

**THE UNEXPECTED LOCALIZATION OF PAD4 IN MONOCYTES: IMPLICATIONS
FOR RHEUMATOID ARTHRITIS PATHOGENESIS**

by

Mekha Ann Thomas

A dissertation submitted to Johns Hopkins University in conformity with the requirement
for the degree of Doctor of Philosophy

Baltimore, Maryland

December 2023

© 2023 Mekha Ann Thomas

All rights reserved

ABSTRACT

Although anti-citrullinated protein autoantibodies (ACPAs) are a hallmark serological feature of rheumatoid arthritis (RA), the mechanisms and cellular sources behind the generation of the RA citrullinome remain incompletely defined. Peptidylarginine deiminase IV (PAD4), one of the key enzymatic drivers of citrullination in the RA joint, is expressed by granulocytes and monocytes; however, the subcellular localization and contribution of monocyte derived PAD4 to the generation of citrullinated autoantigens remain underexplored. In this study, we demonstrate that PAD4 displays a widespread cellular distribution in monocytes, including expression on the cell surface. Surface PAD4 was enzymatically active and capable of citrullinating extracellular fibrinogen and endogenous surface proteins in a calcium dose-dependent manner, and fibrinogen citrullinated by monocyte-surface PAD4 could be recognized by ACPAs. Several novel PAD4 substrates were identified on the monocyte surface via mass spectrometry, with citrullination of the CD11b and CD18 components of the Mac-1 integrin complex being the most abundant. Citrullinated Mac-1 was found to be a target of ACPAs in 25% of RA patients, and Mac-1 ACPAs were significantly associated with HLA-DRB1 shared epitope alleles, higher C-reactive protein and IL-6 levels, and more erosive joint damage. Our findings implicate the monocyte cell surface as a previously undescribed and consequential site of extracellular and cell surface autoantigen generation in RA.

The mechanisms by which PAD4, which lacks conventional secretory signal sequences, traffics to extranuclear localizations are unknown. In this study, we also show

that PAD4 was enriched in the organelle fraction of monocytes with evidence of citrullination of organelle proteins. We demonstrated that PAD4 can bind to several cytosolic, nuclear, and organelle proteins that may serve as binding partners for PAD4 to traffic intracellularly. Additionally, cell surface expression of PAD4 increased with monocyte differentiation into monocyte derived-dendritic cells and co-localized with several endocytic/autophagic and conventional secretory pathway markers, implicating the use of these pathways by PAD4 to traffic within the cell. Our results suggest that PAD4 is expressed in multiple subcellular localizations and may play previously unappreciated roles in physiologic and pathologic conditions.

Material presented is adapted from: Thomas MA, Kim SY, Curran AM, Smith B, Antiochos B, Na CH, Darrah E. An unbiased proteomic analysis of PAD4 in human monocytes: novel substrates, binding partners and subcellular localizations. Philos Trans R Soc Lond B Biol Sci. 2023 Nov 20;378(1890):20220477. doi: 10.1098/rstb.2022.0477. Epub 2023 Oct 2. PMID: 37778379; PMCID: PMC10542449.

Primary Reader and Advisor: Erika L. Darrah, Ph.D.

Secondary Reader: Alan L. Scott, Ph.D.

ACKNOWLEDGEMENTS

This dissertation is dedicated to everyone who has contributed, big and small, to the progress and completion of my Ph.D. While it has been a struggle, without the help of the following people, this work would not be a reality.

To my advisor, Erika Darrah: I have been so lucky to have called you my mentor. You have gone above and beyond for me whether I deserved it or not. Thank you for taking a chance on me and being my biggest advocate. I have not met another who has been so vocal and dedicated to student success and well-being, whether for her own mentees or for those whose thesis committees and qualifying exams she participated in. She has been my advisor, my cheerleader, an amazing listener, and a constant advocate who has always had my back.

Hong Wang is a wizard whose fingers work magic and somehow every experiment she touches, works. She has been so important during my beginning years in lab as a guide. Thank you, Hong, for answering my constant barrage of questions and for being so patient while I learned from you. Eleni Tiniakou has been an amazing source of moral support. She has encouraged my dabbling in science communication and creation of scientific figures and has been an amazing advocate for encouraging me to pursue my passion. Brendan Antiochos has been a role model for me. He has taught me how to be a great scientist, a great physician (even though I am not one), and an amazing person all at once. Thank you so much for being honest and true and for never discouraging me from trying new things. It has been a lot of fun discussing a ton of random topics with you.

Life in the lab would have been so monotonous without all my lab mates. Ashley Curran has been my single greatest supporter and an amazing friend. Without our banter, late night shenanigans, and all that coffee, this Ph.D. would have been dull indeed. Jonathan Crawford has been a constant source of encouragement and a reservoir of the most brilliant ideas. Kusuma Ananth has been an amazing listener with a penchant for injecting some much-needed energy around the lab. And Alexander Girgis, who always knows just the right thing to say, has been instrumental in more ways than one and much more than he realizes.

Thank you to the rest of the Rosen Lab – Livia Casciola-Rosen, Qingyuan Yang, Laura Gutierrez, David Hines, and others – it has been a pleasure working with you. The Division of Rheumatology, especially Shannon Bishop, Ami Shah, and Antony Rosen, has been especially encouraging and supportive during my graduate studies.

To my thesis committee – Alan Scott, Carolyn Machamer, Nicola Heller, and Sergi Regot – thank you for believing in me and in my work. Without you, this dissertation would not be as comprehensive as it is now. I thank the Graduate Program in Immunology administration – Joel Pomerantz, Angela James, and Lori Fountain – for supporting me financially and administratively during my doctoral studies. And a big thank you to my class – Marisa Mitchell-Flack, Shenda Hou, Xiaoxou Wang, Jesus Contreras-Rodriguez, and Joshua Taylor – for all the study sessions and support throughout the years.

The work in this dissertation would not have happened if not for the generous donation of peripheral blood by healthy controls and RA patients. I would also like to thank

Clifton Bingham, Michelle Jones, and Marilyn Towns from the Johns Hopkins Arthritis Center for helping bank patient samples. The various weekly leukopaks obtained by John Schroeder have been instrumental to many of the findings in this work. All mass spectrometry data was obtained by Seok-Young Kim and Chan Hyun Na. Pooja Naik and Hong Wang provided experimental support while Tory Johnson, Ashley Curran, and Jonathan Crawford were key to providing key insights. Alexander Girgis generated the CD11b and CD18 protein structure predictions and Jon Giles provided statistical analysis of patient data. Monoclonal antibodies against citrullinated proteins were generously provided by Shaghayegh Jahanbani and William Robinson while Brendan Antiochos and Barbara Smith were instrumental in providing the confocal and electron microscopy images.

Lastly, this dissertation would not be complete without my family and friends. My foray into research would not have begun if not for Gregory Szeto who first saw my passion and potential during my undergraduate studies. And it would not have continued if not for Ying Zhang, to whom I owe so much. To my extended family, and to my parents, Anita Thomas and Thomas Chandy, thank you for your support while I pursued my doctoral studies. My brother, Melvin Thomas, and my cousins, Kevin and Allen Thomas, have been unwavering in their encouragements. Late-night sessions in lab would not be complete without a call to my loving aunt, Ajita Thomas, who was always up for a midnight chat. My late paternal grandparents, T C Chandy and Annamma Chandy, I thank for their unconditional love, and my late uncle, Anil Alexander, for spoiling me so much when you were still with us. And most importantly, this dissertation is dedicated to my late maternal

grandparents, C M Alexander and Ammini Alexander. Even though you are not here to see the final product of my work, I know that both of you would be so proud of the accomplishments of your granddaughter.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
Chapter 1 – General introduction	1
1.1 Rheumatoid arthritis overview	2
1.2 Risk factors of RA	4
1.3 RA pathophysiology.....	6
1.4 RA serology.....	8
1.5 Citrullination and PADs in RA.....	10
1.6 Monocytes in RA	14
1.7 Thesis overview.....	16
Chapter 2 – The monocyte cell surface is a novel site of autoantigen generation in rheumatoid arthritis.....	17
2.1 Introduction.....	18
2.2 Methods.....	20
2.3 Results	39
2.3.1 <i>PAD4 exists on the surface of monocytes.....</i>	<i>39</i>

2.3.2	<i>Surface PAD4 is enzymatically active</i>	46
2.3.3	<i>Surface PAD4 can citrullinate monocyte cell surface proteins</i>	51
2.3.4	<i>Mass spectrometry reveals novel citrullinated cell surface proteins including integrins</i>	56
2.3.5	<i>Citrullinated Mac-1 is a novel cell surface autoantigen in RA</i>	69
2.3.6	<i>Presence of anti-citrullinated Mac-1 antibodies is associated with worse inflammation</i>	72
2.4	Discussion	83
Chapter 3 – An unbiased proteomic analysis of PAD4 in human monocytes: novel substrates, binding partners, and subcellular localizations		89
3.1	Introduction.....	90
3.2	Methods.....	92
3.3	Results	102
3.3.1	<i>PAD4 is present in vesicular organelles in monocytes</i>	102
3.3.2	<i>PAD substrates are enriched in vesicular organelles and processes</i>	105
3.3.3	<i>PAD4 binds to protein partners found at various organelle compartments</i>	109
3.3.4	<i>Monocyte-derived dendritic cells serve as a model to study cell surface PAD4 trafficking</i>	115
3.4	Discussion	123
Chapter 4 – General conclusions		127
References		133
Curriculum vitae		160

LIST OF TABLES

Table 1	Biological processes enriched in citrullinated plasma membrane proteins identified via mass spectrometry	61
Table 2	Molecular functions enriched in citrullinated plasma membrane proteins identified via mass spectrometry	62
Table 3	Protein class enriched in citrullinated plasma membrane proteins identified via mass spectrometry	63
Table 4	Pathways enriched in citrullinated plasma membrane proteins identified via mass spectrometry	64
Table 5	Top hits from list of citrullinated plasma membrane proteins – mass spectrometry	67
Table 6	Citrullinated peptides in CD18 and CD11b	69
Table 7	Disease associations in ESCAPE RA cohort based on Mac-1 ACPA status.....	74
Table 8	Disease associations in ESCAPE RA cohort based on antibody subset status.....	78
Table 9	Multivariate analysis of disease characteristics in the RA ESCAPE cohort.....	82

LIST OF FIGURES

2 - The monocyte cell surface is a novel site of autoantigen generation in rheumatoid arthritis

2-1	Specificities of anti-PAD4 antibodies.....	23
2-2	Monocytes express PAD4 extranuclearly	40
2-3	Monocytes express PAD4 on their surface.....	42
2-4	Flow cytometry gating strategy.....	44
2-5	Flow cytometry gating strategy for surface PAD4 levels on healthy controls and RA patients.....	45
2-6	PAD4 on the monocyte surface is catalytically active and its ability to citrullinate is comparable to neutrophils.....	47
2-7	Monocyte-citrullinated fibrinogen is recognized by patient autoantibodies	50
2-8	Surface PAD4 citrullinates endogenous surface proteins on monocytes.....	53
2-9	Flow cytometry gating strategy for surface citrullinated protein expression.....	55
2-10	Mass spectrometry identifies several citrullinated cell surface proteins	58
2-11	Statistical enrichment of citrullinated proteins in various protein classifications	61
2-12	Mass spectrometry reveals several citrullination sites on CD11b and CD18.....	66
2-13	CD11b is a novel cell surface autoantigen.....	70
2-14	Disease associations in ESCAPE RA cohort.....	76

3 - An unbiased proteomic analysis of PAD4 in human monocytes: novel substrates, binding partners, and subcellular localizations

3-1	PAD4 is found in the organelle fraction in monocytes	103
3-2	PAD4 is detected in organelles of monocytes	105
3-3	PADs in monocytes.....	106
3-4	Citrullinated proteins are enriched in organelles of monocytes	108
3-5	PAD4 binds protein partners found throughout the cell	112
3-6	PAD4 binding partners.....	114
3-7	PAD4 localization in mo-DCs.....	117
3-8	PAD4 associates with autophagic and exocytic organelles in mo-DCs.....	119
3-9	Larger field of view depicting colocalization of PAD4 with autophagic and exocytic organelles in mo-DCs.....	120
3-10	PAD4 and organelle colocalization in permeabilized mo-DCs	121

4 – General conclusions

4-1	PAD4 cell surface localization contributes to RA pathogenesis.....	130
-----	--	-----

CHAPTER 1

General introduction

Material presented in this chapter is adapted from: Cappelli LC, Thomas MA, Bingham CO 3rd, Shah AA, Darrach E. Immune checkpoint inhibitor-induced inflammatory arthritis as a model of autoimmune arthritis. Immunol Rev. 2020 Mar;294(1):106-123. doi: 10.1111/imr.12832. Epub 2020 Jan 13. PMID: 31930524; PMCID: PMC7047521.

1.1 Rheumatoid arthritis.

Epidemiology. RA is a systemic autoimmune disease most notable for chronic inflammatory arthritis that can lead to erosive joint disease and permanent joint damage. Though joints are most predominantly affected, patients may also experience systemic symptoms and extra-articular disease. The prevalence of RA varies depending on country and study but is estimated to affect 0.5–1% across different populations worldwide.¹ Women are affected more often than men, making up about two-thirds of the population with RA.² Family history is a risk factor for developing RA, with heritability about 40% higher in seropositive patients than in seronegative patients.³ Age of onset can vary, but peaks around 50 years of age in studies from North American and European countries.⁴ A variety of environmental risk factors for the development of RA have been identified including exposure to tobacco smoke, silica dust, air pollution, obesity, and low vitamin D levels.³

Clinical features. Classically, RA is described as symmetric polyarthritis predominantly affecting the hands and wrists. Patients may present differently in terms of number and specific joints involved. RA can affect almost any peripheral joint, while axial disease is typically limited to the cervical spine, particularly the C1–2 articulation. Of the peripheral joints, the most commonly affected include the small joints of the hands and wrists. Patients may experience pain, swelling, or stiffness of their joints. Longstanding or undertreated disease can lead to structural deformities such as flexion contractures in the elbows, ulnar deviation of the phalanges at the metacarpal phalangeal (MCP) joints, and swan neck or boutonniere deformities of the fingers. Tenosynovitis, or inflammation of the

tendon sheath, has become increasingly recognized in early RA using ultrasound and magnetic resonance imaging (MRI).⁵ Evaluation with plain radiographs can show joint space narrowing, erosions, and subchondral cysts, but MRI and ultrasound are more sensitive in detecting erosions and can define synovitis which is not apparent on plain radiography. Ultrasound features of RA include synovial hypertrophy, power Doppler signal in the synovium indicating synovitis, joint effusions, erosions, and tenosynovitis. Before the availability of biologics and the treat to target approach, bone erosions developed in many patients in the first two years of disease.⁶

A variety of extra-articular features are seen in people with RA. Interstitial lung disease (ILD), rheumatoid nodules, rheumatoid vasculitis, and scleritis are some of the conditions that can accompany joint disease. ILD varies in severity and is clinically apparent in 6 to 30% of RA patients depending on how it is diagnosed.⁷ However, asymptomatic subclinical ILD has been reported in up to 60% of patients. Patients with RA also have an increased risk of lymphoma, lung cancer, and cardiovascular disease.^{8,9}

Disease course and prognosis. The advent of more aggressive management with conventional synthetic disease-modifying anti-rheumatic drugs (DMARDs), biological DMARDs, and targeted synthetic DMARDs combined with earlier disease recognition, initiation of definitive DMARD therapy, and treat to target approaches have drastically changed the prognosis of RA. Disability occurred early in older studies of patients with RA. In a study of two cohorts of RA patients from the 1980s and 1990s, 20–30% of patients became unable to work within three to four years of diagnosis.¹⁰ Disease activity levels and disability have fallen in the last 15 years with more intensive treatment.^{11,12} Despite the

advances in therapy, RA is a chronic disease where most patients need some form of DMARD therapy indefinitely.

1.2 Risk factors of RA.

Susceptibility to RA can be driven by a combination of genetic, epigenetic, and environmental factors. Genetic association studies have found a strong correlation between developing RA and individuals harboring HLA-DR alleles that share a common motif, known as shared epitope (SE) alleles.¹³ The motif consists of the region spanning residues 70-74 of the beta chain and has been found to contain the following amino acid sequences: QKRAA, QRRAA, and RRRAA.^{14,15} While SE alleles include *HLA-DRB1*0401*, **0404*, **0405*, **0408*, **0101*, **0102*, **1402*, and **1001* among others,¹⁴ significant associations with the risk of RA development have been found with *HLA-DRB1*01* and *HLA-DRB1*04* alleles.^{13,16} The presence of SE alleles has also been linked to more severe disease and the development of autoantibodies called anti-citrullinated protein antibodies (ACPAs) in patients (see section 1.4 'RA serology' for more about ACPAs).^{17,18}

Other risk loci include mainly those that are associated with inflammatory mechanisms. After *HLA-DRB1* alleles, the second strongest association with the risk of developing RA has been shown with protein tyrosine phosphatase non-receptor type 22 (*PTPN22*). While also associated with the risk of developing other autoimmune diseases, polymorphisms in the *PTPN22* gene has been shown to correlate with seropositive RA.^{19–}²² *PADI4*, the gene that codes for the enzyme peptidylarginine deiminase IV (PAD4) which catalyzes the production of citrullinated proteins, has been shown to associate with RA

development in Asian populations (more about PAD4 can be found in Section 1.4).^{23,24} TNF receptor associated factor 1 (*TRAF1*) is known to be associated with ACPA+ RA and have been implicated with radiological progression.^{22,25} *STAT4* is known to be associated with other autoimmune diseases and is implicated to confer susceptibility to both ACPA+ and ACPA- RA.^{26,27} A V/V polymorphism in the Fc gamma receptor (*FCGR*) gene was found to correlate with the presence of ACPAs while a SNP at the *CD40* locus was shown to associate with increased joint damage in ACPA+ RA patients.^{28,29} Genomic studies have also identified SNPs near *IL6ST*, *SPRED2*, *RBPJ*, *CCR6*, *IRF5*, and *PXK* to be important correlators with RA disease development.³⁰ In addition, epigenetic modifications such as DNA methylation and histone acetylation may act synergistically with environmental factors to contribute to RA development.^{31,32}

Smoking is known to be associated with seropositive RA and is thought to account for about 20-30% of risk for RA.³³ Current smoking status correlates with increased pro-inflammatory cytokines such as IL-2, IL-6, IFN- γ , and TNF- α as well as with higher disease activity.³⁴ In addition to associations with periodontal and lung diseases, smoking is also thought to cause systemic effects that could lead to the development of RA.^{35,36} Synergistic associations of smoking and SE alleles along with the presence of citrullinated proteins in lungs of smokers has led to the development of the theory that smoking causes increased citrullination which leads to the presentation of citrullinated proteins and to the eventual generation of ACPAs.³⁷ Along with high body mass index, smoking could act to increase risk of developing RA.³⁸

Women are two to three times more likely to develop RA and two-thirds of RA patients are women.^{22,39} While it is hypothesized that there may be possible female-specific factors, such as estrogen, that confer risk of developing RA, more studies are still needed to fully understand the factors that lead to higher disease incident rates in women.⁴⁰ Microbiota, especially those in the oral cavity and can cause periodontitis, have also been found to correlate with RA development.⁴¹⁻⁴³ It is hypothesized that breaks in mucosal barriers could kickstart inflammation and, along with other genetic and environmental risk factors, could lead to RA through molecular mimicry or autoimmunity against self-antigens.⁴⁴ A more thorough review of genetic and environmental risk factors of RA can be found in Kurko *et. al.* 2013 and Deane *et. al.* 2017.^{2,22}

1.3 RA pathophysiology.

Neutrophil infiltration is a key characteristic of inflammation seen in RA, with neutrophils comprising the majority of the innate immune cells found in the synovial fluid of RA patients.⁴⁵ Neutrophils are important producers of TNF- α , a central cytokine in RA pathogenesis and target of therapy for the treatment of RA. A recent study showed that treatment of RA patients with anti-TNF- α antibodies reduced TNF- α production by neutrophils to a level similar to that of healthy controls.⁴⁶ In addition to inflicting damage at sites of inflammation through proteases and secreting cytokines which recruit additional immune cells, they also release neutrophil extracellular traps (NETs) comprised of intracellular proteins bound to DNA.⁴⁵ This process may redistribute autoantigens, such as citrullinated proteins, to the extracellular environment where they may be bound by

autoantibodies and drive downstream inflammatory processes such as macrophage or complement activation.^{47,48} These other innate effector mechanisms and cells, including dendritic and natural killer (NK) cells, are critical to RA pathogenesis as they sustain inflammation and contribute to adaptive immune cell activation.⁴⁹

In addition to the clear role for the innate immune system in RA pathogenesis, there is an important role of the adaptive immune system with the presence of antigen-specific T and B cells. Lymphoid structures found in inflamed joints of RA patients, referred to as tertiary or ectopic lymphoid structures, contain distinct T cell zones and germinal centers.^{50,51} B cells present in these structures have been shown to secrete high affinity, class-switched ACPAs locally,⁵² but in vitro studies suggest that B cells from synovial fluid and bone marrow of ACPA-positive RA patients can also secrete ACPAs even without T cell help.^{53,54} While ACPA production may occur with or without T cell help, it is clear that T cells also play a major role in the pathogenesis of RA. CD4+ T cells have been identified that recognize self-peptides presented by HLA-DR molecules carrying the shared epitope motif linked to RA development and can differentiate into Th1 and Th17 cells.^{55,56}

Though not considered to be conventional immune cells, fibroblast-like synoviocytes in the joint augment the immune response via hyperplasia and pannus formation that invades cartilage, production of degradative enzymes that contribute to bone and cartilage damage, and secretion of pro-inflammatory cytokines.⁵⁷ Thus, the inflammatory infiltration not only causes local tissue damage, but also promotes changes to resident cells in the synovium, which creates a feedback loop that propagates inflammation the joint.

1.4 RA serology.

Clinically, seronegativity in RA is defined as being negative for rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPAs; for more information on citrullination, see section 1.5 'Citrullination and PADs in RA'). However, since other autoantibodies have been identified, as described below, it is important to understand if truly seronegative patients exist or if additional biomarkers may help to close the serological gap in patients with RA (reviewed in Trouw and Mahler 2012⁵⁸). This is important since seronegative patients are less likely to fulfill disease classification criteria and may be more difficult to diagnosis, possibly leading to delays in treatment.⁵⁹ Once thought to represent patients with a milder form of RA, recent studies suggest that seronegative patients can have severe treatment-refractory disease, but that remission can be achieved if treatment is initiated early in the disease course.⁶⁰

Seropositivity in RA is characterized clinically by the presence of anti-citrullinated protein antibodies (ACPAs) or rheumatoid factor (RF) and is included in the 2010 ACR classification criteria for RA.⁶¹ RF are IgM autoantibodies against the Fc portion of IgG antibodies and are present in 60–80% of RA population.^{53,62–65} ACPAs are also in 60–80% of patients with RA, depending on the study, and are highly specific for the disease.^{66–71} These antibodies recognize specific self-proteins in which arginine residues have been deiminated by peptidylarginine deiminase (PAD) enzymes, generating the non-classical amino acid citrulline.⁷² ACPAs are present in RA patients years before they show symptoms,^{73,74} and epitope spreading with recognition of an increased number of specific

citrullinated proteins occurs within a 2-year window of disease diagnosis.⁷³ ACPAs are detected clinically by the anti-cyclic-citrullinated peptide (anti-CCP) assay. It is important to highlight that the peptides used in the anti-CCP assay are short synthetic citrullinated peptides, originally based on peptides from citrullinated filaggrin, but are artificial antigens that are able to detect the pool of ACPAs present in most patients with RA.^{75,76} However, studies have shown that a small fraction of patients who are negative by the anti-CCP assay, have detectable ACPAs when specific citrullinated proteins are used as antigens.⁷⁷

[Recently], a growing number of other anti-modified protein antibodies (AMPAs) have been identified in patients with RA, including those to carbamylated (anti-CarP) and acetylated antigens as well as malondialdehyde-acetaldehyde (MAA) adducts.^{78,79} Formed as a result of the lipid peroxidation product malondialdehyde (MDA) chemically reacting with acetaldehyde and proteins, MAA adducts have been seen to co-localize with citrullinated proteins by confocal microscopy and are recognized by autoantibodies in 38% of ACPA-positive and 26% of ACPA-negative patients with RA.⁸⁰ Carbamylation, the conversion of lysine residues to homocitrulline through a reaction with cyanate, is also recognized by autoantibodies in a subset of patients with RA. Anti-CarP antibodies are present in 49–73% of ACPA-positive and 8–16% of ACPA-negative patients, and correlate with worse joint damage in ACPA-negative patients.^{81,82} Anti-acetylated vimentin antibodies have also been found in an average of 52.1% of ACPA-positive and 21.7% of ACPA-negative patients with RA.⁷⁹ Despite the growing study of these new antibody specificities in RA, their diagnostic and prognostic value in the clinical setting has yet to be evaluated.

Other autoantibodies are well described in patients with RA but are present less frequently and are currently not measured in clinical settings. These include members of the PAD enzyme family itself. Although there are five PAD isoenzymes in humans, PAD2 and PAD4 are implicated as the dominant producers of citrullinated autoantigens in RA. Both enzymes have also been shown to be recognized by autoantibodies in different subsets of patients with RA.⁸³⁻⁸⁵ While anti-PAD4 antibodies are associated with severe erosive joint disease and are strongly associated with classic RA genetic and serologic risk factors, anti-PAD2 antibodies appear to identify a serologically and genetically distinct group of patients with milder disease.^{84,86-88} Another well-described target in RA is a nuclear antigen known as RA33. Initially defined by Western blot as a 33 kDa protein in nuclear extracts from HeLa cells that was recognized by sera from RA patients, RA33 was later identified as hnRNP A2/B1, a protein of the spliceosome that can be targeted by autoantibodies in its citrullinated or non-citrullinated form.⁸⁹⁻⁹¹

1.5 Citrullination and PADs in RA.

Citrullination is the post-translational modification in which hydrolysis of the guanidinium group of positively charged arginine residues creates the neutrally charged, non-classical amino acid citrulline.^{92,93} This modification results in changes to protein structure and protein-protein interactions and is detected as a 0.98 Da mass-shift in mass spectrometry. Citrullination is catalyzed by a group of enzymes PADs, of which five isoforms exist.⁷² Each PAD isozyme has a unique tissue localization and some have overlapping substrates. Citrullination by PADs require calcium as a co-factor.⁹⁴⁻⁹⁶ Binding

of calcium induces conformational changes in the protein rendering it in an active state which enables the citrullination reaction to proceed in a reducing environment. Of the five isozymes, PAD2 and PAD4 have been implicated the most in RA pathogenesis.⁹⁷ While PAD2 is expressed by a variety of leukocytes, PAD4 is expressed primarily by granulocytes and monocytes of the immune system.^{72,98,99} PAD2 and PAD4 also have varying intracellular distributions with PAD2 being mainly found in the cytoplasm while PAD4, which contains a nuclear signal sequence and is capable of citrullinating nuclear proteins such as histones, is thought to reside primarily in the granulocyte nucleus.¹⁰⁰⁻¹⁰² The presence of ACPAs and autoantibodies against PAD enzymes as well as their association with more severe disease underscores the importance of PAD2 and PAD4 to RA pathogenesis.

Soluble PAD2 and PAD4 can also be detected in the synovial fluid of RA patients.¹⁰³⁻¹⁰⁵ Increased PAD2 and PAD4 mRNA translation, higher protein expression and increased activity were all noted in RA patients.^{103,106,107} Both PADs were also found in the synovial tissue of RA patients with more PAD4 than PAD2 being found.¹⁰⁶ While PADs are known to be selective of their substrates, interestingly, the RA citrullinome, or the collection of proteins known to be citrullinated by PADs in RA, consists of a wide array of intracellular and extracellular proteins with only a subset proven to be targeted by ACPAs.^{93,108-110} These include extracellular proteins such as fibrinogen and collagen as well as intracellular proteins such as histones, vimentin, NF- κ B, and actin.^{100,101,111-113} Additionally, while PAD2 and PAD4 create distinct patterns of citrullination,¹¹⁴ studies on citrullinated fibrinogen have shown that ACPAs preferentially recognize the PAD4-citrullinated protein.^{115,116} It was also thought that soluble PADs were responsible for extracellular citrullination while

intracellular citrullinated proteins were mainly derived from intrinsic PAD activity in the cell.^{93,117}

Although it is evident that PADs play a key role, genetic risk associations with the *PADI4* gene (see section 1.2 'Risk factors of RA'), autoantibodies against PAD4 correlating with a more severe and erosive disease (see section 1.4 'RA serology'), and amelioration of disease in mouse models of RA after PAD4 inhibition or knock-out demonstrates the importance of PAD4 to RA pathogenesis.^{118,119} However, mechanisms that lead to breaks in tolerance that result in immunogenicity towards citrullinated proteins and in the rise of ACPAs are still not fully understood. It is possible that PAD activation can cause unregulated citrullination creating new citrullinated proteins not previously seen by the immune system, which could lead to the development of ACPAs.⁹³ Another possibility is that hyperactivation of PADs could lead to new sites being citrullinated, which could cause structural changes leading to the processing and presentation of cryptic peptides that were not previously tolerized by the immune system.¹²⁰ However, under physiologic conditions, PAD4 is thought to be inactive because of low intracellular calcium concentrations and lack of a reducing extracellular environment.^{107,117,92} Since intracellular citrullinated proteins can be found extracellularly and are targeted by ACPAs,^{110,121} events that cause increases in intracellular calcium levels, such as differentiation or cell death, were thought to be causative of PAD4 activity.^{92,93}

Cell death pathways, especially those that are membranolytic, have been known to be key contributors to the pathogenic citrullination seen in RA. Membrane attack complexes caused by complement activation and perforin pore-forming pathways can

cause an influx of calcium due to cell lysis.¹²² Compromises in the integrity of the cell membrane can lead to neutrophil cell death accompanied by increased citrullination in a pathway referred to as 'leukotoxic hypercitrullination.'^{122,123} Other cell death pathways such as, NETosis and necrosis as well as autophagic mechanisms have also been implied to induce intracellular PAD activity and release of active PAD enzymes.^{117,123-125} Studies have shown that PAD4 is a required component of NETosis induction,¹²⁶ and citrullinated proteins have been found in NETs.¹¹⁷ Evidence of citrullination of a-enolase and vimentin with increased autophagy markers have been documented,¹²⁴ and necrosis has been shown to release more active PADs than NETosis.¹¹⁷ Furthermore, a subset of autoantibodies, anti-PAD3/4 cross-reactive autoantibodies, have been shown to lower threshold for calcium needed for activation due to inducing conformational changes that hold PAD4 in an active state.¹²⁷ This mechanism allows these antibodies to serve as a driver of citrullination by enhancing extracellular PAD activation.

Despite the advances made in understanding the origin of the RA citrullinome, recent evidence suggests that the generation of citrullinated proteins may be more complicated than previously understood. The presence of a nuclear localization signal sequence on the PAD4 transcript in conjunction with the fact that nuclear proteins such as histones were found to be citrullinated cemented the idea that the role of physiologic PAD4 was mainly nuclear.¹⁰⁰⁻¹⁰² Therefore, it was widely believed that intracellular PAD4 primarily accessed the extracellular environment via cell lytic pathways mentioned above. However, studies have shown that PAD4 plays a role in NADPH oxidase regulation in the cytoplasm of neutrophils and is expressed on the cell surface of both neutrophils and

monocytes, suggesting its physiological role extends beyond the nucleus.^{128,129} In addition, while most of the studies on citrullination and PAD4 were conducted in neutrophils, monocytes remain an underappreciated player in RA pathogenesis.

1.6 Monocytes in RA.

Monocytes are bone-marrow derived, mononuclear, innate immune cells that can be subcategorized based on their expression of CD14 and CD16 into the following: classical (CD14+, CD16-), intermediate (CD14+, CD16+), and non-classical (CD14-, CD16+).¹³⁰ In RA, the proportion of circulating monocytes is thought to correlate with disease activity,¹³¹ and CD14+ monocytes from RA patients are seen to be hyper inflammatory with metabolic capability for maintaining a prolonged immune response.¹³² Studies have shown that the CD14+CD16+ intermediate class of monocytes were found in higher numbers in the peripheral blood of RA patients.¹³³⁻¹³⁵ Monocytes with the intermediate phenotype from inflamed joints of patients with active RA have been shown to induce Th17 differentiation through possibly a direct cell-to-cell contact mechanism.¹³⁶ Synovial monocytes have also been shown to promote both Th1 and Th17 responses in RA.¹³⁴ Increased numbers of CD16-expressing monocytes have also been found in the RA synovial fluid.¹³⁷

Monocytes from RA patients have also been found to express increased markers of cellular adhesion and activation such as integrin subunits, including CD11b, CD11a, CD18, CD35 (complement receptor 1), and Fcγ receptors such as CD32 and CD64.^{138,139} Treatment with low dose prednisolone resulted in normalizing CD11b expression, which

correlated with clinical improvement, suggesting a role for adhesion and phagocytic receptors in inflammation seen in RA.¹³⁹ Healthy monocytes incubated with RA synovial fluid have also been shown to express elevated levels of co-stimulatory molecule, CD86, and culturing these treated monocytes with T cells caused increased production of IFN- γ and TNF- α , indicating the important role of monocytes in modulating adaptive immune responses.¹⁴⁰ RA patient monocytes were also found to secrete significantly higher levels of inflammatory cytokines such as IL-1 β and IL-6 compared to monocytes from healthy donors.¹³⁸

Monocytes are also one of the most versatile immune cells as they can differentiate into macrophages, monocyte-derived dendritic cells (mo-DCs), and osteoclasts depending on the environmental cues.^{141,142} Interestingly, binding of ACPAs to monocytes has been documented to cause their differentiation into osteoclasts and mo-DCs suggesting the presence of citrullinated proteins on the monocyte cell surface.^{143,144} However, the role of monocytes in RA pathogenesis has been mainly studied in the context of phagocytosing immune complexes and apoptotic cells and secreting pro-inflammatory cytokines.

1.7 Thesis overview.

While monocytes are one of the few immune cell subtypes that express PAD4,⁹⁹ their contribution to the RA citrullinome has been understudied. In addition, while some pathways that lead to the production of citrullinated proteins have been well documented, the recent discovery of PAD4 on the surface of neutrophils and monocytes raises the possibility that additional mechanisms of citrullinated autoantigen generation in the RA joint may exist. Therefore, the main objectives of this study were two-fold: (i) to elucidate the role of surface PAD4 on monocytes in generating the RA citrullinome and (ii) to understand the mechanisms of PAD4 trafficking. We hypothesized that surface PAD4 allows monocytes to generate citrullinated autoantigens, circumventing the need for cell death and other cell lytic mechanisms. We also hypothesized that PAD4 may be utilizing proteins found on the cell surface as binding partners and thereby gaining access to the exocytosis pathway to reach the cell surface.

CHAPTER 2

The monocyte surface is a novel site of autoantigen generation in propagating inflammation in rheumatoid arthritis.

Material presented in this chapter is adapted from: Thomas MA, Naik P, Wang H, Giles JT, Girgis AA, Kim SY, Johnson TP, Curran AM, Crawford JD, Jahanbani S, Bingham CO, Robinson WH, Na CH, Darrah E. The monocyte cell surface is a novel site of autoantigen generation in rheumatoid arthritis. (Resubmitted with revisions to the Proceedings of the National Academy of Sciences).

2.1 INTRODUCTION

The citrullinome refers to the set of proteins in which arginine residues can be post-translationally modified to the non-classical amino acid citrulline by the peptidylarginine deiminase (PAD) enzymes.^{113,145} Although this post-translational modification is a natural phenomenon,¹⁴⁶ immune responses are generated against citrullinated proteins in a large proportion of patients with rheumatoid arthritis (RA).⁶⁶⁻⁷¹ RA is a systemic autoimmune disease characterized by chronic inflammation and irreversible damage in the synovial joints, and a hallmark serological finding is the presence of anti-citrullinated protein antibodies (ACPAs).¹⁶ Citrullinated proteins accumulate in the synovial tissue and fluid of RA patients and include a variety of proteins derived from intracellular and extracellular sources, termed the RA citrullinome.^{109,110,113,122} While the RA citrullinome has been well characterized,^{109,110,113,122,147} the mechanistic origins of its generation, *i.e.*, the cellular drivers and the primary reaction sites, remain incompletely defined.

Of the five PAD enzymes found in humans, one of the major contributors to the RA citrullinome is the PAD4 isoform, which is present in high levels and co-localizes with citrullinated proteins in the inflamed RA synovium.^{72,96,106,107} Although its expression has been reported in a variety of cell types, PAD4 is predominantly expressed by granulocytes and monocytes and is unique among the PAD enzymes in its possession of a nuclear localization signal (NLS).^{99,100,148} The presence of an NLS and the observed nuclear staining of PAD4 in granulocytes and cellular transfection systems resulted in the historical view that PAD4 was a nuclear protein.¹⁰⁰ As a result, the mechanism by which a nuclear enzyme

could contribute to citrullination of extracellular autoantigens has been difficult to reconcile. While predominant theories center around the release of intracellular PADs into the extracellular space via cell lytic pathways and NETosis,^{122,149,93,126,117} it was recently found that a fraction of enzymatically active PAD4 can also localize to the plasma membrane of neutrophils,¹⁵⁰ opening unexplored avenues for pathogenic protein citrullination in RA.

Similar to granulocytes, monocytes are a major component of the immune infiltrate present in the RA joint and can express high levels of PAD4,^{16,99} but their contribution to the RA citrullinome remains largely overlooked. With the potential to differentiate into multiple effector immune profiles,¹⁵¹ monocytes represent an understudied but likely important generator of autoantigens in RA. Our study was conducted to assess the localization and role of PAD4 in monocytes and to understand its contribution to RA autoantigen generation. Monocytes were found to have a widespread subcellular localization of PAD4 with prominent surface expression. PAD4 on the monocyte surface was enzymatically active and capable of citrullinating both extracellular and endogenous surface proteins. Characterization of the monocyte surface citrullinome revealed numerous citrullinated plasma membrane proteins and identified citrullinated Mac-1 (CD11b/CD18) as a novel cell surface autoantigen targeted by ACPAs in a subset of RA patients. Thus, PAD4 expression on the monocyte cell surface suggests a new mechanism for citrullinated antigen production as it transforms the monocyte surface into a site of autoantigen generation in RA.

2.2 METHODS

Human subjects. Peripheral blood was obtained from the following sources: 1) healthy donors who provided written informed consent for enrollment in a study of healthy volunteers; 2) de-identified leukopaks classified as medical excess from healthy donors who donated platelets for medical use at the Anne Arundel Blood Donation Center; and (3) rheumatoid arthritis patients belonging to a longitudinal cohort at the Johns Hopkins Arthritis Center.

Healthy control sera used for autoantibody classification studies were obtained from n = 75 volunteers without a history of autoimmune disease, cancer, or chronic infection. Sera from a small convenience cohort of deidentified ACPA⁻ RA (n = 18) and ACPA⁺ RA (n = 18) patients seen at the Johns Hopkins Arthritis Center were used as a discovery cohort for studying autoantibody recognition of CD11b. Sera from 165 patients from the “Evaluation of Subclinical Cardiovascular Disease and Predictors of Events in Rheumatoid Arthritis” (ESCAPE RA) cohort were used in anti-Mac-1 antibody studies.⁸⁴ Patients in the ESCAPE RA cohort met the 1987 ACR-EULAR Classification Criteria for Rheumatoid Arthritis and patients in the convenience cohort were diagnosed with RA by a board-certified rheumatologist.¹⁵² All participants of this study provided informed consent, and the study was approved by the Johns Hopkins Institutional Review Board.

Isolation of monocytes and neutrophils from whole blood. Heparinized whole blood was collected and separated using Ficoll-Paque (GE Healthcare)–based density gradient centrifugation. Monocytes were purified from the peripheral blood mononuclear cell

(PBMC) fraction using either: (1) CD14 positive–selection beads (human CD14 Microbeads, Miltenyi Biotec) according to manufacturer instructions or (2) negative–isolation using EasySep Human Monocyte Isolation Kit (STEMCELL Technologies) for the transmission electron microscopy study. Following density gradient centrifugation of peripheral blood, neutrophils were isolated from the red blood cell layer after erythrocyte lysis using ACK lysis buffer (Thermo Fisher Scientific).

For studies comparing surface PAD4 expression between healthy controls and RA patients, PBMCs isolated from peripheral blood were cryopreserved in either 90% heat-inactivated fetal bovine serum (FBS; Thermo Fisher Scientific) supplemented with 10% dimethyl sulfoxide (DMSO) or in Recovery Cell Culture Freezing Medium (Thermo Fisher Scientific). Frozen cells were thawed, and monocytes were isolated using CD14 positive–selection beads.

Determining anti-PAD4 antibody specificities. Commercial and in-house generated anti-PAD4 antibodies targeting three distinct epitopes were utilized for this study: rabbit anti-PAD4 polyclonal antibody (P4749, MilliporeSigma [targeting the N-terminus, amino acids 1-15]); mouse anti-PAD4 monoclonal antibody (ab128086, Abcam, Clone OTI4H5 [epitope not reported]); and in-house rabbit anti-PAD4 polyclonal antibody [targeting the C-terminus, amino acids 515-528].¹⁵³ Three separate evaluations were conducted to determine the specificity of these antibodies for PAD4 and are reported in **Figure 2-1**. First, 5 µg of cell lysate from healthy control monocytes (known to be PAD4 positive), monocytic cell lines THP-1 and U937 (PAD4 negative), and a non-immune cell line, HSG (a

PAD4 negative, possible HeLa cell line), were resolved by SDS-PAGE and immunoblotted for PAD4 using the three anti-PAD4 antibodies with and without competitive inhibition using recombinant human (rh) PAD4. For blocking, the anti-PAD4 antibodies were pre-incubated with a 5-molar excess of rhPAD4 protein for 1 hour at RT prior to immunoblotting. The second evaluation was performed by immunoblotting 500 ng of rhPAD4 and rhPAD2 proteins, purified as previously described^{153,154}, with all three anti-PAD4 antibodies. The third evaluation was a standard protein BLAST search of the NCBI Protein Reference Sequence (refseq_protein) *Homo sapiens* (taxid: 9606) database using the NCBI BLAST blastp suite against the known N- (MAQGTLIRVTPEQPT) and C- (FEGIKKKKQQKIKN) terminal sequences used as immunogens to generate the MilliporeSigma and in-house anti-PAD4 antibodies, respectively. To analyze homologous proteins, isoforms were collapsed into a single hit based on the parental RefSeq (NP_) ID appearing first in the list, and only proteins with homology between 50-100% were included in the analysis. Overlapping homologous candidates between the two antibody epitopes were visualized using a Venn diagram.

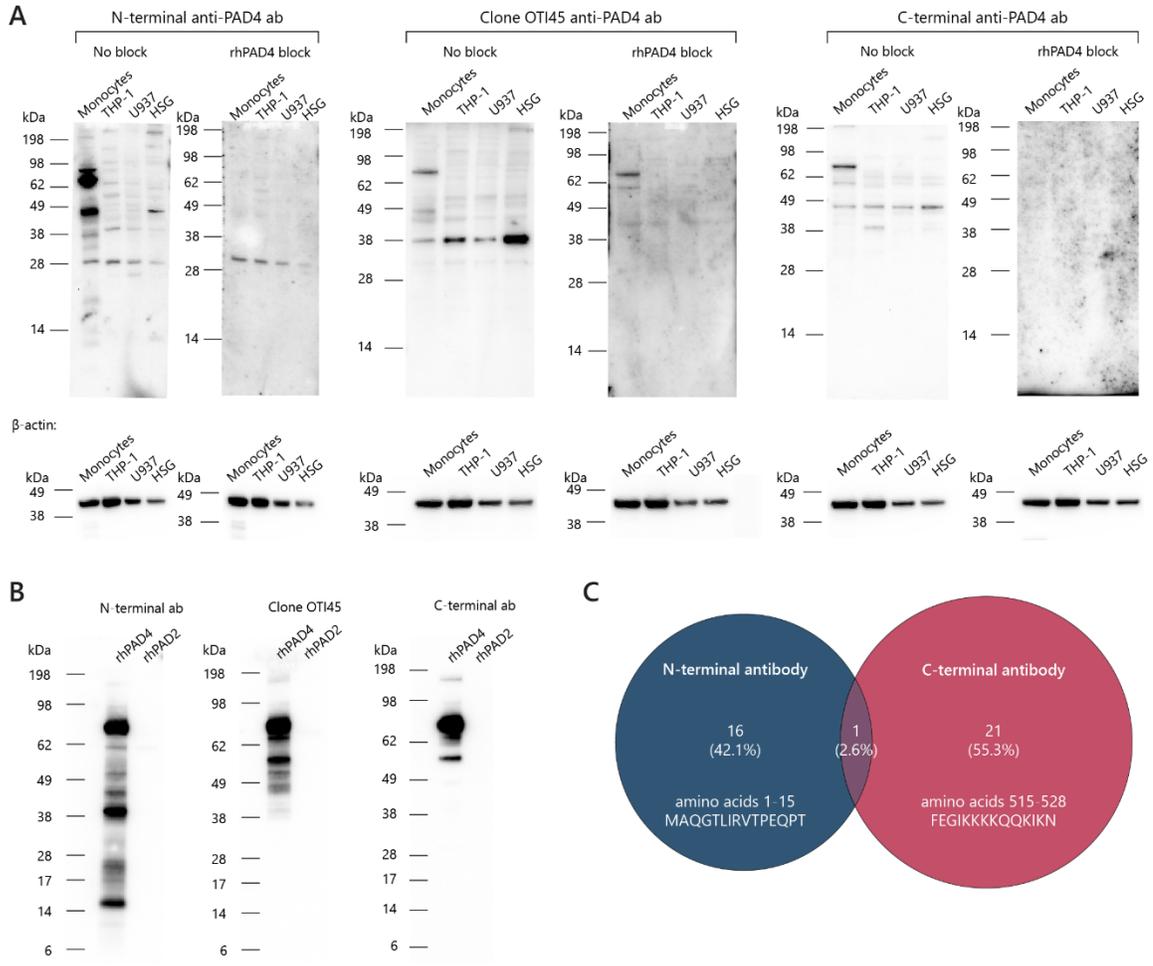


Fig. 2-1. Specificities of anti-PAD4 antibodies. (A) Blocking experiments showing specificity of anti-PAD4 antibodies for human PAD4. 5 μ g of cell lysates of monocytes isolated from peripheral blood of healthy donors (known PAD4 positive) and PAD4-negative cell lines (including monocytic cell lines, THP-1 and U937, as well as a non-monocytic cell line, HSG) were resolved by SDS-PAGE and immunoblotted for PAD4 using three anti-human PAD4 antibodies used in the main manuscript: N-terminal antibody (Sigma), clone OTI45 (Abcam), and C-terminal antibody (in-house). Immunoblotting was also performed in parallel using the same antibodies after incubation with 5-molar excess of rhPAD4 (labeled as 'rhPAD4 block'). β -actin levels served as loading controls. (B) Immunoblots showing specificity of anti-PAD4 antibodies for PAD4 over the highly related protein, PAD2. 500 ng of rhPAD4 or rhPAD2 were immunoblotted using three anti-human PAD4 antibodies: N-terminal antibody (Sigma), clone OT145 (Abcam), C-terminal antibody (in-house). (C) Analysis of homology of known anti-PAD4 epitopes to other proteins in the human genome. Venn diagram displaying the proteins sharing sequence similarity with the immunogen for each respective anti-human PAD4 antibodies (N-terminal antibody in blue; C-terminal antibody in red).

Immunofluorescence studies. To study PAD4 localization, monocytes were positively selected from three healthy donors and fixed directly using 4% paraformaldehyde for 5 minutes at room temperature (RT). Fixation was followed by blocking using 2% (w/v) bovine serum albumin (BSA; MilliporeSigma) in PBS for 30 minutes. Cells were incubated with a 1:100 dilution of each anti-PAD4 antibody (P4749, MilliporeSigma [N-terminus]; ab128086 [Abcam, Clone OT14H5]; or in-house anti-PAD4 [C-terminus]¹⁵³) for 1 hour at RT. After washing, cells were incubated with a 1:200 dilution of Alexa Fluor™ (AF) 488 anti-rabbit or AF594 anti-mouse secondary antibodies (Invitrogen) for 30 minutes at RT. Cells were mounted directly in ProLong Gold Antifade Mounting Media with DAPI (Invitrogen). Slides were dried overnight before visualization via the 20X objective on Zeiss Axioskop 50 fluorescent microscope with a Zeiss AxioCam HRC camera.

To explore PAD4 expression and activity, positively selected monocytes (n = 3 healthy donors) were incubated for 3 hours at 37°C in either 10 mM EDTA or 1.5 mM Ca²⁺ in Hank's Balanced Salt Solution (HBSS) and fixed and blocked using the methods described above. Cells were incubated in 1:200 dilution of anti-PAD4 (P4749; MilliporeSigma [N-terminus]) and anti-peptidyl citrulline IgM (clone F95; MilliporeSigma) antibodies for 1 hour at RT. They were labeled using AF594 anti-rabbit (for PAD4; Invitrogen) and AF488 anti-mouse IgM (heavy chain; for F95; Invitrogen) secondary antibodies for 30 minutes at RT. Slides were prepared as mentioned above and visualized using Zeiss Axioskop 50 fluorescent microscope with a Zeiss AxioCam HRC camera.

To visualize surface citrullination using confocal microscopy, positively selected monocytes treated with 1.5 mM Ca^{2+} were co-stained with anti-GAPDH (clone D16H11 #5174, Cell Signaling Technology) and anti-peptidyl-citrulline IgM (clone F95) antibodies. The cells were washed and incubated with AF594 anti-rabbit (for GAPDH, Invitrogen) and AF488 anti-mouse IgM (heavy chain; for F95; Invitrogen) secondary antibodies. The slides were washed, mounted and counterstained directly in ProLong Gold Antifade Mounting Media with DAPI (Invitrogen), and examined using a Zeiss LSM700 confocal microscope. Brightness and contrast of images from each experiment were equalized to the same values using Fiji software.¹⁵⁵ Images were scaled when appropriate and scale bars were added using Fiji.

Transmission electron microscopy. Negatively selected monocytes (5×10^7) were purified from PBMCs of a healthy donor leukopak using the EasySep Human Monocyte Isolation Kit (STEMCELL Technologies). Samples were fixed in 4% (w/v) paraformaldehyde, 0.1% (w/v) glutaraldehyde, 3 mM MgCl_2 in 0.1 M Sorenson's sodium phosphate buffer (pH 7.2) overnight at 4°C. After buffer rinse, samples were postfixed in 1% (w/v) osmium tetroxide, 0.8% (w/v) potassium ferrocyanide in 0.1 M sodium phosphate for 1 – 2 hours on ice in the dark. Samples were subsequently rinsed in 0.1 M maleate buffer, followed by *en bloc* staining in 2% (w/v) uranyl acetate in maleate buffer for 1 hour in the dark, dehydrated in a graded series of ethanol, and embedded in Eponate 112 resin (Polyscience). Samples were polymerized at 60°C overnight.

Thin sections, 60 to 90 nm, were cut with a diamond knife on a Leica UCT ultramicrotome and picked up with Formvar coated 200 mesh nickel grids. For antibody labeling, grids were wetted with dH₂O and floated on all subsequent steps. Sections were first etched with 1.5% sodium meta periodate, rinsed in dH₂O before 10 mM NH₄Cl in TBS (Tris Buffered Saline), and blocked for 30 minutes in 2% (w/v) BSA in TBS followed by an overnight incubation at 4°C in primary antibody (1:25 for C-terminal antibody, 1:50 for P4749 N-terminal antibody and clone OT14H5 from Abcam, or buffer only). After overnight incubation, grids were brought to room temperature for 1 hour, rinsed with blocking buffer and TBS, and incubated in 12 nm gold nanoparticle-conjugated secondary antibody (AffiniPure goat anti-rabbit IgG [111-205-144] and AffiniPure goat anti-mouse IgG [115-205-146] from Jackson ImmunoResearch Laboratories; 1:40) at room temperature for 2 hours. Samples underwent a primary fixation in glutaraldehyde in sodium cacodylate buffer followed by staining with 2% (w/v) uranyl acetate and rinsed with water. Grids were stored in a cool, dark place until imaged with a Hitachi 7600 TEM at 80 kV. Images were captured with an AMT CCD XR80 (8-megapixel camera – side mount AMT XR80 – high-resolution high-speed camera). Scale bars were added in Fiji.

Flow cytometry analysis of surface PAD4, monocyte viability, and surface

citrullination. Rabbit anti-human PAD4 antibody [generated in-house¹⁵³] was directly labeled with the AF647 antibody labeling kit (Invitrogen) according to manufacturer's instructions. 1.5x10⁶ positively selected monocytes isolated from PBMCs were plated per condition and immediately subjected to live/dead stain (blue fluorescent reactive dye, Invitrogen). Cells were then washed and blocked using human BD Fc Block (BD

Biosciences) in 3% (w/v) BSA in PBS for 10 minutes at RT. Cells were stained for flow cytometry using CD14-PE (555398, BD Biosciences) or CD14-BV605 (301834; BD Biosciences), CD16-BV785 (563689, BD Biosciences) and co-stained with 0.01 $\mu\text{g}/\mu\text{L}$ of either AF647-labeled rabbit IgG isotype control (3452S, Cell Signaling Technology) or AF64-conjugated C-terminal anti-PAD4 antibody for 30 minutes at 4°C. Studies analyzing surface PAD4 expression compared to isotype control were conducted using monocytes from three healthy donors. Monocytes from six healthy controls and six RA patients were used to compare surface PAD4 levels.

To analyze viability following a 3-hour incubation, monocytes positively isolated from peripheral blood and neutrophils of healthy controls ($n = 3$) were incubated in HBSS and 10 mM HEPES supplemented with 5 mM Ca^{2+} for three hours at 37°C. Cells were harvested and stained using live/dead stain (blue fluorescent reactive dye, Invitrogen).

For surface citrullination analysis, monocytes were suspended in HBSS and cultured in a flat bottom 96-well plate at a density of 2×10^6 cells/mL in the presence of Ca^{2+} (0-1.5 mM) at 37°C for 3 hours. Cells were washed and blocked as described earlier and then labeled with 1 $\mu\text{g}/\text{mL}$ of anti-peptidyl-citrulline IgM (clone F95) primary antibody in cold 3% BSA in PBS for 30 minutes at 4°C. They were subsequently stained with AF488 anti-mouse IgM (heavy chain) secondary antibody (Invitrogen) in cold 3% BSA in PBS for 30 minutes. Experiments were conducted using cells from three healthy donors.

Following staining, cells were washed with cold 3% BSA in PBS supplemented with 1% sodium azide and resuspended in PBS for flow cytometry using a FACS Aria II (BD

Biosciences) or Cytex Aurora (Cytex Biosciences) at the Johns Hopkins Bayview Immunomics Core Facility. All flow cytometry data were analyzed using FCS Express 7 Research Edition (De Novo Software).

Surface PAD4 citrullination assay using exogenous fibrinogen. Monocytes and neutrophils were seeded in a 96-well plate at a density of 2×10^6 cells/mL in HBSS per experimental condition. Cells were incubated with 0.7 μ M fibrinogen (341578, MilliporeSigma) in the presence of Ca^{2+} (0-1.5 mM) at 37°C for 3 hours. To assess the dependency of fibrinogen citrullination on PAD4, monocytes were incubated with the PAD4-specific inhibitor GSK484 (17488, Cayman Chemical Company) for 15 minutes at RT before the subsequent addition of 1.5 mM Ca^{2+} . Cell-free supernatants from the extracellular citrullination assay were obtained, denatured at 95°C for 5 minutes, and separated using SDS-PAGE (4–12% gradient gel), and immunoblotted with an anti-citrullinated fibrinogen antibody (1:1000; clone 20B2, Modiquest) followed by HRP-conjugated anti-rabbit secondary antibody (Invitrogen). Chemiluminescence was recorded using the Protein Simple FluorChem M detection system. Data from three independent experiments were analyzed.

Determining levels of extracellular and intracellular PAD4. $3 - 8 \times 10^6$ monocytes positively isolated from peripheral blood and equal numbers of neutrophils from three healthy controls were incubated in 100 μ L HBSS and 10 mM HEPES with 5 mM Ca^{2+} in a 96-well plate at 37°C for 3 hours. Plates were centrifuged following incubation, and cell-free supernatants were removed. Cell pellets were lysed in NP40 lysis buffer supplemented

with 10 mM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitors including chymostatin, antipain, leupeptin, and pepstatin (Calbiochem). PAD4 levels in supernatants and cell pellets were determined using PAD4 (human) ELISA kit (Cayman Chemical Company) according to manufacturer's instructions.

ACPA monoclonal specificity and recognition of citrullinated fibrinogen. Antibody sequences of plasmablasts collected from PBMCs of ACPA⁺ RA patients were barcoded and the corresponding monoclonal antibodies (mAbs) were produced recombinantly as described previously.^{156,157} Monoclonal antibodies (n = 8) were characterized for reactivity against various citrullinated autoantigens, including human fibrinogen.^{157,158}

500 ng of fibrinogen (MilliporeSigma) was coated on high-binding ELISA plates (Corning) overnight and blocked using 2% BSA in PBS for 2 hours at 37°C. As a positive control, 20 ng of rhPAD4 in 5 mM Ca²⁺, 1 mM DTT, and 100 mM Tris-HCl (pH = 7.5) were added to one set of wells. For the cell-based citrullination conditions, 1x10⁷ positively selected monocytes in HBSS with 10 mM HEPES and either 5 mM Ca²⁺ or 10 mM EDTA were added. Cells and rhPAD4 were allowed to citrullinate at 37°C for 3 hours. Wells were washed using 0.05% Tween-20 in PBS (PBST) and incubated with 10 µg of ACPA mAb or 0.5 µg anti-citrullinated fibrinogen antibody (clone 10E9.3, Cayman Chemical Company) in 0.2% BSA in PBST per well overnight at 4°C. Wells were then incubated with a 1:7500 dilution of HRP-conjugated anti-human IgG or anti-mouse IgG in 0.2% BSA in PBST at RT for 1 hour. TMB substrate (KPL) was added, and the reaction was stopped using 1 M HCl. Absorbance was read at 450 nm with background at 560 nm on a Wallac Victor 3 1420 Multilabel Counter plate reader (PerkinElmer).

To define the human ACPA mAb specificity against citrullinated fibrinogen, reactivity of the ACPA mAbs to citrullinated fibrinogen, generated as described above, was compared to reactivity to 500 ng of fibrinogen (MilliporeSigma) or carbamylated fibrinogen (Cayman Chemical Company). 10 µg of ACPA mAbs (n = 7), 0.5 µg of anti-citrullinated fibrinogen antibody (clone 10E9.3, Cayman Chemical Company), or 0.5 µg of anti-carbamylated fibrinogen antibody (clone 1C6, Cayman Chemical Company) were added in 0.2% BSA in PBST overnight at 4°C. Species-specific HRP-conjugated secondary antibodies were added at 1:7500 dilution in 0.2% BSA in PBST for 1 hour at RT, followed by addition of substrate and data acquisition as described above.

Plasma membrane isolation for mass spectrometry. Positively selected monocytes resuspended at 5×10^7 cells/mL in HBSS with 10 mM HEPES were incubated at 37°C for 3 hours in either 5 mM Ca^{2+} or 10 mM EDTA. Plasma membrane and cytosol fractions were isolated using the Minute Plasma Membrane Protein Isolation and Cell Fractionation Kit (Invent Biotechnologies). Plasma membrane pellets were resuspended in 50 µL of NP40 lysis buffer supplemented with 10 mM PMSF and protease inhibitors. A total cell lysate sample was created by lysing 1×10^7 monocytes in the same supplemented NP40 lysis buffer. Protein amounts were quantified by following the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) protocol. 5 µg of protein per condition was immunoblotted using antibodies against GAPDH (1:1000, rabbit, Cell Signaling), Na^+/K^+ -ATPase (1:1000, mouse, Abcam), and lamin B1 (1:500, mouse, Abcam) to ensure fraction enrichment. Plasma membrane fractions from n = 3 independent experiments were analyzed via mass spectrometry for citrullinated proteins.

Preparing samples for mass spectrometry studies. Detergents in lysates of plasma membranes were removed by methanol-chloroform precipitation method followed by reconstituting proteins in lysis buffer consisting of 8 M urea, 10 mM tris(2-carboxyethyl)phosphine (TCEP), and 40 mM chloroacetamide (CAA) in 50 mM triethylammonium bicarbonate (TEAB). After reduction and alkylation of proteins for 1 hour at room temperature, 8 M of urea was diluted to 2 M by adding 50 mM TEAB and proteins were digested with trypsin (sequencing grade modified trypsin, Promega, Fitchburg, WI, USA) at 10 ng/ μ L (v/v) at 37°C overnight. After digestion, tryptic peptides were acidified with 1% trifluoroacetic acid (TFA) to the final concentration, and then desalted using C₁₈ StageTips. The eluted peptides were vacuum dried and stored at -80°C before mass spectrometry analysis.

Mass spectrometry data analysis. The peptides were analyzed on an Orbitrap Fusion Lumos Tribrid mass spectrometer interfaced with an Ultimate 3000 RS Autosampler nanoflow liquid chromatography system (Thermo Scientific). The acquired mass spectra files were fed into Proteome Discoverer (version 2.5.0.4, Thermo Scientific) to identify and quantify proteins by searching against the human UniProt database (released in Jan. 2021) containing common contaminant proteins using SEQUEST HT algorithms. Parameters for data search is as follows: a) trypsin as a proteolytic enzyme with two maximum missed cleavage sites; b) precursor mass error tolerance of 10 ppm; c) fragment mass error tolerance of 0.02 Da; d) carbamidomethylation (+57.02146 Da) at cysteine as fixed modifications; and e) oxidation at methionine (+15.99492 Da), deamidation (+0.98402 Da) at arginine, asparagine and glutamine, acetylation (+42.01057 Da) at protein N-terminus,

methionine loss (-131.04049 Da) at methionine, and methionine loss with acetylation (-89.02992 Da) at methionine as variable modifications. The minimum length of peptide was set to six amino acids. Peptides and proteins were filtered at 1% false discovery rate.¹⁵⁹⁻¹⁶¹

Mass spectrometry analysis results were sorted on peptides found to contain deamidated (*i.e.* citrullinated) arginine. Citrullinated arginine occurring at C-terminal residues were removed from downstream analysis. The first accession number corresponding to the citrullinated peptides was used to determine protein identities. Of the 393 proteins identified to be citrullinated, 24 were removed due to low confidence. Abundances of the remaining 369 proteins were determined in the three donors and those shared by all three donors (337 proteins) in the calcium condition were used for subsequent analysis. Abundances of the 369 proteins for all three donors were determined in the EDTA condition for comparison analyses.

The cellular localization of the 337 citrullinated proteins in the calcium condition were determined using the 'subcellular location' data from the UniProt Database.¹⁶² Since the accession numbers of three proteins were labeled as 'deleted', the first accession number for the gene name found in UniProt was used instead. Proteins were sorted according to their first designated location using the following terms: (1) SUBCELLULAR LOCATION terms: "cell membrane", "membrane", "apical cell membrane", "basolateral cell membrane", "cell junction"; or (2) Gene Ontology terms: "cell surface", "membrane", "cell junction", "cell-cell junction". Seven proteins did not have subcellular localization information in the database and were therefore removed from further analysis. The

remaining proteins were then manually refined to remove those that appear mainly on organelle membranes. The abundances of proteins in the final list were averaged across the three donors.

All proteins found in the human proteome with a status designation of 'reviewed' in the UniProt Database were sorted according to their first designated location using the following terms: (1) SUBCELLULAR LOCATION terms: "cell membrane", "membrane"; or (2) Gene Ontology terms: "cell surface", "membrane" to create a list of all the proteins found on the plasma membrane. Statistical overrepresentation analysis comparing citrullinated membrane proteins to the list of all plasma membrane proteins found in the human proteome was conducted using the PANTHER classification system.^{163,164} Accession numbers of 13 proteins that could not be found in PANTHER were substituted for the first accession number found for the same gene in UniProt. Annotation sets used in this analysis were: PANTHER GO Biological Processes, PANTHER GO Molecular Function, PANTHER protein class, and PANTHER pathways. Fisher's exact test with Benjamini-Hochberg false discovery rate (FDR) correction was used for statistical analysis. Threshold values were set at FDR *p value* = 0.01 and at enrichment value of ± 2 . Factors with a fold enrichment value equal to zero or factor labeled 'unclassified' were omitted from graphical depiction.

Mapping citrullination sites. Predicted structures of CD18 (AF-A0A494COX7-F1) and CD11b (AF-P11215-F1) were downloaded from AlphaFold Protein Structure Database.^{165,166} Citrullinated arginines were manually recolored using PyMOL Molecular Graphics System v2.5.4 (Schrödinger). AlphaFold v2.3.1 was used to generate predicted

native and citrullinated structures for CD18 and CD11b. Due to computational constraints, a truncated form of CD11b was folded comprising amino acids 1-400 of the I-domain, which encompassed three of the four citrullination sites. Glutamine residues were substituted for citrulline at mass spectrometry-confirmed citrullination sites on both proteins for two reasons: (i) glutamine and citrulline share the same side chain and charge and (ii) glutamine has been historically used to model changes caused by citrulline.^{167,168} The AlphaFold Docker script with the following parameters was used to fold protein sequences: "max_template_date = 2020-05-14", "model_preset = monomer", and "db_preset = full_dbs". Folding was performed locally on Ubuntu 22.04 'Jammy Jellyfish' using an AMD Ryzen 7 5800X CPU, 32Gb RAM, and an MSI NVIDIA GeForce GTX 1660 GPU. Of the predicted models, those with the first rank following Amber relaxation were used for downstream analysis. The 'align' command was used to align the native and citrullinated structures, and the 'rms' command was used to calculate the RMSD score between the structures in PyMOL.

In-vitro transcription translation (IVTT)-immunoprecipitation (IP) studies. Human Mac-1 alpha (plasmid #8631, Addgene) was used to generate S³⁵-labelled CD11b by following the TnT Quick Coupled Transcription/Translation Systems kit (L1171, Promega) protocol. The mixture was either citrullinated using 600 µM rhPAD4 in 100 mM Tris-HCl with 5 mM Ca²⁺ and 1 mM DTT at 37°C or kept on ice in 100mM Tris-HCl for 2 hours . Sera from 18 healthy controls, 18 ACPA⁻ RA patients, and 18 ACPA⁺ RA patients were used to immunoprecipitate native and citrullinated CD11b. Briefly, 1 µL of serum in NP40 lysis buffer supplemented with PMSF, protease inhibitors, and 0.2% (w/v) BSA was incubated

with native and citrullinated CD11b for 2 hours at 4°C. Antibody bound-CD11b was pulled down using Dynabeads Protein G magnetic beads (Thermo Fisher Scientific) at 4°C for 1 hour. Samples were denatured at 70°C for 15 minutes, separated using gel electrophoresis, and visualized using radiography. Radiographs were analyzed using Alpha View SA software (ProteinSimple) and bands were normalized by first subtracting background binding to beads and then dividing the amount of CD11b immunoprecipitated by healthy control and RA patient sera by the positive control, anti-human CD11b (clone EP1345Y; ab52478, Abcam). The cut-off value for positivity was set to the 90th percentile of healthy control recognition of citrullinated CD11b.

Autoantibody reactivity against native vs citrullinated Mac-1. 100 ng of recombinant Mac-1 (R&D Systems) diluted in PBS was coated per well on high-binding ELISA plates (Corning) overnight at 4°C. Wells were washed using PBST and blocked with Pierce Protein-Free T20 (PBS) Blocking Buffer (PFTB; Thermo Fisher Scientific) at 37°C for 1 hour. Wells were incubated in 50 ng rhPAD4 and 1 mM DTT with either 5 mM Ca²⁺ or 10 mM EDTA in 100 mM Tris-HCl (pH = 7.5) for 4 hours at 37°C. Background wells that did not contain fibrinogen were also incubated with 50 ng rhPAD4 and 1 mM DTT with either 5 mM Ca²⁺ in 100 mM Tris-HCl following blocking. Healthy control or RA patient sera were diluted 1:200 in 1% PFTB block in PBST and incubated at RT with shaking for 2 hours. Anti-CD11b antibody (clone EPR1344; ab133357; Abcam) was used to create the standard curve, and a 1:550 dilution of RA82 ACPA mAb (previously used in the ELISA to detect extracellular citrullinated fibrinogen) was used as positive control. HRP-labeled secondary antibody (1:7500 dilution of anti-rabbit IgG [Invitrogen] or 1:2000 dilution of anti-human Fc gamma

[Zymed]) was diluted in 1% PFTB in PBST and incubated for 1 hour at RT with shaking. TMB substrate (KPL) was added, and the reaction was stopped using 1M HCl. Absorbances at 450 nm and 560 nm (background) were read using a Wallac Victor 3 1420 Multilabel Counter plate reader (PerkinElmer).

Data analysis was performed by normalizing to the positive control and cut-off for positivity was set to the 90th percentile of healthy control recognition of citrullinated Mac-1. Briefly, raw absorbances were compared to an 8-point standard curve created using the anti-human CD11b antibody (ab133357, Abcam). Four-parameter analysis curves were fit to the data and relative values with respect to the antibody dilutions used to create the standard curve were assigned. Raw absorbances of reactivity from sera to native and citrullinated Mac-1 as well as to the background wells were then converted into relative units using the standard curve generated for each plate. Background units were then subtracted from the native and citrullinated units. To account for inter-assay variation, an ACPA mAb (RA82; 0.5 µg per well) of known positivity was run on each plate as an internal control. All reactivity from sera to native and citrullinated Mac-1 were then divided by the ACPA mAb positive control run on its respective plate resulting in "Normalized anti-Mac-1 units".

To calculate autoantibody preference for citrullinated Mac-1, binding to native Mac-1 was subtracted from the binding to citrullinated Mac-1. The new cut-off for positivity was set to 90th percentile of healthy control preference for citrullinated Mac-1. Antibody levels that exceeded the threshold for positivity were referred to as 'Mac-1

ACPAs'. Thirteen RA patients with Mac-1 ACPA levels that met the cut-off for positivity, but were negative in the original analysis, were considered to be negative for Mac-1 ACPA as well. Of the 165 RA patients in the ESCAPE cohort, one patient did not have CCP information and had to be removed from further analysis. Three CCP⁻ patients had anti-citrullinated Mac-1 antibodies and also had to be removed for consistency in further analysis. Comparisons of disease characteristics were performed using two-tailed t-tests and graphed using R (version 4.0.4).

Statistical analyses. Data from the flow cytometry experiments and antibody binding assays are expressed as mean \pm standard deviation. An ordinary two-way analysis of variance (ANOVA) with a Dunnett's correction for multiple comparisons was performed to evaluate PAD4 activity on the monocyte surface at increasing calcium concentrations relative to 0 mM Ca²⁺. All other graphs were analyzed using two-tailed Student's t-tests except when stated. Analyses between groups of autoantibody binding to citrullinated CD11b and citrullinated Mac-1 were performed using one-tailed Student's t-tests. Fisher's exact test was used as the contingency test in analyzing the relationship between autoantibodies and RA, and the Baptista-Pike test was used to compute the odds ratio. Disease characteristics according to Mac-1 ACPA status were compared using t-tests for normally distributed continuous characteristics, the Kruskal-Wallis test for non-normally distributed continuous variables, and the chi-square goodness-of-fit or Fisher's exact test, as appropriate, for categorical variables. We constructed linear and logistic regression models, according to the characteristic, to explore the association of Mac-1 ACPA status with RA characteristics, adjusting for potential confounders identified from univariate

models. Non-normally distributed continuous variables were transformed for linear regression as required. Adjusted means and percentages and their 95% confidence intervals (Cis) were calculated and transformed variables were back-transformed for ease of interpretation. Clinical data were analyzed using Stata/SE (version 16) and all other analyses were performed using GraphPad Prism 9 Software Inc. (La Jolla, CA, USA). A p-value < 0.05 was considered statistically significant throughout.

Data, Materials, and Software availability. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD044370.¹⁶⁹

2.3 RESULTS

2.3.1 PAD4 exists on the surface of monocytes.

Since the subcellular localization of PAD4 in monocytes is underexplored, we first examined the cellular distribution of PAD4 in healthy donor monocytes by indirect immunofluorescence using antibodies raised against three distinct epitopes of PAD4: N-terminus, full-length recombinant PAD4, and C-terminus. Images revealed widespread expression of PAD4 throughout the monocyte with strong extranuclear staining, irrespective of the antibody used (**Figure 2-2A**). To further define the localization of PAD4 in monocytes at high resolution, we performed transmission electron microscopy (TEM) using the same three antibodies against distinct epitopes of PAD4. TEM imaging supported the widespread expression of PAD4 seen by immunofluorescence, including its expected nuclear expression (**Figure 2-2B**),¹⁰⁰ as well as its localization at the monocyte cell membrane (**Figure 2-2C**). Surface expression of PAD4 could be seen with all three antibodies tested, suggesting that the full-length PAD4 protein is present at the monocyte plasma membrane.

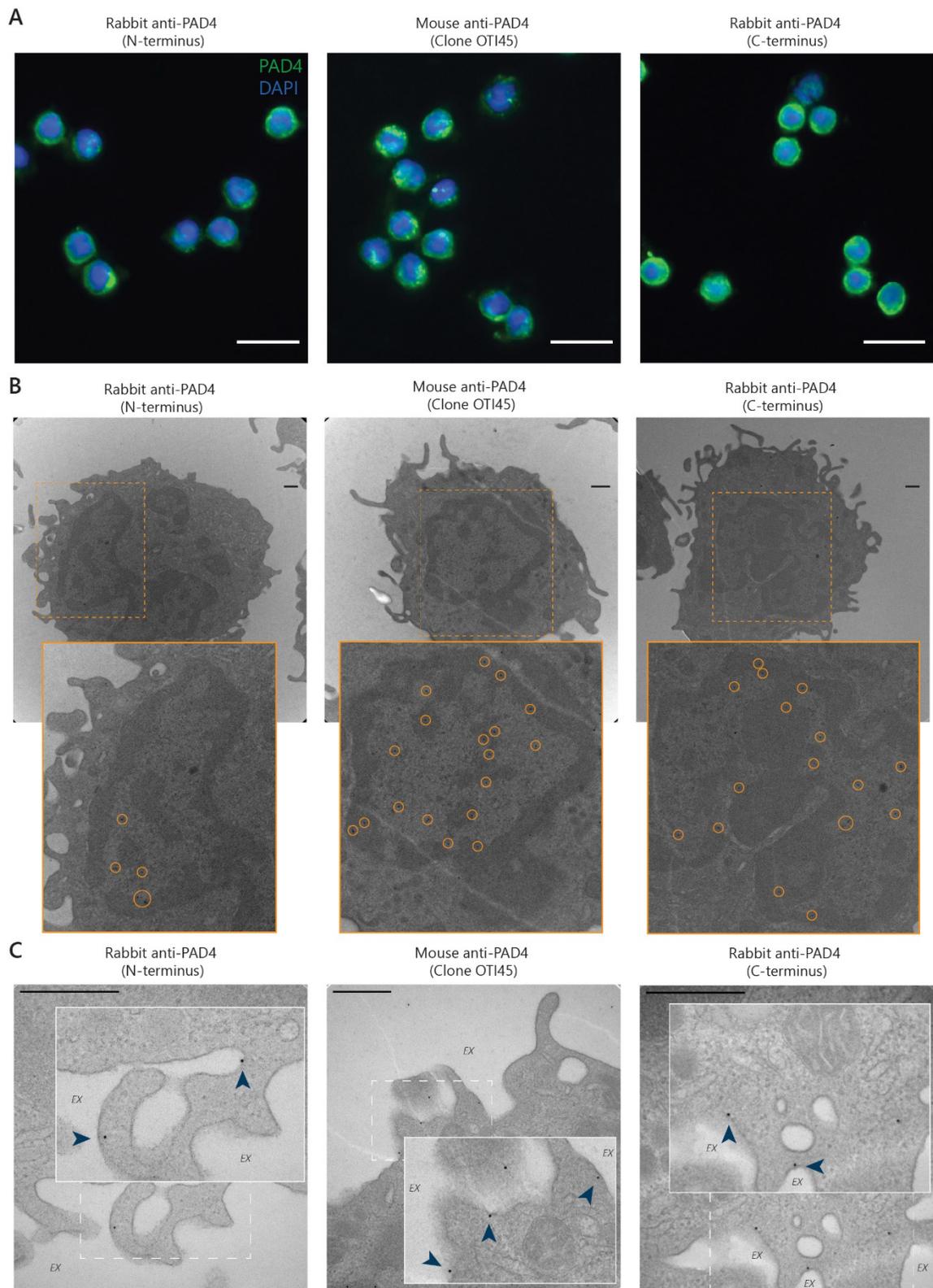


Fig. 2-2. Monocytes express PAD4 extracellularly. (A) Monocytes were isolated from peripheral blood of healthy donors. Cells were imaged using three different PAD4 antibodies (green;

N-terminal, clone OTI45, and C-terminal) and nuclei were counter stained with DAPI (blue). Scale bars represent 10 μm . **(B-C)** Negatively isolated monocytes were fixed and stained for PAD4. Representative images from transmission electron microscopy are shown for each antibody **(B)** Localizations of PAD4 in nuclear regions (zoomed insets) are shown in orange circles. **(C)**. Blue arrowheads on the zoomed-in insets indicate surface PAD4. Scale bars represent 500 nm. Extracellular space is indicated by 'EX.'

Since all three antibodies showed very similar localization via immunofluorescence and electron microscopy, we decided to use the C-terminal antibody for further studies. To determine whether PAD4 can exist on the outer surface of the monocyte membrane, unpermeabilized live monocytes were stained with fluorescently labeled anti-PAD4 C-terminal antibody. Flow cytometric analysis showed significant binding of the anti-PAD4 antibody to monocytes as compared to binding by the isotype control ($p < 0.0001$; **Figure 2-3A**, gating strategy shown in **Figure 2-4A**), indicating that PAD4 is indeed expressed on the cell surface. Interestingly, surface PAD4 levels between monocytes from healthy controls and RA patients were not significantly different (**Figure 2-3B**, **Figure 2-5A**). Analyzing monocyte subsets found in the peripheral blood revealed that while classical monocytes (CD14⁺ CD16⁻) were found in significant numbers in both healthy and RA individuals, their PAD4 expression was significantly lower than the other subsets (healthy: $p = 0.0017$ compared to CD14⁺CD16⁺ monocytes, $p = 0.0083$ compared to CD14⁻CD16⁺; RA: $p = 0.0076$ compared to CD14⁺CD16⁺ monocytes, $p = 0.0081$ compared to CD14⁻CD16⁺; **Figure 2-3C**). Surprisingly, no significant differences could be seen in surface PAD4 levels when comparing monocyte subsets in healthy individuals and RA patients.

Further analysis revealed high viability of monocytes (between 94 – 100%) in culture (**Figure 2-4**, **2-5**). Additionally, flow cytometry on monocytes incubated with

monocyte lysate (a source of extracellular PAD4) and HSG lysate (not a source of extracellular PAD4 – see **Figure 2-1**) does not show a difference in surface PAD4 levels, indicating that the source of cell surface PAD4 in monocytes is cell-intrinsic (**Figure 2-3C**). Taken together, the data strongly supports that both healthy and RA monocytes can display PAD4 on their plasma membranes.

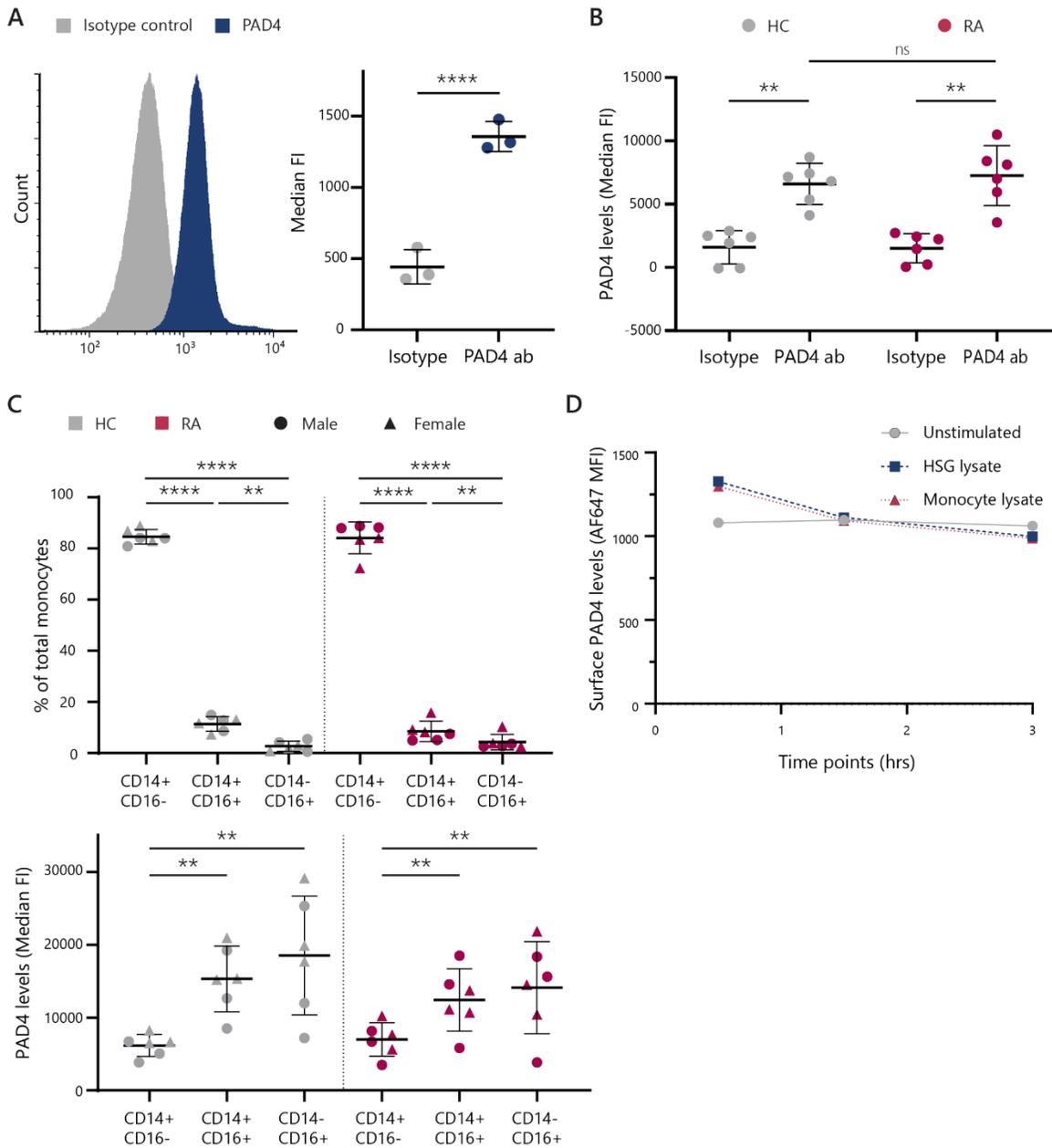


Fig. 2-3. Monocytes express PAD4 on their cell surface. (A) Monocytes were stained using AF647-conjugated IgG isotype control antibody or rabbit anti-PAD4 C-terminal antibody. Levels of surface staining for $n = 3$ healthy donors were analyzed using flow cytometry and a representative histogram shown. Median fluorescence intensity (FI) levels were compared using paired t-test. **(B-C)** Monocytes were isolated from frozen PBMCs of healthy controls (HC; $n = 6$) and RA patients ($n = 6$) and stained using AF647-conjugated isotype control or anti-PAD4 C-terminal antibody. Median FI levels were compared using paired t-tests. Error bars represent mean \pm standard deviation. * = $p < 0.05$, ** = $p \leq 0.01$, **** = $p \leq 0.0001$. **(D)** Monocytes isolated from peripheral blood of a healthy donor ($n = 1$) were incubated in HBSS and 10 mM HEPES supplemented with either monocyte or HSG cell lysates. Cell surface PAD4 expression was interrogated via flow cytometry using AF647-labeled anti-PAD4 C-terminal antibody after 0.5, 1.5, and 3 hours post-incubation with lysates.

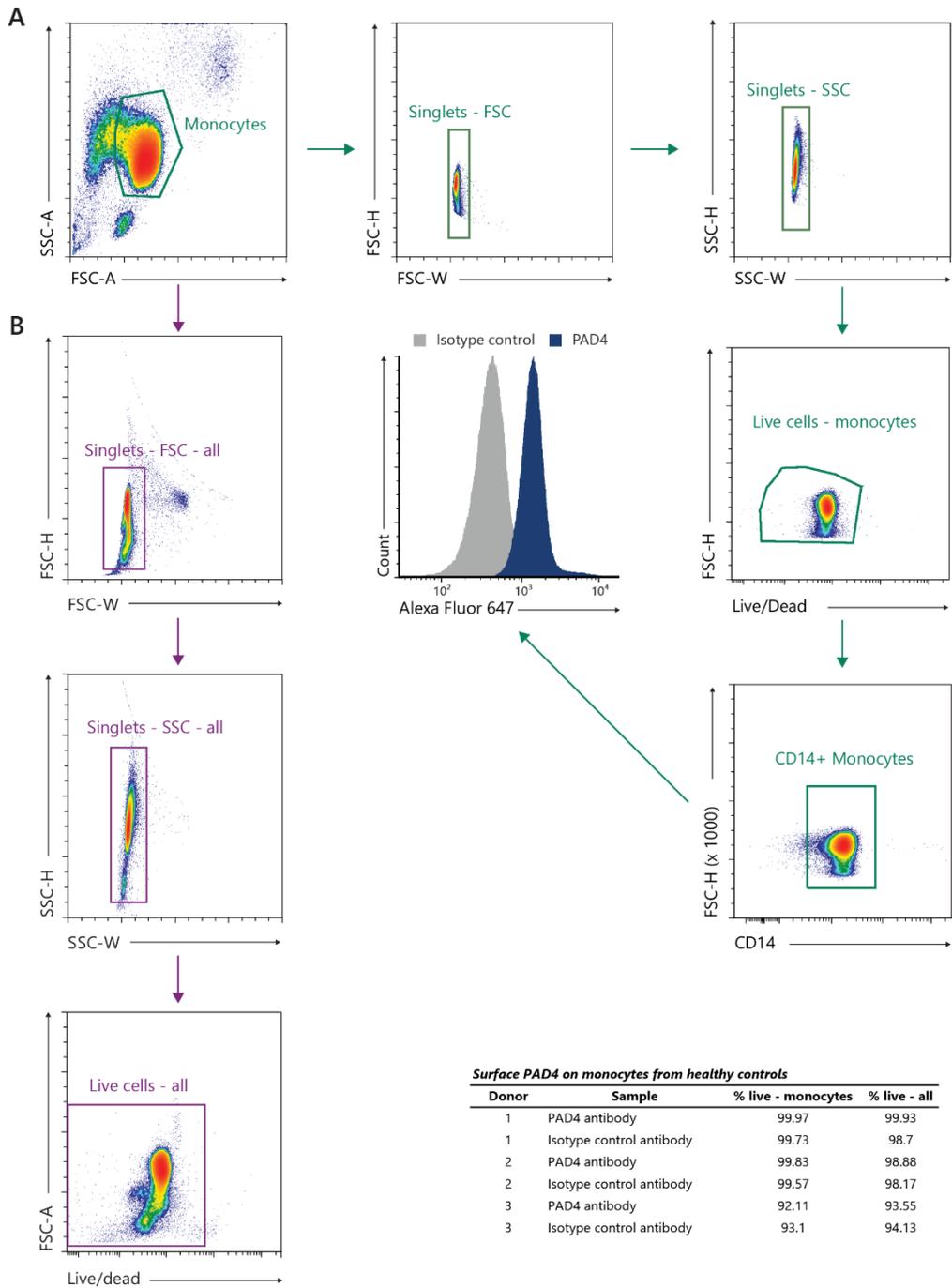


Fig. 2-4. Flow cytometry gating strategy. Monocytes were isolated from peripheral blood of healthy controls (n = 3) and stained for PAD4 using AF647-conjugated anti-PAD4 C-terminal antibody. Gating strategy for analyzing surface expression of PAD4 and viability of **(A)** monocytes (teal arrows) and **(B)** all cells (purple arrows) as well as a table with percent live cells shown.

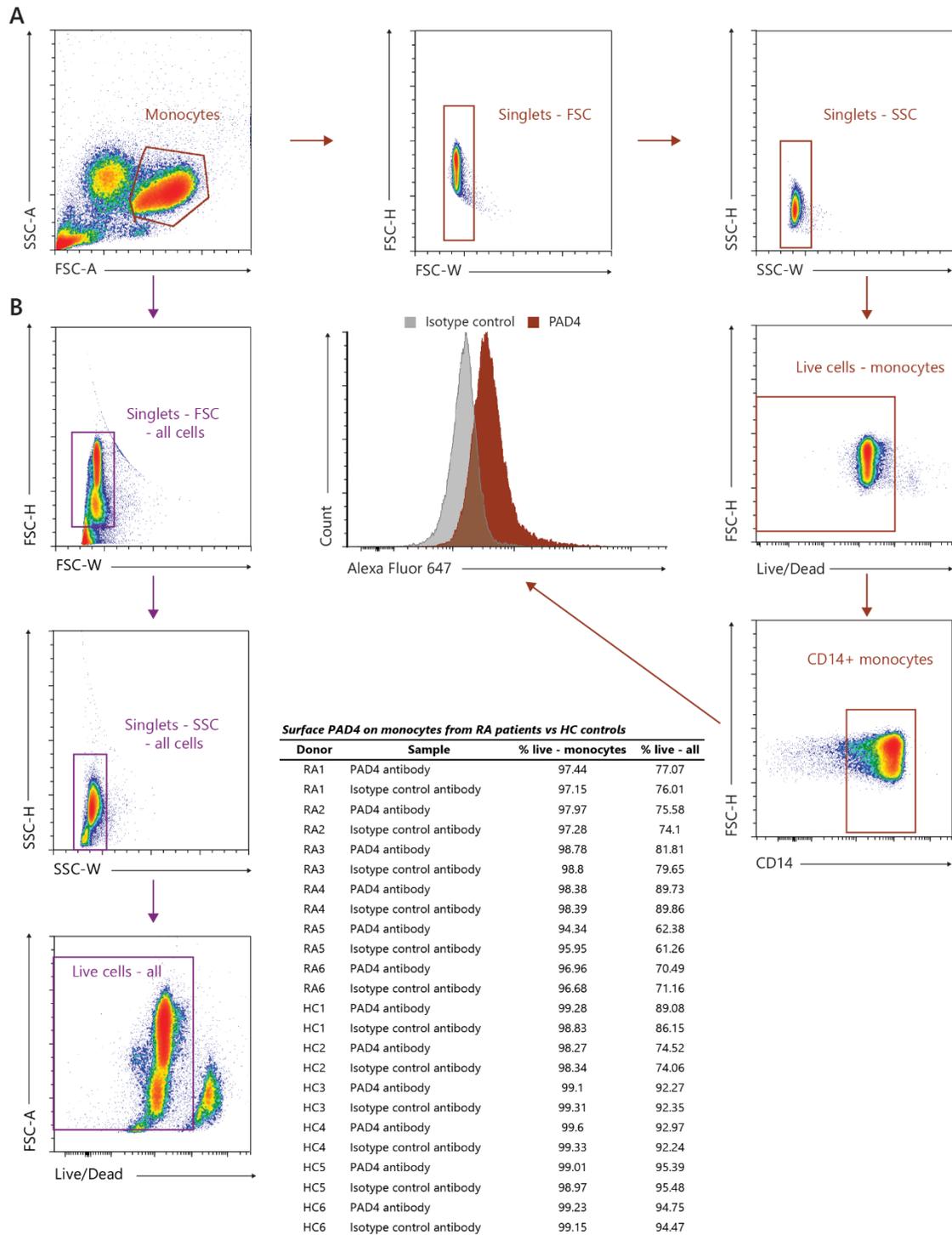


Fig. 2-5. Flow cytometry gating strategy for surface PAD4 levels on healthy controls and RA patients. Monocytes were isolated from peripheral blood of healthy controls (n = 6) and RA patients (n=6) and stained for surface expression of PAD4 using AF647-conjugated anti-PAD4 C-terminal antibody. Gating strategy for analyzing levels of surface PAD4 and viability

of **(A)** monocytes (red arrows) and **(B)** all cells (purple arrows) as well as a table with percent live cells shown.

2.3.2 Surface PAD4 is enzymatically active.

Since we found that PAD4 localizes to the plasma membranes of monocytes, we hypothesized that the monocyte surface may serve as an important source of extracellular citrullination in RA. Fibrinogen, a key component of the extracellular matrix, is commonly found to be citrullinated in the RA joint, and is a known target of ACPAs in its citrullinated form.^{170,171} Therefore, we assessed whether cell surface PAD4 could citrullinate extracellular fibrinogen *in vitro*. Monocytes isolated from peripheral blood of healthy controls were incubated in the presence of fibrinogen and increasing amounts of calcium (Ca^{2+}), an important cofactor for citrullination.⁹⁵ Blotting for citrullinated fibrinogen present in the cell-free supernatants showed a Ca^{2+} dose-dependence in citrullination levels (**Figure 2-6A**, lanes 1-5). Addition of a PAD4 inhibitor, GSK484,¹⁷² revealed a dose-dependent reduction in citrullination, indicating that citrullination of fibrinogen was PAD4-mediated (**Figure 2-6A**, lanes 6-8).

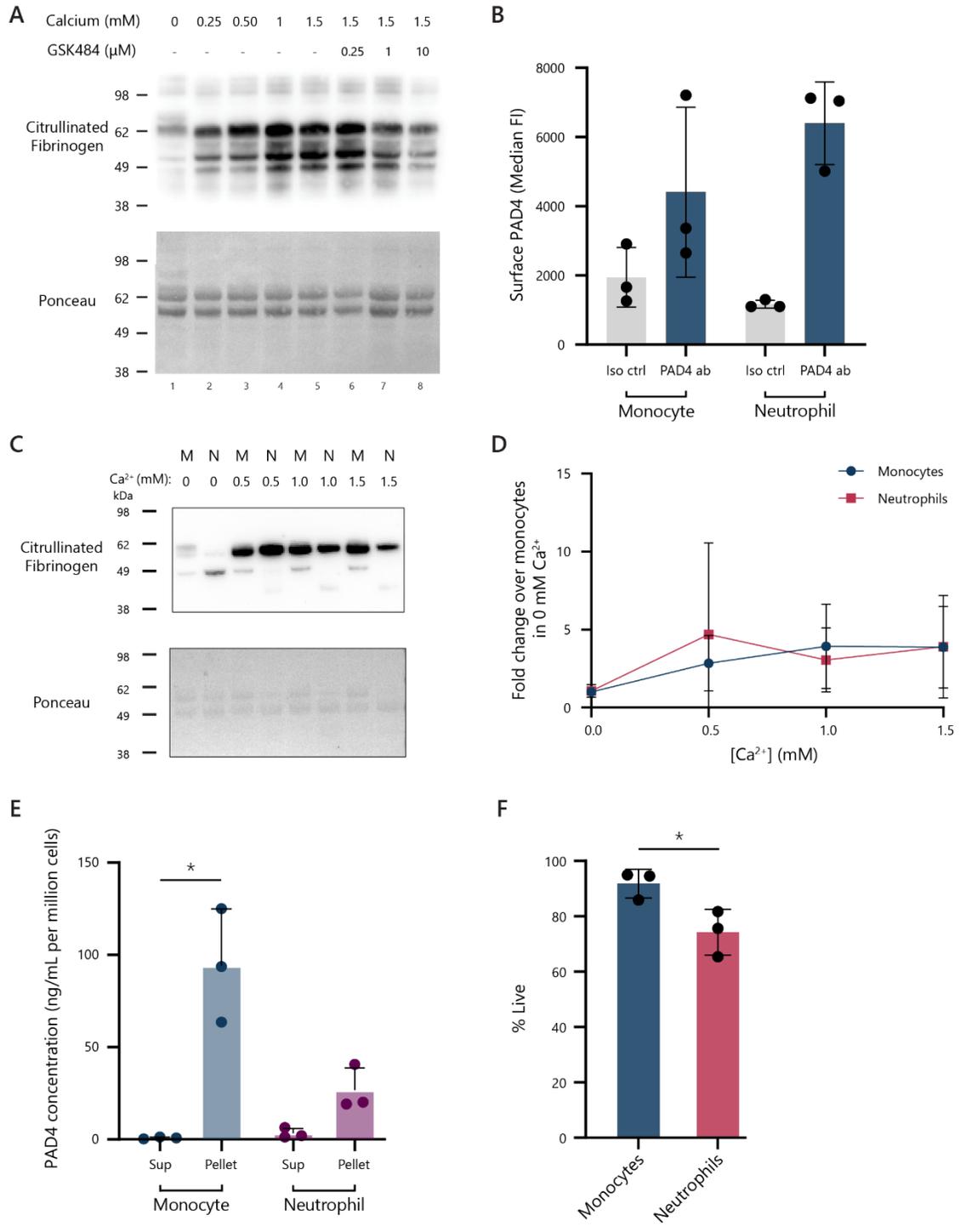


Fig. 2-6. PAD4 on the monocyte surface is catalytically active and its ability to citrullinate is comparable to neutrophils. (A) Monocytes isolated from peripheral blood of healthy donors were incubated at 37 °C for 3 hours in 0-1.5 mM Ca²⁺ (lanes 1-5) and 0.7 μ M of fibrinogen (*top panel*). Cells were also incubated with increasing doses of the PAD4-specific inhibitor, GSK484 (0.25 - 10 μ M; *lanes 6-8*) at 1.5 mM Ca²⁺. Denatured, cell-free

supernatants were resolved by SDS-PAGE and immunoblotted using an anti-citrullinated fibrinogen antibody. Ponceau staining was conducted to visualize protein loading (*bottom panel*). Data is representative of three independent experiments. **(B)** Neutrophils and PBMCs were isolated from peripheral blood of three healthy donors. Cells were subjected to flow cytometry after staining using either AF647-labeled C-terminal anti-PAD4 antibody or isotype control antibody. Surface PAD4 levels (median fluorescence intensity) shown. **(C)** Denatured, cell-free supernatants were immunoblotted for citrullination using anti-citrullinated fibrinogen antibody. Ponceau staining was used to visualize protein loading (*bottom panel*). M = monocyte, N = neutrophil. **(D)** Quantification of results from **(C)** for n = 3 donors shown. Fold change in citrullination of fibrinogen by monocytes and neutrophils was normalized to citrullination found in the condition of monocytes incubated in 0 mM Ca²⁺. Error bars represent mean ± standard deviation. **(E-F)** Monocytes (positively isolated) and neutrophils from peripheral blood of healthy controls (n = 3) were incubated in 5 mM Ca²⁺ for 3 hours at 37 °C. **(E)** PAD4 concentrations in supernatants and cell pellets were determined using commercially available PAD4 ELISA kit. **(F)** Viability of cells were assessed via flow cytometry post-incubation.

Because neutrophils are also major expressors of PAD4 in the RA joint, we compared the ability of PAD4 derived from monocytes and neutrophils to citrullinate extracellular fibrinogen. First, we determined the cell surface expression of PAD4 on monocytes and neutrophils derived from the same donor. Flow cytometry analysis of CD16+ neutrophils and CD14+ monocytes confirmed the presence of PAD4 on neutrophils and on monocytes (**Figure 2-6B**). To analyze the ability of both cell types to citrullinate extracellular fibrinogen, monocytes and neutrophils isolated from healthy donors were incubated with extracellular fibrinogen under increasing calcium concentrations for 3 hours at 37 °C. Immunoblotting revealed that both monocytes and neutrophils had comparable ability to citrullinate extracellular fibrinogen (**Figure 2-6C, D**).

To evaluate the possibility that fibrinogen citrullination was mediated by PAD4 released into the supernatant by dying cells, we incubated monocytes and neutrophils from healthy donors (n = 3) in 5 mM Ca²⁺ for 3 hours and determined the amount of PAD4

in paired supernatants and cell pellets. Our analysis revealed that the concentration of extracellular PAD4 was over 100-fold lower than that which remained cell-associated (0.82 ng/mL vs. 94.11 ng/mL per million monocytes, respectively; **Figure 2-6E**). While neutrophil lysates also contained more PAD4 than the cell-free supernatant, the difference was only eight-fold higher (26.69 ng/mL vs 3.264 ng/mL per million neutrophils, respectively; **Figure 2-6E**). Interestingly, the data revealed that monocytes contain about 3.5-fold more cell-associated PAD4 per million cells than neutrophils. Flow cytometric analysis also showed a high viability of monocytes after incubation (average = 91.8%) and a significantly lower viability of neutrophils (average = 74.3%; **Figure 2-6F**). This data suggests that PAD4 release into the extracellular environment by monocytes is negligible and further supports our conclusion that extracellular fibrinogen in our cellular assay was citrullinated by PAD4 present on the monocyte cell surface.

Given that citrullinated fibrinogen is a prominent target of ACPAs in RA,¹⁷¹ we interrogated whether PAD4 present on the surface of monocytes could generate citrullinated fibrinogen capable of being recognized by human ACPAs. Monocytes isolated from healthy controls were incubated with fibrinogen in the presence of Ca²⁺ to promote citrullination or with EDTA, a Ca²⁺ chelator, to inhibit citrullination. Recombinant human PAD4 (rhPAD4) was added to parallel fibrinogen-coated wells in the presence of Ca²⁺ as a positive control for citrullination. As a source of ACPAs, we used ACPA monoclonal antibodies (mAbs), which were produced from plasmablasts of RA patients and previously shown to recognize citrullinated fibrinogen.¹⁵⁶⁻¹⁵⁸ We found that ACPA mAbs bound equally well to citrullinated fibrinogen generated by incubation with monocytes in Ca²⁺ or

rhPAD4 protein ($p = 0.57$), and significantly better than to native fibrinogen incubated with monocytes in EDTA ($p < 0.0001$ compared to rhPAD4-citrullinated fibrinogen and $p = 0.0136$ compared to monocyte-citrullinated fibrinogen) (**Figure 2-7A**).

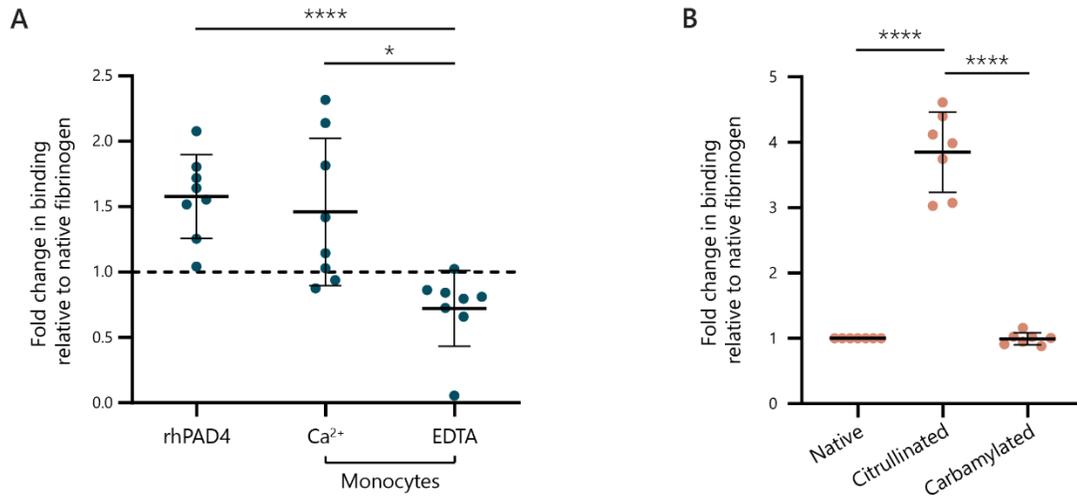


Fig. 2-7. Monocyte-citrullinated fibrinogen is recognized by patient autoantibodies. (A)

Native fibrinogen was coated onto a plate, blocked with BSA, and citrullinated by 20 ng rhPAD4, monocytes in 5 mM Ca²⁺ (Mon + Ca) or monocytes in 10 mM EDTA (Mon + EDTA) for 3 hours. Binding of ACPA mAbs ($n = 8$) with varying specificities against citrullinated fibrinogen were assessed using an ELISA. The fold change in binding of each antibody relative to its binding to native fibrinogen (dashed horizontal line) is shown. **(B)** Binding of ACPA mAbs ($n = 7$) to native, citrullinated, or carbamylated fibrinogen was quantified using ELISA. Error bars represent mean \pm standard deviation. Levels of binding were compared between conditions using paired t-tests. * = $p < 0.05$, **** = $p \leq 0.0001$.

As ACPAs are known to cross-react with carbamylated proteins, especially to homocitrulline residues (or carbamylated lysines),^{173,174} we analyzed the specificity of binding of the ACPA mAbs to citrullinated versus native or carbamylated fibrinogen. ACPA mAb binding to citrullinated fibrinogen was significantly higher than binding to native fibrinogen ($p < 0.0001$) and to carbamylated fibrinogen ($p < 0.0001$), indicating that the ACPA mAbs in our assay were indeed binding to fibrinogen citrullinated by PAD4 (**Figure**

2-7B). These observations demonstrate that PAD4 present on the monocyte surface is enzymatically active and capable of citrullinating extracellular antigens that can be targeted by ACPAs.

2.3.3. Surface PAD4 can citrullinate monocyte cell surface proteins

Since PAD4 was observed to be enzymatically active on the monocyte surface, we sought to understand whether it could citrullinate endogenous membrane proteins. Monocytes were isolated, incubated in the presence or absence of extracellular Ca^{2+} , and interrogated for the localization of citrullinated proteins via immunofluorescence. Background binding by secondary antibodies is shown in the top panel of **Figure 2-8A**. Images revealed that citrullinated proteins occurred as extranuclear punctate clusters after the addition of external Ca^{2+} (**Figure 2-8A**). While PAD4 can be seen throughout the monocyte, citrullinated proteins mainly localize to the outer margins of the cell, suggesting that the citrullinated proteins are being generated by the fraction of PAD4 present on the surface. Consistent with this observation, minimal colocalization of citrullinated proteins was observed with the intracellular protein, GAPDH (**Figure 2-8B**). Flow cytometry on unpermeabilized monocytes further confirmed the presence of citrullinated proteins on the cell surface and revealed a Ca^{2+} dose-dependent increase in surface citrullination in the presence of 0.5 mM Ca^{2+} ($p = 0.006$), 1.0 mM Ca^{2+} ($p = 0.007$), and 1.5 mM Ca^{2+} ($p = 0.0007$; **Figure 2-8C and D**; gating strategy and viability are shown in **Figure 2-9**). Taken together, these observations provide evidence that enzymatically active PAD4 on the monocyte cell surface is capable of citrullinating endogenous plasma membrane proteins.

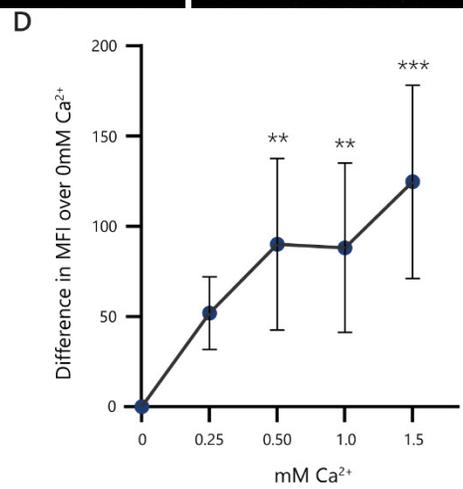
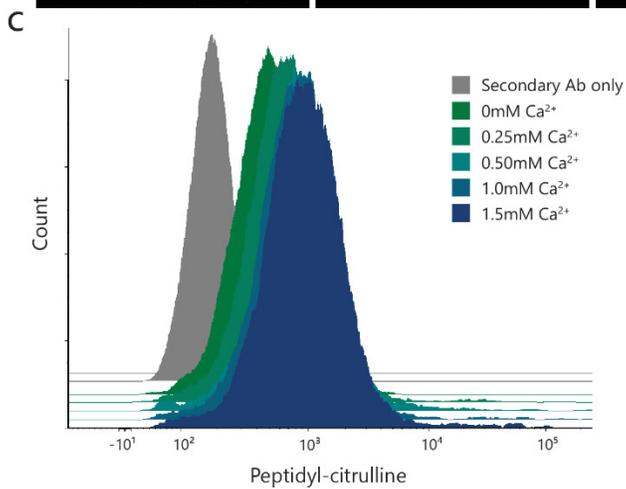
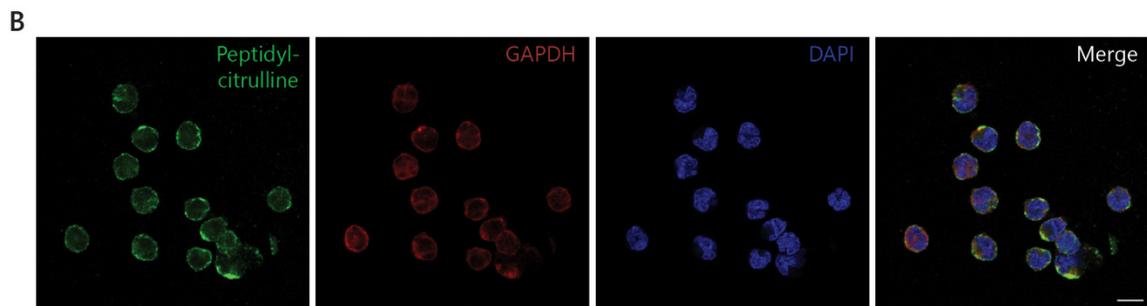
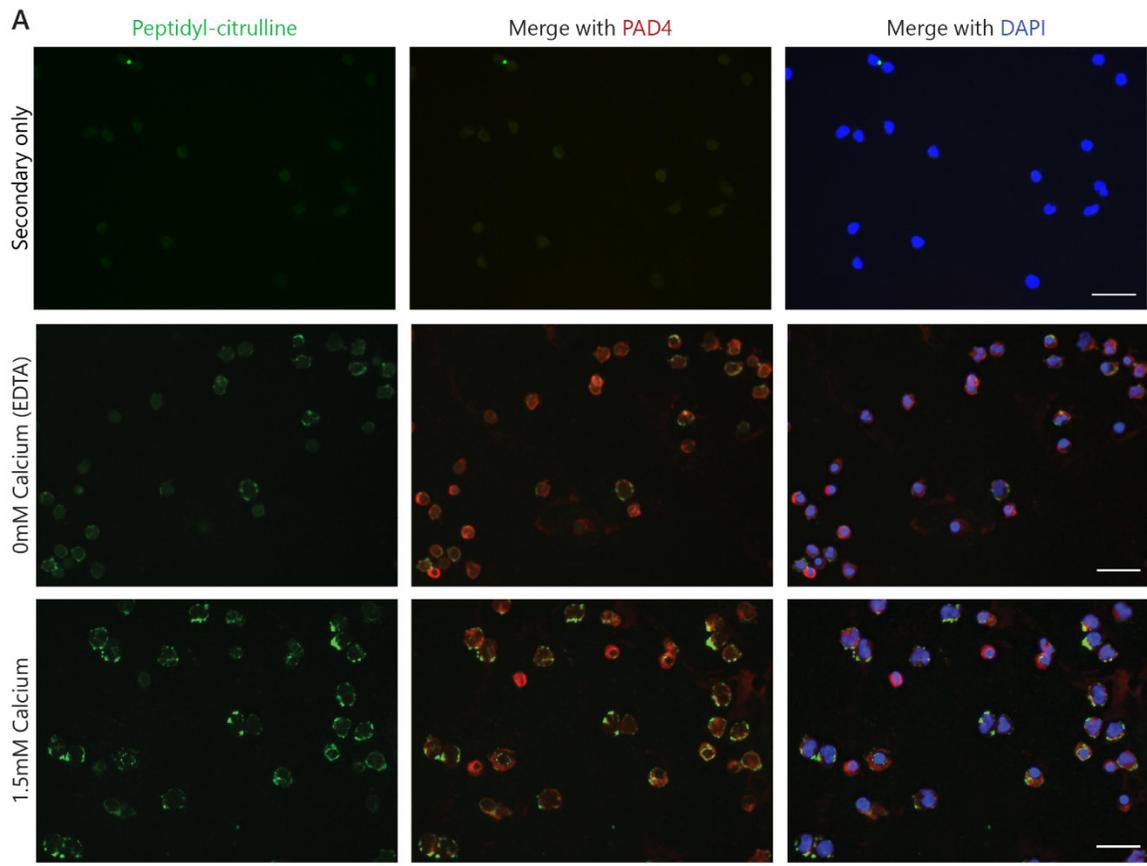
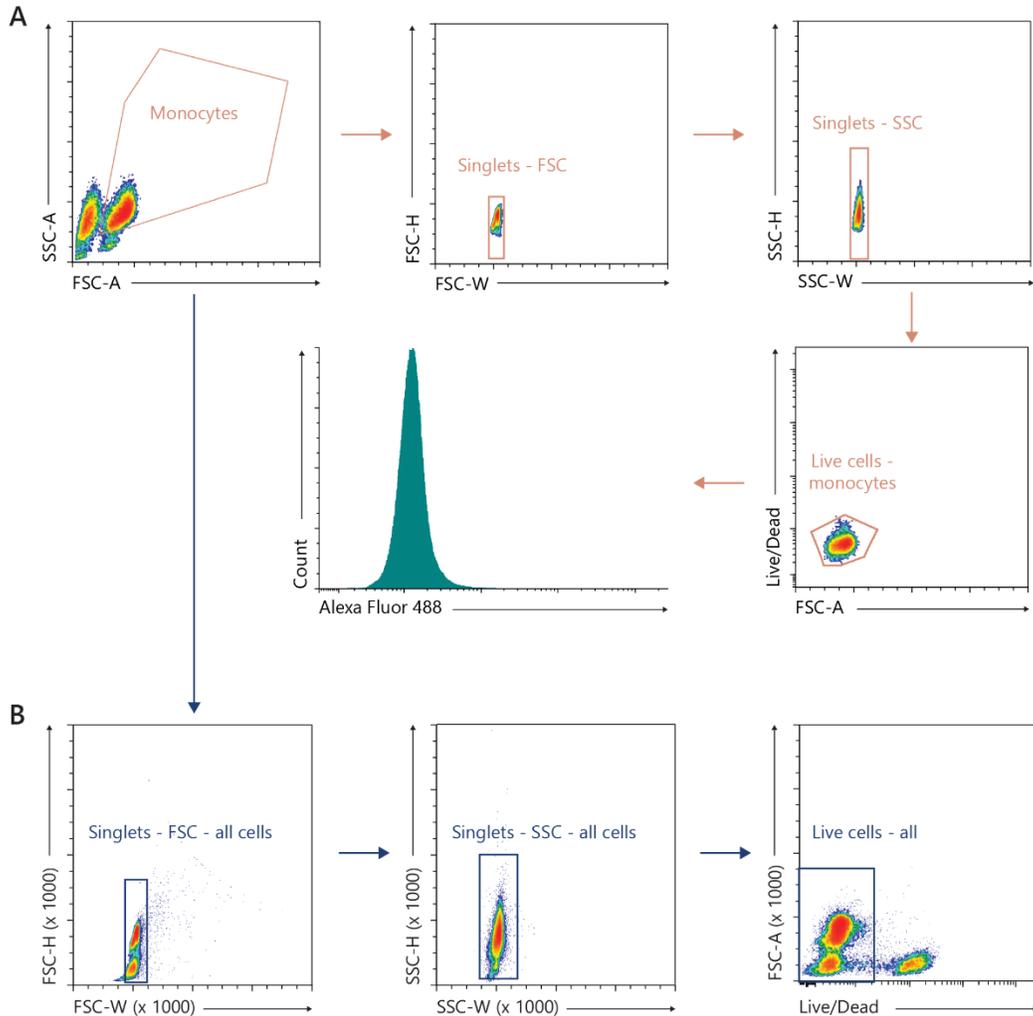


Fig. 2-8. Surface PAD4 citrullinates endogenous surface proteins on monocytes. (A) Representative images of monocytes treated with either EDTA or 1.5 mM Ca^{2+} for 3 hours at 37°C. Cells were stained for citrullinated proteins (green; clone F95) and PAD4 (red; anti-PAD4 N-terminal antibody), and nuclei were counter-stained with DAPI (blue). Background binding by secondary antibodies shown in the top panel. **(B)** Confocal imaging of monocytes demonstrates citrullinated proteins (green; clone F95) surrounding intracellular GAPDH (red). Nuclei were stained using DAPI (blue). Scale bars represent 10 μm . **(C and D)** Monocytes isolated from PBMCs of healthy donors were incubated at 37°C for 3 hours with increasing concentrations of Ca^{2+} (0 - 1.5 mM), and levels of surface citrullination were detected using indirect flow cytometry using an antibody against peptidyl-citrulline (clone F95). Cells were stained with secondary antibody alone as a negative control. **(C)** A histogram of a representative donor with the intensity of surface citrullination plotted on the x-axis, and **(D)** a plot of the mean fluorescence intensities (MFI) from 3 independent donors are shown. Error bars represent mean \pm standard deviation. ** = $p \leq 0.01$, *** = $p \leq 0.001$.



F95 binding to citrullinated proteins on monocyte surface

Donor	Sample	% live - monocytes	% live - all
1	0 mM Ca	98.12	83.79
1	0.25 mM Ca	99.09	95.49
1	0.5 mM Ca	98.34	92.7
1	1 mM Ca	98.84	96.69
1	1.5 mM Ca	98.68	96.92
2	0 mM Ca	93.3	65.83
2	0.25 mM Ca	92.17	78.42
2	0.5 mM Ca	94.24	77.22
2	1 mM Ca	87.41	73.77
2	1.5 mM Ca	90.15	84.64
3	0 mM Ca	93.57	78.19
3	0.25 mM Ca	79.5	54.29
3	0.5 mM Ca	89.12	76.35
3	1 mM Ca	89.71	73.78
3	1.5 mM Ca	88.17	75.38

Fig. 2-9. Flow cytometry gating strategy for surface citrullinated protein expression.

Monocytes were isolated from peripheral blood of healthy controls (n = 3) and incubated in increasing calcium levels. Cells were stained for citrullinated proteins using an antibody against peptidyl-citrulline (clone F95) and analyzed using flow cytometry. Gating strategy

for analyzing surface expression of citrullinated proteins and viability of **(A)** monocytes (orange arrows) and **(B)** all cells (blue arrows) as well as a table quantifying viability of cells shown.

2.3.4 Mass spectrometry reveals several novel citrullinated cell surface proteins including integrins.

To identify the citrullinated proteins present on the cell surface, monocytes from healthy donors were incubated with Ca^{2+} , and plasma membrane fractions were isolated and analyzed via mass spectrometry (see workflow in **Figure 2-10A**). The enrichment of proteins of the plasma membrane fractions was confirmed by the selective enrichment of Na^+/K^+ ATPase and the paucity of the cytosolic and nuclear markers GAPDH and lamin B1, respectively (**Figure 2-10B**). Importantly, in agreement with our earlier findings, PAD4 was detected in the plasma membrane fractions of all three donors by mass spectrometry. After filtering proteins based on the presence of citrulline residues, we found that 95.2% (337/354) of citrullinated proteins were shared by all three donors (**Figure 2-10C**), with 1.4% (5/354) shared only by donors 1 and 2, 1.7% (6/354) shared only by donors 2 and 3, and 0.6% (2/354) shared only by donors 1 and 3. Only 1.1% (4/354) of the citrullinated proteins were donor-specific (0.6% [2 proteins] in donor 1, 0.6% [2 proteins] in donor 2, and 0% [0 proteins] in donor 3). Analysis of the baseline citrullinated proteins present in monocytes incubated in 10 mM EDTA compared to those induced after incubation with 5 mM Ca^{2+} following isolation from peripheral blood of healthy donors revealed that 97.9% (334/341) of citrullinated proteins were shared between the two conditions (**Figure 2-10D**). In total, 0.9% (3/341 proteins) were citrullinated *de novo* by *ex vivo* calcium

incubation while 1.2% (4/341) of citrullinated proteins were lost. Further analysis of the citrullinated peptides showed 4% (19/475) were highly abundant following monocyte incubation with calcium while 2.5% (12/475) had an extremely diminished abundance after incubation in Ca^{2+} (**Figure 2-10E**). This data suggests that some proteins might be selectively citrullinated to a high degree following an increase in extracellular calcium levels.

Analysis of the number and identity of proteins from the plasma membrane fractions of monocytes found to be citrullinated and/or carbamylated, as detected by mass spectrometry, indicates that only 2.6% of proteins were identified to be carbamylated while 95.1% were citrullinated (**Figure 2-10F**). Very few (2.3%) were found to be both citrullinated and carbamylated. Furthermore, analysis of peptide abundances in all three donors showed that the plasma membrane fractions of monocytes incubated in 5 mM Ca^{2+} contained, on average, ten-fold more citrullinated peptides than carbamylated peptides (**Figure 2-10G**). Our data suggests that compared to citrullinated proteins, carbamylated proteins are unlikely to be present on the cell surface. Therefore, monocytes are more likely to be targeted by ACPAs that recognize citrullination rather than those that cross-react to carbamylation or by anti-carbamylated protein antibodies.

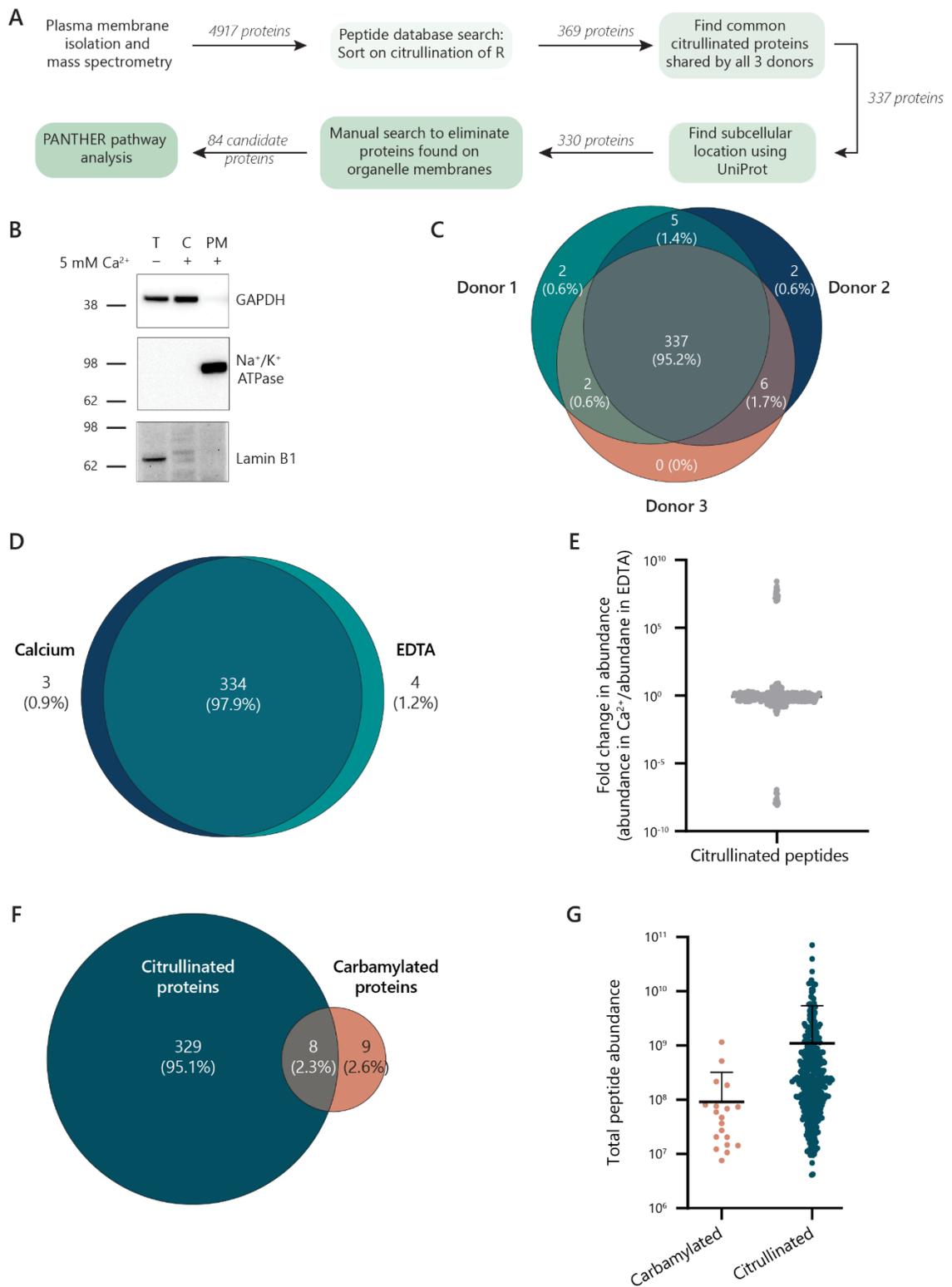


Fig. 2-10. Mass spectrometry identifies several citrullinated cell surface proteins. Monocytes isolated from healthy controls ($n = 3$) were incubated in either 5 mM Ca²⁺ or 10 mM EDTA.

Plasma membrane fractions were isolated and analyzed by mass spectrometry. **(A)** Workflow of mass spectrometry results analysis. The resulting database was sorted on proteins containing deiminated arginines with a greater than 75% probability score. Citrullinated proteins shared by all three donors that had a primary subcellular location on the plasma membrane were then sorted according to their average abundance. **(B)** Western blot of GAPDH (cytosolic marker), Na⁺/K⁺-ATPase (plasma membrane marker), and lamin B1 (nuclear marker) on monocyte cellular fractions. *T* = total cell lysate, *C* = cytosol, *PM* = plasma membrane. **(C)** Venn diagram of citrullinated proteins found in the plasma membrane fractions of healthy donor monocytes. **(D-E)** Monocytes were isolated from peripheral blood of healthy controls (*n* = 3) and incubated in either 5 mM Ca²⁺ or 10 mM EDTA for three hours. Following plasma membrane isolation, citrullinated proteins were identified using mass spectrometry. **(D)** Venn diagram of the number of proteins identified to be citrullinated in either the calcium condition, the EDTA condition, or shared by both. **(E)** Fold change in abundance of citrullinated peptides in the calcium condition relative to their abundances in the EDTA condition. **(F)** Venn diagram of the number of citrullinated and carbamylated proteins found on the cell surface of monocytes (from *n* = 3 healthy donors) incubated in 5 mM Ca²⁺ using mass spectrometry. **(G)** Total abundances (sum of peptide abundance in all three donors) of carbamylated and citrullinated peptides found in the plasma membrane fractions of monocytes incubated in 5 mM Ca²⁺.

The reported subcellular localization of citrullinated proteins common to all three samples in the Ca²⁺ condition were then analyzed using the UniProt database,¹⁶² and the abundances of the proteins predominantly found on the plasma membrane were averaged across the three samples. An analysis of the monocyte surface citrullinome revealed numerous previously undescribed PAD4 substrates including several MHC class I and class II proteins, cell adhesion proteins, and proteins involved in immune regulation such as complement receptors and leukocyte immunoglobulin-like receptors. To understand whether cell surface PAD4 preferentially citrullinates certain functional classes of membrane proteins, we compared the list of citrullinated plasma membrane proteins to all the plasma membrane proteins in the human genome using the same criteria outlined

in **Figure 2-10A**. Statistical overrepresentation analyses using the PANTHER classification system demonstrated significant enrichment of pathways and proteins in four annotation datasets tested: GO Biological Processes, Molecular Function, Protein class, and Pathways.^{163,164} Of the top twenty statistically significant biological processes, cell adhesion mechanisms comprised 20% (4/20) and appeared to be significantly enriched in the citrullinated membrane protein list (**Figure 2-11A, Table 1**). Analysis of the molecular functions of the proteins in our citrullinated protein dataset identified 'protein-containing complex binding' as preferentially enriched (**Figure 2-11B, Table 2**). Interestingly, while proteins belonging to the transmembrane signal receptor and G-protein coupled receptor classes appeared less frequently in our list of citrullinated membrane proteins, actin or actin-binding cytoskeletal proteins were significantly enriched (**Figure 2-11C, Table 3**). Additionally, although not significant, integrins and MHC protein classes had greater than five-fold increase in enrichment in the citrullinated fraction. Furthermore, the integrin signaling pathway was the only one to reach significance and was found to be overrepresented at nearly ten-fold in our dataset (**Figure 2-11D, Table 4**), implicating integrins as important substrates of monocyte surface PAD4.

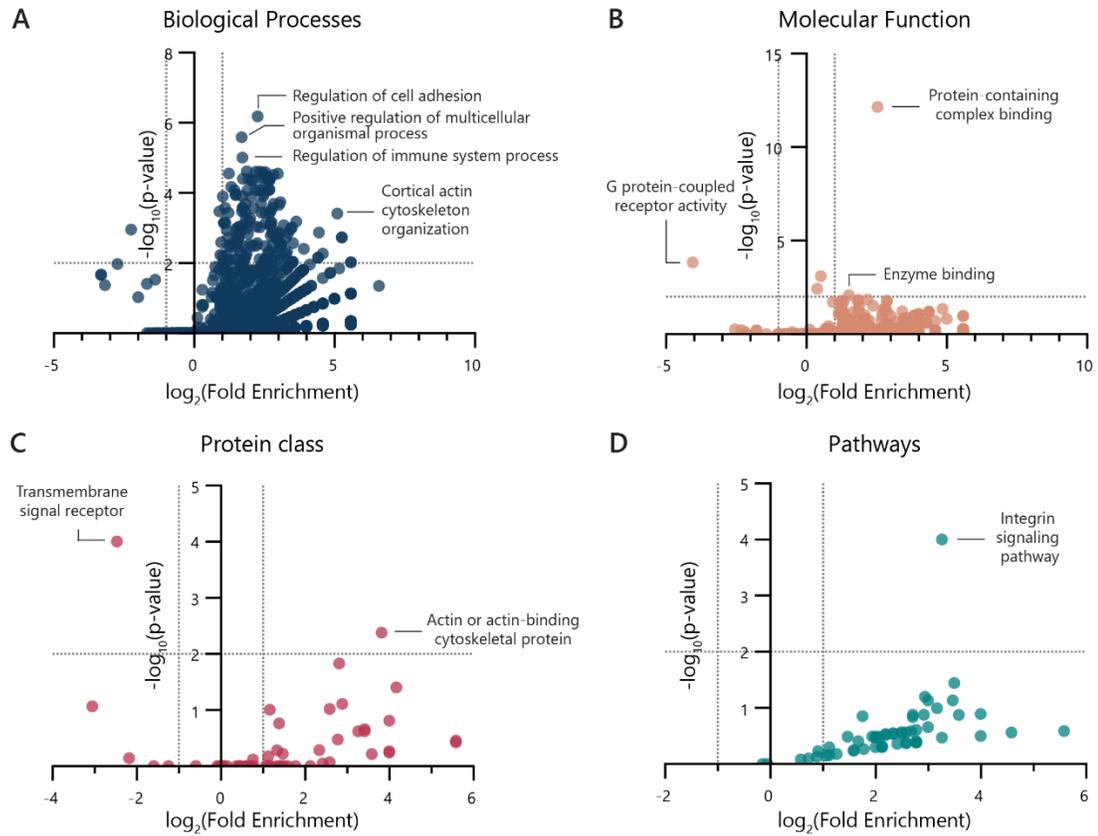


Fig. 2-11. Statistical enrichment of citrullinated proteins in various protein classifications. (A – D) Statistical overrepresentation analyses of citrullinated membrane proteins compared to all plasma membrane proteins from the human proteome using PANTHER analysis software. \log_2 of fold enrichment versus $-\log_{10}$ of false-discovery rate corrected p-values shown. Thresholds for significance were set at $p < 0.01$ and thresholds for fold enrichment were set at ± 2 . Factors with a fold enrichment value greater or less than 0 are shown on the plots.

Table 1: Biological processes enriched in citrullinated plasma membrane proteins identified via mass spectrometry

PANTHER GO Biological Process	Reference List	PM citrullinated proteins	Expected	Over/under	$-\log_{10}$ (FDR)	\log_2 (Fold enrichment)
Regulation of cell adhesion (GO:0030155)	250	25	5.23	+	6.18	2.26
Positive regulation of multicellular organismal process (GO:0051240)	490	33	10.26	+	5.59	1.69
Regulation of immune system process (GO:0002682)	437	30	9.15	+	5.01	1.71
Positive regulation of immune system process (GO:0002684)	326	25	6.83	+	4.62	1.87
Regulation of cell-cell adhesion (GO:0022407)	172	18	3.6	+	4.61	2.32
Regulation of leukocyte activation (GO:0002694)	210	20	4.4	+	4.60	2.19
Positive regulation of leukocyte activation (GO:0002696)	141	17	2.95	+	4.59	2.53
Positive regulation of cell activation (GO:0050867)	151	17	3.16	+	4.58	2.43
Regulation of multicellular organismal process (GO:0051239)	834	41	17.46	+	4.55	1.23
Regulation of T cell proliferation (GO:0042129)	79	13	1.65	+	4.55	2.97
Regulation of leukocyte mediated immunity (GO:0002703)	95	14	1.99	+	4.51	2.82
Immune effector process (GO:0002252)	118	15	2.47	+	4.46	2.60
Regulation of leukocyte cell-cell adhesion (GO:1903037)	137	16	2.87	+	4.43	2.48
Defense response (GO:0006952)	360	25	7.54	+	4.40	1.73
Regulation of immune response (GO:0050776)	275	22	5.76	+	4.39	1.93
Negative regulation of immune system process (GO:0002683)	143	16	2.99	+	4.39	2.42
Positive regulation of leukocyte cell-cell adhesion (GO:1903039)	106	14	2.22	+	4.39	2.66
Regulation of cell activation (GO:0050865)	230	20	4.82	+	4.37	2.05
Regulation of mononuclear cell proliferation (GO:0032944)	105	14	2.2	+	4.37	2.67
Regulation of lymphocyte proliferation (GO:0050670)	104	14	2.18	+	4.36	2.68

PM = plasma membrane

Table 2: Molecular functions enriched in citrullinated plasma membrane proteins identified via mass spectrometry

PANTHER GO Molecular Function	Reference List	PM citrullinated proteins	Expected	Over/under	$-\log_{10}(\text{FDR})$	$\log_2(\text{Fold enrichment})$
Protein-containing complex binding (GO:0044877)	265	32	5.55	+	12.14	2.53
G protein-coupled receptor activity (GO:0004930)	835	1	17.48	-	3.83	-4.06
Protein binding (GO:0005515)	2430	72	50.88	+	3.09	0.51
Binding (GO:0005488)	2839	77	59.44	+	2.42	0.38
Enzyme binding (GO:0019899)	335	20	7.01	+	2.07	1.51
Cytoskeletal protein binding (GO:0008092)	117	11	2.45	+	1.84	2.17
Cell adhesion molecule binding (GO:0050839)	186	14	3.89	+	1.83	1.84
Actin binding (GO:0003779)	46	7	0.96	+	1.79	2.86
Signaling receptor binding (GO:0005102)	468	23	9.8	+	1.78	1.23
Enzyme regulator activity (GO:0030234)	176	13	3.68	+	1.76	1.82

PM = plasma membrane

Table 3: Protein class enriched in citrullinated plasma membrane proteins identified via mass spectrometry

PANTHER Protein Class	Reference List	PM citrullinated proteins	Expected	Over/under	$-\log_{10}(\text{FDR})$	$\log_2(\text{Fold enrichment})$
Transmembrane signal receptor (PC00197)	1046	4	21.9	-	4.00	-2.47
Actin or actin-binding cytoskeletal protein (PC00041)	17	5	0.36	+	2.38	3.81
Cytoskeletal protein (PC00085)	41	6	0.86	+	1.83	2.81
Non-motor actin binding protein (PC00165)	8	3	0.17	+	1.40	4.16
Integrin (PC00126)	26	4	0.54	+	1.11	2.88
G-protein coupled receptor (PC00021)	386	1	8.08	-	1.06	-3.06
Major histocompatibility complex protein (PC00149)	32	4	0.67	+	1.02	2.58
Defense/immunity protein (PC00090)	300	14	6.28	+	1.00	1.16
Kinase modulator (PC00140)	6	2	0.13	+	0.81	3.99
Protein-binding activity modulator (PC00095)	147	8	3.08	+	0.76	1.38

PM = plasma membrane

Table 4: Pathways enriched in citrullinated plasma membrane proteins identified via mass spectrometry

PANTHER Pathways	Reference List	PM citrullinated proteins	Expected	Over/under	$-\log_{10}(\text{FDR})$	$\log_2(\text{Fold enrichment})$
Integrin signaling pathway (P00034)	45	9	0.94	+	4.00	3.26
Dopamine receptor mediated signaling pathway (P05912)	17	4	0.36	+	1.44	3.49
Endothelin signaling pathway (P00019)	25	4	0.52	+	1.20	2.93
B cell activation (P00010)	13	3	0.27	+	1.13	3.46
T cell activation (P00053)	24	4	0.5	+	1.13	2.99
Beta3 adrenergic receptor signaling pathway (P04379)	16	3	0.33	+	0.99	3.16
p53 pathway feedback loops 2 (P04398)	6	2	0.13	+	0.89	3.99
5HT4 type receptor mediated signaling pathway (P04376)	19	3	0.4	+	0.88	2.91
Beta1 adrenergic receptor signaling pathway (P04377)	22	3	0.46	+	0.88	2.70
VEGF signaling pathway (P00056)	8	2	0.17	+	0.87	3.58

PM = plasma membrane

Among the multiple integrins identified as citrullinated, integrin alpha-M (*ITGAM* or CD11b) and integrin beta 2 (*ITGB* or CD18) were found to be the most abundant (a list of the twenty most abundant surface citrullinated proteins are shown in **Table 5**). CD11b and CD18 together comprise the Mac-1 complex, which is responsible for monocyte extravasation, complement binding, and monocyte activation.¹⁷⁵ When the citrullination sites (**Table 6**) were mapped onto the CD11b protein structures predicted by AlphaFold (an artificial intelligence-driven protein structure prediction algorithm), it was revealed that the single citrullination site in CD18 was found to lie in the hybrid domain of the protein that mediates the transition from the low affinity to the high affinity conformation of Mac-1 (**Figure 2-12A**).¹⁷⁶ Interestingly, three of the four citrullination sites in CD11b were found to lie in the ligand-binding domain, or the I-domain, of the protein (**Figure 2-12B**).¹⁷⁷ In order to understand whether structural differences exist between the native and citrullinated versions of the Mac-1 subunits, we calculated the root-mean-squared deviation (RMSD) score between the native and citrullinated predicted models of CD18 and CD11b. Perfect structural alignment is indicated by an RMSD value of 0 Å with values ≤ 2 Å generally accepted to indicate highly similar structures.^{178,179} Comparing the structures of native and citrullinated full-length CD18 resulted in a relatively low RMSD score of 3.983 Å (**Figure 2-12C**), while modeling the I-domain of CD11b revealed moderate structural differences induced by citrullination, with an RMSD value of 8.068 Å (**Figure 2-12D**).

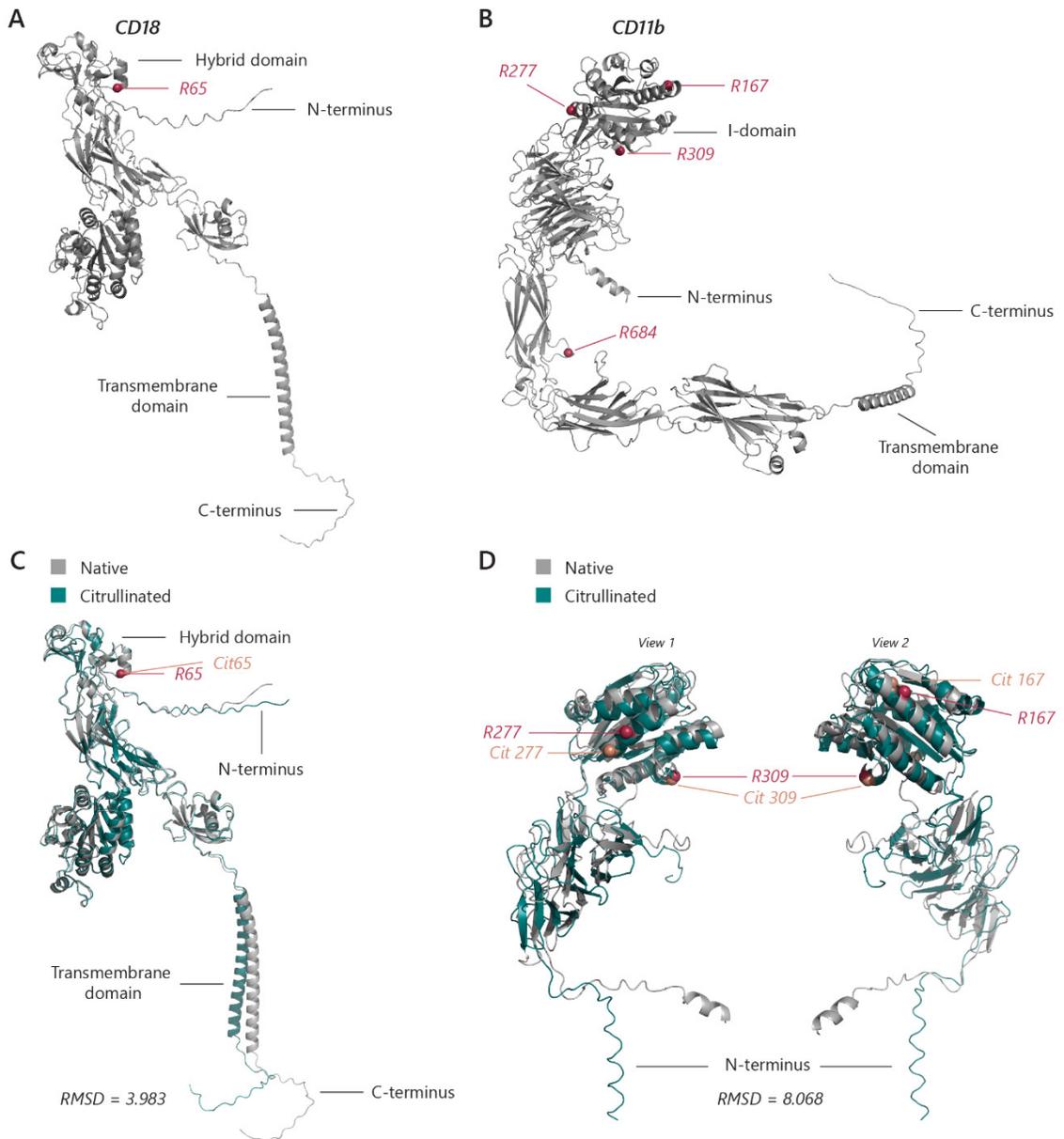


Fig. 2-12. Mass spectrometry reveals several citrullination sites on CD11b and CD18. Sites of citrullinated arginines (in red) identified by mass spectrometry on full-length CD18 (**A**) and CD11b (**B**) structures predicted by AlphaFold. (**C and D**) Comparison of AlphaFold-predicted native (grey) versus citrullinated (teal) full-length CD18 (**C**) and amino acids 1-400, encompassing the I-domain, of CD11b (**D**). Sites of citrullination are marked in orange while its companion native arginine residue site is marked in red on the native structure.

Table 5: Top hits from list of citrullinated plasma membrane proteins – mass spectrometry

UniProt Entry	Protein name	Sample 1 abundance	Sample 2 abundance	Sample 3 abundance	Avg abundance	Citrullination sites [R#(%)]
A0A494C0X7	Integrin beta 2 (CD18)	1.96E+11	3.25E+11	2.40E+11	2.54E+11	Deamidated [R65(100)]
P08575	Receptor-type tyrosine-protein phosphatase C (CD45)	1.29E+11	2.49E+11	2.31E+11	2.03E+11	Deamidated [R621(100), R729(100), R935(99.4), R1208(100)]
P11215	Integrin alpha-M (CD11b)	1.50E+11	2.52E+11	1.87E+11	1.96E+11	Deamidated [R167(99.1), R277(100), R309(100), R684(100)]
Q9Y490	Talin-1	2.45E+10	1.46E+11	7.70E+10	8.24E+10	Deamidated [R35(100), R606(100)]
P46940	Ras GTPase-activating-like protein IQGAP1	3.08E+10	3.32E+10	7.01E+10	4.47E+10	Deamidated [R14(100)]
P20702	Integrin alpha-X (CD11c)	3.08E+10	5.10E+10	4.19E+10	4.12E+10	Deamidated [R616(100)]
O75923	Dysferlin	4.42E+10	5.21E+10	1.57E+10	3.73E+10	Deamidated [R919(100), R1737(100), R1768(100), R2020(100)]
P20701	Integrin alpha-L (CD11a)	2.13E+10	4.73E+10	4.10E+10	3.65E+10	Deamidated [R755(100)]
Q53Z42	HLA class I antigen (HLA-A)	2.41E+10	4.74E+10	3.30E+10	3.49E+10	Deamidated [R89(100), R258(100)]
P17927	Complement receptor type 1 (CD35)	1.88E+10	4.97E+10	3.50E+10	3.45E+10	Deamidated [R530(100) or R980(100), R1957(100)]
Q07954	Prolow-density lipoprotein receptor-related protein 1 (CD91)	2.96E+10	4.24E+10	2.08E+10	3.09E+10	Deamidated [R3274(100), R3573(100)]
Q10588	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2 (CD157)	1.48E+10	4.02E+10	2.77E+10	2.76E+10	Deamidated [R124(98.6)]
Q9NZM1	Myoferlin	3.36E+10	2.89E+10	1.00E+10	2.42E+10	Deamidated [R333(100)]
P01911	HLA class II histocompatibility antigen, DRB1 beta chain (HLA-DRB1)	1.83E+10	3.59E+10	1.44E+10	2.29E+10	Deamidated [R84(100)]

UniProt Entry	Protein name	Sample 1 abundance	Sample 2 abundance	Sample 3 abundance	Avg abundance	Citrullination sites [R#(%)]
Q86UX7	Fermitin family homolog 3	1.48E+10	3.52E+10	1.32E+10	2.11E+10	Deamidated [R200(100)]
B011T2	Unconventional myosin-Ig	1.73E+10	3.97E+10	4.70E+09	2.06E+10	Deamidated [R74(100), R996(100)]
P32942	Intercellular adhesion molecule 3 (CD50)	1.15E+10	2.16E+10	1.74E+10	1.68E+10	Deamidated [R163(100), R545(100)]
P48960	Adhesion G protein-coupled receptor E5 (CD97)	1.61E+10	2.00E+10	9.48E+09	1.52E+10	Deamidated [R478(100)]
P09769	Tyrosine-protein kinase Fgr	1.29E+10	2.04E+10	6.54E+09	1.33E+10	Deamidated [R268(100)]
P43405	Tyrosine-protein kinase SYK	1.43E+10	1.92E+10	5.39E+09	1.30E+10	Deamidated [R197(100), R359(100), R273(100)]

Table 6: Citrullinated peptides in CD18 and CD11b

Annotated sequence	Average abundance	Citrullination sites [R#(%)]
Integrin beta-2 (CD18)		
[R].CDTRPQLLMR.[G]	5.28E+07	R65(100)
Integrin alpha-M (CD11b)		
[R].GCPQEDSDIAFLIDGSGSIIPHDFRR.[M]		R167(99.1)
[K].ILVVITDGEKFGDPLGYEDVIPEADREGVIR.[Y]	2.20E+07	R277(100)
[K].SRQELNTIASKPPRDHVFQVNNFEALK.[T]	5.17E+07	R309(100)
[R].LREGQIQSVVTYDLALDSGRPHSR.[A]	1.65E+07	R684(99.3)
[R].EGQIQSVVTYDLALDSGRPHSR.[A]	2.07E+07	R684(100)

*R = citrullinated arginine

2.3.5 Citrullinated Mac-1 is a novel cell surface autoantigen in RA.

Given that the CD11b subunit contained the majority of the citrullination sites in the Mac-1 complex and was predicted to undergo the greatest structural changes following citrullination, we determined whether ACPAs from RA patients recognized citrullinated CD11b. Native and citrullinated CD11b were *in vitro* transcribed/translated and were immunoprecipitated using sera from a small convenience cohort consisting of healthy controls (n = 18), ACPA⁻ RA patients (n = 18), and ACPA⁺ RA patients (n = 18). Densitometry analysis revealed that 61% (11/18) of ACPA⁺ RA patients had antibodies against the citrullinated version of CD11b, while none had reactivity to native CD11b (representative radiographic images shown in **Figure 2-13A; Figure 2-13B**). Recognition of citrullinated CD11b by ACPA⁺ patients was significantly higher than that of native CD11b ($p = 0.0004$) and significantly higher compared to healthy control ($p = 0.0021$) and ACPA⁻ patient sera ($p \leq 0.0001$). The presence of anti-citrullinated CD11b antibodies were

significantly associated with ACPA⁺ RA ($p = 0.0009$ and $p = 0.001$ compared to healthy control and ACPA⁻ RA populations, respectively).

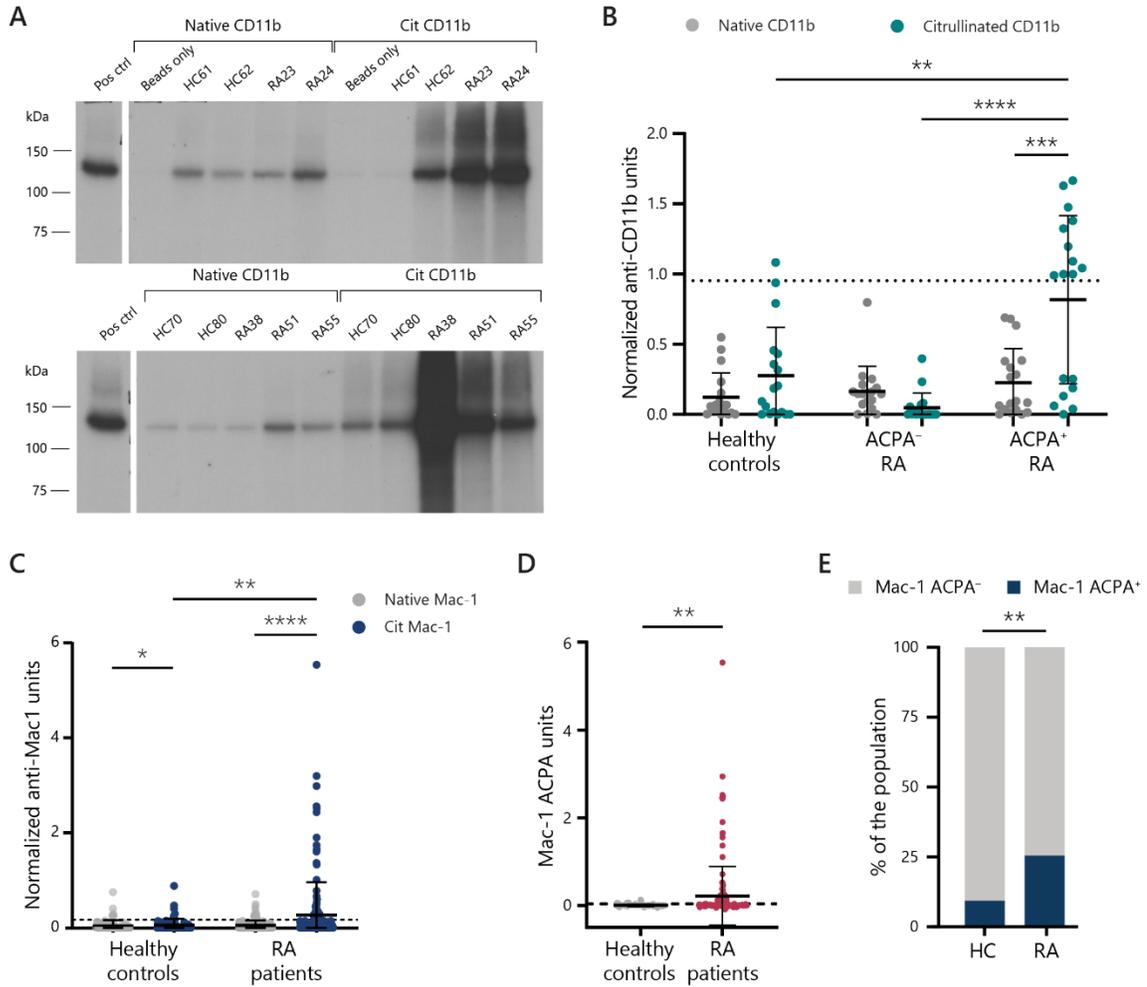


Fig. 2-13. CD11b is a novel cell surface autoantigen. CD11b protein created using IVTT (**A, B**) or recombinant Mac-1 (**C**) were citrullinated in 5 mM Ca²⁺ using rhPAD4. Human sera from healthy controls (n = 18 in **A and B** and n = 75 in **C**) and RA patients (n = 18 ACPA⁻ RA patients, n = 18 ACPA⁺ RA patients in **A and B** and n = 165 patients in **C**) were used to immunoprecipitate either native or citrullinated protein. (**A**) Representative radiographs showing IVTT-immunoprecipitation of native and citrullinated CD11b by healthy control (n = 18) and RA patient sera (n = 18 ACPA⁺ and n = 18 ACPA⁻). Amount of antibody-bound protein was quantified and levels of anti-CD11b (**B**) or anti-Mac-1 (**C**) (native or citrullinated) normalized to positive control is shown. (**D**) Specificity in recognition of citrullinated Mac-1 (*i.e.* Mac-1 ACPAs) was determined by subtracting the binding to the native protein from the binding to the citrullinated version found in **C**. The 90th-percentile

of healthy control reactivity against citrullinated protein was set as the cut-off for positivity. Error bars represent mean \pm standard deviation. Paired t-tests were used to compare within groups while unpaired t-tests were used to compare antibody levels between disease groups. **(E)** Chi-square contingency test on the percentage of healthy controls or RA patients positive for Mac-1 ACPAs. * = $p < 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

As CD11b is mainly found to be complexed with CD18 *in vivo* to form Mac-1, we next analyzed recognition of citrullinated Mac-1 by sera from a larger cohort of healthy controls ($n = 75$) and from a large, well-characterized cross-sectional RA cohort ($n = 165$ patients; ESCAPE cohort). Commercially available recombinant Mac-1 was coated onto wells and citrullinated *in situ* with rhPAD4, and then binding of IgG from healthy control or RA patient sera was quantified. About 30% (49/165) of the RA patients recognized citrullinated Mac-1 while about 10% (17/165) recognized native Mac-1, with significantly higher antibody levels to citrullinated Mac-1 than to the native protein ($p < 0.0001$; **Figure 2-13C**). While the degree of recognition of citrullinated Mac-1 over native protein was also significantly higher in the healthy control population ($p = 0.02$), the levels of antibody binding to citrullinated Mac-1 were much higher in the RA population than in the healthy controls ($p = 0.008$).

Since binding to native Mac-1 was observed for some RA sera, we analyzed the proportion of patients who harbored antibodies that specifically bound to the citrullinated form of the antigen, which we termed Mac-1 ACPAs. The mean level of Mac-1 ACPAs was significantly higher in RA patients than healthy controls ($p = 0.007$; **Figure 2-13D**). This analysis revealed that 25% (42/165) of the RA patients preferred the citrullinated version

of Mac-1, compared to 9% (7/75) of healthy controls ($p = 0.003$; odds ratio = 3.317; **Figure 2-13E**). Overall, our data reveals citrullinated Mac-1 (CD11b/CD18) to be a novel cell surface target of ACPAs in a subset of patients with RA.

2.3.6 Presence of anti-citrullinated Mac-1 antibodies is associated with worse inflammation.

Since Mac-1 ACPAs were present in a subset of patients with RA, we analyzed their clinical significance in the ESCAPE cohort ($n = 165$). Stratifying patients by Mac-1 ACPA status revealed that antibody-positive patients had significantly longer RA duration (median = 14.5 vs. 7 years, $p < 0.001$) and were more likely to be seropositive for anti-cyclic citrullinated peptide (CCP) antibodies (93% vs. 70%; $p = 0.0032$; **Table 7**). Consistent with this observation, Mac-1 ACPA-positive patients harbored a higher number of ACPA fine specificities ($p = 0.006$). Mac-1 ACPAs were also positively associated with anti-PAD4 antibodies ($p < 0.001$) and anti-PAD3/4 cross-reactive antibodies ($p < 0.001$) but were not associated with anti-PAD2 antibodies or rheumatoid factor (RF). Having shared epitope (SE) alleles was linked to the presence of Mac-1 ACPAs (any SE allele, $p = 0.044$; two SE alleles, $p = 0.025$), with the HLA-DRB1*0404 allele being the strongest driver of the association ($p = 0.011$). Clinically, patients with Mac-1 ACPAs did not have higher disease activity on exam or an elevated DAS28 score but did have higher C-reactive protein (CRP, $p = 0.037$) and IL-6 levels ($p = 0.02$). Patients with Mac-1 ACPAs were also more likely to have extraarticular manifestations including rheumatoid nodules ($p = 0.031$) and ground glass opacifications in their lungs on high-resolution chest computed tomography (CT; p

= 0.037) but did not have more clinically apparent interstitial lung disease ($p = 0.20$). Radiographic joint damage was significantly worse in Mac-1 APCA positive patients, as measured by higher total erosion ($p = 0.019$), joint space narrowing (JSN, $p = 0.049$), and Sharp van der Heijde scores (SHS, $p = 0.028$) than Mac-1 ACPA⁻ patients.

Table 7: Disease associations in ESCAPE RA cohort based on Mac-1 ACPA status

Characteristics	Mac-1 ACPA negative	Mac-1 ACPA positive	<i>p</i>-value (neg vs pos)
	n = 123	n = 42	
Age, years, mean ± SD	59 ± 9	61 ± 9	0.14
Male gender, n (%)	45 (37)	20 (48)	0.21
Caucasian, n (%)	104 (85)	38 (90)	0.44
RA duration, years	7 (4-15)	14.5 (7-25)	<0.001
RF seropositivity > 40 units, n (%)	75 (61)	28 (67)	0.51
Anti-CCP seropositivity > 20 units, n (%)	85 (69)	39 (93)	0.003
Anti-CCP units among seropositive; median (IQR)	143 (90-171)	148 (98-187)	0.26
Number of ACPA+ specificities (0-17)	1 (0-6)	4 (1-11)	0.006
Anti-PAD2 positive, n (%)	22 (18)	5 (12)	0.37
Any anti-PAD4 positive, n (%)	17 (14)	38 (90)	<0.001
Anti-PAD3/4 XR positive, n (%)	4 (3)	13 (31)	<0.001
Any HLA-DRB1 shared epitope alleles	81 (66)	34 (81)	0.044
One SE allele	56 (46)	25 (60)	0.15
Two SE alleles	25 (20)	15 (36)	0.025
DRB1*0101	31 (25)	15 (36)	0.17
DRB1*0401	40 (33)	16 (38)	0.47
DRB1*0404	10 (8)	10 (24)	0.011
DRB1*1001	7 (6)	4 (10)	0.47
DAS28, median (IQR)	3.6 (2.8-4.3)	3.7 (2.8-4.4)	0.51
Swollen joint count, median (IQR)	7 (3-10)	7 (2-11)	0.97
Tender joint count, median (IQR)	6 (2-13)	5 (2-13)	0.78
HAQ score (0 – 3), median (IQR)	0.62 (0.25-1.25)	0.81 (0.12-1.38)	0.76
CRP at visit, median (IQR)	2.1 (0.9-5.5)	3.8 (1.5-9.6)	0.037
Average CRP during study, median (IQR)	2.9 (1.1-5.4)	4.5 (1.4-9.2)	0.098
IL-6 at visit, median (IQR)	3.5 (1.6-6.4)	4.9(3.0-24.0)	0.02
Average IL-6 during study, median (IQR)	4.2 (2.1-9.2)	7.0 (2.6-21.5)	0.12
Nodules, n (%)	18 (15)	12 (29)	0.031
Pain (100mm VAS), median (IQR)	21 (8-41)	23 (9-47)	0.89
Non-biologic DMARDs, n (%)	103 (84)	34 (81)	0.68
Biologic DMARDs, n (%)	53 (43)	22 (52)	0.30
Glucocorticoids, n (%)	46 (37)	19 (45)	0.37
Cumulative prednisone, grams, med (IQR)	3.1 (0-8.0)	3.2 (0.1-10.5)	0.78
Number of prior DMARDs, n (%)	1 (0-2)	2 (1-3)	0.047
Any ILD, n (%)	34 (28)	16 (38)	0.20
ILD Score>3 points, n(%)	21 (17)	6 (14)	0.67
GGO, n (%)	11 (9)	9 (21)	0.037
Retic/TB/HC, n (%)	21 (17)	10 (24)	0.37

Characteristics	Mac-1 ACPA negative	Mac-1 ACPA positive	p-value (neg vs pos)
CAC, median (IQR)	3 (0-155)	39 (0-275)	0.90
CAC>0, n (%)	63 (51)	27 (64)	0.14
CAC>100, n (%)	41 (33)	18 (43)	0.27
Baseline SHS>0, n (%)	90 (73)	34 (81)	0.21
Total SHS, median (IQR)	7 (0-32)	20 (2-66)	0.028
Total erosion score, median (IQR)	2 (0-11)	6 (1-22)	0.019
Total JSN score, median (IQR)	4 (0-20)	12 (0-55)	0.049
Δ SHS (per year), median (IQR)*	0.30 (0-1.55)	0.60 (0-3.31)	0.20
Any increase in SHS, n (%)*	54 (44)	23 (55)	0.21

† high level ACPA defined \geq 75th percentile

SD = standard deviation; IQR = interquartile range; BMI = body mass index; RF = rheumatoid factor; XR = cross-reactive; SE = shared epitope; DAS28 = disease activity score-28; HAQ = health assessment questionnaire-disability index; CRP = C-reactive protein; VAS = visual analog score; DMARDs = disease-modifying antirheumatic drugs; ILD = interstitial lung disease; GGO = ground glass opacity; TB/HC = traction bronchiolectasis/honeycombing; CAC = coronary artery calcium; SHS = Sharp van der Heijde score; JSN = joint space narrowing

Given that Mac-1 APCAs were strongly associated with anti-CCP positivity, we next examined whether Mac-1 APCAs identified a clinically distinct subgroup of patients within the anti-CCP⁺ RA population. The RA cohort consisted of 23% anti-CCP–negative patients (37/161; ‘CCP⁻’) and 77% anti-CCP–positive patients (124/161; ‘CCP⁺’) as determined by the CCP assay (**Figure 2-14A**). Of the CCP⁺ patient group, 31.5% (39/124) harbored Mac-1 APCAs and 68.5% (85/124) were negative for Mac-1 APCAs (referred to as ‘CCP only’ population). To define whether Mac-1 APCAs identify a clinically distinct RA subgroup, disease characteristics were compared in the following populations: CCP⁻, CCP only, and Mac-1 ACPA. Comparisons between the CCP⁻ population and the whole CCP⁺ population are shown in **Figure 2-14B** as a reference.

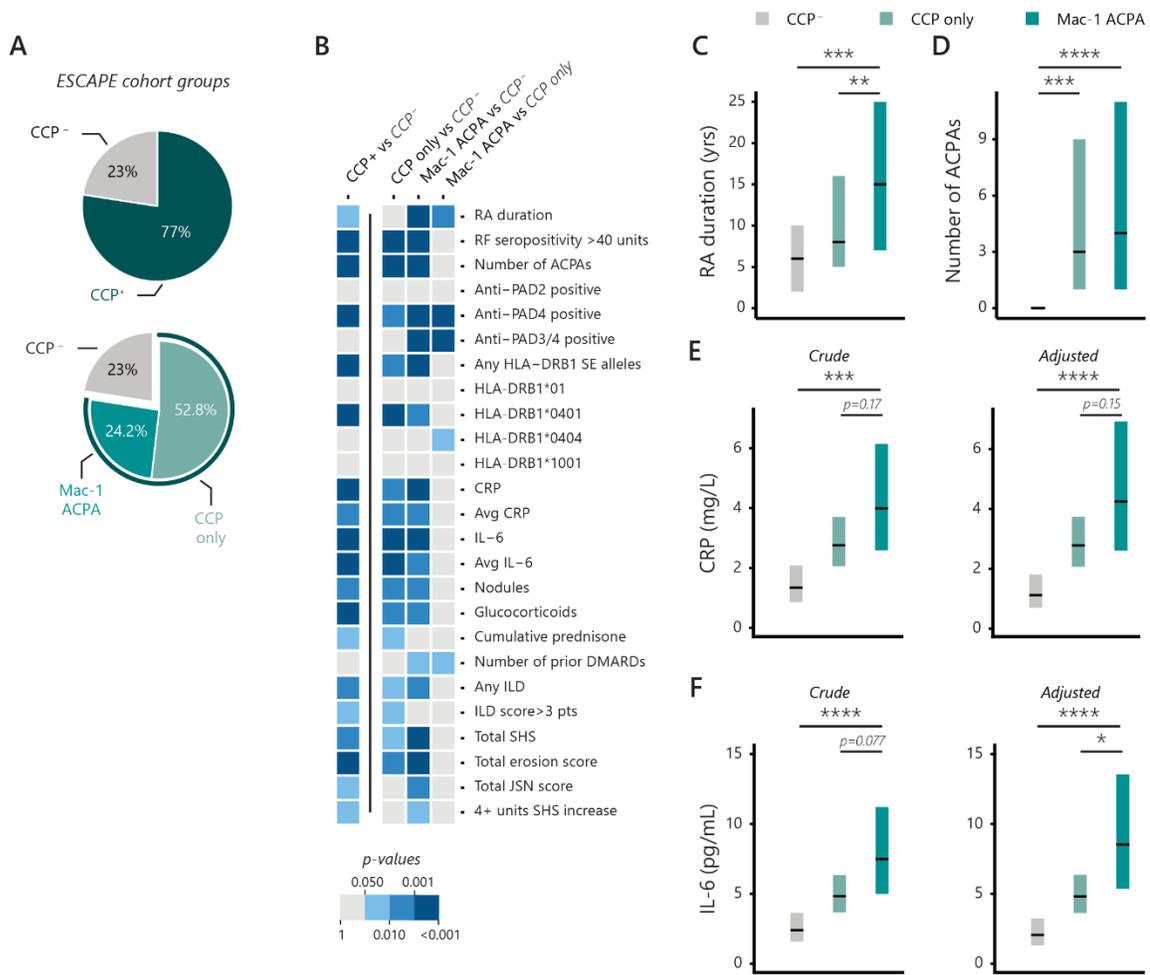


Fig. 2-14. Disease associations in ESCAPE RA cohort. (A) ESCAPE RA cohort subgroup composition for n=161 patients. For this analysis, we excluded three CCP negative patients who were positive for Mac-1 ACPAs and one individual whose CCP status was unknown. (B) Heatmap comparing disease characteristics between RA patient groups. The reference group for each comparison is shown in light, italicized text. (C and D) Median and interquartile ranges of RA duration (C) and number of ACPAs (D) in the various RA cohort subpopulations. (E and F) Univariate and multivariate analyses of CRP levels (E) and IL-6 levels (F) in the RA cohort subpopulations shown. Data represents mean and 95% confidence intervals. * = $p < 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$. CCP = cyclic citrullinated peptide, RF = rheumatoid factor; SE = shared epitope; CRP = C-reactive protein; DMARDs = disease-modifying anti-rheumatic drugs; ILD = interstitial lung disease; SHS = Sharp/van der Heijde score; JSN = joint space narrowing.

The demographic characteristics of the three serologic subgroups were similar (**Table 8**). However, compared to the CCP⁻ population, patients in the CCP only and Mac-1 ACPA groups were more likely to be positive for RF and HLA-DRB1 SE alleles, with HLA DRB1*0401 being the most enriched (**Table 8, Figure 2-14B**). Patients with CCP only or Mac-1 ACPAs had an increase in inflammatory markers compared with CCP⁻ patients, including CRP and IL-6 levels, although there was no difference in their disease activity scores (**Figure 2-14B**). Compared to CCP⁻ patients, CCP only and Mac-1 ACPA patients also had increased joint damage with higher erosion scores and total SHS, as well as a higher prevalence of rheumatoid nodules and ILD (**Table 8, Figure 2-14B**).

Table 8: Disease associations in ESCAPE RA cohort based on antibody subset status

Characteristics	CCP- (n = 37)	CCP only (n = 85)	Mac-1 ACPA (n = 39)	p-value (CCP only vs CCP-)	p-value (Mac-1 ACPA vs CCP-)	p-value (Mac-1 ACPA vs CCP only)
Age, years	58 ± 9	58 ± 9	61 ± 9	0.97	0.20	0.11
Male gender, n (%)	13 (35)	32 (38)	18 (46)	0.79	0.33	0.37
Caucasian, n (%)	34 (92)	69 (81)	35 (90)	0.18	0.99	0.30
RA duration, years	6 (2-10)	8 (5-16)	15 (7-25)	0.10	0.001	0.003
RF seropositivity > 40 units, n (%)	6 (16)	69 (81)	28 (72)	<0.001	<0.001	0.24
Number of ACPA+ (0-17)	0 (0-0)	3 (1-9)	4 (1-11)	<0.001	<0.001	0.39
Anti-PAD2 positive, n (%)	6 (16)	16 (19)	5 (13)	0.73	0.67	0.41
Any anti-PAD4 positive, n (%)	0 (0)	17 (20)	35 (90)	0.003	<0.001	<0.001
Anti-PAD3/4 XR positive, n (%)	0 (0)	4 (5)	12 (31)	0.31	<0.001	<0.001
Any HLA-DRB1 shared epitope alleles, n(%)	17 (47)	64 (75)	32 (84)	0.003	0.001	0.27
DRB1*01	11 (31)	20 (24)	14 (37)	0.42	0.57	0.13
DRB1*0401	3 (8)	37 (44)	15 (39)	<0.001	0.002	0.67
DRB1*0404	4 (11)	6 (7)	9 (24)	0.48	0.22	0.015
DRB1*1001	1 (3)	6 (7)	4 (11)	0.67	0.36	0.50
DAS28, median (IQR)	3.5 (2.6-4.6)	3.6 (3.0-4.1)	3.8 (2.8-4.4)	0.87	0.55	0.43
Swollen joint count, median (IQR)	8 (4-10)	6 (3-9)	7 (2-11)	0.65	0.80	0.94
Tender joint count, median (IQR)	7 (3-18)	6 (2-11)	5 (2-13)	0.20	0.34	0.83
HAQ score (0 - 3), median (IQR)	0.88 (0.5-1.25)	0.50 (0.12-1.25)	0.88 (0.12-1.50)	0.056	0.58	0.32
CRP, median (IQR)	1.3 (0.6-3.5)	2.6 (1.1-7.0)	3.9 (1.6-12.0)	0.006	0.001	0.16

Characteristics	CCP- (n = 37)	CCP only (n = 85)	Mac-1 ACPA (n = 39)	p-value (CCP only vs CCP-)	p-value (Mac-1 ACPA vs CCP-)	p-value (Mac-1 ACPA vs CCP only)
Average CRP, median (IQR)	1.6 (0.7-3.7)	3.2 (1.7-7.2)	5.0 (1.4-9.4)	0.007	0.006	0.30
IL-6, median (IQR)	1.8 (1.1-3.5)	4.4 (2.0-8.3)	5.2 (3.0-24.0)	0.001	<0.001	0.11
Average IL-6, median (IQR)	2.3 (1.8-4.3)	4.9 (3.1-12.3)	7.2 (2.9-21.6)	0.001	0.005	0.28
Nodules, n (%)	2 (6)	16 (20)	12 (33)	0.007	0.006	0.16
Pain (100mm VAS), median (IQR)	23 (10-41)	20 (7-37)	23 (9-47)	0.38	0.59	0.79
Non-biologic DMARDs, n (%)	31 (84)	72 (85)	31 (79)	0.90	0.63	0.47
Biologic DMARDs, n (%)	19 (51)	34 (40)	22 (56)	0.24	0.66	0.088
Glucocorticoids, n (%)	6 (16)	39 (46)	18 (46)	0.002	0.005	0.98
Cumulative prednisone, grams, med (IQR)	1.6 (0.5-4)	4.3 (0.8-8.7)	3.1 (0.1-11.4)	0.03	0.14	0.67
Number of prior DMARDs, median (IQR)	1 (0-2)	1 (0-2)	2 (1-3)	0.90	0.037	0.024
Any ILD, n (%)	5 (15)	29 (38)	15 (43)	0.014	0.01	0.64
ILD Score>3 points, n (%)	2 (6)	19 (25)	6 (17)	0.019	0.26	0.47
GGO, n (%)	3 (9)	8 (11)	9 (26)	0.99	0.11	0.052
Retic/TB/HC, n (%)	4 (12)	17 (24)	9 (26)	0.14	0.22	0.81
CAC, median (IQR)	3 (0-134)	0 (0-161)	66 (0-330)	0.98	0.10	0.068
CAC>0, n (%)	20 (54)	42 (49)	26 (67)	0.64	0.26	0.073
CAC>100, n (%)	11 (30)	30 (35)	18 (46)	0.55	0.14	0.25
Baseline SHS>0, n (%)	25 (68)	64 (75)	31 (82)	0.38	0.16	0.44
Total SHS, median (IQR)	4 (0-11)	10 (1-53)	20 (2-66)	0.02	0.001	0.22

Characteristics	CCP- (n = 37)	CCP only (n = 85)	Mac-1 ACPA (n = 39)	p-value (CCP only vs CCP-)	p-value (Mac-1 ACPA vs CCP-)	p-value (Mac-1 ACPA vs CCP only)
Total erosion score, median (IQR)	0 (0-3)	3 (0-21)	6 (1-22)	0.003	<0.001	0.29
Total JSN score, median (IQR)	2 (0-9)	6 (0-29)	12 (0-55)	0.068	0.008	0.22
Δ SHS (per year), median (IQR)*	0 (0-1.2)	0.6 (0-1.6)	0.3 (0-3.5)	0.43	0.20	0.56
Any increase in SHS, n (%)*	16 (48)	38 (57)	20 (62)	0.44	0.26	0.59
4+ unit SHS increase, n (%)	1 (3)	10 (15)	7 (22)	0.095	0.027	0.40

CCP = cyclic citrullinated peptide; IQR = interquartile range; BMI = body mass index; RF = rheumatoid factor; XR = cross-reactive; SE = shared epitope; DAS28 = disease activity score-28; HAQ = health assessment questionnaire-disability index; CRP = C-reactive protein; VAS = visual analog score; DMARDs = disease-modifying antirheumatic drugs; ILD = interstitial lung disease; GGO = ground glass opacity; TB/HC = traction bronchiolectasis/honeycombing; CAC = coronary artery calcium; SHS = Sharp van der Heijde score; JSN = joint space narrowing

Of the 165 patients in the ESCAPE RA cohort, one patient lacked CCP information and three were positive for Mac-1 ACPA but were CCP negative. These patients were removed from the analysis in Table 8.

When comparing the CCP only and Mac-1 ACPA groups to each other, we observed subtle but significant differences in several domains. Mac-1 ACPA patients had significantly longer RA duration ($p = 0.003$) with a substantially higher number of patients possessing anti-PAD4 antibodies ($p < 0.001$), anti-PAD3/4 ($p < 0.001$) antibodies, and HLA-DRB1*0404 alleles ($p = 0.015$) as compared to the CCP only group (**Figure 2-14C and D, Table 8**). Mac-1 ACPA patients also had a greater history of prior treatment with disease modifying antirheumatic drugs (DMARDs; $p = 0.024$) and were more likely to be currently undergoing treatment with biologic DMARDs ($p = 0.088$). Additionally, while not significant, ground glass opacifications in the lungs ($p = 0.052$) and increased coronary artery calcium (CAC) scores ($p = 0.068$) approached significance in the Mac-1 ACPA group compared to the CCP only group (**Table 8**). Notably, the median erosion, JSN, and total SHS were 2-fold higher in patients with Mac-1 ACPAs compared to CCP only patients. Although the levels did not significantly differ between the two groups, the Mac-1 ACPA group differed more significantly from the CCP⁻ population than did the CCP only group, and JSN scores were significantly different only between Mac-1 ACPA and CCP⁻ individuals ($p = 0.008$).

Univariate analyses revealed a trend toward the highest levels of CRP and IL-6 in the Mac-1 ACPA group compared to the CCP only group ($p = 0.17$ for CRP and $p = 0.077$ for IL-6; **Figure 2-14E and F, Table 9**). In multivariate analyses, after adjusting for RA duration, CCP positivity, anti-PAD3/4 antibody positivity, the presence of SE alleles, and number of prior failed DMARDs, this trend toward higher CRP levels in Mac-1 ACPA positive compared to CCP only patients was maintained, and significantly higher IL-6 levels were observed in the Mac-1 ACPA positive group compared to the CCP only group ($p =$

0.15 for CRP and $p = 0.043$ for IL-6; **Figure 2-14E and F, Table 9**). These findings suggest that patients with autoantibodies targeting citrullinated Mac-1 may represent a clinically and serologically distinct subgroup within the CCP⁺ RA population.

Table 9: Multivariate analysis of disease characteristics in the RA ESCAPE cohort

Characteristics	CCP⁻	CCP only	Mac-1 ACPA	<i>p</i>-value (Mac-1 ACPA vs CCP⁻)	<i>p</i>-value (Mac-1 ACPA vs CCP only)
CRP, mean (95% CI)	1.12 (0.70, 1.81)	2.78 (2.07, 3.73)	4.25 (2.61, 6.92)	<0.001	0.15
IL-6, mean (95% CI)	2.06 (1.31, 3.24)	4.81 (3.63, 6.36)	8.53 (5.37, 13.54)	<0.001	0.043
GGO, % (95% CI)	6 (1, 21)	15 (7, 29)	29 (11, 58)	0.064	0.27
SHS, mean (95% CI)	8 (5, 14)	19 (14, 26)	18 (11, 31)	0.056	0.94
CAC, % (95% CI)	59 (39, 77)	45 (32, 58)	67 (44, 84)	0.61	0.11

CRP = C-reactive protein; GGO = ground glass opacity; SHS = Sharp van der Heijde score; CAC = coronary artery calcium, CI = confidence interval

Adjusted for the following variables: RA duration, CCP positivity, anti-PAD3/4 antibody positivity, the presence of SE alleles, and number of prior failed DMARDs

2.4 DISCUSSION

While it is known that PAD4 plays an important role in autoantigen generation in RA pathogenesis, the cellular mechanisms are still being elucidated. As one of the major immune infiltrators of the RA joint,¹⁴¹ monocytes represent a key driver of inflammation. Although monocytes express PAD4,⁹⁹ the subcellular localization of PAD4 and role of monocytes in citrullinated autoantigen generation remained unclear. In this study, we found that PAD4 localizes to the cell surface of monocytes in an enzymatically active state and is capable of citrullinating extracellular as well as cell surface substrates. We also identified multiple citrullinated cell surface proteins and demonstrated that citrullinated Mac-1 is a novel monocyte surface target of ACPAs in RA patients.

After the discovery of its nuclear localization signal, studies demonstrating its localization in the nuclei, and the presence of citrullinated nuclear proteins, PAD4 was thought to localize and be functionally active primarily in the nucleus.^{91,100,180-183} However, the RA synovial fluid has been known to contain a broad range of citrullinated proteins, both intra- and extracellular, in addition to soluble PADs.^{107,109,110,113,122} Therefore, theories about the origins of extracellular citrullination centered around lytic cell death mechanisms that allowed for PADs to be released and to be active in the extracellular space.^{93,117,184} Yet, studies have shown that PADs require a reducing environment to be catalytically active; the addition of reducing agents and sometimes supraphysiologic calcium levels are necessary for soluble PADs from RA synovial fluid to citrullinate substrates *in vitro*.^{107,117,185} As a result, the levels of citrullination seen in the RA joint cannot be fully addressed by

current conceptual hypotheses, suggesting additional unknown mechanisms involved in the citrullination of extracellular proteins.

Recently, PAD4 was discovered in cellular compartments outside of the nucleus, but its functions in other subcellular locations are not completely understood.^{128,150} While PAD4 was found to be involved in the regulation of NADPH oxidase and to be active on the cell surface, these studies were mainly focused on neutrophils. In addition to the well-appreciated role of neutrophils in contributing citrullinated antigens and extracellular PAD4 enzymes to the RA joint via the release of PADs during cell death,^{117,122,126} our finding of enzymatically active PAD4 on the monocyte surface along with citrullination of both extracellular and surface antigens offers a parallel and complementary mechanism for RA autoantigen generation. Since monocytes are capable of surviving for long periods of time in inflammatory environments,¹⁸⁶ they may sustain prolonged extracellular autoantigen generation without the requirement for cell death. Therefore, our finding that PAD4 is present on the cell membrane of monocytes and can citrullinate extracellular fibrinogen as well as a newly described endogenous monocyte membrane autoantigen, expands the mechanisms for generation of the RA citrullinome.

Although the synovial fluid of RA patients contains calcium concentrations close to physiologic levels, osteoclast-driven bone erosion sites have much higher local levels of calcium.^{187,188} While neutrophils are mostly found in the synovial fluid, infiltrating monocytes are also present in the pannus, a thickened and invasive feature of the RA synovial membrane, close to the sites of bone erosion.^{16,189} Therefore, monocytes may be

poised to take advantage of this unique opportunity at the interface between the synovium and eroding bone, thereby presenting a means for the constant generation of citrullinated autoantigens. Whether such a mechanism holds true *in vivo* should be the subject of future investigation.

We have found multiple cell surface proteins to be citrullinated by endogenous surface PAD4, suggesting a period of close contact with the catalytically active enzyme. Interactions between PAD4 and nuclear transcription factors such as Elk-1 have been known to be critical for their translocation to promoter sequences.¹⁹⁰ Therefore, it can be speculated that surface proteins may also serve as binding partners that aid in PAD4 trafficking to the plasma membrane. Such binding partners may also enable PAD4 to be held in a structurally active conformation, similarly to anti-PAD3/4 antibodies that lower the calcium threshold needed for PAD4 activation.¹²⁷ In addition, association with plasma membrane lipids are known to lower the calcium requirements of rabbit musculoskeletal PAD.⁹⁴ Our observation of the presence of citrullinated surface proteins on monocytes incubated in both calcium and EDTA suggests that circulating monocytes may carry citrullinated protein cargo on their plasma membranes. A basal level of surface protein citrullination may exist *in vivo*, of which a set were seen to have augmented citrullination following *ex vivo* exposure to a calcium-rich environment. Therefore, the surface citrullinome may be influenced via several factors, such as monocyte activation and altered protein trafficking and signaling, due to high extracellular calcium levels. In addition, while we observed the ability of PAD4 to catalyze citrullination at the monocyte cell surface, it is also possible that some proteins may be citrullinated intracellularly and then trafficked to

the plasma membrane under specific conditions. Future studies on human PAD4 interactions between protein binding partners and membrane lipid components will further our understanding of the mechanisms of PAD4 trafficking, activity, and expression at the cell surface.

In this study, we have identified multiple citrullinated monocyte cell surface proteins. We have also shown that the citrullinated proteins generated by surface PAD4 can be targeted by autoantibodies present in patients, implicating a role for cell surface PAD4 in pathogenic citrullination in RA. This finding suggests that the presence of active PAD4 on the cell surface may render monocytes cellular targets of ACPAs. Indeed, previous studies have shown that ACPAs can trigger osteoclastogenesis and pro-inflammatory cytokine production by monocytes.^{143,191} However, the autoantigenic targets and precise mechanisms were largely unknown, with candidate antigens identified as either citrullinated vimentin or citrullinated Grp78. The mass spectrometry results in our study identified multiple additional monocyte cell surface proteins to be citrullinated, most notably integrins. As an integrin with several ligands, including the PAD4 substrate fibrinogen,¹⁷⁵ citrullination of Mac-1 and the consequences of its interactions with native and citrullinated ligands are important topics for future study to fully understand their contribution to inflammatory mechanisms seen in RA.

The implications of Mac-1 ACPAs in disease pathogenesis are suggested by our findings that patients with Mac-1 ACPAs had overall worse disease characteristics compared to Mac-1 ACPA negative individuals. An association of Mac-1 ACPA with longer

RA duration was notable. Whether Mac-1 ACPAs may have evolved in part due to epitope spreading caused by chronic inflammation should be evaluated in future studies.⁷³ Despite accounting for interrelated disease factors that can confound analyses, we found that markers of systemic inflammation, such as CRP and IL-6, were consistently elevated in the Mac-1 ACPA group. IL-6 serves as a nexus in acute and chronic inflammation as it is produced by innate immune cells and can stimulate adaptive immune responses by T and B cells, as well as promote the production of additional inflammatory cytokines.¹⁹² Monocytic recruitment in response IL-6 has also been associated with the transition from acute to chronic inflammation, and increased IL-6 levels are associated with more erosive disease.^{193,194} Interestingly, inhibition of IL-6 signaling via blockade of the IL-6 receptor is an effective treatment for RA. The ESCAPE RA cohort used in our study pre-dated the commercial introduction of IL-6 inhibitors for RA and the patients were primarily treated with non-biologic DMARDs and TNF- α inhibitors. Given our findings showing higher IL-6 levels and larger number of prior DMARDs used in patients with Mac-1 ACPAs, IL-6 inhibitor therapy may be a particularly effective option in the Mac-1 ACPA positive RA subset. Future longitudinal and treatment studies will be important to elucidate the clinical value of these novel autoantibodies.

In summary, we show that PAD4 on the monocyte surface can citrullinate extracellular and surface proteins, generating both known and novel citrullinated RA autoantigens. This represents a previously undescribed site of protein citrullination, which may contribute uniquely to the RA citrullinome and disease pathogenesis by circumventing the need for cell death as a mechanism for extracellular protein

citrullination, transforming the monocyte into a direct cellular target of autoantibodies, and potentially altering monocyte adhesion. The monocyte surface therefore presents a novel site of autoantigen generation in RA with important implications for understanding the role of PAD4 in monocyte function and RA pathology.

CHAPTER 3

An unbiased proteomic analysis of PAD4 in human monocytes: novel substrates, binding partners, and subcellular localizations

Material presented in this chapter is adapted from: Thomas MA, Kim SY, Curran AM, Smith B, Antiochos B, Na CH, Darrah E. An unbiased proteomic analysis of PAD4 in human monocytes: novel substrates, binding partners and subcellular localizations. Philos Trans R Soc Lond B Biol Sci. 2023 Nov 20;378(1890):20220477. doi: 10.1098/rstb.2022.0477. Epub 2023 Oct 2. PMID: 37778379; PMCID: PMC10542449.

3.1 INTRODUCTION

Peptidylarginine deiminases (PADs) have been implicated in playing a key role in several diseases including sepsis, thrombosis, cancer and autoimmunity.^{195,196} Their pathogenic role is perhaps best understood in the context of rheumatoid arthritis (RA), a common systemic autoimmune disease characterized by chronic inflammation and irreversible damage to synovial joints.¹⁶ A hallmark characteristic of RA is the presence of anti-citrullinated protein antibodies (ACPAs), which target proteins in which arginine residues have been post-translationally deiminated by PAD enzymes in a process known as citrullination. Of the five PAD isoforms (PAD1-4, and 6), PAD4 is expressed primarily by granulocytes and monocytes and is present at high levels in the inflamed RA joint.^{97,100,99,98,106,107} Polymorphisms in the *PADI4* gene have been associated with RA development, and knocking out or inhibiting PAD4 has been shown to ameliorate disease in mouse models of RA.^{23,118,119} While the importance of PAD4 to RA pathogenesis is clear, the subcellular localization and functions of the enzyme during normal physiology, which may be exploited in pathogenic states, remain understudied.

Although PAD4 was originally thought to be a nuclear protein due to the presence of a nuclear localization signal, new evidence suggests a more nuanced cellular localization.¹⁰⁰ More recent studies have also described PAD4 in the neutrophil cytosol as well as on the cell surface of neutrophils and monocytes, and importantly, have identified citrullinated proteins at these locations.^{128,129} Additionally, the list of known PAD4 substrates has been expanding to include proteins from several subcellular locations including nuclear, cytosolic, cell surface, and extracellular domains.^{102,111,146,147,197,198}

Importantly, however, it remains unknown how PAD4 traffics to and functions in extranuclear locations. In this study, we conducted an unbiased proteomic investigation of the cellular localization, organelle substrates, and binding partners of PAD4, and identified monocyte differentiation into dendritic cells as a cellular system to study PAD4 trafficking. This systematic analysis revealed a previously undescribed subcellular localization of PAD4 to vesicular organelles and identified numerous novel substrates and putative binding partners involved in vesicular trafficking, suggesting a multi-faceted role for PAD4 in both health and disease.

3.2 METHODS

Human subjects. De-identified leukopaks from healthy donors who gave platelets for medical purposes at the Anne Arundel Blood Donation Center were used as a source of leukocytes.

Monocyte isolation. Peripheral blood mononuclear cells (PBMCs) were separated using Ficoll-Pacque (GE Healthcare) density gradient centrifugation. Monocytes from the PBMC fraction were positively isolated using human CD14 Microbeads (Miltenyi Biotec) according to manufacturer's instructions. Negatively isolated monocytes using the EasySep Human Monocyte Isolation Kit (STEMCELL Technologies) were used for transmission electron microscopy.

Monocyte subcellular fractionation and Western blot. 5×10^7 positively isolated monocytes were fractionated into subcellular compartments using the Minute Plasma Membrane Protein Isolation and Cell Fractionation Kit (Invent Biotechnologies) according to the manufacturer's instructions. Total cell lysates were created using 1×10^7 monocytes lysed in NP40 lysis buffer supplemented with PMSF and protease inhibitors. Nuclear lysates were created in RIPA buffer supplemented with PMSF and protease inhibitors. Plasma membrane and organelle fractions were saved in NP40 lysis buffer with PMSF and protease inhibitors. Lysate protein concentrations were determined using bicinchoninic acid (BCA) assay (Pierce BCA Protein Assay Kit, Thermofisher) by following manufacturer's protocol, and absorbance at 560 nm was read using Wallac Victor 3 1420 Multilabel Counter (PerkinElmer) and 5 μ g of protein was denatured in Laemmli buffer, loaded per

condition, and separated on 4-12% Bis-Tris SDS-PAGE gels (Invitrogen) at 200V for 30 minutes. Proteins were transferred onto nitrocellulose membranes and immunoblotted using rabbit anti-human PAD4 (N-terminal P4749, Sigma), rabbit anti-human Na⁺/K⁺ ATPase 1 (1:1000; 3010S, Cell Signaling), rabbit anti-human α -tubulin (1:1000, ab52866, Abcam), rabbit anti-human GAPDH (1:2000; 51745, Cell Signaling), and mouse anti-human lamin-B1 (1:500; ab8982, Abcam). Images of blots were taken using the FluorChem M detection system (ProteinSimple), and PAD4 levels were quantified using AlphaView SA software (ProteinSimple).

To quantify PAD4 and PAD2 levels, monocytes were positively isolated from the peripheral blood of three healthy controls and cell lysates in NP40 lysis buffer were prepared and quantified as described above. Lysates (5 μ g) were resolved using SDS-PAGE and immunoblotted using mouse anti-PAD4 antibody (1:1000, ab128086, Abcam), rabbit anti-PAD2 antibody (1:1000; ab16478, Abcam), and mouse anti- β -actin antibody (1:1000; clone: AC-15, MilliporeSigma). Following HRP-labelled secondary antibody incubation, blots were imaged as mentioned above and protein levels were quantified using AlphaView SA software (ProteinSimple). PAD4 and PAD2 levels were normalized to β -actin levels present in each donor and graphed using GraphPad Prism 10.0.1 (Dotmatics).

Transmission electron microscopy. 5×10^7 negatively isolated monocytes using the EasySep Human Monocytes Isolation Kit (STEMCELL Technologies) were fixed using 4% paraformaldehyde, 0.1% glutaraldehyde, 3 mM MgCl₂ in 0.1 M Sorenson's sodium phosphate buffer (pH 7.2) overnight at 4°C. Following a buffer wash, 1% osmium tetroxide, 0.8% potassium ferrocyanide in 0.1 M sodium phosphate was used to postfix samples for

1 to 2 hours on ice in the dark. Maleate buffer (0.1 M) was used to rinse samples after osmium treatment. The cells were stained *en bloc* using 2% uranyl acetate in maleate buffer for 1 hour in the dark, dehydrated in a graded series of ethanol, and embedded in Eponate 112 (Polyscience) resin. Sample polymerization was conducted overnight at 60°C.

Sections measuring 60-90 nm were cut using a diamond knife on a Leica UCT ultramicrotome and picked up with Formvar coated 200 mesh nickel grids. To label sections with PAD4 antibodies, grids were hydrated using dH₂O and floated for the rest of the steps. 1.5% sodium meta periodate was used to etch samples, which were subsequently rinsed using dH₂O and 10 mM NH₄Cl in Tris Buffered Saline (TBS). Samples were blocked using 2% BSA in TBS for 30 minutes and incubated in primary antibody (1:25 dilution for C-terminal antibody [generated in-house¹⁵³] and P4749 N-terminal antibody [Sigma], 1:50 for clone OT14H5 [Abcam], or buffer) overnight at 4°C.

Following overnight incubation, grids were allowed to warm to room temperature for 1 hour and then rinsed with blocking buffer and TBS. 12 nm-gold labelled secondary antibodies (AffiniPure goat anti-rabbit IgG [111-205-144] and AffiniPure goat anti-mouse IgG [115-205-146] from Jackson ImmunoResearch Laboratories) were added at 1:40 dilution at room temperature for 2 hours, and samples were fixed using glutaraldehyde in sodium cacodylate buffer. Samples were then stained with 2% uranyl acetate and rinsed in water. Grids were imaged using a Hitachi 7600 TEM at 80 kV. An AMT CCD XR80 (8-megapixel camera – side mount AMT XR80 – high-resolution high-speed camera) was used to capture images.

Preparing samples for mass spectrometry. 5×10^7 positively isolated monocytes from three healthy donors were incubated in 5 mM calcium in Hank's Balanced Salt Solution (HBSS) for 3 hours at 37°C. Cells were subjected to subcellular fractionation as described above. Lysates of organelle fractions were treated to remove detergents via methanol-chloroform precipitation. They were subsequently reconstituted in lysis buffer containing 8 M urea, 10 mM tris(2-carboxyethyl)phosphine (TCEP), and 40 mM chloroacetamide (CAA) in 50 mM triethylammonium bicarbonate (TEAB). Proteins were reduced and alkylated at room temperature for 1 hour. 8 M urea was diluted to 2 M via the addition of three volumes of 50 mM TEAB before digestion. Trypsin (sequencing grade modified trypsin, Promega) was used to digest proteins at 10 ng/ μ L (v/v) at 37°C overnight. Following digestion, peptides were acidified using 1% trifluoroacetic acid (TFA) and desalted using C₁₈ StageTips. Peptides were eluted, vacuum dried, and stored at -80°C before mass spectrometry analysis.

Mass spectrometry analysis. An Orbitrap Fusion Lumos Tribrid mass spectrometer interfaced with an Ultimate-3000 RS LCnano nanoflow liquid chromatography system (Thermo Scientific) was used to analyse peptides. To identify proteins from mass spectrometry spectra, files were analysed using Proteome Discoverer (version 2.4.1.15, Thermo Scientific). Mass spectra results were searched against the human UniProt database (released in Jan. 2021) containing common contaminant proteins using SEQUEST HT algorithms. The following parameters were used for database search: 1) trypsin with a maximum of two missed cleavage sites; 2) 10 ppm precursor mass error tolerance; 3) 0.02 Da fragment mass error tolerance; 4) fixed modification: carbamidomethylation

(+57.02146 Da) at cysteines; 5) variable modifications: oxidation at methionine (+15.99492 Da), deimidation at arginine, asparagine, and glutamine (+0.98402 Da), protein N-terminus acetylation (+42.01057 Da), loss of methionine (-131.04049 Da), and methionine loss with acetylation (-89.02992 Da). Minimum and maximum peptide length were set at six and 35 amino acids, respectively. 1% false discovery rate was used to filter proteins and peptides.¹⁵⁹⁻¹⁶¹ Data processing workflow and representative spectra are provided. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD044188.¹⁶⁹

Peptide data obtained from mass spectrometry was filtered based on citrullination of arginines (excluding those occurring only at C-terminal arginine residues), and the first accession number listed was used to find the corresponding protein for downstream analysis. Fifteen proteins were excluded due to low confidence, and the abundances of the resulting 271 proteins were analysed for each donor. The 257 proteins shared by all three donors were subsequently analysed for subcellular localization using the UniProt database.¹⁶² Accession numbers of three proteins were listed as 'deleted', so they were substituted with the first protein accession number in UniProt under the same gene name. Proteins were sorted based on the first designated localization listed under 'Subcellular localization' and 'Gene Ontology (cellular component)', and proteins predominantly associated with nuclei and cytosol were removed from analysis. The resulting list of 99 citrullinated proteins were then analysed for statistical overrepresentation compared to the human proteome using the PANTHER classification system.^{163,164} Ten proteins were unable to be mapped by PANTHER, so their accession numbers were substituted for the

first listed accession number in UniProt under the same gene name. Statistical significance was determined using Fisher's exact test with Benjamini-Hochberg false discovery rate correction. Thresholds were set at false discovery rate corrected $p = 0.01$ and ± 2 fold enrichment.

Citrullination site amino acid sequence motif analysis. Potential PAD4 citrullination site sequence motifs were evaluated by aligning the P6-P6' positions around each deiminated arginine (i.e., citrulline) detected in the peptides corresponding to the 271 citrullinated proteins identified in the subcellular organelle fractions by mass spectrometry. Citrullination sites containing fewer than six amino acids on either side of the citrulline were excluded from motif analysis. The frequency of each amino acid at each position in the citrullination site was assessed using pLogo v1.2.0 (Schwartz Lab, University of Connecticut) to generate a probability logo to visualize enrichment of amino acids within the citrullination site.¹⁹⁹ Log-odds of the binomial probabilities for each amino acid at each position were used to plot amino acid residues in proportion to their respective frequencies. Residue frequencies were assessed for over- or underrepresentation at each position compared to the human proteome. Significance thresholds were set at $p < 0.05$, represented as log-odds of ± 3.64 .

HuProt binding chip assay. To identify potential surface binding partners of PAD4, PAD4 protein, purified as previously described,¹⁵³ was labelled with Alexa Fluor 555 Protein Labeling Kit as per manufacturer's instructions (ThermoFisher). Fluorescently labelled PAD4 protein was incubated on a blocked HuProt™ protein microarray chip (CDI Laboratories) that contained over 16,000 GST-tagged proteins for 1 hour at room temperature with

shaking. Since calcium binding to PAD4 induces conformation changes and catalysis, the incubation was performed at increasing calcium concentrations (0, 0.2, and 2 mM). Binding to candidate proteins was visualized by fluorimetry and normalized to fluorescence obtained with a Cy5-labeled anti-GST antibody. An A-score was calculated for each normalized protein value, representing the difference in PAD4 binding to each candidate protein minus the array average divided by the standard deviation. Binding partners with an A-score of >2 were considered to be putative binding partners for further analysis. The resulting list of putative PAD4 binding partners was then analysed for statistical overrepresentation compared to the human proteome using the PANTHER classification system.^{163,164} Statistical significance was determined using Fisher's exact test with Benjamini-Hochberg false discovery rate correction. Thresholds were set at false discovery rate corrected $p = 0.01$ and ± 2 fold enrichment.

Analysing binding of PAD4 to putative binding partners. To confirm binding of PAD4 to candidate binding partners, 500 ng of recombinant human (rh)-CEACAM1 (SinoBiologicals), rh-enolase (purified as described⁴²), rh-nucleophosmin (purified as previously described²⁰⁰), rh-MNDA (purified as previously described²⁰¹), or 12.5 ng rhPAD4 (generated in-house as previously described¹⁵³) were coated per well in triplicates onto high-binding ELISA plates (Corning) overnight at 4°C. Wash steps were performed using 0.02% Tween-20 in PBS. Plates were blocked using 2% BSA in PBS for 1 hour at 37°C, followed by serial dilution (2600 ng to 0 ng) of rh-PAD4 in 0.2 mM Ca²⁺ and 100 mM Tris-HCl (pH = 7.5) for 2.5 hours at room temperature with shaking. Wells were subsequently incubated with 1:250 dilution of rabbit anti-PAD4 (clone: P4749, MilliporeSigma) for 2

hours and then with 1:7500 dilution of HRP-labelled anti-rabbit secondary antibody for 1 hour at room temperature with shaking. TMB substrate (KPL) was used to develop wells, and the reaction was stopped by the addition of 1 M HCl. A Wallac Victor 3 1420 Multilabel Counter plate reader (PerkinElmer) was used to read absorbances at 450 nm and 560 nm (background). Absorbance at 0 ng PAD4 was subtracted from each PAD4 concentration for each binding partner and the adjusted absorbance was normalized to anti-PAD4 antibody binding to 12.5 ng of rhPAD4. Binding affinity plots and K_d values were generated using the 'One-site – total binding' model in GraphPad Prism 10.0.1 (Dotmatic).

Monocyte derived-dendritic cell (mo-DC) differentiation. Monocytes isolated using positive selection were cryopreserved in 90% heat-inactivated FBS and 10% DMSO until use. Thawed cells were resuspended at 1×10^6 cells/mL in Mo-DC Differentiation Medium containing granulocyte monocyte-colony stimulating factor (GM-CSF) and interleukin-4 (IL-4; Miltenyi Biotec) for three days. On the third day, cells were supplemented with equal volume of media and cultured for 3 more days for a total of 6 days. On the sixth day, cells were harvested, counted, and used for downstream assays.

Confocal microscopy. Monocytes (day 0) and mo-DCs (day 3 and day 6) were resuspended at 10×10^6 cells/mL in PBS, 100 μ L of cells were added dropwise onto coverslips, and allowed to adhere for a minimum of 30 minutes in 1 mL PBS. Cells were subsequently fixed in 4% paraformaldehyde for 30 minutes and washed three times using PBS. When indicated, cells were permeabilized by incubating coverslips into 100% acetone for 30 seconds. 5% BSA in PBS was used to block for 30 minutes at room temperature. Following three rinses in PBS, primary antibody dilutions were created in 1% BSA in PBS

and added to coverslips for 1 hour at room temperature. The following primary antibody dilutions were used in this study: mouse anti-human HLA-DR (1:50; clone: L243, Biolegend), mouse anti-human Lamp1 (1:100; clone: eBioH4A3, Thermofisher), mouse anti-human Rab7 (1:50; clone: E907E, Cell Signaling), mouse anti-human Rab5a (1:100; clone: 2E8B11, Thermofisher), mouse anti-human KDEL (1:100; SPA-827, Stressgen Technologies), and rabbit anti-human PAD4 (1:100; in-house generated). As a positive control for protein colocalization, cells were stained using rabbit anti-CD11b (1:100; clone: EPR1344, Abcam) and mouse anti-CD18 (1:50; clone: MEM-148, Thermofisher) antibodies. Coverslips were rinsed in PBS and incubated with 1:200 dilution of AF488 goat anti-mouse IgG1 (Invitrogen) or IgG2a (Invitrogen), or AF594 Plus donkey anti-rabbit (Invitrogen) secondary antibodies for 30 minutes at room temperature. Following rinses in PBS and dH₂O, coverslips were mounted using ProLong Gold AntiFade DAPI mounting medium (Thermofisher) and sealed. Samples were imaged using a Zeiss AxioObserver with 880-Quasar confocal module & Airyscan FAST module, and images were converted using ZEN microscopy software (Zeiss). For colocalization analysis, four to five random fields were captured at 20X magnification, and colocalizations were quantified using Volocity version 7.0.0.

Flow cytometry. Treated and untreated cells were stained for viability using blue fluorescent reactive dye (Invitrogen). Cells were washed and blocked using human BD Fc block (BD Biosciences) in 3% BSA in PBS and stained with anti-CD11c-BV421 (clone: Bu15; Biolegend). In-house generated rabbit anti-human PAD4 antibody was fluorescently labelled using AF647 antibody labeling kit (Invitrogen) and used at 0.01 µg/µL. Equal

amounts of AF647-labeled rabbit IgG (3452S, Cell Signaling Technology) served as isotype control. Cells were incubated with labelled antibodies for 30 minutes at 4°C and flow cytometry was conducted at the Johns Hopkins Bayview Immunomics Core Facility using Cytex Aurora (Cytex Biosciences). Data was analysed using FCS Express 7 Research Edition (De Novo Software).

3.3 RESULTS

3.3.1 PAD4 is present in vesicular organelles in monocytes

While PAD4 has been documented in the cytosol, plasma membrane, and nucleus of both granulocytes and monocytes, the mechanism by which PAD4 traffics to these localizations is still unknown. To better understand the subcellular localizations of PAD4, we fractionated monocytes from three healthy donors into various subcellular compartments and immunoblotted for the presence of PAD4 (**Figure 3-1A**). Purity of the subcellular fractions was demonstrated by immunoblotting for markers predominantly found in these localizations.

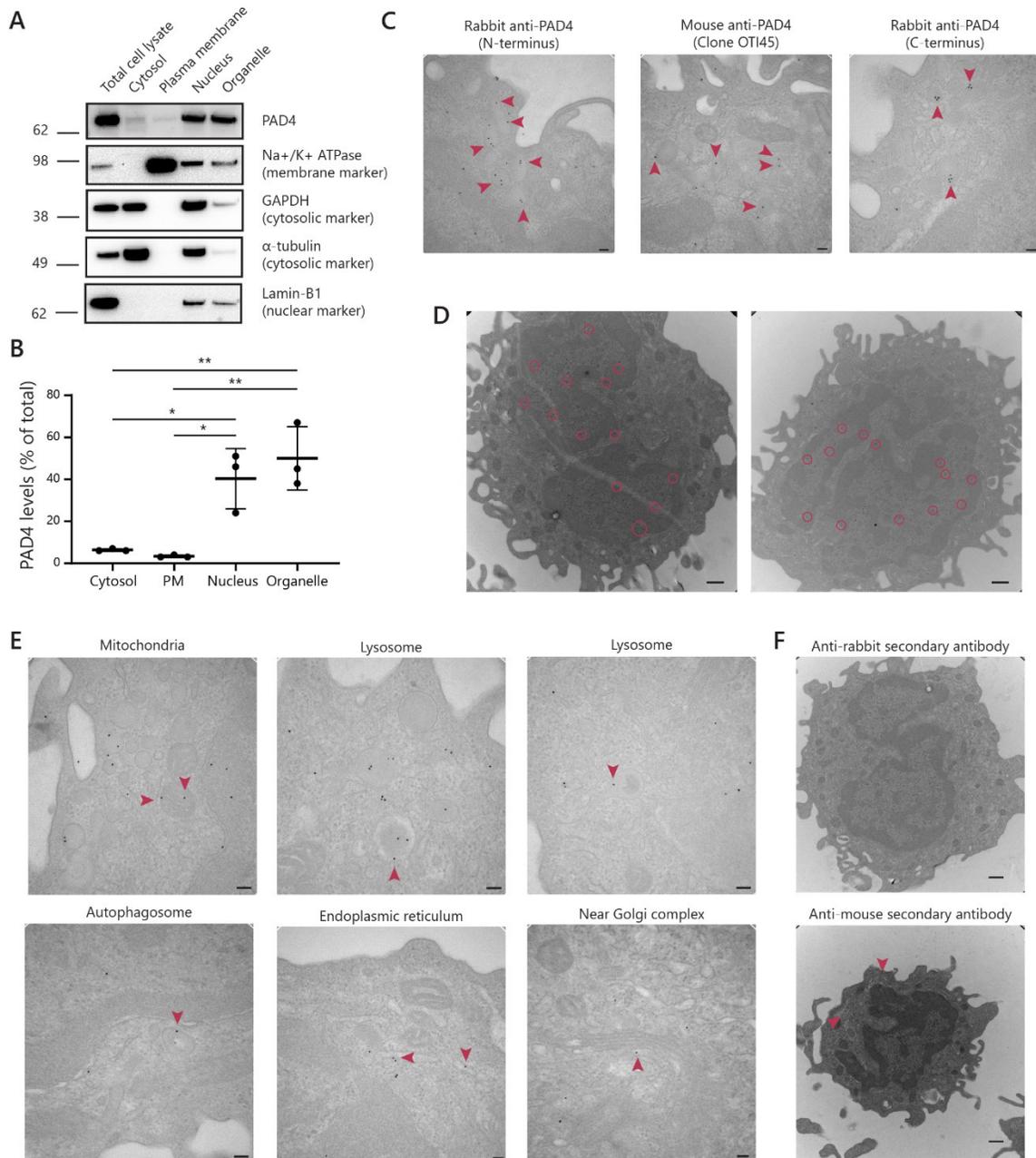


Fig. 3-1. PAD4 is found in the organelle fraction in monocytes. (A and B) Monocytes isolated from peripheral blood of three healthy donors were subjected to subcellular fractionation and immunoblotted. **(A)** Representative blot of PAD4 and Na⁺/K⁺ ATPase (plasma membrane marker), GAPDH and α-tubulin (cytosolic markers), and lamin-B1 (nuclear marker) in the subcellular fractions from monocytes shown. **(B)** PAD4 levels from monocytes isolated from the peripheral blood of three healthy donors were determined by dividing the amount of PAD4 expressed in a fraction by the sum of total PAD4 expression in all the subcellular fractions and compared using one-way ANOVA with Tukey's multiple comparisons test. Error bars represent mean ± standard deviation. **(C, D, E, and F)**

Monocytes isolated using negative selection from a single healthy donor were stained for transmission electron microscopy using gold-labelled secondary antibodies against PAD4 targeting three distinct epitopes. **(C)** Red arrowheads point to vesicular localizations of PAD4. Scale bars represent 100 nm. **(D)** Transmission electron microscopy images demonstrating nuclear expression of PAD4 (red circles). Scale bars represent 500 nm. **(E)** Transmission electron microscopy images depicting organelle and nuclear localizations of PAD4 as marked by the red arrowheads. Scale bars represent 100 nm. **(F)** Transmission electron microscopy images of secondary only staining marked by red arrowheads. Scale bars represent 500 nm.

PAD4 was detected in all cellular compartments including cytosol, plasma membrane, nucleus, and organelles. While differences in PAD4 levels between the cytosol and the plasma membrane fractions were minimal, levels of PAD4 in the organelle fraction were significantly higher when compared to PAD4 levels in the cytosol ($p = 0.0397$) and the plasma membrane ($p = 0.0350$; **Figure 3-1B**). The levels of PAD4 in the nuclear fraction were also significantly increased compared to the amount of PAD4 in the plasma membrane ($p = 0.0446$), as expected. However, there was no significant difference between PAD4 levels in the nuclear and the organelle fractions indicating that organelles are a previously unappreciated reservoir for PAD4. To visualize the organelle localization of PAD4 in more detail, monocytes isolated from healthy donors were imaged via transmission electron microscopy using antibodies against three different epitopes of PAD4 (**Figure 3-1C**; secondary only staining can be seen in **Figure 3-1F**). Imaging confirmed a nuclear localization of PAD4 in monocytes, in accordance with previous studies (**Figure 3-1D**). Electron microscopy also revealed the frequent presence of PAD4 in vesicular organelles, with occasional staining observed in subcellular localizations suggestive of lysosomes, endoplasmic reticulum (ER), and mitochondria, regardless of the antibody used, confirming the enrichment of PAD4 in vesicular organelles (**Figure 3-1E**).

Immunofluorescence of permeabilized monocytes further validated colocalization of PAD4 with lysosomes (Lamp1; average Pearson's coefficient = 0.5758; **Figure 3-2A and C**) and an even greater extent with the ER (KDEL; average Pearson's coefficient = 0.6595; **Figure 3-2B and C**).

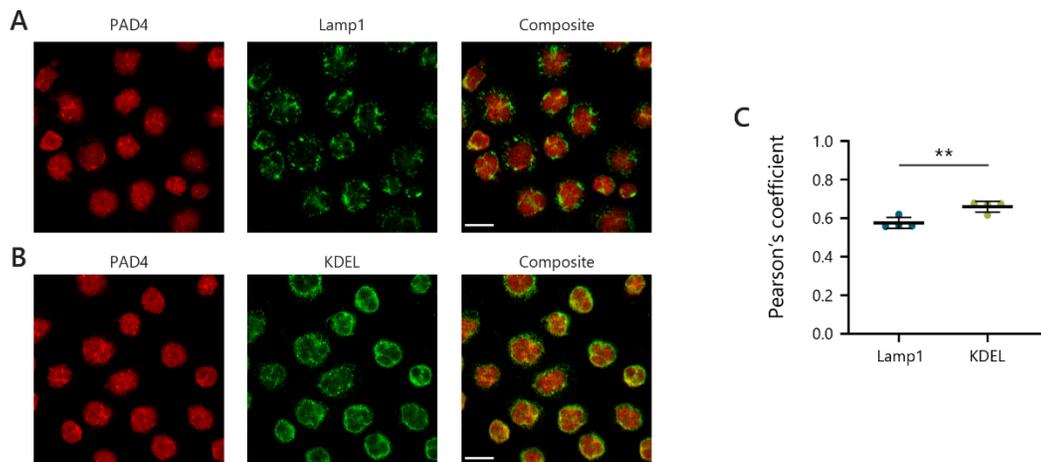


Fig. 3-2. PAD4 is detected in organelles of monocytes. (A-B) Permeabilized monocytes stained for PAD4 (red) and Lamp1 (green; **A**) or KDEL (green; **B**). Scale bars represent 10 μ m. **(C)** Colocalization between PAD4 and organelle markers in monocytes was analysed by averaging the Pearson's coefficients for four fields of view. Statistical analysis was performed using Student's t test. Error bars represent mean \pm standard deviation. ** = $p \leq 0.01$.

3.3.2 PAD substrates are enriched in vesicular organelles and processes

To further elucidate the organelle localization of PAD4 and to examine whether citrullination of organelle proteins could be detected, monocytes from three healthy donors were incubated in calcium, fractionated, and assessed for proteins in the isolated

organelle fractions by mass spectrometry. We first searched the identified proteins for the presence of peptides derived from PAD4 or other PAD isoenzymes, since monocytes have been reported to also express PAD2 in some studies.^{106,202} PAD4 peptides were identified in the organelle fraction of monocytes spanning 34% coverage of the PAD4 protein sequence. Interestingly, PAD2 was the only other PAD family member identified, with peptides spanning 9% of the PAD2 sequence detected. PAD2 was slightly less abundant in organelles than PAD4, and this trend was confirmed in whole monocyte lysate by immunoblotting (**Figure 3-3 A-C**).

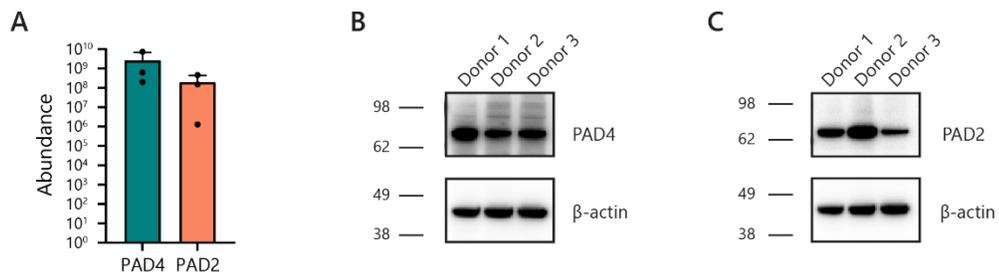


Fig. 3-3. PADs in monocytes. (A) Abundance of PAD4 and PAD2 in monocytes as determined via mass spectrometry. Error bars represent mean \pm standard deviation. (B-C) Western Blot of PAD4 (B) and PAD2 (C) in monocyte lysate from three healthy donors and their respective β -actin levels.

Mass spectrometry results were subsequently filtered on proteins with citrullination of arginines determined with high confidence (see workflow in **Figure 3-4A**). This analysis identified 271 proteins containing citrulline residues. Analysis of their citrullination site sequences revealed an overrepresentation of prolines in the P1' (frequency = 43.53%, $p =$

3.54e-80) and P2' (frequency = 12.30%, $p = 0.006$) sites and an enrichment of aspartic acid (P4, P2, P1', and P2' sites) and glutamic acid (P2' and P3' sites) residues surrounding the citrulline (**Figure 3-4B**). Of the 271 proteins found to be citrullinated, nine did not have an abundance above the detection limit and were discarded from subsequent analysis. Of the remaining 262 citrullinated proteins, 98.1% were shared by all three donors (**Figure 3-4C**). Very few were donor-specific with 0.8% (2/262) being unique to Donor 1, 0% (0/262) found only in Donor 2, and 0.4% (1/262) seen only in Donor 3. The subcellular localizations of the 257 citrullinated organelle proteins shared by all three donors were then analysed using the UniProt database and proteins predominantly assigned to nuclear, and cytosolic locations were excluded from subsequent analysis. Statistical enrichment analysis of citrullinated organelle proteins against the human proteome using the PANTHER database revealed multiple subcellular localizations of citrullinated proteins throughout various organelles (**Figure 3-4D**). Especially predominant were proteins belonging to vesicles and the endomembrane system, which parallels the observations made using electron microscopy (**Figure 3-1C**). Proteins found in organelle membranes, mitochondria, vesicles, endosomes, and lysosomes were also significantly enriched. Interestingly, proteins found in the ER and Golgi were also citrullinated, suggesting an enrichment of PADs within components of the secretory pathway.

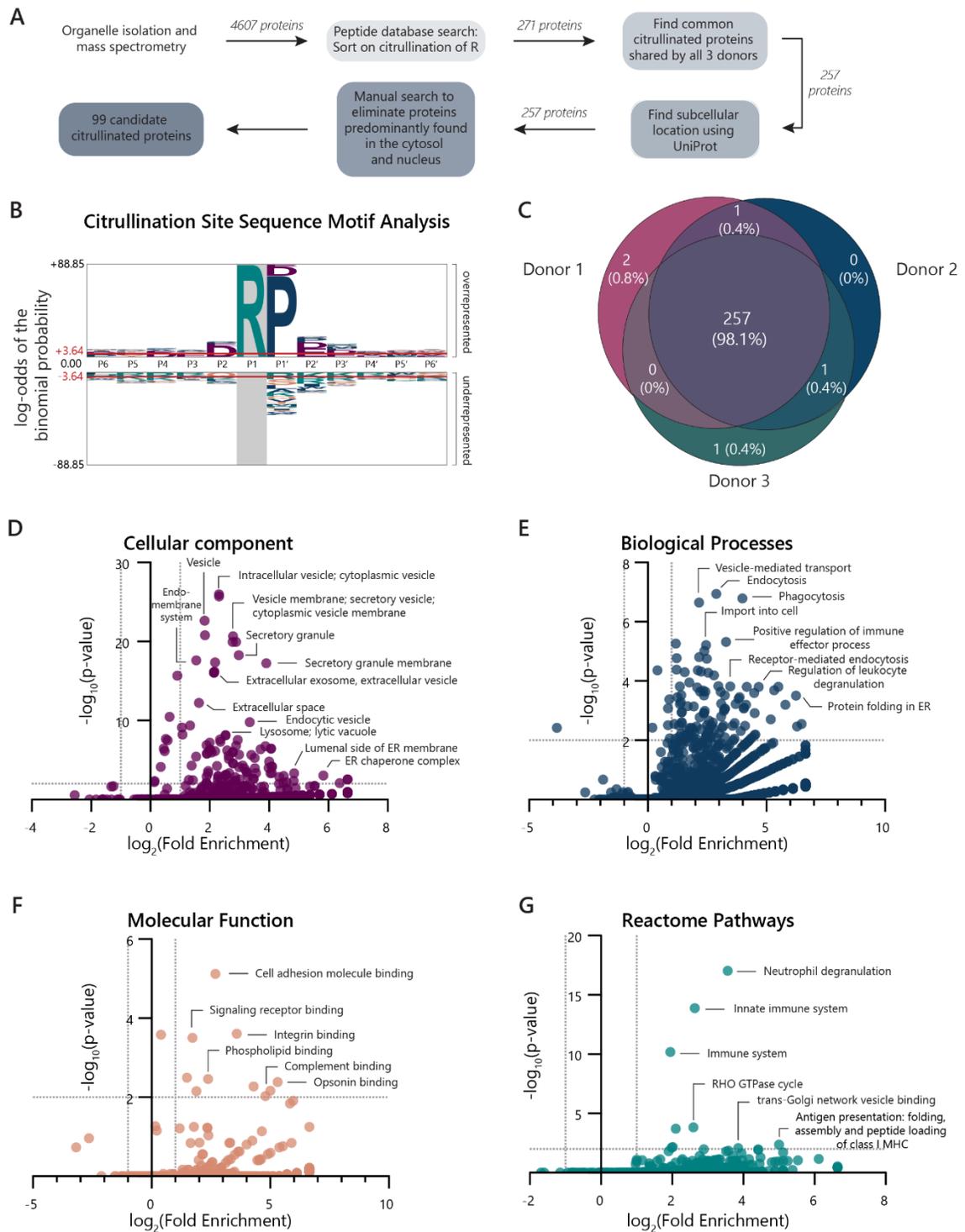


Fig. 3-4. Citrullinated proteins are enriched in organelles of monocytes. (A) Workflow of mass spectrometry data analysis. **(B)** PAD4 citrullination site amino acid sequence motif analysis (at positions P6-P6') for all sites containing deiminated arginine residues detected by mass spectrometry in subcellular organelle fractions using pLogo. Log-odds equal to ± 3.64

represent a p-value < 0.05, and red lines indicate significance thresholds. The size of each letter is proportional to the log-odds binomial probability of that amino acid at its respective position, shown as over- or underrepresented at citrullination sites compared to the human proteome. **(C)** Venn-diagram of citrullinated organelle proteins identified by mass spectrometry on organelle fractions of monocytes (n = 3 donors) incubated in 5 mM calcium that were unique to or shared by donors. **(D – G)** PANTHER statistical overrepresentation analysis of citrullinated organelle proteins compared to the human proteome. log₂ of fold enrichment versus -log₁₀ of FDR-corrected p-values shown. Significance thresholds were set at ±2 fold enrichment and a p < 0.01. Factors with fold enrichment greater than 0.01 are shown. * = p < 0.05, ** = p ≤ 0.01

Analysing the enrichment of proteins participating in various biological processes revealed a dominant presence of citrullinated proteins involved in trafficking of vesicles, endocytosis, phagocytosis, and protein folding in the ER (**Figure 3-4E**). Analysing citrullinated proteins involved in various molecular functions revealed a significant abundance of proteins involved in adhesion, especially those that participate cell adhesion and integrin binding (**Figure 3-4F**). Intriguingly, proteins involved in signalling receptor, phospholipid, complement, and opsonin binding were also enriched, indicating citrullination of proteins trafficking to and present at or near vesicular and plasma membranes. Proteins involved in vesicle-mediated transport to various organelles, membrane trafficking, and antigen presentation were also significantly enriched when analysing citrullinated proteins in reactome pathways (**Figure 3-4G**). Interestingly, neutrophil degranulation and pathways related to the innate immune system were also overrepresented. These data indicate that citrullination of several organelle proteins can be detected in monocytes, suggesting involvement of PADs in diverse cellular functions, with an enrichment in secretory pathway vesicles and functions.

3.3.3 PAD4 binds to protein partners found at various organelle compartments

The presence of PAD4 in vesicles and citrullination of proteins involved in secretory pathway organelles suggested that PAD4 may interact with protein binding partners in different subcellular locations to mediate its trafficking throughout the cell, possibly via an unconventional trafficking mechanism. Since PAD4 lacks conventional ER import signals and membrane-association domains, we hypothesized that PAD4 might be trafficking via binding to protein partners that escort it to various extranuclear compartments including secretory vesicles. To identify putative protein binding partners that may be involved in PAD4 trafficking, we analysed PAD4 binding to 16,000 human proteins using the HuProt protein microarray. Because PAD4 is a calcium-dependent enzyme, which may interact differentially with binding partners in the apo- or calcium bound conformations,⁹⁵ we probed for putative binding partners in the presence of 0, 0.2, and 2 mM calcium.

Analysis of putative PAD4 protein binding partners, which bound with an A score >2, revealed 234 candidates in 0 mM calcium, 280 candidates in 0.2 mM calcium, and 298 candidates in 2mM calcium (**Figure 3-5A**). Subsequent pathway analysis was then performed focusing on the 180 proteins that were bound in all calcium conditions (**Figure 3-5B**). Querying statistical overrepresentation of the 180 putative PAD4 binding partners in various cellular localizations compared to the entire human proteome revealed a significant enrichment of proteins found in the cytosol, extracellular environment, and cell junctions (**Figure 3-5C and D**). Interestingly, proteins found in the peroxisomal matrix, microbody lumen, and vesicles were also enriched, suggesting a trafficking mechanism via binding to transport proteins. Analysing statistical overrepresentation of biological processes revealed that proteins involved in metabolic processes, glycolysis in particular,

were especially enriched as PAD4 binding partners (**Figure 3-5E and F**). Proteins with several catalytic activities were also enriched as binding partners of PAD4 (**Figure 3-5G and H**). This included proteins capable of kinase activity, transferase activity, lyase activity, and oxidoreductase activity. Proteins involved in metabolic processes such as biological oxidations, glycolysis, peroxisome protein import, and synthesis of bile acids were especially enriched in the reactome pathways analysis (**Figure 3-5I and J**). Corroborating this evidence was the significant enrichment of metabolite interconversion enzymes, oxidoreductases, reductases, and dehydrogenases as binding partners of PAD4 (**Figure 3-5K and L**).

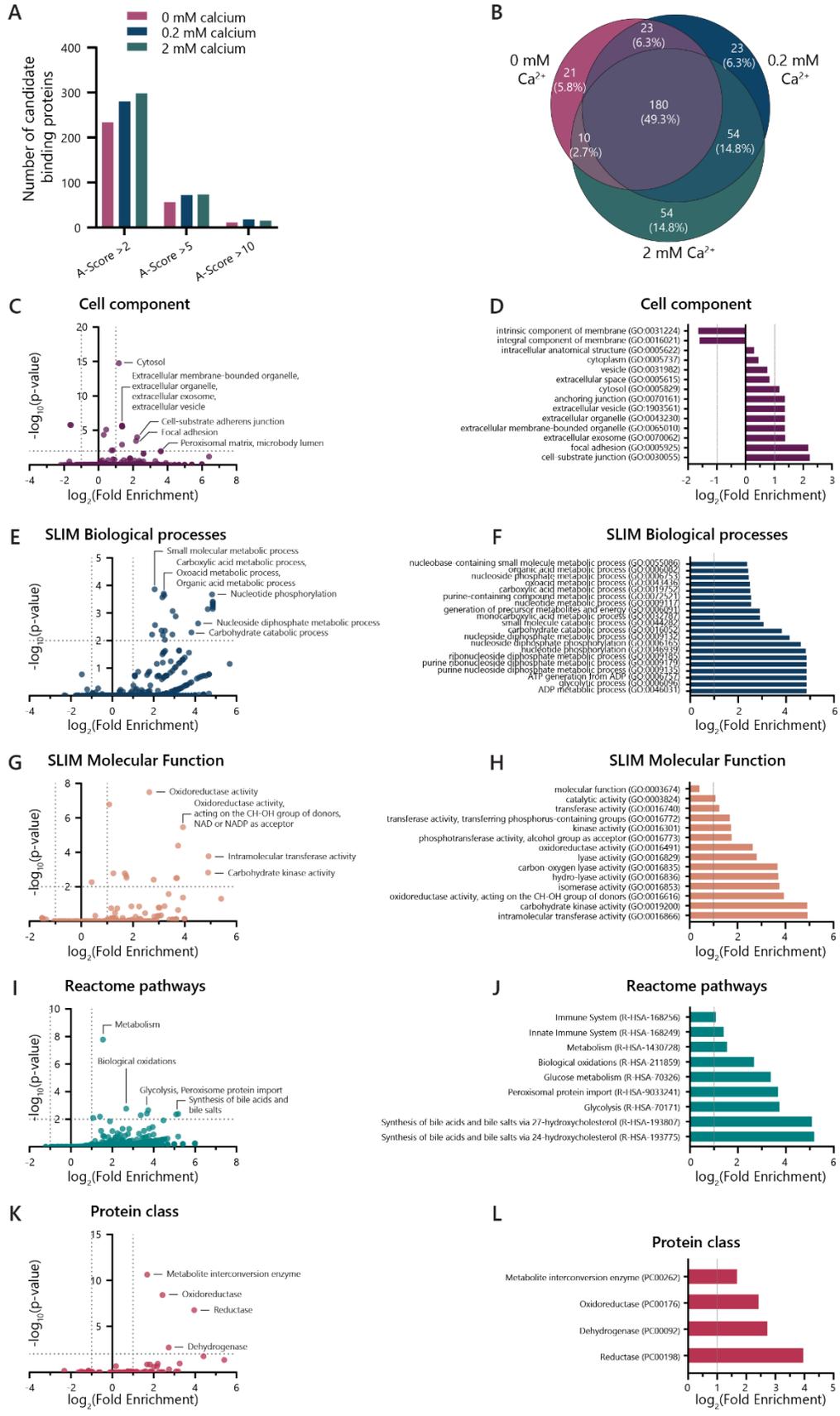


Fig. 3-5. PAD4 binds protein partners found throughout the cell. (A) Binding partners of PAD4 identified from HuProt protein microarray under 0 mM, 0.2 mM, and 2 mM calcium conditions versus their A-scores. **(B)** Venn diagram of putative binding partners identified under 0 mM, 0.2 mM, and 2 mM calcium conditions. PANTHER statistical overrepresentation analysis of PAD4 binding partners compared to the human proteome. **(C, E, G, I, K)** \log_2 of fold enrichment versus $-\log_{10}$ of FDR-corrected p-values shown. Significance thresholds were set at ± 2 -fold enrichment and a $p < 0.01$. Factors with fold enrichment greater than 0.01 are shown. Top twenty factors with $p < 0.01$ for **(D)** Cell component, **(F)** SLIM Biological processes, **(H)** SLIM Molecular Function, **(J)** Reactome pathways, and **(L)** Protein class are shown. Fold enrichment cut-off values of ± 2 are indicated by dotted grey lines.

A deeper analysis revealed several known protein binding partners and substrates of PAD4 including NF- κ B1, nucleophosmin (NPM1), and enolase (ENO1),^{101,182,203} in addition to novel candidate surface binding partners with known expression on the monocyte surface including CEACAM1, OSMR, and C5AR1 **(Figure 3-6A and B)**.²⁰⁴⁻²⁰⁶ In addition, multiple Rab GTPase proteins involved in intracellular vesicular trafficking and signalling were identified as putative PAD4 binding partners including RAB7L1, RABL3, and the Rab regulatory protein GDI2 as well as proteins involved in autophagy (ATG4C) and calcium dependent membrane binding (CPNE2) **(Figure 3-6C)**.²⁰⁷⁻²⁰⁹

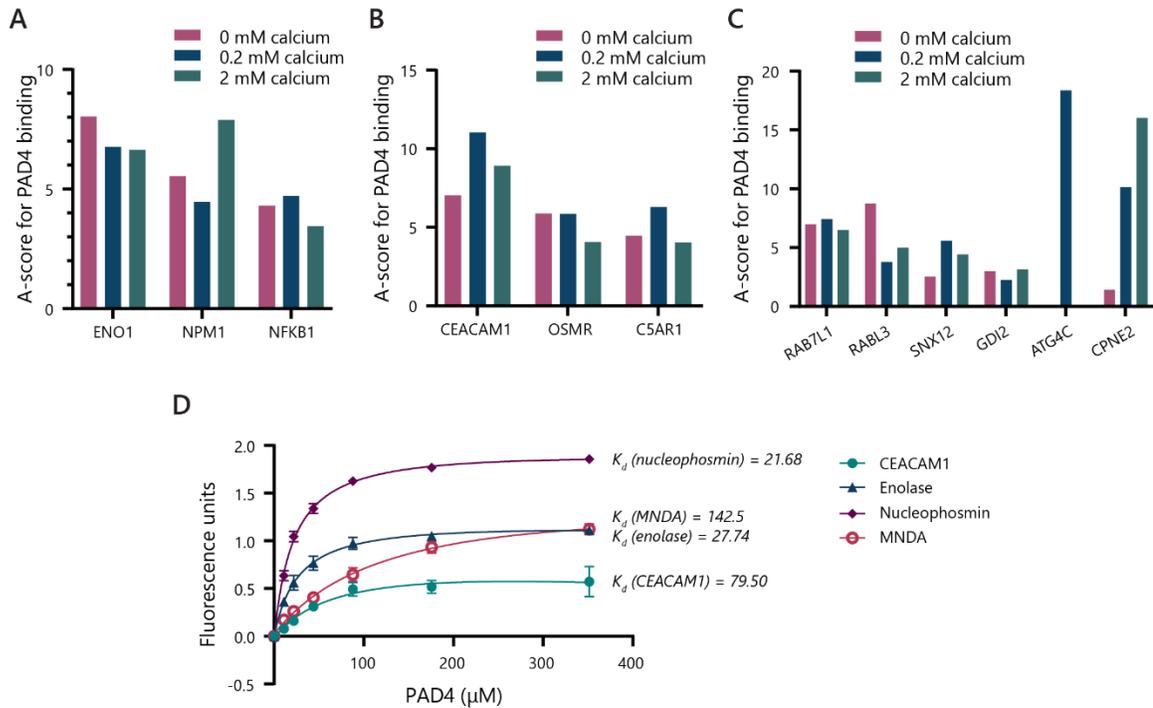


Fig. 3-6. PAD4 binding partners. (A) Known PAD4 substrates identified as binding partners. **(B)** Candidate cell surface binding partners of PAD4 and their A-scores under increasing calcium conditions. **(C)** Proteins involved in autophagy, intracellular vesicular trafficking, and membrane binding identified to be putative PAD4 binding partners with varying A-scores under increasing calcium conditions. **(D)** Binding affinities of PAD4 with putative binding partners. Varying molar concentrations of rhPAD4 was added to wells coated with CEACAM1 (teal), nucleophosmin (purple), α -enolase (blue), and myeloid cell nuclear differentiation antigen (MNDA; red). Fluorescence units corresponding to PAD4 binding is shown with calculated binding constants. Error bars represent mean \pm standard deviation.

To validate the HuPROT chip array findings, we determined affinities of PAD4 to a subset of proteins identified as putative binding partners *in vitro* using an ELISA. We found PAD4 to have increased affinity for nucleophosmin ($K_d = 21.68$), enolase ($K_d = 27.74$), and CEACAM1 ($K_d = 79.50$) compared to myeloid cell nuclear differentiation antigen (MNDA; $K_d = 142.5$), which was not found to interact with PAD4 on the HuPROT array (**Figure 3-6D**). Together, these data suggest that PAD4 is capable of binding to diverse classes of

proteins, with a preference for those with metabolic activity and transport functions across cellular compartments.

3.3.4 Monocyte derived-dendritic cells serve as a model to study cell surface PAD4 trafficking

Because electron microscopy imaging and unbiased analysis of PAD4 substrates and binding partners revealed a putative role for PAD4 in intracellular vesicular compartments in monocytes, we utilized the cellular model of monocyte differentiation into mo-DCs to further understand the mechanisms involved in PAD4 trafficking. To first interrogate the surface expression of PAD4 during monocyte differentiation, monocytes were differentiated into mo-DCs for 6 days and the surface PAD4 expression was assessed using flow cytometry at days 0 (D0), 3 (D3), and 6 (D6). Increased surface expression of PAD4 during the differentiation period was observed with the greatest surface expression seen on D6 mo-DCs (**Figure 3-7A and B**). Differentiation into mo-DCs was confirmed by the increased expression of CD11c (**Figure 3-7C**). Binding of the anti-PAD4 antibody to the cell surface was significantly higher than binding of the isotype control antibody to D0 monocytes ($p = 0.0339$), D3 mo-DCs ($p = 0.0351$), and D6 mo-DCs ($p = 0.0215$), indicating the presence of PAD4 on the surface of monocytes and mo-DCs (**Figure 3-7A and B**). While D3 mo-DCs had similar levels of surface PAD4 to D0 monocytes ($p = 0.5052$), D6 mo-DCs had significantly higher expression of surface PAD4 ($p = 0.0132$ compared to D0 and $p = 0.0093$ compared to D3). Confocal imaging further revealed differences in PAD4

subcellular localization at different stages of monocyte differentiation to mo-DCs (**Figure 3-7D**). D0 monocytes and D3 mo-DCs had primarily extranuclear expression of PAD4. However, D6 mo-DCs had a highly polarized expression of PAD4 toward the cell surface, most commonly located apical to the nucleus.

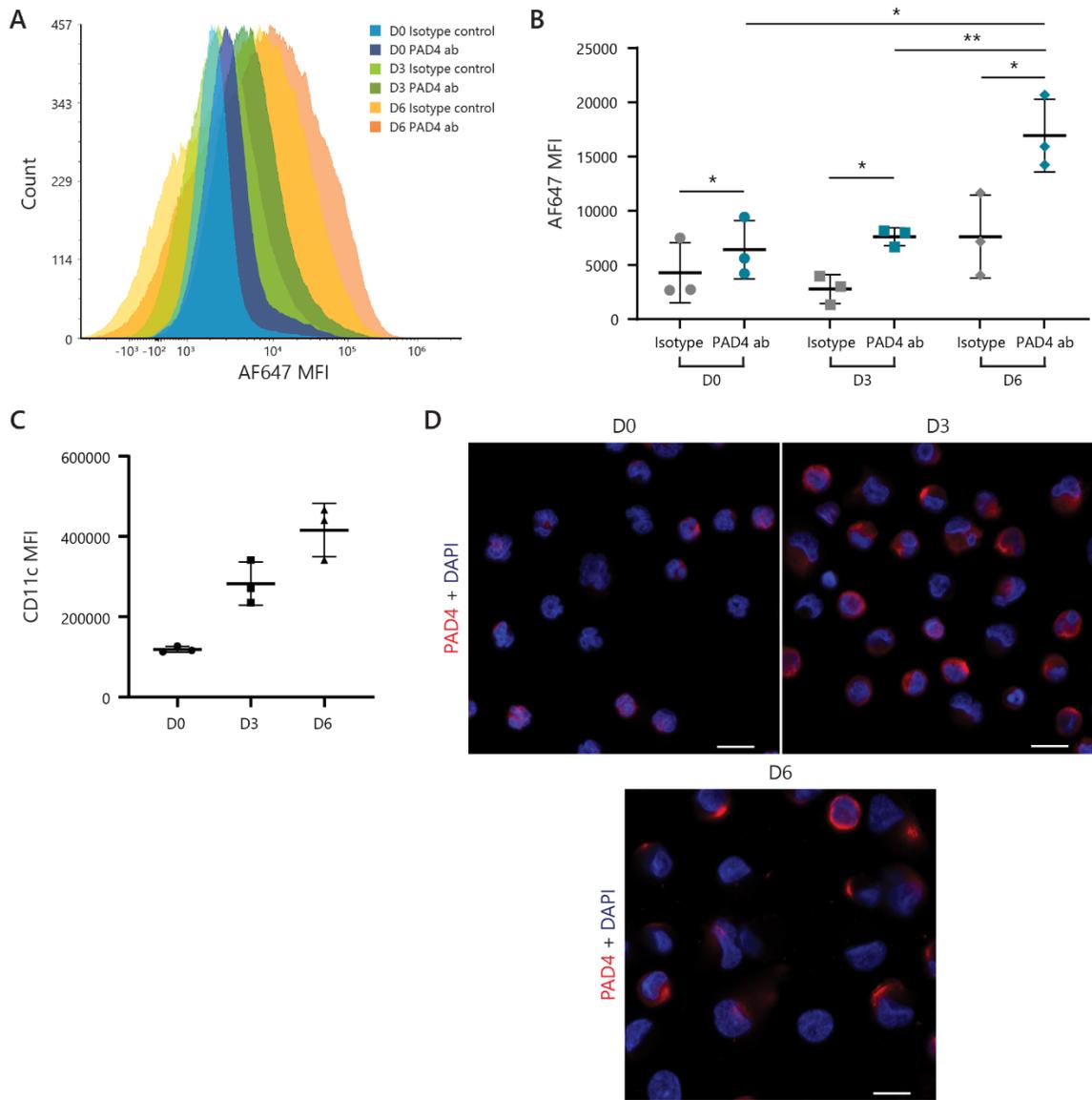


Fig. 3-7. PAD4 localization in mo-DCs. (A–B) Monocytes were differentiated into mo-DCs for six days and surface PAD4 levels were analysed using flow cytometry. **(A)** Representative histogram of surface PAD4 expression during mo-DC differentiation. **(B)** Quantitative analysis of mean fluorescence intensity (MFI) of isotype control (grey) or PAD4 antibody (blue) during mo-DC differentiation. MFI between isotype control and PAD4 antibody were compared using paired t-tests while MFI of PAD4 antibody between mo-DCs of different culture days were compared via unpaired t-tests. Error bars represent mean \pm standard deviation. **(C)** CD11c mean fluorescence intensity (MFI) levels on D0, D3, and D6 mo-DCs. **(D)** PAD4 (red) distribution in D0, D3, and D6 mo-DCs with nuclei counterstained with DAPI (blue). Scale bars represent 10 μ m.

Since D6 mo-DCs had the greatest expression of surface PAD4 and a polarized cellular distribution, they were used as a cellular system to further evaluate the mechanism of PAD4 trafficking, with a focus on evaluating secretory, endolysosomal, and autophagy pathways.²¹⁰ Confocal imaging was performed to visualize and quantify colocalization of PAD4 with canonical endolysosomal markers in D6 mo-DCs using the Pearson's correlation coefficient (**Figure 3-8A-F, Figure 3-9**). Pearson's correlation coefficient is a measure of the overlap between two markers of interest with a coefficient of 1 indicating perfect overlap and values close to 0 suggesting no correlation between the two signals of interest.²¹¹ Lamp1 staining, a marker of lysosomes, could be seen as large clusters present in D6 mo-DCs (**Figure 3-8A, Figure 3-9**). Although PAD4 and Lamp1 exhibited some colocalization (average Pearson's coefficient = 0.5314), cells with distinct localization of PAD4 or Lamp1 alone can also be seen (**Figure 3-8A and F, Figure 3-9**). Extranuclear localization of Rab5 (indicator of early endosomes and regulator of early autophagosome formation²¹⁰) was seen to be dispersed throughout the cell, and PAD4 and Rab5 colocalized to a greater degree with an average Pearson's coefficient of 0.6086 (**Figure 3-**

8B and F, Figure 3-9). Similar to Rab5, staining of Rab7 (a marker of late endosomes and mediator of autophagosome maturation²¹⁰) was seen to be more similar to PAD4 staining with the greatest average Pearson's coefficient of 0.6442 (**Figure 3-8C and F, Figure 3-9**). Interestingly, staining using KDEL to visualize proteins retained in the ER resulted in small punctate spots spread throughout the cytosol of mo-DCs (**Figure 3-8D, Figure 3-9**). Moderate colocalization of PAD4 and KDEL could also be seen with a mean Pearson's coefficient of 0.6082 (**Figure 3-8F, Figure 3-9**). Strong cell surface and some intracellular compartmentalization staining could be seen with HLA-DR, however, the average Pearson's colocalization coefficient was only 0.1750, suggesting minimal colocalization with PAD4 (**Figure 3-8E and F, Figure 3-9**). Analysing the average Pearson's coefficients of all markers with PAD4 revealed significantly stronger colocalization of PAD4 with Rab5 ($p = 0.0010$), Rab7 ($p < 0.0001$), and KDEL ($p = 0.0011$) compared to Lamp1 and for all markers compared to HLA-DR ($p < 0.0001$). To confirm that these findings remained accurate while using conditions that permitted more thorough cellular permeabilization, we repeated this analysis using PFA fixed and acetone permeabilized cells and found that the overall spatial relationship between PAD4 and markers of subcellular localizations remained largely unchanged with increased Pearson's coefficients (**Figure 3-10A-G**). Taken together, our data indicates the presence of PAD4 in endolysosomal compartments involved in autophagy as well as in conventional secretory compartments.

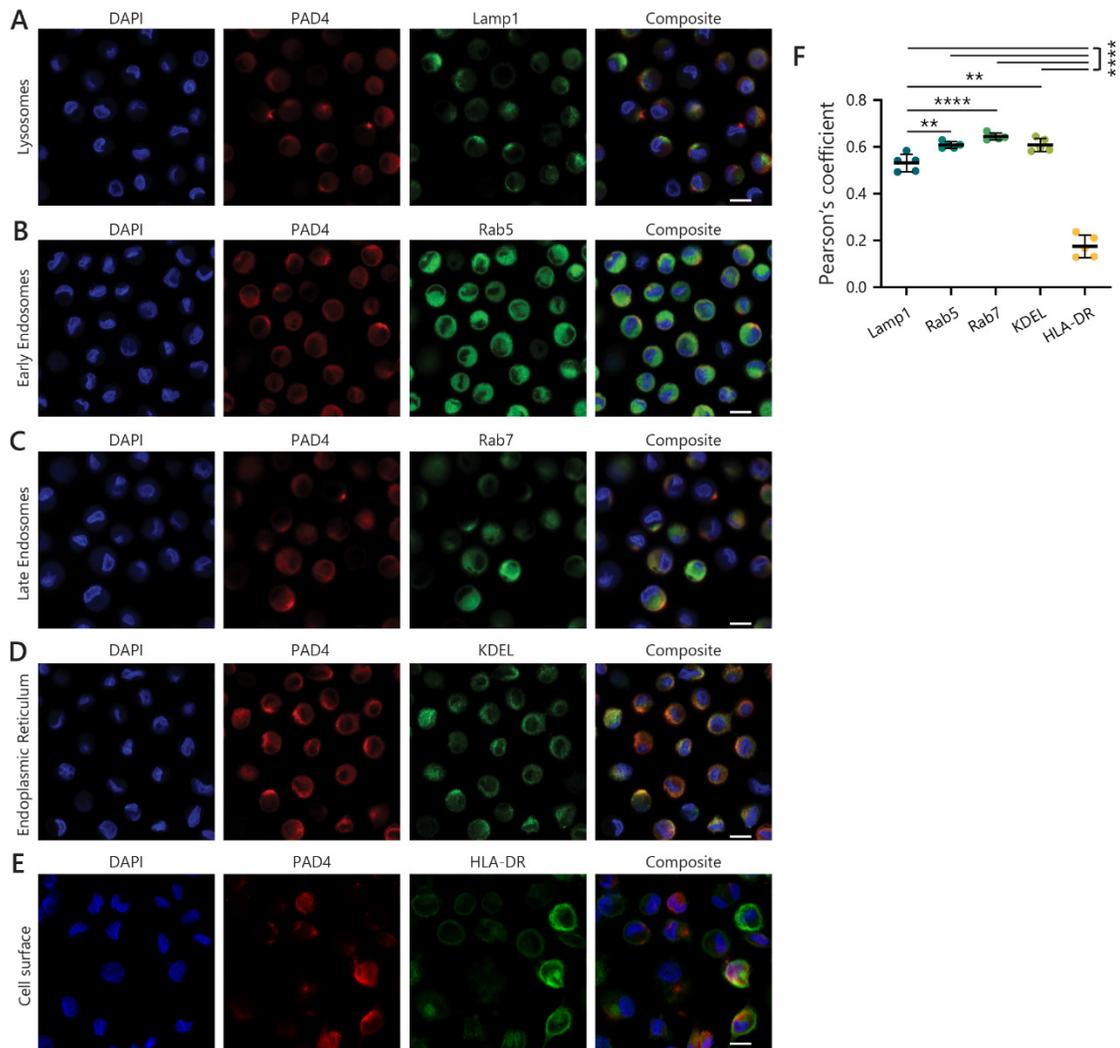


Fig. 3-8. PAD4 associates with autophagic and exocytic organelles in mo-DCs. (A-E) mo-DCs after six days of culture were allowed to adhere to coverslips and co-stained for nuclei (blue), PAD4 (red), and various organelle markers (green). Scale bars represent 10 μ m. (F) Colocalization between PAD4 and organelle markers were analysed by averaging the Pearson's coefficients for five representative fields of view. Statistical analysis using one-way ANOVA and Tukey's multiple comparisons test was performed. Error bars represent mean \pm standard deviation. * = $p < 0.05$, ** = $p \leq 0.01$, **** = $p \leq 0.0001$.

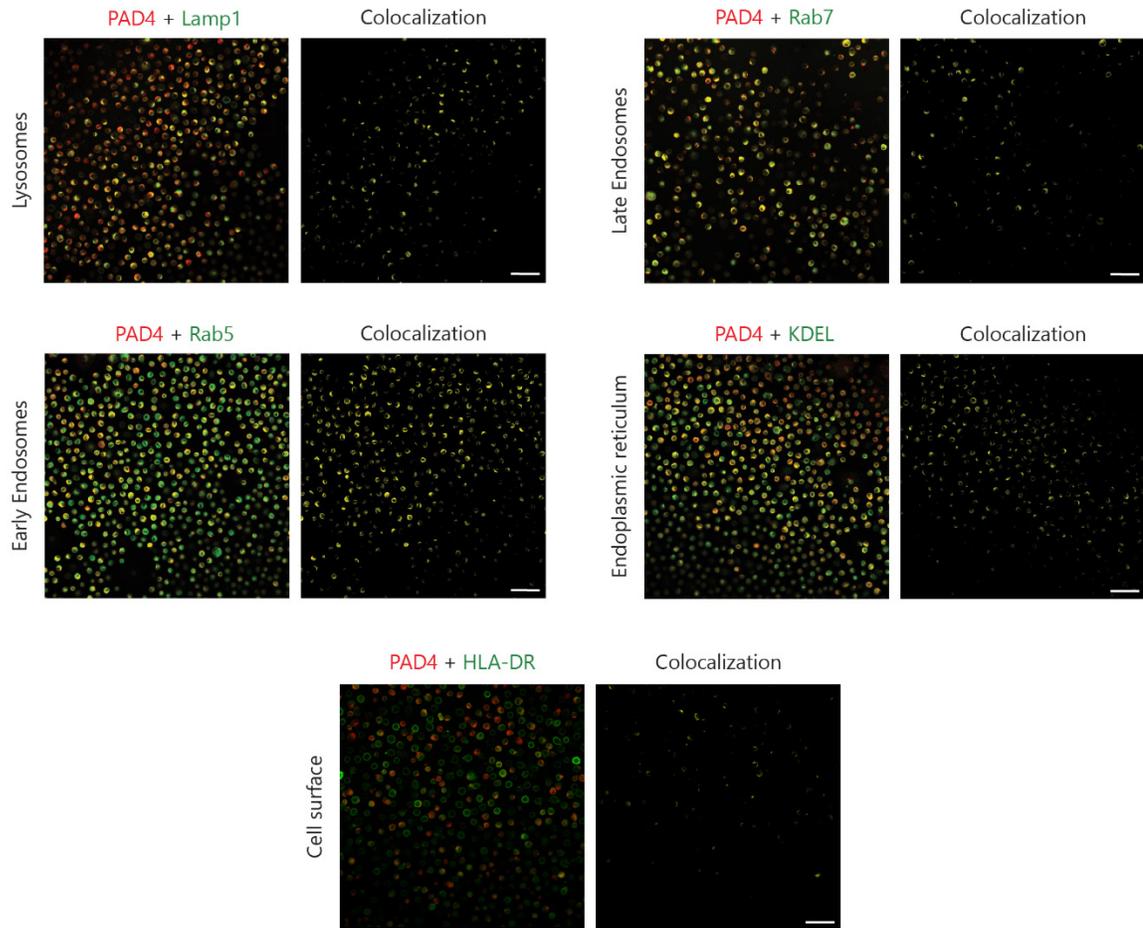


Fig. 3-9. Larger field of view depicting colocalization of PAD4 with autophagic and exocytic organelles in mo-DCs. Representative images of PAD4 (red) with various endolysosomal compartments, ER, and HLA-DR (green) visualized using a large field of view for colocalization analysis. *Right panels:* Colocalization of PAD4 with organelle markers shown as a separate channel in yellow. Scale bars represent 50 μ m.

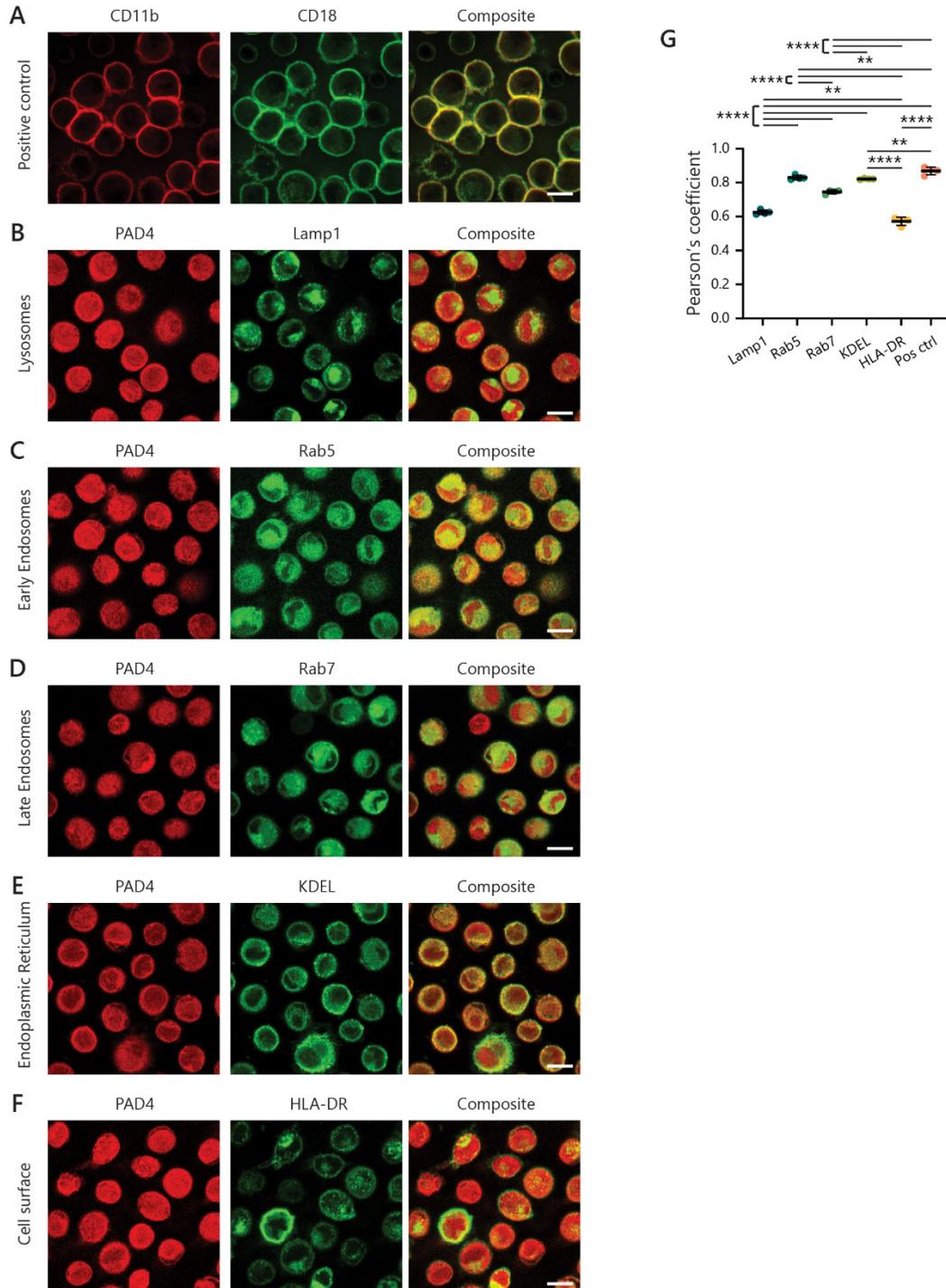


Fig. 3-10. PAD4 and organelle colocalization in permeabilized mo-DCs. Mo-DCs after six days of culture were allowed to adhere to coverslips and permeabilized using acetone. **(A)** Cells were co-stained for CD11b (red) and CD18 (green) as a positive control. **(B-F)** mo-DCs were co-stained for PAD4 (red), and various organelle markers (green). Scale bars represent 10

μm . **(G)** Colocalization analysis was performed by averaging the Pearson's coefficients for four representative fields of view. Statistical analysis was done using one-way ANOVA and Tukey's multiple comparisons test. Error bars represent mean \pm standard deviation. * = $p < 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

3.4 DISCUSSION

While PAD4 has historically been categorized as a primarily nuclear enzyme, recent evidence has revealed its presence in other subcellular compartments including the cytosol and the cell surface.^{128,129} A major question that remains unanswered is the mechanism by which PAD4, containing a nuclear localization signal but lacking a canonical secretory signal sequence, traffics to the plasma membrane of neutrophils and monocytes.¹⁰⁰ In this comprehensive study, we observed that PAD4 associates with proteins found in a variety of subcellular localizations with unexpected enrichment and activity in vesicular organelle compartments. We also found that PAD4 expression on the cell surface increases as monocytes differentiate into mo-DCs and illuminated a probable trafficking mechanism via the endolysosomal and autophagic pathways.

Early studies conducted on PAD4 in granulocytes and transfected HeLa cells were the first to report the presence of a nuclear localization signal sequence and subsequent nuclear localization of PAD4.¹⁰⁰ In support of this observation, various nuclear proteins have been documented to be subject to citrullination in a physiologic environment including histones.^{100,102,182,180,91} However, the presence of PAD4 in the extranuclear space has been recently reported with PAD4 additionally found in cytosolic and cell surface localizations.^{128,129} We found that PAD4 is not only present in the cytosol, plasma membranes, and nuclei of monocytes, but also in the organelle fractions in relatively large quantities. Surprisingly, we also found the presence of PAD2 in monocytes from peripheral blood of healthy controls via mass spectrometry and Western Blot. While PAD4 was known to be present in monocytes, protein expression of PAD2 in these cells was unclear with

discrepant reports regarding its expression.^{99,106,202} Therefore, determining the localizations of PAD2 and its activity in monocytes should be subject to future studies to define the relative contribution of PAD2 and PAD4 to the organelle citrullinome in monocytes.

Our finding that PADs are capable of citrullinating various organelle proteins, including those involved in signalling, vesicular and protein trafficking, and a myriad of other biological processes implicates that PADs may play a larger role in physiologic monocyte cell function than previously appreciated. Our work also expands the human PAD-dependent citrullinome to include a large number of organelle proteins in addition to the cytosolic and nuclear proteins that were previously known. Understanding whether the citrullinated organelle proteins identified in this study are present in other subcellular localizations within monocytes and determining whether citrullination occurs at extranuclear and extracytosolic sites should be subject to further studies. Additionally, the effects of citrullination on the function and structure of organelle proteins and the role they play in a healthy versus disease states need to be elucidated further in future studies.

PAD4 expression in the cytosol may be explained by the process whereby nuclear proteins are translated by free ribosomes present in the cytoplasm prior to nuclear import and that nuclear versus cytosolic distribution of proteins can be regulated by post-translational modifications.^{212,213} However, the mechanism by which PAD4, which lacks ER import signal sequences and transmembrane domains, traffics to the cell surface and is retained there has been a mystery. While PAD4 is known to influence the function of cytoplasmic proteins (e.g. NADPH oxidase complex) and the translocation of nuclear

proteins (e.g. Elk-1),^{128,190} our data shows that PAD4 may have a larger array of binding partners, which may influence its multiregional localization within the cell. We found, via an unbiased evaluation of organelle substrates and binding partners, that PAD4 can associate with proteins from several subcellular localizations, including the plasma membrane, cytosol, vesicles, and cell junctions. Proteins participating in a wide range of metabolic activities, including glycolysis, were also statistically overrepresented as binding partners of PAD4, suggesting a possible novel role of PAD4 in regulating metabolism in cells. Further work is needed to validate these putative binding partners and to evaluate how PAD4 influences the functional aspects of these proteins in physiologic and pathologic conditions.

Monocytes are a major infiltrating leukocyte in the rheumatoid joint and are high expressors of PAD4 with the capacity to differentiate into macrophages, osteoclasts, and mo-DCs.^{141,214} Mo-DCs play a key role in sampling the environment and presenting antigens to T cells at sites of inflammation, and previous studies had implicated associations between autophagy and PAD4-dependent citrullination in antigen presenting cells.^{215,216} Our findings showed that surface PAD4 levels increase during mo-DC differentiation and that PAD4 distribution varies depending on the stage of differentiation. The colocalization of PAD4 with endolysosomal/autophagic vesicles and the ER implicate autophagy as well as the conventional secretory pathway, respectively, in PAD4 cellular trafficking. Importantly, the exact mechanisms of PAD4 association with the endomembrane system, the utilization of exocytosis and autophagy by PAD4, and

implications of dysregulated PAD4 trafficking for the pathogenesis of diseases associated with citrullination, such as RA, need to be further explored.

Our study sought to use unbiased proteomic approaches to shed light on the complexities and nuances of PAD4 expression within monocytes and its implications for citrullination. Taken together, our results suggest that PAD4 has numerous potential binding partners, which it may utilize to traffic to various subcellular localizations. Citrullination of organelle proteins implies that PAD4 is active in these subcellular regions and is perhaps being held in an active conformation via association with its binding partner(s). PAD4 may then be exploiting conventional exocytosis/endocytosis pathways in an unconventional manner, in addition to autophagic processes, to traffic and modify substrates leading to citrullination of a large and diverse set of organelle-associated proteins. Understanding the role of PAD4 in subcellular compartments in healthy and disease states is an important next step in identifying novel targets for preventing aberrant PAD4 activation in diseases marked by dysregulated protein citrullination.

CHAPTER 4

General conclusions

While the importance of PAD4 and citrullination in driving inflammation in RA is known, the mechanisms behind generation of citrullinated autoantigens have not been fully elucidated. In addition, recent evidence about the extranuclear localization of PAD4 has introduced new complexities into understanding the role of PADs and citrullination in physiology and pathology. Our studies revealed that citrullination is more widespread than previously appreciated with both catalytically active PAD4 and citrullinated proteins found on the surface of monocytes. Additionally, PAD4 was found to localize throughout various subcellular compartments in monocytes and bind to a wide range of protein partners from throughout the cell. Our findings, therefore, have important implications for developing methods to therapeutically target PAD4 and citrullination in RA.

Our findings implicate the monocyte surface as a novel site of autoantigen generation. We found PAD4 present on the surface of monocytes to not only be catalytically active and capable of citrullinating extracellular proteins, but the citrullinated proteins produced by surface PAD4 were found to be recognized by patient autoantibodies (**Figure 4-1**). We also observed the presence of citrullinated proteins on the cell surface of monocytes and identified several, previously unknown citrullinated plasma membrane proteins, including citrullinated Mac-1, an integrin critical for monocyte adhesion and activation. A subset of RA patients also harbored autoantibodies that recognized citrullinated Mac-1. These patients were found to have longer RA and elevated levels of inflammatory markers, especially IL-6, and may therefore represent a subpopulation that may benefit from treatment with IL-6 inhibitors.

The presence of citrullinated proteins on the surface of monocytes, as well as the finding that autoantibodies in patient sera can bind citrullinated cell surface proteins, suggest that monocytes may serve as cellular targets of ACPAs. Taken together with the evidence from previous studies that show monocyte differentiation into osteoclasts and mo-DCs after ACPA binding,^{143,144} our findings indicate that monocytes may play a larger role in promoting inflammation and joint damage in RA than previously thought. The presence of PAD4 on the monocyte surface and the prevalence of anti-PAD4 antibodies in a subset of RA patients with more severe disease suggests additional mechanisms by which monocytes exacerbate inflammation in RA. Further studies are needed to fully elucidate the downstream signaling and phenotypic effects of autoantibody binding on infiltrating monocytes and whether monocytes from RA patients are more prone to adopt a pro-inflammatory profile than those from healthy donors.

In trying to understand the mechanisms by which PAD4, which lacks canonical secretory signal sequences and transmembrane domains, traffics to the cell membrane and is retained there, we discovered important insights into the generation of the citrullinome. Not only was PAD4 found to be associated with vesicular, subcellular structures in monocytes, but we also found citrullinated proteins scattered across several intracellular compartments, implicating that PAD4 may be more ubiquitously present and active throughout the cell than previously appreciated (**Figure 4-1**). In addition to previously known (mainly nuclear) binding partners, our *in vitro* studies showed that PAD4 can bind to proteins that function in a wide variety of roles from various subcellular compartments. Of particular interest were those that are found on the cell surface and

those that take part in the secretory and autophagic pathways. Differentiation of monocytes into mo-DCs also caused increased surface PAD4 expression, with co-localization with proteins found in endocytic vesicles, the ER, and lysosomes, indicating that PAD4 is gaining access to conventional secretory pathways through unconventional means.

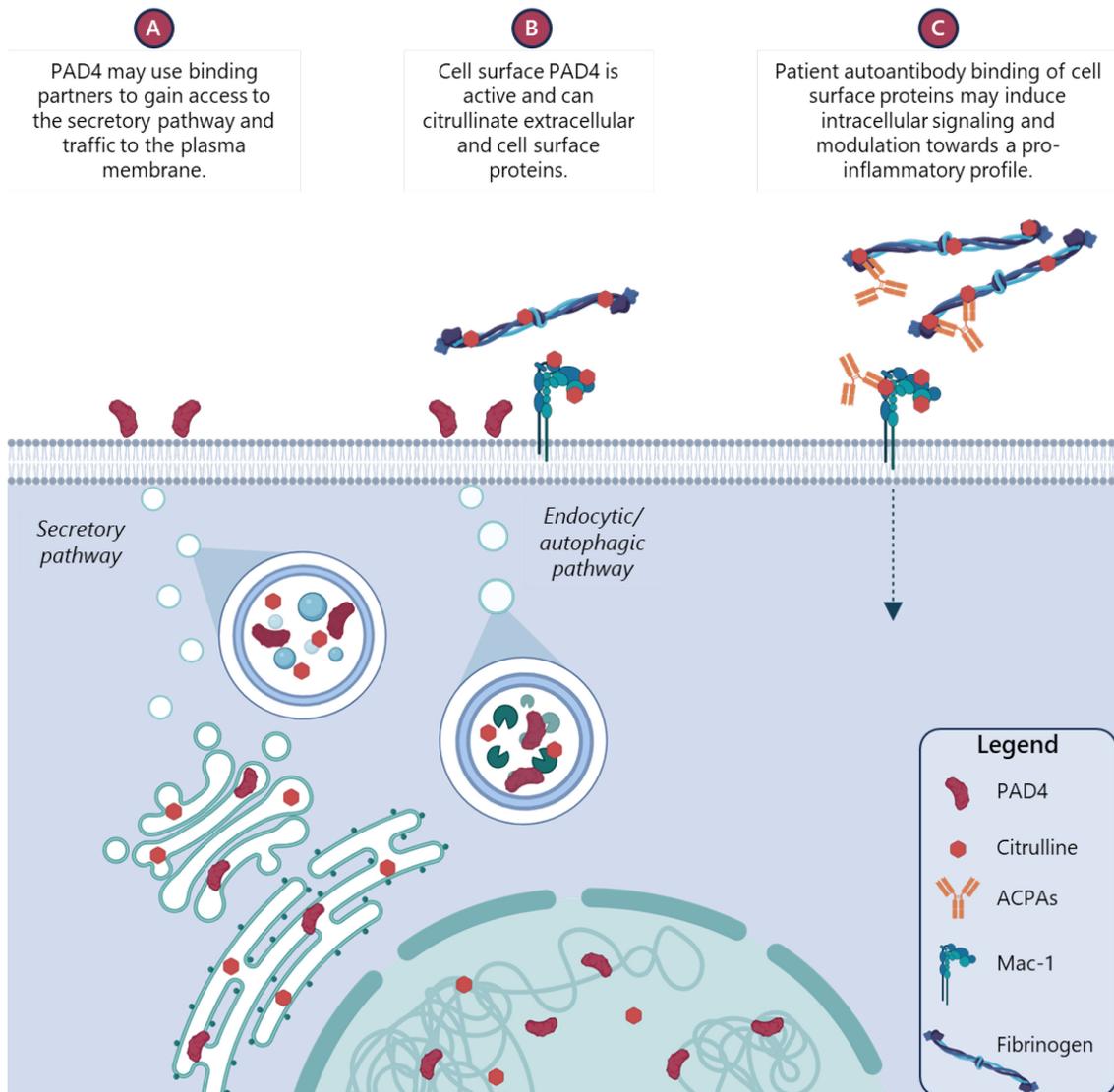


Fig. 4-1 PAD4 cell surface localization turns the monocyte surface into a site of autoantigen generation and makes monocytes a cellular target of ACPAs. (A) PAD4 traffics to the cell surface and other subcellular compartments via binding partners. The presence of citrullinated proteins in these compartments indicates that PAD4 is catalytically active there.

(B) At the cell surface, PAD4 can citrullinate extracellular (*e.g.* fibrinogen) and cell surface proteins (*e.g.* Mac-1). **(C)** Patient autoantibody binds to citrullinated residues on extracellular and cell surface proteins. ACPA binding to citrullinated cell surface proteins such as Mac-1 may induce intracellular signaling that cause cellular modulation towards a more pro-inflammatory profile.

Taken together, our findings provide important mechanistic insights into PAD4 cellular localization and citrullinated autoantigen generation, two pathways with high relevance to RA pathogenesis. We show that monocytes are important reservoirs of PAD4 with an ability to retain and increase surface PAD4 expression even after differentiation into mo-DCs, which suggests the possible ability of monocytes to continue citrullination well after infiltrating the synovium. Additionally, in this study, we reported the presence of PAD2 in monocytes isolated from peripheral blood. Therefore, the relative contribution of both PADs to the RA citrullinome should be probed further to fully understand their individual roles in disease pathogenesis. Citrullination can affect both protein structure and function.²¹⁷⁻²¹⁹ Since citrullinated proteins were found throughout the cell, including the cell membrane and other organelles, the role of citrullination in regulating protein activity in normal physiology as well as in disease pathology should also be subject to further investigation.

Finding PAD4 on the cell membrane and throughout subcellular compartments has important therapeutic implications. Since citrullination is present intracellularly under physiologic conditions, perhaps therapeutic inhibition should be limited to targeting the extracellular and cell surface fractions of PAD4. This more targeted approach might enable amelioration of disease symptoms without adversely affecting important intracellular

physiologic functions performed by PAD4. Currently, PAD4-selective inhibitors are small molecules that preferably bind to the low-calcium bound conformation of PAD4.¹⁷² Consequently, their binding is reversible and under high calcium concentrations in which PAD4 is able to overcome inhibition and stay active. However, greater understanding of how PAD4 is retained at the cell surface, possibly via interactions with protein or lipid binding partners, may inform better PAD4 inhibitor development. Further studies are still needed to fully understand the mechanisms behind pathogenic citrullination. Elucidating the role of monocytes and extranuclear PAD4 in contributing to disease pathogenesis will pave the way for better treatment options for patients suffering from RA.

References

1. Alamanos, Y., Voulgari, P. V. & Drosos, A. A. Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. *Semin. Arthritis Rheum.* **36**, 182–188 (2006).
2. Deane, K. D. *et al.* Genetic and environmental risk factors for rheumatoid arthritis. *Best Pract. Res. Clin. Rheumatol.* **31**, 3–18 (2017).
3. Frisell, T., Saevarsdottir, S. & Askling, J. Family history of rheumatoid arthritis: an old concept with new developments. *Nat. Rev. Rheumatol.* **12**, 335–343 (2016).
4. Alamanos, Y. & Drosos, A. A. Epidemiology of adult rheumatoid arthritis. *Autoimmun. Rev.* **4**, 130–136 (2005).
5. Niemantsverdriet, E. & van der Helm-van Mil, A. H. M. Imaging detected tenosynovitis of metacarpophalangeal and wrist joints: an increasingly recognised characteristic of rheumatoid arthritis. *Clin. Exp. Rheumatol.* **36 Suppl 114**, 131–138 (2018).
6. Hulsmans, H. M. *et al.* The course of radiologic damage during the first six years of rheumatoid arthritis. *Arthritis Rheum.* **43**, 1927–1940 (2000).
7. Brito, Y., Glassberg, M. K. & Ascherman, D. P. Rheumatoid Arthritis-Associated Interstitial Lung Disease: Current Concepts. *Curr. Rheumatol. Rep.* **19**, 79 (2017).
8. Simon, T. A., Thompson, A., Gandhi, K. K., Hochberg, M. C. & Suissa, S. Incidence of malignancy in adult patients with rheumatoid arthritis: a meta-analysis. *Arthritis Res. Ther.* **17**, 212 (2015).

9. Avina-Zubieta, J. A., Thomas, J., Sadatsafavi, M., Lehman, A. J. & Lacaille, D. Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. *Ann. Rheum. Dis.* **71**, 1524–1529 (2012).
10. Reisine, S., Fifield, J., Walsh, S. & Dauser, D. Work disability among two cohorts of women with recent-onset rheumatoid arthritis: a survival analysis. *Arthritis Rheum.* **57**, 372–380 (2007).
11. Gullick, N. J. *et al.* Real world long-term impact of intensive treatment on disease activity, disability and health-related quality of life in rheumatoid arthritis. *BMC Rheumatol.* **3**, 6 (2019).
12. Matsumoto, T. *et al.* Trends in Treatment, Outcomes, and Incidence of Orthopedic Surgery in Patients with Rheumatoid Arthritis: An Observational Cohort Study Using the Japanese National Database of Rheumatic Diseases. *J. Rheumatol.* **44**, 1575–1582 (2017).
13. Gregersen, P. K., Silver, J. & Winchester, R. J. Genetic susceptibility to rheumatoid arthritis and human leukocyte antigen class II polymorphism: The role of shared conformational determinants. *Am. J. Med.* **85**, 17–19 (1988).
14. Holoshitz, J. The Rheumatoid Arthritis HLA-DRB1 Shared Epitope. *Curr. Opin. Rheumatol.* **22**, 293–298 (2010).
15. Raychaudhuri, S. *et al.* Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat. Genet.* **44**, 291–296 (2012).
16. Smolen, J. S. *et al.* Rheumatoid arthritis. *Nat. Rev. Dis. Primer* **4**, 1–23 (2018).

17. Moreno, I. *et al.* Association of the shared epitope with radiological severity of rheumatoid arthritis. *J. Rheumatol.* **23**, 6–9 (1996).
18. Mewar, D. *et al.* Association between radiographic severity of rheumatoid arthritis and shared epitope alleles: differing mechanisms of susceptibility and protection. *Ann. Rheum. Dis.* **67**, 980–983 (2008).
19. Begovich, A. B. *et al.* A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* **75**, 330–337 (2004).
20. Goëb, V. *et al.* Contribution of PTPN22 1858T, TNFR11 196R and HLA-shared epitope alleles with rheumatoid factor and anti-citrullinated protein antibodies to very early rheumatoid arthritis diagnosis. *Rheumatology* **47**, 1208–1212 (2008).
21. Farago, B. *et al.* Protein tyrosine phosphatase gene C1858T allele confers risk for rheumatoid arthritis in Hungarian subjects. *Rheumatol. Int.* **29**, 793–796 (2009).
22. Kurkó, J. *et al.* Genetics of Rheumatoid Arthritis — A Comprehensive Review. *Clin. Rev. Allergy Immunol.* **45**, 170–179 (2013).
23. Suzuki, A. *et al.* Functional haplotypes of PADI4 , encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat. Genet.* **34**, 395–402 (2003).
24. Cha, S. *et al.* Association of anti-cyclic citrullinated peptide antibody levels with PADI4 haplotypes in early rheumatoid arthritis and with shared epitope alleles in very late rheumatoid arthritis. *Arthritis Rheum.* **56**, 1454–1463 (2007).

25. Plenge, R. M. *et al.* TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study. *N. Engl. J. Med.* **357**, 1199–1209 (2007).
26. Remmers, E. F. *et al.* STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N. Engl. J. Med.* **357**, 977–986 (2007).
27. Tong, G., Zhang, X., Tong, W. & Liu, Y. Association between polymorphism in STAT4 gene and risk of rheumatoid arthritis: A meta-analysis. *Hum. Immunol.* **74**, 586–592 (2013).
28. Lee, Y. H., Ji, J. D. & Song, G. G. Associations between FCGR3A polymorphisms and susceptibility to rheumatoid arthritis: a metaanalysis. *J. Rheumatol.* **35**, 2129–2135 (2008).
29. van der Linden, M. P. M. *et al.* Association of a single-nucleotide polymorphism in CD40 with the rate of joint destruction in rheumatoid arthritis. *Arthritis Rheum.* **60**, 2242–2247 (2009).
30. Stahl, E. A. *et al.* Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.* **42**, 508–514 (2010).
31. Liu, Y. *et al.* Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat. Biotechnol.* **31**, 142–147 (2013).
32. Meng, W. *et al.* DNA methylation mediates genotype and smoking interaction in the development of anti-citrullinated peptide antibody-positive rheumatoid arthritis. *Arthritis Res. Ther.* **19**, 71 (2017).

33. Klareskog, L., Gregersen, P. K. & Huizinga, T. W. J. Prevention of autoimmune rheumatic disease: state of the art and future perspectives. *Ann. Rheum. Dis.* **69**, 2062–2066 (2010).
34. Sokolove, J. *et al.* Increased inflammation and disease activity among current cigarette smokers with rheumatoid arthritis: a cross-sectional analysis of US veterans. *Rheumatol. Oxf. Engl.* **55**, 1969–1977 (2016).
35. Johannsen, A., Susin, C. & Gustafsson, A. Smoking and inflammation: evidence for a synergistic role in chronic disease. *Periodontol. 2000* **64**, 111–126 (2014).
36. Arnsen, Y., Shoenfeld, Y. & Amital, H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. *J. Autoimmun.* **34**, J258–J265 (2010).
37. Klareskog, L. *et al.* A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum.* **54**, 38–46 (2006).
38. de Hair, M. J. H. *et al.* Smoking and overweight determine the likelihood of developing rheumatoid arthritis. *Ann. Rheum. Dis.* **72**, 1654–1658 (2013).
39. Ngo, S. T., Steyn, F. J. & McCombe, P. A. Gender differences in autoimmune disease. *Front. Neuroendocrinol.* **35**, 347–369 (2014).
40. Alpízar-Rodríguez, D., Pluchino, N., Canny, G., Gabay, C. & Finckh, A. The role of female hormonal factors in the development of rheumatoid arthritis. *Rheumatol. Oxf. Engl.* **56**, 1254–1263 (2017).

41. Smolik, I., Robinson, D. & El-Gabalawy, H. S. Periodontitis and rheumatoid arthritis: epidemiologic, clinical, and immunologic associations. *Compend. Contin. Educ. Dent. Jamesburg NJ 1995* **30**, 188–190, 192, 194 passim; quiz 198, 210 (2009).
42. Konig, M. F., Paracha, A. S., Moni, M., Bingham, C. O. & Andrade, F. Defining the role of Porphyromonas gingivalis peptidylarginine deiminase (PPAD) in rheumatoid arthritis through the study of PPAD biology. *Ann. Rheum. Dis.* **74**, 2054–2061 (2015).
43. Kharlamova, N. *et al.* Antibodies to Porphyromonas gingivalis Indicate Interaction Between Oral Infection, Smoking, and Risk Genes in Rheumatoid Arthritis Etiology. *Arthritis Rheumatol. Hoboken NJ* **68**, 604–613 (2016).
44. Demoruelle, M. K., Deane, K. D. & Holers, V. M. When and where does inflammation begin in rheumatoid arthritis? *Curr. Opin. Rheumatol.* **26**, 64–71 (2014).
45. Wright, H. L., Moots, R. J. & Edwards, S. W. The multifactorial role of neutrophils in rheumatoid arthritis. *Nat. Rev. Rheumatol.* **10**, 593–601 (2014).
46. Wright, H. L., Chikura, B., Bucknall, R. C., Moots, R. J. & Edwards, S. W. Changes in expression of membrane TNF, NF- κ B activation and neutrophil apoptosis during active and resolved inflammation. *Ann. Rheum. Dis.* **70**, 537–543 (2011).
47. Sokolove, J., Zhao, X., Chandra, P. E. & Robinson, W. H. Immune complexes containing citrullinated fibrinogen costimulate macrophages via Toll-like receptor 4 and Fc γ receptor. *Arthritis Rheum.* **63**, 53–62 (2011).
48. Holers, V. M. & Banda, N. K. Complement in the Initiation and Evolution of Rheumatoid Arthritis. *Front. Immunol.* **9**, (2018).

49. Gierut, A., Perlman, H. & Pope, R. M. Innate Immunity and Rheumatoid Arthritis. *Rheum. Dis. Clin. North Am.* **36**, 271–296 (2010).
50. Takemura, S. *et al.* Lymphoid neogenesis in rheumatoid synovitis. *J. Immunol. Baltim. Md 1950* **167**, 1072–1080 (2001).
51. Silverman, G. J. & Carson, D. A. Roles of B cells in rheumatoid arthritis. *Arthritis Res. Ther.* **5**, S1–S6 (2003).
52. Humby, F. *et al.* Ectopic Lymphoid Structures Support Ongoing Production of Class-Switched Autoantibodies in Rheumatoid Synovium. *PLoS Med.* **6**, (2009).
53. Song, Y. W. & Kang, E. H. Autoantibodies in rheumatoid arthritis: rheumatoid factors and anticitrullinated protein antibodies. *QJM Int. J. Med.* **103**, 139–146 (2010).
54. Reparon-Schuijt, C. C. *et al.* Secretion of anti-citrulline-containing peptide antibody by B lymphocytes in rheumatoid arthritis. *Arthritis Rheum.* **44**, 41–47 (2001).
55. Cope, A. P., Schulze-Koops, H. & Aringer, M. The central role of T cells in rheumatoid arthritis. *Clin. Exp. Rheumatol.* **25**, S4–11 (2007).
56. Yap, H.-Y. *et al.* Pathogenic Role of Immune Cells in Rheumatoid Arthritis: Implications in Clinical Treatment and Biomarker Development. *Cells* **7**, (2018).
57. Bottini, N. & Firestein, G. S. Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors. *Nat. Rev. Rheumatol.* **9**, 24–33 (2013).
58. Trouw, L. A. & Mahler, M. Closing the serological gap: promising novel biomarkers for the early diagnosis of rheumatoid arthritis. *Autoimmun. Rev.* **12**, 318–322 (2012).
59. Ajeganova, S. & Huizinga, T. W. J. Rheumatoid arthritis: Seronegative and seropositive RA: alike but different? *Nat. Rev. Rheumatol.* **11**, 8–9 (2015).

60. Wevers-de Boer, K. *et al.* Remission induction therapy with methotrexate and prednisone in patients with early rheumatoid and undifferentiated arthritis (the IMPROVED study). *Ann. Rheum. Dis.* **71**, 1472–1477 (2012).
61. Aletaha, D. *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann. Rheum. Dis.* **69**, 1580–1588 (2010).
62. Rose, H. M. & Ragan, C. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. *Proc. Soc. Exp. Biol. Med. Soc. Exp. Biol. Med. N. Y. N* **68**, 1–6 (1948).
63. Edelman, G. M., Kunkel, H. G. & Franklin, E. C. INTERACTION OF THE RHEUMATOID FACTOR WITH ANTIGEN-ANTIBODY COMPLEXES AND AGGREGATED GAMMA GLOBULIN. *J. Exp. Med.* **108**, 105–120 (1958).
64. Mellors, R. C., Nowoslawski, A., Korngold, L. & Sengson, B. L. RHEUMATOID FACTOR AND THE PATHOGENESIS OF RHEUMATOID ARTHRITIS. *J. Exp. Med.* **113**, 475–484 (1961).
65. Dörner, T., Egerer, K., Feist, E. & Burmester, G. Rheumatoid factor revisited. *Curr. Opin. Rheumatol.* **16**, 246–253 (2004).
66. Klareskog, L., Rönnelid, J., Lundberg, K., Padyukov, L. & Alfredsson, L. Immunity to Citrullinated Proteins in Rheumatoid Arthritis. *Annu. Rev. Immunol.* **26**, 651–675 (2008).

67. Hill, J. A., Al-Bishri, J., Gladman, D. D., Cairns, E. & Bell, D. A. Serum autoantibodies that bind citrullinated fibrinogen are frequently found in patients with rheumatoid arthritis. *J. Rheumatol.* **33**, 2115–2119 (2006).
68. Nikolaisen, C., Rekvig, O. P. & Nossent, H. C. Diagnostic impact of contemporary biomarker assays for rheumatoid arthritis. *Scand. J. Rheumatol.* **36**, 97–100 (2007).
69. Vallbracht, I. *et al.* Diagnostic and clinical value of anti-cyclic citrullinated peptide antibodies compared with rheumatoid factor isotypes in rheumatoid arthritis. *Ann. Rheum. Dis.* **63**, 1079–1084 (2004).
70. Silveira, I. G., Burlingame, R. W., von Mühlen, C. A., Bender, A. L. & Staub, H. L. Anti-CCP antibodies have more diagnostic impact than rheumatoid factor (RF) in a population tested for RF. *Clin. Rheumatol.* **26**, 1883–1889 (2007).
71. Dejaco, C. *et al.* Diagnostic value of antibodies against a modified citrullinated vimentin in rheumatoid arthritis. *Arthritis Res. Ther.* **8**, R119 (2006).
72. Vossenaar, E. R., Zendman, A. J. W., Venrooij, W. J. van & Pruijn, G. J. M. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *BioEssays* **25**, 1106–1118 (2003).
73. Sokolove, J. *et al.* Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. *PLoS One* **7**, e35296 (2012).
74. Nielen, M. M. J. *et al.* Specific autoantibodies precede the symptoms of rheumatoid arthritis: A study of serial measurements in blood donors. *Arthritis Rheum.* **50**, 380–386 (2004).

75. Schellekens, G. A. *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum.* **43**, 155–163 (2000).
76. Pruijn, G. J., Wiik, A. & van Venrooij, W. J. The use of citrullinated peptides and proteins for the diagnosis of rheumatoid arthritis. *Arthritis Res. Ther.* **12**, 203 (2010).
77. Rönnelid, J. *et al.* Anticitrullinated protein/peptide antibody multiplexing defines an extended group of ACPA-positive rheumatoid arthritis patients with distinct genetic and environmental determinants. *Ann. Rheum. Dis.* **77**, 203–211 (2018).
78. Ospelt, C. *et al.* Carbamylation of vimentin is inducible by smoking and represents an independent autoantigen in rheumatoid arthritis. *Ann. Rheum. Dis.* **76**, 1176–1183 (2017).
79. Juarez, M. *et al.* Identification of novel antiacetylated vimentin antibodies in patients with early inflammatory arthritis. *Ann. Rheum. Dis.* **75**, 1099–1107 (2016).
80. Thiele, G. M. *et al.* Malondialdehyde-acetaldehyde adducts (MAA) and anti-MAA antibody in rheumatoid arthritis. *Arthritis Rheumatol. Hoboken NJ* **67**, 645–655 (2015).
81. Jiang, X. *et al.* Anti-CarP antibodies in two large cohorts of patients with rheumatoid arthritis and their relationship to genetic risk factors, cigarette smoking and other autoantibodies. *Ann. Rheum. Dis.* **73**, 1761–1768 (2014).
82. Shi, J. *et al.* Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc. Natl. Acad. Sci.* **108**, 17372–17377 (2011).

83. Nissinen, R. *et al.* Peptidylarginine deiminase, the arginine to citrulline converting enzyme, is frequently recognized by sera of patients with rheumatoid arthritis, systemic lupus erythematosus and primary Sjögren syndrome. *Scand. J. Rheumatol.* **32**, 337–342 (2003).
84. Harris, M. L. *et al.* Association of autoimmunity to peptidyl arginine deiminase type 4 with genotype and disease severity in rheumatoid arthritis. *Arthritis Rheum.* **58**, 1958–1967 (2008).
85. Kolfenbach, J. R. *et al.* Autoimmunity to Peptidyl Arginine Deiminase Type 4 Precedes Clinical Onset of Rheumatoid Arthritis. *Arthritis Rheum.* **62**, 2633–2639 (2010).
86. Zhao, J., Zhao, Y., He, J., Jia, R. & Li, Z. Prevalence and significance of anti-peptidylarginine deiminase 4 antibodies in rheumatoid arthritis. *J. Rheumatol.* **35**, 969–974 (2008).
87. Halvorsen, E. H. *et al.* Serum IgG antibodies to peptidylarginine deiminase 4 in rheumatoid arthritis and associations with disease severity. *Ann. Rheum. Dis.* **67**, 414–417 (2008).
88. Darrah, E. *et al.* Autoantibodies to Peptidylarginine Deiminase 2 Are Associated With Less Severe Disease in Rheumatoid Arthritis. *Front. Immunol.* **9**, (2018).
89. Hassfeld, W. *et al.* Demonstration of a new antinuclear antibody (anti-RA33) that is highly specific for rheumatoid arthritis. *Arthritis Rheum.* **32**, 1515–1520 (1989).

90. Steiner, G. *et al.* Purification and partial sequencing of the nuclear autoantigen RA33 shows that it is indistinguishable from the A2 protein of the heterogeneous nuclear ribonucleoprotein complex. *J. Clin. Invest.* **90**, 1061–1066 (1992).
91. Konig, M. F., Giles, J. T., Nigrovic, P. A. & Andrade, F. Antibodies to native and citrullinated RA33 (hnRNP A2/B1) challenge citrullination as the inciting principle underlying loss of tolerance in rheumatoid arthritis. *Ann. Rheum. Dis.* **75**, 2022–2028 (2016).
92. Witalison, E. E., Thompson, P. R. & Hofseth, L. J. Protein Arginine Deiminases and Associated Citrullination: Physiological Functions and Diseases Associated with Dysregulation. *Curr. Drug Targets* **16**, 700–710 (2015).
93. Darrah, E. & Andrade, F. Rheumatoid arthritis and citrullination. *Curr. Opin. Rheumatol.* **30**, 72–78 (2018).
94. Takahara, H., Okamoto, H. & Sugawara, K. Calcium-dependent Properties of Peptidylarginine Deiminase from Rabbit Skeletal Muscle. *Agric. Biol. Chem.* **50**, 2899–2904 (1986).
95. Arita, K. *et al.* Structural basis for Ca²⁺-induced activation of human PAD4. *Nat. Struct. Mol. Biol.* **11**, 777–783 (2004).
96. Bicker, K. L. & Thompson, P. R. The protein arginine deiminases (PADs): Structure, Function, Inhibition, and Disease. *Biopolymers* **99**, 155–163 (2013).
97. Curran, A. M., Naik, P., Giles, J. T. & Darrah, E. PAD enzymes in rheumatoid arthritis: pathogenic effectors and autoimmune targets. *Nat. Rev. Rheumatol.* **16**, 301–315 (2020).

98. Asaga, H., Nakashima, K., Senshu, T., Ishigami, A. & Yamada, M. Immunocytochemical localization of peptidylarginine deiminase in human eosinophils and neutrophils. *J. Leukoc. Biol.* **70**, 46–51 (2001).
99. Vossenaar, E. R. *et al.* Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages. *Ann. Rheum. Dis.* **63**, 373–381 (2004).
100. Nakashima, K., Hagiwara, T. & Yamada, M. Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *J. Biol. Chem.* **277**, 49562–49568 (2002).
101. Hagiwara, T., Nakashima, K., Hirano, H., Senshu, T. & Yamada, M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem. Biophys. Res. Commun.* **290**, 979–983 (2002).
102. Christophorou, M. A. *et al.* Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **507**, 104–108 (2014).
103. Chang, X., Zhao, Y., Sun, S., Zhang, Y. & Zhu, Y. The expression of PADI4 in synovium of rheumatoid arthritis. *Rheumatol. Int.* **29**, 1411–1416 (2009).
104. Kinloch, A. *et al.* Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. *Arthritis Rheum.* **58**, 2287–2295 (2008).
105. Umeda, N. *et al.* Prevalence of soluble peptidylarginine deiminase 4 (PAD4) and anti-PAD4 antibodies in autoimmune diseases. *Clin. Rheumatol.* **35**, 1181–1188 (2016).

106. Foulquier, C. *et al.* Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation. *Arthritis Rheum.* **56**, 3541–3553 (2007).
107. Damgaard, D., Senolt, L., Nielsen, M. F., Pruijn, G. J. & Nielsen, C. H. Demonstration of extracellular peptidylarginine deiminase (PAD) activity in synovial fluid of patients with rheumatoid arthritis using a novel assay for citrullination of fibrinogen. *Arthritis Res. Ther.* **16**, 498 (2014).
108. Tarcsa, E. *et al.* Protein Unfolding by Peptidylarginine Deiminase: SUBSTRATE SPECIFICITY AND STRUCTURAL RELATIONSHIPS OF THE NATURAL SUBSTRATES TRICHOHYALIN AND FILAGGRIN*. *J. Biol. Chem.* **271**, 30709–30716 (1996).
109. Wang, F. *et al.* Identification of citrullinated peptides in the synovial fluid of patients with rheumatoid arthritis using LC-MALDI-TOF/TOF. *Clin. Rheumatol.* **35**, 2185–2194 (2016).
110. Tutturen, A. E. V., Fleckenstein, B. & de Souza, G. A. Assessing the Citrullinome in Rheumatoid Arthritis Synovial Fluid with and without Enrichment of Citrullinated Peptides. *J. Proteome Res.* **13**, 2867–2873 (2014).
111. Sharma, M. *et al.* Expanding the citrullinome of synovial fibrinogen from rheumatoid arthritis patients. *J. Proteomics* **208**, 103484 (2019).
112. Fan, L. Y. *et al.* Citrullinated vimentin stimulates proliferation, pro-inflammatory cytokine secretion, and PADI4 and RANKL expression of fibroblast-like synoviocytes in rheumatoid arthritis. *Scand. J. Rheumatol.* **41**, 354–358 (2012).

113. van Beers, J. J. B. C. *et al.* The rheumatoid arthritis synovial fluid citrullinome reveals novel citrullinated epitopes in apolipoprotein E, myeloid nuclear differentiation antigen, and β -actin. *Arthritis Rheum.* **65**, 69–80 (2013).
114. Darrah, E., Rosen, A., Giles, J. T. & Andrade, F. Peptidylarginine deiminase 2, 3 and 4 have distinct specificities against cellular substrates: Novel insights into autoantigen selection in rheumatoid arthritis. *Ann. Rheum. Dis.* **71**, 92–98 (2012).
115. Damgaard, D. *et al.* Relative efficiencies of peptidylarginine deiminase 2 and 4 in generating target sites for anti-citrullinated protein antibodies in fibrinogen, alpha-enolase and histone H3. *PloS One* **13**, e0203214 (2018).
116. Blachère, N. E. *et al.* High-Titer Rheumatoid Arthritis Antibodies Preferentially Bind Fibrinogen Citrullinated by Peptidylarginine Deiminase 4. *Arthritis Rheumatol. Hoboken NJ* **69**, 986–995 (2017).
117. Spengler, J. *et al.* Release of Active Peptidyl Arginine Deiminases by Neutrophils Can Explain Production of Extracellular Citrullinated Autoantigens in Rheumatoid Arthritis Synovial Fluid. *Arthritis Rheumatol. Hoboken Nj* **67**, 3135–3145 (2015).
118. Suzuki, A. *et al.* Decreased severity of experimental autoimmune arthritis in peptidylarginine deiminase type 4 knockout mice. *BMC Musculoskelet. Disord.* **17**, 205 (2016).
119. Willis, V. C. *et al.* Protein arginine deiminase 4 inhibition is sufficient for the amelioration of collagen-induced arthritis. *Clin. Exp. Immunol.* **188**, 263–274 (2017).
120. Curran, A. M. *et al.* Citrullination modulates antigen processing and presentation by revealing cryptic epitopes in rheumatoid arthritis. *Nat. Commun.* **14**, 1061 (2023).

121. Janssen, K. M. J. *et al.* Autoantibodies against citrullinated histone H3 in rheumatoid arthritis and periodontitis patients. *J. Clin. Periodontol.* **44**, 577–584 (2017).
122. Romero, V. *et al.* Immune-mediated pore-forming pathways induce cellular hypercitrullination and generate citrullinated autoantigens in rheumatoid arthritis. *Sci. Transl. Med.* **5**, 209ra150 (2013).
123. König, M. F. & Andrade, F. A Critical Reappraisal of Neutrophil Extracellular Traps and NETosis Mimics Based on Differential Requirements for Protein Citrullination. *Front. Immunol.* **7**, 461 (2016).
124. Sorice, M. *et al.* Autophagy generates citrullinated peptides in human synoviocytes: a possible trigger for anti-citrullinated peptide antibodies. *Rheumatol. Oxf. Engl.* **55**, 1374–1385 (2016).
125. Khandpur, R. *et al.* NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. *Sci. Transl. Med.* **5**, 178ra40 (2013).
126. Li, P. *et al.* PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps. PAD4 in NET-mediated bacterial killing. *J. Exp. Med.* **207**, 1853–1862 (2010).
127. Darrah, E. *et al.* Erosive rheumatoid arthritis is associated with antibodies that activate PAD4 by increasing calcium sensitivity. *Sci. Transl. Med.* **5**, 186ra65 (2013).
128. Zhou, Y. *et al.* Evidence for a direct link between PAD4-mediated citrullination and the oxidative burst in human neutrophils. *Sci. Rep.* **8**, 15228 (2018).
129. Zhou, Y. *et al.* Spontaneous Secretion of the Citrullination Enzyme PAD2 and Cell Surface Exposure of PAD4 by Neutrophils. *Front. Immunol.* **8**, 1200 (2017).

130. Marimuthu, R. *et al.* Characterization of Human Monocyte Subsets by Whole Blood Flow Cytometry Analysis. *J. Vis. Exp. JoVE* (2018) doi:10.3791/57941.
131. Coulthard, L. R. *et al.* Differential effects of infliximab on absolute circulating blood leucocyte counts of innate immune cells in early and late rheumatoid arthritis patients. *Clin. Exp. Immunol.* **170**, 36–46 (2012).
132. McGarry, T. *et al.* Rheumatoid arthritis CD14⁺ monocytes display metabolic and inflammatory dysfunction, a phenotype that precedes clinical manifestation of disease. *Clin. Transl. Immunol.* **10**, e1237 (2021).
133. Davignon, J.-L. *et al.* Targeting monocytes/macrophages in the treatment of rheumatoid arthritis. *Rheumatology* **52**, 590–598 (2013).
134. Yoon, B. R. *et al.* Functional Phenotype of Synovial Monocytes Modulating Inflammatory T-Cell Responses in Rheumatoid Arthritis (RA). *PLOS ONE* **9**, e109775 (2014).
135. Ruiz-Limon, P. *et al.* Molecular Characterization of Monocyte Subsets Reveals Specific and Distinctive Molecular Signatures Associated With Cardiovascular Disease in Rheumatoid Arthritis. *Front. Immunol.* **10**, 1111 (2019).
136. Rossol, M., Kraus, S., Pierer, M., Baerwald, C. & Wagner, U. The CD14^{bright}CD16⁺ monocyte subset is expanded in rheumatoid arthritis and promotes expansion of the Th17 cell population. *Arthritis Rheum.* **64**, 671–677 (2012).
137. Wijngaarden, S., van Roon, J. A. G., Bijlsma, J. W. J., van de Winkel, J. G. J. & Lafeber, F. P. J. G. Fcγ receptor expression levels on monocytes are elevated in rheumatoid

- arthritis patients with high erythrocyte sedimentation rate who do not use anti-rheumatic drugs. *Rheumatology* **42**, 681–688 (2003).
138. Lioté, F., Boval-Boizard, B., Weill, D., Kuntz, D. & Wautier, J.-L. Blood monocyte activation in rheumatoid arthritis: increased monocyte adhesiveness, integrin expression, and cytokine release. *Clin. Exp. Immunol.* **106**, 13–19 (1996).
139. Torsteinsdóttir, I., Arvidson, N.-G., Hällgren, R. & Håkansson, L. Monocyte activation in rheumatoid arthritis (RA): increased integrin, Fcγ and complement receptor expression and the effect of glucocorticoids. *Clin. Exp. Immunol.* **115**, 554–560 (1999).
140. Chimenti, M. S. *et al.* Synovial fluid from patients with rheumatoid arthritis modulates monocyte cell-surface phenotype. *J. Int. Med. Res.* **44**, 15–21 (2016).
141. Rana, A. K., Li, Y., Dang, Q. & Yang, F. Monocytes in rheumatoid arthritis: Circulating precursors of macrophages and osteoclasts and, their heterogeneity and plasticity role in RA pathogenesis. *Int. Immunopharmacol.* **65**, 348–359 (2018).
142. Coutant, F. Shaping of Monocyte-Derived Dendritic Cell Development and Function by Environmental Factors in Rheumatoid Arthritis. *Int. J. Mol. Sci.* **22**, 13670 (2021).
143. Harre, U. *et al.* Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. *J. Clin. Invest.* **122**, 1791–1802 (2012).
144. Krishnamurthy, A. *et al.* Identification of a novel chemokine-dependent molecular mechanism underlying rheumatoid arthritis-associated autoantibody-mediated bone loss. *Ann. Rheum. Dis.* **75**, 721–729 (2016).

145. Curran, A. M., Naik, P., Giles, J. T. & Darrah, E. PAD enzymes in rheumatoid arthritis: pathogenic effectors and autoimmune targets. *Nat. Rev. Rheumatol.* 1–15 (2020)
doi:10.1038/s41584-020-0409-1.
146. Lee, C.-Y. *et al.* Mining the Human Tissue Proteome for Protein Citrullination. *Mol. Cell. Proteomics MCP* **17**, 1378–1391 (2018).
147. Tilvawala, R. *et al.* The Rheumatoid Arthritis-Associated Citrullinome. *Cell Chem. Biol.* **25**, 691-704.e6 (2018).
148. Asaga, H., Nakashima, K., Senshu, T., Ishigami, A. & Yamada, M. Immunocytochemical localization of peptidylarginine deiminase in human eosinophils and neutrophils. *J. Leukoc. Biol.* **70**, 46–51 (2001).
149. Vossenaar, E. R. & van Venrooij, W. J. Citrullinated proteins: sparks that may ignite the fire in rheumatoid arthritis. *Arthritis Res. Ther.* **6**, 107–111 (2004).
150. Zhou, Y. *et al.* Spontaneous Secretion of the Citrullination Enzyme PAD2 and Cell Surface Exposure of PAD4 by Neutrophils. *Front. Immunol.* **8**, (2017).
151. Fukui, S. *et al.* M1 and M2 Monocytes in Rheumatoid Arthritis: A Contribution of Imbalance of M1/M2 Monocytes to Osteoclastogenesis. *Front. Immunol.* **8**, 1958 (2018).
152. Arnett, F. C. *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* **31**, 315–324 (1988).
153. Andrade, F. *et al.* Autocitrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation. *Arthritis Rheum.* **62**, 1630–1640 (2010).

154. Darrah, E. *et al.* Citrullination Is Not a Major Determinant in the Recognition of Peptidylarginine Deiminase 2 and 4 by Autoantibodies in Rheumatoid Arthritis. *Arthritis Rheumatol. Hoboken NJ* **72**, 1476–1482 (2020).
155. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
156. Tan, Y.-C. *et al.* Barcode-enabled sequencing of plasmablast antibody repertoires in rheumatoid arthritis. *Arthritis Rheumatol. Hoboken NJ* **66**, 2706–2715 (2014).
157. Elliott, S. E. *et al.* Affinity Maturation Drives Epitope Spreading and Generation of Proinflammatory Anti-Citrullinated Protein Antibodies in Rheumatoid Arthritis. *Arthritis Rheumatol. Hoboken NJ* **70**, 1946–1958 (2018).
158. Kongpachith, S. *et al.* Affinity Maturation of the Anti-Citrullinated Protein Antibody Paratope Drives Epitope Spreading and Polyreactivity in Rheumatoid Arthritis. *Arthritis Rheumatol. Hoboken NJ* **71**, 507–517 (2019).
159. Ma, S.-X. *et al.* Complement and Coagulation Cascades are Potentially Involved in Dopaminergic Neurodegeneration in α -Synuclein-Based Mouse Models of Parkinson's Disease. *J. Proteome Res.* **20**, 3428–3443 (2021).
160. Ramachandran, K. V. *et al.* Activity-Dependent Degradation of the Nascentome by the Neuronal Membrane Proteasome. *Mol. Cell* **71**, 169-177.e6 (2018).
161. Khan, S. Y. *et al.* Proteome Profiling of Developing Murine Lens Through Mass Spectrometry. *Invest. Ophthalmol. Vis. Sci.* **59**, 100–107 (2018).
162. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023).

163. Thomas, P. D. *et al.* PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
164. Mi, H. *et al.* Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **14**, 703–721 (2019).
165. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
166. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
167. Hill, J. A. *et al.* Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J. Immunol. Baltim. Md 1950* **171**, 538–541 (2003).
168. Travers, T. S. *et al.* Extensive Citrullination Promotes Immunogenicity of HSP90 through Protein Unfolding and Exposure of Cryptic Epitopes. *J. Immunol. Baltim. Md 1950* **197**, 1926–1936 (2016).
169. Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
170. Masson-Bessière, C. *et al.* The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. *J. Immunol. Baltim. Md 1950* **166**, 4177–4184 (2001).

171. Takizawa, Y. *et al.* Citrullinated fibrinogen detected as a soluble citrullinated autoantigen in rheumatoid arthritis synovial fluids. *Ann. Rheum. Dis.* **65**, 1013–1020 (2006).
172. Lewis, H. D. *et al.* Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation. *Nat. Chem. Biol.* **11**, 189–191 (2015).
173. Reed, E. *et al.* Antibodies to carbamylated α -enolase epitopes in rheumatoid arthritis also bind citrullinated epitopes and are largely indistinct from anti-citrullinated protein antibodies. *Arthritis Res. Ther.* **18**, 96 (2016).
174. Lac, P., Racapé, M., Barra, L., Bell, D. A. & Cairns, E. Relatedness of Antibodies to Peptides Containing Homocitrulline or Citrulline in Patients with Rheumatoid Arthritis. *J. Rheumatol.* **45**, 302–309 (2018).
175. Ehlers, M. R. W. CR3: a general purpose adhesion-recognition receptor essential for innate immunity. *Microbes Infect.* **2**, 289–294 (2000).
176. Campbell, I. D. & Humphries, M. J. Integrin Structure, Activation, and Interactions. *Cold Spring Harb. Perspect. Biol.* **3**, a004994 (2011).
177. Morgan, J. *et al.* Structural basis of the leukocyte integrin Mac-1 I-domain interactions with the platelet glycoprotein Ib. *Blood Adv.* **3**, 1450–1459 (2019).
178. Kufareva, I. & Abagyan, R. Methods of protein structure comparison. *Methods Mol. Biol. Clifton NJ* **857**, 231–257 (2012).
179. Castro-Alvarez, A., Costa, A. M. & Vilarrasa, J. The Performance of Several Docking Programs at Reproducing Protein–Macrolide-Like Crystal Structures. *Mol. J. Synth. Chem. Nat. Prod. Chem.* **22**, 136 (2017).

180. Cuthbert, G. L. *et al.* Histone Deimination Antagonizes Arginine Methylation. *Cell* **118**, 545–553 (2004).
181. Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylination. *Science* **306**, 279–283 (2004).
182. Sun, B. *et al.* Citrullination of NF- κ B p65 promotes its nuclear localization and TLR-induced expression of IL-1 β and TNF α . *Sci. Immunol.* **2**, eaal3062 (2017).
183. Neira, J. L. *et al.* Human Enzyme PADI4 Binds to the Nuclear Carrier Importin α 3. *Cells* **11**, 2166 (2022).
184. Fert-Bober, J., Darrah, E. & Andrade, F. Insights into the study and origin of the citrullinome in rheumatoid arthritis. *Immunol. Rev.* **294**, 133–147 (2020).
185. Damgaard, D., Bjørn, M. E., Steffensen, M. A., Pruijn, G. J. M. & Nielsen, C. H. Reduced glutathione as a physiological co-activator in the activation of peptidylarginine deiminase. *Arthritis Res. Ther.* **18**, 102 (2016).
186. Prame Kumar, K., Nicholls, A. J. & Wong, C. H. Y. Partners in crime: neutrophils and monocytes/macrophages in inflammation and disease. *Cell Tissue Res.* **371**, 551–565 (2018).
187. Jäger, E. *et al.* Calcium-sensing receptor-mediated NLRP3 inflammasome response to calcein particles drives inflammation in rheumatoid arthritis. *Nat. Commun.* **11**, 4243 (2020).
188. Silver, I. A., Murrills, R. J. & Etherington, D. J. Microelectrode studies on the acid microenvironment beneath adherent macrophages and osteoclasts. *Exp. Cell Res.* **175**, 266–276 (1988).

189. Hasegawa, T. Updating the pathophysiology of arthritic bone destruction: identifying and visualizing pathological osteoclasts in pannus. *Immunol. Med.* **44**, 246–251 (2021).
190. Zhang, X. *et al.* Genome-wide analysis reveals PADI4 cooperates with Elk-1 to activate c-Fos expression in breast cancer cells. *PLoS Genet.* **7**, e1002112 (2011).
191. Lu, M.-C. *et al.* Anti-citrullinated protein antibodies activated ERK1/2 and JNK mitogen-activated protein kinases via binding to surface-expressed citrullinated GRP78 on mononuclear cells. *J. Clin. Immunol.* **33**, 558–566 (2013).
192. Gabay, C. Interleukin-6 and chronic inflammation. *Arthritis Res. Ther.* **8**, S3 (2006).
193. Gottenberg, J.-E. *et al.* Serum IL-6 and IL-21 are associated with markers of B cell activation and structural progression in early rheumatoid arthritis: results from the ESPOIR cohort. *Ann. Rheum. Dis.* **71**, 1243–1248 (2012).
194. Smolen, J. S. *et al.* Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. *Lancet Lond. Engl.* **371**, 987–997 (2008).
195. Alghamdi, M. *et al.* An Overview of the Intrinsic Role of Citrullination in Autoimmune Disorders. *J. Immunol. Res.* **2019**, 7592851 (2019).
196. Jones, J., Causey, C., Knuckley, B., Slack-Noyes, J. L. & Thompson, P. R. Protein arginine deiminase 4 (PAD4): current understanding and future therapeutic potential. *Curr. Opin. Drug Discov. Devel.* **12**, 616–627 (2009).
197. Darrah, E. & Andrade, F. Rheumatoid arthritis and citrullination. *Curr. Opin. Rheumatol.* **30**, 72–78 (2018).

198. Fert-Bober, J. *et al.* Mapping Citrullinated Sites in Multiple Organs of Mice Using Hypercitrullinated Library. *J. Proteome Res.* **18**, 2270–2278 (2019).
199. O’Shea, J. P. *et al.* pLogo: a probabilistic approach to visualizing sequence motifs. *Nat. Methods* **10**, 1211–1212 (2013).
200. Duan-Porter, W. D., Jr, V. L. W., Maurer, K. D., Li, S. & Rosen, A. Dynamic Conformations of Nucleophosmin (NPM1) at a Key Monomer-Monomer Interface Affect Oligomer Stability and Interactions with Granzyme B. *PLOS ONE* **9**, e115062 (2014).
201. Antiochos, B. *et al.* The DNA sensors AIM2 and IFI16 are SLE autoantigens that bind neutrophil extracellular traps. *eLife* **11**, e72103 (2022).
202. Liu, Y. *et al.* Peptidylarginine deiminases 2 and 4 modulate innate and adaptive immune responses in TLR-7–dependent lupus. *JCI Insight* **3**, (2018).
203. Kinloch, A. *et al.* Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis. *Arthritis Res. Ther.* **7**, R1421-1429 (2005).
204. Yu, Q. *et al.* CEACAM1 (CD66a) promotes human monocyte survival via a phosphatidylinositol 3-kinase- and AKT-dependent pathway. *J. Biol. Chem.* **281**, 39179–39193 (2006).
205. Dillon, S. R. *et al.* Interleukin 31, a cytokine produced by activated T cells, induces dermatitis in mice. *Nat. Immunol.* **5**, 752–760 (2004).
206. Furebring, M., Håkansson, L., Venge, P. & Sjölin, J. Differential Expression of the C5a Receptor and Complement Receptors 1 and 3 after LPS Stimulation of Neutrophils and Monocytes. *Scand. J. Immunol.* **60**, 494–499 (2004).

207. Wang, S. *et al.* A role of Rab29 in the integrity of the trans-Golgi network and retrograde trafficking of mannose-6-phosphate receptor. *PloS One* **9**, e96242 (2014).
208. Nissim, S. *et al.* Mutations in RABL3 alter KRAS prenylation and are associated with hereditary pancreatic cancer. *Nat. Genet.* **51**, 1308–1314 (2019).
209. Zhang, B. *et al.* GSK3 β -Dzip1-Rab8 cascade regulates ciliogenesis after mitosis. *PLoS Biol.* **13**, e1002129 (2015).
210. Szatmári, Z. & Sass, M. The autophagic roles of Rab small GTPases and their upstream regulators. *Autophagy* **10**, 1154–1166 (2014).
211. Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. - Cell Physiol.* **300**, C723–C742 (2011).
212. Lu, J. *et al.* Types of nuclear localization signals and mechanisms of protein import into the nucleus. *Cell Commun. Signal.* **19**, 60 (2021).
213. Stewart, M. Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* **8**, 195–208 (2007).
214. Marzaioli, V. *et al.* Monocyte-Derived Dendritic Cell Differentiation in Inflammatory Arthritis Is Regulated by the JAK/STAT Axis via NADPH Oxidase Regulation. *Front. Immunol.* **11**, 1406 (2020).
215. Ireland, J. M. & Unanue, E. R. Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells. *J. Exp. Med.* **208**, 2625–2632 (2011).

216. León, B., López-Bravo, M. & Ardavín, C. Monocyte-Derived Dendritic Cells Formed at the Infection Site Control the Induction of Protective T Helper 1 Responses against Leishmania. *Immunity* **26**, 519–531 (2007).
217. Mortier, A. *et al.* Posttranslational modification of the NH₂-terminal region of CXCL5 by proteases or peptidylarginine Deiminases (PAD) differently affects its biological activity. *J. Biol. Chem.* **285**, 29750–29759 (2010).
218. Moelants, E. A. V. *et al.* Citrullination of TNF- α by peptidylarginine deiminases reduces its capacity to stimulate the production of inflammatory chemokines. *Cytokine* **61**, 161–167 (2013).
219. Pritzker, L. B., Joshi, S., Gowan, J. J., Harauz, G. & Moscarello, M. A. Deimination of myelin basic protein. 1. Effect of deimination of arginyl residues of myelin basic protein on its structure and susceptibility to digestion by cathepsin D. *Biochemistry* **39**, 5374–5381 (2000).

Curriculum vitae

Mekha A. Thomas

EDUCATION

Ph.D. expected	2023	Program in Immunology Mentor: Erika Darrah PhD	Johns Hopkins School of Medicine
B. S.	2018	Chemical Engineering	University of Maryland, Baltimore County

PROFESSIONAL EXPERIENCE

Doctoral research	2023-2019	Lab of Dr. Erika Darrah	Johns Hopkins School of Medicine
Research rotation	2019	Lab of Dr. Sergi Regot	Johns Hopkins School of Medicine
Research rotation	2018	Lab of Dr. Jonathan Schneck	Johns Hopkins School of Medicine
Undergraduate research	2016-2018	Lab of Dr. Gregory Szeto	University of Maryland, Baltimore County

ACADEMIC HONORS

2023	Won issue cover for journal: <i>Philosophical Transactions of the Royal Society B</i>
2022	Best Poster Presentation Award - 2022 Annual Immunology Retreat Source: Johns Hopkins School of Medicine
2022	Trainee Abstract Award - \$500 Source: American Association of Immunologists
2019	Won issue cover for journal: <i>Immunological Reviews</i>
2018	Outstanding Senior in Chemical Engineering Award Source: University of Maryland, Baltimore County
2017	Maximizing Access to Research Careers Undergraduate Student Training in Academic Research (MARC U*STAR) Scholar Source: NIH
2016	Undergraduate Research Award (URA) Recipient - \$1500 Source: University of Maryland, Baltimore County
2018-2014	President's List Source: University of Maryland, Baltimore County

PUBLICATIONS

1. **Thomas MA**, Naik P, Wang H, Giles JT, Girgis AA, Kim SY, Johnson TP, Curran AM, Crawford JD, Jahanbani S, Bingham CO, Robinson WH, Na CH, Darrah E. The monocyte cell surface is a novel site of autoantigen generation in rheumatoid arthritis. *Under review at Proceedings of the National Academy of Sciences (PNAS)*.
2. **Thomas MA**, Kim SY, Curran AM, Smith B, Antiochos B, Na CH, Darrah E. An unbiased proteomic analysis of PAD4 in human monocytes: novel substrates, binding partners and subcellular localizations. *Philos Trans R Soc Lond B Biol Sci*. 2023 Nov 20;378(1890):20220477.
3. Crawford JD, Wang H, Trejo-Zambrano D, Cimbro R, Talbot CC Jr, **Thomas MA**, Curran AM, Girgis AA, Schroeder JT, Fava A, Goldman DW, Petri M, Rosen A, Antiochos B, Darrah E. The XIST lncRNA is a sex-specific reservoir of TLR7 ligands in SLE. *JCI Insight*. 2023 Sep 21;8(20):e169344.

4. Curran AM, Girgis AA, Jang Y, Crawford JD, **Thomas MA**, Kawalerski R, Collier J, Bingham CO, Na CH, Darrah E. Citrullination modulates antigen processing and presentation by revealing cryptic epitopes in rheumatoid arthritis. *Nat Commun.* 2023; 14, 1061.
5. Eaton WW, Rodrigues KM, **Thomas MA**, Johnson J, Talor MV, Dohan C, Bingham CO, Musci R, Roth K, Kelly DL, Cihakova D, Darrah E. Immunologic profiling in schizophrenia and rheumatoid arthritis. *Psychiatry Res.* 2022 Nov; 317: 114812.
6. Cappelli LC, **Thomas MA**, Bingham CO 3rd, Shah AA, Darrah E. Immune checkpoint inhibitor-induced inflammatory arthritis as a model of autoimmune arthritis. *Immunol Rev.* 2020 Mar; 294(1):106-123.

PRESENTATIONS

1. **Mekha A. Thomas**, "The monocyte cell surface as a novel site of autoantigen generation in Rheumatoid Arthritis." Block Symposium – Autoantigen Identification, Autoantigen Specificity, and Central Tolerance, IMMUNOLOGY2022, American Association of Immunologists, May 2022. Oral presentation.
2. **Mekha A. Thomas**, et al., "The monocyte cell surface as a novel site of autoantigen generation in Rheumatoid Arthritis." IMMUNOLOGY2022, American Association of Immunologists, May 2022. Poster presentation.
3. **Mekha A. Thomas**, et al., "The monocyte cell surface as a novel site of autoantigen generation in Rheumatoid Arthritis." Annual Immunology Retreat, Johns Hopkins Immunology Program, JHU, September 2022. Poster presentation.
4. **Mekha A. Thomas**, "Cell surface PAD4 on monocytes generates autoantigens in Rheumatoid Arthritis." Annual Immunology Retreat, Johns Hopkins Immunology Program, JHU, September 2021. Oral presentation.

SERVICE AND LEADERSHIP

2021 - Present	Communications Specialist	Division of Rheumatology, Johns Hopkins School of Medicine
2021 – 2022	Chief Operating Graduate Officer	Graduate Program in Immunology, Johns Hopkins School of Medicine
2017 – 2015	Undergraduate Teaching Assistant	University of Maryland, Baltimore County

SKILLS

Technical skills: Primary human immune cell isolation and culture, cell line maintenance, cell fractionation and plasma membrane isolation, antibody affinity purification and labeling, flow cytometry, autoantibody detection via ELISAs, immunofluorescence microscopy, animal care, murine organ isolation (bone marrow, spleen, lymph nodes), murine immune cell culture and maintenance, fluorescent confocal microscopy

Programming languages: MATLAB, R, PyMol

Software: GraphPAD Prism, FlowJo, FCS Express, Fiji

Communication/Illustration: BioRender, Adobe Creative Suite (Illustrator, InDesign), American Society of Biochemistry and Molecular Biology (ASBMB) The Art of Science Communication course

Microsoft Office Suite: Word, PowerPoint, Excel, Outlook, Teams