG response in RF+ RA serum to multiple K10 epitopes (chapter 5), supporting the premise of molecular mimicry between the virus and IgG1 Fc, and an association with RF binding epitopes. Moreover, our *in silico* alignment studies (chapter 3) confirmed that there is a considerable degree of homology between human IgG1 Fc and other RA autoantigens, which may contribute to cross-reactivity. Because citrullination occurs extensively within RA synovium, we set out to assess how this process affected the antigen : antibody interactions, and how serum from patients with RA reacted with modified versus native epitopes.

6.2. Results

6.2.1. RA autoantigen selection and detection of citrullinated targets

To further explore the premise of autoantigenic mimicry and potential epitope spreading in RA we next wished to determine the level of sequence and structural homology between HERV K10, IgG1 Fc and other abundant autoantigens in the RA joint. In chapter 3 we have observed that K10 epitopes mimic those on IgG1 Fc (Table 3.06) and that fibrinogen contains numerous highly homologous amino acid sequences (of 5 resides and more) with IgG1 Fc (Table 3.04). To determine sequence similarities, we therefore aligned these fragments (Table 6.01) and demonstrated two peptide sequences on IgG1 Fc, each composed of 5 linear residues (KPREE, DELTK), with 80% and 60% identity, consecutively, in terms of alignment scores with fibrinogen, and further 4- and 6-residue IgG1 Fc sequence (KSRW, HEDPEV) with 75% and 50% identity to fibrinogen, respectively. Overall, the respective substitutions (e.g. proline/lysine, threonine/asparagine, lysine/asparagine and serine/threonine) of these particular peptides scored reasonably in terms of amino acid replacement values (-2 to +34). We also showed two epitopes on HERV-K10, each composed of 5 amino acids (GKELK, IDKSR), with 60% homology to those on fibrinogen and two further K10 epitopes (PPSR, KVSTKNL) with 50% identity to fibrinogen with reasonable replacement score (-14 to +46). In addition, to assess the conformational similarity between these viral and host antigens, we converted them into 3D molecular graphic models (Figure 6.01). Minor structural differences in the overall mesh structure of the compared antigen pairs were revealed particularly between fibrinogen and IgG1 Fc.

Moreover, within these data we observed multiple arginine residues between the homologous sequences identifying them as targets for citrullination in RA (Table 6.01 and Figure 6.01, red highlighted resides), which may lead to autoantibody cross-

reactivity by generating neo-epitopes and epitope spreading (Valesini *et al*., 2015). It is well established that citrullination has several different implications for protein conformation such as change in protein folding due to the loss of arginine positive charge (Badillo-Soto *et al*., 2016). To this end we verified the effect of citrullination on the three-dimensional structures of these epitopes using *in silico* molecular visualisation, which revealed subtle conformational alterations within tertiary structure of the homologous pairs that did not perturb the overall shape of the molecules (Figure 6.02).

	In silico sequence alignments			
IgG1 Fc	190-K**S**RW	66-K**P**REE	132-DEL**TK**	44-H**E**DP**EV**
Fibrinogen	399-K**T**RW (γ)	51-K**K**REE (β)	128-DEL**NN** (β)	97-H**A**DP**DL**
Alignment Score (%)	3/4 (75%)	4/5 (80%)	3/5 (60%)	3/6 (50%)
Replac.Score	+24	-2	+12 +17	+12 +34 +26
HERV-K10 peptide	MAG1: RIGKELKQAGRKGNI	MAG2: PMAPPSRQGSE	MAG3: IIDKSRKEGDT	MAG4: KRGGVKVSTKNL
HERV-K10 epitope	3-GKE**LK**	24-P**PS**R	36-I**DKS**R	29-**K**VS**T**K**NL**
Fibrinogen	691-GKE**CE**	895-P**NG**R	502-**LDKK**R	688-**V**VS**G**K**EC**
Alignment Score (%)	3/5 (60%)	2/4 (50%)	3/5 (60%)	3/6 (50%)
Replac.Score	+15 +24	+8 +9	+46 -10	-14 -13 +13 +15

Table 6.01. Regions of sequence homology established *in silico* for IgG1 Fc, fibrinogen and HERV-K10. Matching homologous amino acid pairs were established using publicly available bioinformatics software ExPASy SIM and LALIGN. Numerals indicated amino acid position and sequence alignment with reference to protein accession numbers: HERV-K10 (AAA88030.1), IgG1 Fc (AF150959.1), fibrinogen β-chain (P02675) and fibrinogen γ-chain (P02679). Bold characters indicate amino acid substitutions and characters highlighted in red identify arginine residues susceptible to citrullination. Residues Alignment score shown as number of identical residues and % homology. Tudos Replac. (replacement) score for amino acid substitution. Amino acid single letter code: A-Alanine, S-Serine, T-Threonine, V-Valine, L-Leucine, D-Asparate, E-Glutamate N-Asparagine, P-Proline, K-Lysine, R-Arginine, H-Histidine, W-Tryptophan.

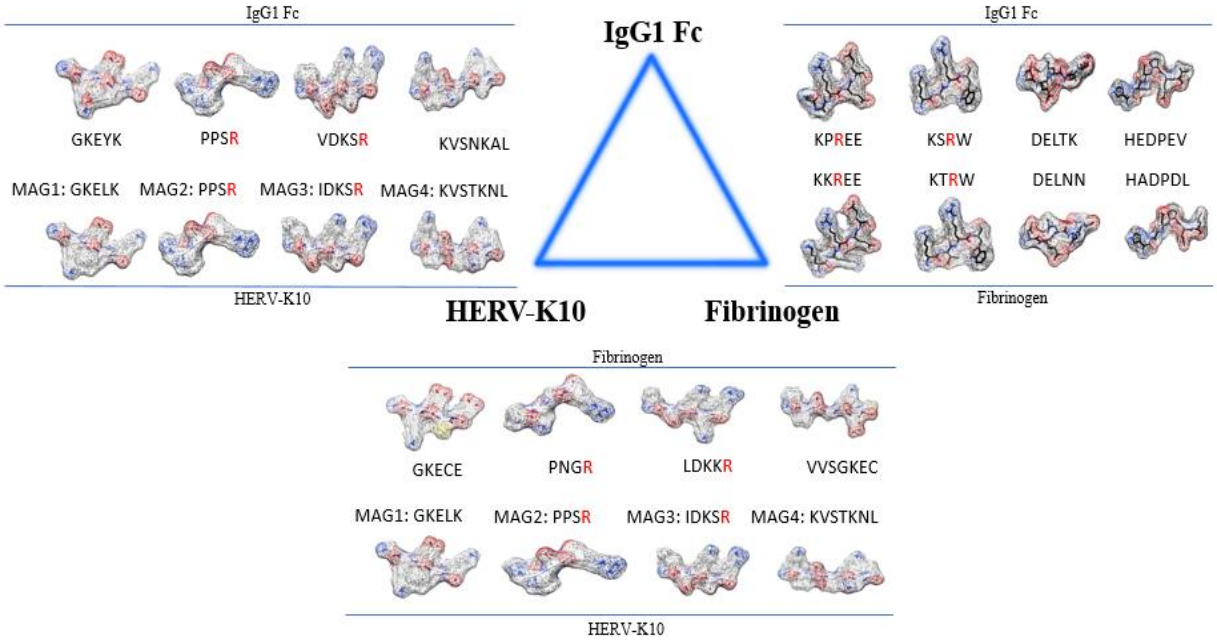


Figure 6.01. Molecular models demonstrating structural mimicry between immunodominant epitopes of HERV-K10 and RA autoantigens IgG1 Fc and fibrinogen. Three dimensional structures of homologous regions between HERV-K10, IgG1 Fc and fibrinogen were modelled in PyMOL and UCSF Chimera molecular graphics systems and presented in mesh mode together with their amino acid sequences. Characters highlighted in red identify arginine residues susceptible to citrullination.

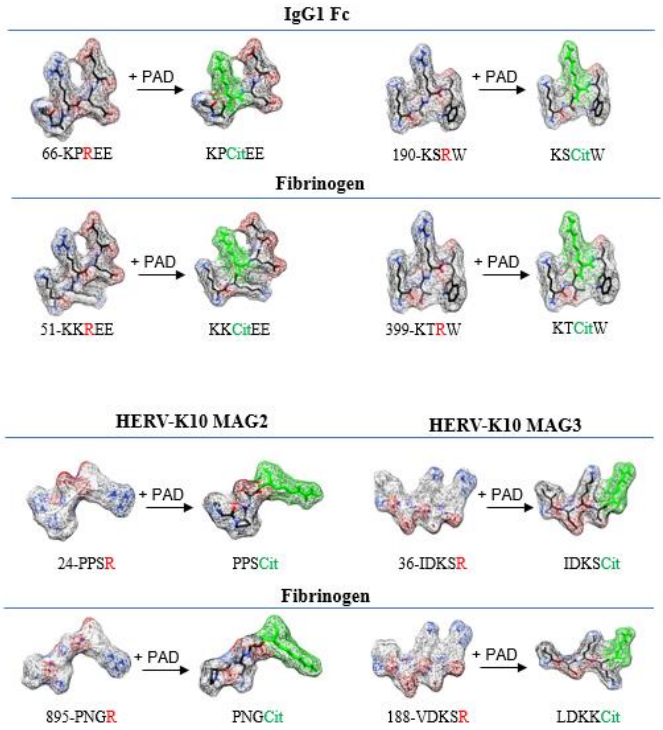


Figure 6.02. Molecular models demonstrating structural mimicry between HERV-K10, IgG1 Fc and fibrinogen in their native and citrullinated form. Three dimensional structures of regions of homology of IgG1 Fc versus fibrinogen and HERV-K10 MAG2 and MAG3 peptides versus fibrinogen were modelled in PyMOL and UCSF Chimera molecular graphics systems and modified using a PyTMs citrullination plugin. Numerals indicate amino acid starting position. Characters highlighted in red identify arginine residues susceptible to modification into citrulline residues (green). Structures to the left of the arrows indicate native epitopes and structures to the right of the arrows indicate their *in silico* modified counterparts. Citrullination is represented here with the symbol of Peptidyl Arginine Deiminase enzyme, PAD with citrulline residues highlighted in green.

To identify citrulline-specific probes, which could serve as positive controls for citrullination, we therefore assessed the reactivity of three commercial anti-citrulline polyclonal antibodies to commercially-sourced citrullinated fibrinogen by ELISA (Figure 6.03). We observed very limited reactivity of all citrulline-specific antibodies in this system (OD < 0.5) when compared to the anti-fibrinogen antibody response (OD = 2.6) which could result from the assay restriction phenomenon. Therefore, an antibody-independent chemical method, using a rhodamine-labelled phenylglyoxal probe that specifically detects citrullinated proteins, was employed in western blotting to identify any protein-bound citrulline residues (Figure 6.04). We found that commercial citrullinated fibrinogen and in-house citrullinated (i.e. using a PAD enzyme cocktail) fibrinogen and IgG1 Fc were detected in this system whilst their native (non-citrullinated) counterparts were not. This assay confirmed that the chemical probe was sensitive and specific for citrullinated residues, and that our in-house citrullination protocol was successful.

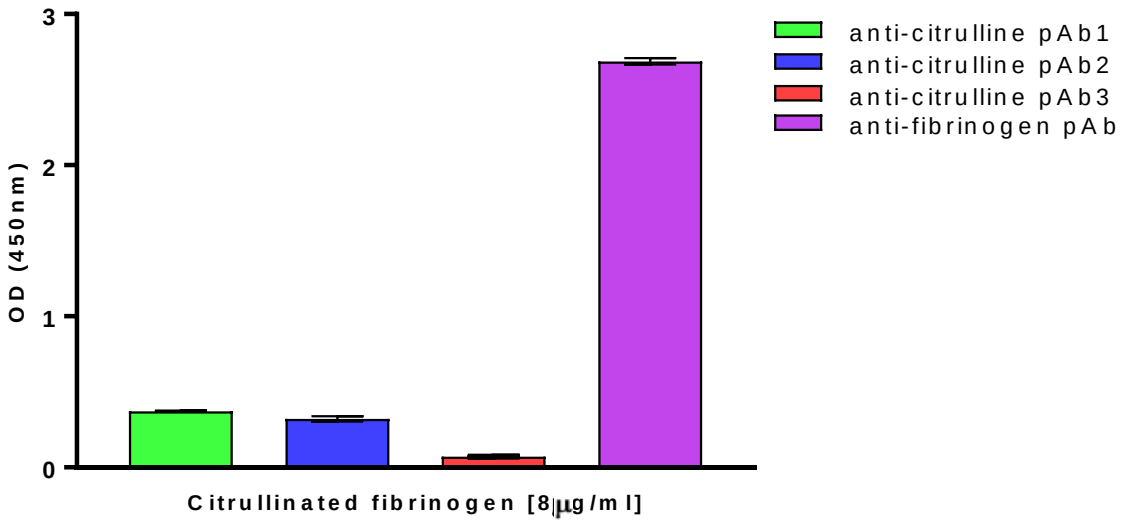


Figure 6.03. Poor reactivity of three commercial anti-citrulline polyclonal antibodies with citrullinated fibrinogen. Three citrulline-specific antibodies and fibrinogen-specific control antibody were incubated with citrullinated fibrinogen (coating antigen; 8 µg/ml) in ELISA system. Error bars show SEM for 3 experiments.

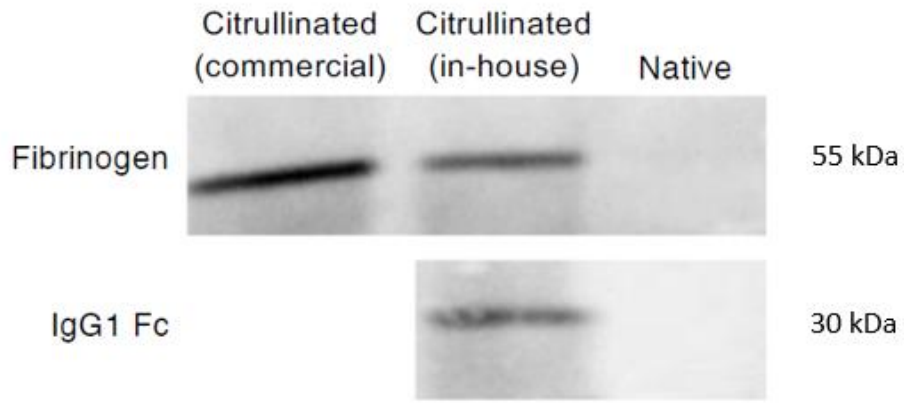


Figure 6.04. Verification of citrullination protocol with citrulline-specific probe on modified and native fibrinogen and IgG1 Fc. Native fibrinogen and IgG1 Fc were citrullinated in-house using a PAD enzyme cocktail (at 1:10 enzyme:protein ratio) or commercially (fibrinogen). Proteins were separated via SDS-PAGE electrophoresis and probed upon protein transfer onto a nitrocellulose membrane using a citrulline-specific reporter (rhodamine-phenylglyoxal; 100μM).

6.2.2. Serum reactivity to native and citrullinated autoantigens

Having established methods to induce and detect citrullination, we next wished to address the effect of this process on antibody binding. We considered that this change in peptide sequence and structure might have consequences for cross-reactivity in RA, and thus we set about investigating antibody responses to these potentially citrullinated targets in RA. To ensure that the tertiary structure of our antigens of interest (HERV-K10 MAG1 – 4, IgG1 Fc and fibrinogen) was not perturbed by our in-house citrullination protocol, we used peptide/protein specific polyclonal antibodies for detection of their corresponding targets in their native and citrullinated forms. We observed a slight decrease in anti-MAG1 antibody reactivity to its citrullinated versus unmodified target (difference in OD < 0.2) whereas anti-MAG2, anti-MAG3 and anti-MAG4 antibodies showed a significant decrease in reactivity (**, $p \leq 0.01$, ***, $p \leq 0.001$ and ***, $p \leq 0.001$, respectively) to their citrullinated targets (Figure 6.05). On the other hand, no such effect was observed for IgG1 Fc and fibrinogen.

Since citrullination has been shown to increase protein immunogenicity as a result of steric shift (Badillo-Soto *et al*., 2016), we considered the possibility that this could induce cross-reactivity of autoantibodies in RA. To assess this, we screened the reactivity of six RA serum samples to native and in-house citrullinated MAG1-4, IgG1 Fc and fibrinogen by ELISA. We found significantly increased RA antibody responses to all citrullinated targets, when compared with their native counterparts (Figure 6.06), except for IgG1 Fc which showed no such effect. The difference in immunoreactivity to native versus citrullinated targets was particularly noticeable for MAG1, MAG2 and MAG3 peptides (****, $p \leq 0.0001$, ****, $p \leq 0.0001$, ***, $p \leq 0.001$, respectively). Moreover, high levels of reactivity to the citrullinated antigens were predominantly exhibited by autoantibodies in the diseased group as serum antibodies from six healthy individuals showed generally low reactivity across all antigens (Figure 6.07). The

increased response within RA serum samples to the majority of modified epitopes containing citrulline residues suggested that they may be more favourably orientated towards RA autoantibodies, possibly due to the change in their charge, than their unmodified counterparts. We therefore studied the effect of this post-translational modification on the cross-reactivity profiles of RF and ACPA autoantibodies in greater detail.

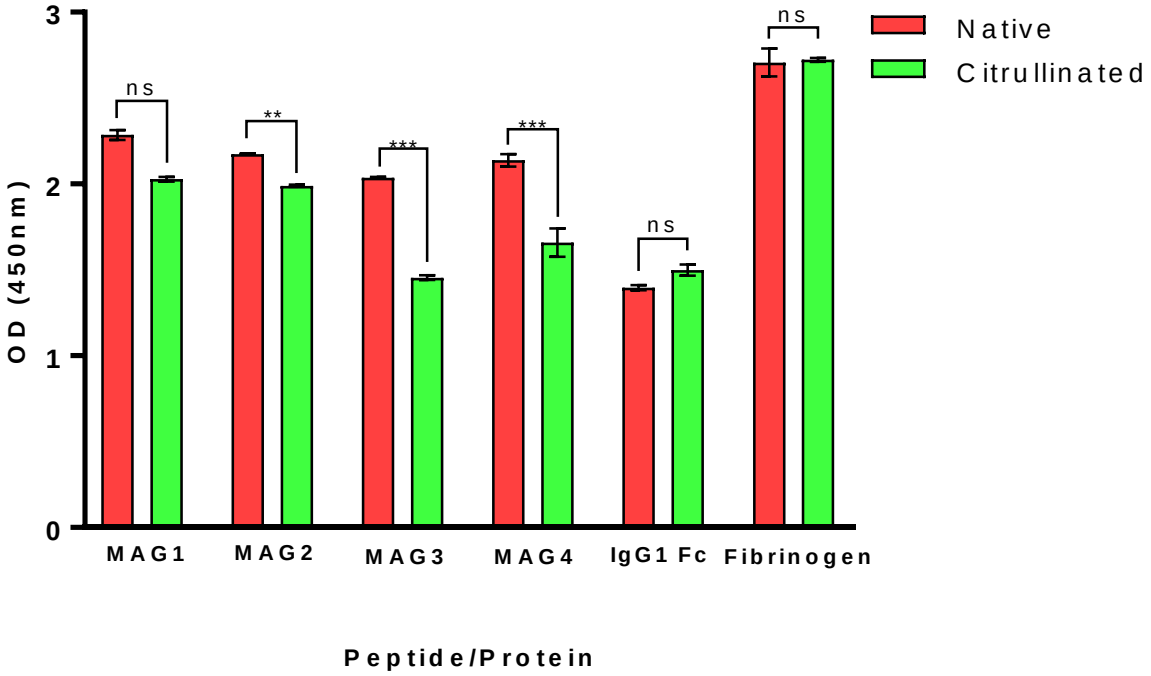


Figure 6.05. Reactivity of peptide- and protein-specific polyclonal antibodies to MAG1 – 4, IgG1 Fc and fibrinogen in their native and citrullinated forms. Peptide and protein coating antigens (8 µg/ml), pre- and post-treatment with a PAD enzyme cocktail, were incubated with their corresponding antibodies in ELISA system. SEM is shown for 3 experiments. **, $p \leq 0.01$, ***, $p \leq 0.001$, *ns* = not significant.

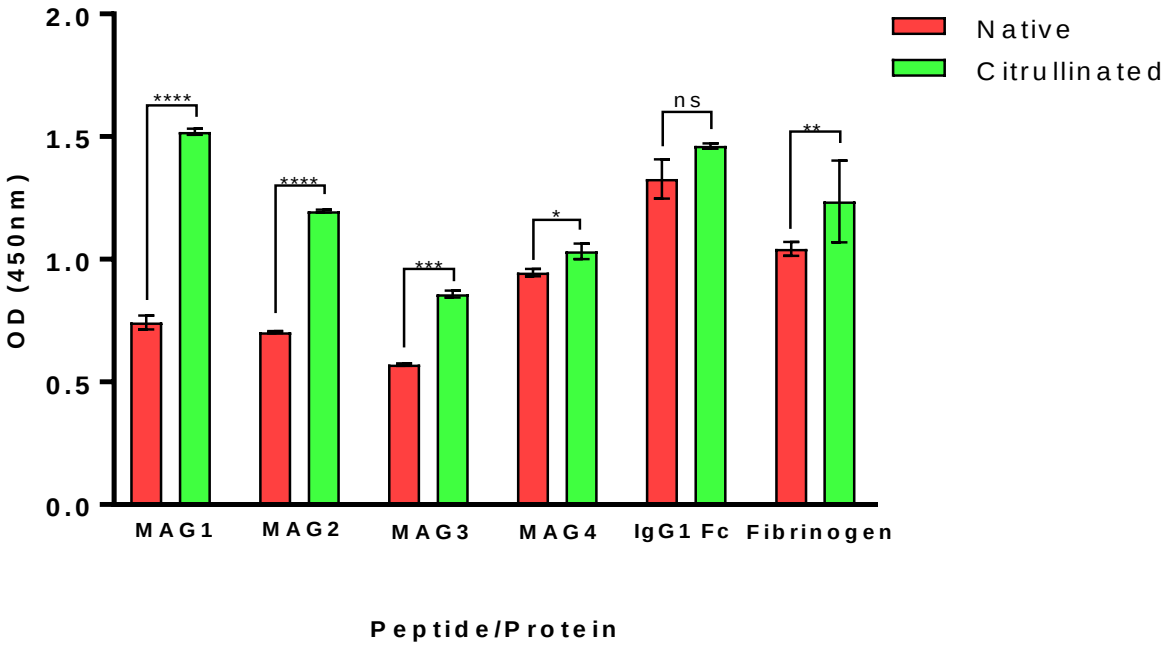


Figure 6.06. Increased mean serological response of six pooled RA patient serum samples to citrullinated versus native forms of MAG1 – 4 peptides, IgG1 Fc and fibrinogen. Crude pooled sera from RA patients were incubated with peptide and protein coating antigens (8 µg/ml), pre- and post-treatment with a PAD enzyme cocktail, in ELISA system. Bars represent means and SEM is shown for 3 experiments. *, $p \leq 0.05$, **, $p \leq 0.01$, ***, $p \leq 0.001$, ****, $p \leq 0.0001$, *ns* = not significant.

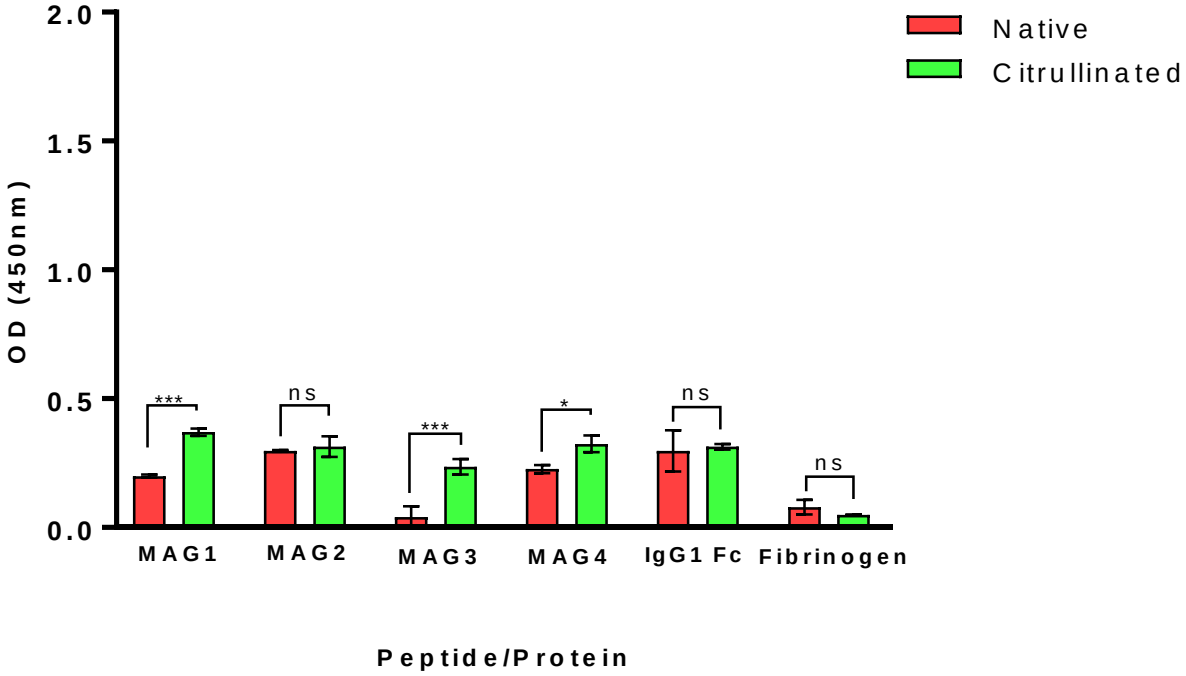


Figure 6.07. Low mean serological response of six pooled serum samples from healthy individuals to native and citrullinated forms of MAG1 – 4 peptides, IgG1 Fc and fibrinogen. Crude pooled sera from healthy controls were incubated with peptide and protein coating antigens (8 µg/ml), pre- and post-treatment with a PAD enzyme cocktail, in ELISA system. Bars represent means and SEM is shown for 3 experiments. *, $p \leq 0.05$, ***, $p \leq 0.001$, *ns* = not significant.

6.2.3. Serological response of RF and ACPA antibodies to native and citrullinated RA autoantigens

In chapter 5, we reported increased levels of reactivity of IgG antibody in RA sera to viral sequences homologous with the principal RF target (Figure 5.06), IgG1 Fc, which showed positive correlation with RF titres (Figure 5.08). This supported the notion of an augmented secondary immune response through mechanisms of molecular mimicry in RA. Moreover, earlier in this chapter we demonstrated considerable sequence and structural homology between the epitopes of HERV-K10 and IgG1 Fc with one of the major ACPA targets, fibrinogen, through *in silico* analysis (Figure 6.01). We therefore set out to analyse the cross-reactivity of RF and ACPA antibodies to IgG1 Fc and fibrinogen autoantigens in their native and citrullinated forms. We first sought to stratify RA patients based on RF and ACPA titres to explore if RF interacts with citrullinated residues, in the context of no ACPA. Thus, we recruited 42 RA patients with clinically pre-determined ACPA and RF status (positive / negative) and quantified the titres of both IgG autoantibodies by ELISA (Figures 6.08 and 6.09). Based on these data, sera were stratified into 9 samples containing ACPA and RF (ACPA+RF+), 9 with ACPA only (ACPA+RF-), 5 with RF only (ACPA-RF+), and 19 with neither autoantibody (ACPA-RF-) (Table 6.02). Determination of autoantibody positivity was based on manufacturer set cut-off values of 5 U/ml for ACPAs and 20 U/ml for RF.

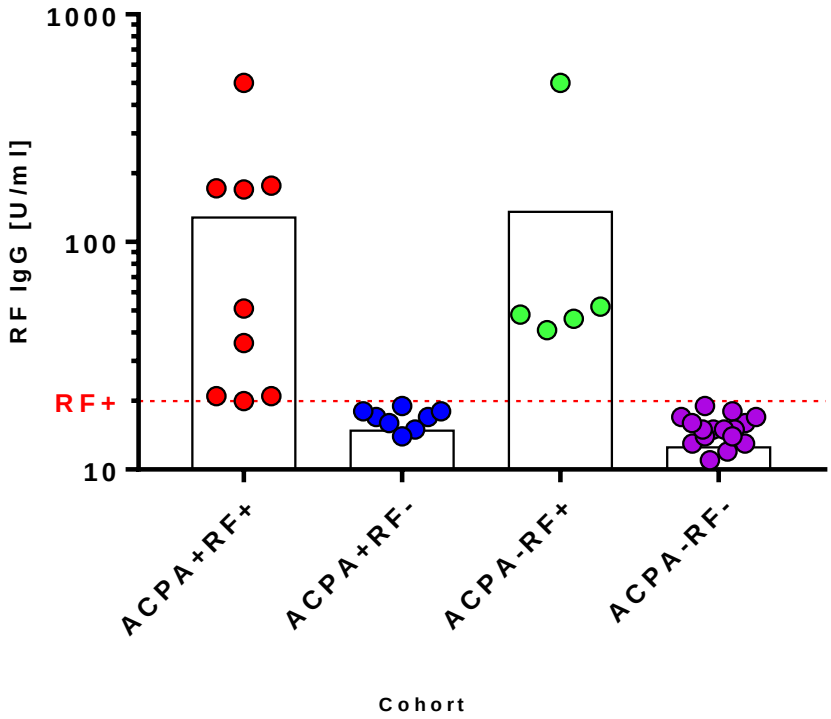


Figure 6.08. Rheumatoid factor IgG titres determined for the four RA sera sub-types. The presence of RF autoantibody titres in patient serum samples was assessed through the verification of reactivity against IgG Fc using a commercial sandwich ELISA kit**.** Threshold for RF+ was set at 20 U/ml (red-dashed line). Bars represent means.

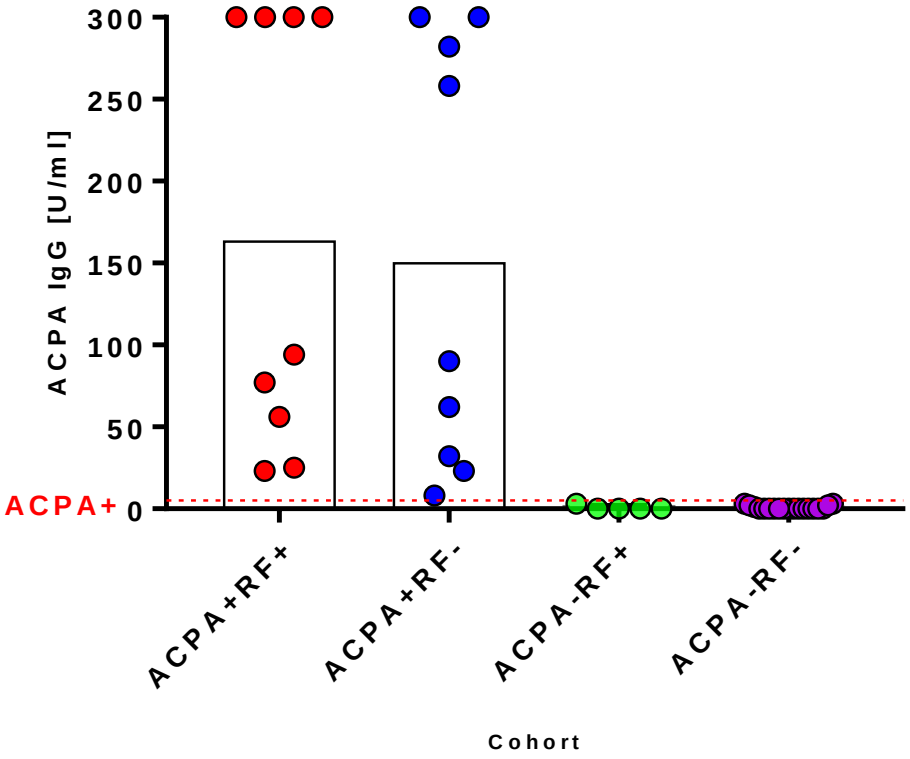


Figure 6.09. Anti-citrullinated protein antibody IgG titres determined for the four RA sera sub-types. The presence of ACPA titres in patient serum samples was assessed through the verification of reactivity against cyclic citrulline peptide using a commercial ELISA kit**.** Threshold for ACPA+ was set at 5 U/ml (red-dashed line). Bars represent means.

	ACPA+ RF+	ACPA+ RF-	ACPA- RF+	ACPA- RF-
Number of subjects	9	9	5	19
Females	7	5	2	16
Males	2	4	3	3
Age, mean years (SD)	61.1 (7.4)	64 (7.7)	60.8 (9.9)	63.05 (10.2)

Table 6.02. Stratification of rheumatoid arthritis patients according to serum titres of RF and ACPAs. Patient characteristics amongst the four sera sub-types (ACPA+RF+, ACPA+RF-, ACPA-RF+, ACPA-RF-) determined by protein-based commercial ELISA kits.

Having stratified the patients based on their RF and ACPA profiles, we next wished to assess the cross-reactivity of these sera subtypes with citrullinated and native fibrinogen by western blotting (Figure 6.10). As expected, sera containing ACPAs (ACPA+RF+) reacted strongly with citrullinated fibrinogen (top row). Despite the significant sequence and structural mimicry between IgG1 Fc and fibrinogen revealed in our *in silico* analysis, no detectable reactivity was demonstrated for sera containing RF alone (ACPA-RF+) with native fibrinogen. However, when fibrinogen was citrullinated we observed binding of RF+ sera, even in the absence of detectable ACPAs (ACPA-RF+; second top row), suggesting that citrullination facilitates cross-reactivity of RF with fibrinogen. This effect was RF-dependent, as reactivity with citrullinated fibrinogen was abolished in ACPA-RF- sera from RA patients or healthy controls. Based on the above findings, we proceeded to confirm these data by ELISA using fibrinogen or IgG1 Fc in their native or citrullinated forms as coating antigens. As expected, we found that ACPA-RF+ sera reacted with both unmodified and citrullinated IgG1 Fc, with marginally enhanced reactivity reported for the latter form (Figure 6.11). Whereas, unsurprisingly, ACPA+RF- sera exhibited significantly higher reactivity with citrullinated versus native IgG1 Fc (****, $p \leq 0.0001$). Accordingly, no serological response with IgG1 Fc was demonstrated by the ACPA-RF- patient cohort or healthy controls. We also observed limited reactivity of ACPA+RF- and ACPA-RF+ sera with native fibrinogen (Figure 6.12). In agreement with the data from western blotting, citrullination facilitated the significant increase in binding of ACPA-RF+ antibody sera with non-target fibrinogen. Consistent with this effect being RF-dependent, no serological response with citrullinated fibrinogen was observed for the ACPA-RF- cohort or healthy controls. Consequently, these findings indicate that RF is able to cross-react with the ACPA target fibrinogen, and that this response is critically dependent on protein citrullination.

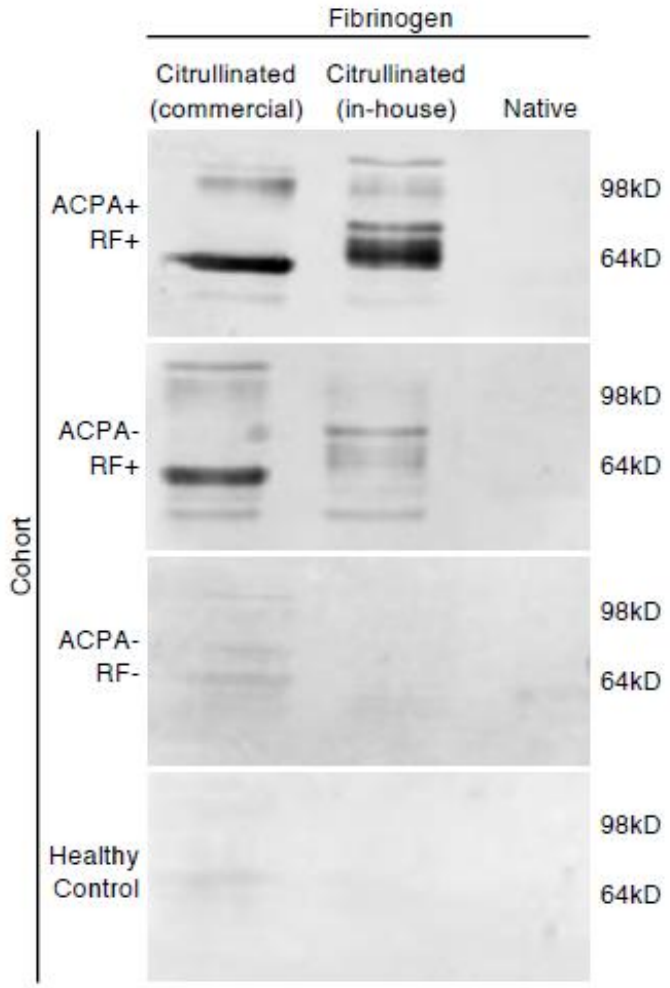


Figure 6.10. C**ross-reactivity of RF+ serum with citrullinated fibrinogen in the absence of ACPAs.** Native fibrinogen was citrullinated in-house using a PAD enzyme cocktail (at 1:10 enzyme:protein ratio) or commercially (fibrinogen). Native and citrullinated fibrinogen was separated via SDS-PAGE electrophoresis and probed upon protein transfer onto a nitrocellulose membrane using sera from RA patient cohorts with different autoantibody profiles and healthy controls.

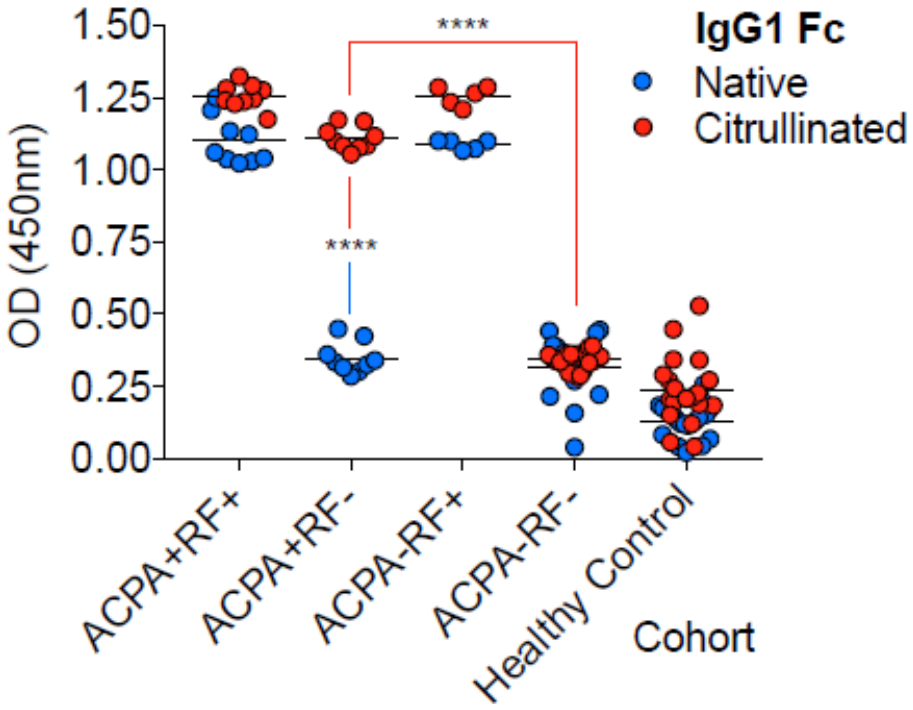


Figure 6.11. Citrullination facilitates cross-reactivity of ACPA+ serum with IgG1 Fc in the absence of RF in ELISA. IgG1 Fc coating antigens (at 8 µg/ml), pre- and post-treatment with a PAD enzyme cocktail, were incubated with sera from RA patient cohorts with different autoantibody profiles and healthy controls. Black lines represent means. Coloured lines indicate statistical comparison between particular datapoints. ****, $p \leq 0.0001$.

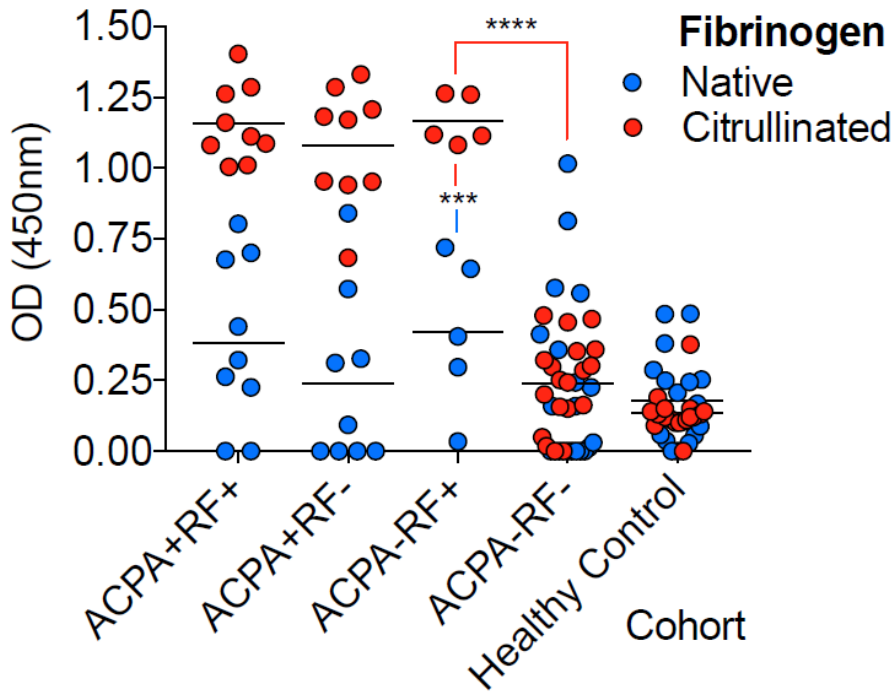


Figure 6.12. Citrullination facilitates cross-reactivity of RF+ serum with fibrinogen in the absence of ACPAs in ELISA. Fibrinogen coating antigens (at 8 µg/ml), pre- and post-treatment with a PAD enzyme cocktail, were incubated with sera from RA patient cohorts with different autoantibody profiles and healthy controls. Black lines represent means. Coloured lines indicate statistical comparison between particular datapoints. ***, $p \leq 0.001$, ****, $p \leq 0.0001$.

We have previously identified the homologous sequences 44-HEDPEV / 97-HADPDL and 132-DELTK / 128-DELNN between IgG1 Fc and fibrinogen. Since these regions do not contain arginine residues, they are not susceptible to citrullination, and therefore are not targeted by ACPAs. This implies that when the arginine-containing sequences 190-KSRW / 399-KTRW and 66-KPREE / 51-KKREE are citrullinated, there remain accessible epitopes even in the presence of ACPAs which RF may target. To determine potential RF reactivity with arginine-free homologous regions, we established sequential ELISAs coated with citrullinated fibrinogen or IgG1 Fc that were first incubated with sera containing ACPAs alone, to reveal maximum reactivity of ACPA to all available citrulline residues. To saturate any remaining unbound citrulline residues and identify potential binding to citrulline-free regions, we then added sera containing either ACPA or RF, respectively. Results from this experiment revealed that for citrullinated IgG1 Fc (Figure 6.13 A), sequential addition of ACPA+RF- sera followed by ACPA-RF+ sera resulted in significantly higher reactivity than was reported with RF- sera. Similar effect was observed for citrullinated fibrinogen (Figure 6.13 B). Nevertheless, sequential incubation with primary ACPAs and secondary RFs in this manner could lead to RF binding with the Fc portion of ACPA autoantibodies, which could be responsible for the increase in reactivity reported. To this end, we addressed this possibility by repeating these assays using ACPA+RF- sera in the primary step that had been digested with pepsin to cleave Fc regions generating ACPA $F(ab')_2$ fragments. In these assays, sequential incubation with ACPA $F(ab')_2$+ sera and sera containing ACPAs alone lead to no significant increase in response, implying that all free ACPA target epitopes were saturated in the primary step (Figure 6.13 C and D). However, when ACPA-RF+ sera were used in the secondary step we observed a significant increase in reactivity for both IgG1 Fc (**, $p \leq 0.01$) and fibrinogen (**, $p \leq 0.01$). For fibrinogen, these data implied that RF is able to interact with citrullinated

fibrinogen even when in competition with ACPAs, by targeting additional non-ACPA target epitopes. This may explain the small but detectable increase in reactivity of ACPA-RF+ sera versus ACPA-RF- sera with native fibrinogen (blue dots).

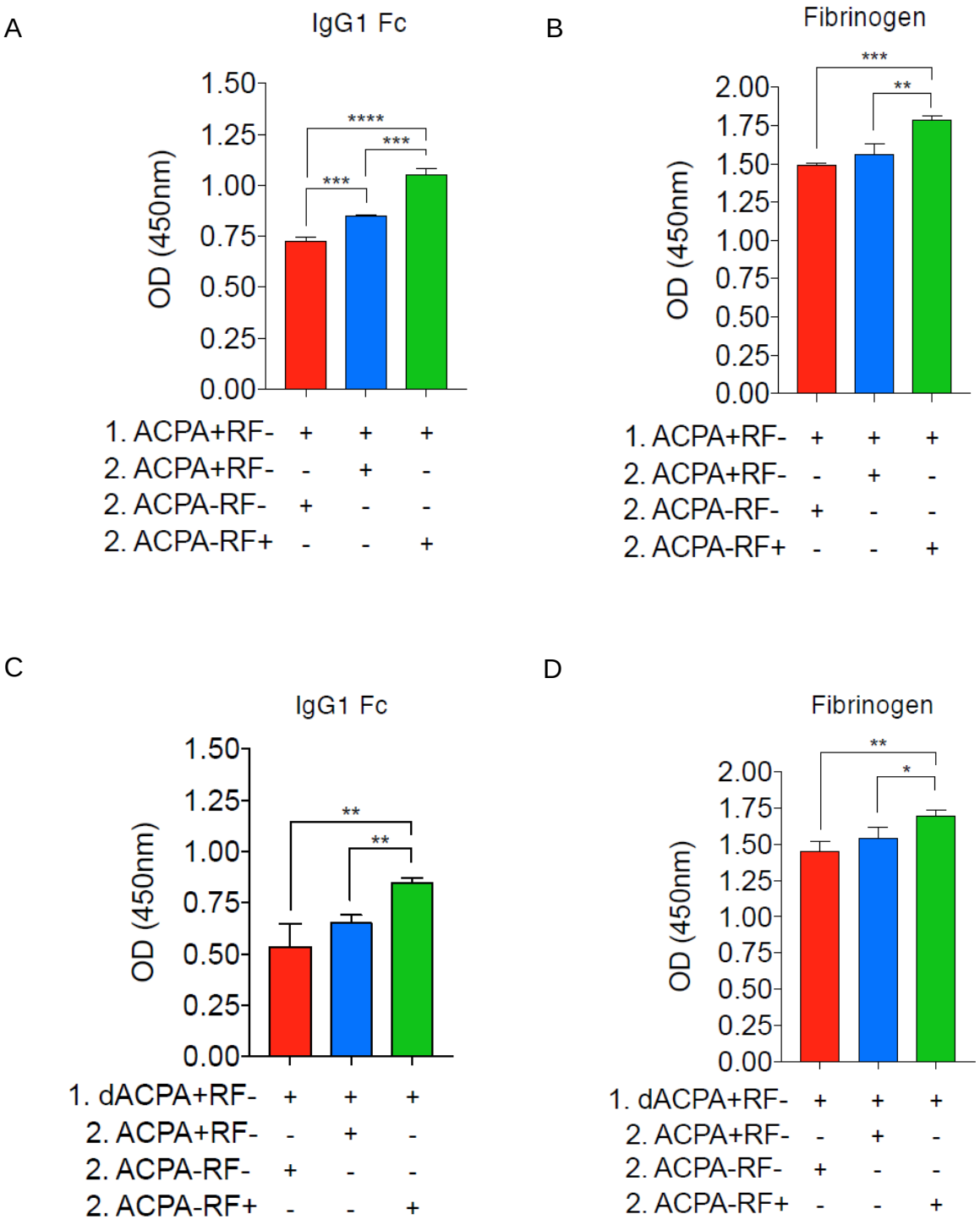


Figure 6.13. RF reacts with additional epitopes on citrullinated fibrinogen and IgG1 Fc in the presence of ACPAs. Graphs show serological response to citrullinated IgG1 Fc (**A**) or fibrinogen (**B**) coating antigens (8 µg/ml) when incubated with ACPA+ sera (step 1), followed by incubation with the indicated RA patient sera-subtypes (step 2), determined by sequential ELISA. The same experiments were repeated using ACPA+ sera digested with pepsin in step 1 (dACPA+RF-) for citrullinated IgG1 Fc (**C**) or fibrinogen (**D**). Bars represent means and SEM is shown for 3 experiments. **** $p \leq 0.0001$; *** $p \leq 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$.

6.3. Discussion

Although the pathogenesis of RA has not been completely elucidated, genetic predisposition, environmental insults and viral pathogens are considered to be contributory factors in the destruction of the joint. This can be mediated through RA autoantibodies reactive with distinct antigen targets such as IgG1 Fc (RF Ab) and citrullinated proteins (ACPA Ab), which form immune complexes (IC) of a pathological nature in the context of impaired clearance mechanisms (Van Steendam *et al*., 2010). Moreover, molecular mimicry and cross-reactivity have the potential to trigger auto-reactivity in RA leading to epitope spreading and the loss of self-tolerance (Cornaby *et al*., 2015).

We have previously demonstrated significant sequence and structural mimicry between HERV-K10 viral epitopes and those on the Fc region of IgG (chapter 3; Table 3.06 & 3.12). This finding was subsequently confirmed *in vitro* where we observed immunological cross-reactivity of anti-K10 antibody with IgG1 Fc (chapter 4; Fig. 4.07) and increased serum antibody reactivity profiles among RA patients to HERV-K10 matrix peptides (chapter 5; Fig. 5.20).

In this study we used our optimised *in silico* system to assess potential homology of HERV-K10 Gag peptides and IgG1 Fc autoantigen with fibrinogen, which constitutes one of the major ACPA targets and is present in the IC of RA joints (Sokolove *et al*., 2011). In total, four regions of homology were found for HERV-K10 versus fibrinogen (GKEL/CK/E, PP/NS/GR, I/LDKS/KR, K/VVST/GKN/EL/C) and for IgG1 Fc versus fibrinogen (KP/KREE, KS/TRW, DELT/NK/N, HE/ADPE/DV/L) which possessed between one and three amino acid substitutions (Fig. 6.01). Interestingly, our group has previously identified one of these regions in IgG1 Fc (KPREE) to be a potential major immunodominant fragment targeted by RF (Nelson *et al*., 2003). Despite minor

sequence variations between the mimics, molecular modelling of the peptides revealed very little difference in overall mesh structure which was in agreement with our previous report (Tugnet *et al*., 2013). Among the predicted regions of homology, four paired peptide alignments contained arginine residues in the sequences of HERV-K10, IgG1 Fc and fibrinogen, identifying them as targets for citrullination. Strikingly, we found no significant effect of citrullination on the peptide tertiary structure when arginine residues were converted into citrulline residues and displayed as molecular models (Fig. 6.02). This finding prompted our *in vitro* studies, which confirmed that citrullination of protein-bound arginine residues did not perturb the structure of immunoreactive epitopes and might in fact lead to neoepitope generation and enhanced antibody response (Kidd *et al*., 2008). Indeed, in our preliminary assay we found that RA serum reacted more readily with citrullinated forms of all HERV-K10 peptides, IgG1 Fc and fibrinogen (Fig. 6.06), while this phenomenon was not reported with serum from healthy individuals (Fig. 6.07). Therefore, our observation implied that antigen citrullination in RA might broaden the effect of epitope spreading and molecular mimicry between the virus and host's proteins with respect to autoantibody reactivity. This is consistent with the findings of Tsuda *et al*. (2015), who demonstrated extensive cross-reactivity of a monoclonal ACPA antibody with a variety of targets such as fungal, plant and viral antigens including EBV-derived EBNA1 epitopes, which are homologous with several RA autoantigens. Moreover, it supports the notion that RA patient's autoantibody repertoire undergoes epitope spreading between citrullinated antigens as reported by Turunen *et al*. (2015), who demonstrated antibody binding to both citrulline and homocitrulline-containing type I and II collagen carboxyterminal telopeptides in RA serum.

Owing to the difference in antibody responses observed in our study towards native and citrullinated forms of HERV-K10 and the two established RA autoantigens, IgG1

Fc and fibrinogen, we wondered whether the arginine-rich regions of homology between these autoantigens would allow for cross-reactivity of RF with citrullinated ACPA target proteins. To test whether reactivity of patient sera with citrullinated fibrinogen could be specifically attributed to RF, we aimed to isolate RF+ sera from RA patients that had no detectable ACPAs. This approach allowed us to demonstrate the binding of RF+ sera to citrullinated but not native fibrinogen (Fig. 6.12), even in the absence of detectable ACPAs, which implies that citrullination facilitates cross-reactivity of RF with fibrinogen. This reactivity was reported for both linear and conformational citrullinate-modified epitopes of fibrinogen and relied on the presence of RFs in RA serum as we did not find any response in sera from ACPA-RF- patients. Therefore, our finding provides direct evidence that RF is cross-reactive with fibrinogen and that this response relies on citrullination. Others have also reported preferential antibody responses to citrullinated sequences with altered peptide binding affinity (James *et al*., 2010), which may explain why we found more limited reactivity of RF+ RA sera with native fibrinogen, despite its mimicry with IgG1 Fc. Moreover, we also demonstrated that RF is able to interact with citrullinated fibrinogen even when in competition with ACPAs, by targeting additional non-ACPA target epitopes (Fig. 6.13 D). As anticipated ACPA+RF- sera exhibited strong reactivity with fibrinogen in its citrullinated form only, which confirmed that modified fibrinogen is indeed a target of ACPAs in RA (Tsuda *et al*., 2015; Steiner Gunter, 2006) and that the RF response was specific. Consequently, levels of reactivity to both forms of fibrinogen were redundant in seronegative patients and comparable to those observed for healthy individuals.

Clearly, given the complex nature of autoimmune reactions and the heterogeneity of human sera, certain amount of autoantibodies could be present even in RA serum samples assessed to be ‘seronegative’ by means of commercially optimised autoantibody assays (Ajeganova and Huizinga, 2014). As reported by Fernandes-

Cerqueira *et al*. (2015) in their study only between 4 and 8% of CCP-negative patients showed reactivity against citrullinated fibrinogen peptides. To our knowledge, no such observations were made for RF assay. Since our study focused on established (secondary) immune response we determined serum IgG-RF levels by means of a standardised assay, however it cannot be excluded that IgM-RF and/or IgA-RF interfere with IgG-RF binding. In the case of ACPA, autoantibody levels are detected using synthetic cyclic citrullinated peptides (CCPs) of unknown sequence (Ossipova *et al*., 2014), for which biological relevance has not been elucidated. Nevertheless, affinity-purified anti-CCP antibodies have been demonstrated to be in fact a collection of different ACPAs and are able to recognise citrullinated fibrinogen (Ioan-Facsinay *et al*., 2011). Moreover, since CCPs do not correspond to any human peptide sequence (Ossipova *et al*., 2014), it may explain the lack of reactivity of ACPA-RF+ sera in the anti-CCP test despite the citrullinated target being present. Although it is plausible that RA autoantibody levels may change over time, Barra *et al*. (2011) reported that there was a minimal change in RF and ACPA positivity up to five years of follow-up whilst Ursum *et al*. (2010) and Mjaavatten *et al*. (2011) demonstrated that RF and ACPA levels are stable within the first year after RA onset.

Interestingly, we also observed increased reactivity of sera containing ACPAs alone with conformational epitopes on citrullinated IgG1 Fc but not on native IgG1 Fc (Fig. 6.11). Although ACPAs are recognised as partly cross-reactive antibodies binding other citrullinated proteins such as vimentin or α-enolase (Valesini *et al*., 2015), there are no previous reports in the literature of their reactivity to IgG1 Fc. IgG has been recently reported to undergo citrullination in RA synovium and therefore could constitute an ACPA target (Badillo-Soto *et al*., 2016). Recognition of numerous citrullinated autoantigens might suggest that no traditional antigenic determinant is targeted by ACPAs and that the culpable autoantigens can indeed change over time

instigating new reactivities in the ACPA responses (Trier *et al*., 2015). Therefore, antibodies generated against citrullinated epitopes could sustain the immune response via epitope spreading and cross-reactivity with broad array of citrullinated proteins of endogenous and exogenous origin. In addition, we observed an enhanced response of ACPA-RF+ serum antibodies to citrullinated versus unmodified IgG1 Fc, which is consistent with other reports demonstrating that citrullination can improve protein immunogenicity as a result of steric shift and effect on relaxation dynamics of citrullinated proteins (Badillo-Soto *et al*., 2016). Moreover, Pratesi *et al*. (2013) reported that citrullinated peptides fit better in the antigen binding grooves than the corresponding arginine containing sequences and that RA-associated HLA DRB1*0401 (shared epitope) has a higher affinity for citrulline-modified peptides than for their native counterparts. Accordingly, we found no reactivity with citrullinated IgG1 Fc in the absence of ACPAs in RF- serum samples.

Over the past decade, a model for the initiation of RA pathogenesis has been proposed which implies that ACPAs react with citrullinated autoantigens, forming ICs containing Fc fragments that stimulate the generation of RF (Tan and Smolen, 2016). According to this model, RF relies on ACPAs and should be expressed after the ACPA reactivity is initiated. However, this model does not consider other RA autoantibodies, in particular RF seropositivity in ACPA- patients. Extending the model to include immune complexes containing viral and other non-citrullinated antigens reverses this order, with RF generated in response to molecular mimics, which lead to tolerance breakdown to other autoantigens, inflammation, release of PAD enzymes, hypercitrullination and onset of the ACPA response. In this model, ACPAs are therefore produced as an indirect result of the RF response. Soluble complexes containing RFs have been previously suggested to facilitate ACPA access to RA joint or to intensify ACPA effector functions (Ioan-Facsinay *et al*., 2011; Laurent *et al*.,

2015). RFs have also been reported to increase the capacity of ACPA-ICs to promote macrophage cytokine generation (Sokolove *et al*., 2014). There are numerous reports which investigate the generation of autoantibodies in the years before the onset of symptoms in RA, and it is certain that both RF and ACPAs can precede clinical manifestations of the disease (Masi, Aldag and Sipes, 2001; Nielen *et al*., 2004; Brink *et al*., 2013). Nevertheless, there are conflicting reports regarding RF/ACPA time of occurrence before the disease onset (Song and Kang, 2010a; Rantapää-Dahlqvist *et al*., 2003; Nielen *et al*., 2004), which may be due to different sensitivities and specificities of the assays used to detect these autoantibodies (Barra *et al*., 2011; Hoovels *et al*., 2018). Moreover, limited longitudinal analyses address IgG RF and ACPA together (Brink *et al*., 2016). Although, there is no direct evidence as to whether either antibody precedes the other, implying that both of the above scenarios likely transpire to some degree, several reports suggested that RF may give rise to subsequent ACPA response in RA, particularly in patients with shared epitope allele and in smokers (Hutchinson *et al*., 2016; Murphy, Mattey and Hutchinson, 2017).

In this chapter, we have demonstrated that citrullination results in enhanced antibody responses to viral and host's antigens which were shown to contain multiple arginine-rich regions of sequence and structural homology. This observation supports the premise of molecular mimicry and neoantigen production in RA that perpetuate the autoimmune response due to epitope spreading. We also reported that protein citrullination resulted in significant cross-reactivity of RF autoantibodies with a common ACPA target, fibrinogen. Importantly, this finding suggests the possibility that RF could act directly as an ACPA precursor, making it a potential therapeutic target to limit the ACPA response and inhibit inflammation.

7. FINAL DISCUSSION AND FUTURE PERSPECTIVES

In this thesis we have explored molecular mimicry mechanisms between endogenous proteins of viral and host origin underlying the immune response in RA. These have significant implications for our understanding of how antigen-driven reactivity to viral epitopes can trigger disease pathogenesis, and perpetuate the autoimmune reactions observed in rheumatoid arthritis (RA).

As highlighted in chapter 1, RA is characterised by breakdown in immunological tolerance to self-antigens in B- and T-cell populations, which is detectable years prior to the onset of symptoms. This may be a consequence of several mechanisms including molecular mimicry, epitope spreading and bystander activation of autoaggressive immune cells by microbial agents, which by priming the cells might provide a 'fertile field' but no apparent disease (Cusick, Libbey and Fujinami, 2012). Together with genetic predisposition, further exposure to environmental insults could trigger self-reactive responses and result in autoimmune conditions. Due to the limitations of existing animal models, this process is still unclear. Understanding endogenous pathognomonic mechanisms which lead to auto-reactivity might shed some light on RA development and contribute to reestablishment and maintenance of tolerance to self. This thesis aimed to investigate the role of endogenous retrovirus K10 in RA autoimmunity by taking advantage of *in silico* and *in vitro* approaches which have been previously used to establish a relationship between HERVs and autoimmunity in other disorders such as MS and type 1 diabetes (Curtin *et al*., 2018; do Olival *et al*., 2013).

In previous chapters we have suggested that the sequence and structural homology between the human endogenous retrovirus (HERV) K10 Gag matrix and IgG1 Fc, the

principal target of rheumatoid factor autoantibodies in RA, might play a significant role in disease pathogenesis. HERVs, which constitute 8% of the human genome as a result of DNA-invading ancestral infections passed in the germline through successive generations (Yu, Zhao and Zhu, 2013; Krzysztalowska-Wawrzyniak *et al*., 2011), have been implicated in autoimmune diseases such as multiple sclerosis (Ramasamy, Joseph and Whittall, 2017) and systemic lupus erythematosus (Nelson *et al*., 2014b). Endogenous retroviruses have also been repeatedly suggested as aetiological factors of autoimmune rheumatic conditions and antibodies to HERVs have been detected in patients with RA (Freimanis *et al*., 2010). The work presented in this thesis suggests that HERV-K10 Gag may contribute to the aetiology and pathogenesis of rheumatoid arthritis, via molecular mimicry mechanisms, which could have significant implications for our understanding of the role of HERVs in other autoimmune diseases.

As a component of the human genetic material, HERV-derived polypeptides should be regarded as autoantigens and tolerated by immune cells. Nevertheless, if not expressed in the thymus during tolerance acquisition, they could be treated as neoepitopes upon reactivation and may therefore represent an early element in the chain of events that leads to tolerance breakdown to other self-antigens (Gröger and Cynis, 2018; Balada, Ordi-Ros and Vilardell-Tarrés, 2009; Herve *et al*., 2002). In this light, the first question that was posed in this thesis was related to the homology between the immunodominant peptides on HERV-K10 matrix and those on IgG1 Fc autoantigen, targeted by rheumatoid factor antibody. This revealed six linear B cell epitopes on HERV-K10, coinciding with our T cell epitope predictions, with four epitopes (MAG1-4) corresponding to established RF binding regions. This could suggest that K10, which was reported to produce viral particles and products in RA joint (Löwer *et al*., 1993; Stransky *et al*., 1993), had the potential to evoke autoimmune antibody responses observed in RA patients. On the other hand, the structure-

dependent nature of antigen:antibody binding could affect the performance of current predictive approaches for B-cell epitopes, which are typically mapped for linear motifs using *in silico* sequence analysis. In addition, our current understanding of protein immunogenicity and B-cell epitope characteristics is limited, which can lead to suboptimal results and erroneous predictions (Kringelum *et al*., 2013). In comparison to its parallel field of T-cell epitope prediction, mapping of B-cell epitopes is less traceable and in need of standardisation (El-Manzalawy and Honavar, 2010). However, with the use of multiple platforms, the performance of *in silico* methods have been reported to provide the preliminary groundwork for subsequent *in vitro* investigations of HERV mimics in autoimmunity (Freimanis *et al*., 2010; do Olival *et al*., 2013; Nelson *et al*., 2014a).

Accordingly, upon identification of highly antigenic epitopes on HERV-K10 (MAG1-4), which exhibited significant homology with IgG1 Fc autoantigen in RA, serological methods were employed in chapters 4 and 5 to investigate the association of K10 with the disease. Due to the paucity of antibody reagents to HERV-K10, we generated an anti-MAG1 polyclonal rabbit antibody and validated it for specificity, selectivity and reproducibility in several immunoassays including ELISA, western blotting, immunocytochemistry and flow cytometry. This probe enabled us to demonstrate significant cross-reactivity with IgG1 Fc which prompted subsequent serological studies with patient serum samples. Evidently, a panel of standardised antibody reagents to this viral region, would find useful applications in the development of diagnostic immunoassays and research work. It would be of interest to generate and standardise a monoclonal reagent of stringent specificity to this epitope which could provide novel approaches to investigating and treating potential K10-induced conditions, comparable to the technology recently introduced in MS with anti-HERV-W monoclonal antibody (GNbAC1) (Curtin *et al*., 2016). Moreover, co-detection of anti-

HERV-K10 antibodies together with the current RF and anti-CCP tests could provide additional clinical value in the diagnosis of RA patients and contribute to RA profiling. The ability to directly assess anti-HERV-K10 specificity in a high-throughput test could facilitate the identification of RA subtypes, and therefore the HERV pathogenic involvement together with the disease prognosis and response to treatment. Like in the case of anti-HERV-W GNbAC1 antibody in MS, anti-HERV-K10 antibody of unique specificity could be employed to diminish or prevent the K10-mediated immune responses to self.

Our serological investigation revealed a significantly elevated mean IgG antibody reactivity to all synthesised HERV-K10 peptides MAG1 – 4 in patients with RA as compared to patients with OA, PsA, SLE and healthy individuals. These data support the notion of molecular mimicry between K10 virus and IgG1 Fc, and an association with RF binding epitopes. Moreover, as suggested in chapter 5 and by Herve and colleagues (Herve *et al*., 2002), the observed mature immune response to multiple K10 epitopes is a collectively sustained antigen-driven response, which could contribute to epitope spreading to other autoantigens. This may have a bearing on the titre and affinity of RF, therefore further analyses of the kinetics of interactions between K10 peptides and RF autoantibodies in patients with RA using surface plasmon resonance technology should prove interesting.

Furthermore, pathological changes and the site of inflammation in RA are mainly reported within the affected joints, therefore it would be of interest to broaden the investigation by assessing the extent of anti-HERV-K10 reactivity in RA synovial fluid. It has been previously reported that serum autoantibody levels may decrease with time as they are progressively sequestered within the cartilage following continuous synovial damage (Whittingham, Stockman and Rowley, 2017). Our findings in a cohort of 20 RA patients, who donated their serum samples at several time points, were

inconclusive in this respect. Therefore, changes in anti-MAG1 levels should be explored in longitudinal studies of a larger cohort of matched serum and synovial fluid samples to identify any differences or similarities between circulating and synovium-localised viral antibodies and their potential correlation with RA disease activity scores.

The processes that lead to the breakdown of immune tolerance, progression to active RA and joint destruction are highly dynamic and are likely to be co-dependent. Understanding the interaction of these mechanisms is vital as it may facilitate suitable diagnostic and prognostic opportunities. In fact, clarifying the links between HERV-K10 and the breach of tolerance may have implications relating to both RA aetiology and therapeutic strategies. Indeed, based on findings from this investigation, it could be possible that the initial trigger of autoimmune reactions, most likely occurring in the secondary lymphoid tissues, is mediated by structural homology between the K10 epitopes and host antigens such as IgG1 Fc which can lead to the production of RF autoantibodies. These antibodies could become cross-reactive, first with unmodified homologous autoantigens (e.g. fibrinogen) and subsequently with their citrullinated counterparts as a result of epitope spreading upon hypercitrullination. This may promote preferential binding of RFs with citrulline-containing protein targets even in the apparent absence of ACPA antibodies (chapter 6). RF could then lead to the initiation and propagation of ACPA responses (Fig. 7.01). Therapeutic targeting of RF and anti-HERV-K10 response may therefore represent a route to limiting RA-associated pathology. As soluble blocking peptides have been previously shown to reduce antibody-mediated damage in MS and cardiomyopathy without generating harmful immune complexes (Fridkis-Hareli, 2013; Münch *et al*., 2012), the inhibitory effect of MAG peptide mimics on RA autoantibodies could be explored.

Collectively, these findings suggest that HERV-K10 epitopes through molecular mimicry with self-antigens could result in breakdown of immune tolerance, leading to

increased immune reactivity to the virus and predisposing to autoimmunity in RA. The effect of these mimics on the autoantibody response might in part explain the variation of the phenotype of RA. Moreover, it is conceivable that the mimicry between viral and self-proteins occurs at a 'disease-related' epitope. Without this, autoimmune responses may be triggered but no pathology transpires. The heterogenous nature of anti-MAG response observed from our data suggests that HERV-K10 epitopes might be immunogenic only in a specific subset of RA patients, who are genetically predisposed and have been exposed to the right environmental factors. This is also in agreement with the previously reported HERV-K polymorphism in the human population and the premise of specific disease-associated unfixed viral loci among individuals (Marchi *et al*., 2014), which could contribute to the observed differences in anti-K10 responses. Therefore, our findings imply that RA does not constitute a discrete clinical entity but an umbrella term for a number of syndromes.

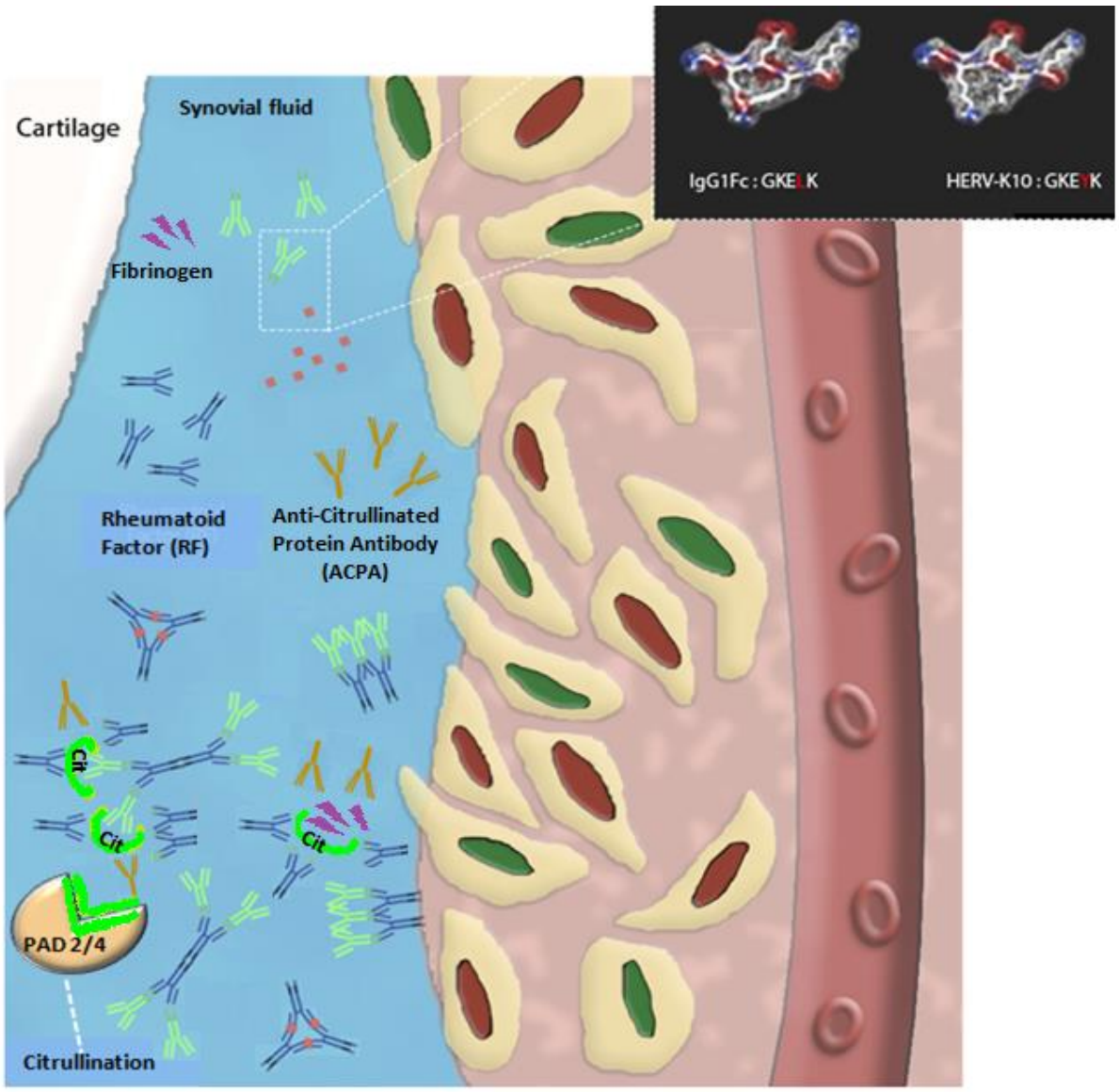


Figure 7.01. A model for the initiation of rheumatoid pathology in RA synovium proposed in the current study. Molecular mimicry between HERV-K10 and host IgG1 Fc proteins triggers the anti-HERV-K10 response and the production of RF autoantibodies. Subsequent formation of immune complexes leads to inflammation, influx of peptidyl arginine deiminase (PAD) enzymes, hypercitrullination of other autoantigens (e.g. fibrinogen) and epitope spreading. This promotes cross-reactivity of RF with citrullinated targets and onset of the ACPA response.

APPENDIX

List of publications:

Trela, M., Nelson, P. N., Perera, S. A., Rylance, P. B. and Attridge, K. (2019) Citrullination facilitates cross-reactivity of rheumatoid factor with non-IgG1 Fc epitopes in rheumatoid arthritis, *Scientific Reports* [Accepted]

Trela, M., Nelson, P. N. and Rylance, P. B. (2016) The role of molecular mimicry and other factors in the association of Human Endogenous Retroviruses and autoimmunity, *Apmis*, **124**(1–2), pp. 88–104.

Nelson, P. N., Roden, D., Nevill, A., Freimanis, G. L., Trela, M., Ejtehadi, H. D., Bowman, S., Axford, J., Veitch, A. M., Tugnet, N. and Rylance, P. B. (2014) Rheumatoid arthritis is associated with IgG antibodies to human endogenous retrovirus Gag matrix: A potential pathogenic mechanism of disease?, *Journal of Rheumatology*, **41**(10), pp. 1952–1960.

Nelson, P., Rylance, P., Roden, D., Trela, M. and Tugnet, N. (2014) Viruses as potential pathogenic agents in systemic lupus erythematosus, *Lupus*, **23**(6), pp. 596–605.

Tugnet, N., Rylance, P., Roden, D., Trela, M. and Nelson, P. (2013) Human Endogenous Retroviruses (HERVs) and Autoimmune RheumaticDisease: Is There a Link?, *The Open Rheumatology Journal*, **7**(1), pp. 13–21.

LIST OF REFERENCES

Abdel Fattah, N. S. A., Hassan, H. E., Galal, Z. A. and El Okda, E. S. E. (2009) Assessment of anti-cyclic citrullinated peptide in psoriatic arthritis., *BMC research notes*,**2**,pp.44-50.[online] Available at: http://www.ncbi.nlm.nih.gov/pubmed/19298673 (Accessed 14 September 2017).

Adtani, P. and Malathi, N. (2015) Epstein-Barr virus and its association with rheumatoid arthritis and oral lichen planus., *Journal of oral and maxillofacial pathology*, **19**(3), pp. 282–285.

Agmon-Levin, N., Mosca, M., Petri, M. and Shoenfeld, Y. (2012) Systemic lupus erythematosus one disease or many?, *Autoimmunity Reviews*, **11**(8), pp. 593–595.

Ahmed, A. U. (2011) An overview of inflammation: mechanism and consequences, *Frontiers in Biology*, **6**(4), pp. 274-281.

Ajeganova, S. and Huizinga, T. W. J. (2014) Rheumatoid arthritis: Seronegative and seropositive RA: alike but different?, *Nature Reviews Rheumatology*, **11**(1), pp. 8–9.

Ajeganova, S., Humphreys, J. H., Verheul, M. K., van Steenbergen, H. W., van Nies, J. A. B., Hafström, I., Svensson, B., Huizinga, T. W. J., Trouw, L. A., Verstappen, S. M. M. and van der Helm-van Mil, A. H. M. (2016) Anticitrullinated protein antibodies and rheumatoid factor are associated with increased mortality but with different causes of death in patients with rheumatoid arthritis: a longitudinal study in three European cohorts, *Annals of the Rheumatic Diseases*, **75**(11), pp. 1924–1932.

Alemany, M., Concord, E., Garin, J., Vincon, M., Giles, A., Marguerie, G. and Gulino, D. (1996) Sequence 274-368 in the beta 3-subunit of the integrin alpha IIb beta 3 provides a ligand recognition and binding domain for the gamma-chain of fibrinogen that is independent of platelet activation, *Blood*, **87**(2), pp. 592-601.

Aletaha, D., Alasti, F. and Smolen, J. S. (2015) Rheumatoid factor, not antibodies against citrullinated proteins, is associated with baseline disease activity in rheumatoid arthritis clinical trials, *Arthritis Research and Therapy*, **17**(1), pp. 229-238.

Aletaha, D., Neogi, T., Silman, A. J., Funovits, J., Felson, D. T., Bingham, C. O., Birnbaum, N. S., Burmester, G. R., Bykerk, V. P., Cohen, M. D., Combe, B., Costenbader, K. H., Dougados, M., Emery, P., Ferraccioli, G., Hazes, J. M. W., Hobbs,

K., Huizinga, T. W. J., Kavanaugh, A., Kay, J., Kvien, T. K., Laing, T., Mease, P., Ménard, H. A., Moreland, L. W., Naden, R. L., Pincus, T., Smolen, J. S., Stanislawska-Biernat, E., Symmons, D., Tak, P. P., Upchurch, K. S., Vencovsky, J., Wolfe, F. and Hawker, G. (2010) 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative, *Annals of the rheumatic diseases*, **69**(9), pp. 1580–1588.

Ali, F., Rowley, M., Jayakrishnan, B., Teuber, S., Gershwin, M. E. and Mackay, I. R. (2011) Stiff-person syndrome (SPS) and anti-GAD-related CNS degenerations: Protean additions to the autoimmune central neuropathies, *Journal of Autoimmunity*, **37**(2), pp. 79–87.

de Almeida, D. E., Ling, S. and Holoshitz, J. (2011) New insights into the functional role of the rheumatoid arthritis shared epitope., *FEBS letters*, **585**(23), pp. 3619–3626.

Alonso-Padilla, J., Lafuente, E. M. and Reche, P. A. (2017) Computer-Aided Design of an Epitope-Based Vaccine against Epstein-Barr Virus, *Journal of Immunology Research*, **2017**, pp. 1–15. doi: 10.1155/2017/9363750

Alpízar-Rodríguez, D., Pluchino, N., Canny, G., Gabay, C. and Finckh, A. (2016) The role of female hormonal factors in the development of rheumatoid arthritis, *Rheumatology*, **56**(8), pp. 1254-1263.

Alves, C. H., Farrell, E., Vis, M., Colin, E. M. and Lubberts, E. (2016) Animal Models of Bone Loss in Inflammatory Arthritis: from Cytokines in the Bench to Novel Treatments for Bone Loss in the Bedside—a Comprehensive Review, *Clinical Reviews in Allergy and Immunology*, **51**(1), pp. 27–47.

Alves, P. M. S., Lévy, N., Stevenson, B. J., Bouzourene, H., Theiler, G., Bricard, G., Viatte, S., Ayyoub, M., Vuilleumier, H., Givel, J.-C. R., Rimoldi, D., Speiser, D. E., Jongeneel, C. V., Romero, P. J. and Lévy, F. (2008) Identification of tumor-associated antigens by large-scale analysis of genes expressed in human colorectal cancer, *Cancer immunity*, **8**, pp. 11-21.

Amur, S., Parekh, A. and Mummaneni, P. (2012) Sex differences and genomics in autoimmune diseases, *Journal of Autoimmunity*, **38**(2), pp. 254–265.

Andersson, G., Svensson, A. C., Setterblad, N. and Rask, L. (1998) Retroelements in the human MHC class II region, *Trends in genetics*, **14**(3), pp. 109–114.

Annadurai, B. (2008) *A textbook of immunology and immunotechnology*, Ram Nagar: S. Chand.

Anquetil, F., Clavel, C., Offer, G., Serre, G. and Sebbag, M. (2015) IgM and IgA Rheumatoid Factors Purified from Rheumatoid Arthritis Sera Boost the Fc Receptor– and Complement-Dependent Effector Functions of the Disease-Specific Anti–Citrullinated Protein Autoantibodies, *The Journal of Immunology*, **194**(8), pp. 3664–3674.

Apoil, P. A., Kuhlein, E., Robert, A., Rubie, H. and Blancher, A. (2006) HIGM syndrome caused by insertion of an AluYb8 element in exon 1 of the CD40LG gene, *Immunogenetics*, **59**(1), pp. 17–23.

Araki, Y. and Mimura, T. (2016) The Mechanisms Underlying Chronic Inflammation in Rheumatoid Arthritis from the Perspective of the Epigenetic Landscape, *Journal of Immunology Research*, **2016**. doi: 10.1155/2016/6290682

Arnoux, F., Fina, F., Lambert, N., Balandraud, N., Martin, M., Ouafik, L., Kanaan, S. B., Roudier, J. and Auger, I. (2016) Newly Identified BRAF Mutation in Rheumatoid Arthritis, *Arthritis and Rheumatology*, **68**(6), pp. 1377–1383.

Ashley, N. T., Weil, Z. M. and Nelson, R. J. (2012) Inflammation: Mechanisms, Costs, and Natural Variation, *Annual Review of Ecology, Evolution, and Systematics*, **43**, pp. 385–406.

Assinger, A., Yaiw, K.-C., Göttesdorfer, I., Leib-Mösch, C. and Söderberg-Nauclér, C. (2013) Human Cytomegalovirus (HCMV) induces Human Endogenous Retrovirus (HERV) transcription, *Retrovirology*, **10**(1), pp. 132-138.

Aterido, A., Palacio, C., Marsal, S., Ávila, G. and Julià, A. (2014) Novel Insights into the Regulatory Architecture of CD4+ T Cells in Rheumatoid Arthritis, *PLoS ONE*, **9**(6), p. e100690. doi: 10.1371/journal.pone.0100690

Auger, I., Balandraud, N., Rak, J., Lambert, N., Martin, M. and Roudier, J. (2009) New autoantigens in rheumatoid arthritis (RA): screening 8268 protein arrays with sera from patients with RA, *Annals of the Rheumatic Diseases*, **68**(4), pp. 591–594.

Auger, I., Charpin, C., Balandraud, N., Martin, M. and Roudier, J. (2012) Autoantibodies to PAD4 and BRAF in rheumatoid arthritis, *Autoimmunity Reviews*, **11**(11), pp. 801–803.

Badillo-Soto, M. A., Rodríguez-Rodríguez, M., Pérez-Pérez, M. E., Daza-Benitez, L., Bollain-Y-Goytia, J. J., Carrillo-Jiménez, M. A., Avalos-Díaz, E. and Herrera-Esparza, R. (2016) Potential protein targets of the peptidylarginine deiminase 2 and peptidylarginine deiminase 4 enzymes in rheumatoid synovial tissue and its possible meaning, *European journal of rheumatology*, AVES, **3**(2), pp. 44–49.

Balada, E., Ordi-Ros, J. and Vilardell-Tarrés, M. (2009) Molecular mechanisms mediated by human endogenous retroviruses (HERVs) in autoimmunity, *Reviews in Medical Virology*, **19**(5), pp. 273–286.

Balada, E., Vilardell-Tarrés, M. and Ordi-Ros, J. (2010) Implication of Human Endogenous Retroviruses in the Development of Autoimmune Diseases, *International Reviews of Immunology*, **29**(4), pp. 351–370.

Balestrieri, E., Pica, F., Matteucci, C., Zenobi, R., Sorrentino, R., Argaw-Denboba, A., Cipriani, C., Bucci, I. and Sinibaldi-Vallebona, P. (2015) Transcriptional Activity of Human Endogenous Retroviruses in Human Peripheral Blood Mononuclear Cells, *BioMed Research International*, **2015**, pp. 1–9. doi: 10.1155/2015/164529

Bankó, Z., Pozsgay, J., Gáti, T., Rojkovich, B., Ujfalussy, I. and Sármay, G. (2017) Regulatory B cells in rheumatoid arthritis: Alterations in patients receiving anti-TNF therapy, *Clinical Immunology*, **184**, pp. 63–69.

Barra, L., Bykerk, V., Pope, J. E., Haraoui, B. P., Hitchon, C. A., Thorne, J. C., Keystone, E. C., Boire, G. and Catch Investigators (2013) Anticitrullinated Protein Antibodies and Rheumatoid Factor Fluctuate in Early Inflammatory Arthritis and Do Not Predict Clinical Outcomes, *The Journal of Rheumatology*, **40**(8), pp. 1259–1267.

Barra, L., Pope, J., Bessette, L., Haraoui, B. and Bykerk, V. (2011) Lack of seroconversion of rheumatoid factor and anti-cyclic citrullinated peptide in patients with early inflammatory arthritis: a systematic literature review, *Rheumatology*, **50**(2), pp. 311–316.

Bartok, B. and Firestein, G. S. (2010) Fibroblast-like synoviocytes: key effector cells in rheumatoid arthritis, *Immunological Reviews*, **233**(1), pp. 233–255.

Baum, H., Brusic, V., Choudhuri, K., Cunningham, P., Vergani, D. and Peakman, M. (1995) MHC molecular mimicry in diabetes, *Nature medicine*, **1**(5), p. 388.

van Beers, J. J. B. C., Raijmakers, R., Alexander, L.-E., Stammen-Vogelzangs, J., Lokate, A. M. C., Heck, A. J. R., Schasfoort, R. B. M. and Pruijn, G. J. M. (2010) Mapping of citrullinated fibrinogen B-cell epitopes in rheumatoid arthritis by imaging surface plasmon resonance, *Arthritis research and therapy*, **12**(6), pp. R219-228. doi: 10.1186/ar3205

van Beers, J. J. B. C., Schwarte, C. M., Stammen-Vogelzangs, J., Oosterink, E., Božič, B. and Pruijn, G. J. M. (2013) The rheumatoid arthritis synovial fluid citrullinome reveals novel citrullinated epitopes in apolipoprotein E, myeloid nuclear differentiation antigen, and β-actin, *Arthritis and Rheumatism*, **65**(1), pp. 69–80.

Benson, R. A., Patakas, A., Conigliaro, P., Rush, C. M., Garside, P., McInnes, I. B. and Brewer, J. M. (2010) Identifying the Cells Breaching Self-Tolerance in Autoimmunity, *The Journal of Immunology*, **184**(11), pp. 6378–6385.

Bhaskaran, R. and Ponnuswamy, P. K. (2009) Positional flexibilities of amino acid residues in globular proteins, *International Journal of Peptide and Protein Research*, **32**(4), pp. 241–255.

Bhattacharjee, M., Balakrishnan, L., Renuse, S., Advani, J., Goel, R., Sathe, G., Keshava Prasad, T. S., Nair, B., Jois, R., Shankar, S. and Pandey, A. (2016) Synovial fluid proteome in rheumatoid arthritis, *Clinical Proteomics*, **13**(1), pp. 12-22. doi: 10.1186/s12014-016-9113-1

Bièche, I., Laurent, A., Laurendeau, I., Duret, L., Giovangrandi, Y., Frendo, J.-L., Olivi, M., Fausser, J.-L., Evain-Brion, D. and Vidaud, M. (2003) Placenta-Specific INSL4 Expression Is Mediated by a Human Endogenous Retrovirus Element1, *Biology of Reproduction*, **68**(4), pp. 1422–1429.

Bilate, A. M. and Lafaille, J. J. (2012) Regulatory T Cells in Immune Tolerance, *Annual Review of Immunology*, **30**, pp. 733–758.

Biswas, S., Sharma, S., Saroha, A., Bhakuni, D. S., Malhotra, R., Zahur, M., Oellerich, M., Das, H. R., Asif, A. R., Afzal, N., Karim, S., Mahmud, T., Sami, W., Arif, M., Goldbach-Mansky, R., Lee, J., McCoy, A., Hoxworth, J., Yarboro, C., Smolen, J., Boekel, M. van, Vossenaar, E., Hoogen, F. van den, Venrooij, W. van, Asif, A., Oellerich, M., Amstrong, V., Riemenschneider, B., Monod, M., Lernmark, G., Chambers, R., MacFarlane, D., Whicher, J., Dieppe, P., Li, G., Arora, P., Chen, Y., McCulloch, C., Liu, P., Osborn, T., Verdrengh, M., Stossel, T., Tarkowski, A.,

Bokarewa, M., Moore, A., Appelboam, A., Kawabata, K., Silva, J. Da, D'Cruz, D., Gowland, G., Shadforth, M. and McNaughton, D. (2013) Identification of Novel Autoantigen in the Synovial Fluid of Rheumatoid Arthritis Patients Using an Immunoproteomics Approach, *PLoS ONE*, **8**(2), pp. e56246-56254. doi: 10.1371/journal.pone.0056246

Blanco, F. J., Möricke, R., Dokoupilova, E., Codding, C., Neal, J., Andersson, M., Rohrer, S. and Richards, H. (2017) Secukinumab in Active Rheumatoid Arthritis: A Phase III Randomized, Double-Blind, Active Comparator- and Placebo-Controlled Study, *Arthritis and Rheumatology*, **69**(6), pp. 1144–1153.

Blaurock, N., Schmerler, D., Hünniger, K., Kurzai, O., Ludewig, K., Baier, M., Brunkhorst, F. M., Imhof, D. and Kiehntopf, M. (2016) C-Terminal Alpha-1 Antitrypsin Peptide: A New Sepsis Biomarker with Immunomodulatory Function, *Mediators of Inflammation*, **2016**, pp. 1–13. doi: 10.1155/2016/6129437.

Boissier, M.-C., Semerano, L., Challal, S., Saidenberg-Kermanac'h, N. and Falgarone, G. (2012) Rheumatoid arthritis: From autoimmunity to synovitis and joint destruction, *Journal of Autoimmunity*, **39**(3), pp. 222–228.

Boller, K., Schonfeld, K., Lischer, S., Fischer, N., Hoffmann, A., Kurth, R. and Tonjes, R. R. (2008) Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles, *Journal of General Virology*, **89**(2), pp. 567–572.

Bombardieri, M., Lewis, M. and Pitzalis, C. (2017) Ectopic lymphoid neogenesis in rheumatic autoimmune diseases, *Nature Reviews Rheumatology*, **13**(3), pp. 141–154.

Bondt, A., Nicolardi, S., Jansen, B. C., Kuijper, T. M., Hazes, J. M. W., van der Burgt, Y. E. M., Wuhrer, M. and Dolhain, R. J. E. M. (2017) IgA N- and O-glycosylation profiling reveals no association with the pregnancy-related improvement in rheumatoid arthritis, *Arthritis Research and Therapy*, **19**(1), pp. 160-167.

Bonilla, F. A. and Oettgen, H. C. (2010) Adaptive immunity, *Journal of Allergy and Clinical Immunology*, **125**(2), pp. S33–S40.

Boots, A. M. H., Hubers, H., Kouwijzer, M., den Hoed-van Zandbrink, L., Westrek-Esselink, B. M., van Doorn, C., Stenger, R., Bos, E. S., van Lierop, M. C., Verheijden, G. F., Timmers, C. M. and van Staveren, C. J. (2007) Identification of an altered peptide ligand based on the endogenously presented, rheumatoid arthritis-associated, human cartilage glycoprotein-39(263-275) epitope: an MHC anchor variant peptide for

immune modulation., *Arthritis research and therapy*, **9**(4), p. R71-81.

Brahmkshatriya, P. P., Mehta, A. A., Saboo, B. D. and Goyal, R. K. (2012) Characteristics and Prevalence of Latent Autoimmune Diabetes in Adults (LADA), *ISRN Pharmacology*, **2012**, pp. 1–8.

Brink, M., Hansson, M., Mathsson-Alm, L., Wijayatunga, P., Verheul, M. K., Trouw, L. A., Holmdahl, R., Rönnelid, J., Klareskog, L. and Rantapää-Dahlqvist, S. (2016) Rheumatoid factor isotypes in relation to antibodies against citrullinated peptides and carbamylated proteins before the onset of rheumatoid arthritis, *Arthritis research and therapy*, **18**(1), pp. 43-53.

Brink, M., Hansson, M., Mathsson, L., Jakobsson, P.-J., Holmdahl, R., Hallmans, G., Stenlund, H., Rönnelid, J., Klareskog, L. and Rantapää-Dahlqvist, S. (2013) Multiplex Analyses of Antibodies Against Citrullinated Peptides in Individuals Prior to Development of Rheumatoid Arthritis, *Arthritis and Rheumatism*, **65**(4), pp. 899–910.

Brookes, S. M., Pandolfino, Y. A., Mitchell, T. J., Venables, P. J., Shattles, W. G., Clark, D. A., Entwistle, A. and Maini, R. N. (1992) The immune response to and expression of cross-reactive retroviral gag sequences in autoimmune disease, *British journal of rheumatology*, **31**(11), pp. 735–742.

Brooks, W. H. and Renaudineau, Y. (2015) Epigenetics and autoimmune diseases: the X chromosome-nucleolus nexus, *Frontiers in Genetics*, **6**, pp. 22-41.

Brown, M. C., Joaquim, T. R., Chambers, R., Onisk, D. V., Yin, F., Moriango, J. M., Xu, Y., Fancy, D. A., Crowgey, E. L., He, Y., Stave, J. W. and Lindpaintner, K. (2011) Impact of Immunization Technology and Assay Application on Antibody Performance – A Systematic Comparative Evaluation, *PLoS ONE*, **6**(12), p. e28718. doi: 10.1371/journal.pone.0028718

Brown, S. M. (2000) *Bioinformatics : a biologist's guide to biocomputing and the internet*. Natick, MA: Eaton.

Buch, M. H., Smolen, J. S., Betteridge, N., Breedveld, F. C., Burmester, G., Dörner, T., Ferraccioli, G., Gottenberg, J.-E., Isaacs, J., Kvien, T. K., Mariette, X., Martin-Mola, E., Pavelka, K., Tak, P. P., van der Heijde, D., van Vollenhoven, R. F., Emery, P. and Rituximab Consensus Expert Committee, for the R. C. E. (2011) Updated consensus statement on the use of rituximab in patients with rheumatoid arthritis, *Annals of the rheumatic diseases*, **70**(6), pp. 909–920.

Bugatti, S., Vitolo, B., Caporali, R., Montecucco, C. and Manzo, A. (2014) B cells in rheumatoid arthritis: from pathogenic players to disease biomarkers., *BioMed research international*, **2014**, p. 681678. doi: 10.1155/2014/681678

Bull, H., Choy, M., Manyonda, I., Brown, C. A., Waldron, E. E., Holmes, S. D., Booth, J. C. and Nelson, P. N. (1999) Reactivity and assay restriction profiles of monoclonal and polyclonal antibodies to acid phosphatases: a preliminary study, *Immunology letters*, **70**(3), pp. 143–149.

Bürgler, S. (2017) Malignant Interaction between B Cells and T Helper Cells. In Isvoranu, G. *Lymphocyte Updates - Cancer, Autoimmunity and Infection*, InTech, [online] Available from: http://www.intechopen.com/books/lymphocyte-updates-cancer-autoimmunity-and-infection/malignant-interaction-between-b-cells-and-t-helper-cells (Accessed 14 January 2018).

Burska, A. N., Hunt, L., Boissinot, M., Strollo, R., Ryan, B. J., Vital, E., Nissim, A., Winyard, P. G., Emery, P., Ponchel, F., Burska, A. N., Hunt, L., Boissinot, M., Strollo, R., Ryan, B. J., Vital, E., Nissim, A., Winyard, P. G., Emery, P. and Ponchel, F. (2014) Autoantibodies to Posttranslational Modifications in Rheumatoid Arthritis, *Mediators of Inflammation*, **2014**, pp. 1–19. doi: 10.1155/2014/492873

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J. and Munafò, M. R. (2013) Power failure: why small sample size undermines the reliability of neuroscience, *Nature Reviews Neuroscience*, **14**(5), pp. 365–376.

Cantley, M. D., Smith, M. D. and Haynes, D. R. (2009) Pathogenic bone loss in rheumatoid arthritis: mechanisms and therapeutic approaches, *International Journal of Clinical Rheumatology*, **4**(5), pp. 561–582.

Caplazi, P., Baca, M., Barck, K., Carano, R. A. D., DeVoss, J., Lee, W. P., Bolon, B. and Diehl, L. (2015) Mouse Models of Rheumatoid Arthritis, *Veterinary Pathology*, **52**(5), pp. 819–826.

Carvalheiro, H., Duarte, C., Silva-Cardoso, S., da Silva, J. A. P. and Souto-Carneiro, M. M. (2015) CD8+ T Cell Profiles in Patients With Rheumatoid Arthritis and Their Relationship to Disease Activity, *Arthritis and Rheumatology*, **67**(2), pp. 363–371.

Carvalheiro, H., da Silva, J. A. P. and Souto-Carneiro, M. M. (2013) Potential roles for CD8+ T cells in rheumatoid arthritis, *Autoimmunity Reviews*, **12**(3), pp. 401–409.

Castro-Sánchez, P. and Roda-Navarro, P. (2017) Physiology and Pathology of Autoimmune Diseases: Role of CD4+ T cells in Rheumatoid Arthritis, In Rezaei, N. *Physiology and Pathology of Immunology*, InTech, [online] Available from: http://www.intechopen.com/books/physiology-and-pathology-of-immunology/physiology-and-pathology-of-autoimmune-diseases-role-of-cd4-t-cells-in-rheumatoid-arthritis (Accessed 11 February 2018).

Catrina, A. I., Joshua, V., Klareskog, L. and Malmström, V. (2016) Mechanisms involved in triggering rheumatoid arthritis, *Immunological Reviews*, **269**(1), pp. 162–174.

Catrina, A. I., Svensson, C. I., Malmström, V., Schett, G. and Klareskog, L. (2017) Mechanisms leading from systemic autoimmunity to joint-specific disease in rheumatoid arthritis, *Nature Reviews Rheumatology*, **13**(2), pp. 79–86.

Challa, A. A. and Stefanovic, B. (2011) A Novel Role of Vimentin Filaments: Binding and Stabilization of Collagen mRNAs, *Molecular and Cellular Biology*, **31**(18), pp. 3773–3789.

Chamaillard, M., Girardin, S. E., Viala, J. and Philpott, D. J. (2003) Nods, Nalps and Naip: intracellular regulators of bacterial-induced inflammation, *Cellular microbiology*, **5**(9), pp. 581–592.

Chang, E. Y., Chen, K. C., Huang, B. K. and Kavanaugh, A. (2016) Adult Inflammatory Arthritides: What the Radiologist Should Know, *RadioGraphics*, **36**(6), pp. 1849–1870.

Chang, K. and Pastan, I. (1996) Molecular cloning of mesothelin, a differentiation antigen present on mesothelium, mesotheliomas, and ovarian cancers, *Proceedings of the National Academy of Sciences of the United States of America*, **93**(1), pp. 136–140.

Chang, K., Yang, S. M., Kim, S. H., Han, K. H., Park, S. J. and Shin, J. Il (2014) Smoking and rheumatoid arthritis, *International journal of molecular sciences*, **15**(12), pp. 22279–22295.

Chang, X., Zhao, Y., Wang, Y., Chen, Y. and Yan, X. (2013) Screening Citrullinated Proteins in Synovial Tissues of Rheumatoid Arthritis Using 2-dimensional Western Blotting, *The Journal of Rheumatology*, **40**(3), pp. 219–227.

Charpin, C., Martin, M., Balandraud, N., Roudier, J. and Auger, I. (2010) Autoantibodies to BRAF, a new family of autoantibodies associated with rheumatoid arthritis, *Arthritis Research and Therapy*, **12**(5), pp. R194-200.

Cheung, Y.-K., Cheng, S. C.-S., Ke, Y. and Xie, Y. (2010) Two novel HLA-A*0201 T-cell epitopes in avian H5N1 viral nucleoprotein induced specific immune responses in HHD mice, *Veterinary Research*, **41**(2), pp. 24-36.

Chiang, C-Y., Chang, J-T., Lin, M-S., Wang, S-R. and Chang, H-Y. (2005) Characterization of a monoclonal antibody specific to the Gag protein of porcine endogenous retrovirus and its application in detecting the virus infection, *Virus Research*, **108**(1-2), pp. 139-148.

Chou, P. Y. and Fasman, G. D. (1974) Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins, *Biochemistry*, **13**(2), pp. 211–22.

Choy, E. (2012) Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis, *Rheumatology*, **51**(suppl 5), pp. v3–v11.

Claverie, J.-M. and Notredame, C. (2007) *Bioinformatics for dummies*. Hoboken, NJ: Wiley.

Cohen, C. J., Lock, W. M. and Mager, D. L. (2009) Endogenous retroviral LTRs as promoters for human genes: A critical assessment, *Gene*, **448**(2), pp. 105–114.

Conley, A. and Hinshelwood, M. (2001) Mammalian aromatases, *Reproduction* **121**(5), pp. 685–695.

Conrad, B., Weissmahr, R. N., Böni, J., Arcari, R., Schüpbach, J. and Mach, B. (1997) A Human Endogenous Retroviral Superantigen as Candidate Autoimmune Gene in Type I Diabetes, *Cell*, **90**(2), pp. 303–313.

Cooper, D. L., Martin, S. G., Robinson, J. I., Mackie, S. L., Charles, C. J., Nam, J., Consortium, Y., Isaacs, J. D., Emery, P. and Morgan, A. W. (2012) FcγRIIIa Expression on Monocytes in Rheumatoid Arthritis: Role in Immune-Complex Stimulated TNF Production and Non-Response to Methotrexate Therapy, *PLoS ONE*, **7**(1), p. e28918. doi: 10.1371/journal.pone.0028918

Cornaby, C., Gibbons, L., Mayhew, V., Sloan, C. S., Welling, A. and Poole, B. D. (2015) B cell epitope spreading: Mechanisms and contribution to autoimmune diseases, *Immunology Letters*, **163**(1), pp. 56–68.

Costenbader, K. H. and Karlson, E. W. (2006) Epstein-Barr virus and rheumatoid arthritis: is there a link?, *Arthritis research and therapy*, **8**(1), pp. 204-210.

Cristaldi, E., Malaguarnera, G., Rando, A. and Malaguarner, M. (2011) A Possible Link Between Autoimmunity and Cancer, In Mavragani, C.P*. Autoimmune Disorders - Pathogenetic Aspects*, InTech, [online] Available from: http://www.intechopen.com/books/autoimmune-disorders-pathogenetic-aspects/a-possible-link-between-autoimmunity-and-cancer (Accessed 24 January 2018).

Crocker, J. and Murray, P. (2003) *Molecular biology in cellular pathology.* Chichester: John Wiley & Sons.

Crofford, L. J. (2013) Use of NSAIDs in treating patients with arthritis, *Arthritis Research and Therapy* **15**(3), pp. S2-12.

Cuesta, Á. M., Sainz-Pastor, N., Bonet, J., Oliva, B. and Álvarez-Vallina, L. (2010) Multivalent antibodies: when design surpasses evolution, *Trends in Biotechnology*, **28**(7), pp. 355–362.

Cunningham, M. (2004) T cell mimicry in inflammatory heart disease, *Molecular Immunology*, **40**(14–15), pp. 1121–1127.

Curtin, F., Bernard, C., Levet, S., Perron, H., Porchet, H., Médina, J., Malpass, S., Lloyd, D., Simpson, R. and RAINBOW-T1D investigators (2018) A new therapeutic approach for type 1 diabetes: Rationale for GNbAC1, an anti-HERV-W-Env monoclonal antibody, *Diabetes, Obesity and Metabolism,* **20**(9), pp. 2075-2084.

Curtin, F., Perron, H., Kromminga, A., Porchet, H. and Lang, A. B. (2015) Preclinical and early clinical development of GNbAC1, a humanized IgG4 monoclonal antibody targeting endogenous retroviral MSRV-Env protein, *mAbs*, **7**(1), pp. 265–275.

Curtin, F., Vidal, V., Bernard, C., Kromminga, A., Lang, A. B. and Porchet, H. (2016) Serum pharmacokinetics and cerebrospinal fluid concentration analysis of the new IgG4 monoclonal antibody GNbAC1 to treat multiple sclerosis: A Phase 1 study, *mAbs*, **8**(5), pp. 854–860.

Cusick, M. F., Libbey, J. E. and Fujinami, R. S. (2012) Molecular mimicry as a mechanism of autoimmune disease., *Clinical reviews in allergy and immunology*, **42**(1), pp. 102–111.

Dam, C. E., Houen, G., Hansen, P. R. and Trier, N. H. (2014) Identification and fine mapping of a linear B cell epitope of human vimentin, *Scandinavian Journal of Clinical and Laboratory Investigation*, **74**(6), pp. 506–514.

Damgaard, D., Bjørn, M. E., Jensen, P. Ø. and Nielsen, C. H. (2017) Reactive oxygen species inhibit catalytic activity of peptidylarginine deiminase, *Journal of Enzyme Inhibition and Medicinal Chemistry*, **32**(1), pp. 1203–1208.

Dangel, A. W., Mendoza, A. R., Baker, B. J., Daniel, C. M., Carroll, M. C., Wu, L. C. and Yu, C. Y. (1994) The dichotomous size variation of human complement C4 genes is mediated by a novel family of endogenous retroviruses, which also establishes species-specific genomic patterns among Old World primates, *Immunogenetics*, **40**(6), pp. 425–36.

Derfuss, T., Curtin, F., Guebelin, C., Bridel, C., Rasenack, M., Matthey, A., Du Pasquier, R., Schluep, M., Desmeules, J., Lang, A. B., Perron, H., Faucard, R., Porchet, H., Hartung, H.-P., Kappos, L. and Lalive, P. H. (2015) A phase IIa randomized clinical study testing GNbAC1, a humanized monoclonal antibody against the envelope protein of multiple sclerosis associated endogenous retrovirus in multiple sclerosis patients — A twelve month follow-up, *Journal of Neuroimmunology,* 285, pp. 68-70

Deuel, T. F., Zhang, N., Yeh, H.-J., Silos-Santiago, I. and Wang, Z.-Y. (2002) Pleiotrophin: A Cytokine with Diverse Functions and a Novel Signaling Pathway, *Archives of Biochemistry and Biophysics*, **397**(2), pp. 162–171.

Deyev, S. M. and Lebedenko, E. N. (2009) Modern Technologies for Creating Synthetic Antibodies for Clinical application, *Acta naturae*, **1**(1), pp. 32–50.

Dieterich, W., Ehnis, T., Bauer, M., Donner, P., Volta, U., Riecken, E. O. and Schuppan, D. (1997) Identification of tissue transglutaminase as the autoantigen of celiac disease, *Nature medicine*, **3**(7), pp. 797–801.

Doria-Rose, N. A. and Joyce, M. G. (2015) Strategies to guide the antibody affinity maturation process, *Current opinion in virology*, **11**, pp. 137–147.

Douville, R. N. and Nath, A. (2014) Human endogenous retroviruses and the nervous system, *Handbook of clinical neurology*, **123**, pp. 465–485.

Ehlhardt, S., Seifert, M., Schneider, J., Ojak, A., Zang, K. D. and Mehraein, Y. (2006) Human endogenous retrovirus HERV-K(HML-2) Rec expression and transcriptional activities in normal and rheumatoid arthritis synovia, *The Journal of rheumatology*, **33**(1), pp. 16–23.

Ejtehadi, H. D., Freimanis, G. L., Ali, H. A., Bowman, S., Alavi, A., Axford, J., Callaghan, R., Nelson, P. N. and Ejtehadi, H. D. (2006) The potential role of human endogenous retrovirus K10 in the pathogenesis of rheumatoid arthritis: a preliminary study, *Annals of Rheumatic Diseases*, **65**(5), pp. 612–616.

Ejthadi, H. D., Martin, J. H., Junying, J., Roden, D. A., Lahiri, M., Warren, P., Murray, P. G. and Nelson, P. N. (2005) A novel multiplex RT-PCR system detects human endogenous retrovirus-K in breast cancer, *Archives of Virology*, **150**(1), pp. 177–184.

El-Manzalawy, Y. and Honavar, V. (2010) Recent advances in B-cell epitope prediction methods, *Immunome research*, **6 Suppl 2**(Suppl 2), pp. S2-10.

Elshahaly, M., Wheater, G., Tuck, S., Datta, H. and van Laar, J. (2012) The role of B cells in bone turnover in rheumatoid arthritis, *International Journal of Clinical Rheumatology*, **7**(2), pp. 167–177.

Emini, E. A., Hughes, J. V, Perlow, D. S. and Boger, J. (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide, *Journal of virology*, **55**(3), pp. 836–839.

Emmer, A., Staege, M. S. and Kornhuber, M. E. (2014) The Retrovirus/Superantigen Hypothesis of Multiple Sclerosis, *Cellular and Molecular Neurobiology*, **34**(8), pp. 1087–1096.

Endl, J., Otto, H., Jung, G., Dreisbusch, B., Donie, F., Stahl, P., Elbracht, R., Schmitz, G., Meinl, E., Hummel, M., Ziegler, A. G., Wank, R. and Schendel, D. J. (1997) Identification of naturally processed T cell epitopes from glutamic acid decarboxylase presented in the context of HLA-DR alleles by T lymphocytes of recent onset IDDM patients, *The Journal of clinical investigation*, **99**(10), pp. 2405–2415.

Escalera-Zamudio, M. and Greenwood, A. D. (2016) On the classification and evolution of endogenous retrovirus: human endogenous retroviruses may not be 'human' after all, *APMIS*, **124**(1–2), pp. 44–51.

Falconer, J., Mahida, R., Venkatesh, D., Pearson, J. and Robinson, J. H. (2016) Unconventional T-cell recognition of an arthritogenic epitope of proteoglycan aggrecan released from degrading cartilage, *Immunology*, **147**(4), pp. 389–398.

Fan, S. S., Zong, M., Zhang, H., Lu, Y., Lu, T. B. and Fan, L. Y. (2015) Decreased expression of alpha-enolase inhibits the proliferation of hypoxia-induced rheumatoid arthritis fibroblasts-like synoviocytes, *Modern Rheumatology*, **25**(5), pp. 701–707.

Farber, D. L., Yudanin, N. A. and Restifo, N. P. (2014) Human memory T cells: generation, compartmentalization and homeostasis, *Nature Reviews Immunology*, **14**(1), pp. 24–35.

Fernandes-Cerqueira, C., Ossipova, E., Gunasekera, S., Hansson, M., Mathsson, L., Catrina, A. I., Sommarin, Y., Klareskog, L., Lundberg, K., Rönnelid, J., Göransson, U. and Jakobsson, P.-J. (2015) Targeting of anti-citrullinated protein/peptide antibodies in rheumatoid arthritis using peptides mimicking endogenously citrullinated fibrinogen antigens, *Arthritis Research and Therapy*, **17**(1), p. 155. doi: 10.1186/s13075-015-0666-6

Fessler, J., Raicht, A., Husic, R., Ficjan, A., Schwarz, C., Duftner, C., Schwinger, W., Graninger, W. B., Stradner, M. H. and Dejaco, C. (2017) Novel Senescent Regulatory T-Cell Subset with Impaired Suppressive Function in Rheumatoid Arthritis, *Frontiers in Immunology*, **8**, pp. 300-310.

Feuchter-Murthy, A. E., Freeman, J. D. and Mager, D. L. (1993) Splicing of a human endogenous retrovirus to a novel phospholipase A2 related gene, *Nucleic acids research*, **21**(1), pp. 135–143.

Firestein, G. S. and Mcinnes, I. B. (2017) Immunopathogenesis of Rheumatoid Arthritis, *Immunity*, **46**(2), pp. 183-196.

Fischer, B. D., Adeyemo, A., O'Leary, M. E. and Bottaro, A. (2017) Animal models of rheumatoid pain: experimental systems and insights, *Arthritis Research and Therapy*, **19**(1), pp. 146-154.

Flockerzi, A., Ruggieri, A., Frank, O., Sauter, M., Maldener, E., Kopper, B., Wullich, B., Seifarth, W., Müller-Lantzsch, N., Leib-Mösch, C., Meese, E. and Mayer, J. (2008) Expression patterns of transcribed human endogenous retrovirus HERV-K(HML-2) loci in human tissues and the need for a HERV Transcriptome Project, *BMC Genomics*, **9**(1), pp. 354-370.

Flores-Borja, F., Bosma, A., Ng, D., Reddy, V., Ehrenstein, M. R., Isenberg, D. A. and Mauri, C. (2013) CD19+CD24hiCD38hi B Cells Maintain Regulatory T Cells While Limiting TH1 and TH17 Differentiation, *Science Translational Medicine*, **5**(173), p. 173ra23. doi: 10.1126/scitranslmed.3005407.

Foley, J. F. (2013) Focus issue: understanding mechanisms of inflammation, *Science signaling*, **6**(258), p. eg2. doi: 10.1126/scisignal.2003938.

Forsström, B., Axnäs, B. B., Rockberg, J., Danielsson, H., Bohlin, A. and Uhlen, M. (2015) Dissecting antibodies with regards to linear and conformational epitopes., *PloS one*, **10**(3), p. e0121673. doi: 10.1371/journal.pone.0121673

Freimanis, G. (2008) *The Detection and Role of Human Endogenous Retrovirus K (HML-2) In Rheumatoid Arthritis*. PhD. Thesis, University of Wolverhampton.

Freimanis, G., Hooley, P., Ejtehadi, H. D., Ali, H. A., Veitch, A., Rylance, P. B., Alawi, A., Axford, J., Nevill, A., Murray, P. G. and Nelson, P. N. (2010) A role for human endogenous retrovirus-K (HML-2) in rheumatoid arthritis: investigating mechanisms of pathogenesis, *Clinical and Experimental Immunology*, **160**(3), pp. 340–347.

Fridkis-Hareli, M. (2013) Design of Peptide Immunotherapies for MHC Class-II-Associated Autoimmune Disorders, *Clinical and Developmental Immunology*, **2013**, pp. 1–9. doi: 10.1155/2013/826191.

Fujinami, R. S., von Herrath, M. G., Christen, U. and Whitton, J. L. (2006) Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease, *Clinical microbiology reviews*, **19**(1), pp. 80–94.

Fukuta, S., Miyamoto, K., Suzuki, K., Maehara, H., Inoue, T., Kikuike, K. and Shimizu, K. (2011) Abundance of calpain and aggrecan-cleavage products of calpain in degenerated human intervertebral discs, *Osteoarthritis and Cartilage*, **19**(10), pp. 1254–1262.

Futosi, K., Fodor, S. and Mócsai, A. (2013) Neutrophil cell surface receptors and their intracellular signal transduction pathways, *International immunopharmacology*, **17**(3), pp. 638–650.

Gabrielli, F., Donadel, G., Bensi, G., Heguy, A. and Melli, M. (1995) A nuclear protein, synthesized in growth-arrested human hepatoblastoma cells, is a novel member of the short-chain alcohol dehydrogenase family, *European journal of biochemistry*, **232**(2), pp. 473–477.

García, D. M., Weigum, S. E. and Koke, J. R. (2003) GFAP and nuclear lamins share an epitope recognized by monoclonal antibody J1-31, *Brain research*, **976**(1), pp. 9–21.

Gerlag, D. M., Norris, J. M. and Tak, P. P. (2016) Towards prevention of autoantibody-positive rheumatoid arthritis: from lifestyle modification to preventive treatment, *Rheumatology*, **55**(4), pp. 607–614.

Germolec, D., Kono, D. H., Pfau, J. C. and Pollard, K. M. (2012) Animal models used to examine the role of the environment in the development of autoimmune disease: findings from an NIEHS Expert Panel Workshop, *Journal of autoimmunity*, **39**(4), pp. 285–293.

Gesteira, T. F., Coulson-Thomas, V. J., Yuan, Y., Zhang, J., Nader, H. B. and Kao, W. W.-Y. (2017) Lumican Peptides: Rational Design Targeting ALK5/TGFBRI., *Scientific reports*, **7**, p. 42057. doi: 10.1038/srep42057.

Gibb, E. A., Warren, R. L., Wilson, G. W., Brown, S. D., Robertson, G. A., Morin, G. B. and Holt, R. A. (2015) Activation of an endogenous retrovirus-associated long non-coding RNA in human adenocarcinoma, *Genome Medicine*, **7**(1), pp. 22-36.

di Giovine, F. S., Bailly, S., Bootman, J., Almond, N. and Duff, G. W. (1994) Absence of lentiviral and human T cell leukemia viral sequences in patients with rheumatoid arthritis, *Arthritis and rheumatism*, **37**(3), pp. 349–358.

Glaant, T. T., Ocsko, T., Markovics, A., Szekanecz, Z., Katz, R. S., Rauch, T. A. and Mikecz, K. (2016) Characterization and Localization of Citrullinated Proteoglycan Aggrecan in Human Articular Cartilage, *PLOS ONE*, **11**(3), p. e0150784. doi: 10.1371/journal.pone.0150784

Gonzalez-Hernandez, M. J., Swanson, M. D., Contreras-Galindo, R., Cookinham, S., King, S. R., Noel, R. J., Kaplan, M. H. and Markovitz, D. M. (2012) Expression of Human Endogenous Retrovirus Type K (HML-2) Is Activated by the Tat Protein of HIV-1, *Journal of Virology*, **86**(15), pp. 7790–7805.

Grantham, R. (1974) Amino acid difference formula to help explain protein evolution., *Science,* **185**(4154), pp. 862–864.

Griffiths, D. J., Cooke, S. P., Herve, C., Rigby, S. P., Mallon, E., Hajeer, A., Lock, M., Emery, V., Taylor, P., Pantelidis, P., Bunker, C. B., Du Bois, R., Weiss, R. A. and Venables, P. J. W. (1999) Detection of human retrovirus 5 in patients with arthritis and systemic lupus erythematosus, *Arthritis and Rheumatism*, **42**(3), pp. 448–454.

Gröger, V. and Cynis, H. (2018) Human Endogenous Retroviruses and Their Putative Role in the Development of Autoimmune Disorders Such as Multiple Sclerosis, *Frontiers in Microbiology*, **9**, pp. 265-272.

Guedes, R. L. M., Rodrigues, C. M. F., Coatnoan, N., Cosson, A., Cadioli, F. A., Garcia, H. A., Gerber, A. L., Machado, R. Z., Minoprio, P. M. C., Teixeira, M. M. G. and de Vasconcelos, A. T. R. (2018) A comparative in silico linear B-cell epitope prediction and characterization for South American and African Trypanosoma vivax strains, *Genomics*, [online] Available from:https://www.sciencedirect.com/science/article/pii/S0888754318301265 (Accessed 15 May 2018).

Gururajan, M., Sindhava, V. and Bondada, S. (2014) B Cells and Immunological Tolerance, *Antibodies*, **3**(1), pp. 116–129.

Güven, E., Duus, K., Lydolph, M. C., Jørgensen, C. S., Laursen, I. and Houen, G. (2014) Non-specific binding in solid phase immunoassays for autoantibodies correlates with inflammation markers, *Journal of Immunological Methods*, **403**(1–2), pp. 26–36.

Haji-Ghassemi, O., Blackler, R. J., Young, N. M. and Evans, S. V (2015) Antibody recognition of carbohydrate epitopes, *Glycobiology*, **25**(9), pp. 920–952.

Hajishengallis, G. and Lambris, J. D. (2010) Crosstalk pathways between Toll-like receptors and the complement system, *Trends in immunology*, **31**(4), pp. 154–163.

Hamze, M., Desmetz, C., Guglielmi, P. and Guglielmi, P. (2013) B cell-derived cytokines in disease, *European Cytokine Network,* **24**(241), pp. 20–26.

Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S. and Delannoy, P. (2001) The human sialyltransferase family, *Biochimie*, **83**(8), pp. 727–737.

Harwood, N. E. and Batista, F. D. (2010) Antigen presentation to B cells., *F1000 biology reports*, **2**, pp. 87-91.

Hassan, R. and Ho, M. (2008) Mesothelin targeted cancer immunotherapy, *European Journal of Cancer*, **44**(1), pp. 46–53.

Havasi, A., Lu, W., Cohen, H. T., Beck, L., Wang, Z., Igwebuike, C. and Borkan, S. C. (2017) Blocking peptides and molecular mimicry as treatment for kidney disease, *American Journal of Physiology-Renal Physiology*, **312**(6), pp. F1016–F1025.

He, C., Feng, P., Su, H., Gu, A., Yan, Z. and Zhu, X. (2017) Disrupted Th1/Th2 balance in patients with rheumatoid arthritis (RA), *International Journal of Clinical and Experimental Pathology*, **10**(2), pp. 1233–1242.

Heesters, B. A., van der Poel, C. E., Das, A. and Carroll, M. C. (2016) Antigen Presentation to B Cells, *Trends in Immunology*, **37**(12), pp. 844–854.

Herve, C. A., Lugli, E. B., Brand, A., Griffiths, D. J. and Venables, P. J. W. (2002) Autoantibodies to human endogenous retrovirus-K are frequently detected in health and disease and react with multiple epitopes, *Clinical and experimental immunology*, **128**(1), pp. 75–82.

Hewagama, A. and Richardson, B. (2009) The genetics and epigenetics of autoimmune diseases, *Journal of autoimmunity*, **33**(1), pp. 3–11.

Hiemstra, H. S., Schloot, N. C., van Veelen, P. A., Willemen, S. J. M., Franken, K. L. M. C., van Rood, J. J., de Vries, R. R. P., Chaudhuri, A., Behan, P. O., Drijfhout, J. W. and Roep, B. O. (2001) Cytomegalovirus in autoimmunity: T cell crossreactivity to viral antigen and autoantigen glutamic acid decarboxylase, *Proceedings of the National Academy of Sciences*, **98**(7), pp. 3988–3991.

Hohn, O., Hanke, K. and Bannert, N. (2013) HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease, *Frontiers in Oncology*, **3**, pp. 246-257.

Hoovels, L. Van, Jacobs, J., Cruyssen, B. Vander, Bremt, S. Van Den, Verschueren, P. and Bossuyt, X. (2018) Performance characteristics of rheumatoid factor and anti-cyclic citrullinated peptide antibody assays may impact ACR / EULAR classification of rheumatoid arthritis, *Clinical and epidemiological research,* **77**(5), pp. 1–11.

Hopp, T. P. (1993) Retrospective: 12 years of antigenic determinant predictions, and more, *Peptide research*, **6**(4), pp. 183–190.

Hopp, T. P. and Woods, K. R. (1981) Prediction of protein antigenic determinants from amino acid sequences, *Proceedings of the National Academy of Sciences of the United States of America*, **78**(6), pp. 3824–3828.

Hosokawa, H. and Rothenberg, E. V. (2017) Cytokines, Transcription Factors, and the Initiation of T-Cell Development, *Cold Spring Harbor Perspectives in Biology*, **10**(5), p. a028621. doi: 10.1101/cshperspect.a028621.

Hsiao, F. C., Lin, M., Tai, A., Chen, G. and Huber, B. T. (2006) Cutting edge: Epstein-Barr virus transactivates the HERV-K18 superantigen by docking to the human complement receptor 2 (CD21) on primary B cells, *Journal of immunology,* **177**(4), pp. 2056–2060.

Huang, C., Liu, Y., Wu, H., Sun, D. and Li, Y. (2017) Characterization of IgG glycosylation in rheumatoid arthritis patients by MALDI-TOF-MSn and capillary electrophoresis, *Analytical and Bioanalytical Chemistry*, **409**(15), pp. 3731–3739.

Hurst, T. and Magiorkinis, G. (2017) Epigenetic Control of Human Endogenous Retrovirus Expression: Focus on Regulation of Long-Terminal Repeats (LTRs), *Viruses*, **9**(6), pp. 130-142.

Hutchinson, D., Murphy, D., Clarke, A. and Eggleton, P. (2016) Are Rheumatoid Factor, Anti-Citrullinated Protein Antibodies, and Anti-Carbamylated Protein Antibodies Linked by Posttranslational Modification of IgG? Comment on the Article by Koppejan *et al*, *Arthritis & Rheumatology*, **68**(11), pp. 2825–2826.

Ingegnoli, F., Castelli, R. and Gualtierotti, R. (2013) Rheumatoid Factors: Clinical Applications, *Disease Markers*, **35**(6), pp. 727–734.

Innala, L., Berglin, E., Möller, B., Ljung, L., Smedby, T., Södergren, A., Magnusson, S., Rantapää-Dahlqvist, S. and Wållberg-Jonsson, S. (2014) Age at onset determines severity and choice of treatment in early rheumatoid arthritis: a prospective study,

Arthritis Research and Therapy, **16**(2), pp. R94-102.

Ioan-Facsinay, A., el-Bannoudi, H., Scherer, H. U., van der Woude, D., Menard, H. A., Lora, M., Trouw, L. A., Huizinga, T. W. J. and Toes, R. E. M. (2011) Anti-cyclic citrullinated peptide antibodies are a collection of anti-citrullinated protein antibodies and contain overlapping and non-overlapping reactivities, *Annals of the Rheumatic Diseases*, **70**(1), pp. 188–193.

Isshiki, S., Togayachi, A., Kudo, T., Nishihara, S., Watanabe, M., Kubota, T., Kitajima, M., Shiraishi, N., Sasaki, K., Andoh, T. and Narimatsu, H. (1999) Cloning, expression, and characterization of a novel UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase (beta3Gal-T5) responsible for synthesis of type 1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom, *The Journal of biological chemistry*, **274**(18), pp. 12499–12507.

Iwasaki, A. and Medzhitov, R. (2015) Control of adaptive immunity by the innate immune system, *Nature Immunology*, **16**(4), pp. 343–353.

Iwasaki, A. and Medzhitov, R. (2010) Regulation of adaptive immunity by the innate immune system, *Science,* **327**(5963), pp. 291–295.

James, E. A., Moustakas, A. K., Bui, J., Papadopoulos, G. K., Bondinas, G., Buckner, J. H. and Kwok, W. W. (2010) HLA-DR1001 presents 'altered-self' peptides derived from joint-associated proteins by accepting citrulline in three of its binding pockets, *Arthritis and rheumatism*, **62**(10), pp. 2909–2918.

Janciauskiene, S. M., Bals, R., Koczulla, R., Vogelmeier, C., Köhnlein, T. and Welte, T. (2011) The discovery of α1-antitrypsin and its role in health and disease, *Respiratory Medicine*, **105**(8), pp. 1129–1139.

Janda, A., Bowen, A., Greenspan, N. S. and Casadevall, A. (2016) Ig Constant Region Effects on Variable Region Structure and Function, *Frontiers in microbiology*, **7**, pp. 22-31.

Janin, J. (1979) Surface and inside volumes in globular proteins, *Nature*, **277**(5696), pp. 491–492.

Jeffery, R. C. (2010) Clinical features of rheumatoid arthritis, *Medicine*, **38**(4), pp. 167–171.

Jern, P. and Coffin, J. M. (2008) Effects of Retroviruses on Host Genome Function, *Annual Review of Genetics*, **42**(1), pp. 709–732.

Jilani, A. A. and Mackworth-Young, C. G. (2015) The Role of Citrullinated Protein Antibodies in Predicting Erosive Disease in Rheumatoid Arthritis: A Systematic Literature Review and Meta-Analysis, *International Journal of Rheumatology*, **2015**, pp. 1–8. doi.org/10.1155/2015/728610

Jones, D. B. and Armstrong, N. W. (1995) Coxsackie virus and diabetes revisited., *Nature medicine*, **1**(4), p. 284.

Jones, G. W. and Jones, S. A. (2016) Ectopic lymphoid follicles: inducible centres for generating antigen-specific immune responses within tissues, *Immunology*, **147**(2), pp. 141–151.

Jones, J. D., Shyu, I., Newkirk, M. M. and Rigby, W. F. (2013) A rheumatoid factor paradox: inhibition of rituximab effector function, *Arthritis Research and Therapy*, **15**(1), pp. R20-30.

de Jong, H., Berlo, S. E., Hombrink, P., Otten, H. G., van Eden, W., Lafeber, F. P., Heurkens, A. H. M., Bijlsma, J. W. J., Glant, T. T. and Prakken, B. J. (2010) Cartilage proteoglycan aggrecan epitopes induce proinflammatory autoreactive T-cell responses in rheumatoid arthritis and osteoarthritis, *Annals of the Rheumatic Diseases*, **69**(01), pp. 255–262.

Juarez, M., Bang, H., Hammar, F., Reimer, U., Dyke, B., Sahbudin, I., Buckley, C. D., Fisher, B., Filer, A. and Raza, K. (2016) Identification of novel antiacetylated vimentin antibodies in patients with early inflammatory arthritis, *Annals of the Rheumatic Diseases*, **75**(6), pp. 1099–1107.

Kadomatsu, K. and Muramatsu, T. (2004) Midkine and pleiotrophin in neural development and cancer, *Cancer Letters*, **204**(2), pp. 127–143.

Kaer, K. and Speek, M. (2013) Retroelements in human disease, *Gene*, **518**(2), pp. 231–241.

Kalden, J. R. and Gay, S. (1994) Retroviruses and autoimmune rheumatic diseases, *Clinical and Experimental Immunology*, **98**(1), pp. 1–5.

Kanazawa, S., Nishizawa, S., Takato, T. and Hoshi, K. (2017) Biological Roles of Glial Fibrillary Acidic Protein as a Biomarker in Cartilage Regenerative Medicine, *Journal of Cellular Physiology*, **32**(11), pp. 3182-3193.

Karplus, P. A. and Schulz, G. E. (1985) Prediction of chain flexibility in proteins, *Naturwissenschaften*, **72**(4), pp. 212–213.

Kassiotis, G. (2014) Endogenous retroviruses and the development of cancer, *Journal of immunology* **192**(4), pp. 1343–1349.

Kay, E., Scotland, R. S. and Whiteford, J. R. (2014) Toll-like receptors: Role in inflammation and therapeutic potential, *BioFactors*, **40**(3), pp. 284–294.

Kidd, B. A., Ho, P. P., Sharpe, O., Zhao, X., Tomooka, B. H., Kanter, J. L., Steinman, L. and Robinson, W. H. (2008) Epitope spreading to citrullinated antigens in mouse models of autoimmune arthritis and demyelination, *Arthritis Research and Therapy*, **10**(5), pp. R119-130.

Kiecker, F., Streitz, M., Ay, B., Cherepnev, G., Volk, H.-D., Volkmer-Engert, R. and Kern, F. (2004) Analysis of antigen-specific T-cell responses with synthetic peptides—what kind of peptide for which purpose?, *Human Immunology*, **65**(5), pp. 523–536.

Kim, H. J., Moon, B.-I., Lee, J. W., Kim, S. C. and Kim, H.-J. (2016) Age-related reduction of antibody response against the human endogenous retrovirus K envelope in women, *Oncotarget*, **7**(14), pp. 17327–17337.

Kita, H., Matsumura, S., He, X., Ansari, A. A., Lian, Z., Van de Water, J., Coppel, R. L., Kaplan, M. M. and Gershwin, M. E. (2002) Analysis of TCR antagonism and molecular mimicry of an HLA-A*0201–restricted CTL epitope in primary biliary cirrhosis, *Hepatology*, **36**(4), pp. 918–926.

Klein, K. and Gay, S. (2015) Epigenetics in rheumatoid arthritis, *Current Opinion in Rheumatology*, **27**(1), pp. 76–82.

Klein, K., Ospelt, C. and Gay, S. (2012) Epigenetic contributions in the development of rheumatoid arthritis, *Arthritis research and therapy*, **14**(6), pp. 227-234.

Klein, L., Kyewski, B., Allen, P. M. and Hogquist, K. A. (2014) Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see), *Nature reviews. Immunology*, **14**(6), pp. 377–391.

Koh, T. J. and DiPietro, L. A. (2011) Inflammation and wound healing: the role of the macrophage, *Expert reviews in molecular medicine*, **13**, p. e23. doi: 10.1017/S1462399411001943.

Kolaczkowska, E. and Kubes, P. (2013) Neutrophil recruitment and function in health and inflammation, *Nature Reviews Immunology*, **13**(3), pp. 159–175.

Kollias, G., Papadaki, P., Apparailly, F., Vervoordeldonk, M. J., Holmdahl, R., Baumans, V., Desaintes, C., Di Santo, J., Distler, J., Garside, P., Hegen, M., Huizinga, T. W. J., Jüngel, A., Klareskog, L., McInnes, I., Ragoussis, I., Schett, G., Hart, B. 't, Tak, P. P., Toes, R., van den Berg, W., Wurst, W. and Gay, S. (2011) Animal models for arthritis: innovative tools for prevention and treatment, *Annals of the rheumatic diseases*, **70**(8), pp. 1357–1362.

Koziel, J., Mydel, P. and Potempa, J. (2014) The Link Between Periodontal Disease and Rheumatoid Arthritis: An Updated Review, *Current Rheumatology Reports*, **16**(3), pp. 408-414.

Kraus, B., Boller, K., Reuter, A. and Schnierle, B. S. (2011) Characterization of the human endogenous retrovirus K Gag protein: identification of protease cleavage sites, *Retrovirology, **8**,* pp*.* 21-28*.*

Kringelum, J. V., Nielsen, M., Padkjær, S. B. and Lund, O. (2013) Structural analysis of B-cell epitopes in antibody:protein complexes, *Molecular Immunology*, **53**(1), pp. 24–34.

Krueger, A., Ziętara, N. and Łyszkiewicz, M. (2017) T Cell Development by the Numbers, *Trends in Immunology*, **38**(2), pp. 128–139.

Krzysztalowska-Wawrzyniak, M., Ostanek, M., Clark, J., Binczak-Kuleta, A., Ostanek, L., Kaczmarczyk, M., Loniewska, B., Wyrwicz, L. S., Brzosko, M. and Ciechanowicz, A. (2011) The distribution of human endogenous retrovirus K-113 in health and autoimmune diseases in Poland, *Rheumatology*, **50**(7), pp. 1310–1314.

Kumanogoh, A. and Kikutani, H. (2004) Biological functions and signaling of a transmembrane semaphorin, CD100/Sema4D, *Cellular and Molecular Life Sciences (CMLS)*, **61**(3), pp. 292–300.

Kurkó, J., Besenyei, T., Laki, J., Glant, T. T., Mikecz, K. and Szekanecz, Z. (2013) Genetics of rheumatoid arthritis - a comprehensive review, *Clinical reviews in allergy and immunology*, **45**(2), pp. 170–179.

Kurosaki, T., Kometani, K. and Ise, W. (2015) Memory B cells, *Nature Reviews Immunology*, **15**(3), pp. 149–159.

Kurowska, W., Kuca-Warnawin, E. H., Radzikowska, A. and Maśliński, W. (2017) The role of anti-citrullinated protein antibodies (ACPA) in the pathogenesis of rheumatoid arthritis, *Central-European journal of immunology*, **42**(4), pp. 390–398.

van der Kuyl, A. C. (2012) HIV infection and HERV expression: a review, *Retrovirology*, **9**(1), pp. 615.

Larek-Rapala, A., Zaba, R., Kowalczyk, M., Szramka-Pawlak, B. and Schwartz, R. (2011) Herpes simplex virus infection as a possible modulator of autoimmune diseases facilitated by human endogenous retroviruses, *Postepy dermatologii i alergologii*, **28**, pp. 313–316.

Laska, M. J., Brudek, T., Nissen, K. K., Christensen, T., Moller-Larsen, A., Petersen, T. and Nexo, B. A. (2012) Expression of HERV-Fc1, a Human Endogenous Retrovirus, Is Increased in Patients with Active Multiple Sclerosis, *Journal of Virology*, **86**(7), pp. 3713–3722.

Laurent, L., Anquetil, F., Clavel, C., Ndongo-Thiam, N., Offer, G., Miossec, P., Pasquali, J.-L., Sebbag, M. and Serre, G. (2015) IgM rheumatoid factor amplifies the inflammatory response of macrophages induced by the rheumatoid arthritis-specific immune complexes containing anticitrullinated protein antibodies, *Annals of the Rheumatic Diseases*, **74**(7), pp. 1425–1431.

Lee, B. and Richards, F. M. (1971) The interpretation of protein structures: estimation of static accessibility, *Journal of molecular biology*, **55**(3), pp. 379–400.

Lee, J. Y., Choi, I. A., Kim, J.-H., Kim, K.-H., Lee, E. Y., Lee, E. B., Lee, Y.-M. and Song, Y. W. (2015) Association between anti-Porphyromonas gingivalis or anti-α-enolase antibody and severity of periodontitis or rheumatoid arthritis (RA) disease activity in RA, *BMC Musculoskeletal Disorders*, **16**(1), pp. 190-197.

Lemaître, C., Harper, F., Pierron, G., Heidmann, T. and Dewannieux, M. (2014) The HERV-K human endogenous retrovirus envelope protein antagonizes Tetherin antiviral activity, *Journal of virology*, **88**(23), pp. 13626–13637.

Lerner, A., Aminov, R. and Matthias, T. (2016) Dysbiosis May Trigger Autoimmune Diseases via Inappropriate Post-Translational Modification of Host Proteins, *Frontiers in microbiology*, **7**, pp. 84-89.

Levet, S., Medina, J., Joanou, J., Demolder, A., Queruel, N., Réant, K., Normand, M., Seffals, M., Dimier, J., Germi, R., Piofczyk, T., Portoukalian, J., Touraine, J.-L. and Perron, H. (2017) An ancestral retroviral protein identified as a therapeutic target in type-1 diabetes, *JCI insight*, **2**(17), p. 94387. doi: 10.1172/jci.insight.94387.

Li, S., Yu, Y., Yue, Y., Zhang, Z. and Su, K. (2013) Microbial Infection and Rheumatoid Arthritis, *Journal of clinical and cellular immunology*, **4**(6), pp. 174-187.

Li, Y., Jiang, L., Zhang, S., Yin, L., Ma, L., He, D. and Shen, J. (2012) Methotrexate attenuates the Th17/IL-17 levels in peripheral blood mononuclear cells from healthy individuals and RA patients, *Rheumatology International*, **32**(8), pp. 2415–2422.

Libbey, J. E., McCoy, L. L. and Fujinami, R. S. (2007) Molecular Mimicry in Multiple Sclerosis, *International review of neurobiology*, **79**, pp. 127–147.

Libusová, L., Sulimenko, T., Sulimenko, V., Janisch, R., Hozák, P. and Dráber, P. (2005) Distinct localization of a beta-tubulin epitope in the Tetrahymena thermophila and Paramecium caudatum cortex, *Protoplasma*, **225**(3–4), pp. 157–167.

Lindh, I., Snir, O., Lönnblom, E., Uysal, H., Andersson, I., Nandakumar, K. S., Vierboom, M., 't Hart, B., Malmström, V. and Holmdahl, R. (2014) Type II collagen antibody response is enriched in the synovial fluid of rheumatoid joints and directed to the same major epitopes as in collagen induced arthritis in primates and mice, *Arthritis research and therapy*, **16**(4), pp. R143-152.

Ling, S., Cheng, A., Pumpens, P., Michalak, M. and Holoshitz, J. (2010) Identification of the Rheumatoid Arthritis Shared Epitope Binding Site on Calreticulin, *PLoS ONE*, **5**(7), p. e11703. doi.org/10.1371/journal.pone.0011703

Lins, L., Thomas, A. and Brasseur, R. (2003) Analysis of accessible surface of residues in proteins, *Protein science,* **12**(7), pp. 1406–1417.

Liston, P., Fong, W. G. and Korneluk, R. G. (2003) The inhibitors of apoptosis: there is more to life than Bcl2, *Oncogene*, **22**(53), pp. 8568–8580.

Liu, M. and Eiden, M. V (2011) Role of human endogenous retroviral long terminal repeats (LTRs) in maintaining the integrity of the human germ line, *Viruses*, **3**(6), pp. 901–905.

Lock, F. E., Rebollo, R., Miceli-Royer, K., Gagnier, L., Kuah, S., Babaian, A., Sistiaga-Poveda, M., Lai, C. B., Nemirovsky, O., Serrano, I., Steidl, C., Karimi, M. M. and Mager, D. L. (2014) Distinct isoform of FABP7 revealed by screening for retroelement-activated genes in diffuse large B-cell lymphoma, *Proceedings of the National Academy of Sciences of the United States of America*, **111**(34), pp. E3534-3543.

López-Santalla, M., Salvador-Bernáldez, M., González-Alvaro, I., Castañeda, S., Ortiz, A. M., García-García, M. I., Kremer, L., Roncal, F., Mulero, J., Martínez-A, C. and Salvador, J. M. (2011) Tyr^{323}-dependent p38 activation is associated with rheumatoid arthritis and correlates with disease activity, *Arthritis and rheumatism*, **63**(7), pp. 1833–1842.

Losina, E. and Katz, J. N. (2017) Improvement at Any Cost? The Art and Science of Choosing Treatment Strategies for Rheumatoid Arthritis, *Annals of Internal Medicine*, **167**(1), pp. 55-56.

Lourenço, E. V and La Cava, A. (2011) Natural regulatory T cells in autoimmunity., *Autoimmunity*, **44**(1), pp. 33–42.

Löwer, R., Boller, K., Hasenmaier, B., Korbmacher, C., Müller-Lantzsch, N., Löwer, J. and Kurth, R. (1993) Identification of human endogenous retroviruses with complex mRNA expression and particle formation, *Proceedings of the National Academy of Sciences of the United States of America*, **90**(10), pp. 4480–4484.

Luckashenak, N., Schroeder, S., Endt, K., Schmidt, D., Mahnke, K., Bachmann, M. F., Marconi, P., Deeg, C. A. and Brocker, T. (2008) Constitutive Crosspresentation of Tissue Antigens by Dendritic Cells Controls CD8+ T Cell Tolerance In Vivo, *Immunity*, **28**(4), pp. 521–532.

Luckheeram, R. V., Zhou, R., Verma, A. D. and Xia, B. (2012) $CD4^+$ T Cells: Differentiation and Functions, *Clinical and Developmental Immunology*, **2012**, pp. 1–12. doi: 10.1155/2012/925135

Lunardi, C., Bason, C., Leandri, M., Navone, R., Lestani, M., Millo, E., Benatti, U., Cilli, M., Beri, R., Corrocher, R. and Puccetti, A. (2002) Autoantibodies to inner ear and endothelial antigens in Cogan's syndrome, *The Lancet*, **360**(9337), pp. 915–921.

Lunardi, C., Bason, C., Navone, R., Millo, E., Damonte, G., Corrocher, R. and Puccetti, A. (2000) Systemic sclerosis immunoglobulin G autoantibodies bind the human cytomegalovirus late protein UL94 and induce apoptosis in human endothelial cells, *Nature Medicine*, **6**(10), pp. 1183–1186.

Lunardi, C., Tinazzi, E., Bason, C., Dolcino, M., Corrocher, R. and Puccetti, A. (2008) Human parvovirus B19 infection and autoimmunity, *Autoimmunity Reviews*, **8**(2), pp. 116–120.

Machado-Alba, J. E., Ruiz, A. F. and Medina Morales, D. A. (2015) The epidemiology of rheumatoid arthritis in a cohort of Colombian patients, *Revista Colombiana de Reumatología*, **22**(3), pp. 148–152.

Madeira, A., Burgelin, I., Perron, H., Curtin, F., Lang, A. B. and Faucard, R. (2016) MSRV envelope protein is a potent, endogenous and pathogenic agonist of human toll-like receptor 4: Relevance of GNbAC1 in multiple sclerosis treatment, *Journal of Neuroimmunology*, **291**, pp. 29–38.

Magiorkinis, G. and Hurst, T. P. (2015) Activation of the innate immune response by endogenous retroviruses, *Journal of General Virology*, **96**(6), pp. 1207–1218.

Magistrelli, C., Samoilova, E., Agarwal, R. K., Banki, K., Ferrante, P., Vladutiu, A., Phillips, P. E. and Perl, A. (1999) Polymorphic genotypes of the HRES-1 human endogenous retrovirus locus correlate with systemic lupus erythematosus and autoreactivity, *Immunogenetics*, **49**(10), pp. 829–834.

Malati, T. (2001) An overview of monoclonal gammapathies, *Indian Journal of Clinical Biochemistry*, **16**(1), pp. 1–8.

Malmström, V., Catrina, A. I. and Klareskog, L. (2017) The immunopathogenesis of seropositive rheumatoid arthritis: from triggering to targeting, *Nature Reviews Immunology*, **17**(1), pp. 60–75.

Mameli, G., Erre, G. L., Caggiu, E., Mura, S., Cossu, D., Bo, M., Cadoni, M. L., Piras, A., Mundula, N., Colombo, E., Buscetta, G., Passiu, G. and Sechi, L. A. (2017) Identification of a HERV-K env surface peptide highly recognized in Rheumatoid

Arthritis (RA) patients: a cross-sectional case-control study, *Clinical and Experimental Immunology*, **189**(1), pp. 127–131.

Marchi, E., Kanapin, A., Magiorkinis, G. and Belshaw, R. (2014) Unfixed Endogenous Retroviral Insertions in the Human Population, *Journal of Virology*, **88**(17), pp. 9529–9537.

Marchionini, D. M., Lehrmann, E., Chu, Y., He, B., Sortwell, C. E., Becker, K. G., Freed, W. J., Kordower, J. H. and Collier, T. J. (2007) Role of heparin binding growth factors in nigrostriatal dopamine system development and Parkinson's disease, *Brain Research*, **1147**, pp. 77–88.

Marston, B., Palanichamy, A. and Anolik, J. H. (2010) B cells in the pathogenesis and treatment of rheumatoid arthritis, *Current Opinion in Rheumatology*, **22**(3), pp. 307–315.

Masi, A. T., Aldag, J. C. and Sipes, J. (2001) Do elevated levels of serum C-reactive protein predict rheumatoid arthritis in men: correlations with pre-RA status and baseline positive rheumatoid factors, *The Journal of rheumatology*, **28**(10), pp. 2359–2361.

Mastrangelo, A., Colasanti, T., Barbati, C., Pecani, A., Sabatinelli, D., Pendolino, M., Truglia, S., Massaro, L., Mancini, R., Miranda, F., Spinelli, F. R., Conti, F., Alessandri, C., Mastrangelo, A., Colasanti, T., Barbati, C., Pecani, A., Sabatinelli, D., Pendolino, M., Truglia, S., Massaro, L., Mancini, R., Miranda, F., Spinelli, F. R., Conti, F. and Alessandri, C. (2015) The Role of Posttranslational Protein Modifications in Rheumatological Diseases: Focus on Rheumatoid Arthritis, *Journal of Immunology Research*, **2015**, pp. 1–10. doi.org/10.1155/2015/712490

Matzaraki, V., Kumar, V., Wijmenga, C. and Zhernakova, A. (2017) The MHC locus and genetic susceptibility to autoimmune and infectious diseases, *Genome Biology*, **18**(1), pp. 76-96.

Mazumdar, B., Meyer, K., Ray, R., Rippe, R. and Friedman, S. (2012) N-Terminal Region of Gelsolin Induces Apoptosis of Activated Hepatic Stellate Cells by a Caspase-Dependent Mechanism, *PLoS ONE*, **7**(8), p. e44461. doi.org/10.1371/journal.pone.0044461

McClain, M. T., Heinlen, L. D., Dennis, G. J., Roebuck, J., Harley, J. B. and James, J. A. (2005) Early events in lupus humoral autoimmunity suggest initiation through molecular mimicry, *Nature Medicine*, **11**(1), pp. 85–89.

McDougall, C., Hurd, K. and Barnabe, C. (2017) Systematic review of rheumatic disease epidemiology in the indigenous populations of Canada, the United States, Australia, and New Zealand, *Seminars in Arthritis and Rheumatism*, **46**(5), pp. 675–686.

McInnes, I. B. and Schett, G. (2017) Pathogenetic insights from the treatment of rheumatoid arthritis, *The Lancet*, **389**(10086), pp. 2328–2337.

Mcnamee, K., Williams, R. and Seed, M. (2015) Animal models of rhematoid arthritis: how informative are they?, *European Journal of Pharmacology*, **759**, pp. 278-286.

Medina, K. L. (2016) Overview of the immune system, *Handbook of clinical neurology*, **133**, pp. 61–76.

Medstrand, P., Landry, J.-R. and Mager, D. L. (2001) Long Terminal Repeats Are Used as Alternative Promoters for the Endothelin B Receptor and Apolipoprotein C-I Genes in Humans, *Journal of Biological Chemistry*, **276**(3), pp. 1896–1903.

Medzhitov, R. and Janeway, C. A. (2002) Decoding the patterns of self and nonself by the innate immune system, *Science,* **296**(5566), pp. 298–300.

Meednu, N., Zhang, H., Owen, T., Sun, W., Wang, V., Cistrone, C., Rangel-Moreno, J., Xing, L. and Anolik, J. H. (2016) Production of RANKL by Memory B Cells: A Link Between B Cells and Bone Erosion in Rheumatoid Arthritis, *Arthritis and Rheumatology*, **68**(4), pp. 805–816.

Meischl, C., de Boer, M., Åhlin, A. and Roos, D. (2000) A new exon created by intronic insertion of a rearranged LINE-1 element as the cause of chronic granulomatous disease, *European Journal of Human Genetics*, **8**(9), pp. 697–703.

Menard, L., Samuels, J., Ng, Y.-S. and Meffre, E. (2011) Inflammation-independent defective early B cell tolerance checkpoints in rheumatoid arthritis, *Arthritis and rheumatism*, **63**(5), pp. 1237–1245.

Metzger, T. C. and Anderson, M. S. (2011) Control of central and peripheral tolerance by Aire, *Immunological reviews*, **241**(1), pp. 89–103.

Meyer, T. J., Rosenkrantz, J. L., Carbone, L. and Chavez, S. L. (2017) Endogenous Retroviruses: With Us and against Us, *Frontiers in Chemistry*, **5**, pp. 23-30. doi: 10.3389/fchem.2017.00023

Minami, Y., Kono, T., Miyazaki, T. and Taniguchi, T. (1993) The IL-2 Receptor Complex: Its Structure, Function, and Target Genes, *Annual Review of Immunology*, **11**(1), pp. 245–268.

Mjaavatten, M. D., van der Heijde, D. M., Uhlig, T., Haugen, A. J., Nygaard, H., Bjørneboe, O. and Kvien, T. K. (2011) Should anti-citrullinated protein antibody and rheumatoid factor status be reassessed during the first year of followup in recent-onset arthritis? A longitudinal study, *The Journal of rheumatology*, **38**(11), pp. 2336–2341.

Montecino-Rodriguez, E. and Dorshkind, K. (2012) B-1 B cell development in the fetus and adult, *Immunity*, **36**(1), pp. 13–21.

Morgan, D. and Brodsky, I. (2004) Human endogenous retrovirus (HERV-K) particles in megakaryocytes cultured from essential thrombocythemia peripheral blood stem cells, *Experimental Hematology*, **32**(6), pp. 520–525.

Morita, T., Shima, Y., Wing, J. B., Sakaguchi, S., Ogata, A. and Kumanogoh, A. (2016) The Proportion of Regulatory T Cells in Patients with Rheumatoid Arthritis: A Meta-Analysis, *PLoS ONE*, **11**(9), p. e0162306. doi: 10.1371/journal.pone.0162306

Morozov, V. A., Dao Thi, V. L. and Denner, J. (2013) The Transmembrane Protein of the Human Endogenous Retrovirus - K (HERV-K) Modulates Cytokine Release and Gene Expression, *PLoS ONE*, **8**(8), p. e70399. doi.org/10.1371/journal.pone.0070399

Motea, E. A. and Berdis, A. J. (2010) Terminal Deoxynucleotidyl Transferase: The Story of a Misguided DNA Polymerase, *Biochimica et biophysica acta*, **1804**(5), pp. 1151-1166.

Moult, J. (2005) A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction, *Current Opinion in Structural Biology*, **15**(3), pp. 285–289.

Moyes, D., Griffiths, D. J. and Venables, P. J. (2007) Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease, *Trends in Genetics*, **23**(7), pp. 326–333.

Mueller, D. L. (2010) Mechanisms maintaining peripheral tolerance, *Nature Immunology*, **11**(1), pp. 21–27.

Mukonyora, M. (2015) A Review of Important Discontinuous B-Cell Epitope Prediction Tools, *Journal of Clinical and Cellular Immunology*, **6**(5), pp. 1–5.

Müllbacher, A., Lobigs, M., Hla, R. T., Tran, T., Stehle, T. and Simon, M. M. (2002) Antigen-dependent release of IFN-gamma by cytotoxic T cells up-regulates Fas on target cells and facilitates exocytosis-independent specific target cell lysis, *Journal of immunology,* **169**(1), pp. 145–150.

Muller, W. A. (2013) Getting leukocytes to the site of inflammation., *Veterinary pathology*, **50**(1), pp. 7–22.

Münch, G., Boivin-Jahns, V., Holthoff, H.-P., Adler, K., Lappo, M., Truöl, S., Degen, H., Steiger, N., Lohse, M. J., Jahns, R. and Ungerer, M. (2012) Administration of the cyclic peptide COR-1 in humans (phase I study): *ex vivo* measurements of anti-β_1-adrenergic receptor antibody neutralization and of immune parameters, *European Journal of Heart Failure*, **14**(11), pp. 1230–1239.

Murphy, D., Mattey, D. and Hutchinson, D. (2017) Anti-citrullinated protein antibody positive rheumatoid arthritis is primarily determined by rheumatoid factor titre and the shared epitope rather than smoking per se, *PLOS ONE*, **12**(7), p. e0180655. doi: 10.1371/journal.pone.0180655

Nakagawa, K., Brusic, V., McColl, G. and Harrison, L. C. (1997) Direct evidence for the expression of multiple endogenous retroviruses in the synovial compartment in rheumatoid arthritis, *Arthritis and rheumatism*, **40**(4), pp. 627–638.

National Institute for Clinical Excellence (NICE) (2009) *Rheumatoid arthritis in adults: management.* [online] Available from: https://www.nice.org.uk/guidance/cg79/resources/rheumatoid-arthritis-in-adults-management-pdf-975636823525 (Accessed 3 February 2018).

Nauclér, C. S., Larsson, S. and Möller, E. (1996) A novel mechanism for virus-induced autoimmunity in humans, *Immunological reviews*, **152**, pp. 175–192.

Nelson, P., Freimanis, G. and Roden, D. (2008) Human Endogenous Retroviruses; Evolutionary Dynamics, Chromosomal Location and Host Benefit, In *Encyclopedia of Life Sciences*. Chichester: John Wiley & Sons.

Nelson, P. N., Fletcher, S. M., de Lange, G. G., van Leeuwen, A. M., Goodall, M. and Jefferis, R. (1990) Evaluation of monoclonal antibodies with putative specificity for human IgG allotypes, *Vox sanguinis*, **59**(3), pp. 190–197.

Nelson, P. N., Hooley, P., Roden, D., Davari Ejtehadi, H., Rylance, P. B., Warren, P., Martin, J. and Murray, P. G. (2004) Human endogenous retroviruses: transposable elements with potential?, *Clinical and Experimental Immunology*, **138**(1), pp. 1–9.

Nelson, P. N., Lever, A. M. L., Bruckner, F. E., Isenberg, D. A., Kessaris, N., Hay, F. C., Bruckner, F. E. and Nelson, P. N. (1994) Polymerase chain reaction fails to incriminate exogenous retroviruses HTLV-I and HIV-1 in rheumatological diseases although a minority of sera cross react with retroviral antigens, *Annals of the Rheumatic Diseases*, **53**, pp. 749–754.

Nelson, P. N., Lever, A. M. L., Smith, S., Pitman, R., Murray, P., Perera, S. A., Westwood, O. M. R., Hay, F. C., Ejtehadi, H. D. and Booth, J. C. (1999) Molecular Investigations Implicate Human Endogenous Retroviruses as Mediators of Anti-Retroviral Antibodies in Autoimmune Rheumatic Disease, *Immunological Investigations*, **28**(4), pp. 277–289.

Nelson, P. N., Roden, D., Nevill, A., Freimanis, G. L., Trela, M., Ejtehadi, H. D., Bowman, S., Axford, J., Veitch, A. M., Tugnet, N. and Rylance, P. B. (2014a) Rheumatoid arthritis is associated with IgG antibodies to human endogenous retrovirus Gag matrix: A potential pathogenic mechanism of disease?, *Journal of Rheumatology*, **41**(10), pp. 1952–1960.

Nelson, P. N., Westwood, O. M. R., Soltys, A., Jefferis, R., Goodall, M., Baumforth, K. R. N., Frampton, G., Tribbick, G., Roden, D. and Hay, F. C. (2003) Characterisation of epitopes of pan-IgG/anti-G3m(u) and anti-Fc monoclonal antibodies, *Immunology Letters*, **88**(1), pp. 77–83.

Nelson, P., Rylance, P., Roden, D., Trela, M. and Tugnet, N. (2014b) Viruses as potential pathogenic agents in systemic lupus erythematosus, *Lupus*, **23**(6), pp. 596–605.

Nelson, P., Westwood, O., Jefferis, R., Goodall, M. and Hay, F. (1997) Characterisation of anti-IgG monoclonal antibody A57H by epitope mapping, *Biochemical Society Transactions*, **25**(2), p. 373.

Nemazee, D. (2017) Mechanisms of central tolerance for B cells, *Nature Reviews Immunology*, **17**(5), pp. 281–294.

Nielen, M. M. J., van Schaardenburg, D., Reesink, H. W., van de Stadt, R. J., van der Horst-Bruinsma, I. E., de Koning, M. H. M. T., Habibuw, M. R., Vandenbroucke, J. P. and Dijkmans, B. A. C. (2004) Specific autoantibodies precede the symptoms of rheumatoid arthritis: A study of serial measurements in blood donors, *Arthritis and Rheumatism*, **50**(2), pp. 380–386.

Nikoloudis, D., Pitts, J. E. and Saldanha, J. W. (2014) A complete, multi-level conformational clustering of antibody complementarity-determining regions, *PeerJ*, **2**, p. e456. doi: 10.7717/peerj.456

Noack, M. and Miossec, P. (2014) Th17 and regulatory T cell balance in autoimmune and inflammatory diseases, *Autoimmunity Reviews,* **213**(6), pp. 668-677.

Oduwole, O. O., Li, Y., Isomaa, V. V., Mäntyniemi, A., Pulkka, A. E., Soini, Y. and Vihko, P. T. (2004) 17β-Hydroxysteroid Dehydrogenase Type 1 Is an Independent Prognostic Marker in Breast Cancer, *Cancer Research*, **64**(20), pp. 7604–7609.

Oldstone, M. B. (1998) Molecular mimicry and immune-mediated diseases., *FASEB journal,* **12**(13), pp. 1255–1265.

do Olival, G. S., Faria, T. S., Nali, L. H. S., de Oliveira, A. C. P., Casseb, J., Vidal, J. E., Cavenaghi, V. B., Tilbery, C. P., Moraes, L., Fink, M. C. S., Sumita, L. M., Perron, H. and Romano, C. M. (2013) Genomic analysis of ERVWE2 locus in patients with multiple sclerosis: absence of genetic association but potential role of human endogenous retrovirus type W elements in molecular mimicry with myelin antigen, *Frontiers in microbiology*, **4**, pp. 172-178.

Ono, M., Yasunaga, T., Miyata, T. and Ushikubo, H. (1986) Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome, *Journal of virology*, **60**(2), pp. 589–598.

Ossipova, E., Cerqueira, C. F., Reed, E., Kharlamova, N., Israelsson, L., Holmdahl, R., Nandakumar, K. S., Engström, M., Harre, U., Schett, G., Catrina, A. I., Malmström, V., Sommarin, Y., Klareskog, L., Jakobsson, P.-J. and Lundberg, K. (2014) Affinity purified anti-citrullinated protein/peptide antibodies target antigens expressed in the rheumatoid joint, *Arthritis research and therapy*, **16**(4), pp. R167-177.

Overgaard, M. T., Boldt, H. B., Laursen, L. S., Sottrup-Jensen, L., Conover, C. A. and Oxvig, C. (2001) Pregnancy-associated Plasma Protein-A2 (PAPP-A2), a Novel Insulin-like Growth Factor-binding Protein-5 Proteinase, *Journal of Biological Chemistry*, **276**(24), pp. 21849–21853.

Panda, S. and Ding, J. L. (2015) Natural antibodies bridge innate and adaptive immunity, *Journal of immunology,* **194**(1), pp. 13–20.

Parker, J. M. R., Guo, D. and Hodges, R. S. (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites, *Biochemistry*, **25**(19), pp. 5425–5432.

Parkin, J. and Cohen, B. (2001) An overview of the immune system, *The Lancet*, **357**(9270), pp. 1777–1789.

Pascual, M., Martin, J., Nieto, A., Giphart, M. J., van der Slik, A. R., de Vries, R. R. and Zanelli, E. (2001) Distribution of HERV-LTR elements in the 5'-flanking region of HLA-DQB1 and association with autoimmunity, *Immunogenetics*, **53**(2), pp. 114–118.

Patience, C., Takeuchi, Y., Cosset, F. L. and Weiss, R. A. (1998) Packaging of endogenous retroviral sequences in retroviral vectors produced by murine and human packaging cells, *Journal of virology*, **72**(4), pp. 2671–2676.

Patra, K. P., Saito, M., Atluri, V. L., Rolán, H. G., Young, B., Kerrinnes, T., Smits, H., Ricaldi, J. N., Gotuzzo, E., Gilman, R. H., Tsolis, R. M. and Vinetz, J. M. (2014) A Protein-Conjugate Approach to Develop a Monoclonal Antibody-Based Antigen Detection Test for the Diagnosis of Human Brucellosis, *PLoS Neglected Tropical Diseases*, **8**(6), p. e2926. doi: 10.1371/journal.pntd.0002926

Pawłowska, J., Smoleńska, Z., Daca, A., Witkowski, J. M. and Bryl, E. (2011) Older age of rheumatoid arthritis onset is associated with higher activation status of peripheral blood CD4(+) T cells and disease activity, *Clinical and experimental immunology*, **163**(2), pp. 157–164.

Pelanda, R. and Torres, R. M. (2012) Central B-cell tolerance: where selection begins, *Cold Spring Harbor perspectives in biology*, **4**(4), p. a007146. doi: 10.1101/cshperspect.a007146.

Pellequer, J. L., Westhof, E. and Van Regenmortel, M. H. (1993) Correlation between the location of antigenic sites and the prediction of turns in proteins, *Immunology letters*, **36**(1), pp. 83–99.

Perl, A. (1999) Mechanisms of viral pathogenesis in rheumatic disease, *Annals of the Rheumatic Diseases*, **58**, pp. 454–461.

Perl, A. (2003) Role of endogenous retroviruses in autoimmune diseases, *Rheumatic Disease Clinics*, **29**(1), pp. 123–143

Perl, A. (2012) Pathogenesis and Spectrum of Autoimmunity, *Methods in molecular biology*, **900**, pp. 1–9.

Petrelli, A. and van Wijk, F. (2016) CD8+ T cells in human autoimmune arthritis: the unusual suspects, *Nature Reviews Rheumatology*, **12**(7), pp. 421–428.

Pi, W., Zhu, X., Wu, M., Wang, Y., Fulzele, S., Eroglu, A., Ling, J. and Tuan, D. (2010) Long-range function of an intergenic retrotransposon, *Proceedings of the National Academy of Sciences*, **107**(29), pp. 12992–12997.

Pieper, K., Grimbacher, B. and Eibel, H. (2013) B-cell biology and development, *The Journal of allergy and clinical immunology*, **131**(4), pp. 959–971.

Piper, K. E., Hanssen, A. D., Lewallen, D. G., Matteson, E. L., Osmon, D. R., Duffy, M. C., Hagan, R. A., Steckelberg, J. M. and Patel, R. (2006) Lack of detection of human retrovirus-5 proviral DNA in synovial tissue and blood specimens from individuals with rheumatoid arthritis or osteoarthritis, *Arthritis and rheumatism*, **55**(1), pp. 123–125.

Plot, L. and Amital, H. (2009) Infectious associations of Celiac disease, *Autoimmunity Reviews*, **8**(4), pp. 316–319.

Ponchel, F., Vital, E., Kingsbury, S. and El-Sherbiny, Y. (2012) CD4+ T cell subsets in rheumatoid arthritis, *International Journal of Clinical Rheumatology*, **7**(1), pp. 37–53.

Ponnuswamy, P. K., Prabhakaran, M. and Manavalan, P. (1980) Hydrophobic packing and spatial arrangement of amino acid residues in globular proteins, *Biochimica et biophysica acta*, **623**(2), pp. 301–316.

Ponomarenko, J. V and Van Regenmortel, M. H. V (2009) B-cell epitope prediction, In Gu, J. and Bourne, P.E. (eds.) *Structural Bioinformatics*, 2nd edn. Chichester: John Wiley & Sons, pp. 849–872.

Potocnakova, L., Bhide, M. and Pulzova, L. B. (2016) An Introduction to B-Cell Epitope Mapping and In Silico Epitope Prediction, *Journal of Immunology Research*, **2016**, pp. 1–11. doi: 10.1155/2016/6760830.

Pratesi, F., Dioni, I., Tommasi, C., Alcaro, M. C., Paolini, I., Barbetti, F., Boscaro, F., Panza, F., Puxeddu, I., Rovero, P. and Migliorini, P. (2014) Antibodies from patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps, *Annals of the Rheumatic Diseases*, **73**(7), pp. 1414–1422.

Pratesi, F., Petit Teixeira, E., Sidney, J., Michou, L., Puxeddu, I., Sette, A., Cornelis, F. and Migliorini, P. (2013) HLA shared epitope and ACPA: Just a marker or an active player?, *Autoimmunity Reviews*, **12**(12), pp. 1182–1187.

Pruijn, G. J. M. (2015) Citrullination and Carbamylation in the Pathophysiology of Rheumatoid Arthritis, *Frontiers in Immunology*, **6**, pp. 192-196.

Quintero, O. L., Amador-Patarroyo, M. J., Montoya-Ortiz, G., Rojas-Villarraga, A. and Anaya, J.-M. (2012) Autoimmune disease and gender: Plausible mechanisms for the female predominance of autoimmunity, *Journal of Autoimmunity*, **38**(2–3), pp. J109–J119.

Quirke, A.-M., Lugli, E. B., Wegner, N., Hamilton, B. C., Charles, P., Chowdhury, M., Ytterberg, A. J., Zubarev, R. A., Potempa, J., Culshaw, S., Guo, Y., Fisher, B. A., Thiele, G., Mikuls, T. R. and Venables, P. J. (2014) Heightened immune response to autocitrullinated Porphyromonas gingivalis peptidylarginine deiminase: a potential mechanism for breaching immunologic tolerance in rheumatoid arthritis, *Annals of the rheumatic diseases*, **73**(1), pp. 263–269.

Rahman, K. S., Chowdhury, E. U., Sachse, K. and Kaltenboeck, B. (2016) Inadequate Reference Datasets Biased toward Short Non-epitopes Confound B-cell Epitope Prediction, *Journal of Biological Chemistry*, **291**(28), pp. 14585–14599.

Raker, V. K., Domogalla, M. P. and Steinbrink, K. (2015) Tolerogenic Dendritic Cells for Regulatory T Cell Induction in Man, *Frontiers in Immunology*, **6**, pp. 569-579.

Ramasamy, R., Joseph, B. and Whittall, T. (2017) Potential molecular mimicry between the human endogenous retrovirus W family envelope proteins and myelin proteins in multiple sclerosis, *Immunology Letters*, **183**, pp. 79–85.

Ramwadhdoebe, T. H., Hähnlein, J., van Kuijk, B. J., Choi, I. Y., van Boven, L. J., Gerlag, D. M., Tak, P. P. and van Baarsen, L. G. (2016) Human lymph-node CD8(+) T cells display an altered phenotype during systemic autoimmunity, *Clinical and translational immunology*, **5**(4), p. e67. doi: 10.1038/cti.2016.8

Rantapää-Dahlqvist, S., de Jong, B. A. W., Berglin, E., Hallmans, G., Wadell, G., Stenlund, H., Sundin, U. and van Venrooij, W. J. (2003) Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis, *Arthritis and Rheumatism*, **48**(10), pp. 2741–2749.

Rashid, T. and Ebringer, A. (2012) Autoimmunity in Rheumatic Diseases Is Induced by Microbial Infections via Crossreactivity or Molecular Mimicry, *Autoimmune Diseases*, **2012**, pp. 1–9. doi: 10.1155/2012/539282

Rawlings, D. J., Schwartz, M. A., Jackson, S. W. and Meyer-Bahlburg, A. (2012) Integration of B cell responses through Toll-like receptors and antigen receptors, *Nature reviews. Immunology*, **12**(4), pp. 282–294.

Raza, K. and Filer, A. (2015) The therapeutic window of opportunity in rheumatoid arthritis: does it ever close?, *Annals of the rheumatic diseases*, **74**(5), pp. 793–794.

De Re, V., Sansonno, D., Simula, M. P., Caggiari, L., Gasparotto, D., Fabris, M., Tucci, F. A., Racanelli, V., Talamini, R., Campagnolo, M., Geremia, S., Dammacco, F. and De Vita, S. (2006) HCV-NS3 and IgG-Fc crossreactive IgM in patients with type II mixed cryoglobulinemia and B-cell clonal proliferations, *Leukemia*, **20**(6), pp. 1145–1154.

Regenmortel, M. H. V. (2009) What Is a B-Cell Epitope?, *Methods in molecular biology*, **524**, pp. 3–20.

Rein, P. and Mueller, R. B. (2017) Treatment with Biologicals in Rheumatoid Arthritis: An Overview, *Rheumatology and Therapy*, **4**(2), pp. 247–261.

Renaudineau, Y., Vallet, S., Le Dantec, C., Hillion, S., Saraux, A. and Youinou, P. (2005) Characterization of the human CD5 endogenous retrovirus-E in B lymphocytes, *Genes and Immunity*, **6**(8), pp. 663–671.

Reynier, F., Verjat, T., Turrel, F., Imbert, P. E., Marotte, H., Mougin, B. and Miossec, P. (2009) Increase in Human Endogenous Retrovirus HERV-K (HML-2) Viral Load in Active Rheumatoid Arthritis, *Scandinavian Journal of Immunology*, **70**(3), pp. 295–299.

Roberts, C. A., Dickinson, A. K. and Taams, L. S. (2015) The Interplay Between Monocytes/Macrophages and CD4+ T Cell Subsets in Rheumatoid Arthritis, *Frontiers in Immunology*, **6**, pp. 571-589.

Robson, S. C., Sévigny, J. and Zimmermann, H. (2006) The E-NTPDase family of ectonucleotidases: Structure function relationships and pathophysiological significance, *Purinergic Signalling*, **2**(2), pp. 409–430.

Rohrbach, A. S., Hemmers, S., Arandjelovic, S., Corr, M. and Mowen, K. A. (2012) PAD4 is not essential for disease in the K/BxN murine autoantibody-mediated model of arthritis, *Arthritis research and therapy*, **14**(3), pp. R104-113.

Rohrer, J., Minegishi, Y., Richter, D., Eguiguren, J. and Conley, M. E. (1999) Unusual Mutations in Btk: An Insertion, a Duplication, an Inversion, and Four Large Deletions, *Clinical Immunology*, **90**(1), pp. 28–37.

Romagnani, S. (2006) Immunological tolerance and autoimmunity, *Internal and Emergency Medicine*, **1**(3), pp. 187–196.

Romanish, M. T., Lock, W. M., de Lagemaat, L. N. van, Dunn, C. A. and Mager, D. L. (2007) Repeated Recruitment of LTR Retrotransposons as Promoters by the Anti-Apoptotic Locus NAIP during Mammalian Evolution, *PLoS Genetics*, **3**(1), p. e10. doi: 10.1371/journal.pgen.0030010

Romanowska-Prochnicka, K., Rzodkiewicz, P., Olesinska, M., Szukiewicz, D. and Maslinski, S. (2013) Extraskeletal Manifestations in Rheumatoid Arthritis - Clinical Cases, In Matsuno, H. *Innovative Rheumatology*, InTech, [online] Available from: http://www.intechopen.com/books/innovative-rheumatology/extraskeletal-manifestations-in-rheumatoid-arthritis-clinical-cases (Accessed 8 February 2018).

Rosas, E. C., Correa, L. B. and Henriques, M. das G. (2017) Neutrophils in Rheumatoid Arthritis: A Target for Discovering New Therapies Based on Natural Products, In Khajah, M.A. *Role of Neutrophils in Disease Pathogenesis*, InTech, [online] Available from:http://www.intechopen.com/books/role-of-neutrophils-in-disease-pathogenesis/neutrophils-in-rheumatoid-arthritis-a-target-for-discovering-new-therapies-based-on-natural-products (Accessed 5 February 2018).

lRosenberg-Hasson, Y., Hansmann, L., Liedtke, M., Herschmann, I. and Maecker, H. T. (2014) Effects of serum and plasma matrices on multiplex immunoassays, *Immunologic research*, **58**(2–3), pp. 224–233.

Rosenblum, M. D., Remedios, K. A. and Abbas, A. K. (2015) Mechanisms of human autoimmunity, *Journal of Clinical Investigation*, **125**(6), pp. 2228–2233.

Ruprecht, K., Obojes, K., Wengel, V., Gronen, F., Kim, K. S., Perron, H., Schneider-Schaulies, J. and Rieckmann, P. (2006) Regulation of human endogenous retrovirus W protein expression by herpes simplex virus type 1: Implications for multiple sclerosis, *Journal of Neurovirology*, **12**(1), pp. 65–71.

Ryan, F. P. (2016) Viral symbiosis and the holobiontic nature of the human genome, *APMIS*, **124**(1–2), pp. 11–19.

Saha, S. and Raghava, G. P. S. (2004) BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties, In Cutello, N.G., Bentley, P.J. and Timmis, J. (eds) *Artificial Immune Systems*. Berlin: Springer, pp. 197–204.

Samuelson, L. C., Phillips, R. S. and Swanberg, L. J. (1996) Amylase gene structures in primates: retroposon insertions and promoter evolution, *Molecular Biology and Evolution*, **13**(6), pp. 767–779.

Sanchez-Trincado, J. L., Gomez-Perosanz, M. and Reche, P. A. (2017) Fundamentals and Methods for T- and B-Cell Epitope Prediction, *Journal of Immunology Research*, **2017**, pp. 1–14. doi: 10.1155/2017/2680160

Saper, C. B. (2009) A guide to the perplexed on the specificity of antibodies, *The journal of histochemistry and cytochemistry,* **57**(1), pp. 1–5.

Saroha, A., Kumar, S., Chatterjee, B. P. and Das, H. R. (2012) Jacalin Bound Plasma O-Glycoproteome and Reduced Sialylation of Alpha 2-HS Glycoprotein (A2HSG) in Rheumatoid Arthritis Patients, *PLoS ONE*, **7**(10), p. e46374. doi: 10.1371/journal.pone.0046374

Sauter, M., Schommer, S., Kremmer, E., Remberger, K., Dölken, G., Lemm, I., Buck, M., Best, B., Neumann-Haefelin, D. and Mueller-Lantzsch, N. (1995) Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas, *Journal of virology*, **69**(1), pp. 414–421.

Scherer, H. U., van der Woude, D., Ioan-Facsinay, A., el Bannoudi, H., Trouw, L. A., Wang, J., Häupl, T., Burmester, G.-R., Deelder, A. M., Huizinga, T. W. J., Wuhrer, M. and Toes, R. E. M. (2010) Glycan profiling of anti-citrullinated protein antibodies

isolated from human serum and synovial fluid, *Arthritis and Rheumatism*, **62**(6), pp. 1620–1629.

Schiffenbauer, J. (1999) Superantigens and Their Role in Autoimmune Disorders, *Archivum Immunologiae et Therapiae Experimentalis*, **47**, pp. 17–24.

Schroeder, D. I., Jayashankar, K., Douglas, K. C., Thirkill, T. L., York, D., Dickinson, P. J., Williams, L. E., Samollow, P. B., Ross, P. J., Bannasch, D. L., Douglas, G. C. and LaSalle, J. M. (2015) Early Developmental and Evolutionary Origins of Gene Body DNA Methylation Patterns in Mammalian Placentas, *PLOS Genetics*, **11**(8), p. e1005442. doi.org/10.1371/journal.pgen.1005442

Schwimmbeck, P. L., Dyrberg, T., Drachman, D. B. and Oldstone, M. B. (1989) Molecular mimicry and myasthenia gravis. An autoantigenic site of the acetylcholine receptor alpha-subunit that has biologic activity and reacts immunochemically with herpes simplex virus., *Journal of Clinical Investigation*, **84**(4), pp. 1174–1180.

Seidl, C., Donner, H., Petershofen, E., Usadel, K. H., Seifried, E., Kaltwasser, J. P. and Badenhoop, K. (1999) An endogenous retroviral long terminal repeat at the HLA-DQB1 gene locus confers susceptibility to rheumatoid arthritis, *Human immunology*, **60**(1), pp. 63–68.

Sela-Culang, I., Kunik, V. and Ofran, Y. (2013) The Structural Basis of Antibody-Antigen Recognition, *Frontiers in Immunology*, **4**, pp. 302-314.

Selders, G. S., Fetz, A. E., Radic, M. Z. and Bowlin, G. L. (2017) An overview of the role of neutrophils in innate immunity, inflammation and host-biomaterial integration., *Regenerative biomaterials*, **4**(1), pp. 55–68.

Selmi, C., Bowlus, C. L., Gershwin, M. E. and Coppel, R. L. (2011) Primary biliary cirrhosis, *The Lancet*, **377**(9777), pp. 1600–1609.

Serra, M. B., Barroso, W. A., Silva, N. N. da, Silva, S. do N., Borges, A. C. R., Abreu, I. C. and Borges, M. O. da R. (2017) From Inflammation to Current and Alternative Therapies Involved in Wound Healing, *International Journal of Inflammation*, **2017**, pp. 1–17. doi: 10.1155/2017/3406215

Sfriso, P., Ghirardello, A., Botsios, C., Tonon, M., Zen, M., Bassi, N., Bassetto, F. and Doria, A. (2010) Infections and autoimmunity: the multifaceted relationship, *Journal of leukocyte biology*, **87**(3), pp. 385–395.

Shachter, N. S. (2001) Apolipoproteins C-I and C-III as important modulators of lipoprotein metabolism, *Current opinion in lipidology*, **12**(3), pp. 297–304.

Shafqat, N., Shafqat, J., Eissner, G., Marschall, H.-U., Tryggvason, K., Eriksson, U., Gabrielli, F., Lardy, H., Jörnvall, H. and Oppermann, U. (2006) Hep27, a member of the short-chain dehydrogenase/reductase family, is an NADPH-dependent dicarbonyl reductase expressed in vascular endothelial tissue, *Cellular and Molecular Life Sciences*, **63**(10), pp. 1205–1213.

Shechter, R., London, A. and Schwartz, M. (2013) Orchestrated leukocyte recruitment to immune-privileged sites: absolute barriers versus educational gates, *Nature Reviews Immunology*, **13**(3), pp. 206–218.

Shi, J., van Veelen, P. A., Mahler, M., Janssen, G. M. C., Drijfhout, J. W., Huizinga, T. W. J., Toes, R. E. M. and Trouw, L. A. (2014) Carbamylation and antibodies against carbamylated proteins in autoimmunity and other pathologies, *Autoimmunity Reviews*, **13**(3), pp. 225–230.

Shimoda, S., Nakamura, M., Ishibashi, H., Kawano, A., Kamihira, T., Sakamoto, N., Matsushita, S., Tanaka, A., Worman, H. J., Gershwin, M. E. and Harada, M. (2003) Molecular mimicry of mitochondrial and nuclear autoantigens in primary biliary cirrhosis, *Gastroenterology*, **124**(7), pp. 1915–1925.

Shimode, S., Nakagawa, S. and Miyazawa, T. (2015) Multiple invasions of an infectious retrovirus in cat genomes, *Scientific Reports*, **5**(1), pp. 8164-8173.

Shrake, A. and Rupley, J. A. (1973) Environment and exposure to solvent of protein atoms. Lysozyme and insulin, *Journal of Molecular Biology*, **79**(2), pp. 351–371.

Sicat, J., Sutkowski, N. and Huber, B. T. (2005) Expression of human endogenous retrovirus HERV-K18 superantigen is elevated in juvenile rheumatoid arthritis, *The Journal of rheumatology*, **32**(9), pp. 1821–1831.

Sigdel, T. K., Li, L., Tran, T. Q., Khatri, P., Naesens, M., Sansanwal, P., Dai, H., Hsieh, S.-C. and Sarwal, M. M. (2012) Non-HLA Antibodies to Immunogenic Epitopes Predict the Evolution of Chronic Renal Allograft Injury, *Journal of the Americal Society of Nephrology*, **23**(4), pp. 750-763.

Simmons, W. (2016) The Role of Human Endogenous Retroviruses (HERV-K) in the Pathogenesis of Human Cancers, *Molecular Biology*, **5**(3), pp. 169-174.

Singh, U., Singh, S., Singh, N. K., Verma, P. K. and Singh, S. (2011) Anticyclic citrullinated peptide autoantibodies in systemic lupus erythematosus, *Rheumatology International*, **31**(6), pp. 765–767.

Sivalingam, G. N. and Shepherd, A. J. (2012) An analysis of B-cell epitope discontinuity, *Molecular immunology*, **51**(3–4), pp. 304–309.

Skwarczynski, M. and Tóth, I. (2016) *Micro- and nanotechnology in vaccine development.* Amsterdam: Elsevier, pp. 197–204.

Smith, M. D. (2011) The normal synovium, *The Open rheumatology journal*, **5**, pp. 100–106.

Smolen, J. S., Aletaha, D., Barton, A., Burmester, G. R., Emery, P., Firestein, G. S., Kavanaugh, A., McInnes, I. B., Solomon, D. H., Strand, V. and Yamamoto, K. (2018) Rheumatoid arthritis, *Nature Reviews Disease Primers*, **4**, p. 18001.8. doi: 10.1038/nrdp.2018.1.

Smolen, J. S., Aletaha, D. and McInnes, I. B. (2016) Rheumatoid arthritis, *The Lancet*, **388**(10055), pp. 2023–2038.

Smolen, J. S., Landewé, R., Bijlsma, J., Burmester, G., Chatzidionysiou, K., Dougados, M., Nam, J., Ramiro, S., Voshaar, M., van Vollenhoven, R., Aletaha, D., Aringer, M., Boers, M., Buckley, C. D., Buttgereit, F., Bykerk, V., Cardiel, M., Combe, B., Cutolo, M., van Eijk-Hustings, Y., Emery, P., Finckh, A., Gabay, C., Gomez-Reino, J., Gossec, L., Gottenberg, J.-E., Hazes, J. M. W., Huizinga, T., Jani, M., Karateev, D., Kouloumas, M., Kvien, T., Li, Z., Mariette, X., McInnes, I., Mysler, E., Nash, P., Pavelka, K., Poór, G., Richez, C., van Riel, P., Rubbert-Roth, A., Saag, K., da Silva, J., Stamm, T., Takeuchi, T., Westhovens, R., de Wit, M. and van der Heijde, D. (2017) EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update, *Annals of the rheumatic diseases*, **76**(6), pp. 960–977.

Søe, K., Andersen, T. L., Hobolt-Pedersen, A.-S., Bjerregaard, B., Larsson, L.-I. and Delaissé, J.-M. (2011) Involvement of human endogenous retroviral syncytin-1 in human osteoclast fusion, *Bone*, **48**(4), pp. 837–846.

Sokolove, J., Johnson, D. S., Lahey, L. J., Wagner, C. A., Cheng, D., Thiele, G. M., Michaud, K., Sayles, H., Reimold, A. M., Caplan, L., Cannon, G. W., Kerr, G., Mikuls, T. R. and Robinson, W. H. (2014) Rheumatoid Factor as a Potentiator of Anti-

Citrullinated Protein Antibody-Mediated Inflammation in Rheumatoid Arthritis, *Arthritis and Rheumatology*, **66**(4), pp. 813–821.

Sokolove, J., Zhao, X., Chandra, P. E. and Robinson, W. H. (2011) Immune complexes containing citrullinated fibrinogen costimulate macrophages via Toll-like receptor 4 and Fcγ receptor, *Arthritis and rheumatism*, **63**(1), pp. 53–62.

Solanki, L. S., Srivastava, N. and Singh, S. (2008) Superantigens: a brief review with special emphasis on dermatologic diseases, *Dermatology online journal*, **14**(2), p. 3 [online] Available from: https://escholarship.org/uc/item/47g8w51m#main (Accessed 5 February 2018).

Song, Y. W. and Kang, E. H. (2010a) Autoantibodies in rheumatoid arthritis: rheumatoid factors and anticitrullinated protein antibodies., *QJM : monthly journal of the Association of Physicians*, **103**(3), pp. 139–146.

Song, Y. W. and Kang, E. H. (2010b) The pathogenic role of rheumatoid factor in rheumatoid arthritis, *International Journal of Clinical Rheumatology*, **5**(6), pp. 651–658.

Sophia, S. and Ramesha, M. M. (2017) An updated overview of immune complex mediated rheumatoid arthritis, *International Journal of Research in Medical Sciences*, **2**(2), pp. 398–403.

Soria-Guerra, R. E., Nieto-Gomez, R., Govea-Alonso, D. O. and Rosales-Mendoza, S. (2015) An overview of bioinformatics tools for epitope prediction: Implications on vaccine development, *Journal of Biomedical Informatics*, **53**, pp. 405–414.

Spinelli, F. R., Pecani, A., Conti, F., Mancini, R., Alessandri, C. and Valesini, G. (2016) Post-translational modifications in rheumatoid arthritis and atherosclerosis: Focus on citrullination and carbamylation, *Journal of International Medical Research*, **44**(1_suppl), pp. 81–84.

Stake, M., Bann, D., Kaddis, R. and Parent, L. (2013) Nuclear Trafficking of Retroviral RNAs and Gag Proteins during Late Steps of Replication, *Viruses*, **5**(11), pp. 2767-2795.

Stathopoulou, E. A., Routsias, J. G., Stea, E. A., Moutsopoulos, H. M. and Tzioufas, A. G. (2005) Cross-reaction between antibodies to the major epitope of Ro60 kD autoantigen and a homologous peptide of Coxsackie virus 2B protein, *Clinical and Experimental Immunology*, **141**(1), pp. 148–154.

Stauffer, Y., Marguerat, S., Meylan, F., Ucla, C., Sutkowski, N., Huber, B., Pelet, T. and Conrad, B. (2001) Interferon-alpha-induced endogenous superantigen. a model linking environment and autoimmunity, *Immunity*, **15**(4), pp. 591–601.

Stave, J. W. and Lindpaintner, K. (2013) Antibody and antigen contact residues define epitope and paratope size and structure, *Journal of immunology* **191**(3), pp. 1428–1435.

Steed, A. L. and Stappenbeck, T. (2014) Role of Viruses and Bacteria-Virus Interactions In Autoimmunity, *Current Opinion in Immunology*, **31**, pp. 102–107.

Van Steendam, K., Tilleman, K., De Ceuleneer, M., De Keyser, F., Elewaut, D. and Deforce, D. (2010) Citrullinated vimentin as an important antigen in immune complexes from synovial fluid of rheumatoid arthritis patients with antibodies against citrullinated proteins, *Arthritis Research and Therapy*, **12**(4), pp. R132-141.

Steiner, G. (2006) Autoantibodies in rheumatoid arthritis, *Clinical Laboratory International*, **7**, pp. 8–10.

Stransky, G., Vernon, J., Aicher, W. K., Moreland, L. W., Gay, R. E. and Gay, S. (1993) Virus-like particles in synovial fluids from patients with rheumatoid arthritis, *British journal of rheumatology*, **32**(12), pp. 1044–1048.

Sudoł-Szopińska, I., Jans, L. and Teh, J. (2017) Rheumatoid arthritis: what do MRI and ultrasound show, *Journal of ultrasonography*, **17**(68), pp. 5–16.

Takeuchi, K., Katsumata, K., Ikeda, H., Minami, M., Wakisaka, A. and Yoshiki, T. (1995a) Expression of endogenous retroviruses, ERV3 and lambda 4-1, in synovial tissues from patients with rheumatoid arthritis, *Clinical and experimental immunology*, **99**(3), pp. 338–344.

Tan, E. M. and Smolen, J. S. (2016) Historical observations contributing insights on etiopathogenesis of rheumatoid arthritis and role of rheumatoid factor, *Journal of Experimental Medicine*, **213**(10), pp. 1937-1950.

Tarlinton, R. E., Meers, J. and Young, P. R. (2006) Retroviral invasion of the koala genome, *Nature*, **442**(7098), pp. 79–81.

Tenembaum, S., Chitnis, T., Ness, J., Hahn, J. S. and International Pediatric MS Study Group (2007) Acute disseminated encephalomyelitis, *Neurology*, **68**(Issue 16, Supplement 2), pp. S23–S36.

Terrier, B., Degand, N., Guilpain, P., Servettaz, A., Guillevin, L. and Mouthon, L. (2007) Alpha-enolase: A target of antibodies in infectious and autoimmune diseases, *Autoimmunity Reviews*, **6**(3), pp. 176–182.

Tighe, P. J., Stevens, S. E., Dempsey, S., Le Deist, F., Rieux-Laucat, F. and Edgar, J. D. M. (2002) Inactivation of the Fas gene by Alu insertion: retrotransposition in an intron causing splicing variation and autoimmune lymphoproliferative syndrome, *Genes and Immunity*, **3**(S1), pp. S66–S70.

Ting, C. N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C. and Meisler, M. H. (1992) Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene, *Genes and development*, **6**(8), pp. 1457–1465.

Tobón, G. J., Izquierdo, J. H. and Cañas, C. A. (2013) B lymphocytes: development, tolerance, and their role in autoimmunity-focus on systemic lupus erythematosus, *Autoimmune diseases*, **2013**, p. 827254. doi: 10.1155/2013/827254.

Trela, M., Nelson, P. N. and Rylance, P. B. (2016) The role of molecular mimicry and other factors in the association of Human Endogenous Retroviruses and autoimmunity, *Apmis*, **124**(1–2), pp. 88–104.

Trier, N. H., Dam, C. E., Olsen, D. T., Hansen, P. R. and Houen, G. (2015) Contribution of Peptide Backbone to Anti-Citrullinated Peptide Antibody Reactivity, *PloS one*, **10**(12), p. e0144707. doi: 10.1371/journal.pone.0144707

Trilling, A. K., Beekwilder, J. and Zuilhof, H. (2013) Antibody orientation on biosensor surfaces: a minireview, *The Analyst*, **138**(6), pp. 1619-1627.

Trouw, L. A., Huizinga, T. W. J. and Toes, R. E. M. (2013) Autoimmunity in rheumatoid arthritis: different antigens—common principles, *Annals of the Rheumatic Diseases*, **72**(suppl 2), pp. ii132-ii136.

Tsubata, T. (2017) B-cell tolerance and autoimmunity, *F1000 Research*, **6**, pp. 391-399.

Tsuchiya, M., Nakao, H., Katoh, T., Sasaki, H., Hiroshima, M., Tanaka, T., Matsunaga, T., Hanaoka, T., Tsugane, S. and Ikenoue, T. (2005) Association between endometriosis and genetic polymorphisms of the estradiol-synthesizing enzyme genes HSD17B1 and CYP19, *Human Reproduction*, **20**(4), pp. 974–978.

Tsuda, R., Ozawa, T., Kobayashi, E., Hamana, H., Taki, H., Tobe, K., Sugiyama, E.,

Iwamoto, M., Imura, J., Kishi, H. and Muraguchi, A. (2015) Monoclonal Antibody Against Citrullinated Peptides Obtained From Rheumatoid Arthritis Patients Reacts With Numerous Citrullinated Microbial and Food Proteins, *Arthritis and Rheumatology*, **67**(8), pp. 2020–2031.

Tűdős, E., Cserző, M. and Simon, I. (1990) Predicting isomorphic residue replacements for protein design, *International Journal of Peptide and Protein Research*, **36**(3), pp. 236–239.

Tugnet, N., Rylance, P., Roden, D., Trela, M. and Nelson, P. (2013) Human Endogenous Retroviruses (HERVs) and Autoimmune RheumaticDisease: Is There a Link?, *The Open Rheumatology Journal*, **7**(1), pp. 13–21.

Turcanova, V. L., Bundgaard, B. and Höllsberg, P. (2009) Human herpesvirus-6B induces expression of the human endogenous retrovirus K18-encoded superantigen, *Journal of Clinical Virology*, **46**(1), pp. 15–19.

Turunen, S., Hannonen, P., Koivula, M.-K., Risteli, L. and Risteli, J. (2015) Separate and overlapping specificities in rheumatoid arthritis antibodies binding to citrulline- and homocitrulline-containing peptides related to type I and II collagen telopeptides, *Arthritis Research and Therapy*, **17**(1), pp. 2-8.

Udby, L., Johnsen, A. H. and Borregaard, N. (2010) Human CRISP-3 binds serum α1B-glycoprotein across species, *Biochimica et Biophysica Acta (BBA),* **1800**(4), pp. 481–485.

Umićević Mirkov, M. and Coenen, M. J. (2013) Pharmacogenetics of disease-modifying antirheumatic drugs in rheumatoid arthritis: towards personalized medicine, *Pharmacogenomics*, **14**(4), pp. 425–444.

Union, A., Meheus, L., Humbel, R. L., Conrad, K., Steiner, G., Moereels, H., Pottel, H., Serre, G. and De Keyser, F. (2002) Identification of citrullinated rheumatoid arthritis-specific epitopes in natural filaggrin relevant for antifilaggrin autoantibody detection by line immunoassay, *Arthritis and Rheumatism*, **46**(5), pp. 1185–1195.

Ursum, J., Bos, W. H., van Dillen, N., Dijkmans, B. A. and van Schaardenburg, D. (2010) Levels of anti-citrullinated protein antibodies and IgM rheumatoid factor are not associated with outcome in early arthritis patients: a cohort study, *Arthritis research and therapy*, **12**(1), pp. R8-14.

Valesini, G., Gerardi, M. C., Iannuccelli, C., Pacucci, V. A., Pendolino, M. and Shoenfeld, Y. (2015) Citrullination and autoimmunity, *Autoimmunity Reviews*, **14**(6), pp. 490–497.

Vargiu, L., Rodriguez-Tomé, P., Sperber, G. O., Cadeddu, M., Grandi, N., Blikstad, V., Tramontano, E. and Blomberg, J. (2016) Classification and characterization of human endogenous retroviruses; mosaic forms are common, *Retrovirology*, **13**(1), pp. 7-35.

Venables, P. and Maini, R. (2017) Clinical manifestations of rheumatoid arthritis, *UpToDate*, [online] Available from: https://www.uptodate.com/contents/clinical-manifestations-of-rheumatoid-arthritis (Accessed 8 February 2018).

van Venrooij, W. J. and Pruijn, G. J. (2014) How citrullination invaded rheumatoid arthritis research, *Arthritis Research and Therapy*, **16**(1), pp. 103-107.

Viatte, S., Plant, D. and Raychaudhuri, S. (2013) Genetics and epigenetics of rheumatoid arthritis, *Nature reviews. Rheumatology*, **9**(3), pp. 141–153.

Vihinen, M., Torkkila, E. and Riikonen, P. (1994) Accuracy of protein flexibility predictions, *Proteins: Structure, Function, and Genetics*, **19**(2), pp. 141–149.

Vincendeau, M., Göttesdorfer, I., Schreml, J. M. H., Wetie, A. G. N., Mayer, J., Greenwood, A. D., Helfer, M., Kramer, S., Seifarth, W., Hadian, K., Brack-Werner, R. and Leib-Mösch, C. (2015) Modulation of human endogenous retrovirus (HERV) transcription during persistent and de novo HIV-1 infection, *Retrovirology*, **12**(1), pp. 27-43.

Vojdani, A. (2014) A Potential Link between Environmental Triggers and Autoimmunity, *Autoimmune Diseases*, **2014**, pp. 1–18. doi.org/10.1155/2014/437231

Voskuhl, R. (2011) Sex differences in autoimmune diseases., *Biology of sex differences*, BioMed Central, **2**(1), pp. 1-21.

Waldron, E. E., Murray, P., Kolar, Z., Young, L., Brown, C., Reynolds, G., Baumforth, K., Toomey, S., Astley, S. J., Perera, S. A. and Nelson, P. N. (2002) Reactivity and Isotype Profiling of Monoclonal Antibodies using Multiple Antigenic Peptides, *Hybridoma and Hybridomics*, **21**(5), pp. 393–398.

Walker-Cillo, G. (2010) Myths and misconceptions about rheumatoid arthritis, *American Nurse Today*, **5**(7), [online] Available from:

https://www.americannursetoday.com/myths-and-misconceptions-about-rheumatoid-arthritis/ (Accessed 12 January 2019).

Wallace, T. A., Downey, R. F., Seufert, C. J., Schetter, A., Dorsey, T. H., Johnson, C. A., Goldman, R., Loffredo, C. A., Yan, P., Sullivan, F. J., Giles, F. J., Wang-Johanning, F., Ambs, S. and Glynn, S. A. (2014) Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers, *Carcinogenesis*, **35**(9), pp. 2074–2083.

Wang-Johanning, F., Li, M., Esteva, F. J., Hess, K. R., Yin, B., Rycaj, K., Plummer, J. B., Garza, J. G., Ambs, S. and Johanning, G. L. (2014) Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer, *International Journal of Cancer*, **134**(3), pp. 587–595.

Wang-Lin, S., Balthasar, J., Wang-Lin, S. X. and Balthasar, J. P. (2018) Pharmacokinetic and Pharmacodynamic Considerations for the Use of Monoclonal Antibodies in the Treatment of Bacterial Infections, *Antibodies*, **7**(1), pp. 5-24.

Wang, X., Chen, P., Cui, J., Yang, C. and Du, H. (2015) Keratin 8 is a novel autoantigen of rheumatoid arthritis, *Biochemical and Biophysical Research Communications*, **465**(4), pp. 665–669.

Warnecke, A., Sandalova, T., Achour, A. and Harris, R. A. (2014) PyTMs: a useful PyMOL plugin for modeling common post-translational modifications, *BMC Bioinformatics*, **15**(1), pp. 370-381.

Wegner, N., Lundberg, K., Kinloch, A., Fisher, B., Malmström, V., Feldmann, M. and Venables, P. J. (2010) Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis, *Immunological Reviews*, **233**(1), pp. 34–54.

Wei, L., Chahwan, R., Wang, S., Wang, X., Pham, P. T., Goodman, M. F., Bergman, A., Scharff, M. D. and MacCarthy, T. (2015) Overlapping hotspots in CDRs are critical sites for V region diversification, *Proceedings of the National Academy of Sciences of the United States of America*, **112**(7), pp. E728-37.

Weiss, R. A. (2016) Human endogenous retroviruses: friend or foe?, *APMIS*, **124**(1–2), pp. 4–10.

Welch, N. G., Scoble, J. A., Muir, B. W. and Pigram, P. J. (2017) Orientation and characterization of immobilized antibodies for improved immunoassays, *Biointerphases*, **12**(2), p. 02D301. doi: 10.1116/1.4978435.

van der Werf, N., Kroese, F. G. M., Rozing, J. and Hillebrands, J.-L. (2007) Viral infections as potential triggers of type 1 diabetes, *Diabetes/Metabolism Research and Reviews*, **23**(3), pp. 169–183.

Westergaard, M. W., Draborg, A. H., Troelsen, L., Jacobsen, S. and Houen, G. (2015) Isotypes of Epstein-Barr Virus Antibodies in Rheumatoid Arthritis: Association with Rheumatoid Factors and Citrulline-Dependent Antibodies, *BioMed Research International*, **2015**, pp. 1–9. doi.org/10.1155/2015/472174

Westwood, O. M. R. and Hay, F. C. (2001) *Epitope mapping : a practical approach*, Oxford: Oxford University Press, pp. 197–204.

Westwood, O. M. R., Nelson, P. N., Westwood, O. M. R., Freimanis, G., Roden, D., Sissaoui, S., Rylance, P. and Hay, F. C. (2008) Comparison of Antigenic Regions Identifi ed on IgG1Fc Using Bioinformatics vs Pepscan Analysis, *Clinical Medicine: Arthritis and Musculoskeletal Disorders*, **1**, pp. 33–42.

Whittingham, S., Stockman, A. and Rowley, M. (2017) Collagen Autoantibodies and Their Relationship to CCP Antibodies and Rheumatoid Factor in the Progression of Early Rheumatoid Arthritis, *Antibodies*, **6**(2), pp. 6-21.

Winn, V. D., Gormley, M., Paquet, A. C., Kjaer-Sorensen, K., Kramer, A., Rumer, K. K., Haimov-Kochman, R., Yeh, R.-F., Overgaard, M. T., Varki, A., Oxvig, C. and Fisher, S. J. (2009) Severe Preeclampsia-Related Changes in Gene Expression at the Maternal-Fetal Interface Include Sialic Acid-Binding Immunoglobulin-Like Lectin-6 and Pappalysin-2, *Endocrinology*, **150**(1), pp. 452–462.

Witalison, E. E., Thompson, P. R. and Hofseth, L. J. (2015) Protein Arginine Deiminases and Associated Citrullination: Physiological Functions and Diseases Associated with Dysregulation, *Current drug targets*, **16**(7), pp. 700–710.

Woo, W. H., Shen, L., Leong, S. M. and Koay, E. S.-C. (2014) Prevalence of human endogenous retroviral element associates with Hodgkin's lymphoma incidence rates, *Leukemia research reports*, **3**(1), pp. 1–3.

Wu, Z., Mei, X., Zhao, D., Sun, Y., Song, J., Pan, W. and Shi, W. (2015) DNA methylation modulates HERV-E expression in CD4+ T cells from systemic lupus erythematosus patients, *Journal of Dermatological Science*, **77**, pp. 110–116.

Wucherpfennig, K. W. and Strominger, J. L. (1995) Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein, *Cell*, **80**(5), pp. 695–705.

Xing, Y. and Hogquist, K. A. (2012) T-cell tolerance: central and peripheral, *Cold Spring Harbor perspectives in biology*, **4**(6), p. a006957. doi: 10.1101/cshperspect.a006957

Yi, J. K., Chang, J. W., Han, W., Lee, J. W., Ko, E., Kim, D. H., Bae, J.-Y., Yu, J., Lee, C., Yu, M.-H. and Noh, D.-Y. (2009) Autoantibody to Tumor Antigen, Alpha 2-HS Glycoprotein: A Novel Biomarker of Breast Cancer Screening and Diagnosis, *Cancer Epidemiology Biomarkers & Prevention*, **18**(5), pp. 1357–1364.

Young, G. R., Stoye, J. P. and Kassiotis, G. (2013) Are human endogenous retroviruses pathogenic? An approach to testing the hypothesis, *BioEssays*, **35**(9), pp. 794–803.

Yu-Bin Dong, *, Jun-Yan Cheng, and, Huang, R.-Q., and, M. D. S. and Loye, H.-C. zur (2003) Self-Assembly of Coordination Polymers from AgX (X = SbF6-, PF6-, and CF3SO3-) and Oxadiazole-Containing Ligands, *Inorganic Chemistry*, **42**(18), pp. 5699–5706.

Yu, H.-L., Zhao, Z.-K. and Zhu, F. (2013) The role of human endogenous retroviral long terminal repeat sequences in human cancer, *International Journal of Molecular Medicine*, **32**(4), pp. 755–762.

Zampeli, E., Vlachoyiannopoulos, P. G. and Tzioufas, A. G. (2015) Treatment of rheumatoid arthritis: Unraveling the conundrum, *Journal of Autoimmunity*, **65**, pp. 1–18.

Zeilinger, S., Kühnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., Weidinger, S., Lattka, E., Adamski, J., Peters, A., Strauch, K., Waldenberger, M. and Illig, T. (2013) Tobacco Smoking Leads to Extensive Genome-Wide Changes in DNA Methylation, *PLoS ONE*, **8**(5), p. e63812. doi.org/10.1371/journal.pone.0063812

Zheng, J., Lin, X., Wang, X., Zheng, L., Lan, S., Jin, S., Ou, Z. and Wu, J. (2017) In Silico Analysis of Epitope-Based Vaccine Candidates against Hepatitis B Virus Polymerase Protein, *Viruses*, **9**(5), pp. 112-129.

Zheng, W., Ruan, J., Hu, G., Wang, K., Hanlon, M. and Gao, J. (2015) Analysis of Conformational B-Cell Epitopes in the Antibody-Antigen Complex Using the Depth Function and the Convex Hull, *PLOS ONE*, **10**(8), p. e0134835. doi.org/10.1371/journal.pone.0134835

Ziegler, B., Gay, R. E., Huang, G. Q., Fassbender, H. G. and Gay, S. (1989) Immunohistochemical localization of HTLV-I p19- and p24-related antigens in synovial joints of patients with rheumatoid arthritis, *The American journal of pathology*, **135**(1), pp. 1–5.

Ziegler, B. L. and Thomas, C. A. (1991) Detection of HIV-I related antigens in rheumatoid synovial tissues, *Acta Histochemica et Cytochemica*, **24**(1), pp. 1-9.