Porphyromonas gingivalis peptidylarginine deiminase in the aetiology of rheumatoid arthritis

Anna Montgomery

St Cross College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

November 2016

Abstract

Anti-citrullinated protein antibodies (ACPA) are the main autoantibodies in rheumatoid arthritis (RA). Citrullination is the conversion of arginine to citrulline by peptidylarginine deiminase (PAD) enzymes. Periodontitis (PD) is a risk factor for RA, and there is evidence this may be due to citrullination by the unique prokaryotic PAD enzyme expressed by PD pathogen *Porphyromonas gingivalis* (PPAD). The aim of this work was to characterise PPAD, and investigate its potential role in the aetiology of RA.

The 3D crystal structure of PPAD was solved, and functional residues identified. The catalytic domain comprises five ββαβ motif repeats, which create a canonical binding groove into the active site, and the immunoglobulin-like domain is an anti-parallel β-sandwich. PPAD lacks significant homology to human PADs, and forms distinct citrullinated peptides from those formed by human PAD isoforms PAD2 and PAD4 due to unique C-terminal substrate specificity.

In vivo, immunisation with *P. gingivalis* proteins: PPAD, arginine gingipain (Rgp), and enolase (PgEnolase), induced ACPA responses in mice. In human cohorts, antibodies to PPAD, Rgp, and PgEnolase were increased in RA patients compared to osteoarthritis or healthy controls, and in RA patients with PD compared to those without. Six months of PD treatment saw a significant reduction in antibodies to *P. gingivalis* in RA patients with PD, but no change in ACPA.

These results demonstrate *P. gingivalis* induces ACPA responses comparable to those observed in RA, potentially through formation of novel citrullinated peptides by PPAD. Structural and mechanistic information obtained here can be used in future studies targeting PPAD as a potential therapeutic strategy in RA.

Table of Contents

Abstract i

Table of Contents ii

Acknowledgments v

Declaration of Originality vi

List of Table and Figures vii

Abbreviations xi

Chapter 1 Introduction 1

Rheumatoid arthritis 1

Anti-citrullinated protein antibodies (ACPA) 3

Use of ACPA in the diagnosis of RA 3

Development of ACPA 4

Pathogenicity of ACPA 8

Fine specificity of ACPA 11

Citrullinated proteins in the joint 14

The aetiology of rheumatoid arthritis 16

Genetic associations with RA 16

MHC-II SE and citrulline specific T-cells 19

Environmental influences in RA 24

The lung as a potential priming site for autoimmunity in RA 25

The gingiva as a potential priming site for autoimmunity in RA 27

Periodontitis 29

Aetiology of periodontal disease 29

Periodontitis and RA 31

The oral microbiome in RA 34

Periodontitis and citrullination in the aetiology of RA 35

Porphyromonas gingivalis 36

Porphyromonas gingivalis peptidylarginine deiminase (PPAD) 39

P. gingivalis antibodies in rheumatoid arthritis 41

Animal models in rheumatoid arthritis research 44

Citrullination and ACPA in in vivo models 48

P. gingivalis in murine models 49

Peptidylarginine deiminase enzymes (PADs) 52

Peptidylarginine deiminases in RA 54

PAD inhibitors 57

PAD inhibitors in models of disease 59

Structure and activity of mammalian PADs 61

Activity of bacterial deiminases 62

Aims of the study 63

Chapter 2 Expression, purification and crystallisation of PPAD 64

Background 64

Buffers 66

Materials and methods 68

Development of novel PPAD expression system at the Structural Genomics Consortium (SGC) 69

Site directed mutagenesis 72

PPAD expression 74

PPAD purification 75

Sodium Dodecyl Sulphate-PolyAcrimide Gel Electrophoresis (SDS-PAGE) 76
Tobacco Etch Virus (TEV) protease cleavage 76
Size exclusion chromatography 76
Characterisation of tPPAD^{WT} and tPPAD^{C351A} 77
Standardisation of protein concentration 78
Crystallisation of recombinant tPPAD^{WT} and tPPAD^{C351A} 78
Structural determination of tPPAD^{WT} and tPPAD^{C351A} 79
Statistical analysis of PPAD activity 80
Results *81*
Purification of recombinant full-length PPAD (flPPAD) 81
Expression and purification of tPPAD^{WT} and tPPAD^{C351A} 87
Comparison of tPPAD^{WT} and flPPAD^{WT} 89
Crystallisation of tPPAD^{WT} and tPPAD^{C351A} 92
Mapping of enzymatic activity using tPPAD^{WT} and tPPAD^{C351A} 98
Substrate specificity of tPPAD^{WT}, PAD2 and PAD4 101
Structural basis of substrate specificity of tPPAD^{WT} 102
Discussion *105*

Chapter 3 Characterisation of PPAD *in vitro* **111**
Background *111*
Buffers *113*
Materials and methods *114*
Comparison of PPAD construct solubility and activity 114
SDS-PAGE 115
Immunoblotting: western blotting and Anti-Modified Citrulline (AMC) assay 115
Investigation of PPAD auto-citrullination 116
Functional association of PPAD and RgpB 117
Cleavage and citrullination of human fibrinogen and α-enolase 118
Liquid Chromatography Mass Spectrometry (LC-MS) 118
Identification of citrullinated peptides from PgEnolase 119
Results *121*
The effect of PPAD N- and C-terminal truncation on solubility 121
The effect of PPAD N- and C-terminal truncation on citrullination activity 122
Characterisation of active constructs c007 and c073 124
Auto-citrullination of PPAD 127
Functional association of PPAD and Rgp 129
Formation of peptides from fibrinogen and α-enolase by PPAD and RgpB 130
Sequencing of peptides formed from fibrinogen and α-enolase by LC-MS 133
Identification of homologous citrullinated regions of PgEnolase and α-enolase 135
Formation of citrullinated peptides from fibrinogen by PPAD, PAD2 and PAD4 138
Discussion *140*

Chapter 4 PPAD, RgpB and PgEnolase *in vivo*: murine models **146**
Background *146*
Buffers *148*
Materials and methods *150*
Development of anti-PgEnolase antibody 150
Preparation of *P. gingivalis* fraction lysates 151
Western blotting of *P. gingivalis* fraction lysates 151
Visualisation of citrullination by biotin-phenylglyoxal (Biotin-Pg) probe 152
Murine immunisation with PPAD, RgpB and PgEnolase 153
Enzyme Linked Immunosorbent Assay (ELISA) with murine sera 155
Splenocyte proliferation assay 157
Cytokine ELISA with splenocyte supernatants 158
Statistical analysis 158

Results *159*
Development of anti-PgEnolase antibody 159
Location of PPAD, RgpB and PgEnolase in P. *gingivalis* fractions 161
Location of citrullination in *P. gingivalis* fractions 162
Anti-PPAD responses in PPAD immunised mice 164
The effect of complete Freund's adjuvant on antibody responses 167
Immunisation with PPAD, RgpB and PgEnolase 168
Antibody responses to PPAD and PgEnolase 169
Antibody responses to citrullinated PPAD peptide CPP3 171
Antibody responses to citrullinated PgEnolase peptides 172
Antibody responses to CEP-1 and REP-1 172
Splenocyte proliferation and cytokine release in response to PPAD 176
Discussion *177*

Chapter 5 PPAD, RgpB and PgEnolase in RA: human cohorts **181**
Background *181*
Buffers *183*
Materials and methods *184*
Patient cohorts 184
ELISA with human sera 186
Statistical analysis of patient demographics 188
Results *189*
WOR cohort patient demographics 189
Antibody reactivity with CEP-1 and REP-1 in RA and OA patients 190
Clinical and aetiological associations with anti-CEP-1 positivity 192
Effect of periodontitis on antibody positivity in RA and OA patients 198
OPERA cohort patient demographics 202
Antibody reactivity in RA^+PD^+, RA and PD patients 204
Effect of PD treatment on antibody levels 209
Discussion *212*

Chapter 6 Conclusion and future directions **217**

References **223**

Appendix **255**
Vectors *255*
Primers *261*

Published papers **262**

Acknowledgments

Thanks to my supervisors Patrick Venables and Wyatt Yue, and Jan Potempa and Ewa Cairns for supervising me during my visits to your labs. Patrick, as an extra note on your retirement, I have been incredibly lucky to work with someone as experienced, well respected, and knowledgeable as you. I hope you enjoy your retirement in peace without my pestering, and that I have made you proud as your last student.

Thanks to everyone in the lab: Anne-Marie, Elena, Raquel, Miriam, Tariq, Alasdair, David, for all the advice and fun over the last few years. Extra special thanks to Fiona, Hayley and Kay.

To my sister Beth, and best friends Emma and Alana, you're the most incredible women I've ever met, and inspire me daily. I am so proud to call you my friends. Thanks for keeping me sane with your messages over the last few months!

Thanks to Mummy and Daddy for all the support over the last year at home, and the previous 27. Everything I have achieved is entirely down to you both. Thanks for always having faith in me.

Björn, I have no idea how to thank you for what an incredible support you've been during my DPhil, and particularly the write up. I don't know what I would do without you.

Thanks!

Declaration of Originality

I declare that everything submitted in this thesis is entirely my own work, and help from others has been appropriately acknowledged.

List of Table and Figures

Chapter 1 Introduction

Table 1.1	ACR/EULAR classification criteria for rheumatoid arthritis	2
Table 1.2	Commonly occurring ACPA epitopes	6
Table 1.3	T-cell reactive citrullinated peptides	23
Table 1.4	Prevalence and severity of periodontitis in rheumatoid arthritis in various studies	32
Table 1.5	Characteristics of commonly used mouse models of arthritis	47
Table 1.6	Proteins citrullinated by peptidylarginine deiminases in healthy tissue(s)	54
Table 1.7	Peptidylarginine deiminase enzymes associated with disease	57
Figure 1.1	The conversion of arginine to citrulline by peptidylarginine deiminase (PAD) enzymes	5
Figure 1.2	Hypothesis for development of rheumatoid arthritis	17
Figure 1.3	3D crystal structure of *HLA-DRB1*0401* shared epitope allele	18
Figure 1.4	Chemical structure of Cl-amidine/F-amidine	58
Figure 1.5	Chemical structures of 2-chloroacetamidine and benzoyl arginine amide (BAA)	58
Figure 1.6	Chemical structure of second generation PAD inhibitor BB Cl-amidine	60

Chapter 2 Expression, purification and crystallisation of recombinant PPAD

Table 2.1	Summary of all PPAD constructs created	86
Table 2.2	Refinement statistics for $tPPAD^{WT}$ and $tPPAD^{C351A}$ crystal structures	94
Figure 2.1	Purification of full-length, GST-tagged, recombinant PPAD in house	81
Figure 2.2	High-throughput SGC pipeline for protein crystallisation	82
Figure 2.3	'Phase 1' construct design for PPAD	83
Figure 2.4	Small-scale test expression of 'phase 1' PPAD constructs	84

Figure 2.5	'Phase 2' construct design	87
Figure 2.6	Representative purification of tPPAD^{C351A}	88
Figure 2.7	Expression and His-tag cleavage of tPPAD^{WT} and tPPAD^{C351A}	89
Figure 2.8	Comparison of tPPAD and flPPAD activity	90
Figure 2.9	Characterisation of tPPAD *in vitro*	91
Figure 2.10	Crystals obtained from tPPAD^{WT} and tPPAD^{C351A}	93
Figure 2.11	Crystal structure of tPPAD^{WT} solved to 1.46Å resolution	95
Figure 2.12	Structural homology of tPPAD^{WT} and tPPAD^{C351A}	97
Figure 2.13	The active site of tPPAD	98
Figure 2.14	Enzymatic conversion of arginine to citrulline by tPPAD modelled using 3D structures	99
Figure 2.15	Substrate specificity of PPAD, PAD2 and PAD4	101
Figure 2.16	The effect of Arg152 and Arg154 mutations(s) on tPPAD activity and specificity	104
Figure 2.17	Crystal structure of human PAD2	108
Figure 2.18	Crystal structure of human PAD4	109

Chapter 3 Characterisation of PPAD *in vitro*

Table 3.1	Summary of PPAD constructs solubility and activity	125
Table 3.2	Citrullinated peptides formed from PPAD and RgpB	128
Table 3.3	C-terminally citrullinated peptides from PgEnolase	136
Table 3.4	Citrullination of fibrinogen by human PAD2 and PAD4	139

Figure 3.1	Test expression of all PPAD constructs	122
Figure 3.2	Activity of PPAD constructs	123
Figure 3.3	Citrullination in PPAD elution fractions	123
Figure 3.4	Activity of full-length and truncated PPAD constructs	124
Figure 3.5	N-terminal truncation of PPAD	126
Figure 3.6	Auto-citrullination of PPAD	128
Figure 3.7	The effect of PPAD on RgpB activity	130
Figure 3.8	Protein incubation with PPAD and RgpB	130
Figure 3.9	Degradation of human fibrinogen and α-enolase by RgpB	131
Figure 3.10	Peptides formed from human fibrinogen by PPAD and RgpB	133
Figure 3.11	Citrullinated peptides formed from fibrinogen and α-enolase by PPAD and RgpB	134

Figure 3.12 PgEnolase homology models 138

Chapter 4 PPAD, RgpB and PgEnolase *in vivo:* murine models

Table 4.1 Schedule for anti-PgEnolase antibody development 159

Figure 4.1 Reactivity of rabbit anti-sera with PgEnolase and α-enolase during PgEnolase antibody development 160
Figure 4.2 Location of PPAD, Rgp and PgEnolase in *P. gingivalis* fractions 162
Figure 4.3 Location of citrullination in *P. gingivalis* fractions 163
Figure 4.4 Preliminary immunisations with wild-type or mutant PPAD 164
Figure 4.5 Antibody responses to PPAD in PPAD immunised mice 165
Figure 4.6 IgM and IgG antibody responses to CEP-1 and REP-1 in PPAD immunised mice 166
Figure 4.7 Antibody responses to CEP-1 and REP-1 in CFA or IFA immunised mice 167
Figure 4.8 Immunisation and harvest procedure for mice immunised with PPAD, RgpB and PgEnolase 169
Figure 4.9 Antibody responses to PPAD and PgEnolase 170
Figure 4.10 Antibody responses to citrullinated PPAD peptide CPP3 171
Figure 4.11 Antibody responses to citrullinated PgEnolase peptides CPgE1-6 173
Figure 4.12 Antibody responses to CEP-1 and REP-1 174
Figure 4.13 Anti-CCP2 and anti-cVim antibody responses 175
Figure 4.14 Splenocyte proliferation and cytokine secretion in PPAD, RgpB and PgEnolase immunised mice 176

Chapter 5 PPAD, RgpB and PgEnolase in RA: human cohorts

Table 5.1 WOR cohort patients demographics 190
Table 5.2 Patient demographics and associations with anti-CEP-1 193
Table 5.3 Antibody positivity for ACPA in the WOR cohort 196
Table 5.4 OPERA cohort patients demographics 203
Table 5.5 Antibody positivity in PD^+RA^+ and RA patients at baseline 204
Table 5.6 OPERA cohort patient demographics for PD^+RA^+ patients 207
Table 5.7 Change in antibody positivity following 6 months of PD treatment 211

Figure 5.1	Anti-CEP-1 and anti-REP-1 in RA patients	191
Figure 5.2	Antibody responses to CPgE and RPgE peptides in RA patients	195
Figure 5.3	Correlations between anti-CEP-1 and anti-CPgE1-6 antibodies	196
Figure 5.4	Anti- RgpB and anti-CCP3 in RA patients	197
Figure 5.5	Anti-CPgE1-6 in PD^+ and PD^- individuals	198
Figure 5.6	Anti-RgpB and anti-CPP3 in PD^+ and PD^- individuals	200
Figure 5.7	Anti-CEP-1 and anti-REP-1 in PD^+ and PD^- individuals	201
Figure 5.8	Anti-CCP2 positivity in PD^+ and PD^- individuals	202
Figure 5.9	Antibody levels in PD^+RA^+ and RA patients	205
Figure 5.10	Baseline antibody levels in PD^+RA^+ patients	208
Figure 5.11	Effect of periodontal treatment on antibody levels in PD^+RA^+ patients after 6 months	210

Abbreviations

aa	amino acid
ACPA	anti-citrullinated protein antibodies
ACR	American College of Rheumatology
ADI	arginine deiminase
ADMA	asymmetric dimethylarginine
AIH	agmatine iminohydrolase
AMC	anti-modified citrulline (assay)
AstB	N-succinylarginine dihydrolase
AT	amidinotransferase
BAEE	benzoyl arginine ethyl esther
CAIA	collagen antibody induced arthritis
CAL	clinical attachment loss
CCP/2	cyclic citrullinated peptide/2 (assay)
CEP-1	immunodominant citrullinated enolase peptide
CFA	complete Freund's adjuvant
chlor	chloramphenicol
CIA	collagen induced arthritis
CII	collagen type II
CRP	C-reactive protein
DAS28	disease activity score
DDAH	dimethylarginine dimethylaminohydrolase
DNA	deoxyribonucleic acid
DTT	dithiothreitol

ECL	enhanced chemiluminescence
ELISA	enzyme linked immunosorbent assay
ESR	erythrocyte sedimentation rate
EULAR	The European League Against Rheumatism
FDR	first-degree relative
FT	flow-through
GCF	gingival crevicular fluid
GME	guanidino modifying superfamily of enzymes
GST	glutathione S-transferase
HLA	human leukocyte antigen
IFA	incomplete Freund's adjuvant
IFN	interferon
IL	interleukin
ILD	interstitial lung disease
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin monosulphate
kDa	kilo Dalton
Kgp	lysine gingipain
LB	lysogeny broth
LC-MS	liquid chromatography mass spectrometry
LDS	lithium dodecyl sulphate
LIC	ligation independent cloning
LPS	lipopolysaccharide
MHC	major histocompatibility complex
MOPS	3-(N-morpholino) propanesulfonic acid

MS/MS	tandem mass spectrometry
NETosis	formation/release of neutrophil extracellular traps
NETs	neutrophil extracellular traps
o/n	overnight
OA	osteoarthritis
OR	odds ratio
OD	optical density
PAD	peptidylarginine deiminase
PBS	phosphate buffered saline
PD	periodontitis
pdb	protein data base (file)
PdDp	periodontal pocket probing depth
PorSS	Por secretion system
PPAD	*Porphyromonas gingivalis* peptidylarginine deiminase
PTPN22	protein tyrosine phosphatase, non-receptor type 22
RA	rheumatoid arthritis
REP-1	non-citrullinated immunodominant enolase peptide variant
RF	rheumatoid factor
RgpA/RgpB	arginine gingipain A/B
RNA	ribonucleic acid
rt	room temperature
SCE	soluble cell extract
SE	shared epitope
SNP	single nucleotide polymorphism
TBS	tris buffered saline

TCE	total cell extract
TEV	tobacco etch virus
TNF-α	tumour necrosis factor α
TTBS	tris buffered saline with tween

Chapter 1 Introduction

Rheumatoid arthritis

Rheumatoid arthritis (RA) is an autoimmune disease affecting around 1% of the worlds population, with a 2-3:1 female to male ratio (Tedeschi *et al.*, 2013). RA is characterised by joint pain and stiffness caused by inflammation of the synovial tissue, eventually leading to loss of cartilage and bone erosion. Outside the joint, RA has a number of extra articular manifestations including: Sjörgren's syndrome, pericarditis, pleuritis, Felty's syndrome, vasculitis, neuropathy, glomerulonephritis, subcutaneous rheumatoid nodules, amyloidosis, xerostomia, pulmonary fibrosis, anaemia, bronchiolitis obliterans organising pneumonia (BOOP), and cervical myelopathy (Turesson *et al.*, 2003). The 30-year cumulative index for developing any extra articular manifestation is 46% (Turesson *et al.*, 2003), of which Sjögren's syndrome, anaemia of chronic disease and pulmonary manifestations are most frequent, occurring in 6–10% of patients (Young and Koduri, 2007). RA also carries an increased risk of secondary disorders such as interstitial lung disease (ILD) (O'Dwyer *et al.*, 2013), and an increased risk of cardiovascular deaths (hazard ratio 2.03) (Maradit-Kremers *et al.*, 2005). RA patients also have increased mortality. A meta-analysis of 11 longitudinal studies representing 51,819 patients over 40 years calculated the mean incidence mortality rate was 2.7 per 100 person-years of follow-up, compared to standardised mortality rate of 2.01 per 100 person-years (Dadoun *et al.*, 2013).

RA is heterogeneous, and so an additive system of criteria is used for diagnosis, either the 'American College of Rheumatology (ACR) 1987 classification criteria for rheumatoid arthritis' (Arnett *et al.*, 1988), or the updated '2010 ACR/European League Against Rheumatism (EULAR)' classification criteria (Aletaha *et al.*, 2010). Using the 2010 ACR/EULAR classification, RA is diagnosed based on synovitis in ≥1 joint, and a summated score of ≥6 from the four diagnostic categories, which must include at least 1 positive test for both acute-phase reactants and serology (Table 1.1).

Table 1.1 *The 2010 ACR/EULAR classification criteria for rheumatoid arthritis*

Category	Score				
	0	**1**	**2**	**3**	**5**
Joint(s)	1 (medium to) large joint	2-10 (medium to) large joints	1-3 small joints	4-10 small joints	>10 small joints
Serology	Negative	-	Low positive for RF and/or ACPA	High positive for RF and/or ACPA	-
Acute-phase reactants	Normal CPR and normal ESR	Abnormal CPR and/or abnormal ESR	-	-	-
Duration	<6 weeks	≥6 weeks	-	-	-

Adapted from (Aletaha *et al.*, 2010). CPR = C-reactive protein; ESP = erythrocyte sedimentation rate; RF = rheumatoid factor; ACPA = anti-citrullinated protein antibodies

The first autoantibodies characterised in RA were rheumatoid factor (RF), which are antibodies reactive with the Fc region of IgG. RF were first described by Waaler in 1940, who identified a serum factor in RA patient sera that induced coagulation of IgG coated sheep erythrocytes (reproduced in (Waaler, 2007)). These findings were reproduced soon after by Rose and colleagues (Rose *et al.*, 1948), and named rheumatoid factor by Pike due to the association with RA (Pike *et al.*, 1949). Although the reactivity of this serum factor remained

unknown, the erythrocyte agglutination test or the 'Waaler-Rose test' became used in RA diagnosis (Otten and Boerma, 1959). Franklin, and colleagues then showed the agglutination activity of RF was due to an immunoglobulin component, and the true IgG reactivity of this serum factor was later determined by Mellors and Ragan (Mellors *et al.*, 1959, Ragan, 1961).

Although still a component of the criteria for diagnosis of RA, it appears more likely that RF production is an artefact of immune over activity rather than an aetiological marker of RA. This is supported by the fact RF can be induced by non-specific B-cell activation (Garzelli *et al.*, 1984). The presence of RF in other autoimmune conditions including systemic lupus erythematosus (SLE) (Ippolito *et al.*, 2011) and mixed cryoglobulinaemia associated with hepatitis C infection (Pietrogrande *et al.*, 1995, Sasso, 2000) also limits their specificity to RA.

Anti-citrullinated protein antibodies (ACPA)

Use of ACPA in the diagnosis of RA

The main autoantibodies in RA are anti-citrullinated protein antibodies (ACPA), which are found in ~80% of patients and are highly specific (97%) for RA (Nijenhuis *et al.*, 2004). ACPA were identified after the discovery of two RA associated autoantibodies, 'anti-keratin antibodies' (AKA) and 'anti-perinuclear factor (APF)' (Nienhuis and Mandema, 1964). AKA were identified by indirect immunofluorescence (IIF) of RA patient sera that reacted with the cornified layer of rat oesophagus epithelium (Young *et al.*, 1979). Keratin was thought to be the antibody substrate, due to the high levels of cytokeratin in this tissue. Similarly, APF antibodies were also identified using IIF of RA sera, which bound the

perinuclear granules of human buccal mucosal cells (Nienhuis and Mandema, 1964). Following this, Sebbag and colleagues established AKA and APF shared numerous characteristics, and after detailed analysis established both antibodies were in fact reactive with filaggrin (Sebbag *et al.*, 1995). Simon and colleagues reported that AKA/APF antibodies reacted exclusively with a neutral/acidic form of filaggrin (Simon *et al.*, 1993), and the true substrate citrullinated filaggrin was identified, leading to the first description of 'ACPA' in 1998 (Schellekens *et al.*, 1998, Vincent *et al.*, 1998).

To increase the sensitivity of citrullinated filaggrin for use as a diagnostic marker, a library of cyclic citrullinated peptides was created, and used in the anti-cyclic citrullinated peptide (anti-CCP) assay. A second-generation anti-CCP test (anti-CCP2) was then developed using more highly reactive peptides, identified by screening peptide libraries for reactivity with RA sera (van Venrooij *et al.*, 2002).

Development of ACPA

ACPA are antibodies reactive with proteins that have undergone the post-translational modification of citrullination - conversion of peptidyl-arginine residue(s) to peptidyl-citrulline residue(s), by peptidylarginine deiminase (PAD) enzymes, with concomitant production of ammonia (Figure 1.1).

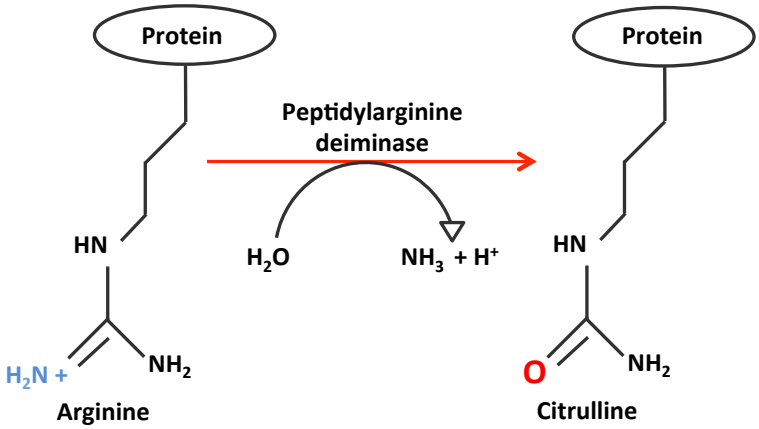


Figure 1.1 *The conversion of arginine to citrulline by peptidylarginine deiminase (PAD) enzymes.*

ACPA develop years before joint pathology occurs (Aho *et al.*, 2000, Rantapaa-Dahlqvist *et al.*, 2003, Nielen *et al.*, 2005, van de Stadt *et al.*, 2011a, Fisher *et al.*, 2015), during which time the titre (Ioan-Facsinay *et al.*, 2011), range of isotype (van der Woude *et al.*, 2010b, Kokkonen *et al.*, 2011) and specificity increases (van der Woude *et al.*, 2010a, Brink *et al.*, 2013, Sokolove *et al.*, 2012), which is predictive of disease development (van de Stadt *et al.*, 2012, van de Stadt *et al.*, 2011b). The most commonly occurring and well-characterised ACPA react with the immunodominant citrullinated peptides from α-enolase, fibrinogen, vimentin, collagen and filaggrin, named CEP-1, cFib, cVim, cCII and cFil respectively, and occur in 30-70% of RA patients (Table 1.2). Anti-CEP-1, anti-cFib, anti-cVim, anti-cCII and anti-cFil antibodies are all increased in patients destined to develop RA (pre-RA), and anti-cFib has the strongest association with developing RA (odds ratio (OR) of 10.4) (Fisher *et al.*, 2015, Aho *et al.*, 2000). In addition to these well-established reactivities, more recently antibodies to citrullinated histones have been found in 63-67% of RA patients (Pratesi *et al.*, 2013), citrullinated myelin basic protein in 34% of RA patients

(Willemze *et al.*, 2012), and citrullinated tenascin C in 47-51% of RA, and 18% of pre-RA patients (Schwenzer *et al.*, 2015).

Table 1.2 *Commonly occurring ACPA epitopes*

Antigen	Sensitivity	Reactive Peptide(s)	Reference
α-enolase	50%	'CEP-1' **^{5}KIHA-Cit-EIFDS-Cit-GNPTVE^{21}** ^{324}NPK-Cit-IAKAVNEKSCNCLL^{341}	*(Lundberg et al., 2008)*
Fibrinogen/fibrin	66%	α- chain **^{36}GP-Cit-VVE-Cit-HQSACKDS^{50}** ^{171}VDIDIKI-Cit-SC-Cit-GSCS^{185} ^{501}SGIGTLDGF-Cit-H-Cit-HPD^{515} ^{621}Cit-GHAKS-Cit-PV-Cit-GIHTS^{635} β- chain **^{60}Cit-PAPPPISGGGY-Cit-A-Cit^{74}**	*(Sebbag et al., 2006)*
Vimentin	40%	**^{59}VYAT-Cit-SSAV-Cit-L-Cit-SSVP^{74}** ^{1}ST-Cit-SVSSSSY-CitCit-MFGG^{16}	*(Snir et al., 2010)*
Collagen Type II	40%	**^{359}A-Cit-GLTG-Cit-PGDA^{369}**	*(Schulte et al., 1998)*
Filaggrin	28%	**^{306}SHQEST-Cit-G-Cit-SRGRSGRSG^{324}**	*(Schellekens et al., 1998)*

Peptide sequence in **bold** = immunodominant; not in bold = reactive

Netosis

NETosis is the production of neutrophil extracellular traps (NETs), and has recently received attention as a potential mechanism for ACPA development. NETs are web like structures composed of extracellular fibres, deoxyribonucleic acid (DNA), histones, chromatin, citrullinated proteins (Sur Chowdhury *et al.*, 2014), and globular protein domains, which aggregate to form protrusions from the cell. Once released, anti-microbial substances such as neutrophil elastase, myeloperoxidase, bactericidal permeability increasing protein (BPI) (Brinkmann *et al.*, 2004), histones (Parseghian and Luhrs, 2006) and proteases on the NET surface enable bacterial capture and killing (Kaplan and Radic, 2012).

Citrullination is required for NET formation (Li *et al.*, 2010a), and NETosis is induced by stimulation of neutrophils with ACPA (Khandpur *et al.*, 2013). During NETosis citrullinated vimentin and α-enolase are internalised by neutrophils (Khandpur *et al.*, 2013), and PAD enzymes PAD2 and PAD4 are released (Spengler *et al.*, 2015). In addition, neutrophils isolated from RA patients undergo increased NETosis compared to those isolated from healthy controls (Khandpur *et al.*, 2013, Sur Chowdhury *et al.*, 2014), and NETosis levels correlate with levels of ACPA and systemic inflammation (Khandpur *et al.*, 2013). Taken together these findings suggest NETs are both induced by ACPA, and may contribute to the formation of ACPA by increasing PAD2 and PAD4 expression, and formation of citrullinated proteins. It has been proposed that during an anti-microbial response, citrullinated proteins may become complexed with microbial proteins within NETs, are internalised by host phagocytes, and presented by antigen presenting cells (APCs), thus priming a citrulline specific ACPA response (Dwivedi and Radic, 2014).

Autophagy

Sorice and colleagues investigated the relationship between citrullination and autophagy in fibroblast-like synoviocytes isolated from RA and OA patients (Sorice *et al.*, 2016). Synoviocytes from RA patients had increased levels of autophagy compared to those from OA patients measured by western blotting of the cell lysates for autophagy marker LC3-II ($p<0.01$). RA patient synoviocytes also showed increased citrullination activity ($p<0.01$), and increased PAD4 expression measured using an ELISA based assay and western blotting respectively. When autophagy was induced in the synoviocytes *in vitro*, PAD4 was induced in cells from both RA and OA patients, and citrullinated proteins

vimentin and enolase were identified in both RA and OA cell lysates irrespective of patient anti-CCP2 status. However, in peripheral blood monocytes from anti-CCP2$^+$ RA patients, levels of LC3-II measured by flow cytometry correlated strongly with anti-CCP2 levels ($p<0.0004$). Although these findings have been reported on one study only, they propose an additional mechanism for the generation of citrullinated proteins in the synovium.

Pathogenicity of ACPA

Of those diagnosed with RA, ACPA status effectively divides patients into two subsets (ACPA$^+$/ACPA$^-$), which differ in clinical outcome (Willemze *et al.*, 2012). Notably, ACPA$^+$ RA is associated with a more severe phenotype when measured by Disease Activity Score (DAS28) and Visual Analog Scale (VAS) (Fisher *et al.*, 2011, van Oosterhout *et al.*, 2008). ACPA$^+$ RA is also more strongly associated with secondary disorders of RA including cardiovascular disease (Cambridge *et al.*, 2013), and lung disorders (Aubart *et al.*, 2011, Giles *et al.*, 2013).

ACPA$^+$ RA patients have increased radiographic progression of affected joints compared to ACPA$^-$ patients, which has been well documented in a number of studies. In a longitudinal study of 125 RA patients enrolled in the Norwegian European Research in Incapacitating Disease and Social Support (EURIDISS) cohort, radiographs of RA patients' hands were scored using sharp van der Heijde score (SVDH) at baseline and 10 years later. ACPA status measured by anti-CCP2 assay was a strongly associated independent predictor of radiographic progression, with OR of 4 (Syversen *et al.*, 2008). Van der Woude and colleagues also found a significant association between anti-CCP2 positivity and radiographic progression in the EURIDISS cohort ($p<0.05$), and replicated these

findings in 171 RA patients from the Leiden Early Arthritis Clinic cohort (Leiden EAC) in a 4 year follow up study ($p<0.05$) (van der Woude *et al.*, 2010b). Anti-CPP2 was also significantly associated with radiographic progression using SVDH score on hand radiographs in a Japanese cohort of 865 RA patients ($p=0.00004$) (Suzuki *et al.*, 2013), and when measured by Larsen score in a Belgian cohort of 180 RA patients ($p<0.001$) (De Rycke *et al.*, 2004). Using radiographic data from both the Leiden EAC and EURIDISS cohorts for analysis, Scherer and colleagues did not find an association between radiographic progression and ACPA fine specificity with citrullinated vimentin, fibrinogen, α-enolase or whole citrullinated MBP in either cohort (Scherer *et al.*, 2011). However, also using the Leiden EAC and EURIDISS cohorts, radiographic progression was positively associated with an increased range of ACPA isotypes (>5, $p<0.05$) and was specifically associated with IgA and IgG2 isotypes in the EURIDISS cohort ($p<0.05$) and IgA, IgG2, IgG3 and IgG4 in the Leiden EAC cohort ($p<0.05$) (van der Woude *et al.*, 2010b). Suwannalai and colleagues used the Leiden EAC cohort to investigate ACPA avidity in relation to radiographic progression. RA patients were grouped based on ACPA avidity into low/medium/or high avidity groups. Surprisingly, radiographic progression was 1.61 fold higher in low avidity ACPA RA patients compared to both medium and high avidity groups. An almost identical finding of a 1.62 fold increase in radiographic progression was observed in a second cohort of 132 RA patients from the Reade Early Arthritis Clinic (Reade EAC) using SVDH scoring of radiographs from a 6 year longitudinal study (Suwannalai *et al.*, 2013). These findings strongly suggest ACPA status influences radiographic progression of

disease in RA patients, which appears to be mediated via an increased range of ACPA isotypes of low avidity.

Although Scherer and colleagues did not find associations with anti-cVim ACPA and radiographic damage, studies from a separate group have proposed a possible explanation for increased bone erosion in ACPA$^+$ patients may be the binding of anti-cVim antibodies to citrullinated vimentin on the surface of osteoclasts, which results in osteoclast activation (Harre *et al.*, 2012). ACPA specific for mutated citrullinated vimentin (MCV) were affinity-purified from 3 RA patients with very high ACPA responses (>1,000 absorbance units (AU)) and incubated with osteoclasts. The purified MCV-ACPA stained both the plasma membrane and cytoplasm of osteoclasts, and increased osteoclast differentiation when added to MCSF and RANKL in osteoclast differentiation from peripheral blood mononuclear cells (PBMCs). Further, lymphocyte-deficient *Rag1*$^{-/-}$ mice challenged with affinity purified MCV-ACPA showed significant reduction in trabecular bone with compared to mice challenged with a control IgG, measured by micro-CT (μCT) 4 weeks post challenge.

Differences in ACPA$^+$/ACPA$^-$ pathology also influence treatment responses in RA patients. Presence of ACPA or RF is predictive of successful response to rituximab (OR 3.5) (Sellam *et al.*, 2011). ACPA$^+$ patients with early RA are more likely to achieve remission with methotrexate at 4 months in comparison to ACPA$^-$ RA (p=0.001) (Wevers-de Boer *et al.*, 2012), and methotrexate therapy can prevent progression to RA in ACPA$^+$ arthralgia patients, although not in those with high ACPA titre (van Dongen *et al.*, 2007).

Fine specificity of ACPA

Within $ACPA^+$ and $ACPA^-$ populations, there remains heterogeneity. To identify further potential sub-populations, Wilemze and colleagues (Willemze *et al.*, 2012) sub-grouped 661 RA patients from the Leiden EAC based on ACPA reactivity to citrullinated myelin basic protein, CEP-1, and cFib, but found no correlation with specific reactivity and clinical characteristics such as swollen joint count, C-reactive protein (CRP) or erythrocyte sedimentation rate (ESR). However, it should be noted that patients in the Leiden EAC were treated differently depending on date of inclusion in the study, which may have had a confounding effect on these negative results.

In one study Harlow and colleagues identified two citrullinated antigens from heat shock protein 90 α and β (HSP-α /-β) in the lungs of RA patients by immunoprecipitation. The presence of these citrullinated antigens was mostly restricted to RA patients with interstitial lung disease (ILD), compared to RA non-ILD patients (Harlow *et al.*, 2013). Thus, it may be more likely that specific ACPA reactivity is associated with extra articular effects, if the reactive antigen is found at a site distal to the joint.

α-enolase as an auto-antigen

CEP-1, the immunodominant citrullinated peptide from α-enolase, is one of the most commonly occurring ACPA reactivities, and was first characterised by our lab (Kinloch *et al.*, 2005). Kinloch and colleagues identified a protein of ~47kDa in citrullinated monocytic and granulocytic HL-60 cell lysates that reacted with RA sera when blotted with RA and non-RA controls. The protein target was then identified as α-enolase by tandem mass spectrometry (MS/MS). To confirm reactivity in a larger RA population, purified α-enolase was citrullinated and

blotted with 52 RA and 40 control sera. In this cohort 46% of RA sera and 6% of controls were reactive, thus confirming citrullinated α-enolase as a novel citrullinated antigen. α-enolase was also later identified in synovial tissues of the RA joint by immunohistochemistry, where it co-localised with citrullinated proteins in synovial cell lysates and synovial fluid (Kinloch *et al.*, 2008).

In order to determine the immunodominant citrullinated epitope(s) from α-enolase, 11 peptides were synthesised spanning 15 of the 17 arginine residues of the molecule, in either native or citrullinated form (Lundberg *et al.*, 2008). Antibody reactivity to these peptides was measured in 102 RA sera and 110 disease controls. Sixty-four per cent of RA patients and 15% of controls had serum reactivity with at least 1 of the 11 peptides, with the highest frequency of reactivity, 37% obtained against the citrullinated peptide spanning amino acids (aa) 5-21, ^{5}C-KIHA-Cit-EIFDS-Cit-GNPTVE-C^{21}. Further studies replacing either or both arginine residues at positions aa9 and aa15 on the peptide with citrulline confirmed it is most reactive when both residues were citrullinated. This peptide was termed CEP-1, and in a second cohort of 81 RA patients and 82 controls, sensitivity and specificity for RA both increased, to 62% and 98% respectively. Antibodies to CEP-1 were then identified in the RA synovial fluid (Kinloch *et al.*, 2008). These findings together confirmed CEP-1 to be a true auto-antigen, as it is both diagnostically sensitive and specific, and present at the site of disease, i.e. the RA synovium.

Since discovery by our group, anti-CEP-1 antibodies have been characterised as one of the most common ACPA which occur pre-RA (van de Stadt *et al.*, 2011a), and have a strong association with carrying major histocompatibility complex type two (MHC-II) 'shared epitope' (SE) allele(s), which is a genetic risk factor

for RA. A detailed epidemiological analysis of three cohorts, Swedish Epidemiological Investigation of RA (EIRA) cohort of 1,000 RA patients and 872 healthy controls, the UK National repository (NR) of 297 RA patients, and the UK Norfolk Arthritis Register (NOAR) of 218 RA patients measured anti-CEP-1 and analysed the association between anti-CCP, anti-CEP-1 and the *HLA-DRB1* SE alleles (Mahdi *et al.*, 2009). Forty-three per cent, 42% and 27% of RA patients were anti-CEP-1 positive in the 3 cohorts respectively, and in all three cohorts, the association of *HLA-DRB1* was stronger in anti-CCP$^+$ anti-CEP-1$^+$ patients than anti-CCP$^+$ anti-CEP-1$^-$ patients. In addition, the combination of smoking and homozygous *HLA-DRB1* had extremely high OR of 53 for development of anti-CCP$^+$ anti-CEP-1$^+$ RA, compared to OR 6 for the corresponding group without anti-CEP-1, thus demonstrating a strong gene-environment interaction involved in anti-CEP-1$^+$ RA. This study established a specific association between anti-CEP-1 and the SE, which has been reproduced in a number of studies. The strongest association was observed in a cohort of 513 Korean RA patients, between anti-CEP-1 positivity and carrying 1 or more SE allele(s) (p=0.000163), which was comparable to CCP positivity (p=0.0002) (Fisher *et al.*, 2013). Studies from 330 early RA patients from the Netherlands and 521 Spanish RA patients also observed a weaker, but still statistically significant association between anti-CEP-1 and one or more SE allele (p=0.02) (Scherer *et al.*, 2011, Montes *et al.*, 2011). In a large cohort of 1,985 early RA patients in Sweden, anti-CEP-1 antibodies and SE had OR 5.1 for development of RA, which greatly increased to OR 39.1 when patients also had SE and were smokers. Anti-CEP-1 was also associated with smoking in the Korean cohort analysed by Fisher (p=0.00128) (Fisher *et al.*, 2013).

Fibrinogen as an auto-antigen

Citrullinated fibrinogen was first identified as an auto-antigen in RA by Masson-Bessiere and colleagues (Masson-Bessiere *et al.*, 2001) when probing RA synovial samples with anti-filaggrin antibodies (AFA), which are now known to cross react with fibrinogen. Sequencing identified the two substrates as peptides from the α and β chains of fibrinogen, and the reactivity of AFA was then confirmed with citrullinated α and β chains of fibrinogen *in vitro*. Two reactive peptides from fibrinogen were identified, from the α-chain ^{36}GP-Cit-VVE-Cit-HQSACKDS^{50}, and the β-chain ^{60}Cit-PAPPPISGGGY-Cit-A-Cit^{74} (Sebbag *et al.*, 2006), which are commonly used in ELISA to measure ACPA specificity.

Citrullinated fibrinogen was identified in RA synovial membranes and synovial fluid (Takizawa *et al.*, 2006), and antibodies to citrullinated fibrinogen are one of the most commonly occurring ACPA, with comparable sensitivity in RA to anti-CCP. One study analysed serum reactivity to citrullinated fibrinogen and anti-CCP in 379 early arthritis patients, 91 rheumatology clinic patients without arthritis, and 91 healthy controls (Nielen *et al.*, 2005). In this cohort anti-citrullinated fibrinogen and anti-CCP had sensitivity of 55.8% and 57.8% respectively and specificity 92.6% and 94.2% respectively.

Citrullinated proteins in the joint

A number of studies have measured the total levels of citrullination and PAD expression in the synovial tissue and synovial fluid of RA patients. De Rycke and colleagues (De Rycke *et al.*, 2005), and Kinloch and colleagues (Kinloch *et al.*, 2008) reported increased total citrullination in RA patient synovial membrane

and fluid respectively, compared to healthy patient controls by immunoblotting with anti-citrulline or anti-modified citrulline (AMC) antibodies. However, with the advancement of bioinformatics approaches and analysis of 'big-data', the next step for understanding citrullination in RA is mapping the 'citrullinome' of sera and tissues from RA patients. One study by Tutturen and colleagues laid the groundwork by using a proteomics approach to analyse the citrullinome of synovial fluid from RA patients (Tutturen *et al.*, 2014). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) before and after IgG and human serum albumin depletion in RA synovial fluid identified 119 and 157 citrullinated peptides in the native and depleted samples respectively. These included citrullinated peptides from fibrinogen, vimentin and α-enolase. However, only one RA sample and no healthy controls were used in this study. More recently Wang and colleagues measured citrullination in synovial fluid from 6 $ACPA^+$ RA patients and 6 OA patients (Wang *et al.*, 2016). Synovial fluid from RA patients contained 200 citrullination sites, in 182 citrullinated peptides, from 83 proteins, compared to 4 citrullination sites, in 3 citrullinated peptides, from 3 proteins in OA patients. Citrullinated PAD2 and PAD4 were detected in synovial fluid from both RA and OA samples, however ACPA antigens citrullinated α-enolase, fibrinogen, histones, TNC and vimentin were found in RA synovial fluid only. These studies show citrullination occurs in the synovium of both RA patients and OA patients, however the range of citrullinated peptides, including commonly occurring ACPA substrates, may be increased in the RA synovium.

The aetiology of rheumatoid arthritis

Although no definitive cause is known for RA, it is thought to fulfill the classical autoimmune paradigm, whereby an environmental factor initially triggers formation of ACPA away from the joint. These ACPA increase in titre and range of specificities until a threshold is breached and disease ensues (Figure 1.3).

Genetic associations with RA

The MHC-II shared epitope

Genetic factors contribute 50-60% to the risk of developing RA (van der Woude *et al.*, 2009, MacGregor *et al.*, 2000). A well-characterised genetic risk factor for RA is the group of MHC-II allelic variants termed the shared epitope (SE). MHC-II molecules are a noncovalent complex of an α- and a β- chain, each composed of two domains (α1, α2, β1, β2), which is expressed on the surface of APCs and B-cells. Interactions between the α- and β-chain form a peptide binding cleft, which binds processed antigens for presentation to T- cells to prime antigen specific responses. Polymorphisms in the peptide-binding cleft enable binding of a wide range of peptides, to optimise host immune responses. The SE alleles are *HLA-DR4* serotype alleles *DRB1*0401*, **0404*, **0405* and **0408*, *HLA-DR1* serotype alleles *DRB1*0101* and **0102* and *HLA-DR10* allele *DRB1*1001* (Holoshitz, 2010). These encode variants of a five amino acid sequence (Q/R-K/R-R-A-A) in the highly polymorphic 3rd hypervariable region of the β-chain binding cleft (residues 70-74) (Gregersen *et al.*, 1987). The 3D crystal structure of *HLA-DRB1*0401* is shown in Figure 1.3.

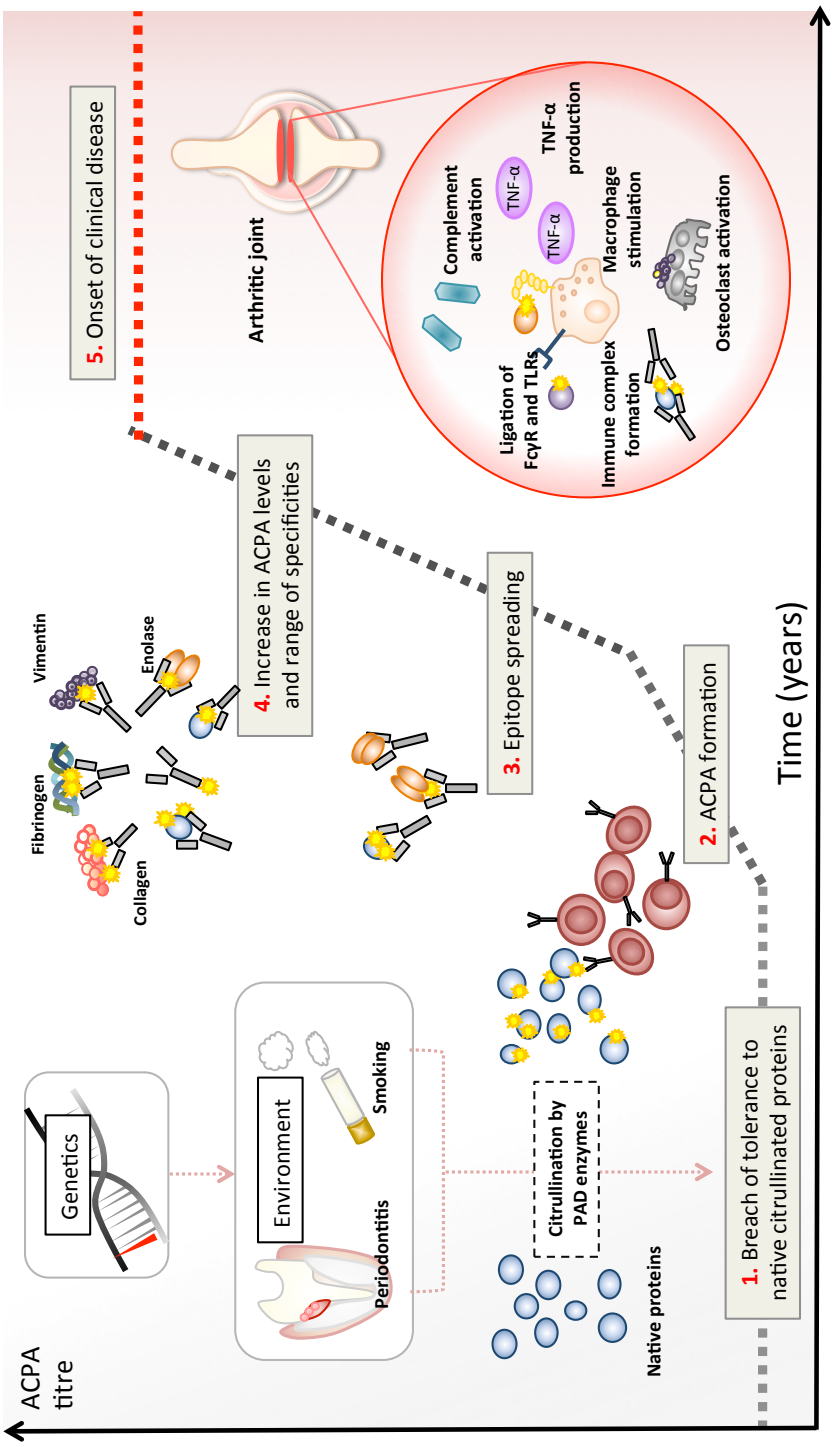


Figure 1.2 *Hypothesis for the development of rheumatoid arthritis.* 1. Breach of tolerance occurs to citrullinated proteins. 2. An ACPA response develops. 3. Epitope spreading leads to an (4.) increase in ACPA titre and range of specificities. 5. Onset of clinical arthritis. Figure designed by myself, and is adapted from the original published in (Montgomery *et al.*, 2013).

The SE was identified by Stastny, who discovered a commonly occurring *HLA-D* alloantigen in mixed lymphocyte cultures from RA patients compared to healthy or non-RA controls (Stastny, 1976, Stastny, 1978). Carrying the SE allele(s) is positively associated with RA and is predictive of a more severe form of disease (Turesson *et al.*, 2005, Gonzalez-Gay *et al.*, 2002). A popular hypothesis for this observation is that the SE favours presentation of self or arthritogenic antigens (Wucherpfennig *et al.*, 1995).

The SE is more strongly associated with ACPA$^+$ RA than ACPA$^-$ RA (Klareskog *et al.*, 2006, Huizinga *et al.*, 2005, Verpoort *et al.*, 2007), and is specifically associated with anti-cVim (Verpoort *et al.*, 2007, Scherer *et al.*, 2011), anti-cCII (Snir *et al.*, 2009), and anti-CEP-1 (Scherer *et al.*, 2011, Fisher *et al.*, 2011, Lundberg *et al.*, 2012, Montes *et al.*, 2011) antibodies.

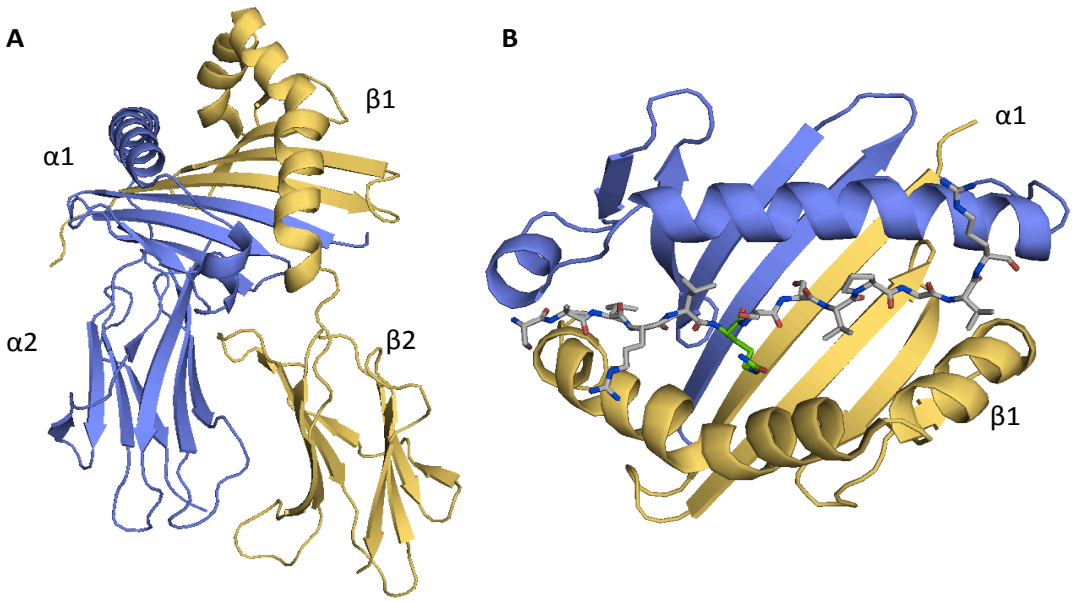


Figure 1.3 *3D crystal structure of HLA-DRB1*0401 shared epitope allele.* Adapted from (Scally *et al.*, 2013) using protein database file (pdb) 4MDI. A) Overall architecture of the molecule comprising α (blue) and β (yellow) chains. B) Peptide binding groove modelled with citrullinated vimentin peptide cVim$_{66\text{-}78}$ (SAVRA-Cit-SSVPGVR) bound. Citrulline residue is green.

MHC-II SE and citrulline specific T-cells

Dessen and colleagues compared binding of different peptides to recombinant *HLA-DRB1* alleles, and concluded the charge of residue 71 in the P4 binding pocket determined disease susceptibility alleles, i.e. SE alleles *HLA-DRB1*0401, *0404, *0101* contain positively charged residues lysine (K) or arginine (R), whereas non-associated allele **0402* contains a negatively charged glutamic acid (E) (Dessen *et al.*, 1997). By extension of this finding, the conversion of arginine to citrulline, which replaces the charged imino side-chain group with an uncharged carbonyl, could positively influence binding of citrullinated peptides to MHC-II SE allelic variants. To investigate this hypothesis Hill and colleagues measured binding of a citrullinated vimentin peptide $cVim_{65-77}$ and the native variant Vim_{65-77} to SE alleles *HLA-DRB1*0101, *0401, *0404* and non SE associated alleles *HLA-DRB1*0301, *0701, *0802, *1101, *1302* using purified HLA molecules from B-cell lines (as described by (Gorga *et al.*, 1987)) in competition assays (Hill *et al.*, 2003). $CVim_{65-77}$ bound the SE alleles **0101, *0401* and **0404* with 100-, 90- and 20- fold higher affinity respectively than Vim_{65-77}, and did not bind to any SE negative alleles. In addition, mice transgenic for the *HLA-DRB1*0401* molecule (DR4-Tg) immunised with $cVim_{65-77}$ developed an HLA-DR restricted immune response *in vivo,* whereby splenocytes purified from mice 10 days post immunisation had proliferative $CD4^+$ T-cell responses and IFN-γ secretion in response to $cVim_{65-77}$. Hill and colleagues proposed the binding of $cVim_{65-77}$ was due to neutral citrulline binding in the P4 binding pocket of the MHC-II molecule, which was later confirmed in 3D crystallography studies using *HLA-DRB1*0401* (Scally *et al.*, 2013). James and colleagues advanced upon these findings by analysing binding affinity of citrulline residues

in the P1, P4, P6, P7 and P9 pockets of SE allele *HLA-DRB*1001* using two peptides, tetanus heavy chain peptide $TT_{560-571}$, and influenza A peptide $HA_{217-236}$ (James *et al.*, 2010). Replacing amino acid residues with citrulline enabled binding in the P4 and P9 pockets, and enhanced binding in the P7 pocket. James and colleagues then measured binding affinities of various amino acid substitutions, and compiled sequences and binding affinities for P1, P4, P6 and P9 to develop a predictive algorithm for *DRB1*1001* binding, which was verified in 18 of 20 test peptides. This algorithm was used to predict binding of peptides from a number of ACPA targets to *HLA-DRB1*1001*, which identified a vimentin peptide from a similar region as used by Hill and colleagues, $cVim_{58-72}$, and two peptides from the fibrinogen α-and β-chains, $cFibA_{737-751}$ and $cFibB_{68-82}$. James then repeated these experiments to confirm the preferred binding of citrulline in the P4, P7 and P9 pockets of the *HLA-DRB1*0401* molecule, and this time identified citrullinated peptides from vimentin $cVim_{59-78}$, $cVim_{418-413}$, fibrinogen β-chain $cFibB_{69-81}$ and α-enolase $cEno_{11-25}$, $cEno_{326-340}$ predicted to bind *HLA-DRB1*0401* using prediction algorithms. These were validated in binding assays with recombinant molecules (James *et al.*, 2014).

A number of others have used similar experiments to measure binding of a citrullinated peptide of interest to *HLA-DRB1* molecules and validate immunogenicity of the peptide in DR4-Tg mice, or using RA patient PBMCs. Snir and colleagues also predicted the $cVim_{59-78}$ peptide used by James (James *et al.*, 2014), would bind MHC-II using the SYFPEITHI database of MHC binding molecules (http://www.syfpeithi.de/), and confirmed preferred binding of $cVim_{59-78}$ to *HLA-DRB1*0401* compared to the non-citrullinated variant Vim_{59-78}, in binding assays (Snir *et al.*, 2011). *In vivo*, $cVim_{59-78}$ induced splenocyte

proliferative responses in DR4-Tg mice immunised with $cVim_{59-78}$, whereas mice immunised with Vim_{59-78} showed no recall responses. Further, 2 PBMC samples from 6 RA patients, and 2 from 6 healthy patients stained positively following $cVim_{59-78}$ *HLA-DRB1*0401* tetramer staining, and upon stimulation PBMCs from RA patients had significantly increased IFN-γ and TNF-α staining. Instead of using epitope prediction, Feitsma and colleagues identified reactive peptides from vimentin by immunising DR4-Tg mice with 33 citrullinated peptides encompassing the protein, and assessing recall responses to the immunising peptides using splenocyte proliferation assays (Feitsma *et al.*, 2010). This approach identified two antigenic citrullinated vimentin peptides, $cVim_{26-44}$ and $cVim_{425-433}$. Both $cVim_{26-44}$ and $cVim_{425-433}$ induced IFN-γ from *HLA-DRB1*0401* T-cells sorted from RA PBMCs in 3 of the 10 samples tested, which was higher than responses to arginine control peptides. In a separate study, 3 citrullinated collagen II peptides were identified from a range of peptides covering the molecule for binding to *HLA-DRB1*1001* using binding assays with recombinant HLA molecules (Chemin *et al.*, 2016). In 15 *HLA-DRB1*1001* RA patients, $CD4^+$ T-cells showed significant ($p=0.0034$) activation in response to one of the CII peptides, $cCII_{311-325}$ compared to media alone, measured by CD154 upregulation. In addition, $cCII_{311-325}$ stimulation induced significantly more TNF, IL-17 and IL-10 from *HLA-DRB1*1001* PBMCs than stimulation with the non-citrullinated $CII_{311-325}$ peptide from both RA and healthy controls. Conversely, in a separate study, binding of 96 peptides spanning the fibrinogen α- and β-chains to *HLA-DRB1*0401, *0404, *0101* SE alleles was not influenced by citrullination of the peptide in *in vitro* binding assays. Further, using these fibrinogen peptides to stimulate PBMCs from RA patients and healthy controls, induced increased

proliferation in RA compared to control patients in response to peptides, but this was not influenced by citrullination of the peptide (Auger *et al.*, 2005).

These studies validate the hypothesis that the MHC-SE alleles may increase presentation of citrullinated epitopes via favourable binding of citrullinated peptides. Thus, an increased density of citrulline bound MHC-II complexes on APCs is more likely to exceed the threshold of T-cell activation to produce citrulline specific T-cell responses (DiPaolo and Unanue, 2001, Deeg *et al.*, 2013). In these studies, citrullinated peptides induced proliferation and cytokine secretion from MHC-II SE restricted T-cells in both RA patients and healthy controls (Chemin *et al.*, 2016, Snir *et al.*, 2011) suggesting the presence of citrulline specific T-cells is not exclusive to RA. The difference between citrulline specific T-cells from RA and healthy individuals may be their frequency, or the responses induced upon stimulation, such as increased cytokine response from RA PBMCs compared to healthy controls (Snir *et al.*, 2011). Peptides validated as T-cell epitopes in these studies are shown in Table 1.3.

Table 1.3 *T-cell reactive citrullinated peptides*

Protein	Region	MHC-II allele	Method of validation	Reference(s)
Vimentin	26-44	*HLA-DRB1*0401*	Murine studies with DR4-Tg mice, T-cells sorted from RA PBMCs	(Feitsma *et al.*, 2010)
	65-77	*HLA-DRB1*0101, *0401, *0404*	Binding assays, murine studies with DR4-Tg mice	(Hill *et al.*, 2003)
	58-72	*HLA-DRB1*1001*	Prediction algorithm	*(James et al., 2010)*
	59-78	*HLA-DRB1*0401*	Prediction algorithm, binding assays	(James *et al.*, 2014)
			Binding assays, DR4-Tg murine studies, RA patients PBMCs	(Snir *et al.*, 2011)
	418-413	*HLA-DRB1*0401*	Prediction algorithm, binding assays	(James *et al.*, 2014)
	425-433	*HLA-DRB1*0401*	Murine studies with DR4-Tg mice, RA T-cells sorted from PBMCs	(Feitsma *et al.*, 2010)
Fibrinogen α	737-751	*HLA-DRB1*1001*	Prediction algorithm	*(James et al., 2010)*
Fibrinogen β	69-81	*HLA-DRB1*0401*	Prediction algorithm, binding assays	(James *et al.*, 2014)
	68-82	*HLA-DRB1*1001*	Prediction algorithm	*(James et al., 2010)*
α-enolase	11-25	*HLA-DRB1*0401*	Prediction algorithm, binding assays	(James *et al.*, 2014)
	326-340	*HLA-DRB1*0401*	Prediction algorithm, binding assays	
Collagen II	311-325	*HLA-DRB1*1001*	Binding assays, RA T-cells from PBMCs	*(Chemin et al., 2016)*

The PTPN22 R620W gene variant

The *PTPN22 R620W* gene variant is also positively associated with RA (Gregersen and Batliwalla, 2005). This variant is a single amino acid substitution from arginine to tryptophan at position 620 of the PTPN22 protein. The *R620W* polymorphism in *PTPN22* leads to decreased T- and B-cell receptor signalling, which may regulate T and B cell tolerance. *PTPN22 R620W* new emigrant/transitional B-cells isolated from healthy donors are significantly more polyreactive than those without the gene variant, indicating an effect on central B-cell tolerance (Menard *et al.*, 2011).

PTPN22 R620W is also more strongly associated with ACPA$^+$ RA (Feitsma *et al.*, 2007) and RF$^+$ RA (Lee *et al.*, 2005, van Oene *et al.*, 2005), and the interactions

between both the SE and *PTPN22 R620W* gene variants in a number of different large cohorts has been collated and summarised in detail by Kallberg and colleagues (Kallberg *et al.*, 2007). In 1,977 RA patients and 2,405 controls from Swedish, Dutch, and North American cohorts they found a multiplicative interaction between the SE and the *PTPN22 R620W* gene variant in ACPA$^+$ RA. Two SE alleles and *PTPN22 R620W* has a significantly raised OR of 25.7, compared to two SE alleles alone (OR 13.7) or *PTPN22 R620W* alone (OR 1.4). In one cohort, smoking increased the association of two SE alleles and *PTPN22 R620W* from OR 18.1 in non-smokers to OR 23.4 in smokers.

Accepting the proposed mechanisms for SE and *PTPN22 R620W* involvement in RA, these findings show combining increased presentation of citrullinated autoantigens by MHC-II SE allele(s), and the disordered deletion of auto-reactive B-cells significantly increases the risk of ACPA$^+$ RA.

Environmental influences in RA

Genetic risk alleles are not solely responsible for RA, as indicated by the low RA concordance in monozygotic twins (~15%) (Aho *et al.*, 1986, Silman *et al.*, 1993). Silica exposure was first linked to RA in the 1950's following reports of rheumatoid lung in coal miners (Caplan, 1953, Miall *et al.*, 1953). Although uncommon in recent years, in a Swedish cohort of 577 male RA patients and 659 controls, prior silica exposure was moderately associated with ACPA$^+$ RA (OR 1.67) but not with ACPA$^-$ RA (OR 0.98) (Stolt *et al.*, 2010).

Surprisingly, moderate alcohol consumption may be protective against RA. In two cohorts of 1,419 Swedish and 515 Danish RA patients, the patient quartile with the highest alcohol consumption was associated with protection from RA

($p<0.001$) (Kallberg *et al.*, 2009). In a separate 7-year longitudinal study of 34,141 women in Sweden, Di Guiseppe and colleagues found a significantly decreased risk of RA (relative risk 0.63) in women who consumed ≥4 units of alcohol per week, compared to <1 unit per week ($p=0.04$), after adjustment for smoking (Di Giuseppe *et al.*, 2012)

The lung as a potential priming site for autoimmunity in RA

Smoking is a well-characterised environmental risk factor for RA, which is more strongly associated with $ACPA^+$ RA (Klareskog *et al.*, 2011). In the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) cohort of 1,205 RA patients and 872 controls, smoking contributed 35% of the total contributed risk of developing RA, which increased to 55% when combined with two SE alleles (*HLA-DRB1*01, *04 or *10*) (Kallberg *et al.*, 2011).

The association between smoking and $ACPA^+$ RA has led to the hypothesis that the lungs may act as a priming site for ACPA pre-RA. The Swedish Lung in early RA (LURA) cohort study was designed to investigate this hypothesis, by investigating changes in the lungs of 105 patients with early, untreated RA and 43 healthy controls, using high-resolution computed tomography (HRCT), and analysis of bronchial tissues and bronchoalveolar lavage (BAL) fluid (collected from 23 RA patients). HRCT showed RA patients had increased parenchymal and airway changes and fibrosis at the point of diagnosis, compared to controls, and $ACPA^+$ RA patients had increased ACPA in the BAL fluid compared to $ACPA^-$ (Reynisdottir *et al.*, 2014). A second study on this cohort found increased lymphocyte infiltration in the bronchial tissue and BAL fluid of $ACPA^+$ RA patients compared to $ACPA^-$, and formation of germinal centres in bronchial

tissue of ACPA$^+$ RA patients only (Reynisdottir *et al.*, 2015). In a separate study of BAL cells from 14 healthy smokers, and 16 healthy non-smokers, cells isolated from the lungs of smokers showed increased expression of PAD2 and citrullinated proteins compared to non-smokers (Makrygiannakis *et al.*, 2008). Further to this, Ytterberg and colleagues identified a citrullinated vimentin peptide in the lungs of RA patients, which was reactive with RA sera. Two citrullinated peptides from vimentin were identified in both bronchial and synovial biopsies of RA patients, cVim$_{446\text{-}466}$ (in 6/6 bronchial samples and 6/7 synovial samples) and cVim$_{440\text{-}445}$ (in 5/6 bronchial samples and 6/6 synovial samples), and a citrullinated vimentin peptide covering citrulline residues from both peptides (cVim$_{435\text{-}455}$) was reactive in 15% of 393 of RA patient sera, compared to 1.8% of 152 healthy controls (Ytterberg *et al.*, 2015).

Taken together, these findings suggest smoking can increase citrullination and PAD expression in the lung, which could prime an ACPA response. Although this hypothesis is appealing, recent work in our lab has shown no differences in PAD expression or citrullination in the lungs of smokers compared to non-smokers, independent of RA status. In addition, we identified the same citrullinated vimentin peptide as Ytterberg and colleagues in a number of healthy lung tissues (Lugli *et al.*, 2015).

The lung remains a viable candidate for loss of tolerance and ACPA formation, although smoking may not be the sole cause. In a study from our lab of 50 RA patients with bronchiectasis (BR RA), 50 RA patients, 122 individuals with BR, 87 with asthma, and 79 healthy controls, levels of RF, anti-CCP2, and antibodies to CEP-1, cVim, cFib and their arginine control peptides REP-1, Vim and Fib were measured (Quirke *et al.*, 2015). BR RA patients had increased RF (82% vs 52%),

anti-CCP2 (88% vs 48%), anti-CEP-1 (60% vs 24%), anti-cVim (56% vs 20%), and anti-cFib (74% vs 40%) compared to RA patients with no lung disease. Antibodies reactive with arginine control peptides REP-1 (12% vs 0%) and Vim (20% vs 2%) were also increased in BR RA patients compared to RA only patients. Further, BR patients without RA had significantly increased titre of RF, anti-cFib, anti-REP-1, anti-Vim and anti-Fib compared to healthy controls, whereas patients with asthma showed no increase in antibodies compared to controls. In addition, BR RA patients had a reduced percentage of ever-smokers compared to RA alone (42% vs 56% p=0.06), indicating these results were not caused by smoking. These findings show bronchiectasis as a model for chronic bacterial infection inducing autoimmunity to antigens associated with ACPA in RA, independently of smoking.

The gingiva as a potential priming site for autoimmunity in RA

The gingiva has been identified as a possible priming site in RA due to the historical association of RA and dental health. In the first published account in 1818, Benjamin Rush cured 'rheumatism' with removal of an infected tooth (catalogued in 'The Rush Family Papers' available from the Library Company of Philadelphia, http://dla.library.upenn.edu). By the early 1900's, focal infection theory was becoming a popular concept, whereby systemic disease was thought to originate from one infected site. As a result, tonsillectomy and tooth removal as a cure for disease became popularised by physician Frank Billings (Billings, 1912). Shortly after in 1914, Stedman and colleagues published on a case of rheumatoid arthritis cured by the removal of septic teeth (Steadman, 1914), and total dental clearance was still used as a treatment for RA into the 1950's and

1960's. Now, in the early 21st century dental hygiene and RA has remained an expanding field of research, and in 2004 Rosenstein and colleagues established the modern hypothesis for RA and periodontitis (PD) (Rosenstein *et al.*, 2004). They proposed PD may be an aetiological link to RA via four key shared pathogenic mechanisms between RA and PD, 1) shared genetic and environmental risk factors, 2) RF formation by oral pathogens (Hara *et al.*, 1996), 3) similarities between local immunologic events, including deposition of immune complexes, complement and activated phagocytes, and 4) citrullination by *Porphyromonas gingivalis* PAD enzyme.

Periodontitis

Periodontitis (PD) affects ~30% of the adult population (Brown and Loe, 1993). It results when bacteria build up from plaque and tartar to cause mild inflammation of the gum surface, known as gingivitis. If unresolved, gingivitis increases beyond a manageable level for the host immune system, resulting in PD. PD is defined by bleeding of gingival tissue on probing, and the formation of periodontal pockets (when the space between teeth and gum increases ≥4mm) caused by gum recession from the tooth surface. Severe inflammation causes breakdown of connective tissues around the teeth, resulting in tooth loss, and dissemination of oral bacterial below the gum line.

Aetiology of periodontal disease

Analysis of subgingival plaque shows groups of bacteria commonly occur together in complexes, e.g. gram negative and motile bacteria form a complex adjacent to the epithelial lining of the periodontal pocket, whilst gram-positive rods and cocci form an adherent band around the tooth surface (Listgarten, 1976, Listgarten, 1994). To investigate bacterial complexes in PD, Socransky and colleagues used cluster analysis of subgingival plaque from 165 PD patients and 20 periodontally healthy individuals to characterise bacterial complexes, and measure their associations with clinical disease (Socransky *et al.*, 1998). They observed a correlation between 'the red complex' of bacteria, consisting of *Bacteroides forsythus, Porphyromonas gingivalis* and *Treponema denticola,* with severity of PD when measured by periodontal pocket depth and bleeding on probing. *P. gingivalis* alone, or complexed with *T.denticola* and/or *B.forsythus* was associated with deepest pocket depth. Since then, the red complex has been

attributed to both the development (Haffajee *et al.*, 1998, Socransky *et al.*, 2002, Ximenez-Fyvie *et al.*, 2000), and progression of PD (Byrne *et al.*, 2009). In an early landmark study *P. gingivalis* alone could induce periodontitis in non-human primates (Holt *et al.*, 1988). Taken together, these findings have implicated *P. gingivalis* as one of the most pathogenic bacteria of the red complex and a 'keystone pathogen' in PD (Hajishengallis *et al.*, 2012). However, PD is a highly polymicrobial disease, and it has more recently been proposed that pathogenesis may be due to dysbiosis of the oral microbiome, rather than a specific causative organism (Darveau, 2010, Hajishengallis and Lamont, 2012). This shift in hypothesis may be attributable to the improvement in bacterial sequencing methods, which has increased understanding of complex polymicrobial ecosystems. Earlier targeted approaches such as DNA hybridisation and polymerase chain reaction (PCR) were sufficient for identification and quantification of specific bacterial targets, but not for exploration or comprehensive overviews of communities and their interactions. Conversely, the development of pyrosequencing 16S ribosomal ribonucleic acid (16S rRNA) genes has allowed the collection of thousands of sequences per sample and provides the power to study bacterial community composition in detail (Keijser *et al.*, 2008). 16S rRNA is a component of prokaryotic ribosomes encoded by 16S ribosomal DNA (rDNA), which can be used to reconstruct bacterial phylogenies due to the slow rate of evolution of 16S rDNA. These advances in characterising bacterial communities still support the pathogenicity of bacteria of the red complex, which increase during changes in the oral microbiota in PD (Camelo-Castillo *et al.*, 2015, Griffen *et al.*, 2012, Li *et al.*, 2014). If development of PD is due to dysbiosis of the oral microbiome, this may be due

to a deregulation in host defence (Hajishengallis and Lamont, 2012) caused by additional risk factors. Smoking is also an environmental risk factor for PD (Tonetti, 1998, Calsina *et al.*, 2002). Further, chronic PD has been associated with the *HLA-DR4* SE alleles **0401, *0404, *0405* and **0408* (Bonfil *et al.*, 1999), however a meta-analysis of 12 studies found no significant association between PD and SE alleles (Stein *et al.*, 2008).

Periodontitis and RA

There is a well documented increased prevalence of PD in RA patients (Tolo and Jorkjend, 1990, Kasser *et al.*, 1997, Mercado *et al.*, 2001, Pischon *et al.*, 2008, de Pablo *et al.*, 2008, Dissick *et al.*, 2010, Abdelsalam *et al.*, 2011, de Smit *et al.*, 2012, Scher *et al.*, 2012, Chen *et al.*, 2013, Rosamma *et al.*, 2013, Susanto *et al.*, 2013, Mikuls *et al.*, 2014), which is independent of smoking status (Kasser *et al.*, 1997, de Pablo *et al.*, 2008, Dissick *et al.*, 2010, de Smit *et al.*, 2012, Scher *et al.*, 2012) and additional cofounders such as age, gender or medication (Tolo and Jorkjend, 1990, Pischon *et al.*, 2008, Abdelsalam *et al.*, 2011, de Pablo *et al.*, 2008, Dissick *et al.*, 2010). The results from these studies are summarised in Table 1.4.

Table 1.4 *Prevalence and severity of periodontitis in rheumatoid arthritis in various studies. Table overleaf.* Table adapted from (Montgomery *et al.*, 2015). PD = periodontitis; CAL = clinical attachment loss; PdDp = probing depth; pockets = gingival pockets; ICD9-CM = International Classification of Diseases, Ninth Revision, Clinical Modification.

Table 1.4 *Prevalence and severity of periodontitis in rheumatoid arthritis.*

Periodontal status	Periodontal Evaluation	Prevalence in RA	Prevalence in control	*p* value	Reference
		n=287	***n=330 (OA)***		
Periodontitis (%)	CAL ≥ 6mm PrDp of ≥ 5mm	35□	26	0.022□	*(Mikuls et al., 2014)*
		n=75	***n=75***		*(Susanto et al., 2013)*
None/mild (%)	No PD / CAL ≥3 mm	29	31	0.547	
Severe (%)	CAL ≥6mm PrDp≥5mm	21	16	0.547	
		n=100	***n=112***		*(Rosamma et al., 2013)*
None (%)	CAL=0mm	0	18	<0.05	
Mild (%)	CAL=1-2mm	42	75	<0.05	
Moderate (%)	CAL=3-4mm	48	5	<0.05	
Severe (%)	CAL≥5mm	10	2	<0.05	
		n=13,779	***n=13,7790***		*(Chen et al., 2013)*
Periodontitis (%)	ICD9-CM:523.3-5	39	35	<0.001	
Severe (%)	ICD9-CM:523.3-4	23	20	<0.001	
		n=34	***n=18***		*(Scher et al., 2012)*
None (%)	No bleeding upon probing.	6	45	<0.01	
Severe (%)	CAL ≥5mm PrDp ≥5mm	53	22	<0.01	
		n=95	**n=420**		*(de Smit et al., 2012)*
None (%)	Pockets ≤3mm	30	70	< 0.001	
Moderate (%)	PrDp 4–5mm without gingival recession	43	18	< 0.001	
Severe (%)	PrDp ≥4mm with gingival recession	27	12	< 0.001	
		n=80	***n=80***		*(Abdelsalam et al., 2011)*
Mild (%)	CAL ≤ 1.17 mm	65	88	0.006	
Moderate/ Severe (%)	CAL > 1.17mm ≤ 3.5mm/ >3.5mm	35	12		
		n=69	***n=34 (OA)***		*(Dissick et al., 2010)*
None (%)	Healthy peridontium	19	37	<0.05	
Moderate/Severe (%)	<50% bone loss, BOP, and tooth mobility <2 to ≥2	51	26	0.03	
		n=103	***n=4358***		*(de Pablo et al., 2008)*
None (%)	Dentate and no PD	28	56	<0.001	
Periodontitis (%)	Dentate, CAL and PrDp ≥ 4mm	16	10	<0.0001	
Edentulism (%)	No natural teeth	58	34	<0.0001	
		n=57	***n=52***		*(Pischon et al., 2008)*
None (%)	CAL <4.0mm	65	90	0.001	
Moderate (%)	CAL >4mm to <6mm	23	6	0.001	
Severe (%)	CAL >6mm	12	4	0.001	
		n=65	***n=65***		*(Mercado et al., 2001)*
No/mild (%)	PrDp 0-6mm	55	75	<0.05	
Moderate/severe (%)	PrDp 6.2 to >8.2mm	45	25	<0.05	
		n=50	***n=101***		*(Kasser et al., 1997)*
Periodontitis	PrDp (mm)	2.9 ± 0.8	2.3 ± 0.4	<0.0001	
	CAL (mm)	2.6 ± 1.7	0.95 ± 0.7	<0.0001	
		n=37	***n=37***		*(Tolo and Jorkjend, 1990)*
Periodontitis	Mean bone loss (%)	23.8	18.9	<0.05	
	Mean tooth loss (n)	17.9	22.6	<0.05	

In addition to prevalence, the association between RA and PD also extends to disease severity. RA patients who have severe PD compared to moderate or no PD have increased DAS28 score ($p<0.001$) (de Smit *et al.*, 2012) and levels of periodontal bone loss correlate with swollen joint counts, health assessment questionnaire, CRP, ESR (Mercado *et al.*, 2001), and ACPA level (Gonzalez *et al.*, 2014). Further to this, periodontal therapy consisting of root scaling and polishing resulted in a reduction in DAS28 score in 17 RA patients with PD (RA PD) compared to 12 untreated RA PD patients after 8 weeks ($p<0.05$) (Al-Katma *et al.*, 2007), and in 20 treated RA PD patients compared to 20 without therapy ($p<0.001$) (Ortiz *et al.*, 2009) in two separate studies. In addition to a significant reduction in mean DAS28 score ($p<0.01$), a third study reported 10 RA PD patients also saw significant reduction in ESR ($p<0.001$) and serum-TNF levels ($p<0.05$) 1 month following PD therapy (Biyikoglu *et al.*, 2013).

These studies can be criticised for the small population size (≤20 samples per group), and short duration of follow up study (≤3 months). In addition, in both studies using RA control groups there was a non-significant trend towards a decrease in DAS28 scores in control groups (Al-Katma *et al.*, 2007, Ortiz *et al.*, 2009), indicating a placebo effect that could have contributed to the significant change in DAS28 score observed in treated groups. The DAS28 score as a marker of RA is particularly susceptible to influence by patient opinion due to the contribution of global scores for pain and overall health. In order to confirm these findings, a larger, controlled study of periodontal treatment in RA PD patients is required. As a non-biased measure of disease, the effect of PD therapy on serology in a cohort of RA patients with PD, over a 6-month period, is presented later in this thesis.

The oral microbiome in RA

The microbiome is the term given to the complex ecosystem of bacteria that survive symbiotically in the human host, and dysbiosis of this complex balance has been implicated in disease. In the gastrointestinal tract, *Helicobacter pylori* is linked to peptic ulcers, mucosal associated lymphoid tissue tumours and gastric adenocarcinoma (McColl, 2010). *Fusobacterium nucleatum* is enriched in colorectal cancers (Castellarin *et al.*, 2012, Kostic *et al.*, 2012), and both Crohn's disease and ulcerative colitis are linked to a lack of microbial diversity in the gut (Manichanh *et al.*, 2006, Hviid *et al.*, 2011, Lepage *et al.*, 2011). However, it remains unknown to what extent these findings are cause or effect of the conditions. An example of cause may be the observation that segmented filamentous bacteria in the lamina propria of the gut can induce Th17 cell production in mice (Ivanov *et al.*, 2009), and so can directly augment immune responses in the host.

Periodontitis as an environmental risk factor for RA could indicate involvement of the oral microbiome in RA. To investigate this, Scher and colleagues measured bacterial composition of subgingival plaque in early RA, chronic RA, and non-RA controls (Scher *et al.*, 2012). Although colonization with *P. gingivalis* correlated with the severity of PD, overall exposure to *P. gingivalis* was similar among the RA and non-RA groups. The role of *Prevotella* and *Leptotrichia* species merits further study, as taxa observed in patients with new-onset RA irrespective of PD status. It may be dissemination of oral bacteria or bacterial products to remote tissues, rather than their presence in the oral cavity, that is involved in the aetiology of RA. This is supported by the presence of oral bacterial DNA in synovial fluid from RA patients (Moen *et al.*, 2006, Reichert *et al.*, 2013).

Periodontitis and citrullination in the aetiology of RA

An aetiological link between PD and RA is supported by the presence of PAD enzymes, citrullinated proteins, and ACPA in the gingiva of PD patients. Immunohistochemistry of gingival tissues showed levels of citrullination, PAD2 and PAD4 were significantly increased in affected tissues from 29 PD patients compared to non-inflamed tissue from 11 periodontally healthy controls ($p<0.05$) (Harvey *et al.*, 2013). Gingival crevicular fluid (GCF) from PD patients in this study also had increased anti-CCP antibodies, although this was not statistically significant ($p=0.097$). In addition, immunohistochemistry of gingival tissues from 15 PD patients, and 6 periodontally healthy controls, using two anti-citrulline antibodies, F95 and Ab5612, identified citrullination in 80% of PD tissues compared to 33% of controls (Nesse *et al.*, 2012). Further, NETs are increased in the GCF of PD patients (Vitkov *et al.*, 2010) and are induced *in vitro* from isolated human neutrophils by *P. gingivalis* (Delbosc *et al.*, 2011). As discussed, NETosis is a potential mechanism for breach of tolerance to citrullinated proteins and so it is possible that enhanced citrullination and NETosis in the gingiva and GCF of PD patients could prime an ACPA response pre-RA. Findings from a study by Lappin and colleagues support this hypothesis, whereby 39 PD patients had increased serum anti-CCP ($p<0.0001$) and anti-CEP-1 ($p<0.001$) (Lappin *et al.*, 2013). However, a similar study suggests serum antibody responses in individuals with PD are less citrulline specific than in RA patients. In 96 PD patients, there was no significant difference in anti-CCP compared to 98 non-PD controls, but PD patients had significantly increased anti-CEP-1 ($p<0.02$) and antibodies to REP-1, and fib ($p<0.001$), the non-citrullinated variants of CEP-1 and cFib peptides (de Pablo *et al.*, 2014).

Porphyromonas gingivalis

P. gingivalis is a gram-negative, rod shaped bacterium of the *Bacteroidetes* phylum. In oral infection, *P. gingivalis* is capable of manipulating host responses to promote bacterial growth and survival using virulence factors outlined below.

Gingipains

P. gingivalis secretes a number of serine- cysteine- and metallo- proteases, including gingipains, collagenase, and gelatinase (Potempa *et al.*, 2000). Of these, the most active are the gingipains, which are responsible for 85% of the total protease activity of *P. gingivalis* (Potempa *et al.*, 1997). Gingipains are a group of three cysteine endopeptidases, specific for either arginine (RgpA and RgpB) or lysine (Kgp) and are encoded by three genes of the same name (*RgpA*, *RgpB*, *Kgp*) (Guo *et al.*, 2010). Gingipains are considered major virulence factors in *P. gingivalis* (Holt *et al.*, 1999), and have been thoroughly characterised (Potempa *et al.*, 2003, Imamura, 2003, Guo *et al.*, 2010). Key to their virulence is degradation of host proteins, which can promote bacterial survival directly, or indirectly. Direct effects of gingipains include degradation of periodontal tissues (Travis *et al.*, 1997) and plasma constituents (Vincents *et al.*, 2011) to provide nutrients for *P. gingivalis* asaccharolytic growth, and degradation of components of the host immune response to enable immune evasion.

Incubation of human sera with *P. gingivalis* reduces antibacterial activity by 60%, which is largely attributed to gingipain activity (Popadiak *et al.*, 2007). Cleavage of antimicrobial peptide LL-37 by gingipains (Bachrach *et al.*, 2008) provides *P. gingivalis* resistance to LL-37 mediated killing (Ouhara *et al.*, 2005, Guo *et al.*, 2010). Similarly cleavage of complement components by gingipains enables *P. gingivalis* to evade complement killing by all complement pathways

(Slaney and Curtis, 2008). Complement killing involves a cascade of events initiated via the classical, lectin, or alternative pathways, resulting in formation of the membrane attack complex (MAC), which binds to pathogens, perforating their membrane. The classical pathway involves binding of the C1 complex to immunoglobulins that have already bound a foreign pathogen. This results in cleavage of complement components C2 and C4 into their active fragments, C2a and C4b, which combine to form C3 convertase. C3 convertase then cleaves C3 into C3a and C3b, which initiates MAC formation from C5b, C6, C7, C8 and C9. The lectin pathway begins when mannose-binding lectin (MBL) binds foreign polysaccharides on the pathogen surface, initiating cleavage of C2 and C4 as before, and alternative activation requires spontaneously formed C3b to bind microbes directly, initiating MAC formation (Walport, 2001a, Walport, 2001b). Gingipains degrade complement components C3 and C4, and prevent deposition of C3b, C4b, and C9, to halt the cascade at various points, inhibiting all 3 pathways (Popadiak *et al.*, 2007). Gingipains also directly degrade cytokines, including membrane-bound TNF-α (Mezyk-Kopec *et al.*, 2005), IL-8 (Dias *et al.*, 2008), IL-6 (Banbula *et al.*, 1999), and the IL-6 receptor (Oleksy *et al.*, 2002), IL-4 (Yun *et al.*, 2003) and INF-γ (Yun *et al.*, 1999), which inhibits MHC-II presentation, CD4 and CD8 molecules from T-cells (Kitamura *et al.*, 2002), and IgG (Vincents *et al.*, 2011).

Further to the direct benefits of degrading components of the host immune response, gingipains mediate complex downstream effects on host pathways through modulation of one or more protein(s) upstream. Gingipains can increase vascular permeability through modulation of the kallikrein/kinin pathway. Rgp releases kallikrein from prekallikrein to generate bradykinin, and

Kgp directly liberates kinin from its precursor (Imamura *et al.*, 1995). This benefits *P. gingivalis* by increasing the acquisition of nutrients in the periodontal pocket, and is thought to contribute to the formation of GCF. Rgp also induce blood coagulation through activation of factors IX and X to activate thrombin from prothrombin (Imamura *et al.*, 1997), which in turn releases fibrin from fibrinogen to activate platelets and form a clot. It is unknown if this mechanism has a direct benefit to *P. gingivalis*, however fibrin released may contribute to alveolar bone loss via fibrin-mediated inflammation.

Fimbriae

P. gingivalis is capable of invading gingival fibroblasts and gingival epithelial cells where it can survive and replicate (Amornchat *et al.*, 2003). Different strains of *P. gingivalis* have different invasive capability measured by PCR and using flow cytometry based assays (Dorn *et al.*, 2000, Baek *et al.*, 2015), and mutation studies have shown invasion is dependent on major fimbriae (Weinberg *et al.*, 1997), which form interactions with β1 integrins for cell entry via integrin receptors (Yilmaz *et al.*, 2002). The invasive index of *P. gingivalis* strains found in infected tissues correlates with severity of PD, measured by probing depth and bone loss (Baek *et al.*, 2015), and the intracellular location of *P. gingivalis* is in close proximity to inflammatory infiltrates (Kim *et al.*, 2010), indicating the invasion of *P. gingivalis* into host cells directly contributes to pathology in PD.

P. gingivalis Enolase

Our group has proposed *P. gingivalis* enolase (PgEnolase) could be involved in autoimmunity in RA through cross reactivity with human α-enolase. PgEnolase and human α-enolase are 51% homologous across the whole protein and 82% homologous at the immunodominant CEP-1 region, and human anti-CEP-1

antibodies cross react with PgEnolase (Lundberg *et al.*, 2008). Fujinama and colleagues established the mechanisms of molecular mimicry in autoimmunity, proposing cross-reactivity between epitopes of invading pathogens and native antigens may stimulate an autoimmune response (Fujinami *et al.*, 1983, Albert and Inman, 1999). A well-characterised example is rheumatic fever induced by *Streptococcus pyogenes*. Following evidence that antibodies from rheumatic fever patients reacted with antigens from *Streptococcus*, Cunningham defined the streptococcal antigens responsible for autoantibodies in rheumatic heart disease, and reproduced symptoms of cardiac disease in rodent models (Cunningham, 2016). The anti-streptococcal antibodies reacted with cardiac myosin protein in heart cells, which was later found to cross-react with the immunodominant epitope of the streptococcal group A carbohydrate, N-acetyl-beta-D-glucosamine (GlcNAc) (Cunningham, 2000). Thus, similar mechanisms between PgEnolase and α-enolase could be responsible for formation of anti-CEP-1 antibodies in RA patients (Lundberg *et al.*, 2008).

Porphyromonas gingivalis peptidylarginine deiminase (PPAD)

P. gingivalis expresses the only known prokaryotic PAD enzyme (PPAD), which has been implicated as the connecting link between RA and PD by a number of groups including ours (Rosenstein *et al.*, 2004, Farquharson *et al.*, 2012, Mikuls *et al.*, 2009, Wegner *et al.*, 2010, Quirke *et al.*, 2014a). However, few studies have suggested a specific mechanistic link between citrullination by PPAD and RA. Our group has proposed the unique C-terminally specificity of PPAD as a potential mechanism for breach of tolerance to citrullinated proteins, and shown formation of C-terminally citrullinated peptides from fibrinogen and α-enolase

by *P. gingivalis in vitro* (Wegner *et al.*, 2010). Both α-enolase and fibrinogen are ubiquitously expressed and so would be available for citrullination by PPAD in the gingiva. In fact, fibrinogen is upregulated in gingivitis (Alexander *et al.*, 2011), making it a likely substrate for PPAD. In addition, a further study from our group showed full-length recombinant PPAD expressed and purified from *E.coli* was citrullinated. Following this finding, a second hypothesis for PPAD activity in ACPA development was proposed, whereby an immune response to 'auto-citrullinated' PPAD causes breach of tolerance to native citrullinated proteins (Quirke *et al.*, 2014a). This was supported by an increased antibody response to the citrullinated form of PPAD compared to a non-citrullinated form in a cohort of RA patients measured by ELISA. A separate study comparing recombinant PPAD purified from *E.coli*, and native PPAD in *P. gingivalis* found evidence of citrullination in the recombinant enzyme only, or following lysis of *P. gingivalis* cells, and concluded auto-citrullination of PPAD was an artefact of recombinant expression, which would not occur *in vivo* (Konig *et al.*, 2014).

PPAD has few known directly pathogenic mechanisms. Citrullination of proteins causes a net loss of charge, and can result in conformational changes to the protein structure that may alter protein function. Thus, citrullination may be used to modify proteins of the host immune system. One characterised example of this is citrullination of the complement component C5a by PPAD, which inactivates the protein and inhibits the chemoattractant capabilities (Bielecka *et al.*, 2014). Concomitant production of ammonia during citrullination also benefits *P. gingivalis* survival by maintaining a high pH to survive acid cleansing cycles in the oral cavity (Casiano-Colon and Marquis, 1988), and by directly inhibiting neutrophils (Niederman *et al.*, 1990).

P. gingivalis antibodies in rheumatoid arthritis

A number of studies have found an increased antibody response to *P. gingivalis* in RA patients. In one such study of 78 RA patients, 39 PD patients, and 40 healthy controls, 67%, 77%, and 40% respectively had serum antibodies to whole lysed *P. gingivalis*, and in RA patients, anti-*P. gingivalis* antibody level correlated with levels of IgM, IgG2 and IgG4 anti-CCP2, however not with total anti-CCP2 levels (Mikuls *et al.*, 2009). The same group then measured antibody responses to a *P. gingivalis* outer membrane antigen (PgOMA), originally characterised by Curtis (Curtis *et al.*, 1991), and *P. gingivalis* lipopolysaccharide (PgLPS) by ELISA, and levels of *P. gingivalis* in subgingival plaque by PCR in a larger cohort of 287 RA patients, and 330 osteoarthritis (OA) controls (Mikuls *et al.*, 2014). In this group, anti-PgOMA, anti-PgLPS and the level of *P. gingivalis* DNA in subgingival plaque were increased in OA or RA patients with PD, compared to periodontally healthy patients. In RA patients, anti-PgOMA antibodies correlated with both anti-CCP2 and RF, however neither anti-PgOMA nor anti-PgLPS antibodies were significantly different between RA and OA groups. In a further study by Mikuls and colleagues, anti-PgOMA antibodies were increased in individuals at risk of developing RA (first degree relatives of RA patients (FDRs) or SE positive individuals), who were anti-CCP2^{+} or RF^{+} (Mikuls *et al.*, 2012). In a separate study of FDRs, Hitchon and colleagues measured antibodies to PgLPS in relation to ACPA status in 82 RA patients and 205 FDRs (Hitchon *et al.*, 2010). Antibodies to PgLPS were significantly increased in RA patients compared to FDRs ($p=0.005$), and were significantly increased in both ACPA^{+} RA ($p<0.04$) and FDRs ($p<0.001$) compared to the corresponding ACPA^{-} group.

These studies could be criticised for using PgOMA and whole *P. gingivalis* as antigen in ELISA assays, as both are likely to contain citrullinated proteins formed by PPAD during the purification process, which may bind serum ACPA. In order to determine the antibody response to *P. gingivalis* more accurately studies in our lab have used purified recombinant proteins, PPAD and RgpB. In a cohort of 80 RA patients, 44 individuals with PD and 82 healthy controls, 25%, 11% and 5% respectively were positive for anti-PPAD antibodies using a recombinant full-length PPAD in ELISA (Quirke *et al.*, 2014a). Antibodies to RgpB in PD patients were significantly increased compared to controls ($p<0.001$) but no difference was observed between RA and controls. In a larger cohort of 1,975 RA patients, 377 non-RA controls, 65 PD patients and 59 non-PD controls, anti-RgpB antibodies were also significantly increased in PD patients compared to non-PD controls ($p<0.0001$). In addition, anti-RgpB antibodies were significantly increased in RA patients compared to non-RA controls ($p<0.0001$), and in $ACPA^+$ RA compared to $ACPA^-$ RA patients ($p<0.003$). When analysed by smoking status, anti-RgpB antibodies were more strongly associated with RA than smoking (OR 2.96 and OR 1.37 respectively) (Kharlamova *et al.*, 2016). However, in a study of serology in 103 patients pre-RA, neither anti-PPAD nor anti-RgpB antibodies were associated with risk of RA, or presence of additional RA autoantibodies (Fisher *et al.*, 2015).

To date, antibodies to PgEnolase have been measured in two studies. When characterising the CEP-1 peptide of α-enolase, Lundberg and colleagues also measured the antibody response to the same region of PgEnolase (PgCEP-1 ^{5}CKIIG-Cit-EILDS-Cit-GNPTVEC^{21}) in 102 RA sera and 130 disease controls. 34% of RA patients had anti-PgCEP-1 antibodies compared to 1% of disease controls

(Lundberg *et al.*, 2008). In a separate study, Li and colleagues isolated $CD19^{low}$ $CD27^{high}$ plasmablasts from anti-$CCP2^{+}$ RA patients, and tested their reactivity with anti-CCP2 and citrullinated peptides cFib, CEP-1 and cVim using ELISA (Li *et al.*, 2016). They found 19.5% of plasmablasts had anti-CCP reactivity, or reactivity with one or more citrullinated peptide(s), and of these, 39.5% were reactive with PgOMA and 52.6% were reactive with PgCEP-1. Anti-PgCEP-1 antibodies also correlated with anti-CEP-1. Taken together, these studies of RA serology show that antibodies to *P. gingivalis* are associated with PD and RA, although were not increased in RA patients compared to controls in all studies, or in pre-RA sera. Further, antibodies to PgEnolase and cross reactivity with human α-enolase may be involved in development of anti-CEP-1 antibodies.

Animal models in rheumatoid arthritis research

The first animal model of arthritis was the adjuvant-induced model (AIA) in rats (Pearson, 1956), following the discovery of complete Freund's adjuvant (CFA) by Jules Freund (Freund, 1956). This was first described by Stoerk who reported arthritis in rats immunised with splenic proteins emulsified in CFA, whilst his contemporaries Lipton & Freund demonstrated a breach of tolerance to the central nervous system (CNS) in rats following a similar immunisation with spinal chord tissue in CFA (Stoerk *et al.*, 1954, Freund and Lipton, 1955). AIA can be transferred to healthy irradiated recipients through passive transfer of $CD4^+$ T-cells (Taurog *et al.*, 1983, Pelegri *et al.*, 1996), and so an evolutionarily conserved protein found in the mycobacterium of CFA and the joint, such as heat shock protein (HSP) or muramyl dipeptide (MDP), may act as the T-cell target. This led to the molecular mimicry hypothesis in RA (Oldstone, 1989), which remains a concept in current RA research, including the proposed mechanism between PgEnolase and α-enolase, which will be investigated further in this thesis. AIA is a severe, rapid and self-limiting form of disease, and is therefore no longer used in the same form today. However, these experiments highlighted the potential to develop *in vivo* models of autoimmune diseases. There are a number of murine models in use today which display comparable pathology to RA including joint swelling, pannus formation, synovitis, cartilage degradation and bone erosion. These are outlined below and summarised in Table 1.5.

Collagen induced arthritis (CIA)

Collagen induced arthritis (CIA) is induced by intradermal injection of foreign type-two collagen (CII) (normally of bovine or chicken origin) emulsified in CFA.

This activates B-cells to produce anti-CII antibodies, which then react with native collagen in the joint. Resulting pathology involves infiltration of cells into synovial tissues, swelling of joints, and erosion of bone and cartilage (Holmdahl *et al.*, 1989). Development of CIA is dependent on mouse strain, e.g. C57Bl/6 have a more chronic and prolonged disease than DBA mice, driven by T-cell proliferative responses (Inglis *et al.*, 2007). CIA also has MHC-II linked genetic associations, whereby mice expressing MHC-II SE transgenes, *HLA-DRB1*0101* or *HLA-DRB1*0401* have increased susceptibility to CIA (Nabozny *et al.*, 1994, Brand *et al.*, 2003). Criticism of the CIA model is that the rate of incidence is highly variable, and can be susceptible to external factors such as animals living conditions (Holmdahl *et al.*, 1992).

Collagen antibody induced arthritis (CAIA)

The collagen antibody induced arthritis model (CAIA) uses immunisation with a commercially available cocktail of anti-collagen antibodies (Khachigian, 2006). Disease is comparable to CIA, but with more rapid onset, increased incidence, homogeneity of disease, and the capacity to use a broader spectrum of mice strains, including those with genetic modifications.

K/BxN Model

This model uses passive transfer of serum from K/BxN mice that spontaneously develop severe inflammatory arthritis (Kyburz and Corr, 2003). K/BxN mice were originally generated by crossing KRN-C57Bl/6 mice that express a transgenic T-cell receptor specific for bovine ribonuclease peptide $RNase_{42-56}$, with non-obese diabetic (NOD) mice (Kouskoff *et al.*, 1996). Transfer of serum results in transfer of autoantibodies reactive with glucose-6-phosphate isomerase (G6PI), a protein found in synovial fluid (Monach *et al.*, 2008), and

symptoms are mediated by anti-G6PI antibodies, innate immune cells, and $CD4^+$ T-cells (Kouskoff *et al.*, 1996). K/BxN mice also develop endocarditis (Binstadt *et al.*, 2009). After establishing the K/BxN model, anti-G6PI antibodies were detected in 64% of sera from 69 RA patients, compared to only 3% of 107 healthy controls, and were identified complexed with G6PI in the RA synovial fluid (Schaller *et al.*, 2001).

Tg-huTNF mice

Tg-huTNF mice were fist described by Kollias and colleagues, who used a targeting vector encoding the human *TNF-α* gene to replace the regulatory AU-rich element (ARE) of the mouse *TNF-α* gene, resulting in chronic overexpression of TNF-α and aggressive arthritis (Keffer *et al.*, 1991). Penetrance in this model is 100%, and disease is highly destructive and erosive to bone and cartilage, with homogeneous progression. However, disease is mediated by cytokines and innate cells, without the specific immune components characteristic of RA such as ACPA or RF.

DR4-Tg mice

DR4-Tg are transgenic mouse strains expressing human *HLA-DR4* SE alleles, most commonly *HLA-DRB1 *0401* (Woods *et al.*, 1994), usually with co-expression of human CD4 or chimeric mouse and human MHC II molecules due to the poor cohesion between human MHC II and mouse CD4 (Vignali *et al.*, 1992). DR4-Tg mice enable investigation of the effect of MHC-II SE alleles in susceptibility to or progression of murine models of arthritis, such as increased susceptibility to CIA (Brand *et al.*, 2003). The susceptibility to arthritis in DR4-Tg mice is particularly striking in models of RA using citrullinated antigen. Hill and colleagues administered citrullinated and native fibrinogen to both C57Bl/6

wild-type mice and DR4-Tg mice, which resulted in paw swelling and increased antibody production in 35% of DR4-Tg mice that received citrullinated fibrinogen. No pathology occurred in wild-type mice, or any mice immunised with native fibrinogen (Hill *et al.*, 2008). This study elegantly models the hypothesis for SE involvement in RA, in confirming increased antigenicity of citrullinated peptides when presented by SE expressing MHC-II molecules.

Table 1.5 *Characteristics of commonly used mouse models of arthritis.*

Model	Pathology seen in RA	B-/T-cell involvement	Disease progression	Genetic associations	Reference(s)
Collagen induced arthritis (CIA)	Yes	Antibody production Anti-collagen antibodies ACPA (contentious) RF (contentious) T-cell involvement	Acute, remitting in late disease	Yes C57Bl/6 less susceptible	*(Trentham et al., 1977)* (Courtenay *et al.*, 1980) *(Holmdahl et al., 1986)* (Inglis *et al.*, 2007)
Collagen antibody induced arthritis (CAIA)	Yes	Anti-collagen antibodies (passive transfer)	100% onset Homogeneous	No	*(Stuart and Dixon, 1983)* (Holmdahl *et al.*, 1986) (Khachigian, 2006)
KBxN	Yes	Requirement for CD4 T-cells Antibody production (not to autoantigens seen in RA)	Self limiting Homogeneous	Transgenic mouse develops spontaneous disease No requirement for passive transfer	*(Kouskoff et al., 1996)* (Matsumoto *et al.*, 1999) *(Korganow et al., 1999)* (Mangialaio *et al.*, 1999)
Tg-HuTNF-α	Yes	No autoantibodies No known T-cell role	Chronic Aggressive	Transgenic mouse develops spontaneous disease	*(Keffer et al., 1991)* (Probert *et al.*, 1996)

'Pathology seen in RA' is considered presence of joint swelling, pannus formation, synovitis, cartilage degradation and bone erosion.

Citrullination and ACPA in in vivo models

It could be argued that a major component lacking in murine models of arthritis is a robust model of ACPA development. In the most frequently used CIA model, anti-CII antibodies and pathology of the joints are commonly used markers of disease, and few studies have attempted to measure ACPA using anti-CCP2. In addition, those that have are conflicting. Vosenaar and colleagues observed citrullination of synovial proteins but no ACPA production in DBA mice (Vossenaar *et al.*, 2003), whereas Kuhn and colleagues did report increased ACPA in DBA mice with CIA (Kuhn *et al.*, 2006). However, variation in results may be due methods used to measure ACPA. Studies by Vossenaar and Kuhn used anti-CCP assays from different manufacturers, which may contain different antigenic peptides, and also used different dilution factors. Kuhn and colleagues used a 1:10 serum dilution, compared to the 1:100 used by Vosenaar, which would be more likely to detect low levels of serum ACPA.

Two studies to date have reported arthritis induction via immunisation with citrullinated CII (cit-CII) without adjuvant. DBA mice immunised with cit-CII developed significantly increased ankle swelling compared to mice immunised with native CII ($p<0.001$), which was comparable to ankle swelling in mice immunised with CII in CFA as for CIA (Thiele *et al.*, 2012). Mice immunised with cit-CII or CFA also had significantly reduced bone surface ($p=0.033$ and $p=0.027$ respectively) and bone volume ($p=0.006$ and $p=0.001$ respectively) measured by μCT compared to mice immunised with CII. Although immunisation with cit-CII did not induce significantly increased antibody responses to CII or cit-CII, these mice did have significantly increased anti-CCP ($p=0.022$) compared to CII mice, which was also measured using a 1:10 serum dilution. In a further study of this

model, mice immunised with cit-CII showed increased cartilage damage (p=0.007) measured by toluidine blue staining compared to CII immunised mice, and cartilage collagen loss (p=0.008) measured by Masson's trichrome (Dusad *et al.*, 2013). In addition, ACPA in immunised mice correlated with parameters of bone erosion. These findings support the increased antigenicity of citrullinated proteins as antigens, and display the increased bone erosion characteristic of $ACPA^{+}$ RA.

The role of citrullination in *in vivo* models of arthritis has also been investigated using PAD4 knock-out mice ($PAD4^{-/-}$). Rohrbach, Seri and Suzuki and colleagues, investigated the onset and severity of K/BxN serum transfer arthritis, glucose-6-phosphate isomerase induced arthritis (GIA) and CIA respectively in $PAD4^{-/-}$ mice compared to wild-type controls on the same genetic background. Arthritis in K/BxN mice had comparable onset and severity, measured by histology, in both $PAD4^{-/-}$ and wild-type mice (Rohrbach *et al.*, 2012), whereas both GIA and CIA induced reduced clinical score and antibody titre in $PAD4^{-/-}$ mice compared to wild-type controls (Seri *et al.*, 2015, Suzuki *et al.*, 2016). These findings suggest citrullinated by PAD4 is not required for the effector phase of passively induced arthritis, but is involved in onset and development of inflammatory arthritis in these *in vivo* models.

P. gingivalis in murine models

To investigate the association between *P. gingivalis*, PD and RA *in vivo*, two studies by Marcheson, Gully and colleagues treated mice orally with *P. gingivalis* to induce PD (for 15 and 14 days respectively) as per the model of PD previously described by Polak (Polak *et al.*, 2009), before induction of CIA or CAIA

(Marchesan *et al.*, 2013, Gully *et al.*, 2014). Marcheson and colleagues observed an increased incidence and severity of CIA in mice that had been previously infected with *P. gingivalis* compared to those with CIA alone. Further, mice administered both *P. gingivalis* and CIA had significantly increased paw swelling ($p=0.006$, $p=0.004$ medial lateral and dorsal ventral measurements respectively), and increased osteoclasts in the paws ($p<0.005$) compared to CIA only mice (Marchesan *et al.*, 2013). In addition to pre-treatment with *P. gingivalis*, Gully and colleagues administered a PPAD deficient *P. gingivalis* strain and a vehicle only as negative controls. In mice previously inoculated with wild-type *P. gingivalis*, CAIA had increased speed of onset, paw swelling ($p=0.0001$), and bone erosion ($p<0.04$) measured by μCT, compared to CIA mice administered the vehicle only control (Gully *et al.*, 2014). No significant differences were observed between mice immunised with vehicle control and PPAD deficient *P. gingivalis*. In addition, mice treated with wild-type *P. gingivalis* had increased anti-*P. gingivalis* antibody formation compared to the PPAD deficient strain ($p=0.02$).

Maresz and colleagues also examined the effect of prior infection with *P. gingivalis* on CIA using the chamber coil model, which uses implantation of a *P. gingivalis* filled chamber subcutaneously in mice, to distribute bacteria continuously (described by (Genco *et al.*, 1991)) (Maresz *et al.*, 2013). In addition to live *P. gingivalis*: heat killed *P. gingivalis*, periodontal pathogen *Prevotella intermedia*, a citrullinated *P. gingivalis* capsule protein, and a PPAD null strain of *P. gingivalis* were implanted into the chambers as controls. In this study active *P. gingivalis* only enhanced incidence ($p<0.001$) and clinical score ($p<0.001$) of CIA compared to CIA alone, thus confirming specific effects of live *P. gingivalis* and PPAD. In addition, mice with live *P. gingivalis* chamber implants

had increased anti-CCP compared to CIA alone ($p<0.05$). This small number of studies shows *P. gingivalis* increases severity of disease in both CIA and CAIA, which is dependent on PPAD expression, and therefore may be due to citrullination.

Peptidylarginine deiminase enzymes (PADs)

There are five known human PAD isoforms, PAD1-4 (previously PAD5) and PAD6 encoded by the *PADI1-4* and *PADI6* genes located on a single gene cluster on chromosome 1 (1p36) (Iida and Nakamura, 2004). Human PADs are constitutively expressed and citrullination of specific protein targets is required for a number of functions. PAD substrates with known functions are briefly outlined in Table 1.6.

PAD1 is found in the epidermis and uterus. Its main substrate is keratin K1, which is citrullinated in the cornification of epidermal tissues into keratinised materials (Nachat *et al.*, 2005). In humans, this process is essential in formation of hair and nails.

PAD2 is expressed in secretory glands, brain, uterus, spleen, and skeletal muscle. It is also found in neutrophils and macrophages. PAD2 substrates are myelin basic protein (MBP) in tissues of the central nervous system, vimentin in skeletal muscle and macrophages, and β-/γ- actin in neutrophils (Darrah *et al.*, 2012). In multiple sclerosis, a neurodegenerative disease defined by destruction of the protective myelin sheaths around nerve axons, citrullination of MBP in brain tissues is increased. This is thought to alter the myelin sufficiently to increase its susceptibility to degradation by proteases, causing direct loss of nerve protection and exposing neoepitopes for T-cells to potentiate disease (Moscarello *et al.*, 2013). In addition, the conversion of arginine to citrulline results in a loss of charge, which may also inhibit nerve axon signal conductivity (Cao *et al.*, 1999). In breast cancer, PAD2 expression correlates with HER2 protein levels in $HER2^+$ breast tumours (McElwee *et al.*, 2012).

PAD3 is expressed in the epidermis where it co-localises with profilaggrin/filaggrin and regulates epidermal homeostasis (Nachat *et al.*, 2005). In hair follicles, it is co-localised with trichohyalin, a structural protein in the follicle inner root sheath and medulla, which provides mechanical strength. PAD3 is also an auto-antigen in RA (Seaman *et al.*, 2016, Darrah *et al.*, 2013).

PAD4 is expressed in granulocytes, monocytes, macrophages, and neutrophils (Zhang *et al.*, 2004) and is required for formation of NETs, indicated by the inhibitory effect of PAD4 inhibitors on NET formation (Lewis *et al.*, 2015). It is the only PAD isoform to possess a nuclear localisation factor found at the N-terminus of the protein and thus is found in the cell nucleus, where it regulates gene expression through citrullination of histones H2A, H3 and H4 (Nakashima *et al.*, 1999).

PAD6 is expressed in cells and tissues of the urinary and reproductive systems including ovary, testis, embryonic stem cells and oocytes. Citrullination by PAD6 regulates oocyte skeletal sheet formation, and is essential in regulation of female fertility and early embryonic development (Esposito *et al.*, 2007).

Table 1.6 *Proteins citrullinated by PADs in healthy tissue.*

Protein	Location	Function of citrullination	Reference(s)
Fibrinogen	Peripheral blood	Homeostasis of the coagulation cascade	*(Nakayama-Hamada et al., 2008)* (Okumura *et al.*, 2009)
Filaggrin	Epidermis	Skin moisturisation	(Rawlings and Harding, 2004)
		Cornification of epidermis	(Senshu *et al.*, 1996)
Histones	Various cells	Regulation of gene expression	*(Wang et al., 2009)* *(Christophorou et al., 2014)* *(Li et al., 2010b)*
		NET formation	*(Wang et al., 2009)*
Keratin	Epidermis	Cornification of epidermis	*(Senshu et al., 1996)*
Myelin basic protein (MBP)	Central nervous system	Modulation of myelin basic protein methylation for correct myelin sheath formation	*(Pritzker et al., 2000)*
Nuclear lamin C	Various cells	Apoptosis pathway	(Tanikawa *et al.*, 2012)
Nucleophosmin	Various cells	Apoptosis pathway	(Hagiwara *et al.*, 2002)
Oocyte cytoplasmic protein (unknown)	Ovary	Embryonic development/fertility	*(Esposito et al., 2007)*
Trichohyalin	Hair follicles	Mechanical strength	*(Steinert et al., 2003)*
Vimentin	Monocytes, macrophages	Cell death	(Majumder *et al.*, 2011)
		Apoptosis pathway	(Asaga *et al.*, 1998)

Peptidylarginine deiminases in RA

Genome wide linkage studies have identified single nucleotide polymorphisms (SNPs) on the *PADI* gene associated with RA. Two separate studies identified the D1S228 and D1S214 regions in genome scans of 114 RA sibling pairs from 97 European Caucasian nuclear families (Cornelis *et al.*, 1998), and 41 sibling pairs from Japanese families (Shiozawa *et al.*, 1998) respectively. Suzuki and colleagues then summated the gene region in close proximity to both marker

regions was 1p36, which contains the *PADI* genes encoding PAD1-4 and PAD6. In 830 RA patients and 736 unaffected Japanese individuals Suzuki and colleagues found *PADI4*, not neighbouring *PADI*s was a susceptibility locus for RA (p=000008). Identifying 8 single nucleotide polymorphisms (SNPs) within the *PADI4* gene associated with RA, which equated to 4 possible haplotypes (Suzuki *et al.*, 2003). Since this study, these four haplotypes have been investigated in a number of other studies and have shown a positive association with RA in Korean (Kang *et al.*, 2006) and Indian (Panati *et al.*, 2012) cohorts but not in Chinese (Chen *et al.*, 2011), Spanish (Martinez *et al.*, 2005), French (Caponi *et al.*, 2005), or British (Barton *et al.*, 2004, Burr *et al.*, 2010) cohorts. A similar investigation into SNPs in the *PADI2* gene identified 3 RA associated SNPs in a cohort of 267 RA, 51 ankylosing spondylitis (AS) and 160 healthy control patients from northern China. An association with RA was then confirmed in a second cohort of 307 RA, 324 AS and 509 control patients for 2 SNPs ($p<0.0008$ and $p<0.0033$) (Chang *et al.*, 2013).

In RA patients, both PAD2 and PAD4 are found in the RA synovium (Kinloch *et al.*, 2008), where PAD2 expression level correlates with DAS28 ($p<0.005$) (Damgaard *et al.*, 2016). PAD2 isolated from the synovium is active, and can citrullinate synovial fluid mast cells (Arandjelovic *et al.*, 2012) and fibrinogen (Damgaard *et al.*, 2014). PAD4 is also an established auto-antigen in RA, with anti-PAD4 antibodies found in 23-40% of RA patients in a number of studies from Japan, Norway, China, and the USA (Takizawa *et al.*, 2005, Halvorsen *et al.*, 2009, Zhao *et al.*, 2008, Harris *et al.*, 2008). PAD4 is auto-citrullinated and anti-PAD4 antibodies bind both native and citrullinated forms (Andrade *et al.*, 2010). In addition, a subset of anti-PAD4 antibodies increases activity of the enzyme

when bound by lowering the Ca^{2+} activation threshold (Darrah *et al.*, 2013). Therefore formation of anti-PAD4 antibodies in RA could increase levels of citrullination, and increase ACPA formation in a self-propagating mechanism. This hypothesis is supported by observations that anti-PAD4 antibodies are found in more severe disease, measured by DAS28 (Zhao *et al.*, 2008) and radiographic progression (Halvorsen *et al.*, 2009, Harris *et al.*, 2008). A subset of anti-PAD4 antibodies that are cross-reactive with PAD3 have been found in 12-18% of RA patients in one study, and are also associated with a more destructive disease phenotype (Darrah *et al.*, 2013).

In addition to RA, increased or disordered citrullination by PADs has been implicated in a number of diseases summarised in Table 1.7.

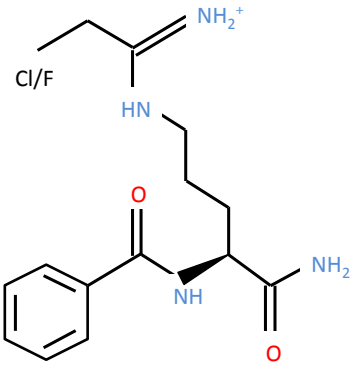
Table 1.7 *Peptidylarginine deiminase enzymes associated with disease.*

Disease	Associated PAD(s)	Mechanisms	Reference(s)
Alzheimers disease	PAD2 & PAD4	Abnormal accumulation of citrullinated proteins in CNS	(Ishigami *et al.*, 2005)
Prion disease	PAD2	Clustering of abnormal citrullinated proteins in CNS, may be due to increase Ca^{2+}	(Jang *et al.*, 2013)
Multiple Sclerosis (MS)	PAD2 & PAD4	Citrullination of myelin basic protein, increases degradation of MBP Loss of charge on citrullinated proteins inhibits nerve axon signal transduction	(Moscarello *et al.*, 2013) (Cao *et al.*, 1999)
Psoriasis	PAD1	Loss of citrullination of keratin leads to excessive cornification	(Ishida-Yamamoto *et al.*, 2000)
Rheumatoid arthritis	PAD2, PAD3 & PAD4	Increased PAD4 translation	(Suzuki *et al.*, 2003)
		Expression of PAD2 and PAD4 in RA joint	(Vossenaar *et al.*, 2004) (Kinloch *et al.*, 2008) (Damgaard *et al.*, 2016)
		Production of anti-PAD3 and anti-PAD4 antibodies associated with severe disease Production of pathogenic ACPA	(Kolfenbach *et al.*, 2010) (Darrah *et al.*, 2013)
Ulcerative colitis	PAD4	Associated with PAD4 haplotype composed of three SNPs *PADI4*_92_G, *PADI4*_96_C, and *PADI4*_102_C	(Chen *et al.*, 2008)
		ACPA formation (pathology unknown)	(Chumanevich *et al.*, 2011)
Cancer	PAD4	Increased PAD4 in blood and tissues of malignant cancer patients Repression of p53 tumour suppressor	(Chang *et al.*, 2009) (Jones *et al.*, 2009)
Glaucoma	PAD2	Citrullination of myelin basic protein disrupts optic nerve myelination	(Bhattacharya *et al.*, 2006)

PAD inhibitors

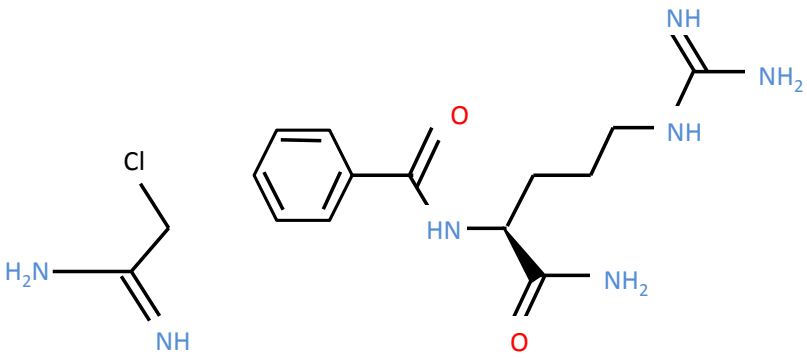
The first described inhibition of PAD was using paclitaxel, a chemotherapeutic agent that targets microtubule formation to inhibit mitosis. Using paclitaxel, Pritzker found a weak inhibition of PAD isolated from the bovine brain (IC_{50} ~5mM) (Pritzker and Moscarello, 1998), which may be due to the similarity between the paclitaxel side chain and Nα-Benzoyl-L-arginine ethyl ester (BAEE), which is an efficient PAD substrate. Benzoyl-arginine derivative Benzyl-N^{γ}-

monomethyl-Arg (Bz- ADMA) has also been shown to inhibit PAD4 with an IC_{50} of 400 μM (Hidaka *et al.*, 2005). However the most potent PAD inhibitors to date have been developed in the lab of Thompson and colleagues. These are F-amidine [N-α-benzoyl-N^5-(2-fluoro-1-iminoethyl)-1-ornithine amide] and Cl-amidine [*N*-α-benzoyl-N^5-(2-chloro-1-iminoethyl)-1-ornithene amide], which inhibit PAD4 with IC_{50} of 21.6μM and 5.9μM respectively (Luo *et al.*, 2006a, Luo *et al.*, 2006b) (Figure 1.4). Both F-amidine and Cl-amidine preferentially inhibit the active calcium bound form of PAD4, which may help limit toxicity *in vivo* and increase PAD4 inhibition at sites of inflammation where Ca^{2+} is increased.



Cl/F-amidine

Figure 1.4 *Chemical structure of PAD inhibitors Cl-amidine/F-amidine.*



2-chloroacetamidine Benzoyl arginine amide (BAA)

Figure 1.5 *Chemical structures of 2-chloroacetamidine and bonzoyl arginine amide (BAA).*

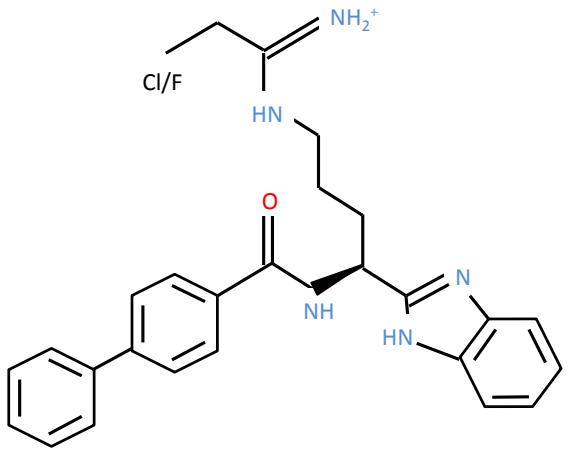
Both F-amidine and Cl-amidine are analogues of benzoylarginine amide (BAA) (Figure 1.5), a small molecule substrate for PAD4 (Kearney *et al.*, 2005), with the substrate guanidinium group exchanged for a haloacetamidine warhead. The warhead alone, 2-chloroacetamidine (Figure 1.5) was shown to inhibit PAD4 by Stone and colleagues who concluded it may act as a suitable chemical scaffold from which to develop more highly selective and high affinity inhibitors (Stone *et al.*, 2005). 2-chloroactetamidine is also the most effective inhibitor of PPAD tested in our lab (unpublished findings). Both F-amidine and Cl-amidine inhibit PAD4 by alkylation of active site residue Cys645, although mechanisms have not been elucidated in detail (Jones *et al.*, 2009).

PAD inhibitors in models of disease

Use of F-amidine and Cl-amidine *in vivo* has confirmed PAD4 involvement in: NET formation (Wang *et al.*, 2009), regulation of tumour suppressor p53 (Li *et al.*, 2008), and apoptosis (Yao *et al.*, 2008). In addition, F-amidine and Cl-amidine have been successful in reducing severity of disease in a number of mouse models. In the dextran sulphate sodium (DSS) mouse model of colitis Cl-amidine given either prophylactically or as a treatment regimen reduced symptoms of colitis by driving apoptosis of immune cells in the associated tissues (Chumanevich *et al.*, 2011). Lupus prone MRL/lpr mice treated with Cl-amidine daily for 8-14 weeks showed reduced NET formation compared to controls, improved vascular function via endothelium mediated vasorelaxation, reduced immune complex deposition in the kidneys and reduced skin involvement (Knight *et al.*, 2013, Knight *et al.*, 2015). In $Apoe^{-/-}$ mice prone to atherosclerosis, mice treated daily with Cl-amidine had reduced atherosclerotic

lesion size, with reduced recruitment of neutrophils and macrophages (Knight *et al.*, 2014). In the CIA mouse model of RA, treatment of mice following disease onset reduced severity by ~50% and reduced synovial citrullination, complement deposition and antibody formation (Willis *et al.*, 2011).

BB Cl-amidine (figure 1.6) is a second-generation form of Cl-amidine, which has the C-terminal replaced with a benzimidazole ring to prevent proteolysis, and the N-terminal replaced with a biphenyl moiety to increase hydrophobicity and enhance cellular uptake of the molecule. In a recent study from our lab, treatment of CIA with BB Cl-amidine post onset reduced clinical score ($p<0.0001$ after 5 days), paw swelling ($p<0.0001$), and citrullination in the lymph nodes of treated mice compared to vehicle treated controls (Kawalkowska *et al.*, 2016). Lymph nodes from mice treated with BB Cl-amidine had reduced Th1 cells and increased Th2 cells compared to naïve mice, and BB Cl-amidine treatment *ex vivo* suppressed Th17 cells. These results show potential for PAD inhibition as a treatment strategy for RA. However, the high sequence homology between PADs may pose a risk of toxicity, or immunosuppression.



BB Cl-amidine

Figure 1.6 *Chemical structure of second-generation PAD inhibitor BB Cl-amidine.*

Structure and activity of mammalian PADs

Of the 5 PAD isoforms, PAD2 and PAD4 are most relevant to RA. PAD2 is a 75kDa protein of 665 amino acids, and PAD4 is 64kDa and 663 amino acids. However, PAD4 exists as a dimer, which increases enzymatic activity (Liu *et al.*, 2011), and so may not exist *in vivo* at 64kDa. The 3D crystal structures of PAD2 and PAD4 have been solved, in active Ca^{2+} bound conformation (Slade *et al.*, 2015, Arita *et al.*, 2004), enabling structural and functional characterisation. Both PAD2 and PAD4 are composed of three domains, two immunoglobulin-like fold (IgLF) domains comprised of aa1-295 in the PAD2 molecule, and aa1-300 in the PAD4 molecule, and a catalytic domain (CD) from aa262-665 and aa301-663 in PAD2 and PAD4 molecules respectively. IgLF domains are made up of β-sheets, turns, and some short α-helices, whereas the CD is an α-/β- propeller structure of ββαβ repeats. The CD houses the active site, composed of a nucleophilic cysteine, a histidine which acts as a general acid and base for proton transfer and two aspartic acid residues involved in substrate stabilising. Both the active site of PAD2 and the active site of PAD4 are surrounded by Ca^{2+} binding sites, which are required for enzymatic turnover. Five Ca^{2+} binding sites are conserved between all PADs, and PAD2 has one additional, i.e. 6 in total. The nomenclatures used in publications by Slade, Arita, and their colleagues differ slightly, but it can be surmised from both studies that initial binding of two (for PAD2) or three (for PAD4) Ca^{2+} ions stabilises the enzyme, but does not enable it catalytically. Binding of the remaining four or two Ca^{2+} ions (PAD2 and PAD4 respectively) then induces a confirmation change to enable enzymatic turnover.

Activity of bacterial deiminases

PPAD belongs to the guanidino modifying superfamily of enzymes (GMEs) with arginine deiminase (ADI), L-arginine:glycine amidinotransferase (AT), N^{ω}, N^{ω}-dimethylarginine dimethylamino hydrolase (DDAH), agmatine deiminase (AIH) and arginine succinyltransferase (AstB). Like the mammalian PAD, GMEs are characterised by a key nucleophilic cysteine (Rodriguez *et al.*, 2010), required to modify the arginine or arginine analogue substrate (Linsky and Fast, 2010). However, PPAD is distinct from human PADs in a number of ways: 1) a lack of requirement for Ca^{2+} ions for enzymatic activity, 2) an ability to citrullinate both free L-arginine and peptidyl-arginine residues, 3) C-terminal substrate specificity, and 4) lack of sequence homology with mammalian PADs.

Aims of the study

The aim of this project was to elucidate the link between *P. gingivalis* infection and RA by characterising PPAD *in vitro* and determining the immune response to PPAD, RgpB, and PgEnolase *in vivo* using patient cohorts and murine models.

Chapter 2: Expression, purification and crystallisation of recombinant PPAD

1. Develop a productive recombinant expression system for PPAD.
2. Solve the 3D crystal structure of PPAD and use structural information to investigate mechanisms of action.

Chapter 3: Characterisation of PPAD in vitro

3. Characterise PPAD activity *in vitro.* Using truncated PPAD isoforms, investigate the domains required for enzymatic function.
4. Compare the activity of PPAD to PAD2 and PAD4.

Chapter 4: PPAD, RgpB and PgEnolase in vivo: murine models

5. Measure antibody production and splenocyte responses in mice immunised with PPAD, RgpB, and PgEnolase.
6. Investigate antibody responses to PgEnolase as a mechanism for breach of tolerance to α-enolase.

Chapter 5: PPAD, RgpB and PgEnolase in RA: human cohorts

7. Measure antibody reactivity to *P. gingivalis* proteins in RA patient sera, and compare antibody responses to PgEnolase and human α-enolase.
8. Determine the effect of non-surgical PD therapy on anti-*P. gingivalis* antibody and ACPA levels in RA and PD patients.

Chapter 2 Expression, purification and crystallisation of PPAD

Background

Citrullination by PPAD during *P. gingivalis* infection has been implicated in the aetiology of RA. Because of this, a small number of studies have aimed to investigate PPAD activity and function using enzymes isolated from the *P. gingivalis* bacterium, or purified from recombinant systems (McGraw *et al.*, 1999, Rodriguez *et al.*, 2009, Abdullah *et al.*, 2013). PPAD is a four-domain protein comprising the N-terminal signal peptide (NtSP); the catalytic domain (CD); the immunoglobulin-like fold (IgLF); and the C-terminal domain (CTD). It contains 556 amino acids, with an expected molecular weight of 62kDa.

McGraw and colleagues purified PPAD from the culture supernatant of a mutant *P. gingivalis* strain, and were the first to demonstrate that PPAD is secreted from *P. gingivalis* in a truncated form of ~46kDa (McGraw *et al.*, 1999). They then characterised the optimal temperature of enzymatic activity of PPAD (55°C), the effects of freezing the enzyme (not detrimental), and efficacy of inhibitors: *N*α-L-arginine methyl ester (L-NAME), thiocitrulline, thiourea, Tosyl-L-lysyl-chloromethane hydrochloride (TLCK), leupeptin, antipain and cysteine. Leupeptin was most effective, completely ablating activity at 5mM, thiocitrulline, cysteine and thiourea also fully inhibited activity at 12.5mM, 25mM and 50mM respectively. TLCK reduced activity by 98% at 12.5mM, and antipain and L-NAME were only 50% effective at the highest concentrations used. Rodriguez and colleagues cloned and expressed both full-length PPAD, and a truncated recombinant form missing 43aa from the N-terminus, comparable to the 46kDa

secreted form. They reported the truncated form was significantly more active and stable than the full-length (Rodriguez *et al.*, 2009).

Although these studies have been successful in characterising aspects of PPAD, they did not provide mechanisms for the properties observed. In comparison, structural determination of mammalian PAD2 and PAD4 (Slade *et al.*, 2015, Arita *et al.*, 2004) has been invaluable in the development of PAD inhibitors, which have been used as successful prophylaxis or treatment in a number of *in vivo* disease models, including colitis, lupus, atherosclerosis and RA (Chumanevich *et al.*, 2011, Knight *et al.*, 2013, Knight *et al.*, 2014, Willis *et al.*, 2011, Kawalkowska *et al.*, 2016).

Essential to the hypothesis that citrullination by PPAD could be involved in development of ACPA are the multiple properties of PPAD that are distinct from human PADs, such as specificity for C-terminal arginine substrates. However, the lack of structural information from PPAD has made the mechanisms involved in these functional differences entirely speculative, as direct comparison to mammalian PADs cannot be made.

To this end, the main aims of this chapter are:

- To develop an efficient expression system for recombinant PPAD and an inactive PPAD mutant in *E.coli.*
- To solve the 3D crystal structure of PPAD.
- To use 3D structural information to propose mechanisms of action for PPAD, and contrast differences to mammalian PADs.

Buffers

Purification of full-length GST-tagged PPAD (flPPAD)

Glutathione elution buffer	50mM Tris HCl pH7.6
	10mM reduced glutathione
Lysis buffer	90% BugBuster Mastermix
	10% glycerol
	20µl DNAse/ 3L LB culture volume

Purification of truncated PPAD (tPPAD) constructs

Binding Buffer (BB)	500mM NaCl
	5% glycerol
	50mM 4-(2-hydroxyethyl)-1-
	piperazineethanesulfonic acid (HEPES) (pH7.5)
	20mM imidazole
Wash Buffer (WB)	500mM NaCl
	5% glycerol
	50mM HEPES (pH7.5)
	40mM imidazole
Elution Buffer (EB)	500mM NaCl
	5% glycerol
	50mM HEPES (pH7.5)
	250mM imidazole
Gel Filtration Buffer (GF)	500 mM NaCl
	5% glycerol
	10mM HEPES (pH 7.5)

Measurement of PPAD/PAD activity

PPAD buffer	50mM N-cyclohexyl-2-aminoethanesulfonic acid (CHES) (pH 9.5)
	10mM dithiothreitol (DTT)
PAD buffer	200mM Tris-HCl (pH 7.4)
	20mM CaCl2
	10mM DTT

Colourimetric detection reagents (CDR)

CDR A	80mM 2,3-Butanedione oxime (DAMO)
	2mM thiosemicarbazide (TSC)
CDR B	3M 85% H_3PO_4
	6M 96-98% H_2SO_4
	2mM NH4Fe(SO4)2·12H2O (ammonium ferric sulphate dodecahydrate)

Western blotting

Loading buffer	1x NuPAGE® LDS sample buffer (LifeTechnologies)
	100mM DTT

Materials and methods

Expression and purification of recombinant full-length GST-tagged PPAD (flPPAD)

A recombinant expression system for wild type and inactive mutant GST-His-tagged full-length PPAD (flPPAD^{WT}, flPPAD^{C351A}) was previously established in the Venables lab (Quirke *et al.*, 2014a). Briefly, an overnight (o/n) culture was prepared by inoculating a scraping from glycerol stocks into sterile lysogeny broth (LB) (ThermoFischer) with 30µg/ml kanamycin monosulphate (Kan) (Sigma), and incubated at 37°C, shaking at 180 rotations per minute (rpm) o/n. Desired volumes of LB (30µg/ml Kan) were inoculated with 10ml/L of o/n culture and incubated at 37°C shaking at 180 rotations per minute (rpm) until the desired optical density at 600nm (OD600) of 0.6-0.8 was reached, indicating bacterial growth was in logarithmic phase. Protein expression was induced using a final concentration (final conc.) of 0.1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Calbiochem) o/n at 18°C. Cells were harvested from LB by centrifugation at 3500rpm for 15min at room temperature (rt). Cell pellets were snap frozen at -80°C and thawed to disrupt cells before lysis using a 5:1 ratio of lysis buffer:pellet (v/w). Soluble cell extract (SCE) was isolated from total cell extract (TCE) by untracentrifugation at 16,000g for 40min at 4°C, collected and filter sterilised.

Stock glutathione sepharose 4B resin slurry (80%) (GE Healthcare) was washed x3 in 1x phosphate buffered saline (PBS) before resuspension to a working concentration of 50% slurry in 1x PBS. 1ml 50% slurry was used per 6L of starting culture. Resin was added to SCE and incubated rotating for 1hour rt to bind GST-tagged-proteins. Resin was filtered from flow through (FT) on

disposable columns (Pierce). Resin was washed x3 with column volumes of PBS. Protein was eluted from the column in 5 steps using 1ml of reduced glutathione elution buffer (see 'Buffers') and collected. Protein concentration was assessed by Bradford assay using Pierce™ Coomassie Plus (Bradford) assay kit (Thermo Scientific).

Development of novel PPAD expression system at the Structural Genomics Consortium (SGC)

Development of PPAD constructs, expression, purification and crystallisation was done in collaboration with the SGC, and carried out in their laboratory at the University of Oxford. Researchers involved are credited where applicable.

Construct design

PPAD regions for cloning were designed by Dr W. W. Yue (SGC) using the mGenTHREADER structural prediction program, and structural information from the enzyme agmatine iminohydrolase, as the nearest comparable protein to PPAD available. Truncations at the N- and C-terminals were designed to produce ordered domains, without affecting protein confirmation or function.

Cloning

Cloning and mutagenesis were carried out with/by Dr N. Burgess-Brown and Ms. L. Shrestha.

The desired PPAD sequence was amplified from the full-length PPAD coding sequence of *P. gingivalis* strain W83 using forward and reverse primers with upstream (5'TACTTCCAATCCATG) and downstream (5'TATCCACCTTTACTGTCA) 5' extensions to enable ligation independent cloning (LIC). A full list of primers used is available in appendix. Forward and reverse primers were diluted to 5μM

(in 20-50µl) and templates to 2.5ng/µl. A PCR 'MasterMix' was prepared as below and 21µl added to each well of a 96 well plate. Primers (1.5µl) and templates (2.5µl) were then added to corresponding wells.

PCR MasterMix 1

	x1	**x100**
Pfx buffer (10x stock)	2.5µl	250µl
50mM $MgSO_4$	0.5µl	50µl
10mM dNTP mix	0.75µl	75µl
Platinum Pfx	0.2µl	20µl
Water	17.05µl	1.705ml
Primers (5uM)	1.5µl	
Template DNA (2.5ng/µl)	2.5µl	

Standard PCR conditions were used:

94°C, 5 min
(94°C, 30 sec; 52°C, 30 sec; 68°C, 1-3 min) x30 cycles
72°C, 10 min
15°C hold

3µl of each reaction was removed and diluted in 12µl of double distilled water (ddH_2O) for analysis on a 1.5% ethidium bromide gel, with 20µl of E-gel marker (Invitrogen) (diluted 1:1 with ddH_2O). Correct PCR products were identified based on correct size, and any necessary repeats carried out until all successful products were obtained. PCR products were purified using a PCR purification plate (Millipore) and eluted in 50µl Tris- ethylenediaminetetraacetic acid (EDTA) (TE).

Plasmid ligation

DpnI Treatment: To degrade the remaining DNA template, 1µl of DpnI (diluted 1:20 in NEBuffer 2 (New England Biolabs) was added to each product and incubated for 1 h at 37°C. The enzyme was then inactivated at 80°C for 20min.

Vector Digestion: 5µg of each plasmid (pNIC28-Bsa4 and pGTVL2) was linearised with 3µl BsaI, 10µl NEBuffer 3 (New England Biolabs), and 1µl BSA

(New England Biolabs) in a final volume of 100µl ddH_2O, and incubated for 2 hours at 50°C.

T4 Treatment: PCR products and digested vectors were treated with T4 polymerase to enable LIC. Products and vectors were mixed as below and incubated for 30 min at 22°C then 20 min at 75°C in a thermocycler.

PCR products	**Vectors**
5µl DNA	50µl BsaI-digested plasmid
1µl NEBuffer 2	10µl NEBuffer 2
0.25µl Novagen T4 DNA polymerase	2.5µl Novagen T4 DNA polymerase
1µl **dCTP** 25mM	10µl **dGTP** (25mM)
0.5µl 100mM DTT	0.5µl 1M DTT
0.2µl BSA (NEB)	1µl BSA (NEB)
2.05µl water	21.5µl water
Total: 10µl	Total: 100µl

Ligation: 1µl of vector was ligated with 2µl of insert for 30 min at rt, then transferred to ice.

Transformation

40µl of competent MACH-1 cells were added to the ligated vector/insert and incubated on ice for 30min, heat shocked at 42°C for 45 seconds and chilled briefly before adding 100µl of super optimal broth with catabolite repression (SOC) medium (in house). Cells were covered with porous adhesive film and incubated for 60min at 37°C.

100µl of transformed cells were plated onto LB agar with 50µg/ml Kan, containing 5% sucrose for SacB gene selection, and incubated o/n at 37°C.

One colony was picked from each plate for colony PCR screening using the following primers: pLICF and pLICR for pNIC28-Bsa4, and pGEX5' and pLICR for pGTVL2.

Colony screen

Forward and reverse primers were mixed to 10uM and MasterMix prepared as below.

PCR MasterMix 2	**x1**	**x100**
Bioline 10x buffer	2µl	200µl
Bioline $MgCl_2$	0.6µl	60µl
DMSO (100%)	0.6µl	60µl
dNTPs (2mM)	2µl	200µl
Primer mix (10uM each)	2µl	200µl
Bioline Tag Red Polymerase	0.4µl	40µl
Water	12.4µl	1.24ml
Total	20µl	2ml

Individual colonies were inoculated into 20µl of reaction mix in wells of a 96 well plate for PCR, and the inoculating loop was immediately transferred to corresponding positions in a 96 well block with 1ml LB with 50µg/ml Kan. Plate was covered with ABgene adhesive film and the block with porous film. Cultures in block were incubated o/n at 37°C 700rpm and 330µl glycerol was added for long-term storage at -80°C.

DNA was amplified from colonies by thermocycling as below:

94°C, 10 min
(94°C, 30 sec; 50°C 30 sec; 72°C, 1-3 min) x25
72°C 5 min
15° hold

10µl of PCR products were analyzed on an ethidium bromide agarose gel to confirm correct product size.

A DNA mini-prep kit (Qiagen) was used to extract DNA from the MACH-1 cloning strain, which was transformed into in house *Escherichia coli* Rosetta cells BL21(DE3)-R3-pRARE2 cells using the transformation protocol as before.

Site directed mutagenesis

To generate $tPPAD^{C351A}$, cysteine residue 351 was replaced with alanine (A), to generate mutants $tPPAD^{R152/4A}$ and $tPPAD^{R152/4Q}$, arginine residues Arg152 and

Arg154 were replaced with alanine (A) or glutamine (G), and to create tPPAD^{R152A} and tPPAD^{R154A} the two arginine residues were mutated to alanine (A) separately. Site directed mutagenesis using the megaprimer method was carried out as described (Sarker & Sommer, 1990) to create these mutants.

Mutagenic primers were designed by Ms. L. Shrestha using vector NTI and ExPASy programs (http://www.expasy.org/). A 20-30mer primer sequence around the mutation was used as the forward primer, and a reverse created using sequence massager program (http://www.attotron.com/cybertory/analysis/seqMassager.htm). These were then synthesised for use in PCR. For the first round of PCR, separate reaction mixtures were prepared with 5' primer & mutagenic reverse or 3' primer & mutagenic forward as below:

HercII buffer (5x stock)	5µl
dNTPs (10mM)	0.75µl
HercII polymerase	0.25µl
Water	15µl
Primers (5µM) (LIC & Mut)	1.5µl
Template (2.5ng/µl)	2.5µl
Total	25µl

Thermocycler conditions were used as below:

94°C, 10 min
(95°C, 30 sec; 50°C, 30 sec; 68°C, 1 min*) x20
68°C, 5 min
15°C hold

50µl of reaction volume was analysed on ethidium bromide agarose gel to confirm correct product size. PCR products from round 1 of PCR create 'Megaprimers' for the second round of PCR, ie Mut forward and LIC reverse act as 'MegaR', LIC forward and Mut reverse act as 'MegaF'. The most successful PCR reaction determined by agarose gel was used for PCR 2. PCR products were purified using the InVitrogen (Life Technologies) PureLink quick PCR

purification kit as per manufacturers instructions. For the second round of PCR the reaction was prepared with MegaF primer & 3' primer or 5' primer & MegaR primer as below:

HercII buffer (5x stock)	5µl
dNTPs (10mM)	0.75µl
HercII polymerase	0.25µl
Water	15µl
Primers (LIC & Megaprimer)	1.5µl (2.5µl of each diluted in 45µl of water)
Template DNA (2.5ng/µl)	2.5µl
Total	25µl

Thermocycler conditions were used as below:

94°C, 10 min
(95°C, 30 sec; 50°C, 30 sec; 68°C, 1.5 min*) x20 cycles
68°C, 5 min
15°C hold

Correct product size was confirmed on ethidium bromide agarose gel as before. PCR products were purified, and further cloning and purification strategies for all constructs and mutants remained the same.

PPAD expression

Following cloning, test expression was carried out on a small scale (50ml) to assess protein expression and solubility levels, before large scale (6-12L) expression and purification of viable constructs. Both follow the same protocol unless stated otherwise.

An o/n culture was made by inoculating 2ml/250ml terrific broth (TB) (with 1:1000 kan & chloramphenicol (chlor)) with an ice scraping from the glycerol stock of desired expression construct and incubated o/n at 37°C.

The o/n culture was checked for dense growth and inoculated 1:100 in TB (with 1:1000 kan/chlor), which was incubated at 37°C shaking at 180rpm until OD600 of 0.6-0.8 was reached (2-3hours (h)), at which point temperature was reduced. When cultures reached 18°C protein expression was induced with a final conc. of

0.1mM IPTG and incubated o/n at 18°C. Cells were harvested by centrifugation at 3000rpm for 15min. Pellets were resuspended in 2:1 (v:v) binding buffer (BB) with 1:1000 (v:v) EDTA free protease inhibitors (Merck-Millipore). Cell pellets from small-scale cultures were lysed by sonication for 10min at 35% power (10sec pulse, 10sec off). To lyse large-scale pellets, liquefied pellet was passed through a mechanical cell breaker at 15KPa (Mark 30) 5-6 times until lysis was confirmed by colour and texture change. A TCE sample was kept for SDS-PAGE before the SCE was isolated by centrifugation at 4,500rpm for 10min (small scale) or 35,000g for 60min (large scale). SCE was collected and a sample kept for Sodium Dodecyl Sulphate-PolyAcrimide Gel Electrophoresis (SDS-PAGE) before purification.

PPAD purification

Talon Co^{2+} resin (Clontech) was used for affinity purification. Required volumes were prepared as detailed below.

Talon resin preparation: Talon resin binds 10mg protein/ml resin on average. The required bed volume was estimated based on similar previous experiments and prepared. Storage ethanol was removed and resin washed x3 by centrifugation (3,000rpm for 1min) before final resuspension 1:1 in BB.

Small scale: Resin was added to the column and allowed to settle before lysate was added and run though. Column was washed with 2x 500µl BB, 2x 500µl wash buffer (WB). Protein was eluted with 250µl elution buffer (EB) and stored.

Large scale: Resin was added to SCE and rotated for 1-2h at 4°C to 'batch bind'. SCE incubated with resin was added to a column and allowed to flow through while the resin settled. Flow through was collected as 'FT'. Two 10x column

volume washes with BB were applied to the column and collected as 'BB1', 'BB2', two 10x column volume washes with WB were applied and collected as 'WB1', 'WB2', and five 2x column volume wash of EB washes were applied and collected as 'E1-E5'.

Sodium Dodecyl Sulphate-PolyAcrimide Gel Electrophoresis (SDS-PAGE)

TCE, SCE, and FT were diluted 1:20 in BB. All other fractions were analysed without dilution. 5µl of loading buffer (see 'Buffers') was added to 15µl of each sample. Samples, were heated to 95°C for 5min before being resolved on 12-14% Novex® NuPAGE® Bis-Tris gels (LifeTechnologies) in 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (ThermoFisher), with 5µl BioRad protein marker, at 175V for 40min. SDS-PAGE gels were stained with Coomassie blue (ThermoFisher) to identify proteins.

Tobacco Etch Virus (TEV) protease cleavage

A 20µl sample of eluted protein was incubated o/n at 4°C with 2µl tobacco etch virus (TEV) protease (in house) to test cleavage of the His-tag. A sample of protein with and without protease was resolved by SDS-PAGE to determine if cleavage was successful.

To cleave His-tag from the whole protein sample, 1mg protease per 20mg protein was added and incubated o/n before Co^{2+}-NTA affinity chromatography was repeated to remove the cleaved fusion tag and TEV protease.

Size exclusion chromatography

For further purification by size exclusion chromatography proteins were concentrated until a volume ≤3ml was reached. A Superdex s200 column was used in an ÄktaExpress system which had been calibrated with gel filtration (GF)

buffer according to manufacturers protocol. Protein samples were fractionated through the column, and the resulting chromatogram was analysed for peaks representing proteins of the target proteins molecular weight. Fractions corresponding to these peaks were resolved by SDS-PAGE and stained with Coomassie (ThermoFisher) to identify protein content. Fractions containing PPAD protein with no major contaminants were pooled and stored.

Characterisation of tPPAD^{WT} and tPPAD^{C351A}

Differential scanning fluorimetry (DSF)

Stability of PPAD in varying buffer conditions (pH, NaCl concentration) and temperatures was measured by differential scanning fluorimetry (DSF). 1µl of tPPAD (5mg/ml) was combined with 1µl of 100x SYPRO® Orange protein stain (Sigma), and added to 18µl of buffer of the desired composition. A buffer system described by Newman (Newman, 2004) was used, which comprises seven broad-range buffer systems, which have stable pH with varying chemistry. The DSF method is detailed in the SGC protocol (Niesen *et al.*, 2007). Briefly, melting temperature (T_m°C) is measured as midpoint of protein denaturation. Measurement of T_m°C uses a fluorescent dye that is quenched in aqueous solution, but becomes highly fluorescent in non-polar environments, such as the hydrophobic sites on unfolded proteins.

PPAD activity assay- colourimetric detection of citrulline

The colourimetric assay for citrullination activity was used as previously described (Rodriguez *et al.*, 2009). Briefly, specified concentrations of PPAD constructs or human PAD2 or PAD4 were incubated in duplicate in 96 well Co-star plates with 20mM N_α- benzoyl-L-arginine ethyl esther hydrochloride (BAEE)

(Sigma) or 5mM concentration of fibrinogen peptides FibA-R (CESSSHHPGIAEFPS-R), or FibA-R-XX (CESSSHHPGIAEFPS-R-GK) (Pepceuticals), or short synthetic peptides (Arg-Ala, Gly-Arg, Arg-Gly-Gly, Met-Arg-Phe) in PPAD or PAD buffer (see 'Buffers) for 1h at 37 °C. Substrate only wells were used for blank measurements and a standard curve of 0-500μM of L-citrulline was included for calculations. Citrulline produced was measured by addition of 200μl/well of developing reagent, prepared using a 1:3 ratio of developing reagent A to reagent B, (see 'Buffers') and incubated at 95°C for 15 min. The plate was cooled briefly on ice before absorbance was measured at 540nm. Citrulline produced was calculated using absorbance reading of the L-citrulline standard curve in Microsoft Office Excel.

Standardisation of protein concentration

For all assays protein concentration was standardised using the NanoDrop A280 module and/or Bradford assay.

Crystallisation of recombinant tPPAD^{WT} and tPPAD^{C351A}

Prior to crystallisation, recombinant proteins were concentrated by centrifugation at 3,500rpm at 4°C using 10kDa molecular weight (mw) cut off spin concentrator columns until concentration >10mg/ml was achieved. Crystallisation plates of PPAD constructs were set up using Mosquito crystallisation robot (TPP Labtech) and incubated at room temperature (20°C) and 4°C. The vapour diffusion method was used with sitting drops in 2:1, 1:1 and 1:2 protein: precipitant ratios, incubated over a 20μl reservoir of precipitant. A range of commercially available precipitant screens with different chemical compositions was used to identify crystallisation conditions. Crystals of

tPPAD^{WT} and tPPAD^{C351A} were obtained by mixing 150nL of protein solution (12mg/ml and 15mg/ml respectively) with 300nL precipitant solution (30% PEG5000MME, 200 mM ammonium sulfate and 100 mM MES pH 6.5) and (22% PEG3350, 100 mM lithium sulfate and 100 mM bis—tris pH 5.5) for tPPAD^{WT} and tPPAD^{C351A} respectively. All crystals were cryoprotected with 25% ethylene glycol solution mixed with reservoir buffer and cryocooled in liquid nitrogen.

Structural determination of tPPAD^{WT} and tPPAD^{C351A}

Structural determination was carried out by Dr J. Kopec at the SGC. All images of structures, analyses made, or conclusions drawn in this work were made by me.

The diffraction data were collected at Diamond Light Source (Didcot, UK). The diffraction data were processed using XIA2 (Evans, 2006, Kabsch, 2010, Sauter *et al.*, 2004, Winter *et al.*, 2013), and scaled with Aimless (Evans and Murshudov, 2013) in the CCP4 suite (Collaborative Computational Project, 1994). The structure was solved by molecular replacement using sequence alignments with >100 proteins, identified using a homology detection program (https://toolkit.tuebingen.mpg.de/hhpred), and fragments identified using the robetta server (https://www.phenix-online.org/documentation/reference/mr_rosetta.html#setting-up-for-a-run-of-mr-rosetta-a-fragment-files-from-the-robetta-server), using phenix.mr_rosetta software. Coot (Emsley and Cowtan, 2004, Emsley *et al.*, 2010) and refmac5 (Murshudov *et al.*, 2011) were used for further building and refinement, respectively.

The tPPAD^{WT} structure shows a covalent bond between C351 and Cζ of arginine, confirmed by mass spectrometry. The refinement restraints for this link were calculated using JLigand and used in the late stage of refinement in refmac.

Structural modeling of tPPAD, PAD2 and PAD4

3D images were created and edited using MacPyMol. For tPPAD^{WT} and tPPAD^{C351A} coordinates, Protein Data Base (pdb) entries deposited by our lab, 5AK7 and 5AK8 respectively were used. For PAD2 and PAD4 models were created based on the pdb entries 4N20 (Slade *et al.*, 2015) and IWD9 (Arita *et al.*, 2004) respectively. For superimposition of PPAD domains with agmatine deiminase, dimethylarginine dimethylaminohydrolase and fibronectin type III protein pdb entries 3HVM, 2CI3 and 4WTX were used respectively.

Sequence alignment

Sequence alignments were carried out using the Kalign algorithm at http://www.ebi.ac.uk/Tools/msa/ or ClustalW at http://www.ebi.ac.uk/Tools/msa/clustalw2.

Statistical analysis of PPAD activity

Statistical differences between activities of various PPAD clones were calculated using an unpaired t-test in Microsoft Excel. Differences were considered statistically significant if $p \leq 0.05$.

Results

Purification of recombinant full-length PPAD (flPPAD)

An *E.coli* based expression system for a full-length, GST-tagged recombinant form of both wild-type (flPPAD^{WT}), and inactive mutant (flPPAD^{C351A}) PPAD was previously established in the Venables group. Initial expression and purification using these flPPAD$^{WT/C351A}$ constructs yielded low quantities of protein with high levels of contaminants (Figure 2.1A). One major contaminant was an excess production of free GST fusion protein, which suggests a mutation in the plasmid, and resulted in significant difficulties in purifying the protein and removing the fusion tag. flPPAD^{WT} purified using this system was active in citrullinating arginine analogue BAEE, producing ~250μM citrulline/μg enzyme (Figure 2.1B).

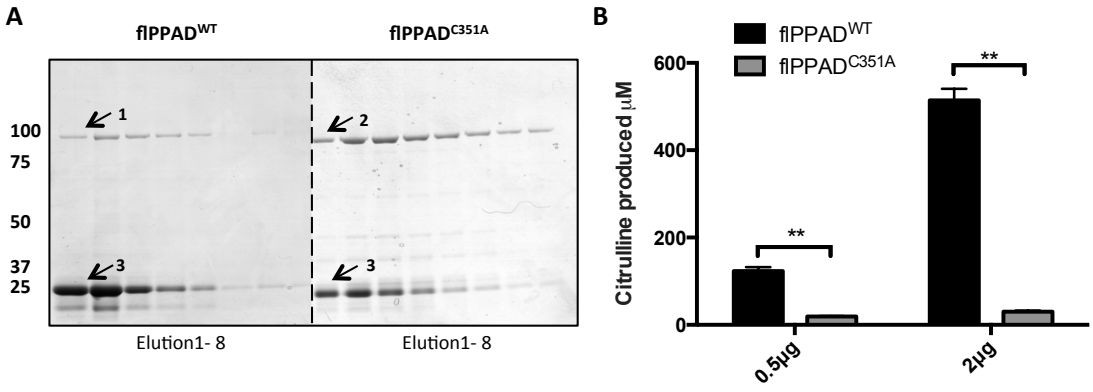


Figure 2.1 *Purification of full-length, GST-tagged, recombinant PPAD.* A) flPPAD^{WT} and flPPAD^{C351A} were purified from lysed *E.coli* using glutathione sepharose resin, and eluted in 8x 1ml washes with reduced glutathione elution buffer (Elution1-8). Protein was analysed on Coomassie stained SDS-PAGE gel. flPPAD^{WT} and flPPAD^{C351A} migrated ~100kDa (arrows 1 and 2) with free GST fusion protein which migrated ~25kDa (arrow 3). B) Activity of flPPAD^{WT} and flPPAD^{C351A} was measured by colourimetric detection of citrulline assay. flPPAD was incubated with 20mM BAEE for 1h at 37°C and citrulline produced measured. Values are n≥2±SD **p≤0.01.

Development of novel PPAD expression systems for structural determination

The flPPAD$^{WT/C351A}$ protein obtained from GST-tagged expression systems was not suitable for use in crystallisation due to poor expression levels and low purity. At this point collaboration was formed with the Structural Genomics Consortium, University of Oxford (SGC), to use their high throughput structural biology platform. The SGC pipeline comprises 7 main component steps, which create a highly efficient path to protein crystallisation (Figure 2.2).

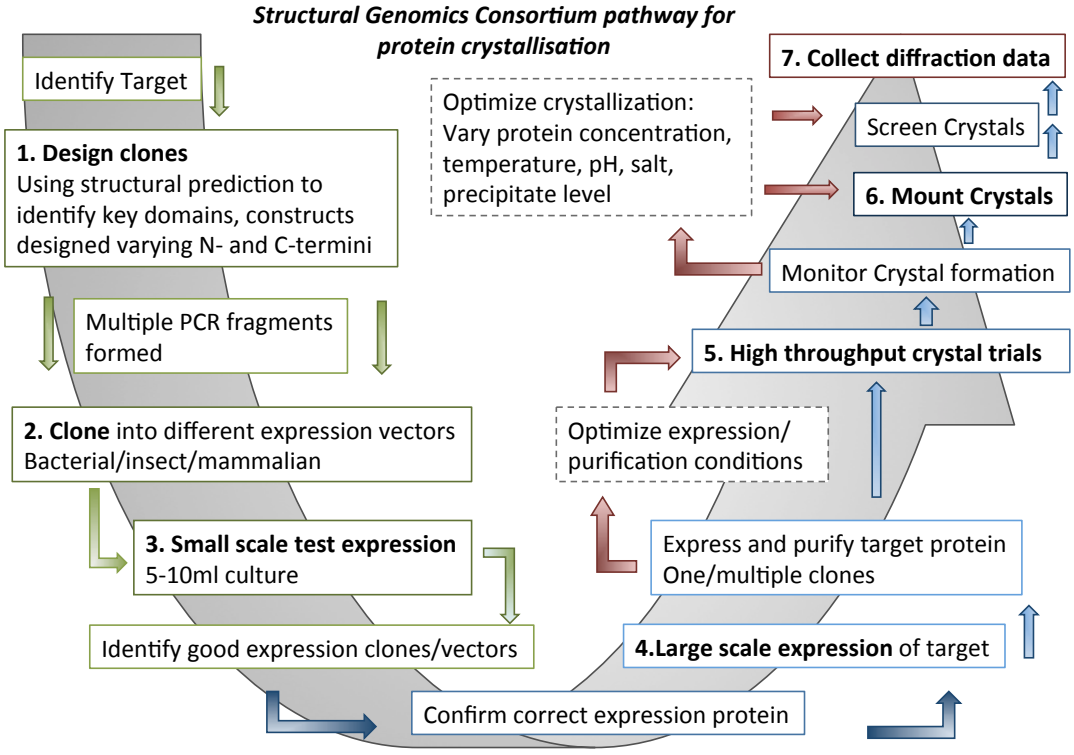


Figure 2.2 *High-throughput Structural Genomics Consortium pipeline for protein crystallisation.*

PPAD construct design

Structural information from agmatine iminohydrolase (AIH) was used to design PCR fragments around the 4 main domains of PPAD, namely the N-terminal signal peptide (NtSP) (aa1-49), catalytic domain (CD) (aa50-360), immunoglobulin-like fold (IgLF) (aa361-463), and C-terminal domain (CTD) (aa464-556) (Figure 2.3). Four forward (blue arrows on Figure 2.3) and 3 reverse (red arrows on Figure 2.3) primers were synthesised to create 7 different N- and C-terminal boundaries. In the first phase of construct design these were used in a number of combinations to create 12 protein constructs (c001-c012), which were either: full-length, NtSP and CD, or the CD only. These constructs were designed with the aim to maintain protein activity but create ordered domains that would crystallise readily. For each of the 12 constructs, corresponding sequences containing the C351A mutation known to inactivate the enzyme were also designed.

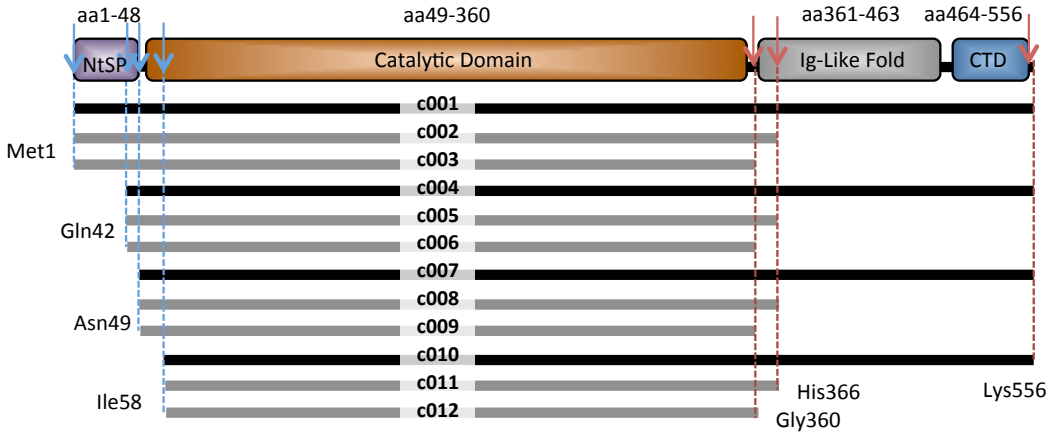


Figure 2.3 *'Phase 1' construct design for PPAD.* Full-length PPAD was used as a template to design a library of novel constructs with varying N- and C-terminal truncations. Combinations of four forward and three reverse primers (marked with blue and red arrows respectively) created twelve PPAD domains.

Cloning, transformation and test expression of PPAD constructs

Following construct design, flPPAD^{WT} was used as a template from which to amplify PCR fragments. To maximise chances of expression, DNA sequences were ligated into two pET-based expression vectors: pGTVL2, which incorporates a His_6 and GST tag at the N-terminus followed by a TEV cleavage site, sites for LIC, and the SacB gene for negative selection on 5% sucrose, and pNIC28-Bsa4, which has a His_6 tag in a 22aa N-terminal fusion peptide with a TEV cleavage site. The vector also includes sites for LIC and a SacB gene. Plasmids were then transformed into competent MACH-1 cells to confirm ligation by colony PCR, and verified clones transformed into Rosetta cells for expression. All constructs were expressed and purified in parallel on a small scale to identify constructs with high expression and solubility, for use in the large-scale purification stages (Figure 2.4).

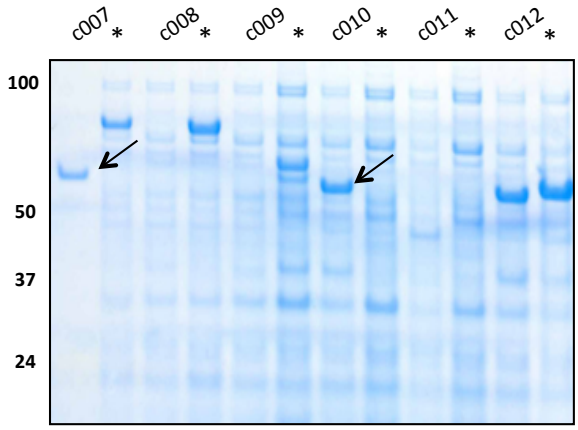


Figure 2.4 *Small-scale test expression of 'phase 1' constructs.* Proteins were expressed in 1ml *E.coli,* purified using Ni^{2+} resin and eluted in one step with elution buffer. Protein expression in elutions was determined by SDS-PAGE and Coomassie staining. Figure is a representative gel depicting a select number of PPAD constructs (c007-c012). * Represents unrelated proteins tested on the same gel. Two constructs that purified successfully are indicated with arrows as construct 7 and 10 (c007 and c010).

PPAD construct design 'Phase 2'

From small-scale expression two near full-length constructs, construct c007 and construct c010, expressed medium/high levels of soluble protein (Figure 2.4) suitable for large-scale purification. Both c007 and c010 and their corresponding C351A mutant constructs were purified readily but did not crystallise. Therefore a second phase of construct design was undertaken ('Phase 2') with the aim to maintain the high expression and solubility of c007 and c010 but promote successful crystallisation.

Based on the results from the first phase of purification trials, the pNIC28-Bsa4 vector only was used. Similarly, the N-terminal truncations were not altered, although additional C-terminal boundaries were designed to crop the CTD. This created an additional 8 constructs (c073-c080) (Figure 2.5). In total, 57 wild type and mutant PPAD constructs were cloned, expressed and purified (Table 2.1). In phase 2, construct c073 was found to yield high levels of soluble protein, which was purified easily, and was used in further studies. Following identification as a candidate for crystallisation, the C351A mutant form of c073 was also synthesised. For clarity, from here throughout this work these constructs are named truncated PPAD wild-type ($tPPAD^{WT}$) and truncated PPAD C351A mutant ($tPPAD^{C351A}$).

Table 2.1 *Summary of PPAD constructs created.* Table overleaf. Protein sequences identified during construct design were cloned into two vectors and transformed into Rosetta expression cells. Expression of all clones was tested in small-scale purification and protein yield from all clones were compared to identify possible candidates for crystallisation. Clones discussed in this chapter are highlighted in yellow.

Table 2.1 *Summary of PPAD constructs created.* Full legend on previous page.

	Amino Acid Sequence	pNIC28-Bsa4 Vector				pGTVL2 Vector			
		Construct ID	Fusion Tagged mass (Da)	Protein mass (Da)	Test Yield	Construct ID	Fusion tagged mass (Da)	Protein mass (Da)	Test Yield
Wild Type Phase I	Met1-Lys556	c001	64283	61817	0: None	c025	90025	61817	0: None
	Met1-His366	c002	43647	41182	0: None	c026	69389	41182	0: None
	Met1-Gly360	c003	42829	40364	0: None	c027	68571	40364	0: None
	Met41-Lys556	c004	59899	57433	N/A	c028	85641	57433	0: None
	Met41-His366	c005	39263	36798	0: None	c029	65005	36798	0: None
	Met41-Gly360	c006	38445	35980	0: None	c030	64187	35980	0: None
	Asn49-Lys556	**c007**	**59038**	**56572**	**3: Med**	c031	84780	56572	0: None
	Asn49-His366	c008	38402	35937	0: None	c032	64144	35937	0: None
	Asn49-Gly360	c009	37584	35119	0: None	c033	63326	35119	0: None
	Ile58-Lys556	**c010**	**58178**	**55712**	**3: Med**	c034	83920	55712	5: High
	Ile58-His366	c011	37542	35077	0: None	c035	63284	35077	0: None
	Ile58-Gly360	c012	36724	34259	0: None	c036	62466	34259	0: None
Cys351-Ala351 Mutant	Met1-Lys556	c013	64251	61785	5: High	c037	89993	61785	3: Med
	Met1-His366	c014	43615	41150	0: None	c038	69357	41150	0: None
	Met1-Gly360	c015	42797	40332	0: None	c039	68539	40332	0: None
	Met41-Lys556	c016	59867	57401	5: High	c040	85609	57401	5: High
	Met41-His366	c017	39231	36766	0: None	c041	64973	36766	0: None
	Met41-Gly360	c018	38413	35948	0: None	c042	64155	35948	0: None
	Asn49-Lys556	c019	59006	56540	5: High	c043	84748	56540	5: High
	Asn49-His366	c020	38370	35905	0: None	c044	64112	35905	0: None
	Asn49-Gly360	c021	37552	35087	0: None	c045	63294	35087	0: None
	Ile58-Lys556	c022	58146	55680	3: Med	c046	83888	55680	5: High
	Ile58-His366	c023	37510	35045	0: None	c047	63252	35045	0: None
	Ile58-Gly360	c024	36692	34227	0: None	c048	62434	34227	0: None
	Asn49-Ala484	**c081/tPPAD^{c351a}**	**51324**	**48858**	**5: High**				
Wild Type Phase II	**Asn49-Ala484**	**c073/tPPAD^{WT}**	**51356**	**48890**	**5: High**				
	Asn49-Thr470	c074	49898	47432	3: Med				
	Asn49-His354	c075	36985	34519	0: None				
	Asn49-Thr341	c076	35496	33030	0: None				
	IIe58-Ala484	c077	50496	48030	1: Low				
	Ile58-Thr470	c078	49038	46572	1: Low				
	Ile58-His354	c079	36125	33659	0: None				
	Ile58-Thr341	c080	34636	32170	0: None				

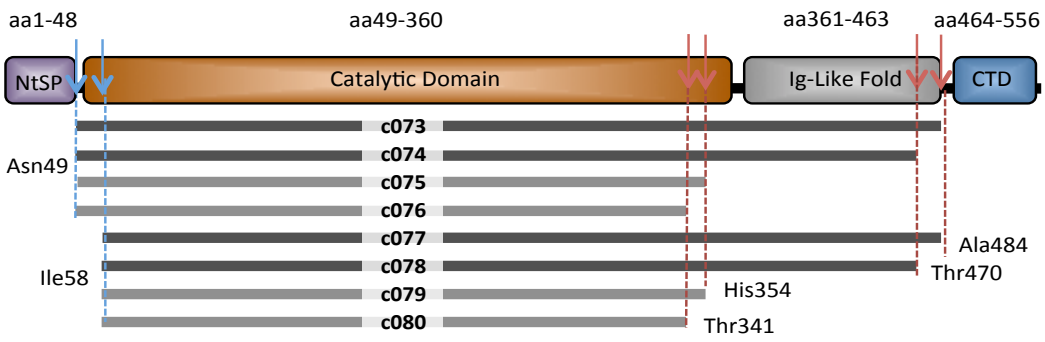


Figure 2.5 *Phase 2 construct design.* Following test purification of constructs from phase one, a number of additional C-terminal truncations were created to produce ordered protein domains lacking the CTD. Eight additional protein constructs were created.

Expression and purification of tPPAD^{WT} and tPPAD^{C351A}

To obtain protein of optimal purity, tPPAD was purified from *E.coli* cells in a multi-step process involving different chromatography procedures (Figure 2.6). Briefly, tPPAD was purified from the soluble cell extract of lysed bacterial cells using Co^{2+} affinity resin which was washed with binding buffer (BB), wash buffer (WB) and elution buffer (EB) which have increasing concentrations of imidazole. tPPAD was eluted from Co^{2+} resin in the moderate levels of imidazole (40mM) in WB and high imidazole (250mM) of EB. The low imidazole in WB is used to remove any non-specific endogenous proteins that may be bound due to naturally occurring histidine clusters, but may also elute some of the target protein. The WB and EB fractions containing most protein with acceptably low levels of contaminants were purified further by size exclusion chromatography (Figure 2.6B). Peaks identified on the chromatogram were analysed by SDS-PAGE (Figure 2.6C). Peak 2 corresponded to a protein of ~50kDa and contained the target protein, whereas peak 1 corresponded to a protein ~100kDa, which

would suggest a high molecular weight protein, so was discarded (Figure 2.6C). Fractions containing protein were treated with TEV protease to remove the His-tag, which was separated from the protein preparation by reverse affinity column. FT, BB, WB and elution fractions were analysed by SDS-PAGE, and pure untagged protein collected for crystallisation (figure 2.6D blue box). Protein of low concentration or containing remaining His-tagged molecules was discarded (figure 2.6D yellow box).

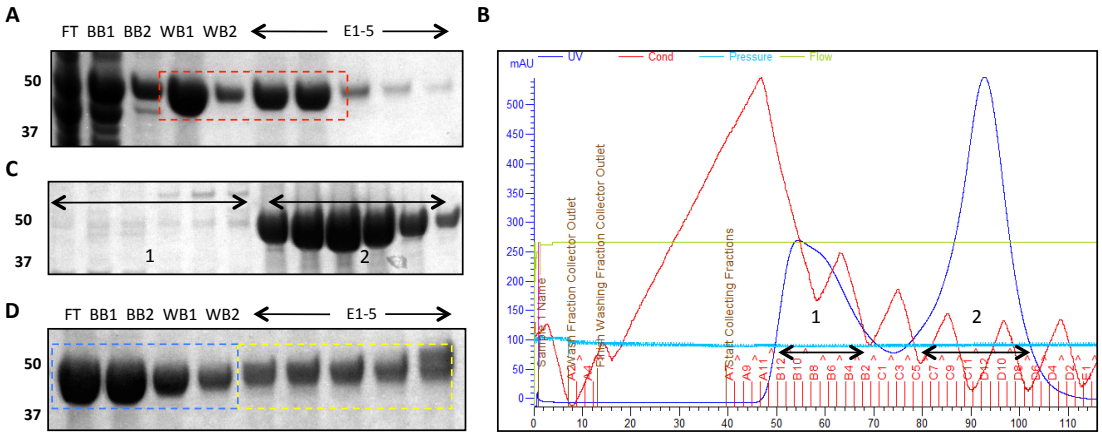


Figure 2.6 *Representative purification of tPPAD^{C351A}.* A) Metal affinity chromatography using Co^{2+} resin. Bacterial cells were lysed and the soluble fraction was applied to Co^{2+} resin loaded column and allowed to flow through. Flow through (FT), wash steps with binding buffer (BB1, BB2), and wash buffer (WB1, WB2), and elution fractions (E1-5) were collected and proteins resolved by SDS-PAGE. Coomassie stained SDS-PAGE gel identified concentrated tPPAD^{C351A} of acceptable purity in fractions WB1-E2, which were pooled and concentrated for size exclusion chromatography (indicated with red box). B) Size exclusion chromatography resolved two main peaks. SDS-PAGE of corresponding fractions identified protein in peak 2 only (C). D) Fractions from peak 2 were pooled and incubated with TEV o/n to cleave His-tag and remove by reverse affinity chromatography. Fractions with highest concentration of pure protein were pooled for crystallisation trials (blue box). Fractions of lower concentration were pooled separately for non-crystallisation based assays (yellow box).

Comparison of tPPAD^{WT} and flPPAD^{WT}

tPPAD comprises amino acid sequence aa49-484, which includes the CD and IgLF, whereas flPPAD comprises all four domains (Figure 2.7A). Compared to flPPAD, both tPPAD^{WT} and tPPAD^{C351A} were readily purified and yielded large amounts of soluble protein, with more mutant protein expressed (Figure 2.7B). tPPAD^{WT} and tPPAD^{C351A} expression yielded an average 2mg and 2.5mg per litre of culture compared to 0.03mg/L and 0.13mg/L for flPPAD wild type and mutant respectively (Figure 2.8). In addition, the His-tag was easily cleaved from both tPPAD constructs, providing protein of higher purity than flPPAD (Figure 2.7C).

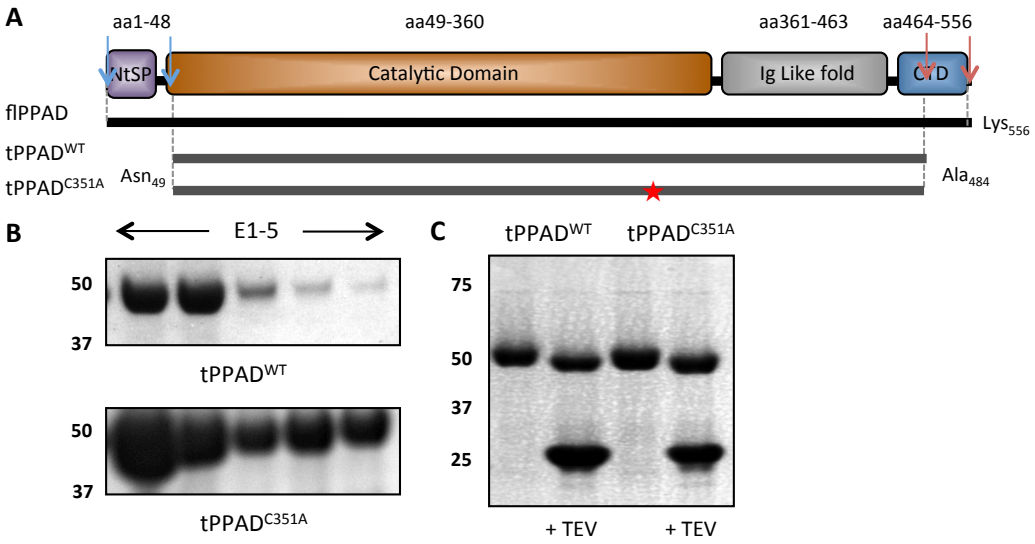
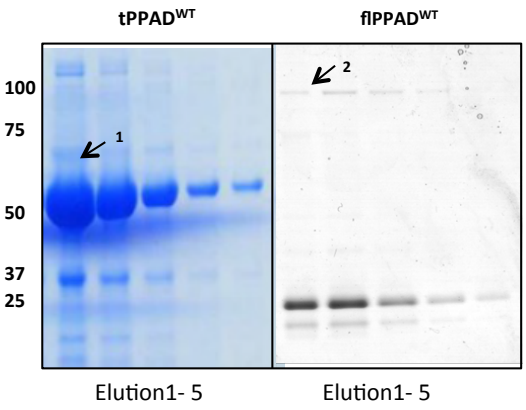


Figure 2.7 *Expression and His-tag cleavage of tPPAD^{WT} and tPPAD^{C351A}* A) tPPAD^{WT} construct from aa49-484 comprises catalytic domain and Ig-Like fold. A mutant form tPPAD^{C351A} has identical sequence with C351A mutation (indicated by red star). B) Representative Coomassie stained SDS-PAGE gel of tPPAD^{WT} and tPPAD^{C351A} expression in eluted fractions (elution 1-5). C) Cleavage of His-tag from tPPAD^{WT} and tPPAD^{C351A} with TEV. tPPAD was incubated o/n with TEV before protein resolution by SDS-PAGE and Coomassie gel staining. Successful cleavage was confirmed by mass shift.



	Average protein yield per L culture	
	Wild-Type	Mutant
flPPAD	0.03mg	0.13mg
tPPAD	~2mg	~2.5mg

Figure 2.8 *Comparison of tPPAD and flPPAD activity*. Image shows expression of tPPAD^{WT} and flPPAD^{WT}. PPAD proteins were expressed and purified from *E.coli*. Eluted fractions containing expressed proteins were resolved by SDS-PAGE and resulting gels stained with Coomassie. tPPAD^{WT} and flPPAD^{WT} in elutions are indicated by arrows labeled 1 and 2 respectively. Table shows average yield of tPPAD of flPPAD, calculated based on a number of purifications (n>5).

Stability of tPPAD^{WT}

Differential scanning fluorimetry (DSF) was used to measure stability of tPPAD^{WT} by determining the temperature of the denaturation midpoint (T_m°C) in a range of buffers: BB, WB, EB and GF buffers used in purification procedures, and buffers of increasing NaCl concentration (50mM - 500mM), and pH (pH4.5 - 10) (Figure 2.9A). tPPAD^{WT} had a similar T_m°C in all NaCl concentration and purification buffers tested. pH appeared to have the greatest influence on stability, with the highest T_m°C value obtained with pH7, although tPPAD^{WT} remained stable in a relatively broad range of pH ~pH6 - 8.5. Stability in GF buffer (used for size exclusion chromatography) allowed safe storage of protein at -80°C following purification, and concentration of protein to the required ≥10mg/ml for crystallisation without buffer exchange.

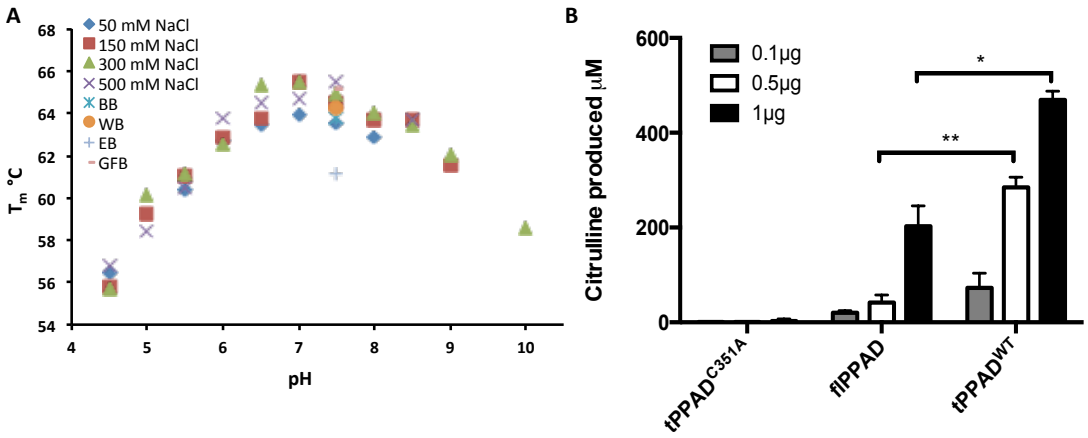


Figure 2.9 *Characterisation of tPPAD in vitro.* A) Differential scanning fluorimetry (DSF) was carried out on tPPAD^{WT} incubated at pH4 - 10, in increasing NaCl (50mM - 500mM) and in purification buffers (binding buffer (BB), wash buffer (WB), elution buffer (BB)) and size exclusion chromatography (gel filtration buffer (GFB)). Midpoint of denaturation was measured by T_m °C. B) Activity of tPPAD^{WT} and tPPAD^{C351A} was measured on synthetic arginine BAEE using a colourimetric assay for detection of citrulline. 0.1-1µg protein was incubated with 20mM BAEE at 37°C for 1 hour before measurement of citrulline. Values are n≥2±SD. *p≤0.05, **p≤0.01.

Activity of tPPAD^{WT} and tPPAD^{C351A}

Activity of tPPAD^{WT} and tPPAD^{C351A} was measured to confirm the enzyme was correctly folded and functional. tPPAD^{WT}, tPPAD^{C351A} and flPPAD were incubated with BAEE as before, and citrulline production was measured using the colourimetric assay for detection of citrulline. Using 0.5µg/1µg of PPAD enzyme, tPPAD^{WT} was significantly more active than flPPAD^{WT}, producing ~2x more citrulline (Figure 2.9B). tPPAD^{C351A} displayed no activity, confirming successful mutation of the catalytic residue.

Crystallisation of tPPAD^{WT} and tPPAD^{C351A}

tPPAD^{WT} and tPPAD^{C351A} were concentrated in centrifugal filter columns to 12mg/ml and 15mg/ml respectively. Crystallisation trials were prepared using a number of commercially available screens that contain various precipitant conditions conducive to crystallisation for a diverse set of proteins. Crystallisation plates were dispensed with precipitant condition and protein solution (tPPAD^{WT} or tPPAD^{C351A}) mixed in a specific volume ratio (2:1, 1:1, 1:2), and incubated at either 20°C or 4°C. Crystals of tPPAD^{WT} and tPPAD^{C351A} were obtained at 20°C by vapour diffusion method when mixing 150nL of protein solution with 300nL precipitant (30% PEG5000MME, 200 mM ammonium sulfate and 100 mM MES pH 6.5) and (22% PEG3350, 100 mM lithium sulfate and 100 mM bis—tris pH 5.5) respectively (Figure 2.10).

Crystals obtained from tPPAD^{WT} and tPPAD^{C351A} were mounted with cryoprotection, and diffraction data were collected at Diamond Light Source[1]. Using the molecular replacement method, 3D structures were determined for both tPPAD^{WT} and tPPAD^{C351A} to 1.46Å and 1.48Å resolution respectively (refinement statistics summarised in Table 2.2). Data collection, processing, model building and refinement were performed by Dr. J. Kopec at the SGC. Coordinates and structure factors were deposited at the Protein Data Bank with pdb accession codes 5AK7 (tPPAD^{WT}) and 5AK8 (tPPAD^{C351A}).

[1] Acknowledgements: *We thank Diamond Light Source for access to beamline I03 (proposal number mx8421) that contributed to the results presented here*

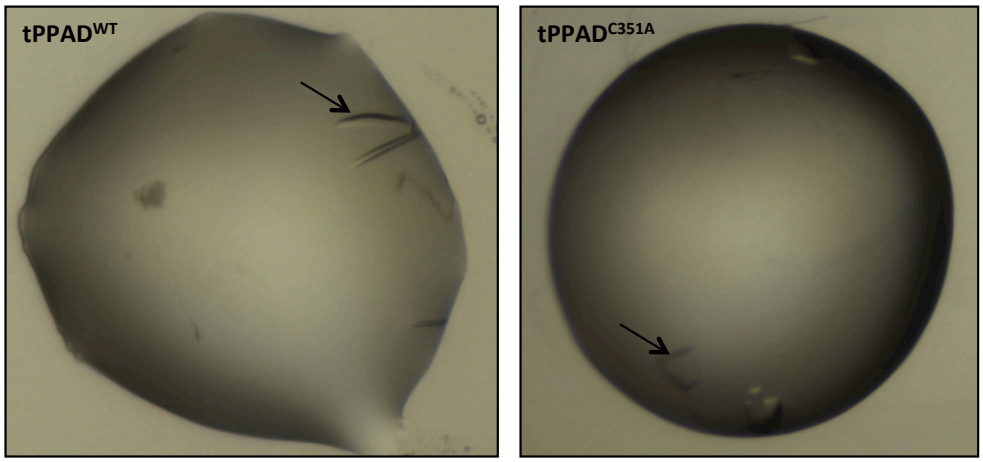


Figure 2.10 *Crystals obtained from tPPAD^{WT} and tPPAD^{C351A}.* Crystals used for data collection are indicated with arrows.

The tPPAD CD and IgLF form a tightly packed structure

The full tPPAD sequence comprises aa49-484, however visible regions revealed by electron density maps of both tPPAD^{WT} and tPPAD^{C351A} depict aa49-463, suggesting that the aa464-484 region of the protein is disordered.

The two-domain structure of PPAD has the dimensions 55Å x 57Å x 50Å (h x w x d) (Figure 2.11). The structure comprises the CD, which adopts a canonical fold of five ββαβ repeat motifs, and the IgLF. From the top face of the CD, the five repeats form five 'blades' of a disfigured propeller, with the entrance to the narrow binding groove in the centre. On the other face, the CD mediates extensive inter-domain contacts with the IgLF, a 4- and 5- antiparallel strand β-sandwich.

Table 2.2 *Refinement statistics for tPPAD^{WT} and tPPAD^{C351A} crystal structures.* Table prepared by Dr J. Kopec.

	tPPAD^{WT}	**tPPAD^{C351A}**
Dataset	PG1424PGA-d001	PG1424PGA-d002
Mutation	wild-type	C351A
Active site ligand	Arg peptide	Arg peptide
Data collection and processing		
Beamline	Diamond I03	Diamond I03
Wavelength	0.97625	0.97625
Unit cell parameters (Å)	101.860 84.660 55.660	101.79 84.09 55.460
(°)	90.0 92.5 90.0	90.0 92.64 90.0
Space group	C2	C2
Resolution range (Å)	65.08-1.46 (1.50-1.46)	64.8-1.48 (1.52-1.48)
Observed/Unique reflections	272195/79913 (18093/5724)	253844/73831 (18948/5330)
Rsym(%)	0.041 (0.497)	0.055 (0.535)
CC 1/2	0.998 (0.744)	0.997 (0.812)
I/sig(I)	14.3 (2.2)	10.4 (2.1)
Completeness	97.9 (95.0)	95.3 (93.9)
Multiplicity	3.4 (3.2)	3.4 (3.6)
Refinement		
Rcryst (%)	12.71	12.72
Rfree (%)	15.32	16.69
Wilson *B* factor (Å^2)	13.92	15.199
Average total *B* factor for protein(Å^2)	18.13	19.03
Average ligand *B* factor (Å^2)	27.9	24.68
Ligand occupancy	1	1
R.m.s.d. bond length (Å^2)	0.0061	0.0127
R.m.s.d. bond angle (°)	1.2697	1.5588
Molprobity analysis		
Clashscore	0.3 (100th percentile)	1.5 (99th percentile)
Ramachandran favoured (%)	97.64	97.64%
Ramachandran outliers (%)	0	0
Rotamer outliers (%)	0	0

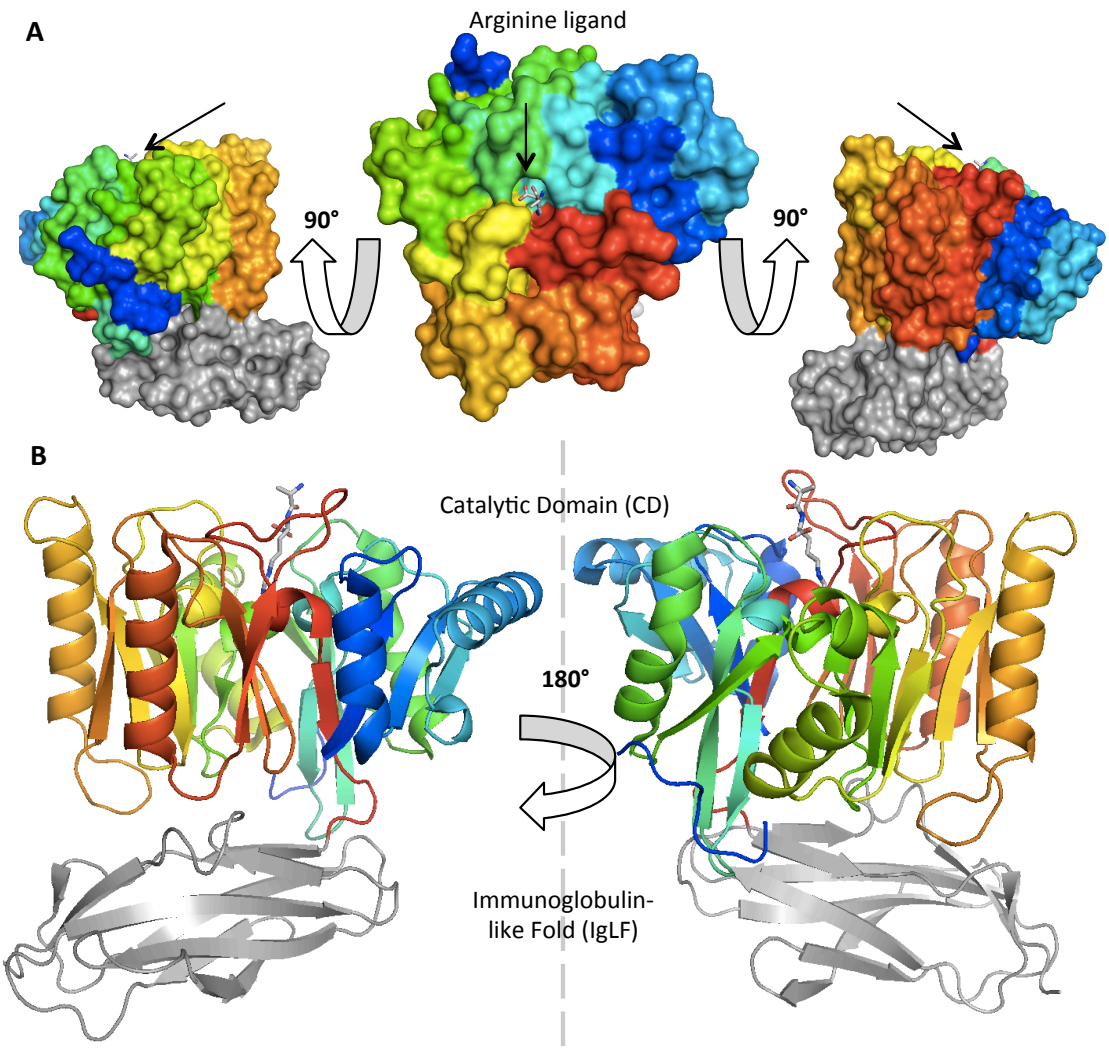


Figure 2.11 *Crystal structure of tPPAD^{WT} solved to 1.46Å resolution.* A) Surface representation of tPPAD^{WT}. Structure comprises the catalytic domain (CD), with five repeat motifs shown in blue, green, yellow, orange, and red, and the immunoglobulin-like fold (IgLF) in grey. Structure was modelled with an arginine containing peptide bound, which is visible at the top face of the CD, highlighted with arrows. B) Detailed ribbon structure of tPPAD^{WT} front face and rotated 180°. CD comprises five ααβα repeats to form canonical binding groove for the arginine ligand (modelled in grey sticks at the top face of the CD).

Homology of tPPAD^{WT} and tPPAD^{C351A} domains

The structures obtained for tPPAD^{WT} and tPPAD^{C351A} are superimposable, with a root mean square deviation (RMSD) of ~0.09Å (Figure 2.12A). This indicates the C351A mutation results in loss of function through mutation of an integral active site residue, rather than perturbation of protein integrity.

The CD adopts a 5-blade propeller structure (Figure 2.12B) that is well conserved between GMEs (Figure 2.12C), as are key residues of the active site (Figure 2.12D). Conversely, the IgLF lacks homology to any other GME and shares only limited sequence and structural homology with a variety of unrelated fibronectin type III proteins (e.g. fibronectin type III from integrin β4) as revealed by DALI structural neighbour search Z-score 7.4, RMSD 3.2Å, sequence identity 13% (Figure 2.12E, F).

The PPAD active site

The base of the CD contains the enzymatic active site, comprising conserved residues Asp130, His236, Asp238, Asn297 and Cys351 (Figure 2.13).

In both tPPAD^{WT} and tPPAD^{C351A} structures, difference Fourier electron density maps clearly reveal an active site bound ligand. Both tPPAD^{WT} and tPPAD^{C351A} were crystallised 'as-purified', without ligand supplementation, and evidently captured arginine ligands from culture media or during cell lysis of *E.coli*. In the tPPAD^{C351A} structure, this is modelled as an arginine containing dipeptide (Figure 2.13B), and in the tPPAD^{WT} structure as an amidino adduct of arginine, where its Cζ atom is covalently linked to the Cys351 sulphur atom of the enzyme (Figure 2.13A).

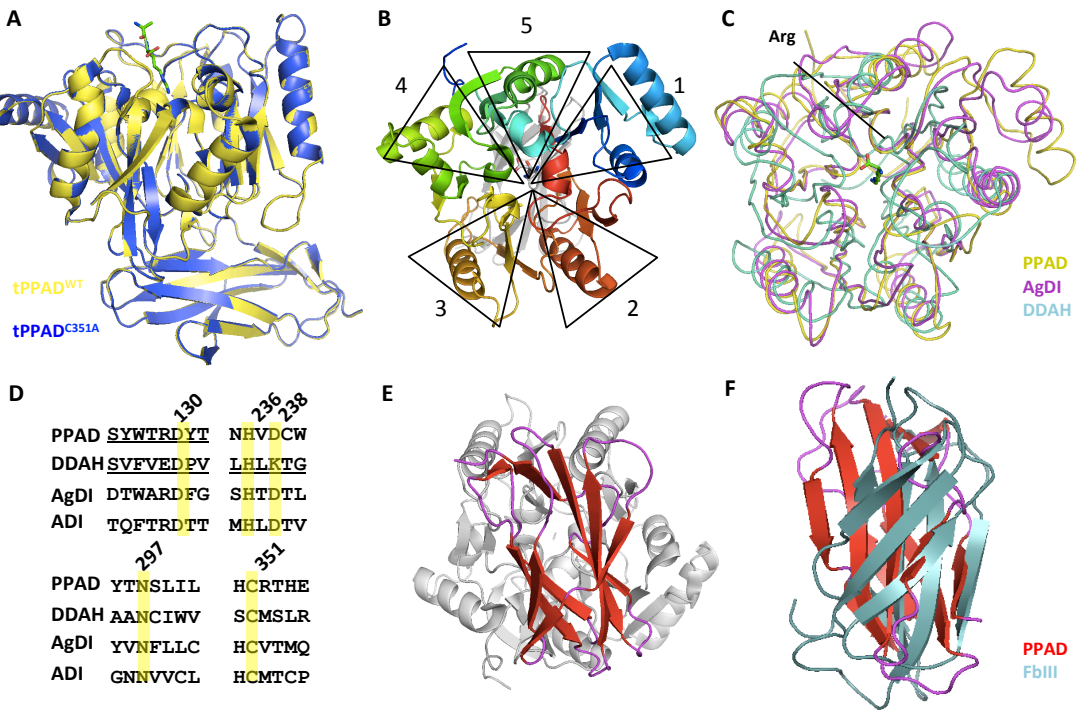


Figure 2.12 *Structural homology of tPPAD^{WT} and tPPAD^{C351A}.* A) Ribbon representations of tPPAD^{WT} (yellow) and tPPAD^{C351A} (blue) show the structures are superimposable. B) The tPPAD catalytic domain (CD) forms a five-blade propeller structure. C) Structure of the PPAD CD (yellow) is conserved amongst members of the guanidino modifying superfamily of enzymes (GMEs) including agmatine deiminase (AgDI) (purple) and N$^{\omega}$, N$^{\omega}$-dimethylarginine dimethylamino hydrolase (DDAH) (blue). Structures of AgDI and DDAH were modelled using pdb files 3HVM and 1RXX respectively. D) Active site residues of PPAD Asp130, His236, Asp238, Asn297 and Cys351 are conserved amongst GME family members DDAH, AgDI and arginine deiminase (ADI). E) The PPAD immunoglobulin like fold (IgLF) (red strands and purple loops) viewed from the bottom face of the PPAD molecule (CD shown in grey). F) The IgLF shares structural homology with fibronectin type III protein modelled using pdb file 4WTX.

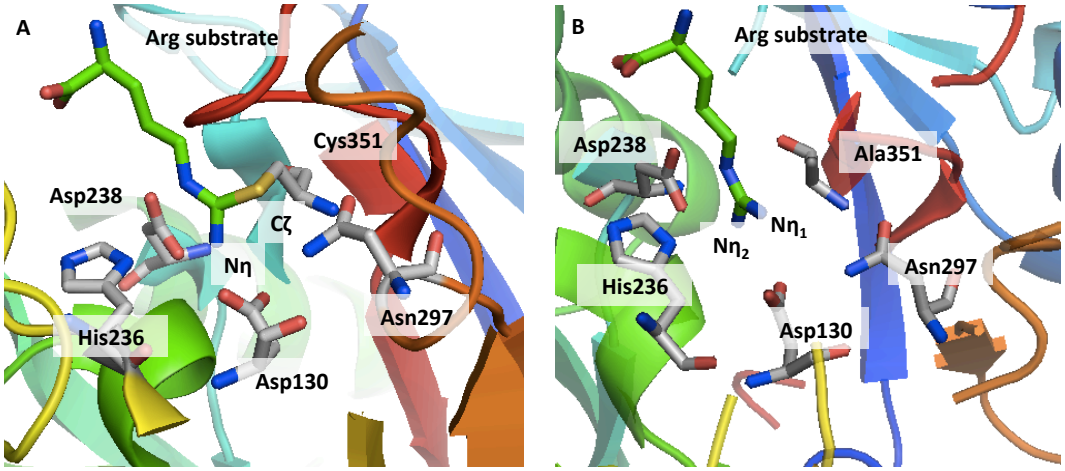


Figure 2.13 *The active site of tPPAD.* The active site is harbored at the base of the narrow binding groove within the CD. Arginine ligand enters at the top face of the CD and is stabilised for enzymatic turnover through ionic interactions with nitrogens (Nη1 and Nη2) of the guanidino group. The active site comprises key residues Asp130, His236, Asp238, Asn297, and Cys351 (A) or mutated to Ala (B). In the tPPAD^{WT} structure a covalent bond is formed between the guanidino carbon Cζ and sulphur group of the Cys residue.

Mapping of enzymatic activity using tPPAD^{WT} and tPPAD^{C351A}

At the entrance to the active site, ionic interactions between Arg152 and Arg154 with the arginine carboxyl group secure the peptide in the binding groove, with additional support from Tyr233 (Figure 2.14Ai). The peptide backbone uses its hydrocarbon moiety to pack against the cavity wall, positioning the arginine guanidino group within the active site, comprising Asp130, His236, Asp238, Asn297 and Cys351 (Figure 2.14Aii). Ionic interactions between the two terminal guanidino nitrogens (Nη1 and Nη2) and Asp238, Asp130 (Fig 2.14Ai) positions the substrate guanidino group between His236 and Cys351, the key residues in catalysis.

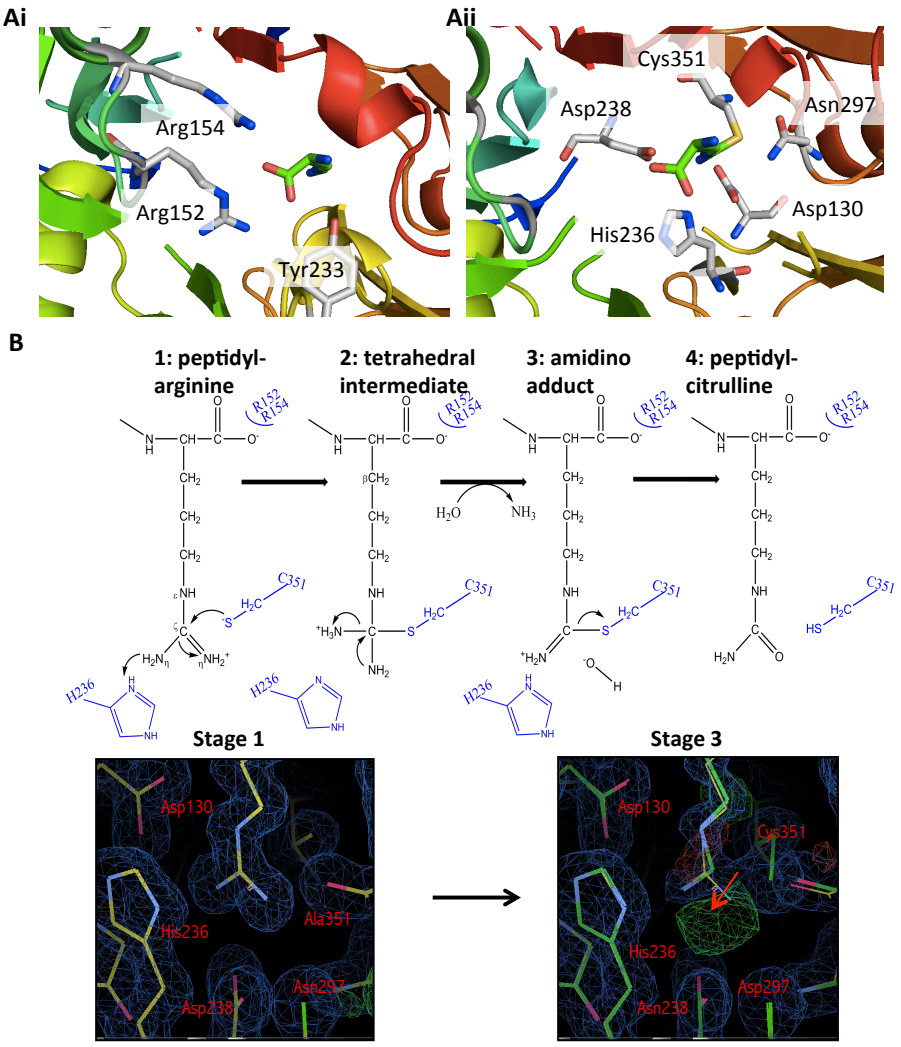


Figure 2.14 *Enzymatic conversion of arginine to citrulline by tPPAD modelled using 3D structures.* Ai) Interactions between tPPAD residues Arg152, Arg154 and Tyr233 and arginine substrates at the active site entrance Aii) Interactions between arginine substrate and key residues in the active site of tPPAD^{WT}. B) Citrullination of peptidyl-arginine to peptidyl-citrulline involves 4 reaction steps with the formation of a tetrahedral intermediate and production of an amidino adduct. Binding of an arginine ligand without enzymatic turnover modelled using the tPPAD^{C351A} structure (stage 1). Binding of an arginine ligand covalently linked between its guanidino carbon atom and Cys351 (electron density map in blue meshing) and the formation of an amidino adduct (electron density map in green meshing with red arrow) modelled using the tPPAD^{WT} structure (stage 3).

Citrullination involves four main steps, of which steps 1 and 3 can be modelled using the tPPAD^{C351A} and tPPAD^{WT} structures respectively (Figure 2.14B). When the arginine guanidino is secured at the active site (stage 1), Cys351 performs a nucleophilic attack on the Cζ atom by donating an electron, to form a tetrahedral reaction intermediate (stage 2). This electron is then shifted to the leaving group, H_2N. His236 acts as a general acid/base for proton transfer throughout the process, and first abstracts a proton from guanidino Nη1, which in turn takes a proton from the catalytic thiol group. The tetrahedral intermediate collapses to a covalent thiol intermediate with production of ammonia from the leaving group, which receives a proton from His236 and leaves the active site (Figure 2.14B stage 3). A free water molecule then enters the space left by ammonia. His236 again abstracts a proton from the water molecule, which in turn performs nucleophilic attack on the thiol intermediate, and then collapses to form a citrullinated product.

Substrate specificity of tPPAD^{WT}, PAD2 and PAD4

To compare PPAD activity to mammalian PADs, the activity of tPPAD^{WT}, PAD2 and PAD4 was measured *in vitro* using the colourimetric detection assay for citrulline. Human PAD2 and PAD4 were used as the two PAD isoforms commonly associated with RA. Recombinant human PAD2 and PAD4 were expressed and purified from *E.coli* expression systems, and the concentration of each enzyme was standardised. To investigate the substrate specificity of each enzyme, citrulline production was measured from a number of short peptides with the reactive arginine residue at the N- terminus (Arg-Gly-Glu), C-terminus (Gly-Arg), or internally positioned (Met-Arg-Phe) (Figure 2.15A).

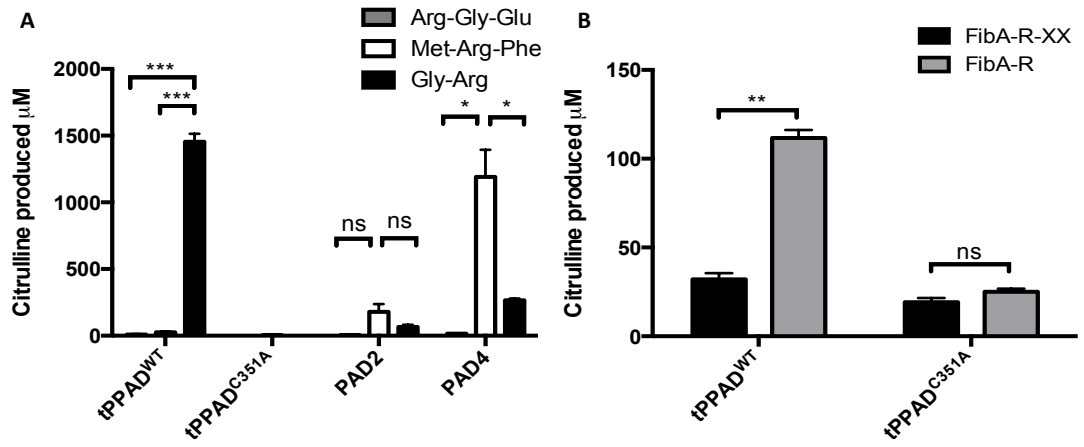


Figure 2.15 *Substrate specificity of tPPAD, PAD2 and PAD4.* A) Production of citrulline from arginine containing peptide substrates using a colourimetric detection assay for citrulline. 1μg tPPAD^{WT}, tPPAD^{C351A}, PAD2 or PAD4 was incubated with 5mM peptide for 1h at 37°C and citrulline produced measured. B) 1μg of tPPAD^{WT} or tPPAD^{C351A} was incubated with 5mM of fibrinogen peptide and citrulline produced measured. Values are mean n≥2±SD. *p≤0.05, **p≤0.01, ***p≤0.005, ns=not statistically significant.

tPPAD^{WT} only produced citrulline from the C-terminally located arginine peptide Gly-Arg, whereas PAD4 showed significantly increased activity on the internal arginine (Met-Arg-Phe) compared to the C-terminal (Gly-Arg) ($p<0.005$), and N-terminal arginines (Arg-Gly-Glu) ($p<0.05$). PAD2 displayed similar trends to PAD4 although this did not reach statistical significance, due to lower activity overall. As control, tPPAD^{C351A} did not display measureable activity on any of the substrates. To confirm the C-terminal specificity of tPPAD^{WT} on physiologically relevant substrates, two peptides from the fibrinogen α-chain were used in the assay. These include FibA-R (CESSSHHPGIAEFPS-R) with a C-terminal arginine, and FibA-R-XX, the same sequence with an additional two residues after the arginine (CESSSHHPGIAEFPS-R-GK). As before, tPPAD^{WT} produced significantly more citrulline from FibA-R than FibA-R-XX ($p<0.01$) (Figure 2.15B). The low reading obtained from tPPAD^{WT} on FibA-R-XX was comparable to tPPAD^{C351A} where no activity is expected, indicating a high background reading, or spontaneous hydrolysis of the substrate.

Structural basis of substrate specificity of tPPAD^{WT}

In the 3D structure, the terminal carboxylate group of the arginine substrate is fixed by ionic interactions with Arg152 and Arg154 (Figure 2.14Aii). This indicates a role for these basic residues in determining the reactivity of PPAD towards C-terminal arginine peptide substrates. I.e. if the arginine is at the C-terminus, the carboxyl groups are free to interact, whereas if the arginine is internal, there are no free carboxylates for interactions. To investigate the importance of these residues, both residues were mutated to either alanine

(tPPAD$^{R152/4A}$) or glutamine (tPPAD$^{R152/4Q}$), and the mutant proteins were expressed and purified as before. Using the same colourimetric detection assay from BAEE, tPPAD$^{R152/4A}$ had significantly reduced activity compared to both tPPAD^{WT} and tPPAD$^{R152/4Q}$ ($p<0.005$) (Figure 2.16A). The activity of tPPAD$^{R152/4Q}$ was not measurably affected.

To investigate these residues further, two additional mutants were synthesised replacing Arg152 and Arg154 with Ala individually, creating tPPAD^{R152A} and tPPAD^{R154A}. Using the same BAEE substrate and colourimetric assay, the activity of tPPAD^{R152A} was significantly reduced compared to tPPAD^{R154A} ($p<0.001$) and tPPAD^{WT} ($p<0.005$). There was no significant difference between tPPAD^{WT} and tPPAD^{R154A} (Figure 2.16B). However, when the arginine containing short peptides were used as substrates (Arg-Gly-Glu, Gly-Arg, Met-Arg-Phe), the activity of both tPPAD^{R152A} and tPPAD^{R154A} was significantly lower than tPPAD^{WT} ($p=0.001$ and $p=0.005$ respectively). As before, all enzymes displayed activity on C-terminal arginine peptide Gly-Arg only (figure 2.16C). The same activity trend and C-terminal specificity were observed using FibA-R and FibA-R-XX as substrates, with activity of tPPAD^{WT}, tPPAD^{R152A}, and tPPAD^{R154A} producing significantly more citrulline from FibA-R than FibA-R-XX ($p<0.01$) (Figure 2.16D). These results confirm that Arg152 and Arg154 are involved in enzymatic activity by stabilising peptide substrate. Mutation of Arg152 had a greater affect on enzymatic activity; suggesting interactions between arginine carboxyl group and Arg152 are more essential than those with Arg154. In addition, these results confirm PPAD C-terminal substrate specificity and show mutation of Arg152 or Arg154 does not enable binding of internal arginines.

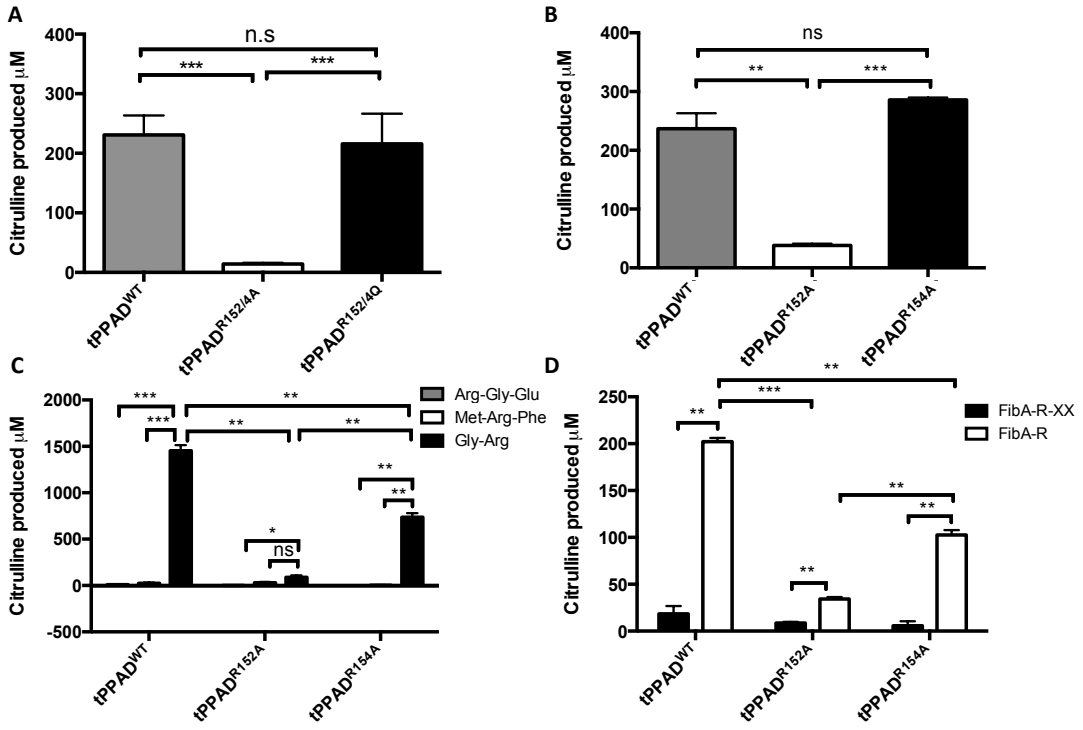


Figure 2.16 *The effect of R152 and R154 mutation(s) on tPPAD^{WT} activity and specificity.* A) Residues R152 and R154 were mutated to alanine (tPPAD$^{R152/4A}$) or glutamine (tPPAD$^{R152/4Q}$), expressed and purified as per tPPAD. 1μg tPPAD$^{R152/4A}$ or tPPAD$^{R152/4Q}$ was incubated with 20mM BAEE for 1 h at 37°C. Citrulline produced was measured by colourimetric detection assay. B/C/D) R152 and R154 were mutated individually to alanine (tPPAD^{R152A}, tPPAD^{R154A}) and expressed and purified as per tPPAD. 1μg tPPAD^{R152A} or tPPAD^{R154A} was incubated with 20mM BAEE (B) or 5mM peptide (C/D) for 1 hour at 37°C. Citrulline produced was measured as before. All graphs are mean n≥2±SD. *p≤0.05, **p≤0.01, ***p≤0.005.

Discussion

The recently characterised structures for both PAD2 (Slade *et al.*, 2015), and PAD4 (Arita *et al.*, 2004) have increased understanding of PAD mechanisms and led to development of novel inhibitors which have shown early success in animal models of RA (Kawalkowska *et al.*, 2016). To that end, the aim of this chapter was to solve the 3D crystal structure of PPAD, and elucidate the distinct enzymatic mechanisms that may make PPAD a suitable target for inhibition.

In order to solve the 3D crystal structure of PPAD, 20 constructs with varying N- and C-terminal boundaries were synthesised, expressed and purified, with construct c073 (tPPAD, aa49-484) identified as suitable for crystallisation. During purification it was observed that full-length protein and constructs comprising the CD only were predominantly insoluble. The full-length result was surprising, given the prior work in our lab using flPPAD constructs (Quirke *et al.*, 2014a), although other studies with PPAD have also found the full-length protein to be insoluble or unstable (Rodriguez *et al.*, 2009). These findings may be explained by the *in situ* role of the NtSP, IgLF and CTD domains flanking the CD. A number of enzymes expressed by *P. gingivalis,* including PPAD, belong to the type IX or Por secretion system (PorSS), which involves truncation of the NtSP and CTD for secretion from the bacterium (Shoji *et al.*, 2011, Zhou *et al.*, 2013). This processing may also explain the increased activity of tPPAD compared to flPPAD. These hypotheses will be investigated more thoroughly in Chapter 3.

Our structures reveals the active site comprises residues Asp130, His236, Asp238, Asn297 and Cys351, which confirmed conclusions from previous

mutation studies (Rodriguez *et al.*, 2010). The structural data showed both tPPAD^{WT} and tPPAD^{C351A} had serendipitously trapped arginine ligands during crystallisation of the 'as-purified' protein sample. Due to the high resolution of diffraction data obtained, the substrate arginine residue could be accurately modelled, however, the second residue of the tPPAD^{WT} bound dipeptide is unknown. By modelling the arginine in the tPPAD^{C351A} as a bound substrate before catalysis, and the arginine in the tPPAD^{WT} as a covalent adduct trapped mid process, the citrullination process was mapped with corresponding structural data to propose full enzymatic mechanisms.

In a separate study, Goulas and colleagues also solved the 3D crystal structure of a truncated form of PPAD (aa44-465), with identical architecture and topology to the findings of this study, and using mutation studies *in vitro,* proposed the same mechanisms of action (Goulas *et al.*, 2015). In addition, Goulas and colleagues identified a further aspect of enzymatic mechanisms involving one of the many loops at the top face of the CD. They proposed a specific surface loop (Val226-Val237) may act as a gate to the entrance of the active site, and appears to change enzyme confirmation to 'open' or 'closed' upon substrate binding. This hypothesis is supported by the findings of this study, in which residue Tyr233 (contained within this loop) stabilises the substrate at the active site entrance.

The α-β-propeller structure of the PPAD CD is conserved with other GMEs (figure 2.13C), and has limited homology the propeller structures of human PAD2 and PAD4, which have less defined rotational symmetry (figures 2.17, 2.18). However, PPAD, PAD2 and PAD4 CDs cannot be superimposed structurally, and there is a lack of sequence homology between the three enzymes (15-17%). Crucially, the binding grooves of the PAD2 and PAD4 CDs

must have specific features to enable binding and turnover of internal arginine residues, which is not possible in PPAD. In the tPPAD structure Arg152 and Arg154 at the entrance to the binding groove are involved in stabilising C-terminal arginine residues through ionic interactions with the free carboxyl group. This enables binding of C-terminal peptidyl-arginines for PPAD, which supports the *in vitro* observations made here and by others, that PPAD displays C-terminal substrate specificity (Wegner *et al.*, 2010, Bielecka *et al.*, 2014, Pyrc *et al.*, 2013). Similarly, in the PAD2 and PAD4 structures, arginine residues are also used to stabilise the main body of the peptide, but instead interact with the upstream peptide carbonyl groups.

Within the CDs, key residues of the PPAD active site are conserved with both GMEs and PAD2 and PAD4. The active site of GMEs comprise a conserved triad consisting of Cys, His and either Asp/Glu (Shirai *et al.*, 2006). The PPAD active site comprises five residues, Asp130, His236, Asp238, Asn297 and Cys351, and so shares Cys and His with the GME triad but replaces Asp/Glu with Asn297, which may provide a similar function. However, PPAD shares more similarity with the PAD2 and PAD4 active sites, which also contain a pair of Asp residues to stabilise the arginine guanidino group for turnover, a key nucleophilic Cys, and a His to act as a general acid or base for proton transfer. The active site residues of PAD2 (Asp351, His471, Asp473 and Cys647) and PAD4 (Asp350, His471, Asp473 and Cys645) are shown in Figures 2.17 and 2.18 respectively.

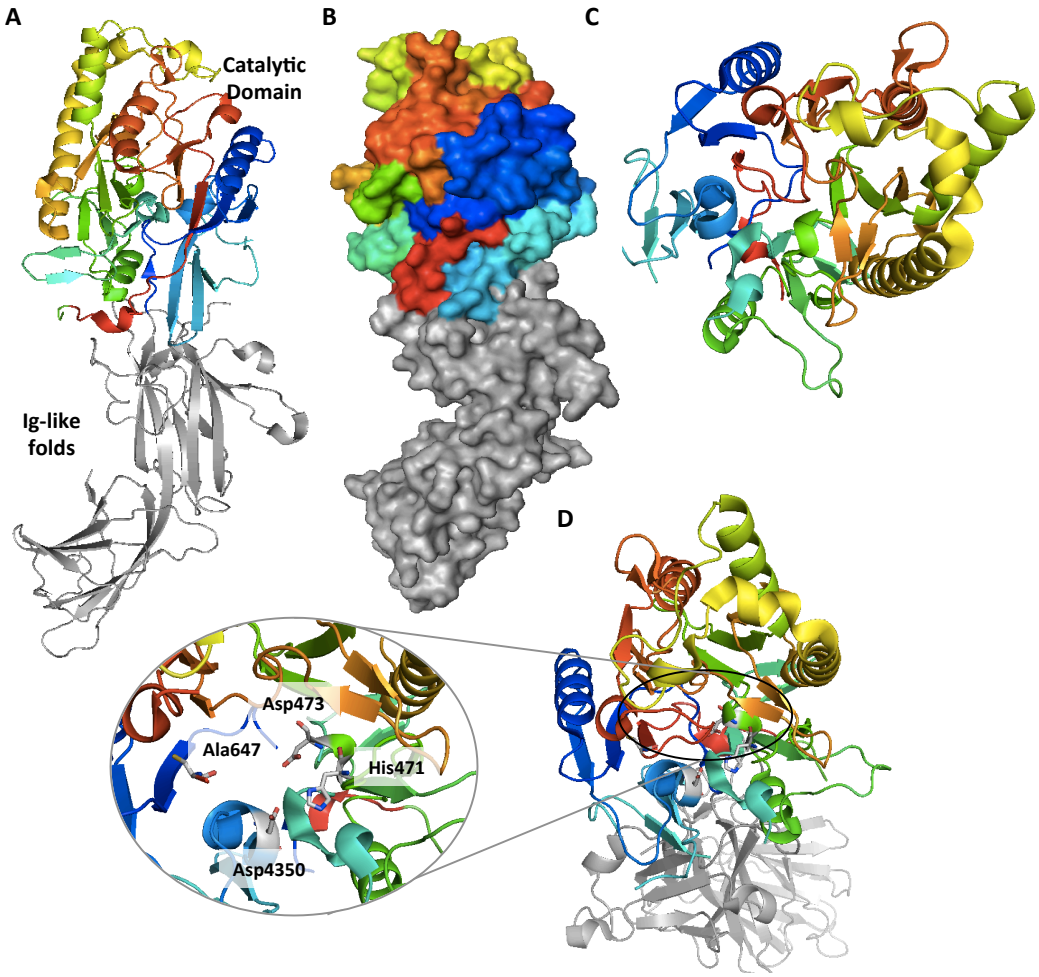


Figure 2.17 *Crystal structure of human PAD2.* Figures were made using protein database file (pdb) 4n20 of a C647A mutant deposited by (Slade *et al.*, 2015). A) Calcium unbound PAD2 in inactive conformation comprising CD (coloured ribbons) and two IgLF domains (grey ribbons), forward facing and rotated 180°. B) Overall surface architecture of PAD2. C) CD viewed from the top face. D) Catalytic domain with active site highlighted. Zoom in on active site shows four key sites for catalysis Asp351, His471, Asp473 and location of Cys647 mutated to Ala in this structure.

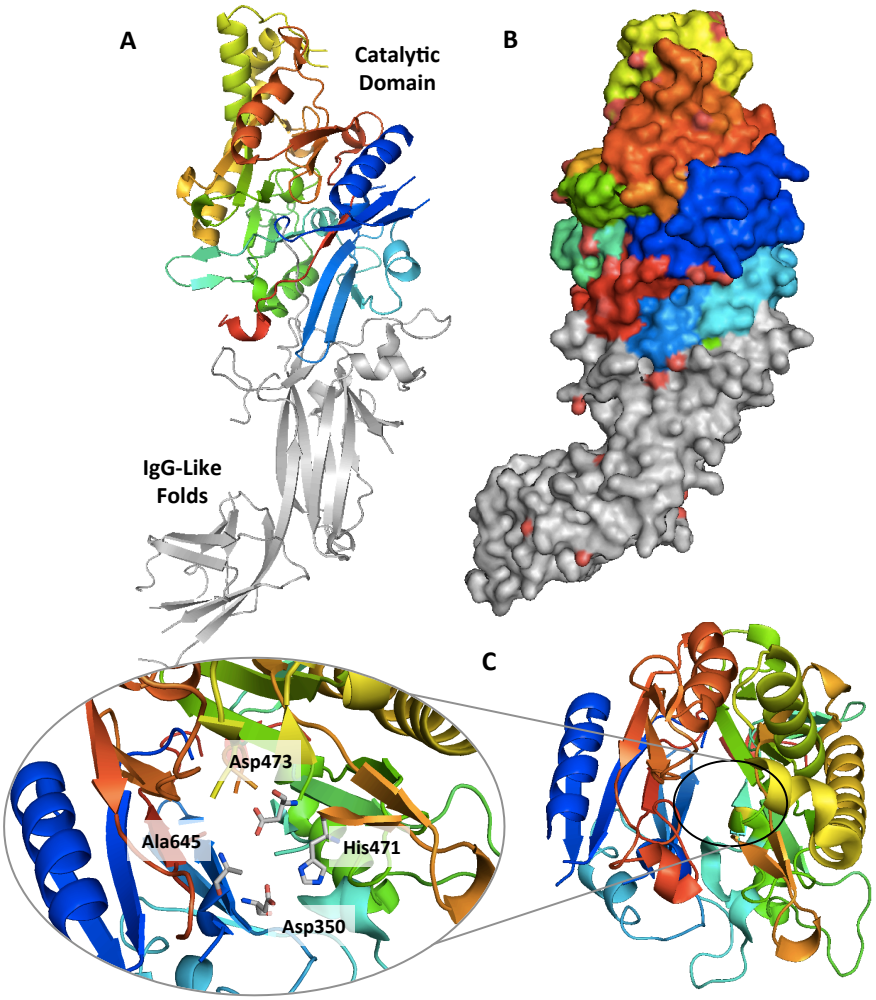


Figure 2.18 *Crystal structure of human PAD4.* These figures were drafted using protein database file (pdb) iwd9 of a C645A mutant deposited by (Arita *et al.*, 2004) A) Calcium bound PAD4 in active conformation comprising CD (coloured ribbons) and two IgLF domains (grey ribbons). B) Overall surface architecture of PAD4. C) CD of PAD4 viewed from the top face. Zoom in on active site shows four key sites for catalysis Asp350, His471, Asp473 and location of Cys645 mutated to Ala in this structure.

Away from the CD and active site, PPAD shares little homology to the GMEs, PAD2 or PAD4. Notably, the IgLF is completely absent from other GMEs, and the IgLF domains of PAD2 and PAD4 are unrelated to that of PPAD. In addition, the CD and IgLF domains of PAD2 and PAD4 are interspersed with 5 and 6 Ca^{2+} binding sites respectively, 2 in the CD and 3 or 4 in the adjoining IgLF, which require ligation to enable enzymatic competency. PAD4 also exists as a dimer which increases enzymatic activity (Liu *et al.*, 2011). By comparison, PPAD is not known to require metal binding or dimerisation for activity.

Although citrullination by PPAD is comparable to that of PAD2 and PAD4, specific structural features of PPAD may confer evolutionary benefits to *P. gingivalis*. In support of this are the comparable protein structures of PPAD and the *P. gingivalis* protease, arginine gingipain B (RgpB) (Eichinger *et al.*, 1999, Potempa and Nguyen, 2007). RgpB structure and function has been well characterised, and some notable features, such as the IgLF domain, which is required for membrane expression of RgpB may also be relevant to PPAD (Slakeski *et al.*, 2011). RgpB may also provide an evolutionary explanation for the C-terminal substrate specificity of PPAD, by providing ideal substrates for citrullination through protein cleavage.

The key findings of this chapter were:

- Solving the 3D crystal structure of $tPPAD^{WT}$ and $tPPAD^{C351A}$.
- Mapping enzymatic turnover with corresponding structural data, and confirming residues of the active site Asp130, His236, Asp238, Asn297 and Cys351.
- Identifying residues Arg152 and Arg154 that determine C-terminal substrate specificity.

Chapter 3 Characterisation of PPAD *in vitro*

Background

In Chapter 2, the 3D crystal structures of tPPAD^{WT} and the inactive mutant tPPAD^{C351A} were solved, providing insight into enzyme mechanisms and specificity. With that in place, the aim of this chapter was to characterise the function of tPPAD^{WT} *in vitro*, and investigate potential ways citrullination by PPAD could be involved in ACPA formation.

During the screening of PPAD constructs cloned for crystallisation, it became evident that truncation of N- and/or C- terminal boundaries affects solubility and activity of PPAD, and the truncated clone tPPAD^{WT} used for crystallisation was significantly more active than full-length flPPAD^{WT}. PPAD is secreted in a truncated form from *P. gingivalis* (McGraw *et al.*, 1999), and so may be activated during secretion. However the specific domain truncation mediating this increase in enzymatic activity is not identified.

In addition, the activity of PPAD requires Rgp to provide C-terminally located arginine substrates. Previous work from our lab has implicated the formation of C-terminally citrullinated peptides by PPAD and Rgp as a potential mechanism for breach of tolerance to citrullinated proteins in RA (Wegner *et al.*, 2010). In the structural study described in the previous chapter, the mechanisms for enzymatic activity were elucidated, and residues involved in determining C-terminal substrate specificity were identified. However this was not demonstrated using protein substrates *in vitro*.

In addition, previous work from our lab studying a recombinant full-length PPAD (flPPAD^{WT}) proposed PPAD is capable of self, or auto-citrullination and therefore

anti-*P. gingivalis* responses in the host could give rise to a citrulline specific response. This was supported by an increased antibody response to citrullinated flPPAD^{WT} compared to non-citrullinated flPPAD^{C351A} in RA patients (Quirke *et al.*, 2014a). However, a subsequent study comparing purified and recombinant forms of PPAD found no citrulline residues in native purified PPAD, and concluded that auto-citrullination is an artefact of recombinant expression and cell lysis (Konig *et al.*, 2014).

In addition to expression of PPAD, *P. gingivalis* may induce autoimmunity in RA through cross reactivity between PgEnolase and human α-enolase, which is a commonly occurring ACPA substrate in human RA (Lundberg *et al.*, 2008). Therefore, the main aims of this chapter are:

- Determine the domains required for PPAD activity.
- Investigate the functional relationship of Rgp and PPAD.
- Measure auto-citrullination of PPAD using a selection of recombinant and purified forms of the enzyme.
- Using fibrinogen and α-enolase as model ACPA substrates, measure citrullination by PPAD, PAD2 and PAD4 to compare citrullinated peptides formed.
- Compare citrullinated peptides formed by PPAD from α-enolase and PgEnolase as a mechanism for cross reactivity in RA.

Buffers

Western blotting

Loading buffer	1x NuPAGE® LDS sample buffer / 1x NuPAGE® tricine sample buffer (LifeTechnologies)
	100mM DTT
Transfer buffer	1x transfer buffer (LifeTechnologies)
	10% methanol

Anti-modified citrulline (AMC) assay reagents

Reagent A	25% H_2SO_4
	20% H_3PO_4
	0.025% $FeCl_3$
Reagent B	0.5% 2,3-butanedione monoxime
	0.25% antipyrine
	0.5M acetic acid

PPAD/PAD activity assays

As in Chapter 2

Colourimetric Detection Reagents (CDR)

As in Chapter 2

Materials and methods

Comparison of PPAD construct solubility and activity

PPAD constructs were expressed and purified as in Chapter 2. For initial comparative studies of all constructs, they were purified in parallel from 1L of starting cultures using one-stage purification of Co^{2+} resin only, as per 'small scale' purification. For detailed comparative and quantitative studies of two or more selected constructs, these were expressed and purified separately from 6L of starting culture as per the 'large-scale' purification method. For analysis of protein expression and solubility of constructs, 15µl of each elution fraction was resolved by SDS-PAGE as below and stained with InstantBlue protein stain Coomassie (Expedion) to identify proteins.

To compare activity of PPAD constructs, the colourimetric assay for citrullination activity was used as previously described in Chapter 2. For preliminary measurement of PPAD activity in elution fractions, estimation of protein concentration was made based on Coomassie stained SDS-PAGE gel. Elutions were divided into 'high', 'medium', and 'low' expression levels, and the volumes used in subsequent incubation varied accordingly (5µl, 10µl, 15µl respectively). For quantification of PPAD activity in specific constructs, protein concentration was calculated by Bradford assay using Pierce™ Coomassie Plus (Bradford) assay kit (ThermoFisher) and the NanoDrop A280 modules. Desired concentration of PPAD elutions were incubated with 20mM BAEE (Sigma) in PPAD buffer for 1 hour at 37 °C. Citrulline produced was calculated as before using absorbance reading of the L-citrulline standard curve in Microsoft Office Excel.

SDS-PAGE

Samples were prepared for electrophoresis with loading buffer (see 'Buffers'). Proteins were denatured by heating at 95°C for 5min before resolution on 12-14% Novex® NuPAGE® Bis-Tris gels (LifeTechnologies) with full-range molecular weight protein marker (LifeTechnologies) in 1x MOPS running buffer (LifeTechnologies) at 125V for 45min. For visualisation of small molecular weight peptides, samples were prepared with tricine loading buffer (see 'Buffers') before heating as before and resolution on Novex® 10-20% Tricine gels with low-range molecular weight protein marker in 1x Tricine SDS running buffer (all LifeTechnologies) at 125V for 90min.

Following electrophoresis, proteins were visualised by gel staining with InstantBlue protein stain Coomassie (Expedion) or Colloidal blue, or transferred to nitrocellulose membranes for immunoblotting.

Immunoblotting: western blotting and Anti-Modified Citrulline (AMC) assay

Nitrocellulose membranes and blotting paper (GE Healthcare) were equilibrated in transfer buffer (see 'Buffers'). SDS-PAGE gels were sandwiched with nitrocellulose membranes and blotting paper within transfer cassettes, and proteins transferred by electrophoresis at 72V for 3 hours at 4°C. For detection of citrullinated peptides the anti-modified citrulline (AMC) kit was used (Upstate, Milipore). Briefly, following protein transfer the membrane was washed with ddH_2O, treated with 0.1% ovalbumin for 15min rt and incubated o/n at 37°C with equal quantities of modification reagents A and B (see 'Buffers). Following modification, the membrane was washed with ddH_2O and blocked with 3% powdered milk in tris buffered saline (TBS) (milk-TBS) for 30min at rt, before a

1:660 dilution of primary antibody (rabbit anti-modified citrulline, Upstate/Milipore) was added in 3% milk-TBS and incubated rt for 6h. From here a standard western blot protocol is observed, using a 1:5000 dilution of goat anti-rabbit horseradish peroxidase conjugated (HRP) secondary antibody (Dako lot:00062104) in 3% milk-TBS. All membrane washes were with ddH_2O until the final wash in TBS with 0.05% tween before development with enhanced chemiluminescence (ECL) plus detection reagent (ThermoFisher).

For all other immunoblotting, a standard western blot protocol was observed using the following antibodies: rabbit anti-fibrinogen (Calbiochem, 1:5000), rabbit anti-enolase H300 (Santa-Cruz, 1:5000), rabbit anti-PPAD (In house, 1:1000), and goat anti-rabbit HRP secondary antibody (Dako lot:00062104, 1:5000). Antibodies were prepared in 3% milk in TBS with 0.1% tween (TTBS) (milk-TTBS). Membranes were blocked for 1h in 3% milk-TTBS at rt, or o/n at 4°C before incubation of primary antibodies for 1h rt in dilutions described above. Membranes were washed in TTBS for 1x 15min, 3x 5min before incubation with secondary antibody for 1h rt, and after. Proteins were visualised on X-ray film (Fuji) following membrane stimulation with 1ml ECL Plus reagent (ThermoFisher) per membrane.

Investigation of PPAD auto-citrullination

PPAD or *P. gingivalis* lysates were resolved by SDS-PAGE, and PPAD identified by Coomassie stained SDS-PAGE gel and western blotting with an anti-PPAD antibody (in house) as above. Protein samples used were: recombinant PPAD expressed and purified from *E.coli* (truncated $tPPAD^{WT}$, and full length $flPPAD^{WT}$), recombinant PPAD expressed and purified from *P. gingivalis*

(truncated and full length) and whole lysed *P. gingivalis.* For western blotting, 0.2µg of recombinant protein or 3µl of lysates was loaded per well, resolved by SDS-PAGE and transferred to nitrocellulose. For analysis of PPAD by liquid chromatography mass spectrometry (LC-MS), 2µg of recombinant PPAD or 10µl of lysed *P. gingivalis* was loaded per gel, resolved by SDS-PAGE and the gel was stained with Colloidal blue. Fixing and staining solutions were prepared as per the Colloidal blue staining kit instructions (IfeTechnologies). Briefly, SDS-PAGE gel was added to 10ml of fixing solution and incubated shaking for 10 min rt. Fixed gel was then added to 10ml staining solution and incubated shaking for 10 min at rt as before. Finally, 0.5ml Solution B was added to the staining solution and gel left to stain for 3-12h. Bands were excised and stored in ddH_2O for analysis.

Functional association of PPAD and RgpB

$tPPAD^{WT}$ or $tPPAD^{C351A}$ was added to fibrinogen at 1:10 enzyme:protein ratio and incubated briefly before the addition of arginine gingipain B (RgpB) also at 1:10 enzyme:protein ratio. Recombinant *P. gingivalis* RgpB was kindly donated by Prof J. Potempa, following purification as described (Potempa & Nguyen, 2007). Proteins and enzymes were incubated at 37°C and a sample removed after 0, 5, 30 and 120min. TLCK protease inhibitor (Sigma) was added to 10mM final conc. to stop the reaction. Samples were then heated at 95°C for 5min, tricine loading buffer (see 'Buffers') added, and heated again 95°C for 5min. Proteins were resolved by SDS-PAGE on Novex® 10-20% tricine gels, and visualised with InstantBlue protein stain Coomassie (Expedion).

Cleavage and citrullination of human fibrinogen and α-enolase

Recombinant human fibrinogen and α-enolase were purchased (Sigma Aldrich Cat no: F3879, MyBioSource Cat no: MBS203689) and prepared to a working concentration of 1mg/ml as per the manufactures instructions. Fibrinogen or α-enolase was incubated with PPAD, RgpB or both PPAD and RgpB using 1:10 enzyme:protein ratios in PPAD buffer (see 'Buffers') o/n at 37°C. To compare citrullination by human PAD2 and PAD4, fibrinogen was incubated with RgpB and PAD2 or PAD4 as for PPAD, using PAD buffer (see 'Buffers'). To measure citrullination of internal arginine residues, the same experimental design was used replacing RgpB with proteinase K. Protein cleavage and citrullination was determined by SDS-PAGE, immunoblotting and LC-MS.

Liquid Chromatography Mass Spectrometry (LC-MS)

Ms M-L. Thezenas and Dr. R. Fisher at the Target Discovery Institute (TDI), University of Oxford carried out preparation of samples and data analysis, analysis of results was carried out by myself. Citrullination in PPAD samples, or protein samples incubated with PPAD, PAD2 or PAD4 was determined by liquid chromatography tandem mass spectrometry (LC-MS). Peptides were suspended in 2% acetonitrile/0.1% trifluoroacetic acid before LC-MS analysis on a Q Exactive mass spectrometer coupled with a Dionex Ultimate 3000 UPLC (both Thermo Fisher Scientific). Peptides were desalted online (PepMAP C18, 300µm x 5mm, 5 µm particle, Thermo) for 1min at a flow rate of 20µl/min and separated on a nEASY column (PepMAP C18, 75µm x 500mm, 2µm particle, Thermo) over 60min using a gradient of 2%-35% acetonitrile in 5% DMSO/0.1% formic acid at 250nl/min. Survey scans were acquired at a resolution of 70,000 at 200m/z and

the 15 most abundant precursors were selected for higher-energy collision dissociation fragmentation with an automatic exclusion time of 27 seconds. Data was analysed in Peaks (Bioinformatics Solutions, V.6) using the PeaksDB search option at mass tolerances of 10ppm for MS1 and 0.05Da for MS2 at 1% FDR and without proteolytic restriction. Citrullination (Arg), deamidation (Asn, Gln) and Oxidation (Met) were used as variable modifications. Protein regions covered in the result spectra were analysed for citrulline and their rate of occurrence calculated as a percentage.

Identification of citrullinated peptides from PgEnolase

Recombinant PgEnolase was purified from *E.coli* as per flPPAD detailed in materials and methods of Chapter 2. Human α-enolase was purchased from MyBioSource (Cat no: MBS203689). Both forms of enolase were incubated with RgpB and PPAD at 1:10 enzyme:protein ratio of both o/n at 37°C in PPAD buffer (see 'Buffers'). Citrullinated peptides formed were sequenced by LC-MS as detailed above. Whole sequences of PgEnolase and α-enolase were compared using a ClustalW sequence alignment, and citrullinated peptides formed from the most highly homologous regions were identified. Six homologous regions were identified with a C-terminal arginine, and peptides synthesised in native conformation, or with a citrulline at the C-terminus (Pepceuticals). To ensure synthesised peptides were in native conformation, acidifications were added to all peptides at the C-terminus. Peptides were named CPgE1-6 for citrullinated peptides and RpgE1-6 for arginine controls. In peptides containing ≥2 arginine residues, synthetic peptides were designed with both residues in native or citrullinated conformation.

Modeling of PgEnolase

Homology models of PgEnolase were made by Dr W. W. Yue using the Swiss Model Server (http://swissmodel.expasy.org/) using either human α-enolase or streptococcal enolase as templates, with pdb codes 3ZLG and 2PSN respectively. Visual representations were made by me using Mac PyMol.

Results

The effect of PPAD N- and C-terminal truncation on solubility

To compare the expression level of soluble protein from all 20 constructs, (c001-c012, c073-c080) (Figures 2.3, 2.5), proteins were expressed and purified in parallel from 1L culture, and PPAD in the elutions resulting from one-step affinity chromatography was quantified by Coomassie staining of SDS-PAGE gel (Figure 3.1). Protein expression was confirmed by gel bands corresponding with the expected mass of the protein, and by MS. Expression level was rated from 0-5, where no expression was rated 0, the lowest observed protein rated 1, and highest observed protein rated 5.

Constructs c007 and c010 that contain the CD, IgLF and CTD, and constructs c073 and c074 that contain the CD and IgLF expressed the highest amount of soluble protein, rated 5. The lowest level of protein was expressed by c078, which comprise CD and IgLF with minor truncations into both domains, and c080, which comprises the CD only, also with minor truncations at both N- and C-terminals. Measurable protein was not identified for constructs c001, c011, or c075. Constructs c004, and c079 were not cloned, and construct c012 expressed protein of incorrect size, probably due to amplification of the wrong PCR product.

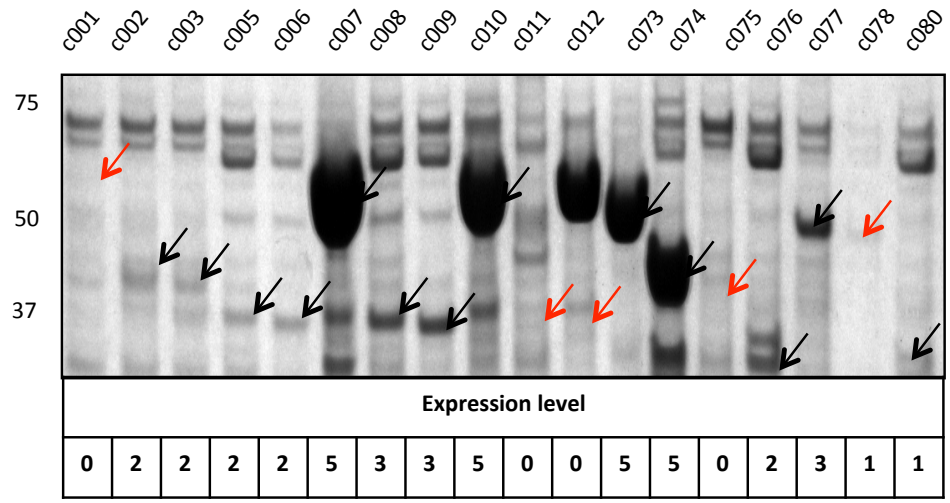


Figure 3.1 *Test expression of all PPAD constructs.* All PPAD constructs that had been cloned successfully were expressed and purified from 1L of *E.coli* culture. PPAD was purified from the soluble cell extract of lysed *E.coli* cells by Co^{2+} ion affinity chromatography and eluted from Co^{2+} resin in one elution buffer fraction. Equal volumes of elution fraction from each construct were resolved by SDS-PAGE gel and stained with Coomassie to identify proteins. Expression of soluble PPAD construct was judged by presence of protein bands at expected molecular weight and rated from 0-5. Black arrows are PPAD of correct molecular weight confirmed by MS. Red arrows represent where PPAD should be present.

The effect of PPAD N- and C-terminal truncation on citrullination activity

Using the above elution fractions, activity was compared using BAEE in the colourimetric citrulline detection assay, alongside previously described $flPPAD^{WT}$ to provide a standard. Of the 18 elution fractions, measurable activity was observed in elutions containing c007 and c073 only (Figure 3.2) which was confirmed by the presence of citrulline using the anti-modified citrulline (AMC) antibody in western blotting (Figure 3.3).

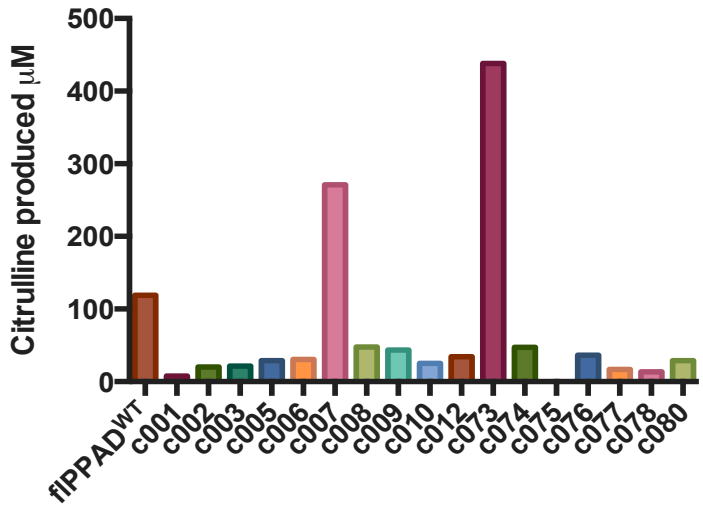


Figure 3.2 *Activity of PPAD constructs.* Elution fractions of PPAD constructs were incubated with 20mM BAEE in PPAD buffer for 1h at 37°C and citrulline production measured using colourimetric detection assay. Values are blank corrected mean n=2.

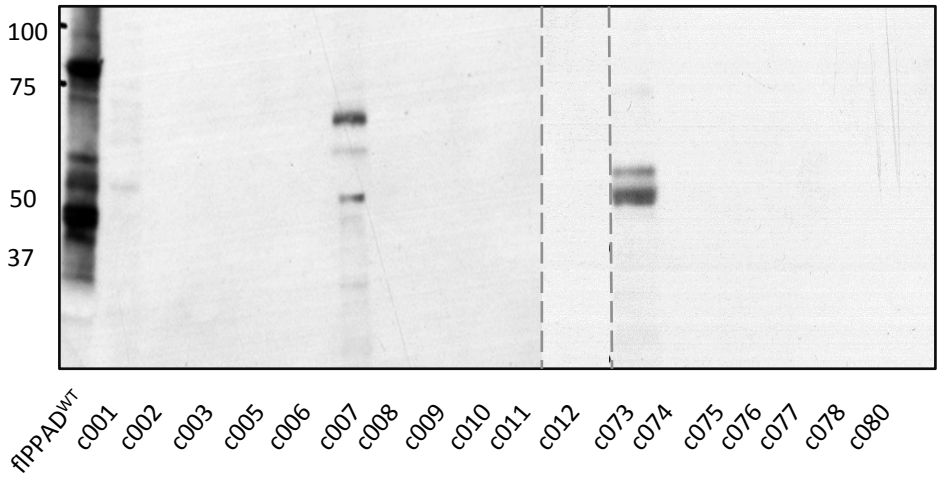


Figure 3.3 *Citrullination in PPAD elution fractions.* Elution fractions of PPAD constructs were resolved by SDS-PAGE and transferred to nitrocellulose. Citrullination in elution fractions was detected by western blotting with the anti-modified citrulline antibody (AMC). flPPAD^{WT} was used as a positive control. Dashed line represents different individual membranes.

Characterisation of active constructs c007 and c073

Protein concentration of c007 and c073 was standardised, and the BAEE assay repeated using 0.1-1µg of enzyme, incubated with substrate for 30 or 60min. Construct c073 was the most active, followed by c007 and flPPAD^{WT} (Figure 3.4). Mean production of citrulline following 60min incubation with 1µg enzyme was 584µM, 369.17µM, and 202.64µM respectively (Figure 3.4).

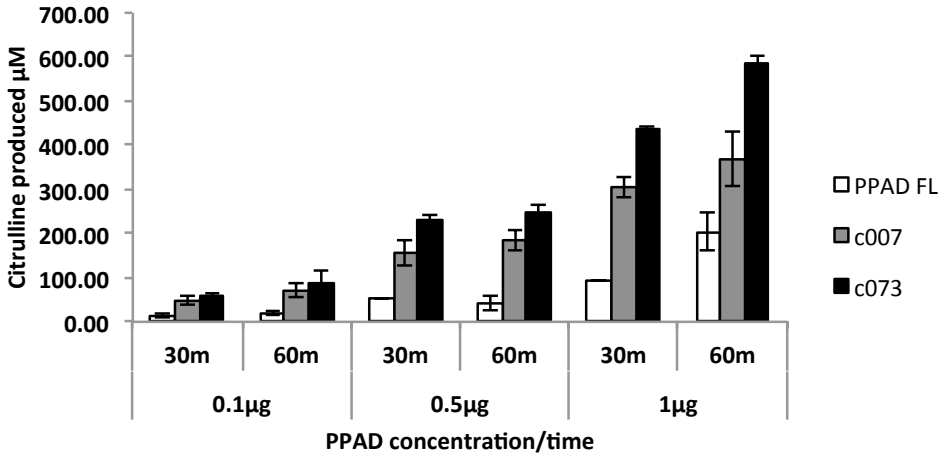


Figure 3.4 *Activity of full-length and truncated PPAD constructs.* Full-length (FL, aa1-556), N-terminally truncated (c007, aa49-556) or both N- and C- terminally truncated (c073, aa49-484) proteins were expressed and purified from *E.coli*. Protein concentration was standardised, and increasing concentrations of PPAD were incubated with 20mM BAEE in PPAD buffer for 30 or 60min at 37°C. Citrulline production was measured using the colourimetric detection assay. Values are blank corrected, mean n=2, ±SD.

Relationship between PPAD solubility and activity

The relationship between solubility of PPAD and activity was compared across all constructs (Table 3.1), which identified a number of constructs that expressed highly but did not display measurable activity. The most striking of these were c010 (Ile_{58}-Lys_{556}) and c077 (Ile_{58}-Ala_{484}), which both expressed soluble protein but did not demonstrate any measurable activity. Intriguingly, c010 and c077

comprise similar sequences to highly active c007 (Asn_{49}-Lys_{556}) and c073 (Asn_{49}-Ala_{484}) respectively, but have a further 9aa truncation at the N-terminus into the CD. Thus, truncating a small region between NtSP and the CD does not affect solubility, but has a significant impact on PPAD activity.

Table 3.1 *Summary of PPAD constructs solubility and activity.*

Construct Number	Sequence	Solubility	Activity
c001	Met_1-Lys_{556}	No expression	N/A
c002	Met_1-His_{366}	++	+
c003	Met_1-Gly_{360}	++	+
c004	Gln_{42}-Lys_{556}	Not cloned	N/A
c005	Gln_{42}-His_{366}	++	+
c006	Gln_{42}-Gly_{360}	++	+
c007	**Asn_{49}-Lys_{556}**	**+++++**	**+++**
c008	Asn_{49}-His_{366}	+++	+
c009	Asn_{49}-Gly_{360}	+++	+
c010	**Ile_{58}-Lys_{556}**	**+++++**	**+**
c011	Ile_{58}-His_{366}	No expression	N/A
c012	Ile_{58}-Gly_{360}	Cloning error	N/A
c073	**Asn_{49}-Ala_{484}**	**+++++**	**+++++**
c074	**Asn_{49}-Thr_{470}**	**+++++**	**+**
c075	Asn_{49}-His_{354}	No expression	N/A
c076	Asn_{49}-Thr_{341}	++	+
c077	Ile_{58}-Ala_{484}	+++	+
c078	Ile_{58}-Thr_{470}	+	+
c079	Ile_{58}-His_{354}	Not cloned	N/A
c080	Ile_{58}-Thr_{341}	+	+

All constructs were expressed and purified from *E.coli*. Elution fractions were analysed for protein expression by resolution by SDS-PAGE and Coomassie staining. Expression of the correct PPAD product was confirmed by molecular weight of the protein band and MS. Solubility of PPAD products was scored on a scale of + - +++++. Activity in elution fractions was measured using the colourimetric detection assay for citrulline as described, and scored on a scale of + - +++++. N/A= not applicable/not measured. Constructs relevant to the text are shown in bold.

The effect of N- and C- terminal truncations on activity and solubility

The 9aa region from Asn_{49}-Ile_{58} forms a loop in the PPAD structure at the base of the CD, distant from the active site (Figure 3.5). Therefore this region does not appear to directly impact catalysis or substrate binding, but may have a role in protein function, e.g. by maintaining overall protein architecture. Similarly, for constructs with identical N-terminals, truncating the C-terminus domains results in reduced expression e.g. c073 Asn_{49}-Ala_{484}, c074 Asn_{49}-Thr_{470}, and c076 Asn_{49}-Thr_{341} had expression level high, medium, and none respectively which corresponds to increased C-terminal truncation. There appears to be a trend for increased solubility when the CTD is removed, but decreased solubility when both IgLF and CTD are removed. Taken together, these data suggest a PPAD construct comprising CD and IgLF only is the most highly active and soluble. One such construct ($tPPAD^{WT}$) was used for crystallisation, and is used throughout this thesis as PPAD standard unless stated otherwise.

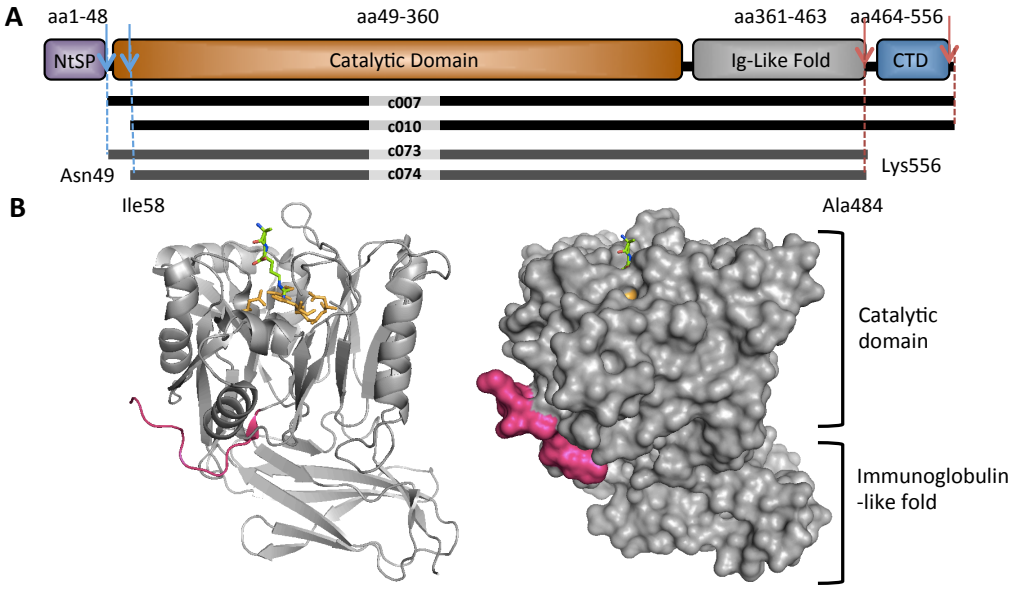
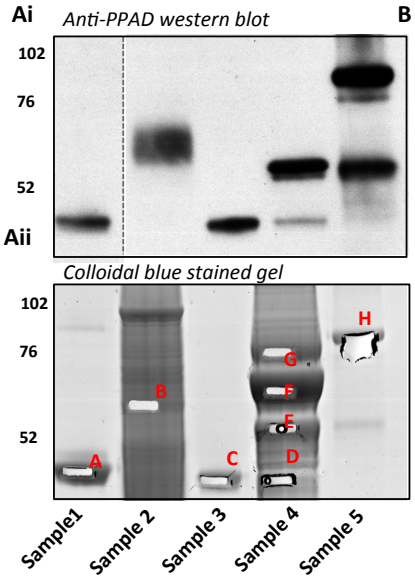


Figure 3.5 *N-terminal truncation of PPAD.* A) Protein c007, c010, c073 and c074, which have identical C-terminal domains but truncated N-terminals from aa49-58. B) Aa49-58 on the 3D structure of PPAD (pink), with an arginine ligand (green) bound in the active site (orange).

Auto-citrullination of PPAD

To investigate the ability for PPAD to auto-citrullinate, LC-MS was used to detect citrulline residues in a number of PPAD samples purified from different systems. These were: flPPAD^{WT} and tPPAD^{WT} expressed and purified from *E.coli,* full-length and truncated His_6 tagged PPAD expressed and purified from *P. gingivalis*, and lysed native *P. gingivalis* cells. PPAD exists in a number of isoforms *in vivo* (Rodriguez *et al.*, 2009, McGraw *et al.*, 1999), so to identify PPAD in respective samples, proteins were resolved by SDS-PAGE, transferred to nitrocellulose and probed with an anti-PPAD antibody (Figure 3.6 Ai). This confirmed the presence of PPAD in all samples, and revealed the molecular weights of the recombinant truncated (~50kDa) (Figure 3.6 Ai, lanes 1 and 3), and full-length forms (~62kDa) (Figure 3.6 Ai, lanes 2, 4 and 5) expressed from both *E.coli* and *P. gingivalis*. The full-length *E.coli* expressed PPAD (flPPAD^{WT}, lane 5) also migrates as an additional band at ~100kDa. PPAD in whole *P. gingivalis* lysate migrated as a diffuse band ~65kDa (Figure 3.6 Ai, lane 2), which is characteristic of native *P. gingivalis* proteins due to heavy glycosylation in the bacterium. Bands corresponding to the molecular weight of PPAD by western blotting were excised for analysis by LC-MS (Figure 3.6 Aii). LC-MS did not identify citrulline in any of the samples (Figure 3.8B). The C-terminal specificity of PPAD suggests auto-citrullination of internal PPAD arginine residues would be impossible. To investigate if PPAD could auto-citrullinate following cleavage of peptides to yield C-terminal arginines, tPPAD^{WT} was incubated with RgpB and peptides formed were analysed by LC-MS. In this instance, a number of citrullinated peptides were identified from both PPAD and RgpB (Table 3.2).



B

	Source of PPAD	Band(s) cut	MS result
Sample 1	tPPAD^{WT} (recombinant)	A	No citrullination
Sample 2	*P.gingivalis* lysate	B	No citrullination
Sample 3	Truncated PPAD purified from *P.gingivalis* (recombinant)	C	No citrullination
Sample 4	Full-length PPAD purified from *P.gingivalis* (recombinant)	D	No citrullination
		E	
		F	
		G	
Sample 5	flPPAD^{WT} (recombinant)	H	No citrullination

Figure 3.6 *Auto-citrullination of PPAD.* A number of recombinant and native forms of PPAD were resolved by SDS-PAGE and Ai) transferred to nitrocellulose membrane for western blotting with anti-PPAD or Aii) stained with colloidal blue to identify PPAD protein bands for LC-MS. Eight bands (A-H) were excised containing different PPAD isoforms from five samples. B) LC-MS results of PPAD bands excised from SDS-PAGE gel.

Table 3.2 *Citrullinated peptides formed from PPAD and RgpB.*

aa region	**tPPAD^{WT}**
130-152	DYTGWFAMYDTNKVGLVDFIYN-Cit
243-252	YLAPNKILI-Cit
277-286	AWGTKYEVY-Cit
287-304	ALATNEQPYTNSLILNN-Cit
421-445	TYSFTGLNKNDKVEYYISAADNSG-Cit
	RgpB
241-263	MIVIVPKKYEEDIEDFVDWKNQ-Cit
492-518	AQKDGKPTGTVAIIASTINQSWASPM-Cit
556-576	KDGEKMLDTWTVFGDPSLLV-Cit

Equal quantities of PPAD and RgpB were incubated together in PPAD buffer at 37°C o/n. Peptides formed were sequenced by LC-MS.

Functional association of PPAD and Rgp

PPAD requires Rgp activity to provide C-terminal arginine substrates. However, it is not clear if PPAD activity also has a beneficial effect on Rgp. To investigate this, RgpB was incubated with whole human fibrinogen with either $tPPAD^{WT}$ or inactive $tPPAD^{C351A}$ (Figure 3.7). Enzymes were treated with TLCK protease inhibitor to stop reaction instantly or after 5, 30, or 120min (T0, T5, T30, T120). Degradation of fibrinogen by RgpB was determined by comparing banding patterns on Coomassie stained gel following SDS-PAGE. At T0, differences in banding patterns between fibrinogen incubated with $tPPAD^{WT}$ and $tPPAD^{C351A}$ were observed, with a clear band identified at ~30kDa in the $tPPAD^{C351A}$ sample (Figure 3.7, T0, blue arrow). At T5, bands in both samples were similar (black and blue arrows), however by T30 and T120 more bands in the $tPPAD^{WT}$ samples compared to the $tPPAD^{C351A}$ samples suggested increased protein cleavage by RgpB in the presence of functional PPAD (Figure 3.7 T30, T120 black vs blue arrows). This may be through augmentation of protein structure or charge by citrullination, providing more cleavage sites for RgpB.

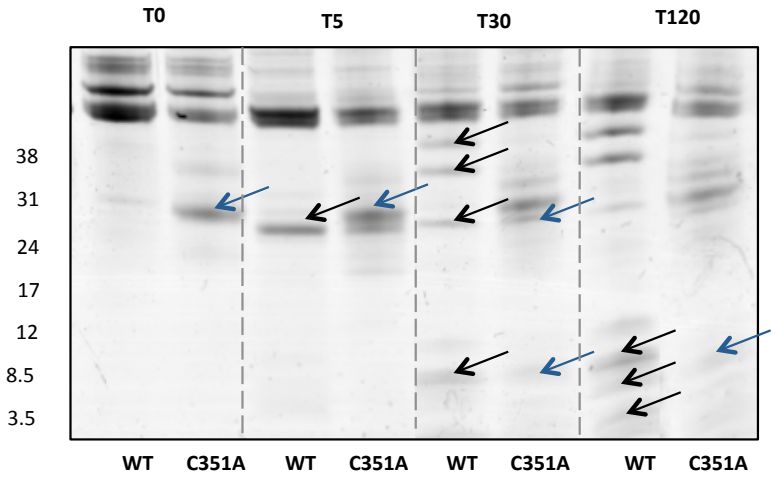


Figure 3.7 *The effect of PPAD on RgpB activity.* Fibrinogen was incubated with RgpB and tPPAD^{WT} (WT) or tPPAD^{C351A} (C351A) o/n at 37°C. RgpB was inhibited by adding TLCK to 10mM final concentration instantly (T0) or after 5 (T5), 30 (T30) or 120 (T120) min. Samples were resolved by SDS-PAGE and proteins identified by Coomassie staining. Peptides formed by WT (black arrows) and C351A (blue arrows) were compared.

Formation of peptides from fibrinogen and α-enolase by PPAD and RgpB

Our lab has previously shown formation of citrullinated peptides from fibrinogen and α-enolase following incubation with *P. gingivalis* (Wegner *et al.*, 2010). To advance upon these findings using recombinant proteins, fibrinogen and α-enolase were incubated with RgpB, with or without tPPAD^{WT} (Figure 3.8).

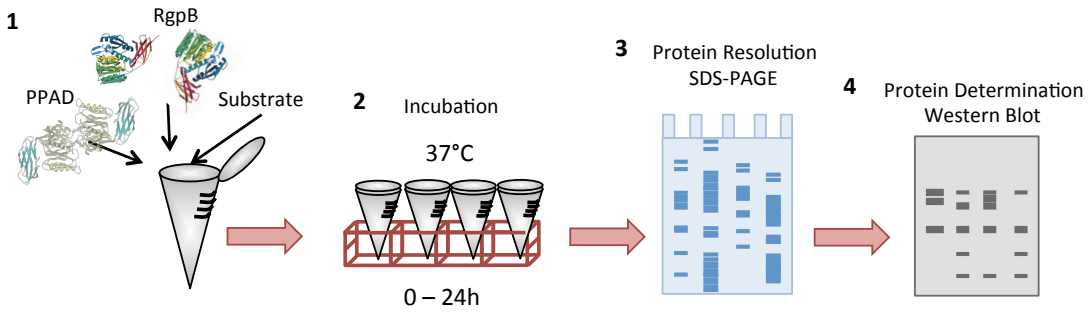


Figure 3.8 *Analysis of peptides formed from human fibrinogen and α-enolase.* Fibrinogen or α-enolase substrate was incubated with PPAD and RgpB (1) o/n at 37°C (2). Peptides formed were identified by SDS-PAGE and Coomassie staining (3), and western blotting (4).

When incubated with RgpB, the degradation of fibrinogen and α-enolase was confirmed by formation of lower molecular weight bands, identified by Coomassie staining of SDS-PAGE gels (Figure 3.9). Due to the large molecular weight of fibrinogen (360kDa), fibrinogen samples were resolved on both 12-14% bis-tris gels and 10-20% tricine gels to visualise peptides formed in a larger molecular weight range. As before, a difference in binding pattern was observed when fibrinogen or α-enolase substrates were incubated with both RgpB and tPPAD^{WT}, as compared to RgpB alone (Figure 3.9, A, B, C, lanes i, vs ii).

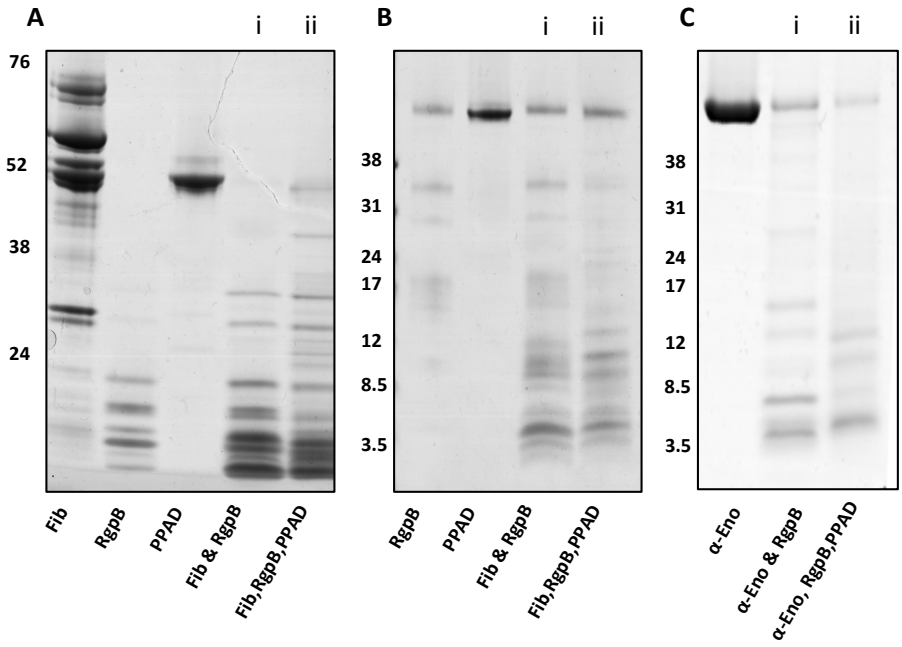


Figure 3.9 *Degradation of human fibrinogen and α-enolase by RgpB.* Fibrinogen (Fib) and α-enolase (α-Eno) were incubated with RgpB or both RgpB and PPAD at 37°C o/n. Fibrinogen, RgpB and PPAD were incubated individually as controls. Fibrinogen samples were resolved by SDS-PAGE on both 12-14% bis-tris (A) and 10-20% tricine gels (B) and α-enolase (α-Eno) was resolved by SDS-PAGE on 10-20% tricine gels (C). Gels were stained with Coomassie.

Fibrinogen incubations were repeated and transferred to nitrocellulose for western blotting. Blotting with anti-fibrinogen antibodies identified formation of lower molecular weight fibrinogen fragments following incubation with RgpB, which are still detected with the anti-fibrinogen antibody (Figure 3.10, A, lane 4 arrow). Interestingly, no fibrinogen fragments were identified following incubation with RgpB and $tPPAD^{WT}$ (Figure 3.10, A, lane 5). This may be due to increased efficiency of RgpB when incubated with $tPPAD^{WT}$, rendering peptides too small for detection, or by citrullination inhibiting antibody-binding sites.

When probed with an anti-PPAD antibody, larger molecular weight bands in the $tPPAD^{WT}$ control were not seen in the fibrinogen, RgpB and $tPPAD^{WT}$ sample (Figure 3.10 B, arrows lane 3 and 5), indicating $tPPAD^{WT}$ was cleaved by RgpB.

Finally, the AMC antibody was used to determine citrullination in the samples. Citrullination was found in the $tPPAD^{WT}$ control sample (Figure 3.10 C, arrow on lane 3), and in the sample containing fibrinogen, RgpB and $tPPAD^{WT}$ (Figure 3.10, C, lane 5 arrow). Bands identified in the fibrinogen, RgpB and $tPPAD^{WT}$ sample appeared to correspond to the bands identified with anti-fibrinogen blotting of fibrinogen and RgpB (Figure 3.10 A, lane 4, and C lane 5).

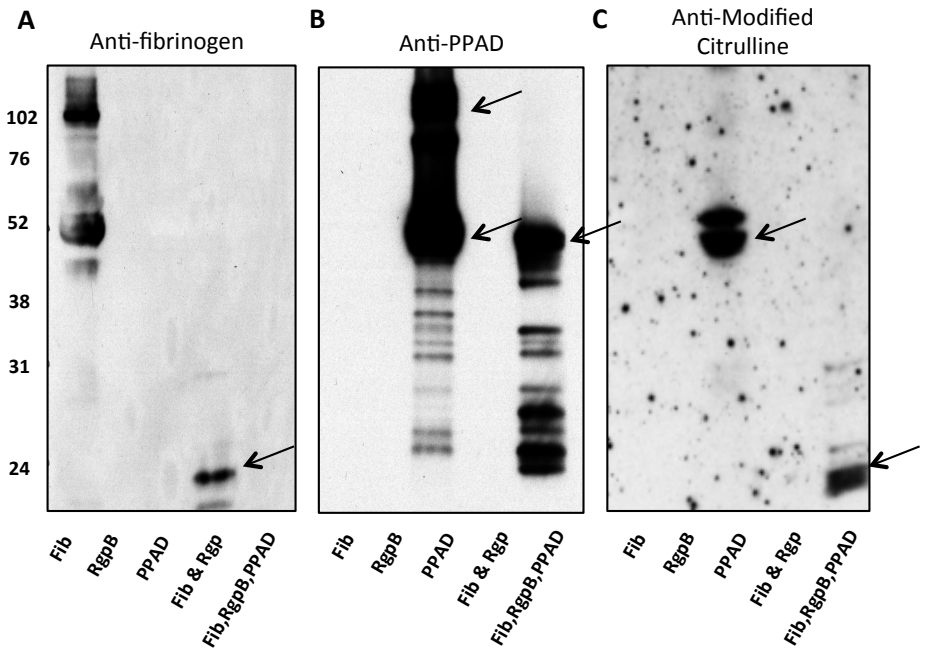


Figure 3.10 *Identification of peptides formed from human fibrinogen by RgpB and PPAD.* Fibrinogen (Fib) was incubated with RgpB or with RgpB and PPAD at 37°C. Fibrinogen, RgpB and PPAD were incubated individually in the same way as controls. All samples were resolved by SDS-PAGE on 12-14% bis-tris gels, and transferred to nitrocellulose membranes for western blotting. Membranes were probed with anti-fibrinogen (A) or anti-PPAD (B) antibody, or modified and probed with anti-modified citrulline antibody (C).

Sequencing of peptides formed from fibrinogen and α-enolase by LC-MS

To confirm citrullination and sequence the peptides formed from fibrinogen and α-enolase, the same incubations were prepared as before and citrullination was measured by LC-MS. Good protein coverage was obtained for all samples, encompassing 25 of 38 arginine residues in the fibrinogen α-chain, 15 of 26 in the β-chain, 8 of 12 in the γ-chain and 11 of 16 from α-enolase. The level of citrullination for each arginine residue in the LC-MS covered region was quantified (Figure 3.11). In total, 41 C-terminally citrullinated peptides were formed from fibrinogen (19 α-chain, 14 β-chain, 8 γ-chain) and 11 from α-

enolase. Of the 19 citrullinated peptides from the fibrinogen α-chain, 16 were citrullinated in the majority of instances (≥50%), whilst 3 towards the C-terminus of the protein were citrullinated less frequently (14-25%), which could be due to intrinsic secondary structure that restricts availability of the residues to PPAD. Of the 11 citrullinated peptides formed from α-enolase, 10 were citrullinated in the majority of instances (≥50%). All peptides were citrullinated at the C-terminus.

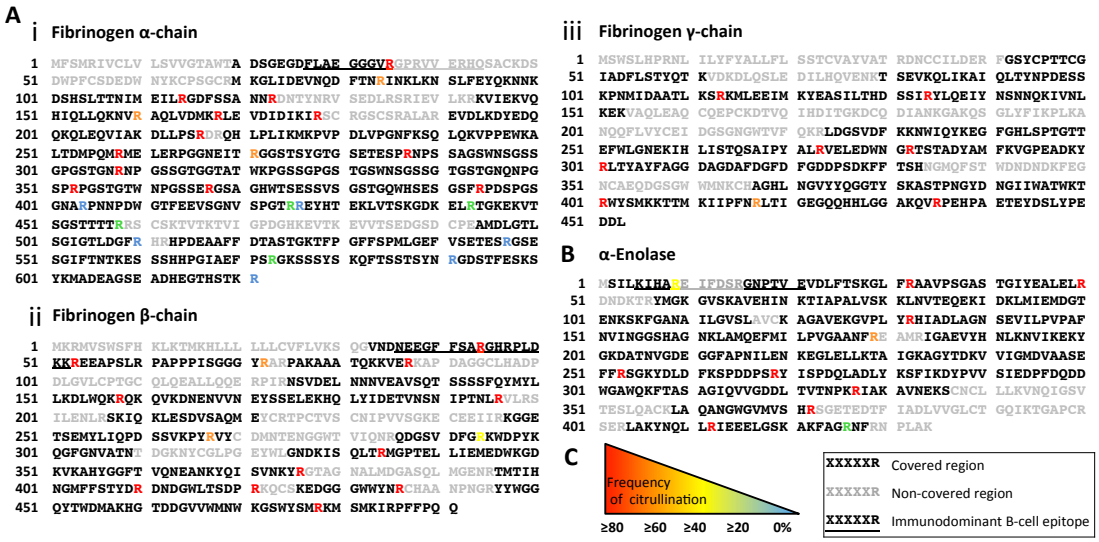


Figure 3.11 *Citrullinated peptides formed from fibrinogen and α-enolase by RgpB and PPAD.* Fibrinogen and α-enolase were incubated o/n at 37°C and resulting peptides were identified by LC-MS. A) Peptides from fibrinogen α-, β- and γ-chains (panels i, ii, and iii respectively). B) Peptides from α-enolase. C) Key: Sequence in black denotes region covered in LC-MS, grey depicts non-covered regions. Underlined sequences denote known immunodominant B-cell epitopes. Arginine residues occurring frequently citrullinated are represented by a red/orange R, occasionally citrullinated by a yellow/green R, and those never detected citrullinated by a blue R.

Identification of homologous citrullinated regions of PgEnolase and α-enolase

Cross reactivity between PgEnolase and α-enolase is a possible mechanism for development of autoimmunity in RA. To identify potentially cross-reactive peptides, PgEnolase was incubated with PPAD and RgpB as for fibrinogen and α-enolase, and peptides formed sequenced by LC-MS. C-terminally citrullinated peptides formed from PgEnolase were then compared to peptides from α-enolase using ClustalW sequence blast. Six peptides were identified from the most highly homologous regions, and were synthesised with either C-terminal arginine or citrulline residues. These were named RPgE1-6 and CPgE1-6 respectively (Table 3.3).

Table 3.3 *C-terminally citrullinated peptides from PgEnolase.*

1	Peptide formed by PPAD & RgpB	PgEnolase	aa1-9	MEIAKIIGR
		α-enolase	aa1-9	MAILKIHAR
	Synthetic peptides		CPgE1	MEIAKIIG-Cit-acid
			RPgE1	MEIAKIIGR-acid
2	Peptide formed by PPAD & RgpB	PgEnolase	aa10-15	EILDSR
		α-enolase	aa10-15	EIFDSR
	Synthetic peptides		CPgE2	IIG-Cit-EILDS-Cit-acid
			RPgE2	IIGREILDSR-acid
3	Peptide formed by PPAD & RgpB	PgEnolase	aa34-51	AAVPSGASTGENEAIELR
		α-enolase	aa33-50	AAVPSGASTGIYEALELR
	Synthetic peptides		CPgE3	TGENEAIEL-Cit-acid
			RPgE3	TGENEAIELR-acid
4	Peptide formed by PPAD & RgpB	PgEnolase	aa168-175	PVGACCFR
		α-enolase	aa 172-179	PVGAANFR
	Synthetic peptides		CPgE4	PVGACCF-Cit-acid
			RPgE4	PVGACCFR-acid
5	Peptide formed by PPAD & RgpB	PgEnolase	aa311-326	QLVGDDLFVTNVEFLR
		α-enolase	aa314-327	QVVGDDLTVTNPKR
	Synthetic peptides		CPgE5	LFVTNVEFL-Cit-acid
			RPgE5	LFVTNVEFLR-acid
6	Peptide formed by PPAD & RgpB	PgEnolase	aa401-409	MAKYNQLLR
		α-enolase	aa404-412	LAKYNQLLR
	Synthetic peptides		CPgE6	MAKYNQLL-Cit-acid
			RPgE6	MAKYNQLLR-acid

PgEnolase and human α-enolase were incubated with PPAD and RgpB o/n at 37°C. C-terminally citrullinated peptides formed were sequenced by LC-MS and compared. Peptides from six PgEnolase regions highly homologous to human α-enolase were synthesised in citrullinated and native arginine containing forms with C-terminal acidifications. Peptides were named citrullinated PgEnolase 1-6 (CPgE1-6) and arginine PgEnolase 1-6 (RPgE1-6).

Structural modelling of PgEnolase

Next, the RPgE1-6 peptides were mapped onto the 3D structure of PgEnolase. Because the 3D crystal structure of PgEnolase has not been solved, two homology models were created using either human α-enolase or streptococcal enolase as structural templates (Figure 3.12). When created using human α-enolase, the PgEnolase structure is modelled as a dimer, with the N-terminus of one molecule forming an interface with the C-terminal of the second molecule (Figure 3.12A). When the RPgE1-6 peptides are highlighted in this structure the RPgE1, RPgE2 and RPgE6 regions become partially or completely embedded in the interface of the dimer. Of note is the RPgE2/CPgE2 peptide ^{6}IIG-R/Cit-EILDS-R/Cit^{15}, which comprises a similar region covering the same 2 citrulline residues, and shares high sequence homology with the α-enolase immunodominant citrullinated peptide CEP-1 ^{5}KIIG-Cit-EILDS-Cit-GNPTVEC^{21}. CEP-1 is also embedded in the centre of the human α-enolase dimer and so may act as a cryptic epitope. When homology modeling was performed based on streptococcal enolase, PgEnolase exists as a monomer, forming a rough canonical configuration of perpendicular pairs of α-helices (Figure 3.12B). In this structure all RPgE1-6 peptides are accessible.

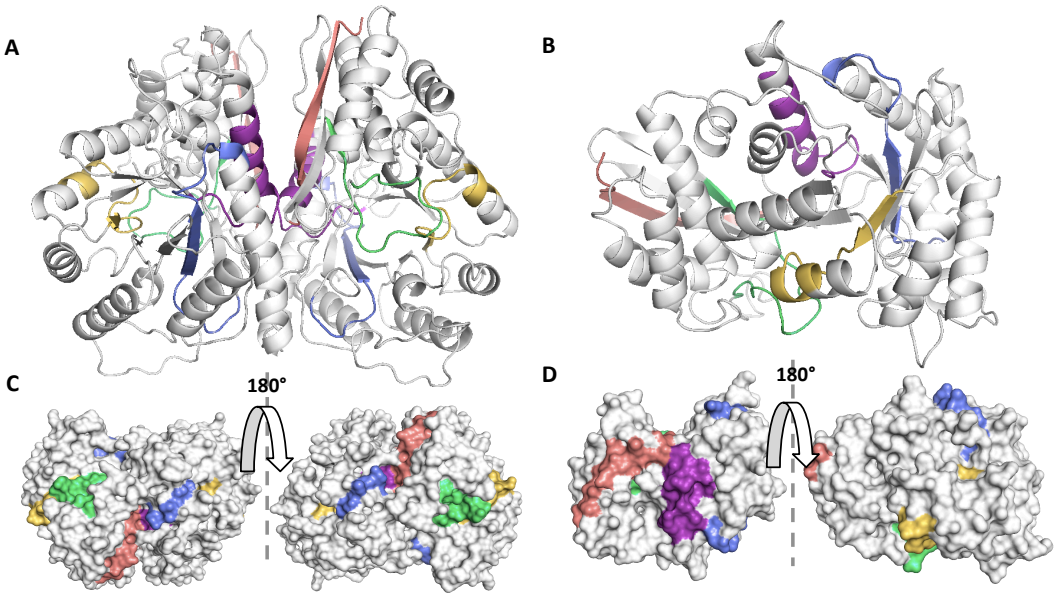


Figure 3.12 *PgEnolase homology models with mapped RPgE1-6 positions.* Homology models for PgEnolase were created by W. W. Yue based on human α-enolase (pdb 3ZLG) (A, C) and streptococcal enolase (pdb 2PSN) (B, D). RPgE1-6 are highlighted in colour: RPgE1/2 red, RPgE3 green, RPgE4 blue, RPgE5 yellow, RPgE6 purple.

Formation of citrullinated peptides from fibrinogen by PPAD, PAD2 and PAD4

Fibrinogen was incubated with RgpB or proteinase K to create C-terminal, or internal arginine peptides respectively, before PAD2 or PAD4 were added. Of the 19 citrullinated peptides formed from the fibrinogen α-chain by PPAD and RgpB, 17 were also identified in the PAD2 and PAD4 MS result spectra, and so were used to compare the levels of citrullination (Table 3.4A). The 17 peptides were citrullinated by PPAD in 14-100% of occurrences in the MS data, compared to 0-60% by PAD2 and 0-20% by PAD4. Overall, PAD2 and PAD4 citrullinated 25% and 4% of C-terminal arginines from all three fibrinogen chains respectively (Table 3.4B). Both PAD2 and PAD4 citrullinated more internal arginine containing peptides formed by incubation with proteinase K, with 80% and 60% of internal arginines from fibrinogen citrullinated.

Table 3.4 *Citrullination of fibrinogen by human PAD2 and PAD4.*

aa	C-terminal arginine peptides from fibrinogen α-chain	Occurrence of Citrulline				
		tPPAD^{WT}	PAD2		PAD4	
20-35	ADSGEGDFLAEGGGVR	100%	0/9	44%	0/6	0%
70-84	MKGLIDEVNQDFTNR	100%	-	-	0/1	0%
84-114	INKLKNSLFEYQKNNKDSHSLTTNIMEILR	100%	1/6	17%	0/4	0%
114-123	GDFSSANNR	100%	0/2	0%	0/1	0%
144-160	KVIEKVQHIQLLQKNVR	100%	-	-	0/1	0%
161-168	AQLVDMKR	100%	-	-	-	-
168-178	LEVDIDIKIR	100%	-	-	-	-
191-216	EVDLKDYEDQQQKQLEQVIAKDLLPSR	100%	0/1	0%	0/2	0%
237-258	FKSQLQKVPPEWKALTDMPQMR	100%	8/19	42%	1/16	6%
259-271	MELERPGGNEITR	100%	0/4	0%	0/2	0%
271-287	GGSTSYGTGSETESPR	100%	0/6	0%	0/3	0%
288-308	NPSSAGSWNSGSSGPGSTGNR	100%	2/12	17%	0/9	0%
309-343	…SGPGSTGSWNSGSSGTGSTGNQNPGSPR	100%	-	-	0/2	0%
354-367	PGSTGTWNPGSSER	100%	1/2	50%	0/2	0%
367-394	GSAGHWTSESSVSGSTGQWHSESGSFR	100%	3/5	60%	1/5	20%
405-425	PNNPDWGTFEEVSGNVSPGTR	17%	2/8	25%	0/1	0%
427-443	EYHTEKLVTSKGDKELR	14%	1/4	25%	1/10	10%
444-458	TGKEKVTSGSTTTTR	25%	-	-	0/1	0%
548-573	GSESGIFTNTKESSSHHPGIAEFPSR	15%	3/7	43%	0/4	0%
574-591	GKSSSYSKQFTSSTSYNR	N/A	3/9	33%	0/2	0%

	Occurrence of citrulline	
PAD2	**C-terminal**	**Internal**
Fib α-chain	4/31	10/10
Fib β-chain	0/3	-
Fib ϒ-chain	28/94	33/41
Total	32/128 (25%)	43/51 (84%)
PAD4	**C-terminal**	**Internal**
Fib α-chain	3/101	3/5
Fib β-chain	2/29	-
Fib ϒ-chain	0/4	0/2
Total	5/134 (4%)	3/5 (60%)

Fibrinogen was incubated with RgpB or proteinase K and PAD2 or PAD4 o/n at 37°C in PAD buffer. Peptides formed were sequenced by LC-MS. Top) Citrullination of C-terminal arginine residues from the fibrinogen α-chain by PPAD, PAD2 and PAD4. Bottom) Citrullination of C-terminal or internal arginine residues from fibrinogen α, β, and γ-chains by PAD2 and PAD4.

Discussion

The aim of this chapter was to characterise the activity of PPAD *in vitro*, in order to evaluate potential mechanisms linking PPAD to autoimmunity in RA, namely formation of novel citrullinated peptides or cross reactivity between citrullinated bacterial and human proteins.

First, the effects of truncating the NtSP, CTD and IgLF around the CD were evaluated. These findings showed a PPAD construct comprising CD and IgLF only was the most soluble and active, and minor truncations to the CD or IgLF could significantly affect activity or solubility. This construct tPPAD^{WT} was used to elucidate the 3D structure of PPAD, and as a PPAD standard in further findings presented in this thesis. The observed effects of PPAD truncation may be due to native processing of PPAD in *P. gingivalis*. Both PPAD and Rgp are processed and secreted through the PorSS, which has been well characterised for RgpB (Zhou *et al.*, 2013). Due to similarities with RgpB in subdomain architecture, namely the presence of a CTD domain, which characterises PorSS proteins, PPAD is thought to undergo the same process. Briefly, the full-length 62kDa immature precursor enzyme is targeted for translocation across the inner membrane via the NtSP, which is then removed to activate the latent enzyme (Veillard *et al.*, 2013). The CTD then targets the active protein to the PorSS for translocation through the outer membrane (Sato *et al.*, 2010, Shoji *et al.*, 2011), during which the CTD is removed by the *P. gingivalis* enzyme PG0026, or PorU (Glew *et al.*, 2012, Gorasia *et al.*, 2015). Removal of the CTD forms the mature enzyme of 46kDa (McGraw *et al.*, 1999), which is anchored to the bacterial membrane through putative binding of the remaining linker region following the IgLF (Zhou *et al.*, 2013), to anionic-LPS on the bacterial surface (Gorasia *et al.*, 2015). Once secreted, *P.*

gingivalis proteins, including PPAD and RgpB, form large multi-unit complexes on the bacterial surface (Glew *et al.*, 2014).

Understanding the mechanisms of PorSS processing provides explanations for the observed solubility and activity variations in our PPAD constructs. tPPAD^{WT} was the most soluble and active construct, and has the most comparable domain boundaries to the mature, virulent form of PPAD secreted from *P. gingivalis*, i.e. the CD and IgLF domains, lacking the NtSP and CTD. Similarly, the full-length precursor enzyme comprising NtSP and CTD is not secreted, and so is unlikely to remain soluble in our recombinant system. In addition, one study using another ProSS protein in *P. gingivalis* (hemin binding protein 35) identified a specific 22 amino acid portion of the CTD that was required for secretion and subsequent membrane attachment (Shoji *et al.*, 2011). Of these 22 amino acids, 20 are conserved amongst other PorSS proteins (Seers *et al.*, 2006), and so are thought to be required for CTD function. This is supported in the current study, which shows that only partial truncation of the CTD from lys556 to ala484 results in a significant increase in activity (c073 compared to c007).

In vivo, naturally occurring substrates for PPAD with C-terminal arginine residues are rare, and so citrullination alone as a defence mechanism for *P. gingivalis* would be inefficient. However, as a functional partner and an active arginine protease, Rgp supplies C-terminally located arginine substrates for PPAD. It is possible that citrullination of host proteins following cleavage by Rgp provides extra bacterial defence to ensure anti-microbial products are completely inactive. In support of this mechanism is the inactivation of complement component C5a following citrullination by PPAD (Bielecka *et al.*, 2014). In addition, in the study presented in this chapter, co-incubation of

fibrinogen with active tPPAD^{WT} and Rgp appeared to increase the degradation of fibrinogen by Rgp, possibly by augmenting protein architecture to expose cleavage sites for Rgp. This supports the hypothesis that PPAD and RgpB function together *in vivo*, and have mutually complementary action.

Fibrinogen and α-enolase are two of the most commonly occurring ACPA reactivities and occur before disease onset (Rantapaa-Dahlqvist *et al.*, 2003, Nielen *et al.*, 2005, van de Stadt *et al.*, 2011a). In this study, PPAD and RgpB formed 41 C-terminally citrullinated peptides from fibrinogen, and 11 from α-enolase. In addition, citrullination of the same C-terminal arginine residues by PAD2 and PAD4 was compared, and was significantly lower than that of PPAD. However, both PAD2 and PAD4 were efficient in citrullinating internal arginines, which had been shown previously for both fibrinogen and α-enolase (Wegner *et al.*, 2010, van Beers *et al.*, 2010). These findings confirm, and advance upon a previous study in our lab, whereby the formation of C-terminally citrullinated peptides from fibrinogen and α-enolase by *P. gingivalis* was dependent on expression of both PPAD and Rgp (Wegner *et al.*, 2010).

These findings demonstrate that citrullinated peptides created during *P. gingivalis* infection would differ from native citrullinated peptides, and so could be recognised as foreign by the immune system to drive an ACPA response. B-cells may directly bind blood borne antigens to produce an antibody response independently of T-cells (Obukhanych and Nussenzweig, 2006), however the increasing titre and avidity of ACPA pre-RA (Suwannalai *et al.*, 2012) indicates T-cell help is required in their development, which is also supported by the strong association of the *HLA-DRB1* SE alleles with ACPA^{+} RA. To this end, numerous studies have examined the binding of citrullinated peptides to MHC SE alleles,

and shown the preferential binding of citrullinated peptides compared to native arginine ligands (Hill *et al.*, 2004, James *et al.*, 2010, James *et al.*, 2014). However, in these studies peptides with internal citrulline residues (as would be generated by PAD2 and PAD4) were used, and to date no comparative studies with C-terminally citrullinated peptides have been documented. Perhaps the most compelling evidence to support antigenicity of citrullines located at or near the C-terminus is the murine model of autoimmunity in mice transgenic for hen egg lysozyme (HEL). In these studies, mice generated to present HEL by murine MHC-II I-A^k were immunised with HEL in CFA, and T-cell hybridomas generated using cells isolated from the draining lymph nodes. The resulting hybridomas were screened for reactivity with HEL peptides, which identified the immunodominant T cell epitope $HEL_{48\text{-}62}$, which has a C-subterminal citrulline (^{48}DGSTDYGILQINS-Cit-W^{62}) (Ireland *et al.*, 2006). In addition, crystallisation of the native $HEL_{48\text{-}62}$ peptide bound to murine MHC-II I-A^K showed that residues Arg61 and Trp62 of the peptide at positions P10 and P11 of the MHC-II molecule were solvent exposed, and thus binding was unaffected by changing Arg61 to Cit61 (Fremont *et al.*, 1998). These findings propose an additional binding conformation for C-terminally citrullinated peptides, distinct from the established model of internal citrulline residues favourably binding in the P4 pocket of MHC-II SE alleles. In this conformation, the non-citrullinated peptide region is secured within the MHC-II binding groove, resulting in solvent exposed citrulline residue(s) at the C-terminus. This may be more likely to induce a citrulline specific T-cell response, as the citrulline residue(s) are positioned to bind the T-cell receptor (TCR). In a later study by Ireland and colleagues using the same model, dendritic cells and macrophages isolated from HEL expressing

transgenic mice presented the citrullinated HEL_{48-62} peptide, however splenic B-cells presented the non-citrullinated form (Ireland and Unanue, 2011). This could explain the observation that PD patients often have an elevated frequency of ACPA, which also react with the non-citrullinated epitope variant (de Pablo *et al.*, 2013, Lappin *et al.*, 2013).

Another potential mechanism for PPAD in RA involves cross-reactivity with auto-citrullinated PPAD to drive an ACPA response. This is supported by findings that full-length recombinant PPAD is auto-citrullinated, and RA patients have an increased antibody response to citrullinated compared to non-citrullinated PPAD (Quirke *et al.*, 2014a), however it has also been suggested that auto-citrullination of PPAD is an artefact of recombinant *E.coli* expression which would not occur *in vivo* (Konig *et al.*, 2014). PPAD auto-citrullination also contradicts the observation that PPAD is exclusively reactive with C-terminal arginine substrates. In this study, no arginine residues were found citrullinated in full-length or truncated PPAD from either recombinant sources or isolated from native *P. gingivalis* when analysed by LC-MS. However, when PPAD and RgpB were incubated together, a number of citrullinated peptides were identified from both enzymes, which could be antigenic when formed *in vivo*. Studies by our group have found significantly increased antibody responses to citrullinated peptides spanning the PPAD molecule in RA patients, which ranged from 15%, to 40% reactivity with citrullinated PPAD peptide CPP3 (^{71}CAKTDSYWT-Cit-DYTGWFAMYDC^{89}) (Quirke *et al.*, 2014a, Quirke *et al.*, 2014b). Although these peptides have internal citrullines, they demonstrate the capability for PPAD to act as a citrullinated antigen, and support the hypothesis that antibodies to PPAD in *P. gingivalis* infection could prime an ACPA response.

In addition to PPAD, cross reactivity between PgEnolase and human α-enolase has also been implicated as a mechanism for autoimmunity in RA (Lundberg *et al.*, 2008). Here, C-terminally citrullinated peptides formed by PPAD and RgpB from PgEnolase were highly homologous to peptides formed from α-enolase and so could result in a cross-reactive antibody response. This will be investigated in further chapters of this thesis.

The main findings from the studies in this chapter were:

- Truncation of NtSP and CTD domains of PPAD increases recombinant PPAD solubility and activity, possibly by replicating PorSS processing in *P. gingivalis.*
- PPAD protein purified from *E.coli* or *P. gingivalis* was not auto-citrullinated by MS analysis, however citrullinated PPAD peptides formed by RgpB and PPAD may still be antigenic *in vivo.*
- PPAD and RgpB effectively form C-terminally citrullinated peptides from RA autoantigens fibrinogen and α-enolase *in vitro*, distinct from those formed by PAD2 and PAD4.
- C-terminally citrullinated peptides from PgEnolase homologous to α-enolase were identified as potential mediators of cross-reactivity between bacterial and human enolase.

Chapter 4 PPAD, RgpB and PgEnolase *in vivo*: murine models

Background

Animal models of disease are integral to biomedical research, and have been particularly valuable in the study of RA, notably in the development of anti-TNF biologic therapy (Williams *et al.*, 1992). A number of *in vivo* systems are regularly used to model RA pathogenesis, either by inducing auto-immunity targeting the joint, e.g. collagen-induced arthritis (CIA) (Inglis *et al.*, 2007), by passive transfer of joint specific antibodies, e.g. collagen antibody induced arthritis (CAIA) (Holmdahl *et al.*, 1986), or by exploiting spontaneous inflammatory symptoms in mutant mouse strains, e.g. KxB/N mice (Kouskoff *et al.*, 1996). However, none have described an aetiological pathway to RA development *in vivo*.

The proposed aetiological association between PD and RA has resulted in a small number of studies investigating the effect of *P. gingivalis* on pathology in murine models of arthritis. Marchesan, Gully, and their colleagues reported increased incidence and severity of CIA and CAIA respectively in mice pre-treated with *P. gingivalis* using the *P. gingivalis* induced model of PD (Marchesan *et al.*, 2013, Gully *et al.*, 2014). Similarly, administration of *P. gingivalis* via the chamber coil model also increased incidence and severity of CIA (Maresz *et al.*, 2013). In these studies, increased severity of disease was dependent on pre-treatment with live *P. gingivalis* expressing PPAD, indicating PPAD may be the cause of this effect.

Further to this, previous studies from our group reported immunisation of DR4-Tg mice with PgEnolase induced inflammatory arthritis, and anti-CEP-1 and anti-

REP-1 antibody responses (Kinloch *et al.*, 2008). These findings suggest PgEnolase is antigenic *in vivo*, and may induce anti-CEP-1 antibodies through cross reactivity with PgEnolase. In RA patient sera, affinity purified anti-CEP-1 antibodies cross-react with citrullinated recombinant PgEnolase (Lundberg *et al.*, 2008).

To that end, the aim of this chapter was to investigate the effects of PPAD, co-immunised with RgpB to maximise function, and PgEnolase in mice to replicate *P. gingivalis* infection in the aetiology of RA.

The main aims were:

- To develop a specific anti-PgEnolase antibody for use with anti-PPAD and anti-Rgp antibodies in identifying the location of PPAD, Rgp and PgEnolase in the *P. gingivalis* bacterium.
- Measure antibody responses to PPAD, RgpB and PgEnolase in immunised mice.
- Investigate antibody responses to citrullinated peptides from PgEnolase, as a mechanism for breach of tolerance to human α-enolase.
- Investigate formation of ACPA in immunised mice.

Buffers

Western blotting

As in Chapter 2

Biotin-phenylglyoxal Probe

Biotin-Pg Probe Buffer	1% SDS
	7% beta mercaptoethanol
	20mM HEPES
	100mM arginine
	100mM NaCl

Enzyme Linked Immunosorbent Assay (ELISA) with murine sera

Carbonate coating buffer	*50mM carbonate buffer pH 9.5*
	15mM Na_2CO_3
	35mM $NaHCO_3$
	0.02% NaN3
Blocking buffer	1x PBS
	5% bovine serum albumin (BSA)
Wash buffer	1x PBS
	0.05% tween
Sample buffer	1x PBS
	0.05% tween
	0.1% BSA
Stop solution	2M H_2SO_4

Splenocyte proliferation assay

Ammonium-Chloride-potassium (ACK) lysis buffer	10mM $KHCO_3$
	150mM NH_4Cl
	0.1mM EDTA
	pH 7.2-7.4
Trypan blue	0.2% trypan blue
	1x PBS
Complete RPMI	10% heat-inactivated foetal bovine serum (FBS)
	1% penicillin streptomycin
	1% HEPES
	1% non-essential amino acids
	1% GlutaMAX
	1% sodium pyruvate
	Roswell Park Memorial Institute (RPMI) media

Materials and methods

Development of anti-PgEnolase antibody

Cambridge Biosciences were commissioned for development of an anti-PgEnolase antibody. Two peptide regions from PgEnolase were predicted to be immunogenic for antibody production in rabbits using epitope prediction software (not disclosed by Cambridge Biosciences). Two rabbits were immunised with peptide 1 ($_{417}$CAVYGYKKV$_{425}$-acid) (rabbits 1.1 and 1.2) or peptide 2 ($_{251}$[C]-GIYDYTKFEGEK$_{262}$-amide) (rabbits 2.1 and 2.2) in CFA, and boosted x4 with the corresponding peptide in incomplete Freund's adjuvant (IFA). A pre-immune and four sequential bleeds were taken from each rabbit (as outlined in the schedule in 'Results') providing 20 anti-sera samples. Anti-sera were tested for reactivity to PgEnolase and cross-reactivity to human α-enolase by standard western blotting protocol, using a 1:500 dilution of anti-sera as primary antibody to probe 10μg each per lane of recombinant PgEnolase (purified in house) and α-enolase (MyBioSource (Cat no: MBS203689)). Anti-sera from rabbit 1.1 was the most highly reactive with PgEnolase, and had no cross reactivity to α-enolase. Antibody was purified by affinity purification from rabbit 1.1 sera, and stored in glycine. This antibody is used as 'anti-PgEnolase' in western blotting throughout this thesis. Anti-human α-enolase antibodies, anti-CEP-1 (in house), and anti-Enolase H300 raised against the N-terminal 300aa protein fragment (Santa Cruz), were blotted against PgEnolase and α-enolase as controls. Anti-CEP-1 was diluted 1:10000 in 3% milk-TTBS, and anti-Enolase H300 was diluted 1:3000 in 3% milk-TTBS and a standard western blotting protocol followed.

Preparation of P. gingivalis fraction lysates

P. gingivalis was cultured and fractionated as previously described (Potempa and Nguyen, 2007) by Dr. K. Gawron, and D. Staniec and kindly gifted from the lab of Prof. J. Potempa. Two sets of fractions: whole cell extract (WCH); inner membrane (IM); periplasmic fraction (PF); outer membrane (OM); and cytoplasmic fraction (CF) were prepared in parallel, and one set made to 10mM final conc. of TLCK protease inhibitor, with the exception of PF which was prepared without inhibitors only. Fractions were stored at -20°C until use. A Bradford assay was carried out to calculate concentration. Briefly, fractions were thawed, vortexed thoroughly, and serial dilutions in duplicate made on a 96-well plate (Costar) with a range of BSA protein standards. PBS was added in duplicate as a blank control. Pierce Coomassie Plus (Bradford) Assay kit (ThermoFisher) was added as per manufacturer's instructions and absorbance read at 595nm. Protein concentration was calculated in Microsoft office excel using the equation of the standard curve. Values were mean n=2 minus blank.

Western blotting of P. gingivalis fraction lysates

Fractions were thawed on ice and 0.5µg prepared in 10µl of ddH_2O with loading buffer (see 'Buffers'). An equal concentration of relevant positive controls was also prepared. Proteins were resolved by SDS-PAGE as before and transferred to nitrocellulose for western blotting as before. Membrane was blocked o/n at 4°C in 3% milk-TTBS and a standard western blotting protocol observed as described. Primary and secondary antibodies were diluted in 3% milk-TTBS and incubated for 1h at rt. Antibodies were used at the following dilutions: anti-PPAD 1:1000 (Rabbit, in house), anti-gingipain 1:15000 (rabbit, kind gift from

Prof. J. Potempa) and anti-PgEnolase 1:1000 (rabbit, in house), anti-rabbit HRP secondary antibody 1:5000 (goat, Dako).

Visualisation of citrullination by biotin-phenylglyoxal (Biotin-Pg) probe

Labeling citrullinated proteins with Biotin-Pg probe

Biotin-Pg probe was kindly gifted from the lab of Prof. P. Thompson. Full details of use for Biotin-Pg probe can be found in (Hensen *et al.*, 2015). Briefly, 5µg of *P. gingivalis* lysates and 1µg of citrullinated α-enolase as a positive control were diluted in 20µl ddH_2O then 20% final conc. of trichloroacetic acid (TCA) was added, before 0.1mM (final conc.) of biotin-PG probe. Samples were incubated at 37°C for 30min. The reaction was quenched with 100mM citrulline (final conc.) in 50mM HEPES, pH7.6. After quenching, samples were incubated for a further 30min at 37°C then 30min on ice to complete precipitation. Proteins were isolated by centrifugation at 16,000g for 1min at 4°C and supernatant discarded. 50µl 4°C acetone was added to each pellet and centrifuged again at 16,000g for 15min at 4°C. Supernatant was discarded and tubes left open for residual acetone to evaporate for 10min at rt.

Eluting citrullinated proteins

Pellets were resuspended in 10µl of Biotin-Pg probe buffer (see 'Buffers'), heated to 100°C for 10min and sonicated for 10min. Steps were repeated until the pellet was completely resuspended. 200µl 0.2% SDS in PBS was added to each sample. High capacity streptavidin-agarose 50% slurry (ThermoFisher) was prepared in PBS. 50µl slurry was added per sample and incubated o/n at 4°C. Resin was isolated by centrifugation at 500g for 2min at 4°C and supernatant collected as flow through for analysis. Resin was washed x3 with 300µl 0.2%

SDS in PBS for 10min at rt, and washed x3 with ddH_2O for 10min at rt. Proteins were eluted with 50µl 2x LDS sample buffer (LifeTechnologies) and heated at 100°C for 5min. Unbound resin was removed by centrifugation at 500g for 2min, and the supernatant collected for analysis. 20µl of supernatant(s) were prepared for SDS-PAGE and protein transfer as previously described. Following transfer, membrane was incubated with 1:5000 dilution of streptavidin HRP in 5% BSA for 1h at rt before washing and protein visualisation as before.

Murine immunisation with PPAD, RgpB and PgEnolase

Primary immunisations with PPAD

Primary immunisations with flPPAD were carried out in house, the Kennedy Institute of Rheumatology, London. Dr K. McNamee and Mr. D. Perachau provided technical guidance during procedures. All experimental procedures were approved by the local ethical review process committee and the UK Home Office in accordance with the Animals (Scientific Procedures) Act 1986.

CFA was prepared by mixing heat-treated *Mycobacterium tuberculosis* (BD Biosciences) in IFA (BD Biosciences) in a mortar and pestle to 3.3mg/ml. Treatment enzymes (flPPAD^{WT} or flPPAD^{C351A}) were expressed and purified as described in Chapter 2, and emulsified in CFA to give a treatment dose of 50µg enzyme in final volume of 150µl (0.33mg/ml). Control groups were immunised with 150µl emulsified CFA or IFA alone. Emulsified preparations were kept on ice until immunisation.

Male 10-week-old DBA/1 (DBA) mice (Harlan) were used due to the previous experience using this strain in house, and their susceptibility to CIA. flPPAD was used as immunising agent before the development of more efficient tPPAD

expression system. Mice were immunised with 150µl of flPPAD antigen preparation at two sites intra-dermally at the base of the tail whilst anaesthetised. Mice were monitored daily for signs of adverse or severe reactions, and scored for symptoms of arthritis for the 14 or 30-day duration of the experiment. Treatment groups were caged separately to prevent caging bias. Mice were housed in groups of 5 in ventilated cages, maintained at a mean temperature of 21°C (±2°C) in a 12-hour light/12-hour dark cycle, with food and water available *ad libitum*.

Mice were sacrificed by CO_2 and cervical dislocation. Blood was collected by cardiac puncture postmortem and allowed to clot on ice for 30min before centrifugation at 2000g for 10min.

Immunisations with PPAD, RgpB & PgEnolase

Immunisations with $tPPAD^{WT}$, PgEnolase & RgpB were carried out in the lab of Prof. E. Cairns, Western University, London, Ontario, Canada. Mr. A. Patel and Ms. S. Saunders assisted on various aspects of experiments. All animal experiments in London, Canada (Cairns laboratory) were performed in accordance with the Canadian Council on Animal Care Guide to the Care and Use of Experimental Animals. The animal protocol (# 2009-101) was approved by the Animal Use Subcommittee at Western University.

$tPPAD^{WT}$ was expressed and purified as described previously in Chapter 2. PgEnolase was expressed and purified using the flPPAD purification protocol described previously in Chapter 2. RgpB was kindly donated by Prof. J. Potempa. Proteins were prepared to desired concentration in final volume 200µl of sterile PBS. Ten-week-old female C57BL/6 (B6) mice were purchased from The Jackson Laboratory (Maine, USA). B6 mice were used due to previous experience with

this strain in house and to enable future experiments in DR4-Tg mice, which have a B6 genetic background. Mice were immunised with either: PPAD 100µg, PgEnolase 100µg, and RgpB 10µg; PPAD 100µg, and RgpB 10µg; PPAD 100µg alone; PgEnolase 100µg alone by intra-peritoneal injection (IP). One mouse from each treatment group was sacrificed day 3 post immunisation and a dissection was carried out to detect internal signs of severe adverse reactions to immunisations, of which there were none. Mice were monitored daily for signs of distress and symptoms of arthritis for the 10 or 28 day duration of the experiments. Mice were housed in the Animal Care and Veterinary Services Barrier Facility at Western University under specific pathogen free conditions, in groups of 4 in ventilated cages, maintained at a mean temperature of 21°C (±2°C) and a 12-hour light/12-hour dark cycle, with food and water available *ad libitum*. Mice were sacrificed on day 10 post-immunisation or boosted (as on day 0) on day 21 and sacrificed 7 days later on day 28. Mice were sacrificed by overdose of anaesthetic. Post-mortem blood was collected by cardiac puncture and allowed to clot for 30min on ice before centrifugation at 2000g for 10min. Spleens were removed aseptically and stored in RPMI (Gibco, Life Tech.) until use. Data shown is from two repeat experiments as indicated in results figures.

Enzyme Linked Immunosorbent Assay (ELISA) with murine sera

Antigens were diluted to 10mg/ml or 5mg/ml for peptides and proteins respectively in carbonate coating buffer (see 'Buffers'). Whole protein antigens used were recombinant tPPAD^{WT}, and recombinant PgEnolase. Peptide antigens used were CPgE1-6 (^{1}MEIAKIIG-Cit^{9}, ^{6}IIG-Cit-EILDS-Cit^{15}, ^{41}TGENEAIEL-Cit^{51}, ^{168}PVGACCF-Cit^{175}, ^{317}LFVTNVEFL-Cit^{326}, ^{401}MAKYNQLL-Cit^{409}), CEP-1 (^{5}CKIHA-

Cit-EIFDS-Cit-GNPTVEC[21]), CPP3 ([71]CAKTDSYWT-Cit-DYTGWFAMYDC[89]), and cVim ([59]VYAT-Cit-SSAV-Cit-L-Cit-SSVP[74]) with their corresponding arginine controls. 96 well plates (Costar) were coated with 100µl/well of antigen and incubated o/n at 4°C. Coating antigen was washed out x3 with wash buffer (see 'Buffers') (Sigma) and tapped dry. Plates were incubated for 2h at rt with 200µl of blocking buffer (see 'Buffers'), and blocking buffer tapped out without washing. Sera were diluted 1:20 in sample buffer (see 'Buffers') and 100µl per well applied in duplicate. Pooled sera from naïve B6 mice was used as a control, and two wells with sample buffer only were used as a blank. Plates were incubated for 1.5h at rt then washed as before. 100µl anti-mouse secondary antibody (Goat anti-mouse IgM, Jackson, Rabbit anti-mouse IgG, Dako) diluted 1:5000 in sample buffer was added and incubated for 1.5h at rt. Where possible, antigens were tested for both IgM and IgG reactivity. When sera sample volume was limited, reactivity was tested with IgG only. Plates were washed as before. 100µl 3,3′,5,5′-Tetramethylbenzidine (TMB) substrate (BioLegend) per well was added as per manufacturer's instructions, incubated for 5min at rt and 50µl ELISA stop acid added (see 'Buffers'). Absorbance was read at OD450nm.

To measure ACPA a modified anti-CCP2 kit (Eurodiagnostica) was used. Manufacturers methods and reagents were used with two adjustments. Murine sera was diluted 1:100 in sample diluent (supplied) and incubated as per manufacturers instructions, and secondary antibody was replaced with 1:5000 Rabbit anti-mouse IgG (Dako) diluted in sample diluent (supplied). Absorbance was read at OD520nm and values converted to absorbance units (AU) using a standard curve. Due to the commercially designed peptides used in the anti-CCP2 assay, no arginine containing control peptides were available.

Splenocyte proliferation assay

Spleens were collected post-mortem and stored separately in 5ml RPMI on ice. Each spleen was homogenised using a sterile glass homogeniser, washed with 3x 5ml PBS between each homogenisation. Homogenates were centrifuged at 450g for 5min, supernatant discarded, and pellet resuspended in remaining buffer. 4ml Ammonium-Chloride-Potassium (ACK) lysis buffer (see 'Buffers') was added to each tube and incubated for 4min with inversion every minute. 11ml PBS was added per tube and inverted to neutralise. Cells were centrifuged at 450g for 5min, supernatant discarded and pellet resuspended as before. Cells were resuspended in 10ml PBS and any traces of fat or debris removed before centrifugation again at 450g for 5min. Supernatant was discarded, pellet resuspended in 2-5ml complete RPMI (depending on original pellet size), and any remaining debris was removed. 80µl of trypan blue (see 'Buffers') was added to 20µl of final cell suspension and the cell number quantified using a haemocytometer. Cells were diluted to $4x10^6$/ml in complete RPMI (see 'Buffers') and 100µl added to each well in triplicate/antigen. Test antigen tPPAD was diluted to 100µg/ml, and positive control concanavalin A (ConA) (Sigma Diagnostics) to 20µg/ml in RPMI. 100µl of each was added to corresponding wells. Triplicate wells without antigens were used as a negative control. Cells were incubated at 37°C in 5% CO_2 for 54h. 100µl of supernatant was removed, aliquoted into corresponding positions in a 96-well plate and stored -20°C for cytokine analysis. 1µCi of [^{3}H]-thymidine (PerkinElmer) in 100µl sterile complete RPMI was added to splenocytes in each well. Cells were incubated for 18h before being harvested, soaked in 'ReadySafe' scintillation liquid (Beckman Coulter) and read on a 150 Microbeta liquid scintillation counter (Wallac).

Means of n=3 were calculated in Microsoft Excel and values expressed as stimulation index (SI): ratio of mean counts per minute of triplicate cultures with antigen to triplicate cultures without antigen.

Cytokine ELISA with splenocyte supernatants

Cell supernatants were harvested from stimulated splenocytes as described, and pooled by treatment group. To measure cytokine production from splenocytes, BD OptEIA™ ELISA kits for IL-6, TNF-α and IFN-γ (BD Biosciences) were used as per manufacturer's instructions. Briefly, a standard curve was prepared by serial dilution of standard provided in sample diluent. 50µl ELISA diluent was pipetted into each well, and 100µl of supernatant pool or standard added in duplicate. Plates were sealed and incubated for 2h at rt. Plates were washed x5 in ELISA wash buffer (supplied) and tapped dry. Working detector was prepared by adding enzyme concentrate to detection antibody provided, and 100µl was added to each well. Plates were sealed and incubated for 1h at rt before washing x7 as before. 100µl TMB One-Step reagent was added and incubated at rt for 30min. 50µl stop reagent was added per well and plates read at 450nm. Data was analysed as per serum ELISA.

Statistical analysis

For all ELISAs and the anti-CCP2 assay, data calibration was carried out in Microsoft Excel. Means are n=2 minus blank and expressed as OD450nm values or Absorbance Units (AU) or Pg/ml where a standard curve was available. Statistical significance was determined using Mann-Whitney testing in GraphPad Prism 6. Values for experimental repeats are indicated with different symbol colouring in graphs as indicated in figure legends.

Results

Development of anti-PgEnolase antibody

Two rabbits each were immunised with PgEnolase peptides 1) CAVYGYKKV-acid and 2) C]-GIYDYTKFEGEK-amide. Rabbits were bled before and sequentially after immunisations, providing 5 serum samples each (Table 4.1). Reactivity to PgEnolase in the antiserum was determined by western blotting in parallel with human α-enolase to determine levels of cross-reactivity (Figure 4.1). All antisera reacted exclusively with PgEnolase, and no cross-reactivity with human α-enolase was observed. In all samples protein bands were identified at ~50kDa, above the expected molecular weight of PgEnolase of 46kDa, and at ~100kDa, which could indicate dimerisation of the protein. Antisera from rabbit 1.1 was most reactive, indicated by strongest protein bands, and so was harvested, affinity purified and stored for further use as 'anti-PgEnolase'.

Table 4.1 *Schedule of anti-PgEnolase antibody development.*

Date	Operation	Method	Sera collected
01-May-15	Pre-immune bleed #1	Intravenous	5ml
01-May-15	Immunisation #1	Subcutaneous nodal in CFA	
15-May-15	Immunisation #2	Subcutaneous dorsal in IFA	
22-May-15	Bleed #2	Intravenous	5ml
29-May-15	Immunisation #3	Subcutaneous dorsal in IFA	
05-Jun-15	Bleed #3	Intravenous	5ml
12-Jun-15	Immunisation #4	Subcutaneous dorsal in IFA	
19-Jun-15	Bleed #4	Intravenous	20ml
26-Jun-15	Immunisation #5	Subcutaneous dorsal in IFA	
03-Jul-15	Bleed #5	Intravenous	20ml
10-Jul-15	Immunisation #6	Subcutaneous dorsal in IFA	
17-Jul-15	Harvest Bleed	Intravenous	50-80ml

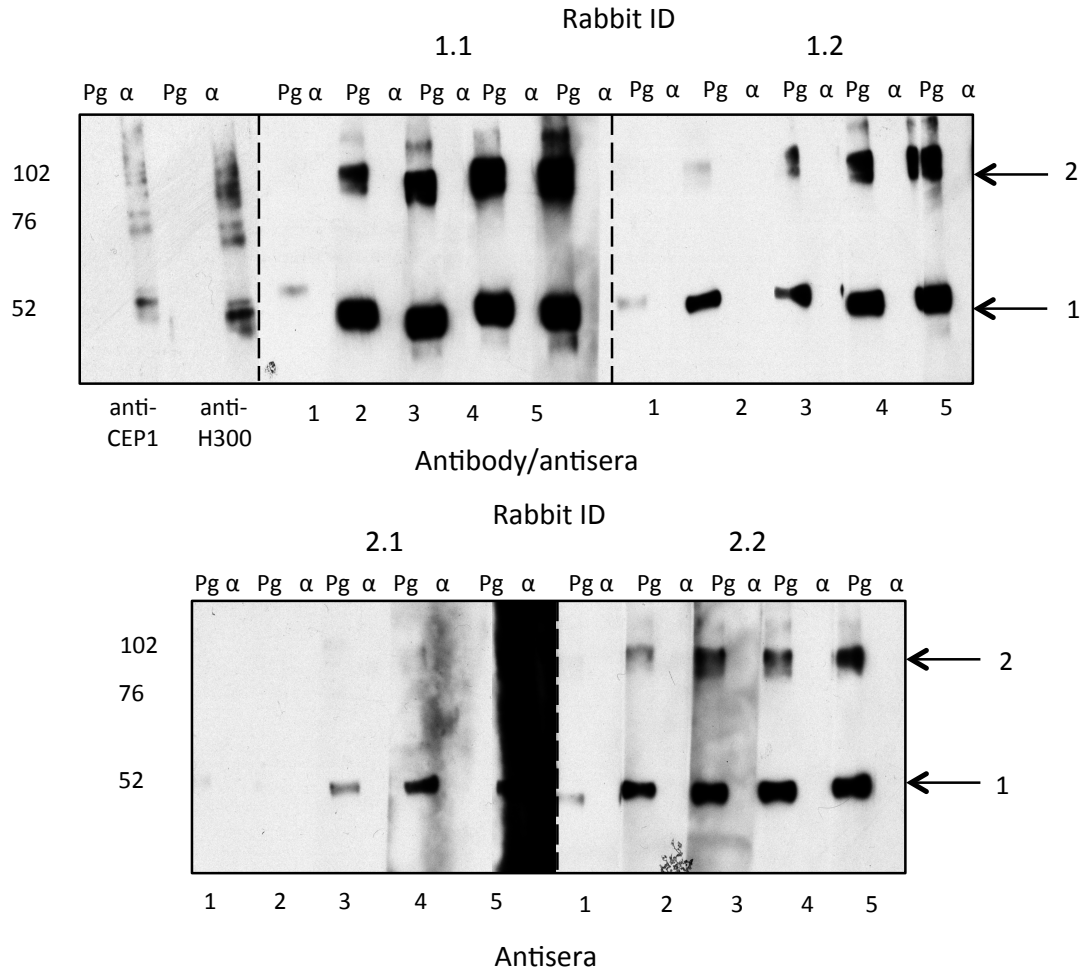


Figure 4.1 *Reactivity of rabbit antisera with PgEnolase and α-enolase.* 10μg per lane of PgEnolase and human α-enolase were resolved by SDS-PAGE in pairs labelled as 'Pg' and 'α' on blots. Proteins were transferred to nitrocellulose membrane, and membranes cut into strips. Each membrane strip with PgEnolase and human α-enolase was incubated with test bleed sera (1:500) or anti-human α-enolase antibodies anti-CEP1 (1:10000) and anti-H300 (1:3000). Pre-immune sera and four test bleeds labelled (1-5) were obtained from two rabbits immunised with peptide 1 (1.1 and 1.2) and two rabbits immunised with peptide 2 (2.1 and 2.2). PgEnolase is indicated with arrow 1, a possible PgEnolase dimer is indicated with arrow 2.

Location of PPAD, RgpB and PgEnolase in P. gingivalis fractions

P. gingivalis fractions: whole cell extract (WCH), inner membrane (IM), periplasmic fraction (PF), outer membrane (OM), and cytoplasmic fraction (CF), were resolved by SDS-PAGE alongside PPAD, RgpB and PgEnolase as positive controls, transferred to nitrocellulose, and probed using anti-PPAD, anti-RgpB and anti-PgEnolase antibodies (Figure 4.2). PPAD, Rgp and PgEnolase were all found in the WCH (with inhibitor), confirming their presence in the bacterium (Figure 4.2). PPAD migrated in two bands, a diffuse upper band at ~70kDa and a clear lower band at 50kDa, alongside the positive control $tPPAD^{WT}$. This is consistent with previous findings in this thesis that may be due to PorSS processing (Figure 4.1 arrows 1 and 2). However, only one Rgp band was identified when probed, at ~70kDa above the control RgpB at ~55kDa (Figure 4.2 arrow 3), which may suggest the anti-Rgp antibody used reacts exclusively with the full-length form of the enzyme, and not the truncated form. PgEnolase migrated slightly above the expected molecular weight of 46kDa, at ~52kDa (Figure 4.2 arrow 4), which may also indicate some modifications in the bacterium. The GST-tagged positive control migrated at 75kDa, due to the added molecular weight of GST fusion protein.

PPAD and RgpB were found in both the IM and in higher amounts in the OM. Neither showed difference in expression with or without protease inhibitor. PgEnolase was present in high levels in the CF and PF, but was identified in the OM fraction with inhibitor only, indicating PgEnolase may be degraded in the OM fraction, which contains high levels of RgpB, without protease inhibitors.

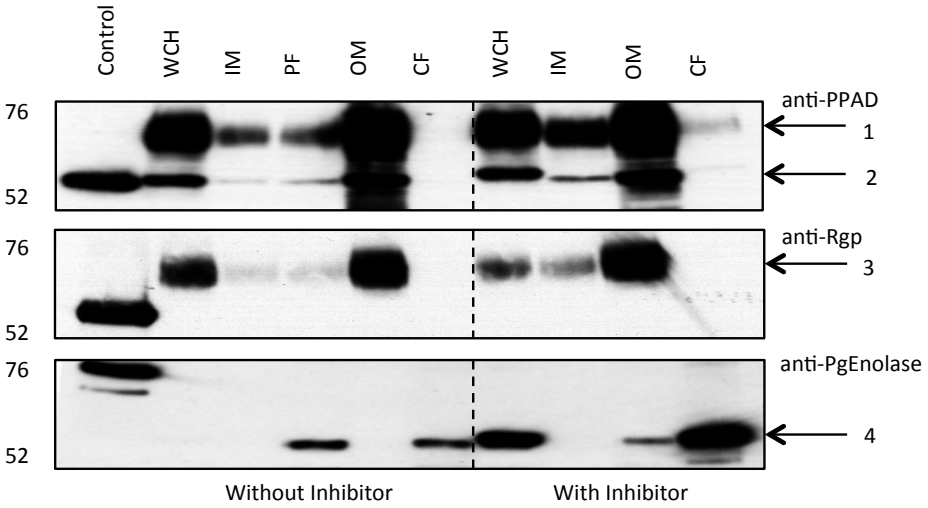


Figure 4.2 *Location of PPAD, Rgp and PgEnolase in P. gingivalis fractions. P. gingivalis* was fractionated into WCH, IM, PF, OM and CF with or without TLCK. Fractions were lysed, and protein concentration measured. 0.5µg of lysed *P. gingivalis* fraction protein and relevant positive controls (PPAD, RgpB or PgEnolase) were resolved by SDS-PAGE and transferred to nitrocellulose. Nitrocellulose membranes were incubated with 1:15000 dilution of anti-Rgp, 1:1000 dilution anti-PgEnolase or 1:1000 dilution anti-PPAD. Proteins were visualised using a streptavidin-HRP secondary antibody and ECL.

Location of citrullination in P. gingivalis fractions

Citrullinated proteins from *P. gingivalis* WCH, IM, PF, OM, and CF were isolated using a citrulline specific biotin-Pg probe and resolved by SDS-PAGE (Figure 4.3). The highest level of citrullination was found in WCH and OM, irrespective of the presence of TLCK protease inhibitor in the sample. The CF fraction with protease inhibitor also has moderate levels of citrullination. No citrullinated proteins were detected in IM or PF. Taken together with the western blotting findings, these data indicate highest levels of citrullination occur in the presence of PPAD, RgpB and PgEnolase.

In addition, three clear bands were identified consistently between WCH and OM, suggesting the OM is the main source of citrullinated proteins (Figure 4.3 arrows 1, 2 and 3). These were diffuse bands migrating at ~46kDa, ~55kDa and ~70kDa, consistent with previous western blots of *P. gingivalis* fractions that show diffuse, heavily glycosylated protein bands. One of these bands was also visible in CF in the presence of TLCK. This could be due to contamination or citrullinated proteins getting degraded beyond recognition by the Biotin-Pg probe in the CF of the bacterium with active proteases.

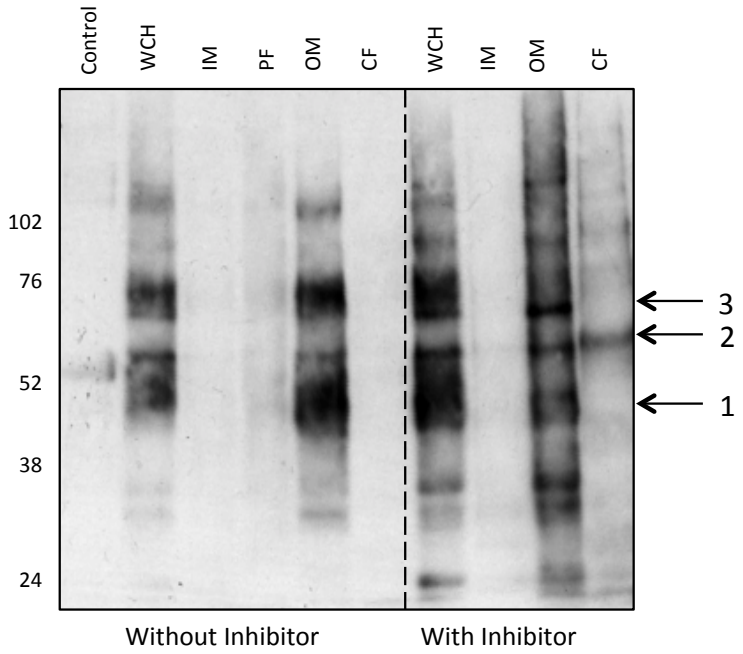


Figure 4.3 *Location of citrullination in P. gingivalis bacterial fractions. P. gingivalis* was fractionated into WCH, IM, PF, OM, and CF with or without TLCK. Fractions were lysed, and protein concentration measured. 5µg of protein was incubated with a Biotin-Pg probe to label citrullinated proteins. Proteins were precipitated from the sample and citrullinated proteins isolated using streptavidin resin. Citrullinated proteins were eluted from resin, resolved by SDS-PAGE with 1µg of citrullinated α-enolase as a positive control, and transferred to nitrocellulose membrane. Nitrocellulose membrane was incubated with 1:5000 dilution of streptavidin-HRP to label citrullinated proteins, which were visualised using ECL. Main protein bands are indicated with arrows 1, 2 and 3.

Anti-PPAD responses in PPAD immunised mice

DBA mice were immunised with flPPAD^{WT} or inactive flPPAD$^{C£51A}$ in CFA (Figure 4.4), sera collected and antibody responses to PPAD measured (Figure 4.5). Measured on day 30, anti-PPAD IgM and IgG responses were observed in both groups, with no significant difference between mice immunised with flPPAD^{WT} or mice immunised with flPPAD^{C351A}. This indicates PPAD itself can act as an antigen *in vivo*, independently of citrullinating activity.

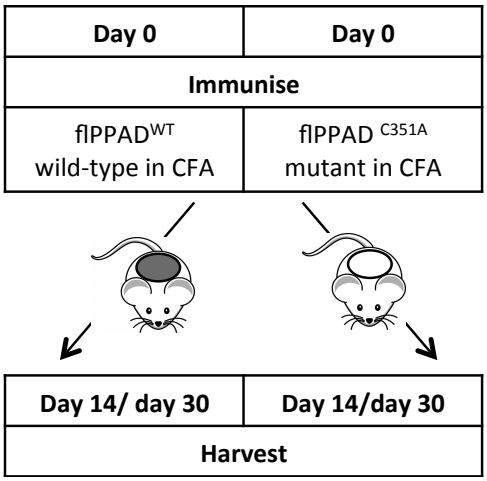


Figure 4.4 *Preliminary immunisations with wild-type or mutant PPAD.* Fifteen DBA mice were immunised with 50µg wild-type PPAD (flPPAD^{WT}) or inactive mutant PPAD (flPPAD^{C351A}) emulsified in CFA intra-dermally at the base of the tail, and monitored for 14 (7 mice) or 30 days (8 mice).

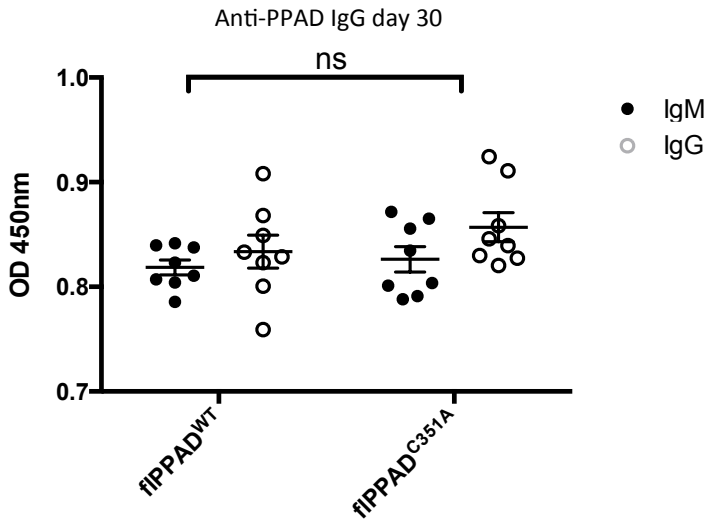


Figure 4.5 *Antibody responses to PPAD in PPAD immunised mice.* Serum was collected from DBA mice immunised with flPPAD^{WT} or flPPAD^{C351A} in CFA, 30 days post-immunisation. IgG and IgM antibody reactivity was measured to whole PPAD using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Statistical differences between groups were calculated by Mann-Whitney test in Graphpad Prism. n.s = not statistically significant.

Antibody responses to human enolase in PPAD immunised mice

Levels of IgM and IgG anti-CEP-1 and anti-REP-1 were measured on day 14 and day 30-post immunisation (Figure 4.6). Both flPPAD^{WT} and flPPAD^{C351A} immunised mice had significantly increased IgM and IgG anti-REP-1 antibodies on day 14 compared to CEP-1, and a trend towards an increased anti-REP-1 response on day 30, which was statistically significant for IgG isotype. No significant differences were observed between antibody production in flPPAD^{WT} or flPPAD^{C351A} immunised mice in any group. These findings could suggest PPAD immunisation induces an antibody response to human α-enolase, but the possibility that this response could be due to the presence of *Mycobacterium* enolase in the CFA adjuvant was investigated.

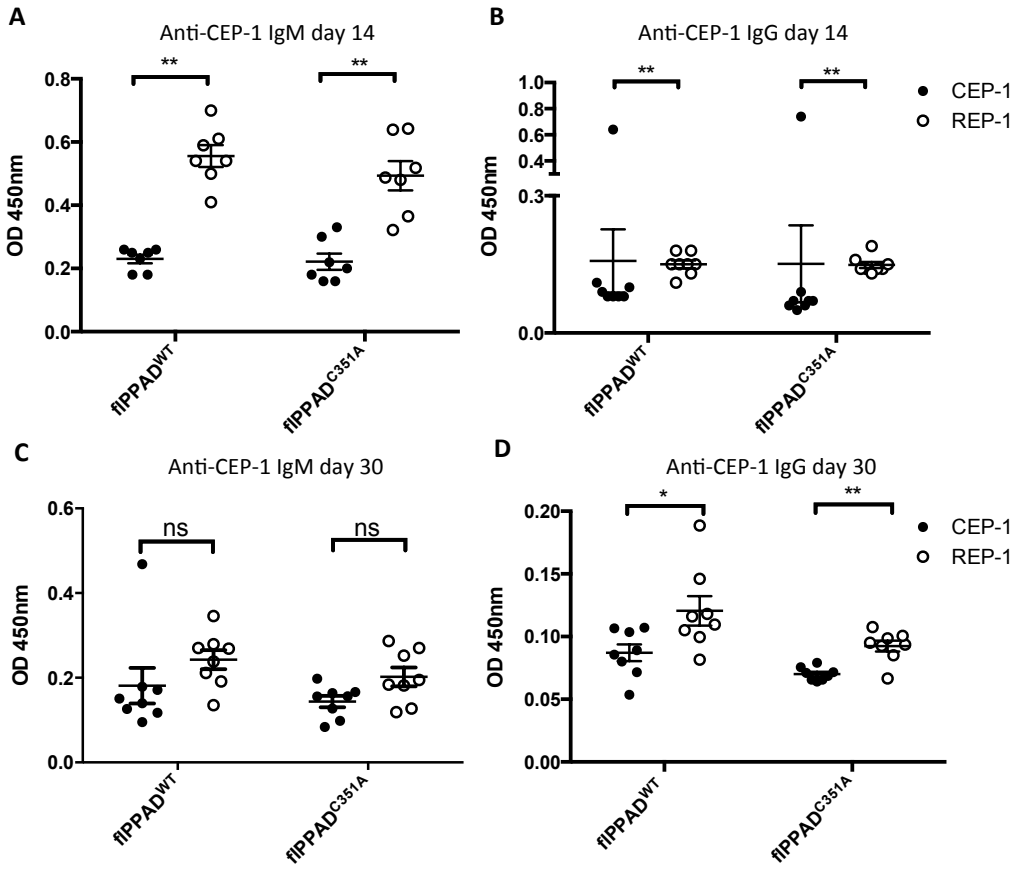


Figure 4.6 *IgM and IgG antibody responses to CEP-1 and REP-1 in PPAD immunised mice.* Serum was collected from DBA mice immunised with flPPAD^{WT} or flPPAD^{C351A} in CFA 14 or 30 days post-immunisation. IgG and IgM antibody reactivity was measured to CEP-1 and REP-1 using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents one mouse. Statistical differences between groups were calculated by Mann-Whitney test in Graphpad Prism. *=P<0.05, **=P<0.01, n.s = not statistically significant.

The effect of complete Freund's adjuvant on antibody responses

To investigate if CFA could induce anti-enolase responses due to the presence of mycobacterial enolase, two groups of mice were immunised with CFA or IFA alone, sacrificed and had serum harvested on day 30 for antibody analysis (Figure 4.7). Overall IgG levels of anti-CEP-1 and anti-REP-1 were low, however a significantly increased anti-CEP-1 response was observed in CFA immunised mice compared to IFA immunised mice without mycobacterium. Thus to avoid development of anti-enolase responses to mycobacterial enolase in CFA, the experimental design was reevaluated, and following experiments carried out without adjuvant.

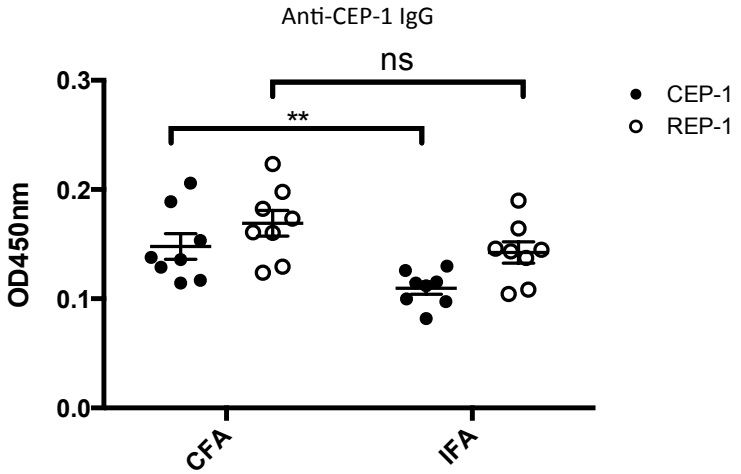


Figure 4.7 *Antibody responses to CEP-1 and REP-1 in CFA or IFA immunised mice.* Serum was collected from DBA mice immunised with CFA or IFA 30 days post-immunisation. IgG antibody reactivity was measured to CEP-1 and REP-1 peptides using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Statistical differences between groups were calculated by Mann-Whitney test in Graphpad Prism. **=P<0.01, n.s = not statistically significant.

Immunisation with PPAD, RgpB and PgEnolase

The hypothesis of this thesis was that PPAD may be involved in ACPA responses in RA, through a number of mechanisms. In addition, the homology of citrullinated peptides formed by PPAD and RgpB from PgEnolase and α-enolase proposes a mechanism for cross-reactivity to produce auto-reactive ACPA. Finally, by confirming the presence of PPAD, Rgp, and PgEnolase on the bacterial OM, all three enzymes were found to be potential targets for the host immune system during *P. gingivalis* infection. Thus, an experimental protocol was designed to investigate the effects of PPAD, RgpB and PgEnolase *in vivo*. In order to avoid potentially confounding effects of *Mycobacterium* enolase, mice were immunised without adjuvant, by IP route. In addition, $tPPAD^{WT}$ replaced $flPPAD^{WT}$ in immunisations for maximum efficacy.

Four groups of female B6 mice were immunised with PPAD, RgpB and PgEnolase, PPAD and RgpB, PPAD alone or PgEnolase alone and either sacrificed for serum and tissue harvest on day 10, or boosted with the original immunisation cocktail on day 21 and sacrificed on day 28 (Figure 4.8). Immunisations were prepared in PBS and delivered by IP. Due to the potency of these bacterial enzymes a test group of 4 mice were immunised with each of the different cocktails and sacrificed after 3 days to check for signs of peritonitis, or other internal adverse reactions, of which there were none.

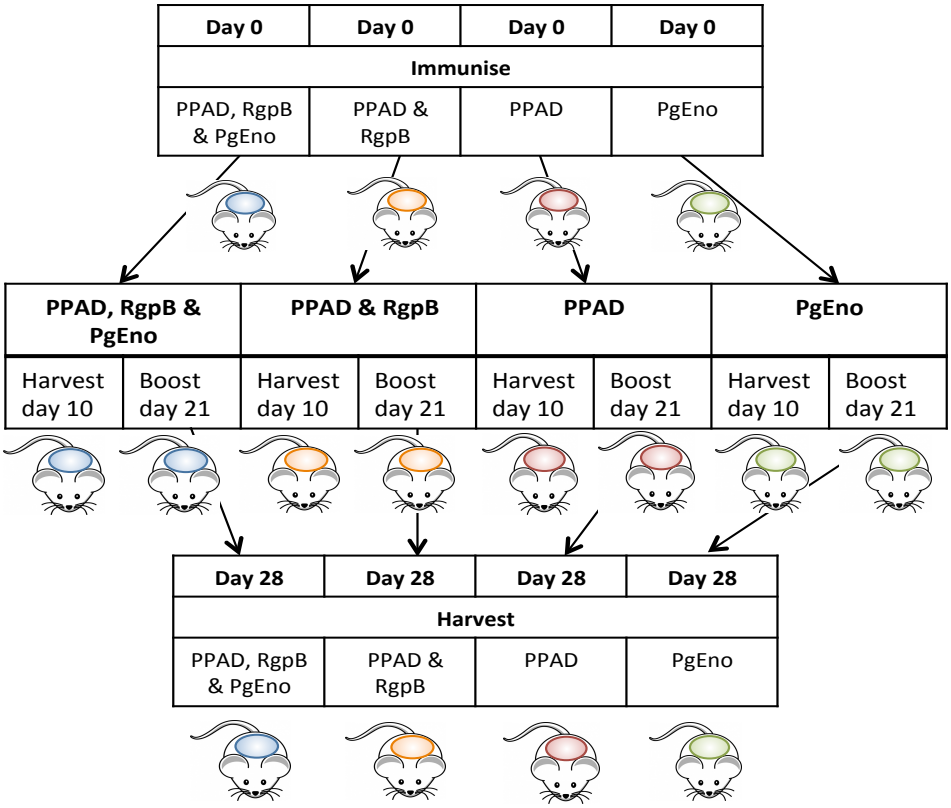


Figure 4.8 *Immunisation procedures for B6 mice with PPAD, RgpB and PgEnolase.*

Antibody responses to PPAD and PgEnolase

First IgM and IgG antibody reactivity with PPAD and PgEnolase was measured (Figure 4.9). Anti-PPAD IgG was significantly increased in all PPAD immunised groups compared to PgEnolase alone, whereas no significant differences were observed between anti-PPAD IgM in any group (figure 4.9A, B). This indicates the IgG response may be more specifically antigen driven, and IgM responses may be broader and more cross-reactive. When measuring anti-PgEnolase reactivity, responses were highest in the PPAD, RgpB, and PgEnolase group, compared to all others, and were significantly higher than the PgEnolase only group (Figure 4.9C, D). This suggests addition of PPAD and RgpB has an adjuvant affect on antibody production. For both anti-PPAD and anti-PgEnolase, there was a trend towards increased antibodies from day 10 -28.

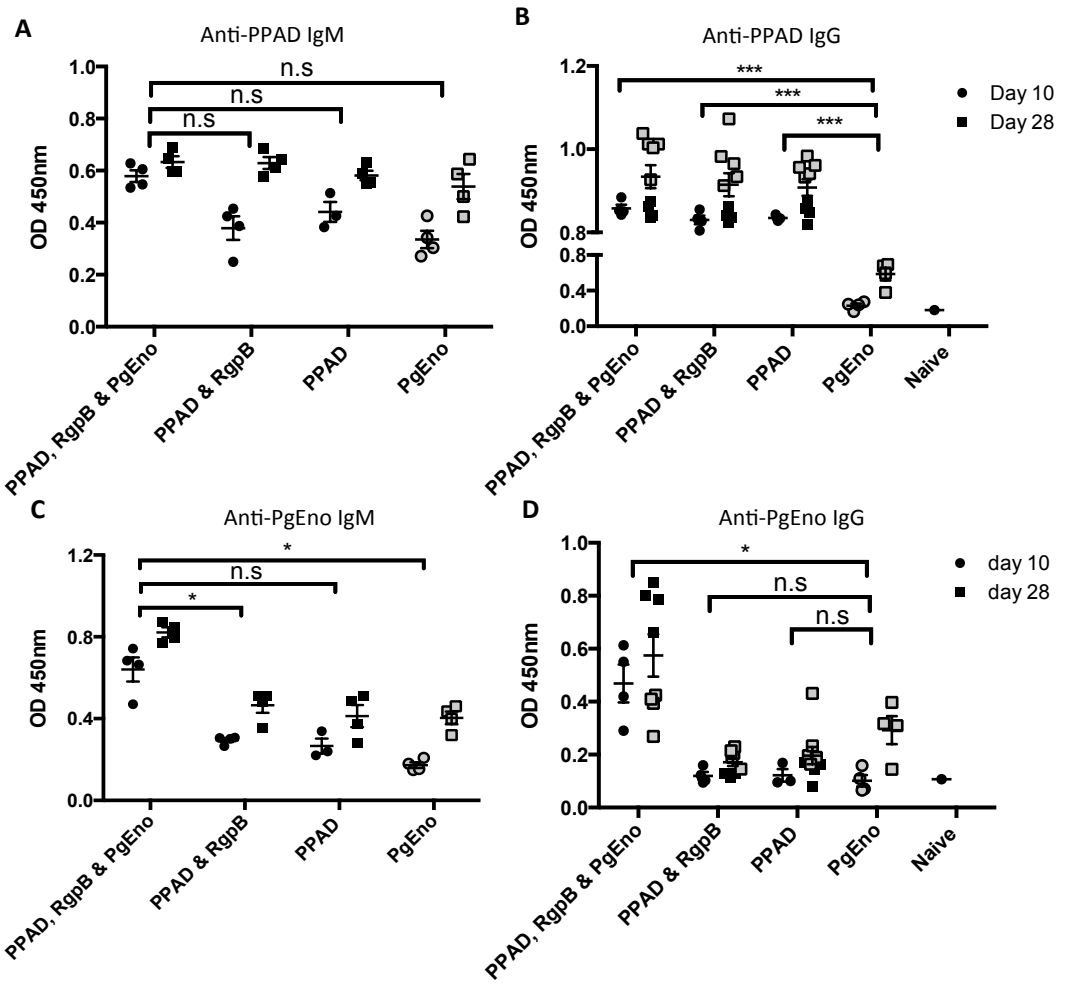


Figure 4.9 *Antibody responses to PPAD and PgEnolase.* Serum was collected from B6 mice immunised with PPAD, RgpB and PgEno, PPAD & RgpB, PPAD alone or PgEno alone, 10 days after immunisation, or 28 days post-immunisation following a booster immunisation on day 21. IgG and IgM antibody reactivity was measured to whole PPAD (A, B) and PgEnolase (PgEno) (C, D) using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Results are pooled from two separate experiments using identical protocols, represented in figures by black (experiment one) or grey (experiment two) icons. Statistical significance between groups was measured by Mann-Whitney test in GraphPad Prism 6, n.s = not statistically significant, *= P<0.05, **=P<0.01, ***=P<0.005.

Antibody responses to citrullinated PPAD peptide CPP3

To investigate if immunisation with PPAD, RgpB and PgEnolase induced antibodies to citrullinated peptides, IgG reactivity with citrullinated PPAD peptide CPP3, which is reactive with 40% of RA patient sera (Quirke *et al.*, 2014a), and arginine control peptide RPP3, was measured in sera from mice 28 days post-immunisation (Figure 4.10). Antibody response to CPP3 and RPP3 in PPAD, RgpB and PgEnolase immunised mice were significantly increased compared to all other groups. No differences were observed between anti-CPP3 and anti-RPP3 IgG antibodies.

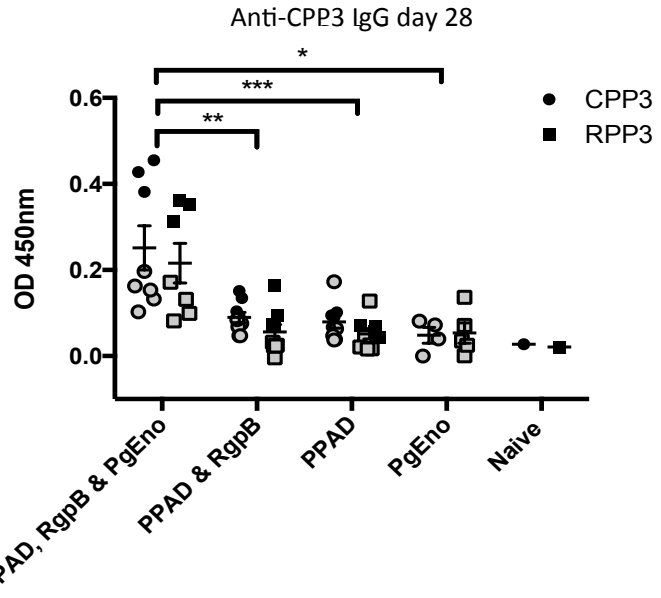


Figure 4.10 *Antibody responses to citrullinated PPAD peptide CPP3.* Serum was collected from B6 mice immunised with PPAD, RgpB and PgEno, PPAD & RgpB, PPAD alone or PgEno alone 28 days post-immunisation. IgG antibody reactivity was measured to citrullinated PPAD peptide CPP3 and the arginine control peptide RPP3 using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Results are pooled from two separate experiments, represented in figures by black (experiment one) or grey (experiment two) icons. Statistical significance between groups was measured by Mann-Whitney test, *= $P<0.05$, **=$P<0.01$, ***=$P<0.005$.

Antibody responses to citrullinated PgEnolase peptides

It has been proposed that cross reactivity between PgEnolase and α-enolase may induce anti-CEP-1 responses in RA. To investigate this, IgG reactivity to the CPgE1-6 peptides described in Chapter 3, and their arginine control peptides RPgE1-6 was measured in sera 28 days post-immunisation (Figure 4.11). Reactivity to CPgE1-6 was similar, with antibody levels in the PPAD, RgpB and PgEnolase group significantly higher than all other groups, with the exception of CPgE3 where this difference did not reach statistical significance. No differences were observed between anti-CPgE1-6 and anti-RPgE1-6 antibody responses.

Antibody responses to CEP-1 and REP-1

By day 10, levels of anti-CEP-1 and anti-REP-1 IgG were raised in the PPAD, RgpB and PgEnolase group, although this increase was only statistically significant compared to the PPAD and RgpB group (Figure 4.12). Due to a lack of available sera, only one sample was measured for anti-CEP-1, and two for anti-REP-1 in the PgEnolase only group, and so the anti-CEP-1 response cannot be accurately characterised. On day 28, anti-CEP-1 and anti-REP-1 levels were significantly higher in the PPAD, RgpB and PgEnolase mice compared to PPAD & RgpB or PPAD alone, whereas levels in PgEnolase mice were not significantly lower, but exhibited a trend towards a decreased response. In all groups anti-CEP-1 and anti-REP-1 were not significantly different, which varies from our earlier results that showed immunisation with flPPAD induced a higher anti-REP-1 response (Figure 4.6). This suggests that the changes made to the experimental design have increased the citrulline specific response.

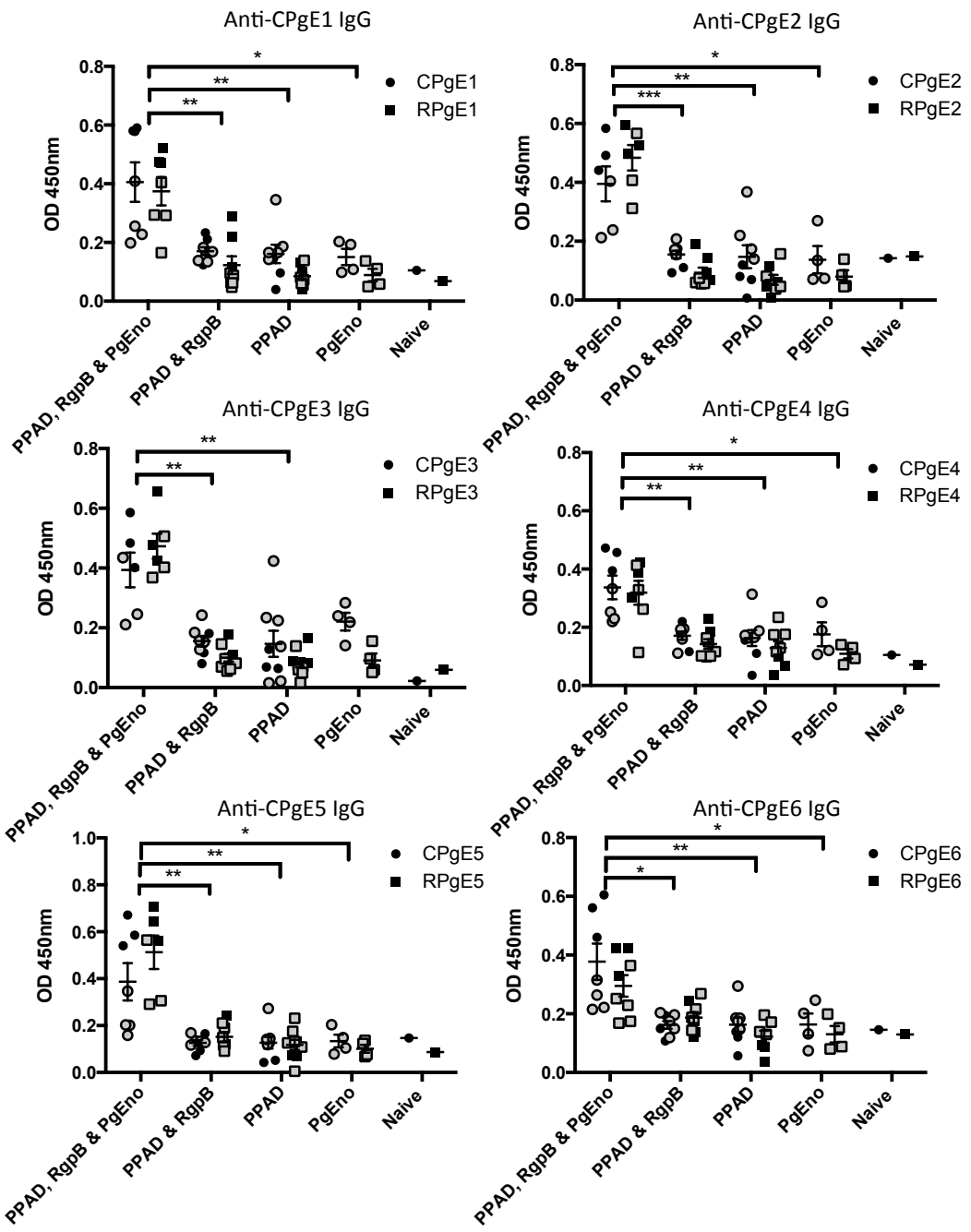


Figure 4.11 *Antibody responses to citrullinated PgEnolase peptides CPgE1-6.* Serum was collected from B6 mice immunised with PPAD, RgpB and PgEno, PPAD & RgpB, PPAD alone or PgEno alone, 28 days post-immunisation. IgG antibody reactivity was measured to CPgE1-6 and arginine control peptides RPgE1-6 using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Results are pooled from two separate experiments, represented in figures by black (experiment one) or grey (experiment two) icons. Difference between groups was measured by Mann-Whitney test, n.s = not statistically significant, *= P<0.05, **=P<0.01, ***=P<0.005.

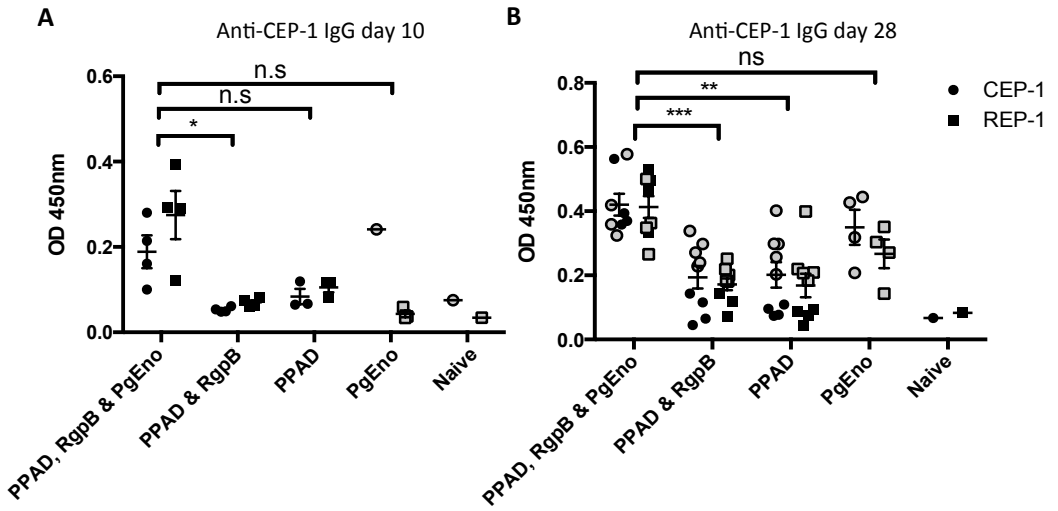


Figure 4.12 *Antibody responses to CEP-1 and REP-1.* Serum was collected from B6 mice immunised with PPAD, RgpB and PgEno, PPAD & RgpB, PPAD alone or PgEno alone 10 days after immunisation, or 28 days post-immunisation following a booster immunisation on day 21. IgG antibody reactivity was measured to whole PPAD and PgEno using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Results are pooled from two separate experiments using identical protocols, represented in figures by black (experiment one) or grey (experiment two) icons. Statistical significance between groups was measured by Mann-Whitney test, n.s = not statistically significant, *= P<0.05, **=P<0.01, ***=P<0.005.

Antibody responses to ACPA epitopes

To investigate if anti-CEP-1 responses are indicative of a general ACPA response in immunised mice, IgG ACPA were measured using a modified anti-CCP2 kit, and a known human ACPA epitope cVim alongside the arginine control peptide rVim (Figure 4.13). As for previously tested antigens, PPAD, RgpB and PgEnolase immunised mice had significantly increased levels of anti-CCP2 compared to PPAD & RgpB or PPAD immunised mice. PgEnolase immunised mice also had raised anti-CCP2 levels, although with a larger variation between individual samples. ACPA responses in all mice increased from day 10 to day 28.

Mice immunised with PPAD, RgpB and PgEnolase had higher anti-cVim and anti-rVim antibodies on day 28 compared to all other groups. Anti-cVim IgG was significantly higher than in PPAD & RgpB immunised mice, and anti-rVim was significantly higher than PPAD immunised mice. No significant difference was observed in PPAD, RgpB and PgEnolase immunised mice compared to PgEnolase immunised mice, although there was a trend towards an increase in the former. These findings suggest immunisation with PgEnolase breaks tolerance and initiates an antibody response, which is increased with the addition of PPAD and RgpB. This may be via a bacterial adjuvant affect, or through citrullination.

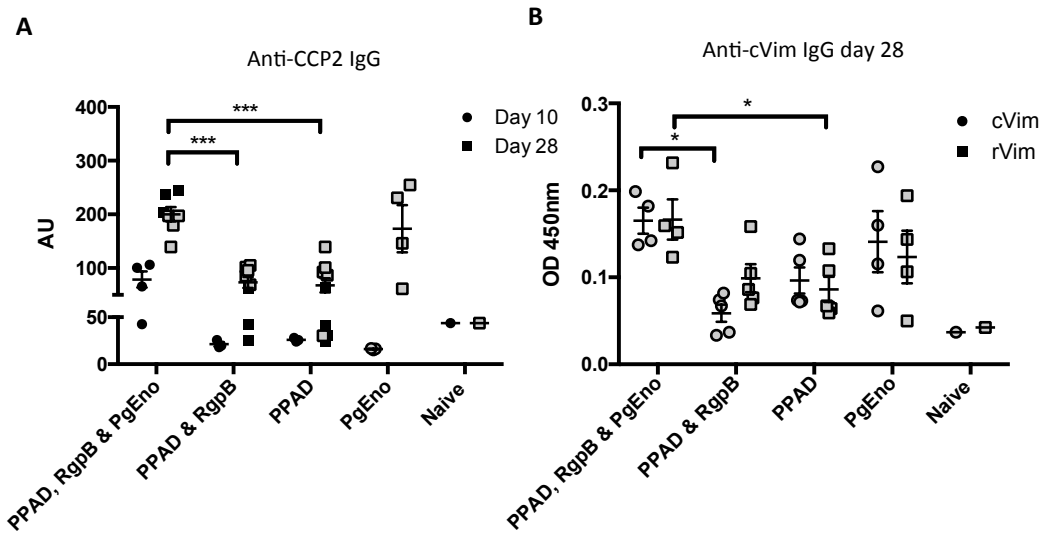


Figure 4.13 *Anti-CPP2 and anti-cVim antibody responses.* Serum was collected from B6 mice immunised with PPAD, RgpB and PgEno, PPAD & RgpB, PPAD alone or PgEno alone, on day 28 post-immunisation. A) Anti-CCP2 IgG antibody reactivity was measured using an adapted human anti-CCP2 kit. Absorbance was read at 540nm and converted to AU in Microsoft Excel using a standard curve. B) Anti-cVim and anti-rVim was measured using a standard ELISA protocol. Absorbance was read at 450nm. Values are blank corrected mean n=2 ±SE. Each dot represents reactivity from one mouse. Results are pooled from two separate experiments, represented in figures by black (experiment one) or grey (experiment two) icons. Statistical significance between groups was measured by Mann-Whitney test, *= P<0.05, ***=P<0.005

Splenocyte proliferation and cytokine release in response to PPAD

Splenocytes were isolated and stimulated with PPAD, and proliferation and cytokine release in response to stimulation was measured (Figure 4.14). When probed with PPAD, a moderate proliferative response was observed from all splenocytes (SI ~4-12), with no significant differences between immunised groups (Figure 4.14 A). Similarly, PPAD induced production of interleukin-6 (IL-6), tumour necrosis factor-α (TNF-α), and interferon-γ (IFN-γ) from splenocytes harvested from all treatment groups, with no significant differences observed (Figure 4.14 B). The lack of significant differences in proliferative or cytokine responses following PPAD stimulation suggests PPAD does not stimulate specific antigen recall responses from splenocytes, or these have been masked by non-specific stimulation from LPS contamination in the samples.

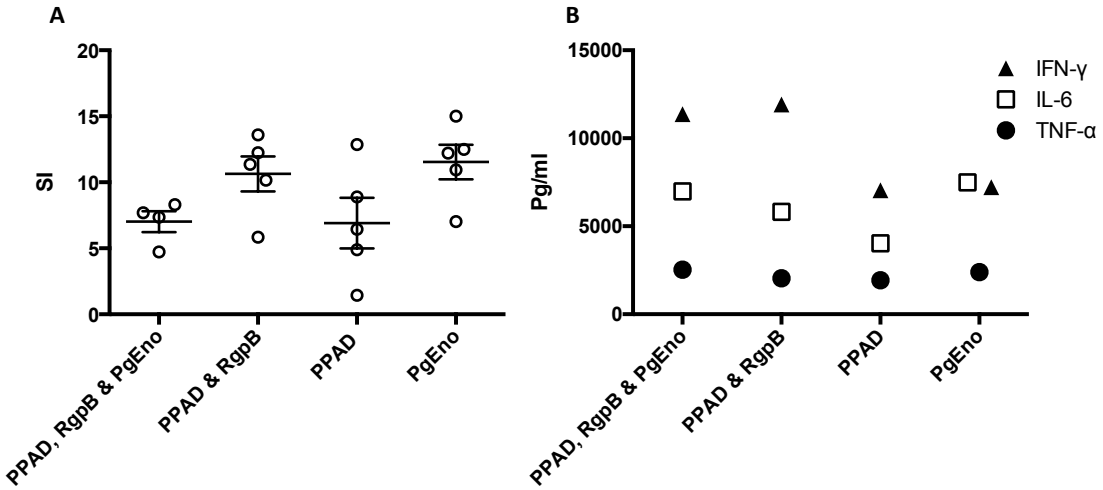


Figure 4.14 *Splenocyte proliferation and cytokine secretion in response to PPAD.* B6 mice were immunised with PPAD, RgpB & PgEno, PPAD & RgpB, PPAD alone or PgEno alone. Mice were sacrificed 28 days post-immunisation and splenocytes isolated from spleens. Splenocytes were stimulated with PPAD, and splenocyte proliferation measured by uptake of tritiated thymidine. A) Values are shown as stimulation index (SI), mean of n=3 ±SE. B) Culture supernatant was collected from stimulated cells and IL-6, TNF-α, and IFN-γ. Levels were determined using a commercial ELISA based kit. Values are blank corrected mean of n=3.

Discussion

The aim of this chapter was to investigate immunisation with PPAD, RgpB and PgEnolase in mice, to replicate *P. gingivalis* infection *in vivo*.

The hypothesis of this study was that the during *P. gingivalis* infection in PD, the host immune response may result in formation of ACPA through the following mechanisms. During PD, *P. gingivalis* and additional periodontal pathogens accumulate around the tooth surface, causing inflammation of the gingiva. In this inflammatory environment, bacterial proteins become degraded by both host proteases and bacterial gingipains, and the resulting peptides become citrullinated. Increasing damage to host tissues by *P. gingivalis* virulence factors and localised inflammatory cells creates a positive feedback loop, increasing protein degradation and citrullination. This increased citrullination, combined with an influx of APCs, increases the rate of processing and presentation of citrullinated antigens, which may be of bacterial origin, or unique C-terminally citrullinated peptides formed by PPAD. Together, these mechanisms prime an antibody response to citrullinated antigens, which may include citrullinated PgEnolase. The findings from the *in vivo* system described in this chapter support this hypothesis.

First, anti-PPAD responses were observed in all groups of mice immunised with PPAD, confirming PPAD itself is immunogenic. Anti-PgEnolase antibodies were also increased in PgEnolase immunised mice. Interestingly, the addition of PPAD and RgpB also significantly increased the anti-PgEnolase response to both whole PgEnolase, and peptides CPgE/RPgE1-6, compared to mice immunised with PgEnolase alone, suggesting local citrullination by PPAD may have increased antibody formation, potentially through presentation of more immunogenic

citrullinated PgEnolase peptides. Mice immunised with PPAD, RgpB and PgEnolase also developed significantly increased anti-CEP-1 and anti-REP-1 responses compared to other groups, providing direct evidence of *P. gingivalis* enzymes inducing antibody responses to RA ACPA epitopes, and supporting the hypothesis that PgEnolase could break tolerance to α-enolase. When measuring antibody responses to CEP-1 and CPgE1-6, and their arginine control peptides, no significant differences were observed between reactivity to the citrullinated peptide compared to the arginine control peptide. This confirms the results from a previous study from our lab by Kinloch and colleagues, who found DR4-Tg mice immunised with PgEnolase had increased anti-CEP-1 and anti-REP-1 IgG compared to naïve mice, however in their study, anti-REP-1 antibody levels were significantly higher than anti-CEP-1 (Kinloch *et al.*, 2011).

Notable in this study was the ACPA response in PPAD, RgpB and PgEnolase immunised mice, although antibodies did not exclusively react with citrullinated peptide variants. In addition to anti-CEP-1 antibodies, mice immunised with PPAD, RgpB and PgEnolase had increased anti-cVim and anti-rVim antibodies, and increased ACPA measured with a modified anti-CCP2 assay. Interestingly, this antibody response to both citrullinated and to non-citrullinated peptide variants is characteristic of PD, rather than RA. In a study of 39 PD patients and 30 healthy controls, anti-CCP2, anti-CEP-1 and anti-REP-1 were all significantly increased in PD patients ($p<0.001$) (Lappin *et al.*, 2013), and in a separate study of 96 PD patients and 98 non-PD controls, anti-CEP-1 ($p=0.02$) and anti-REP-1 ($p<0.001$), and antibodies reactive with a non-citrullinated fibrinogen peptide ($p<0.001$) were significantly increased in PD compared to controls. In addition, both anti-REP-1 and anti-fibrinogen antibody positivity in PD patients was

higher than antibodies to citrullinated variants CEP-1 (16% vs 12%) and cFib (26% vs 5%) (de Pablo *et al.*, 2014).

In addition to antibody responses, splenocyte proliferation and cytokine release was measured. Proliferative responses and secretion of TNF-α, IFN-γ and IL-6 were observed in all immunised groups when stimulated with PPAD. However the lack of significant differences in response from different treatment groups, i.e. those immunised with PPAD compared to those without, indicates the observed responses may be non-specific stimulation from LPS contamination in the PPAD sample, following purification from *E.coli*. LPS can induce innate and specific immune activation through ligation of toll-like receptors (TLRs), most commonly TLR4. Although historically considered a stimulator of innate cells, LPS contamination can also induce proliferative responses in T-cell (McAleer and Vella, 2008), and B-cell (Xu *et al.*, 2008) populations. Marginal zone B-cells of the spleen are particularly susceptible to development of T-cell independent proliferative, polyclonal responses following stimulus (Martin and Kearney, 2002, Batista and Harwood, 2009), in order to quickly eliminate blood borne antigens that enter the spleen. Therefore it is possible a proportion of the proliferative and secretory responses observed from splenocytes in this study were due to LPS stimulation, which may have masked any specific antigen recall responses. In future experiments, LPS should be removed from recall antigens by treatment with polymyxin B. Sorting of splenocytes into lymphocyte populations before stimulation would provide more understanding of the responding cell phenotype(s), and the use of splenocytes harvested from naïve mice as a negative control may provide more information on the antigen specificity of the response.

In addition, stimulation of B-cells by LPS may also have affected antibody responses reported here. LPS contamination is known to confound murine *in vivo* studies due to the high expression of TLRs on murine cells, including B-cells (Applequist *et al.*, 2002). Under inflammatory conditions in the presence of IL-2 and IL-4, human B-cells increase expression of TLRs and can therefore become activated by LPS (Mita *et al.*, 2002), and differentiate into antibody secreting plasma cells through an LPS induced increase of IL-6 (Dumont *et al.*, 2009). This may be a mechanism to sustain an efficient immune response to bacteria for clearance, however may also result in a pathogenic immune response. Thus, it is possible that the additional stimulus from *P. gingivalis* LPS during PD could be involved in the development of antibodies, including ACPA, by increasing proliferation and differentiation of B-cells.

The main findings from this chapter were:

- PPAD and PgEnolase induce robust antibody responses *in vivo.*
- Co-administration of PPAD and RgpB has an adjuvant affect *in vivo*, increasing antibody formation to citrullinated and native peptides.
- Immunisation with PgEnolase induces antibodies to peptides from PgEnolase and human α-enolase, CEP-1 and REP-1.
- Immunisation with PPAD, RgpB and PgEnolase induces anti-CCP2 and antibodies to ACPA epitopes, however the lack of citrulline specificity may be more comparable to antibody responses in PD than RA.

Chapter 5 PPAD, RgpB and PgEnolase in RA: human cohorts

Background

The detection of ACPA years before onset of joint pathology in RA patients suggests that priming of the ACPA response occurs outside the joint (Rantapaa-Dahlqvist *et al.*, 2003, van de Stadt *et al.*, 2011a, Nielen *et al.*, 2005), and the positive association between PD and periodontal pathogen *P. gingivalis* with RA has proposed the gingiva as one such potential priming site.

In support of this, citrullination, PAD2 and PAD4 are increased in inflamed gingival tissues from PD patients compared to healthy tissues from controls (Nesse *et al.*, 2012, Harvey *et al.*, 2013). Individuals with PD also have increased serum ACPA, however these are less citrulline specific than antibody responses in RA (Lappin *et al.*, 2013, de Pablo *et al.*, 2014). Further, RA patients have significantly increased antibodies reactive with *P. gingivalis* compared to non-RA controls (Mikuls *et al.*, 2009), which are positively associated with ACPA (Hitchon *et al.*, 2010, Mikuls *et al.*, 2014). Antibodies to *P. gingivalis* are also found in seropositive ($ACPA^+$ and/or RF^+) individuals considered at risk of developing RA, FDRs of RA patients or *HLA-DR4* SE positive individuals (Mikuls *et al.*, 2012, Hitchon *et al.*, 2010). However in these studies whole lysed bacteria, PgLPS and PgOMA were used as antigen in ELISA assays, which may have non-specifically bound serum ACPA due to the high levels of citrullination in both whole *P. gingivalis* and the outer membrane ((Wegner *et al.*, 2010), and findings from Chapter 4). Using recombinant antigens and peptides, studies from our lab and others have found anti-PPAD, anti-CPP3, anti-Rgp, and anti-PgCEP-1, in 25%,

40%, 32% and 34% of RA patients respectively (Quirke *et al.*, 2014a, Kharlamova *et al.*, 2016, Lundberg *et al.*, 2008). However these findings are yet to be confirmed in independent studies to establish PPAD, RgpB or PgEnolase as consistent antibody substrates in RA.

In Chapter 4 of this thesis mice immunised with PPAD, RgpB and PgEnolase developed antibodies reactive with CPP3 and PgEnolase peptides CPgE1-6, in addition to human ACPA epitopes CEP-1 and cVim, thus demonstrating that immunisation with *P. gingivalis* proteins can induce antibody responses seen in RA. However the lack of citrulline specific responses in immunised mice indicated this model is more representative of antibody responses in PD. In this chapter, antibodies to PPAD, RgpB and PgEnolase will be measured in serum from two cohorts of RA patients with or without PD, including a subset of PD^+ RA^+ patients who received periodontal therapy.

The aims of this chapter are:

- Determine anti-*P. gingivalis* reactivity in RA patients using PPAD, RgpB and PgEnolase as antigens.
- Measure anti-CEP-1 and anti-PgEnolase reactivity in RA patients as a mechanism of cross-reactivity.
- Investigate the effect of PD treatment on the levels of anti-*P. gingivalis* antibodies in RA patients with PD.

Buffers

ELISA with human sera

Carbonate coating buffer	*50mM carbonate buffer pH 9.5*
	15mM Na_2CO_3
	35mM $NaHCO_3$
	0.02% NaN_3
Blocking buffer	1x PBS
	5% BSA
Wash buffer	1x PBS
	0.05% tween
RIA buffer	1% BSA
	350mM NaCl
	10mM tris-HCl (pH 7.6)
	1% triton-X-100
	0.5% Na-deoxycholate
	0.1% SDS
Stop solution	2M H_2SO_4

Materials and methods

Patient cohorts

The Within Our Reach (WOR) cohort

Access to serum samples from the WOR cohort was provided in collaboration with Prof. T. Mikuls, University of Nebraska Medical Centre (UNMC). The work was funded by the Rheumatology Research Foundation (formerly the American College of Rheumatology Research and Education Foundation) (no specific grant number issued).

Serum samples were collected from 287 patients with RA and 330 control patients with OA recruited from the coordinating centre for the study, UNMC, and from Veterans Affairs Medical Centres from four locations: Omaha, Nebraska; Dallas, Texas; Salt Lake City, Utah; and Washington, DC. RA patients were recruited from rheumatology clinics and satisfied the 1987 ACR classification criteria for RA. OA patients were recruited from rheumatology, general medicine, and orthopaedic clinics. The study was approved by Institutional Review Boards (IRB's) at each participating site. The UNMC IRB number assigned to the study is #342-10-FB, with study name 'Periodontal Disease and *P. gingivalis* in Rheumatoid Arthritis'. Enrolment was done over an 18-20 month period, from October of 2010 to June 2012. Blood samples were collected using a standard blood draw procedure, centrifuged to isolate serum, and sent to coordinating centre for banking.

PD status was determined after enrolment by a dentist or periodontist. RA disease markers RF, and CRP were determined by nephelometry. Tender and swollen joint counts (0-28), and patient global well-being (0 to 100 mm visual

analogue scales) were taken and a DAS-28-CRP was calculated. Medical history including smoking status was also recorded (full details in (Mikuls *et al.*, 2014)). All patients were over 18 years of age and females were excluded if pregnant. Patients were also excluded if they had taken tetracycline medication in the last 6 months, had ever taken Dilantin anti-convulsant, Cylcosporine, or required antibiotic premedication before dental visits, if they had history of congenital heart disease or any prior cardiac infections or surgery, or complete dentures. RA patients were excluded following prior total joint replacement surgery.

The Outcomes of Periodontal therapy in Rheumatoid Arthritis (OPERA) cohort

Access to samples from the OPERA cohort was provided in collaboration with Dr T. Dietrich and Dr M. Grant at the University of Birmingham. The project was funded by the National Institute for Health Research (NIHR) (no specific grant number available). Serum samples were collected from 42 PD^+ RA^+ patients, 48 RA patients, and 17 healthy individuals. PD^+ RA^+ patients were split into two groups to receive PD treatment (n=20), or not (n=22) and had serum samples taken again after 6 months. All RA cases fulfilled the 1987 ACR classification criteria, and had DAS28 ≥3.2. PD status was assessed by an experienced dentist. Patients were considered PD^+ if clinical attachment loss (CAL) was ≥4mm on ≥2 non-adjacent teeth, and they had a cumulative probing depth of ≥40mm. In this study and all others compared, periodontal treatment is defined as non-surgical scaling and root surface planning (Heitz-Mayfield and Lang, 2013). This was completed in 2-3 sessions within 2 weeks of enrolment, with additional root planing, supragingival scaling and prophylaxis performed at all sites with ≥4 mm probing depth and bleeding on probing, or sites with ≥5 mm probing depth, at three and six months. Non-treatment group received standard oral hygiene

instructions only. All patients were over 18 years old. Patients were excluded if they had any rheumatic autoimmune disease other than RA (vasculitis, pulmonary fibrosis or Felty's syndrome), history of inflammatory joint disease other than RA (gout, reactive arthritis, seronegative spondyloarthropathy), any other systemic autoimmune disorder (systemic lupus erythematosus, inflammatory bowel disease, scleroderma, inflammatory myopathy, mixed connective tissue disease or any overlap syndrome), or prior diagnosis of juvenile idiopathic arthritis, or juvenile rheumatoid arthritis and/or RA before age 16. Patients were also excluded if intra-articular or parenteral glucocorticoids were taken within 4 weeks prior to study baseline, or if they had any surgical procedure, including bone/joint surgery/synovectomy (including joint fusion or replacement) or periodontal treatment (surgical or non-surgical, excluding supragingival cleanings) within 12 weeks prior to study baseline. Patients may also have been excluded at the investigators discretion due to significant concomitant disease or dental condition that may preclude participation in the trial.

ELISA with human sera

In the WOR cohort, ELISA assays with CEP-1, CPgE peptides, and their relevant controls were carried out by me. ELISA assays with cFibβ, cVim, cTNC5, RgpB, and CPP3 were carried out by Dr A. Schwenzer and Dr A. Quirke in the Venables lab as part of separate studies. Anti-CCP2 ELISA assay was carried out in the lab of Prof. T. Mikuls. All results were donated with permissions for use in this thesis, and all analysis of results, and graphical representations were made by me.

ELISAs were carried out on all 617 patients and grouped either by RA (n=287) or OA status (n=330), PD status (PD^+ n=287, PD^- n=430) or both (RA PD^+ n=100, RA PD^- n=187, OA PD^+ n=87, OA PD^- n=243).

In the OPERA cohort, all ELISA assays, analysis of results, and graphical representations were carried out or made by me. ELISAs were carried out on all groups as described. PD^+ RA^+ patients were tested together (n=42) or split into treatment (n=20) and non treatment (n=22) groups.

Citrullinated peptide antigens from α-enolase (CEP-1) ^{5}CKIHA-Cit-EIFDS-Cit-GNPTVEC^{21}, fibrinogen β-chain (cFibβ) ^{36}NEEGFFSA-Cit-GHRPLDKK^{52}, vimentin (cVim) ^{59}VYAT-Cit-SSAV-Cit-L-Cit-SSVP^{74}, tenascin-C (cTNC5) ^{2183}EHSIQFAEMKL-Cit-PSNF-Cit-NLEG-Cit-Cit-KR^{2200}, PgEnolase (CPgE1-6) ^{1}MEIAKIIG-Cit^{9}, ^{6}IIG-Cit-EILDS-Cit^{15}, ^{41}TGENEAIEL-Cit^{51}, ^{168}PVGACCF-Cit^{175}, ^{317}LFVTNVEFL-Cit^{326}, ^{401}MAKYNQLL-Cit^{409}, and PPAD (CPP3) ^{71}CAKTDSYWT-Cit-DYTGWFAMYDC^{89}, and their corresponding arginine controls were diluted to 10mg/ml, and whole proteins PPAD, RgpB and PgEnolase were diluted to 5mg/ml in carbonate coating buffer. 100µl of antigen was added per well to a 96-well plate (Costar) and plates incubated o/n at 4°C. Plates were washed x3 with wash buffer (see 'Buffers') and tapped dry. Plates were blocked with 200µl/well blocking buffer (see 'Buffers') for 2h at rt. Human sera were diluted 1:100 in RIA buffer (see 'Buffers') and 100µl/well added in duplicate with a standard curve (if available). Two wells with buffer only were used as blank controls. Plates were incubated for 1.5h at rt and washed as before. 200µl/well secondary antibody (Human IgGFc Monoclonal antibody (HRP), Stratech), diluted 1:3000 in RIA buffer was added and incubated for 1h at rt. Plates were washed as before and 100µl/well of TMB (BioLegend) added. The reaction was

stopped with 50µl/well stopping solution (see 'Buffers') after 5min, and absorbance read at 450nm. Data was analysed in Microsoft Excel. Values are mean n=2 –blank. Statistical differences between groups were calculated using Mann-Whitney test in GraphPad Prism 6.

Antibody positivity was determined using a 98% cut-off set by the control group (OA patients in the WOR cohort, healthy individuals in the OPERA cohort). Difference between groups was determined significant if $p \leq 0.05$, calculated by Mann-Whitney testing in GraphPad Prism.

Statistical analysis of patient demographics

Statistical analysis of anti- CEP-1 positivity was carried out by Dr H. Sayles, using data from the ELISA assays carried out by me. Data was donated with full permission for use in this thesis. Statistical analysis of differences between patient demographics in the OPERA cohort was carried out by me. In both cohorts Chi-square tests were used to compare frequency distributions of categorical variables, while t-tests and Wilcoxon rank-sum tests were used to compare all continuous variables. Correlations between antibody levels were calculated by Spearman's rank coefficient in GraphPad Prism 6.

All statistical differences were considered significant if $p \leq 0.05$.

Results

WOR cohort patient demographics

The demographics of individuals in the RA and OA groups of the WOR cohort are shown in Table 5.1. The mean age of participants was 59 years in both groups and 63% and 60% of RA and OA groups respectively were male. This is a higher proportion of males than commonly seen in RA cohorts, which could be due to their recruitment from Veterans Affairs Medical Centres. Seventy-seven per cent of RA patients and 73% of OA controls were Caucasian, 17% and 23% African-American and 6% or 5% were other ethnicities respectively. The RA group had a reduced percentage of never smokers compared to the OA group (38% and 54%) and had increased percentages of former and current smokers (43% and 19% compared to 35% and 11%), which may be because smoking is a risk factor for RA. Frequency of *HLA-DRB1* SE alleles was significantly increased in the RA group (76% compared to 45%, $p<0.001$), which is as expected as an associated genetic risk factor for RA. BMI was significantly lower in RA patients (30) than OA controls (32) ($p=0.001$). This may be as obesity is an associated risk factor for OA (Finckh and Turesson, 2014), and only confers significant risk to developing RA in young women (Heidari, 2011).

Table 5.1 *WOR cohort patient demographics* adapted from (Mikuls *et al.*, 2014).

	Total N=617	**RA N=287**	**Controls N=330**	
Age	59	59	59	
Male gender	62%	63%	60%	
Race				
Caucasian	75%	77%	73%	
African American	20%	17%	23%	
Other	5%	6%	5%	
BMI, kg/m2^2	31	30	32	p=0.001
Smoking status				
Never	46%	38%	54%	p<0.001
Former	39%	43%	35%	
Current	15%	19%	11%	
HLA-DRB1 SE positive	59%	76%	45%	p<0.001

Antibody reactivity with CEP-1 and REP-1 in RA and OA patients

Using a 98% cut-off for positivity set by the OA group, 33% of RA patients were positive for anti-CEP-1 antibodies (98% cut-off 124 AU), and <1% were positive for anti-REP-1 (98% cut off 142 AU) (Figure 5.1). The level of anti-CEP-1 was significantly increased in RA patients compared to OA ($p<0.0001$), but no significant different was observed for anti-REP-1 levels. In RA patients, no correlation was observed between anti-CEP-1 and anti-REP-1 antibodies ($r=0.092$, $p=0.120$), with levels of anti-CEP-1 antibodies higher than anti-REP-1 antibodies, indicating increased production of citrulline specific response in RA patients. Conversely, in OA patients, a strong correlation was observed between anti-CEP-1 and anti-REP-1 antibodies ($r = 0.5195$, $p<0.0001$), which may be representative of a less citrulline specific anti-enolase response.

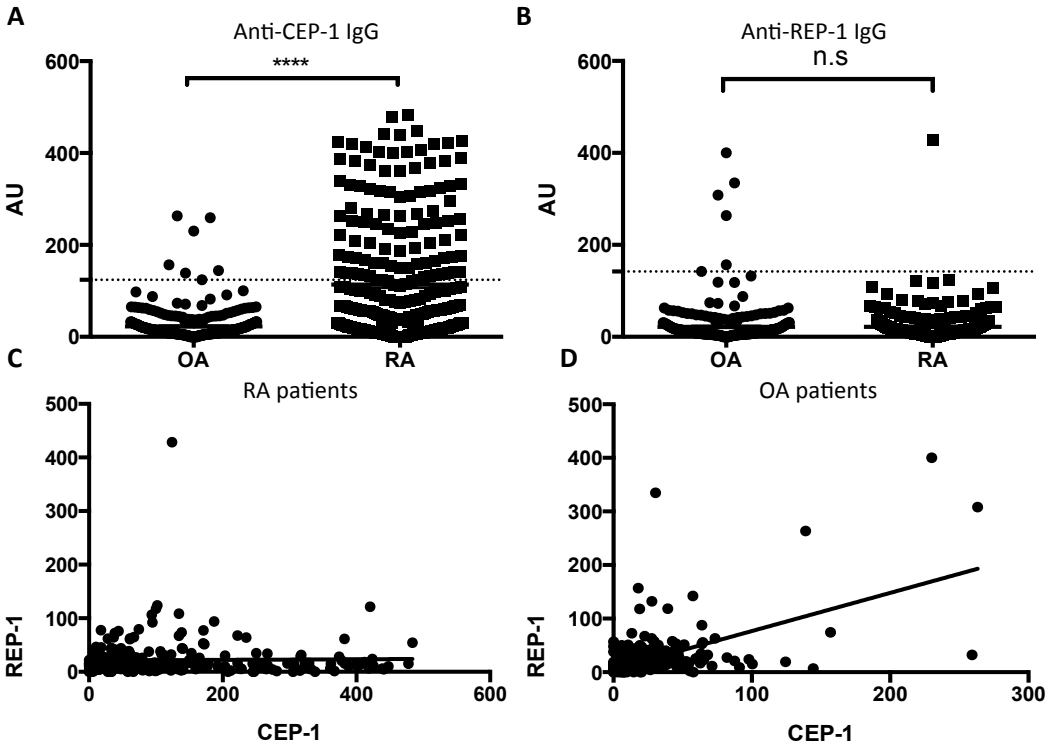


Figure 5.1 *Antibody responses to CEP-1 and REP-1 in RA and OA individuals.* IgG antibody responses to CEP-1 (A) and REP-1 (B) in the WOR cohort of 287 RA and 330 OA sera were measured by ELISA. Absorbance was measured at OD450nm and converted to AU using a standard curve in Microsoft Excel. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Antibody positivity was determined using a 98% cut-off in the OA control group shown with a dashed line. Statistical difference between groups was calculated by Mann-Whitney testing in GraphPad prism, n.s. = not statistically significant, **** = P< 0.0001. Correlation between anti-CEP-1 and anti-REP-1 antibodies were calculated in RA (C) and OA (D) patients using Spearman's rank coefficient in GraphPad Prism 6.

Clinical and aetiological associations with anti-CEP-1 positivity

Anti-CEP-1 positivity was analysed in relation to patient demographics, and measures of RA and PD disease (Table 5.2). Positivity for anti-CEP-1 correlated positively with male gender, smoking, and expression of the *HLA-DRB1* SE allele, which could be due to the higher occurrence of these factors in the RA cohort. Anti-CEP-1 was also associated with hypertension, which could be due to the increased prevalence of smokers, and males, who are more susceptible (Halperin *et al.*, 2008) in this group. Anti-CEP-1 did not correlate with clinical or radiological measures of RA disease, but was significantly associated with bone turnover markers cathepsin K (CTPK) and Bone Alkaline Phosphatase (BAP), and with having >20% alveolar bone loss in PD, which is consistent with previous findings that $ACPA^+$ RA is associated with increased bone erosion (Syversen *et al.*, 2008).

Table 5.2 *Patient demographics and associations with CEP-1 positivity*. Table overleaf. Statistical correlations between CEP-1 positivity, patient demographics, and disease markers were calculated by H. Sayles. Chi-square tests were used to compare categorical variables. T-tests and Wilcoxon rank-sum tests (marked with an "a") were used to compare all continuous variables. Associations were considered statistically significant if $p \leq 0.05$.

Table 5.2 *Patient demographics and associations with CEP-1 positivity*. Full legend on previous page.

Characteristic	CEP-1+ (N=95)	CEP-1- (N=192)	P-value
Sociodemographics			
Age, years	60 (12)	58 (12)	0.362
Male gender	79	56	<0.001
Race			
Caucasian	81	76	0.413
African- American	13	19	
Other	6	5	
Comorbidities			
Body mass index, kg/m^2	30 (6)	30 (7)	0.516
Smoking status‡			
Never	27	43	0.034
Former	48	41	
Current	24	17	
Diabetes mellitus	22	16	0.217
Hypertension	57	40	0.006
Cardiovascular disease	15	12	0.512
Osteoporosis	7	13	0.188
Genetic Risk			
HLA-DRB1 SE status	88	69	0.001
Rheumatoid Arthritis Characteristics			
Disease duration (years)	14 (10)	12 (10)	0.246
Swollen joints	3.6 (4.9)	3.6 (4.0)	0.235[a]
Tender joints	3.4 (4.7)	3.0 (4.6)	0.341[a]
Hs-CRP (mg/l)	9 (16)	8 (23)	0.116[a]
Patient global (0-10)	4.4 (2.9)	4.1 (2.5)	0.51
DAS-28-CRP	4.1 (1.5)	3.9 (1.3)	0.32
Sharp score	17 (19)	20 (25)	0.898[a]
Joint space narrowing	14 (14)	15 (18)	0.822[a]
Erosion score	3 (6)	5 (9)	0.590[a]
Anti-CCP pos.	99	77	<0.001
Anti-CCP (U/ml)	247 (106)	95 (106)	<0.001
RF positivity	92	69	<0.001
RF (U/ml)	274 (291)	247 (566)	<0.001[a]
Periodontitis Characteristics			
PD (Eke criteria)			
None/mild	13	21	0.184
Moderate	51	49	
Severe	37	30	
Anti-PgOMV	122 (98)	98 (84)	0.022[a]
P. gingivalis PCR positive	67	59	0.21
Mean sites/patient with PD > 5mm	3.7 (6.1)	2.0 (3.7)	0.160[a]
Missing teeth	3.2 (3.0)	3.1 (3.2)	0.624[a]
Mean alveolar bone loss	9.7 (8.3)	7.8 (6.7)	0.083[a]
Sites > 20% bone loss (mean %)	13.1 (19.7)	9.3 (18.9)	0.029[a]
Bone Turnover Markers			
TRAP5b	1.7 (1.1)	1.9 (1.1)	0.056[a]
BAP	23.6 (8.4)	22.5 (6.4)	0.499[a]
CTPK	7.0 (12.1)	12.3 (52.3)	0.011[a]

Antibody reactivity to CPgE1-6 in WOR cohort

Antibody reactivity to CPgE1-6 and arginine controls RPgE1-6 was measured in 25 OA and 54 RA serum samples from the WOR cohort by ELISA (Figure 5.2). Using a 98% cut-off for positivity from the OA controls, positivity for CPgE peptides in RA sera ranged from 9% for CPgE4 and CPgE6, to 41% for CPgE2. Similarly, reactivity with the arginine control peptides ranged from 9% to 39%. Interestingly, level of positivity with the arginine control peptides was not always proportionate to positivity for the citrullinated peptide. RPgE3, RPgE4 and RPgE6 all showed higher levels of reactivity than their citrullinated counterparts, suggesting citrullination may have rendered these peptides less immunogenic. CPgE2 had the highest level of positivity, and lowest reactivity with arginine control RPgE2, suggesting a level of citrulline specificity.

Positivity for ACPA in the WOR cohort

To establish the reactivity with other ACPA epitopes in this cohort, ELISA assays were carried out using the anti-CCP2 assay and citrullinated peptides cFibβ, cVim, cTNC-5 (Table 5.3). Using a 98% cut-off as before, 66.2%, 32.1%, and 50.5% of RA patients were positive for cFibβ, cVim and cTNC5 respectively. Using the positive limits defined by the anti-CCP2 kit manufacturers, 84.5% of RA patients were positive.

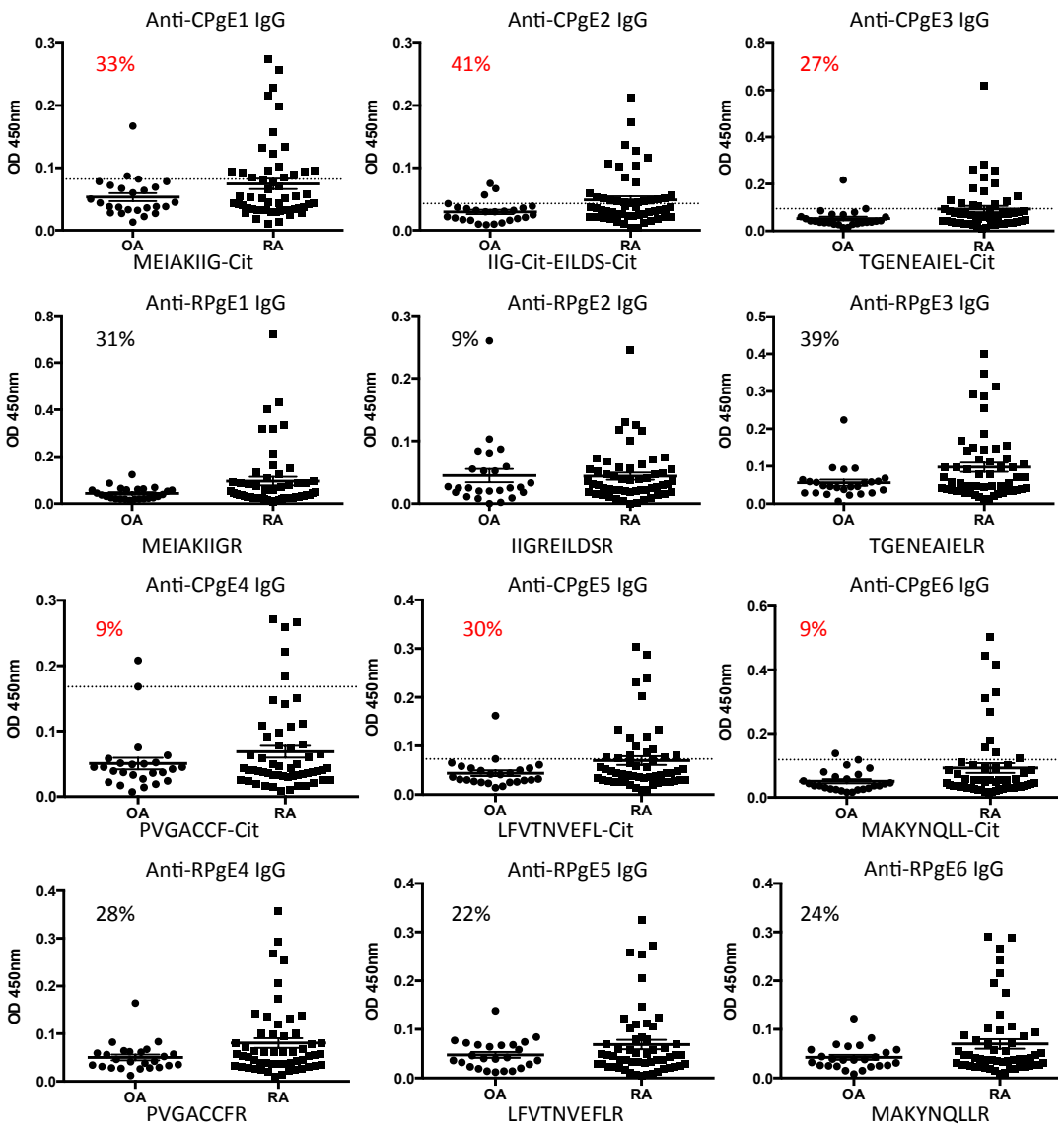


Figure 5.2 *Antibody responses to CPgE1-6 and RPgE1-6 peptides in RA and OA patients.* IgG antibody reactivity to CPgE1-6 and RPgE1-6 was measured in a sample of 25 OA and 54 RA patients randomly selected from the WOR cohort using a standard ELISA protocol. Absorbance was measured at OD450nm. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Antibody positivity was determined by 98% cut-off in the OA control group, which is shown with a dashed line on CPgE1-6 graphs. Antibody positivity is shown as percentage in individual panes.

Table 5.3 *Antibody positivity for ACPA in RA patients in WOR cohort.*

Peptide	Sequence	% positivity
CCP2	Unknown	84.5
cFib β_{36-52}	NEEGFFSA-Cit-GHRPLDKK	66.2
$cVim_{59-74}$	VYAT-Cit-SSAV-Cit-L-Cit-SSVP	32.1
$cTNC5_{2183-2200}$	EHSIQFAEMKL-Cit-PSNF-Cit-NLEG-Cit-Cit-KR	50.5

IgG antibody positivity for citrullinated peptides from fibrinogen β-chain (cFibβ), vimentin (cVim), and tenascin-C (cTNC5) were measured by ELISA in the WOR cohort of 287 RA and 330 OA patients by Dr A. Quirke and Dr A. Schwenzer. Anti-CCP2 was calculated using the anti-CCP2 assay, and positivity determined using manufacturers recommended cut-off.

Relationship between antibodies to CEP-1 and CPgE1-6 peptides

Positive linear correlations were observed between anti-CEP-1 and antibodies to all CPgE peptides, with r-values ranging from 0.443-0.587 (Figure 5.3). All correlations were statistically significant ($p<0.0001$).

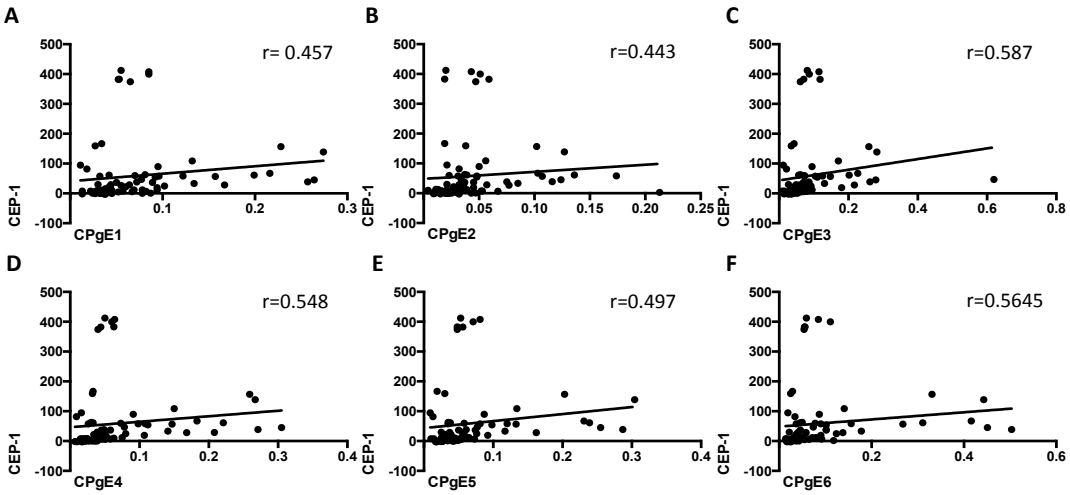


Figure 5.3 *Correlations between CEP-1 and CPgE1-6.* IgG antibody reactivity to CPgE1-6 and CEP-1 was measured in 25 OA and 54 RA patients using a standard ELISA protocol. Absorbance was measured at OD450nm. Values are blank corrected mean n=2 ±SE. Correlation 'r' was calculated using Spearman's rank coefficient in GraphPad Prism.

Anti-PPAD and anti-RgpB in WOR cohort

Antibody reactivity to the citrullinated PPAD peptide CPP3, and whole RgpB was measured in RA and OA patients (Figure 5.4). Using the 98% cut-off (252 AU), 5% of RA patients were positive for anti-CPP3 antibodies, and 1% of RA patients were positive for anti-RgpB antibodies (98% cut-off OD 1.77). These antibody levels are lower than previous findings of 40% positivity for anti-CPP3 (Quirke *et al.*, 2014a), and 23% positivity for anti-Rgp in RA patients (Kharlamova *et al.*, 2016).

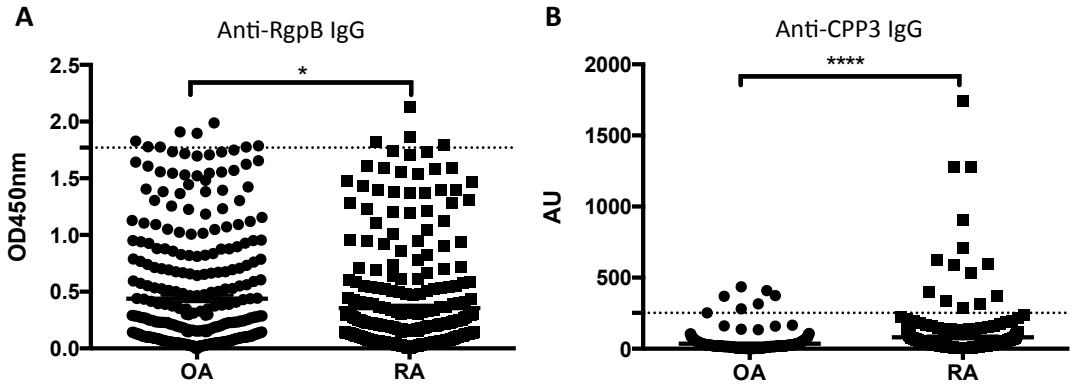


Figure 5.4 *Antibody positivity to CPP3 and RgpB.* IgG antibody reactivity to A) RgpB and B) CPP3 was measured in the WOR cohort of 287 RA and 330 OA patients using a standard ELISA protocol. In all ELISAs, absorbance was measured at OD450nm and converted to AU where available using a standard curve in Microsoft Excel. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Antibody positivity in RA patients was calculated based on a 98% cut-off determined in the OA group and is represented with a dashed line. Statistical difference between groups was calculated by Mann-Whitney testing in GraphPad prism, * = P< 0.05, **** = P< 0.001.

Effect of periodontitis on antibody positivity in RA and OA patients

Given the association of *P. gingivalis* and PD, we investigated the influence of PD status on antibody positivity. First, the CPgE1-6 ELISA results from 79 (25 OA and 54 RA) individuals were recombined, and split into groups based on presence (57 PD^+) or absence (23 PD^-) of PD (Figure 5.5). A slight trend was observed towards increased antibody reactivity in the PD^+ group, which this was statistically significant in the CPgE2 group ($p<0.05$). However prevalence of RA was significantly higher in PD^+ compared to non PD^- patients ($p=0.0221$), which may have had a confounding affect on results.

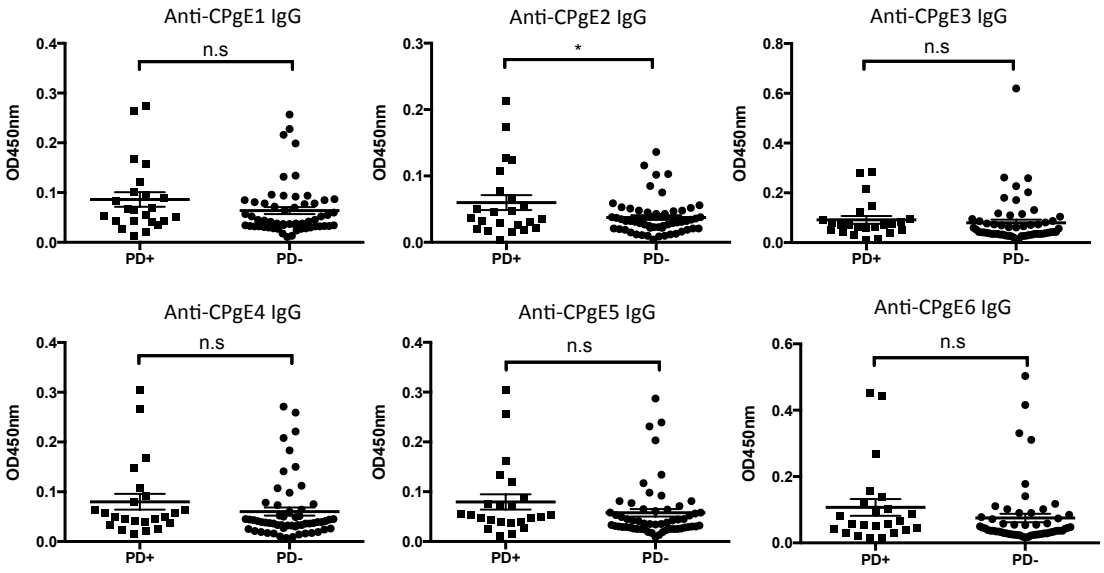


Figure 5.5 *Anti-CPgE1-6 in PD^+ and PD^- individuals.* IgG antibody reactivity to CPgE1-6 was measured in 57 PD negative (PD^-) and 23 PD positive (PD^+) patients from the WOR cohort using a standard ELISA protocol. Absorbance was read at OD450nm. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Statistical differences between groups was calculated by Mann-Whitney test in GraphPad Prism, n.s = no statistically significant difference, * = $P< 0.05$.

Anti-RgpB, anti-CPP3 and anti-CEP-1 antibody responses in the whole WOR cohort were then investigated in relation to PD status. Anti-RgpB was significantly increased in 187 PD^+ compared to 430 PD^- patient sera ($p<0.001$), whereas anti-CPP3 was not statistically significant ($p=0.06$) (Figure 4.6 A, B). To determine the effect of RA status on these findings, the WOR cohort was further divided based on both RA or OA, and PD status to create four groups, OA PD^- (n=243), RA PD^- (n=187), OA PD^+ (n=87) and RA PD^+ (n=87). Within these groups, 9% of RA PD^+, and 8% of OA PD^+ individuals were anti-RgpB positive, which were both significantly increased compared to the corresponding non-PD groups RA PD^- and OA PD^- ($p=0.0061$ and $p=0.0004$ respectively) (Figure 5.6 C). Conversely, anti-CPP3 antibodies were significantly increased in both RA groups compared to OA groups, with or without PD (Figure 5.6 D). RA PD^+ had the highest level of positivity (8.2%) and was significantly higher than OA PD^+ and OA PD^- groups ($p=0.0001$, $p=<0.0001$ respectively). RA PD^- antibody levels were also significantly raised compared to OA PD^- ($p<0.0001$).

Anti-CEP-1 was significantly increased in PD^+ compared to PD^- ($p<0.05$), whereas no significant difference was found between anti-REP-1 levels (Figure 5.7). When further divided as before, anti-CEP-1 antibody level was highest in the RA PD^+ group (55.7% positive), which was significantly higher than both RA PD^- ($p<0.05$) and OA PD^+ ($p<0.0001$) groups. In addition, antibody levels in the RA PD^- group were significantly higher than the OA PD^- group ($p<0.0001$). No significant differences were observed between levels of anti-REP-1 in any group.

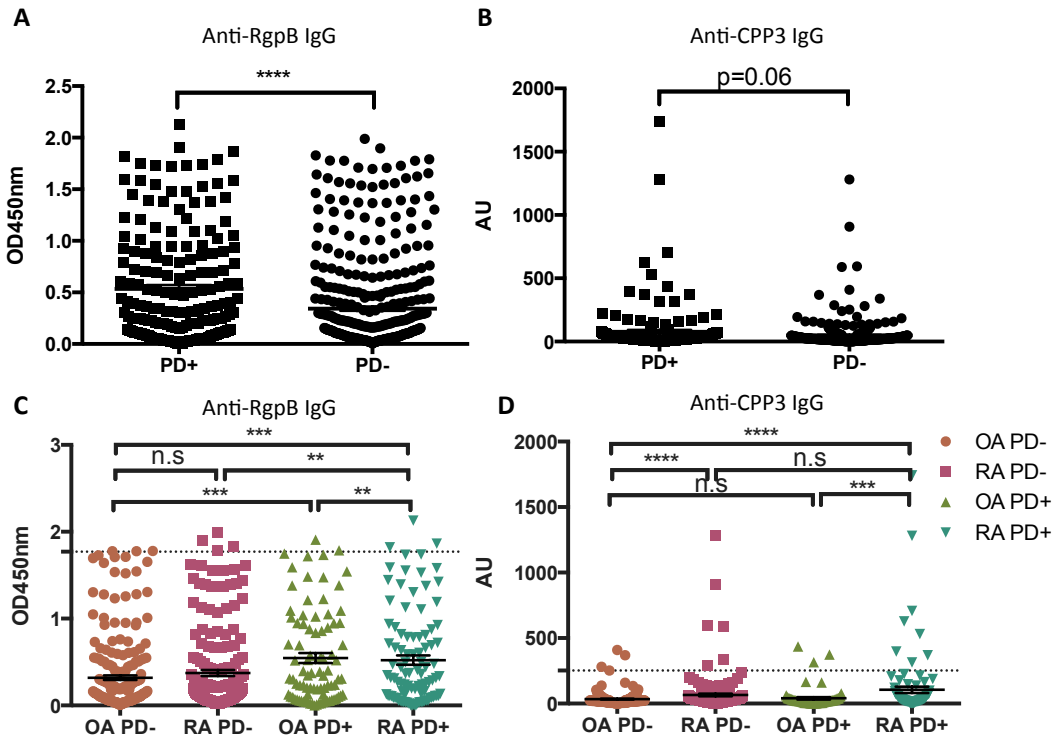


Figure 5.6 *Anti-RgpB and anti-CPP3 in PD⁺ and PD⁻ individuals.* IgG antibody reactivity to RgpB (A) and CPP3 (B) was measured in the WOR cohort using a standard ELISA protocol. Cohort was divided based on PD status: PD⁺ (187) or PD⁻ (430), or both RA and PD status: OA PD⁻ (194), RA PD⁻ (143), OA PD⁺ (73) and RA PD⁺ (70). Absorbance was read at OD450nm and converted to AU in Microsoft Excel using a standard curve. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Statistical differences between groups was calculated by Mann-Whitney test in Microsoft Excel, n.s = not statistically significant, * = P< 0.05, ** = P<0.005, *** = P<0.001 ****=P<0.0001.

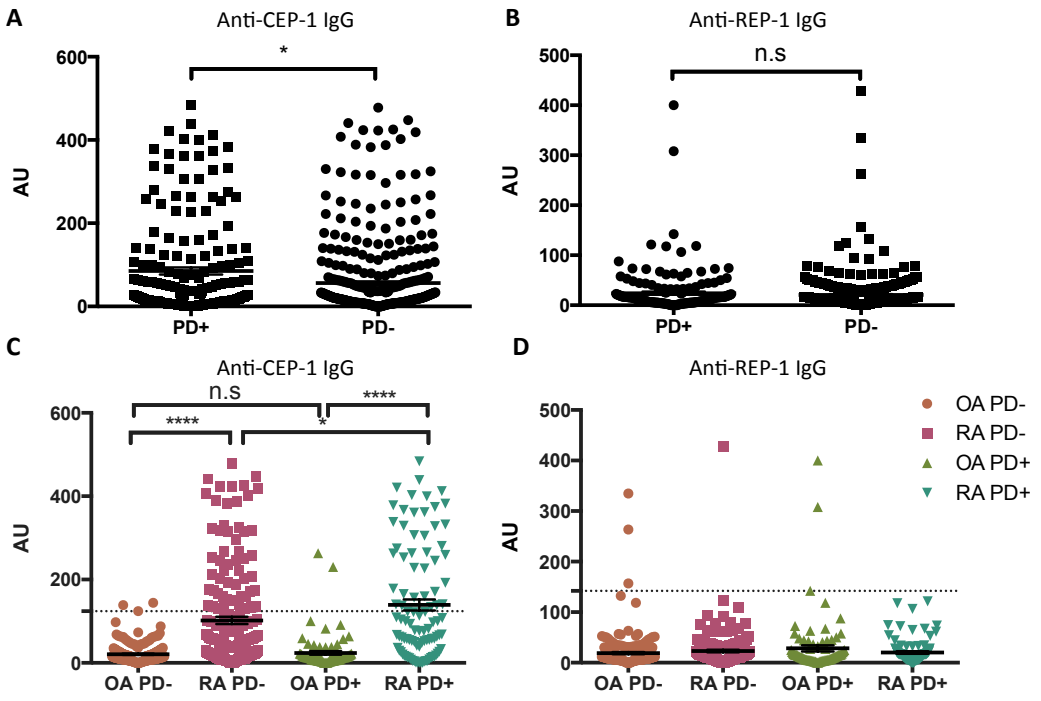


Figure 5.7 *Anti-CEP-1 and anti-REP-1 in PD⁺ PD⁻ individuals.* IgG antibody reactivity to CEP-1 (A) and REP-1 (B) was measured in the WOR cohort of 617 individuals using a standard ELISA protocol and grouped into PD⁺ (187) or PD⁻ (430). Cohort was further divided by RA and PD status into OA PD⁻ (194), RA PD⁻ (143), OA PD⁺ (73) and RA PD⁺ (70) and reactivity to CEP-1 (C) and REP-1 (D) analysed. In all ELISAs, absorbance was measured at OD450nm and converted to AU in Microsoft Excel using a standard curve. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Statistical differences between groups was calculated by Mann-Whitney test in Microsoft Excel, n.s = not statistically significant, * = P< 0.05, ****=P<0.0001.

Finally, anti-CCP2 positivity in RA patients was significantly increased in patients with PD compared to those without ($p<0.05$) (Figure 5.8). Anti-CCP2 was measured in RA patient sera only, and so further subdivisions could not be carried out.

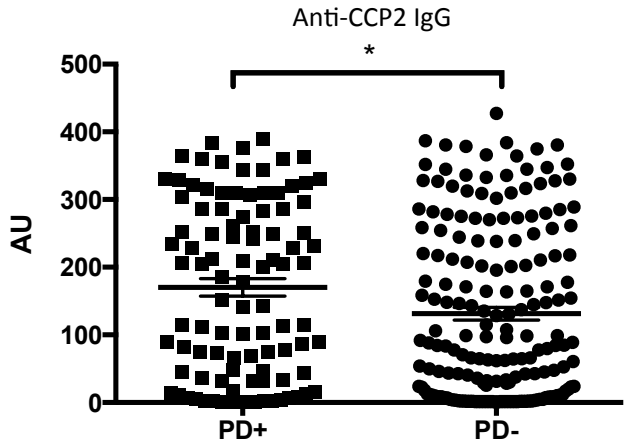


Figure 5.8 *Anti-CCP2 positivity in PD⁺ PD⁻ individuals.* Anti-CCP2 antibody levels were measured in the WOR cohort of 617 individuals using an anti-CCP-2 test kit as per manufacturers instructions and grouped into PD⁺ (187) or PD⁻ (430). Absorbance was measured at OD550nm and converted to AU using a standard curve in Microsoft Excel. Values are blank corrected mean n=2 ±SE. Each dot represents one sample. Statistical difference between PD⁺ and PD⁻ groups was calculated by Mann-Whitney testing in Microsoft Excel, * = P< 0.05.

OPERA cohort patient demographics

The OPERA study was designed to investigate the effect of periodontal treatment on RA, measured by changes in antibody levels and clinical measures (not reported here). A total of 108 patients were enrolled in the study (Table 5.4).

The youngest average participants were in the healthy control group (36.7 years) and oldest in the RA group (58.6 years). The RA group had an increased proportion of females (83.3%), probably due to the female gender bias of RA. Caucasian participants were increased in RA PD⁺ and RA groups compared to RA

and healthy respectively (p=0.050, p<0.0005), however this should not affect treatment outcomes. The percentage of ever smokers was higher in RA PD$^+$ and RA groups (58.4%, 92.9%) than healthy controls, most likely as smoking is an environmental risk factor for both RA and PD. CAL was measured for all participants, and was significantly higher in RA PD$^+$ (3.82) compared to RA (2.86) (p<0.001) and in RA compared to healthy controls (2.86, 1.76) (p<0.001). DAS28 was measured for all RA patients and was increased in RA PD$^+$ group compared to RA alone (mean 4.09 compared to 3.48) (p<0.0001).

Table 5.4 *OPERA cohort patient demographics.*

	Healthy N=18	**RA N=48**	**RA PD$^+$ N=42**	**P value**	
Age (years)	36.7	58.6	57.7	Healthy vs RA	**p<0.0001**
				RA vs RA PD$^+$	p=0.696
Female gender	61.2%	83.3%	76.2%	Healthy vs RA	p=0.053
				RA vs RA PD$^+$	p=0.274
Race					
Caucasian	88.9%	92.9%	58.4%	Healthy vs RA	**p=0.050**
		6 NR	18NR	RA vs RA PD$^+$	**p<0.0005**
BMI, kg/m2^2	26.23	28.69	26.6	Healthy vs RA	p=0.164
		1 NR	2NR	RA vs RA PD$^+$	p=0.917
Smoking status					
Ever Smoker	16.7%	44.7%	47.6%	Healthy vs RA	**p=0.030**
				RA vs RA PD$^+$	p=0.322
CAL	1.76	2.86	3.82	Healthy vs RA	**p<0.0001**
				RA vs RA PD$^+$	**p<0.0001**
DAS28	NA	3.48	4.09	RA vs RA PD$^+$	**p<0.0001**
		14 NR			

Values are expressed as means, excluding non recorded (NR) values. P values for continuous data were calculated by unpaired T-test in Microsoft Excel, p values for discrete data were calculated using Chi-square test in Microsoft excel. Statistically significant differences are shown in bold.

Antibody reactivity in RA⁺PD⁺, RA and PD patients

Antibody positivity was calculated using a 98% cut-off determined by the healthy group. IgG antibody reactivity was measured towards *P. gingivalis* antigens, recombinant PPAD and PgEnolase, CPP3, RPP3, and RgpB, and RA antigens, CEP-1 and arginine control peptide REP-1, and using the anti-CCP2 ELISA kit (Figure 5.9). For all *P. gingivalis* antigens, antibody positivity in the PD⁺ RA⁺ group was higher than the RA group, although a significant difference in level was observed for anti-RgpB only ($p<0.05$). Anti-CEP-1 antibodies were also higher in the PD⁺ RA⁺ group compared to RA, whilst anti-CCP2 and anti-REP-1 antibodies were lower in PD⁺ RA⁺, although none of these trends reached statistical significance. Percentage of antibody positivity is shown in Table 5.5.

Table 5.5 *Antibody positivity in RA⁺ PD⁺ and RA patients at baseline.*

Antigen		Peptide Sequence	Samples positive (%)	
			RA	PD⁺ RA⁺
P. gingivalis	**PgEnolase**	(recombinant protein)	31.9	42.9
	RgpB	(purified protein)	89.4	92.9
	PPAD	(recombinant protein)	38.3	54.8
	CPP3	^{71}CAKTDSYWT-Cit-DYTGWFAMYDC^{89}	46.8	61.9
	RPP3	^{71}CAKTDSYWTRDYTGWFAMYDC^{89}	100	100
ACPA	**CEP-1**	^{5}CKIHA-Cit-EIFDS-Cit-GNPTVEC^{21}	51.1	59.5
	REP-1	^{5}CKIHAREIFDSRGNPTVEC^{21}	17	9.5
	CCP2	(unknown)	93.6	83.3

Antibody levels in all PD⁺RA⁺ (42), RA (48), and healthy (18) individuals were measured at baseline after study enrollment. IgG antibody levels to PgEnolase, RgpB, PPAD, CPP3, RPP3, CEP-1 and REP-1 were measured using a standard ELISA protocol. Absorbance was read at 450nm. Anti-CCP2 IgG was measured using an anti-CCP2 kit as per manufacturers instructions. Absorbance was read at 550nm and OD values converted into AU in Microsoft Excel using a standard curve. Antibody positivity was determined by 98% cut-off in the healthy group.

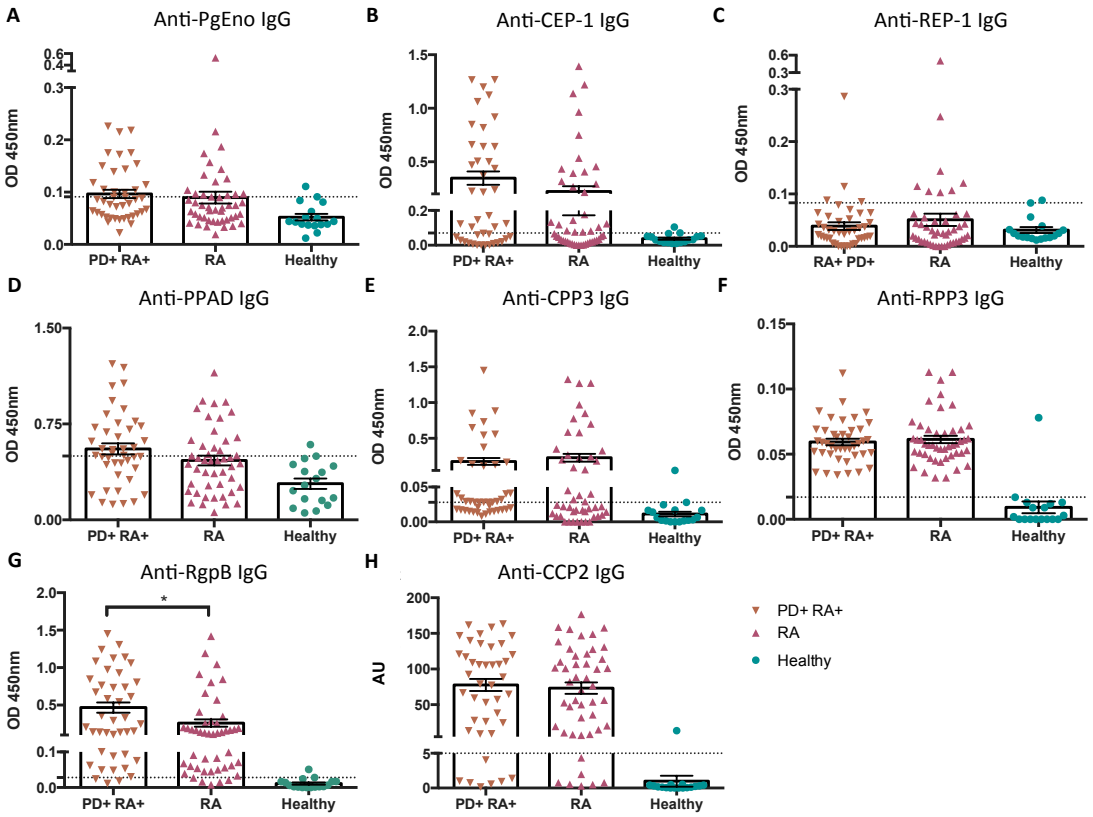


Figure 5.9 *Antibody levels in PD⁺ RA⁺, RA, and healthy individuals.* Antibody levels in all PD⁺RA⁺ (42), RA (48), and healthy (18) individuals were measured at baseline after study enrollment. IgG antibody levels to PgEnolase (A), CEP-1 (B), REP-1 (C), PPAD (D), CPP3 (E), RPP3 (F), and RgpB (G) were measured using a standard ELISA protocol. Absorbance was read at 450nm. Anti-CCP2 IgG (H) was measured using an anti-CCP2 kit as per manufacturers instructions. Absorbance was read at 550nm and OD values converted into AU in Microsoft Excel using a standard curve. For all graphs values are mean n=2 ±SE. Each symbol represents one sample. Antibody positivity was determined by 98% cut-off in the healthy group and is indicated with a dashed line. Statistical difference between PD⁺ and PD⁻ groups was calculated by Mann-Whitney testing in Microsoft Excel. * = P< 0.05.

Antibody positivity in OPERA compared to WOR cohort

From these results, it was observed that reactivity in the OPERA cohort was higher than that of the WOR cohort for all antigens tested, in the majority of patient groups. However it should be recognised that the numbers were much smaller in OPERA and the controls less well matched. In the OPERA cohort 59.5% of PD^+ RA^+ patients and 51.1% of RA patients were positive for anti-CEP-1, compared to 55.7% and 39% respectively in the WOR cohort. Most striking was the increased levels of anti-CPP3 and anti-RgpB in the OPERA cohort. 61.9% of PD^+ RA^+ and 46.8% of RA patients in OPERA were anti-CPP3 positive, compared to 8.6% and 4.2% respectively, in the WOR cohort. Similarly, 92.9% and 89.4% of PD^+ RA^+ and RA patients respectively were anti-RgpB positive, compared to 4.3% and 2.1% in the WOR cohort. However, in both cohorts, these findings consistently indicate that PD has an additive effect on antibody production in RA.

Comparing treatment and non-treatment PD^+ RA^+ groups at baseline

The PD^+ RA^+ participants in the OPERA cohort were split into two groups to receive periodontal treatment (n=20), or no periodontal treatment (n=22) for 6 months to measure the effect of periodontal therapy on antibody titre (Table 5.6) and measures of RA disease (not recorded here). Patients were matched for age and gender in treatment and non treatment groups, however the percentage of ever-smokers, and proportion of Caucasian participants in the PD^+ RA^+ group receiving treatment were significantly higher than those not receiving treatment, ($p=0.027$ and $p<0.0001$ respectively), which should be considered during analysis.

Table 5.6 *OPERA cohort patient demographics for RA PD⁺ treatment groups.*

	RA PD⁺		
	Treatment N=20	**No Treatment N=22**	**P value**
Age (years)	58.9	56.6	p=0.457
Female gender	65%	86.4%	p=0.106
Race			
Caucasian	80.0%	42.9%	p<0.0001
	10 NR	8 NR	
BMI, kg/m2[2]	23.40	29.21	p=0.712
	2 NR		
Smoking status			
Ever Smoker	65%	31.8%	p=0.027
CAL	4.08	3.59	p=0.104
DAS28	4.72	5.32	p=0.089

Values are expressed as means, (excluding non recorded (NR) values). P values for continuous data were calculated by unpaired T-test in Microsoft Excel, p values for discrete data were calculated using Chi-square test in Microsoft excel.

Baseline antibody levels in serum were measured to ensure no differences that could influence analysis (Figure 5.10). No significant differences were observed in antibodies to PgEnolase, CEP-1, REP-1, whole PPAD, CPP3, RPP3, and RgpB, or anti-CCP2 between groups.

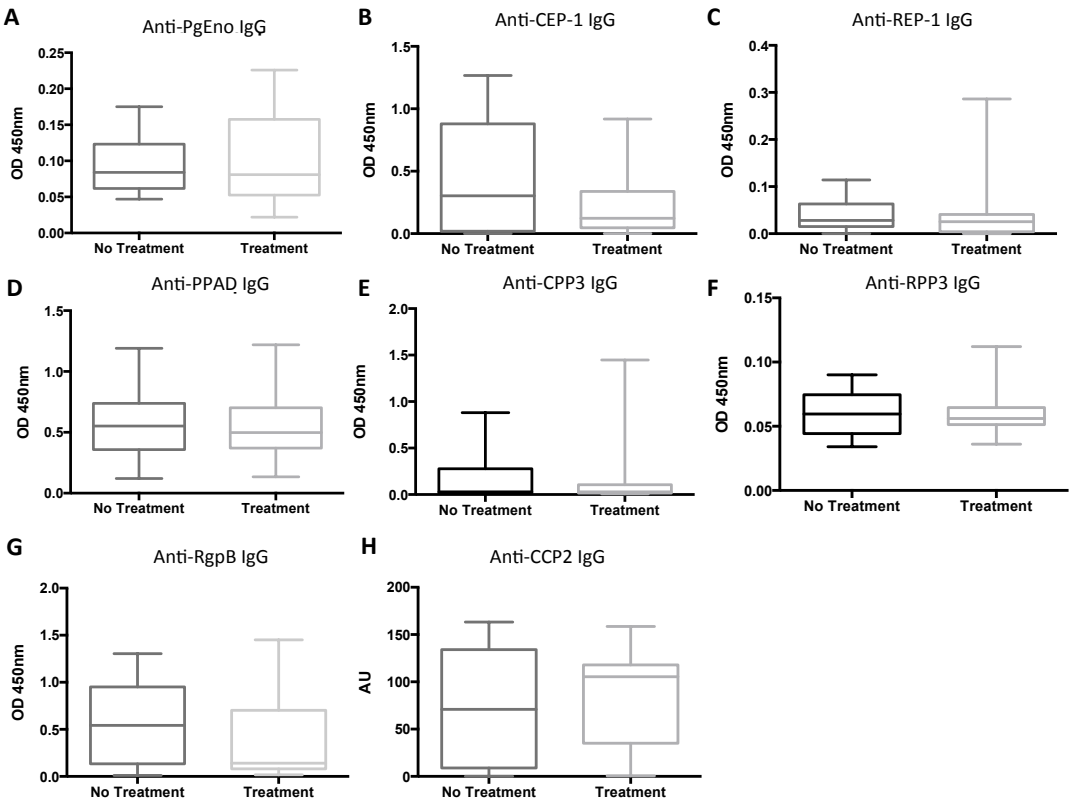


Figure 5.10 *Baseline antibody levels in PD⁺ RA⁺ individuals. A*ntibody levels in PD⁺RA⁺ individuals in PD treatment (20) and non-treatment (22) groups. IgG antibody levels to PgEnolase (A), CEP-1 (B), REP-1 (C), PPAD (D), CPP3 (E), RPP3 (F) and RgpB (G) were measured using a standard ELISA protocol. Absorbance was read at 450nm. Anti-CCP2 IgG (H) was measured using an anti-CCP2 kit as per manufacturers instructions. Absorbance was read at 550nm and OD values converted into AU in Microsoft Excel using a standard curve. Minimum to maximum values are shown as box whiskers, box shows 25th-75th percentile and median.

Effect of PD treatment on antibody levels

Six months following treatment, or without PD treatment, serum reactivity was measured again and no significant differences were observed between antibody levels in patient groups (Figure 5.11). However, it should be noted that data relating to the outcome of PD treatment was not available, thus differences between successful and non-successful treatment of PD were not analysed.

However, when comparing the percentage of seropositivity for antibodies in groups who did or did not receive PD treatment, significant differences were observed for anti-*P. gingivalis* antibodies (Table 5.7). Without PD treatment, positivity for anti-PgEnolase and anti-PPAD increased by 9.1% and 4.5% respectively, and anti-CPP3 saw a decrease of 12.7% in 6 months. However, in patients receiving treatment, positivity of anti-PgEnolase and anti-PPAD decreased by 15% and 5% respectively, which was a statistically significant difference compared to the non-treated group for both antigens ($p<0.001$ and $p=0.002$ respectively). In addition, anti-CCP3 decreased significantly more in the treatment group than the non-treatment group (20% vs 12.7%) ($p=0.012$).

The high occurrence of anti-RPP3 antibodies (100% positivity) was maintained in both groups. Positivity for anti-RgpB saw a slight decrease (-4.7%) in the group without treatment, whereas the group receiving treatment saw no change. No differences were observed in antibody positivity for anti-CCP2, anti-CEP-1 or anti-REP-1. These findings suggest PD treatment reduced anti-*P. gingivalis* antibodies, but not ACPA overall. Clinical status assessment of PD^+ RA^+ patients receiving PD treatment will also be assessed, and results are awaited with interest.

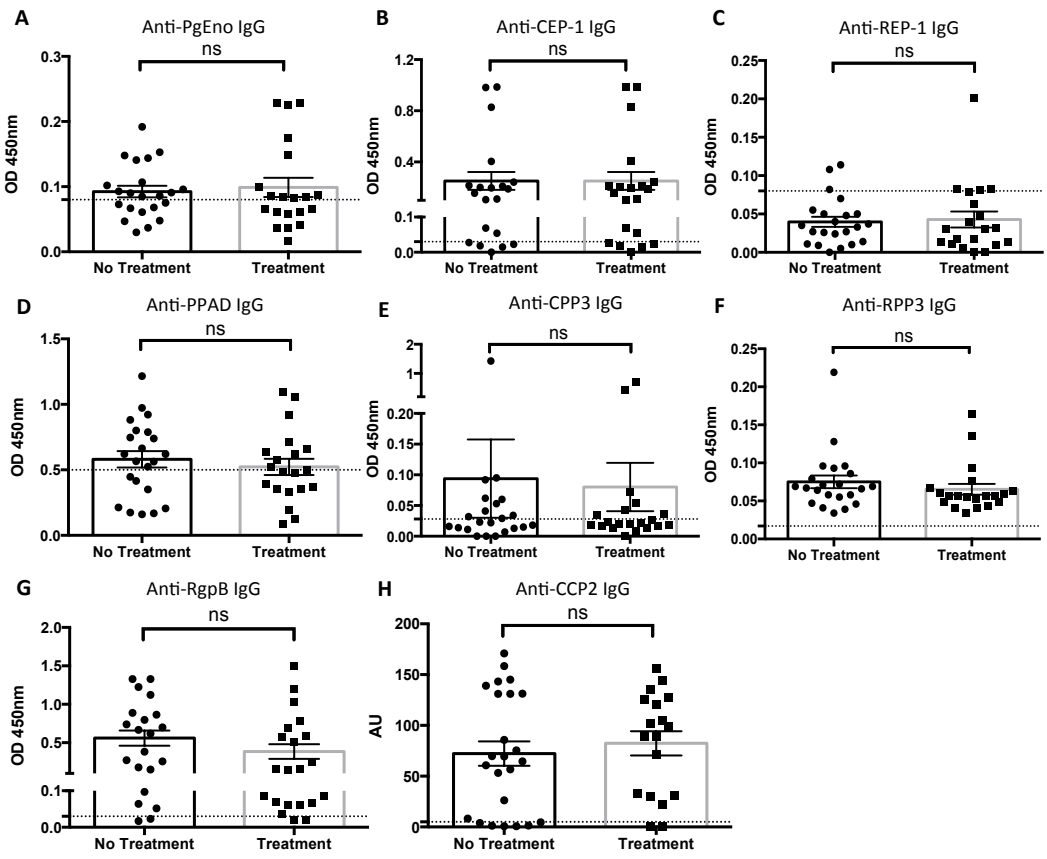


Figure 5.11 *Effect of periodontal treatment on IgG titre in PD⁺ RA⁺ individuals after 6months.* *A*ntibody levels in PD⁺ RA⁺ individuals were measured after 6 months of PD treatment (20) or no-treatment (22). IgG antibody levels to PgEnolase (A), CEP-1 (B), REP-1 (C), PPAD (D), CPP3 (E), RPP3 (F), and RgpB (G) were measured using a standard ELISA protocol. Absorbance was read at 450nm. Anti-CCP2 IgG (H) was measured using an anti-CCP2 kit as per manufacturers instructions. Absorbance was read at 550nm and OD values converted into AU in Microsoft Excel using a standard curve. For all graphs values are mean n=2 ±SE. Each symbol represents one sample. Antibody positivity was determined by 98% cut-off in the healthy group at baseline and is indicated with a dashed line. Statistical differences between groups was calculated by Mann-Whitney testing in Microsoft Excel, n.s = not statistically significant.

Table 5.7 *Change in antibody positivity following 6 months of PD treatment or no treatment.*

	Samples positive (%)							
	RA	No Treatment			Treatment			
Antigen	Baseline	Baseline	6 Months	% change	Baseline	6 Months	% change	
PgEno	31.9	40.9	50.0	+9.1	45.0	30.0	-15	**p<0.001**
RgpB	89.4	95.2	90.5	-4.7	90.5	90.5	±0	p=0.621
PPAD	38.3	59.1	63.6	+4.5	50.0	45.0	-5	**p=0.002**
CPP3	46.8	68.2	45.5	-12.7	55.0	35.0	-20	**p=0.012**
RPP3	100	100.0	100.0	±0	100.0	100.0	±0	p=1.000
CEP-1	51.1	54.5	65.0	+10.5	65.0	65.0	±0	p=0.193
REP-1	17	9.1	13.6	+4.5	10.0	20.0	+10	p=0.145
CCP2	93.6	86.4	81.8	-4.6	80.0	80.0	±0	p=0.457

Antibody levels in PD^+ RA^+ individuals were measured at baseline and after 6 months of PD treatment (20) or no-treatment (22). IgG antibody levels to PgEnolase, CEP-1, REP-1, PPAD, CPP3, RPP3, and RgpB were measured using a standard ELISA protocol. Absorbance was read at 450nm. Anti-CCP2 IgG was measured using an anti-CCP2 kit as per manufacturers instructions. Absorbance was read at 550nm and OD values converted into AU in Microsoft Excel using a standard curve. Antibody positivity was determined by 98% cut-off in the healthy group at baseline and change in antibody positivity calculated after 6 months. Statistical differences between groups was calculated by Chi square testing in Microsoft Excel, values were considered significant if $p \leq 0.05$, and are shown in bold.

Discussion

The aim of this chapter was to investigate the antibody response to *P. gingivalis* enzymes PPAD, RgpB and PgEnolase in RA, compare the relationship between anti-CEP-1 and anti-PgEnolase antibodies as a possible mechanism for cross-reactivity, and investigate the effect of PD therapy on antibody levels.

Initially, the WOR cohort of OA and RA patients was used, which has been extensively characterised already (Mikuls *et al.*, 2014, Gonzalez *et al.*, 2014, Schwenzer *et al.*, 2015). In this cohort 33% of RA patients were anti-CEP-1 positive, which is lower than previously reported (Lundberg *et al.*, 2008), although reactivity was highly citrulline specific with <1% positivity for anti-REP-1. This confirms CEP-1 to be a true citrullinated auto-antigen in RA. In this cohort CEP-1 was associated with SE and smoking, which has been previously reported in Swedish and South Korean cohorts (Mahdi *et al.*, 2009, Fisher *et al.*, 2013), but not previously in a cohort from the USA. In addition, anti-CEP-1 was positively associated with markers of bone loss in RA, and periodontal bone loss in PD, which supports previous findings that $ACPA^+$ RA is associated with increased bone erosions (Syversen *et al.*, 2008). When further divided based on both RA and PD status, anti-CEP-1 positivity was increased in patients with PD, which supports the hypothesis that cross-reactivity with PgEnolase in PD may contribute to anti-CEP-1 antibody production. This was further supported by the strong positive correlations between anti-CEP-1 and anti-CPgE antibodies.

In vivo, mice immunised with PPAD, RgpB, and PgEnolase had increased antibodies reactive with all CPgE1-6 peptides, however in the WOR cohort CPgE2 (^{6}IIG-Cit-EILDS-Cit^{15}) was the most highly reactive, with 41% of RA patients anti-

CPgE2 positive. The CPgE2 peptide (^{6}IIG-Cit-EILDS-Cit^{15}) contains the same citrulline residues as PgCEP-1 (^{5}CKIIG-Cit-EILDS-Cit-GNPTVEC^{21}), which corresponds to the CEP-1 region of α-enolase (^{5}CKIHA-Cit-EIFDS-Cit-GNPTVEC^{21}), and so is an ideal candidate for cross reactivity. In addition, structural modelling suggested CPgE2/PgCEP-1 is buried in the PgEnolase dimer, similarly to CEP-1 in α-enolase (Chapter 3), and so both are likely to act as cryptic epitopes (Warnock and Goodacre, 1997). The level of anti-CPgE2 positivity reported here also corresponds with anti-PgCEP-1 positivity of 34% in an earlier study of 104 RA patients in our lab (Lundberg *et al.*, 2008). In addition, in a study with ACPA reactive plasmablasts isolated from RA patients, 52.6% reacted with PgCEP-1, which correlated with reactivity to CEP-1 (Li *et al.*, 2016). Intriguingly, when the plasmablasts were reverted to germ-line reactivity, ACPA reactivity was lost (measured by anti-CCP2 and reactivity with cFib, CEP-1 and cVim), whilst reactivity with PgCEP-1 remained in two germ line plasmablasts. This suggests reactivity with *P. gingivalis* exists at a germline level, perhaps through evolutionarily conserved anti-bacterial responses. Following somatic mutations induced by presentation of citrullinated epitopes by T-cells, this reactivity could spread to α-enolase and broader ACPA reactivity.

In the WOR cohort, antibodies to CPP3 and RgpB were raised in 5.2% and 2.1% of RA patients without PD, which was lower than seen previously in studies by our group and others. In a previous study, 40% of RA patients had increased anti-CCP3 antibodies in a cohort of 80 RA patients (Quirke *et al.*, 2014a), which was also consistent with positivity of 46.8% in the OPERA cohort. In a study of 1,974 RA patients, 23% had raised anti-Rgp antibodies (Kharlamova *et al.*, 2016). These data suggest antibody reactivity to CPP3 and RgpB in the WOR cohort may

have been low due to external factors in the cohort, such as geographical location- the WOR cohort was recruited from the USA, compared to UK and Swedish cohorts in the other studies, or gender bias- the WOR cohort had a male majority, compared to a female bias in the OPERA study reported here, and studies by Quirke, Kharlamova and their colleagues (Quirke *et al.*, 2014a, Kharlamova *et al.*, 2016). Further to this, some unexpected results in the OPERA cohort could indicate the control group was underpowered. I.e. anti-RPP3 positivity in the RA, and PD^+ RA^+ groups was higher than anti-CPP3, contrary to previous studies showing antibody responses to citrullinated PPAD in RA are higher than to the uncitrullinated form (Quirke *et al.*, 2014a). In future studies, a larger, matched control group should be used to ensure results are valid representations of populations.

In the OPERA cohort, we investigated the effect of PD treatment on antibody titre in PD^+RA^+ patients. First, we noted PD^+ RA^+ patients had a significantly increased DAS28 score compared to RA patients ($p=0.000001$), indicating PD increases severity of RA, which was previously observed by de Smit and colleagues ($p<0.001$) (de Smit *et al.*, 2012).

Prior to the findings reported here, a number of small studies reported a beneficial effect of PD treatment on RA. Al-Katma and colleagues (Al-Katma *et al.*, 2007) showed 58.8% of RA patients had improvement in DAS28 and ESR after PD treatment, compared to 16.7% who saw improvement without PD treatment, while in a separate study improvement in DAS28, ESR and TNF-α levels was seen exclusively in the group receiving PD treatment (Ortiz *et al.*, 2009). In determining the effect of PD on RA, more research has investigated the benefit of antimicrobials in RA patients, and it has been well reported that the

addition of doxyclycline, a common antimicrobial used in the treatment of PD (Prakasam *et al.*, 2012) improves the effect of methotrexate in RA (O'Dell *et al.*, 2006, Smith *et al.*, 2011). To date, only one previous study has reported the effect of PD treatment on serology in RA patients. Lappin and colleagues measured anti-CCP2, antibodies to CEP-1, REP-1 and whole *P. gingivalis* in 39 PD patients and 36 or 30 healthy controls (for anti-CCP2 and peptide ELISA respectively) in addition to PCR detection of *P. gingivalis* DNA in plaque (Lappin *et al.*, 2013). Six months post-treatment anti-*P. gingivalis* titre was significantly decreased ($p=0.002$) irrespective of PCR positivity for *P. gingivalis* DNA. Further, anti-CCP significantly decreased in non-smokers following PD treatment ($p<0.05$), but not current smokers, suggesting smoking was a confounder in this study. These findings support the results presented in this thesis, which showed a reduction in anti-*P. gingivalis* antibodies: anti-PgEnolase, anti-PPAD and anti-CPP3 following PD treatment. Although no change was observed in anti-CCP2, anti-CEP-1 or anti-REP-1 positivity, it should be noted that there were significantly more ever smokers in the treatment arm of the OPERA PD^+ RA^+ group, which may have had a confounding effect on the results.

Taken together, the findings from both cohorts support the hypothesis that *P. gingivalis* infection in PD could provide a source for antibodies in RA, as individuals with both RA and PD had higher anti-*P. gingivalis* and ACPA antibody levels than patients with RA only, and anti-*P. gingivalis* antibodies were reduced following PD treatment.

The main findings were:

- Antibodies to PPAD, RgpB and PgEnolase are increased in RA patients and further increased in PD^+ RA^+ patients compared to controls, however reactivity to PPAD and RgpB was low in the WOR cohort.
- Antibodies to CPgE peptides correlate positively with anti-CEP-1 antibodies, indicating antibody formation in PD may induce antibodies to α-enolase through epitope spreading.
- PD treatment significantly reduced antibody levels to *P. gingivalis* in RA patients with PD.

Chapter 6 Conclusion and future directions

The aim of this thesis was to investigate PPAD in the aetiology of RA. The key findings were:

- Crystal structure of PPAD provides rationale for activity and C-terminal substrate specificity.
- Mice immunised with *P. gingivalis* proteins developed ACPA and antibodies reactive with the non-citrullinated peptide variants, and anti-*P. gingivalis* antibodies.
- RA patients had increased antibodies to *P. gingivalis* proteins, which were further increased in RA patients with PD. Periodontal therapy significantly reduced anti-*P. gingivalis* antibodies in RA patients with PD.

Structural determination of PPAD established the architecture of the protein, which is comprised of the CD and IgLF. The CD is an α/β propeller domain forming a canonical binding groove to the active site, which contains residues Asp130, His236, Asp238, Asn297, and Cys351, and the IgLF is an anti-parallel β-sandwich. In addition, residues Arg152 and Arg154 at the entrance to the binding groove were identified from the structure as key residues determining the C-terminal substrate specificity of PPAD. In a separate study, Goulas and colleagues also determined the 3D structure of a similar PPAD construct composed of the CD and IgLF (Goulas *et al.*, 2015). They modelled a structure identical to ours, confirmed the same residues in the active site, and residues involved in determining C-terminal substrate specificity. Goulas and colleagues also assessed the function of these residues *in vitro* by mutagenesis and

enzymatic studies, and published results comparable to the findings presented here (Chapter 2). Further, Goulas and colleagues were also unsuccessful in crystallisation of the full-length PPAD protein, supporting the hypothesis that full-length PPAD is insoluble due to a requirement for PorSS processing in the bacterium, which truncates PPAD for activation and secretion from *P. gingivalis in vivo* (Chapter 3).

In vitro, the combined activity of PPAD and RgpB created C-terminally citrullinated peptides from human fibrinogen and α-enolase, which were distinct from those formed by PAD2 and PAD4. This supports the hypothesis that PPAD may break tolerance to native citrullinated proteins through formation of citrullinated neoepitopes. In addition, citrullinated peptides formed from PgEnolase had high sequence homology to citrullinated peptides from α-enolase, and had reactivity with RA sera measured by ELISA (Chapter 5). Forty-one per cent of RA patients tested were positive for antibodies to PgEnolase peptide CPgE2, which encompasses the same citrulline residues as CEP-1, and so could induce cross-reactive antibody responses. This was supported by strong correlations between anti-CEP-1 and anti-CPgE antibodies in RA patients (Chapter 5), and the development of anti-CEP-1 antibodies in mice immunised with PgEnolase, which was significantly increased when mice also received PPAD and RgpB (Chapter 4).

Murine studies in this thesis showed immunisation with PPAD, RgpB, and PgEnolase created a model system for ACPA formation *in vivo*. Immunised mice developed anti-CCP2, anti-CEP-1 and anti-cVim, in addition to antibodies reactive with the arginine control peptides REP-1 and vim. These findings support and advance upon a previous study from our group, in which Kinloch

and colleagues immunised DR4-Tg mice with PgEnolase (Kinloch *et al.*, 2011). Mice immunised with either native or citrullinated PgEnolase developed inflammatory arthritis, and antibodies to both CEP-1 and REP-1. Immunisation with citrullinated or native α-enolase also induced anti-CEP-1 and anti-REP-1 antibodies, however arthritis occurred in DR4-Tg mice only, not wild type. Similarly, in a separate study using the same immunisation protocol as Kinloch, with citrullinated fibrinogen, both wild type and DR4-Tg mice developed antibodies to citrullinated fibrinogen, but only DR4-Tg mice developed arthritis (Hill *et al.*, 2003).

Taken together, results in this thesis and previous work from Kinloch, Hill and colleagues indicate wild-type mice can produce ACPA, but development of arthritis requires presentation of citrullinated antigens by MHC-II SE alleles. This supports the strong association of SE expression with $ACPA^+$ RA, and recapitulates the hypothesis that favoured presentation of citrullinated peptides by SE alleles could induce a pathogenic response.

Data obtained from crystallisation of PPAD shows a lack of structural homology with PAD2 or PAD4, making PPAD an ideal therapeutic target for inhibition. Chemical inhibitors known to be effective against PPAD include alkylating agents such as iodacetic acid, iodoacetamide and 2-chloroacetamidine ((Rodriguez *et al.*, 2010) and unpublished findings in our lab). Alkylating agents covalently bind the thiol group of cysteine residues, and so inhibit PPAD through reaction with nucleophilic Cys351 to prevent formation of the covalent intermediate. However, human PADs also function using a nucleophilic cysteine, and are inhibited by alkylating agents. Therefore, to prevent cross-reactivity, a novel PPAD inhibitor could be developed using high throughput screening (HTS)

techniques with small molecule libraries, or by *in silico* design using 3D structural data.

Effective HTS compound screening requires a functional bioassay to detect activity of the enzyme target, and a library of compounds to be screened using the assay. Compounds identified with the desired effect on the enzyme target are first validated for binding and target engagement, and further investigated for mode of action, bioavailability, and potential toxicity. These lead compounds can either be used directly for drug development, or as a chemical starting point from which to optimise.

The content in this thesis identified PPAD as a promising target for a HTS campaign. One avenue to pursue this may be via the European Lead Factory (ELF), a public-private collaboration which aims to identify unique drug starting points using a library of >500,000 compounds[2]. Currently, the colourimetric detection of citrulline assay used to measure PPAD activity is not suitable for HTS due to the requirement for low pH, high concentration of reducing agent, and high temperature. To circumvent this, a novel activity assay is in development, using fluorescent probes to measure concomitant ammonia production during citrullination. However, if a suitable bioassay for PPAD activity cannot be developed, an alternative, binding-based approach by fragment screening may be pursued. Fragment screening uses small 100-300Da compounds in a medium or high-throughput screening system to identify fragment(s) capable of binding to the enzyme active site. The relatively small size of fragments, compared to drug-like molecules in conventional HTS, renders it easier to effectively survey the diverse chemical space using significantly fewer

[2] https://www.europeanleadfactory.eu

compounds. However, fragments usually bind with low affinity, and therefore require sensitive biophysical detection such as surface plasmon resonance (SPR), or structural techniques such as co-crystallisation or nuclear magnetic resonance spectroscopy (NMR). To further the work here using this approach, the crystallography-based fragment-screening platform at the SGC could be utilised. This group uses a bespoke fragment library to soak fragments into crystals prior to structural determination. This identifies bound fragments for in-house optimisation[3].

Finally, the high resolution PPAD structural data may make it possible for structure-guide inhibitor design *in silico* (Verlinde and Hol, 1994), which requires structural data for the target, ideally with a bound lead compound. Using this information, algorithms are designed to predict substrate binding and molecular interactions by virtual screening of chemical structures, from which novel inhibitors can be designed.

In addition to drug treatment, PPAD vaccination could be considered as a preventative strategy for PD and RA. Vaccination with heat or formalin killed *P. gingivalis* in murine models of PD significantly reduced alveolar bone loss, but did not entirely prevent PD (Polak *et al.*, 2010, Wang *et al.*, 2015). This may be ineffective as a preventative strategy due to *P. gingivalis* strain diversity. 7 clinically relevant strains of *P. gingivalis* have been identified to date (Igboin *et al.*, 2009), and so it is unlikely protection could be conferred against all of them by vaccination with a single strain. Thus, vaccination with a conserved domain from *P. gingivalis* such as PPAD (Gabarrini *et al.*, 2015) or PgEnolase may provide more effective protection.

[3] http://www.thesgc.org/science/screening

Reduction in alveolar bone loss in the *P. gingivalis* model of PD has been observed in mice and rats vaccinated prior with gingipains (Gibson and Genco, 2001, Rajapakse *et al.*, 2002, Wilensky *et al.*, 2016), haemagluttinin B (Katz *et al.*, 1999) and a *P. gingivalis* capsular protein (Gonzalez *et al.*, 2003). In addition, PgEnolase is highly immunogenic *in vivo*, and previous studies have shown prior vaccination with enolase from infectious pathogens confers effective protection in mouse models. Vaccination with enolase from *Candida albicans,* and *Leishmania donovani* protected against candiasis (Li *et al.*, 2011, Shibasaki *et al.*, 2014), and visceral leishmaniasis respectively (Joshi *et al.*, 2014). In the murine *in vivo* experiments featured in this thesis, PPAD and PgEnolase had an additive effect on antibody production when immunised together. Thus, a combined approach using PPAD and PgEnolase, or antigenic peptides such as CPP3 and CPgE2, may be a promising vaccine strategy.

The association between PD and RA is well documented, although the mechanisms involved remain unknown. In this thesis, citrullination by PPAD was investigated, in addition to the role of RgpB as a co-functional enzyme, and PgEnolase, which is antigenic *in vivo*, and may be cross-reactive with α-enolase. The findings presented here support the role of *P. gingivalis* in ACPA development, and propose specific mechanisms through which this may occur. Further, the detailed structural and enzymatic characterisation of PPAD provide essential information for development of a therapeutic inhibitor.

References

Abdelsalam, S. K., Hashim, N. T., Elsalamabi, E. M. & Gismalla, B. G. 2011. Periodontal status of rheumatoid arthritis patients in khartoum state. *BMC Res Notes,* 4**,** 460.

Abdullah, S. N., Farmer, E. A., Spargo, L., Logan, R. & Gully, N. 2013. Porphyromonas gingivalis peptidylarginine deiminase substrate specificity. *Anaerobe,* 23**,** 102-8.

Aho, K., Koskenvuo, M., Tuominen, J. & Kaprio, J. 1986. Occurrence of rheumatoid arthritis in a nationwide series of twins. *J Rheumatol,* 13**,** 899-902.

Aho, K., Palosuo, T., Heliovaara, M., Knekt, P., Alha, P. & Von Essen, R. 2000. Antifilaggrin antibodies within "normal" range predict rheumatoid arthritis in a linear fashion. *J Rheumatol,* 27**,** 2743-6.

Al-Katma, M. K., Bissada, N. F., Bordeaux, J. M., Sue, J. & Askari, A. D. 2007. Control of periodontal infection reduces the severity of active rheumatoid arthritis. *J Clin Rheumatol,* 13**,** 134-7.

Albert, L. J. & Inman, R. D. 1999. Molecular mimicry and autoimmunity. *N Engl J Med,* 341**,** 2068-74.

Aletaha, D., Neogi, T., Silman, A. J., Funovits, J., Felson, D. T., Bingham, C. O., 3rd, Birnbaum, N. S., Burmester, G. R., Bykerk, V. P., Cohen, M. D., Combe, B., Costenbader, K. H., Dougados, M., Emery, P., Ferraccioli, G., Hazes, J. M., Hobbs, K., Huizinga, T. W., Kavanaugh, A., Kay, J., Kvien, T. K., Laing, T., Mease, P., Menard, H. A., Moreland, L. W., Naden, R. L., Pincus, T., Smolen, J. S., Stanislawska-Biernat, E., Symmons, D., Tak, P. P., Upchurch, K. S., Vencovsky, J., Wolfe, F. & Hawker, G. 2010. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis,* 69**,** 1580-8.

Alexander, K. S., Madden, T. E. & Farrell, D. H. 2011. Association between gamma' fibrinogen levels and inflammation. *Thromb Haemost,* 105**,** 605-9.

Amornchat, C., Rassameemasmaung, S., Sripairojthikoon, W. & Swasdison, S. 2003. Invasion of Porphyromonas gingivalis into human gingival fibroblasts in vitro. *J Int Acad Periodontol,* 5**,** 98-105.

Andrade, F., Darrah, E., Gucek, M., Cole, R. N., Rosen, A. & Zhu, X. 2010. Auto-citrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation. *Arthritis Rheum,* 62**,** 1630-40.

Applequist, S. E., Wallin, R. P. & Ljunggren, H. G. 2002. Variable expression of Toll-like receptor in murine innate and adaptive immune cell lines. *Int Immunol,* 14**,** 1065-74.

Arandjelovic, S., Mckenney, K. R., Leming, S. S. & Mowen, K. A. 2012. ATP induces protein arginine deiminase 2-dependent citrullination in mast cells through the P2X7 purinergic receptor. *J Immunol,* 189**,** 4112-22.

Arita, K., Hashimoto, H., Shimizu, T., Nakashima, K., Yamada, M. & Sato, M. 2004. Structural basis for Ca(2+)-induced activation of human PAD4. *Nat Struct Mol Biol,* 11**,** 777-83.

Arnett, F. C., Edworthy, S. M., Bloch, D. A., Mcshane, D. J., Fries, J. F., Cooper, N. S., Healey, L. A., Kaplan, S. R., Liang, M. H., Luthra, H. S. & Et Al. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum,* 31**,** 315-24.

Asaga, H., Yamada, M. & Senshu, T. 1998. Selective deimination of vimentin in calcium ionophore-induced apoptosis of mouse peritoneal macrophages. *Biochem Biophys Res Commun,* 243**,** 641-6.

Aubart, F., Crestani, B., Nicaise-Roland, P., Tubach, F., Bollet, C., Dawidowicz, K., Quintin, E., Hayem, G., Palazzo, E., Meyer, O., Chollet-Martin, S. & Dieudé, P. 2011. High levels of anti-cyclic citrullinated peptide autoantibodies are associated with co-occurrence of pulmonary diseases with rheumatoid arthritis. *J Rheumatol,* 38**,** 979-82.

Auger, I., Sebbag, M., Vincent, C., Balandraud, N., Guis, S., Nogueira, L., Svensson, B., Cantagrel, A., Serre, G. & Roudier, J. 2005. Influence of HLA-DR genes on the production of rheumatoid arthritis-specific autoantibodies to citrullinated fibrinogen. *Arthritis Rheum,* 52**,** 3424-32.

Bachrach, G., Altman, H., Kolenbrander, P. E., Chalmers, N. I., Gabai-Gutner, M., Mor, A., Friedman, M. & Steinberg, D. 2008. Resistance of Porphyromonas gingivalis ATCC 33277 to direct killing by antimicrobial peptides is protease independent. *Antimicrob Agents Chemother,* 52**,** 638-42.

Baek, K. J., Ji, S., Kim, Y. C. & Choi, Y. 2015. Association of the invasion ability of Porphyromonas gingivalis with the severity of periodontitis. *Virulence,* 6**,** 274-81.

Banbula, A., Bugno, M., Kuster, A., Heinrich, P. C., Travis, J. & Potempa, J. 1999. Rapid and efficient inactivation of IL-6 gingipains, lysine- and arginine-specific proteinases from Porphyromonas gingivalis. *Biochem Biophys Res Commun,* 261**,** 598-602.

Barton, A., Bowes, J., Eyre, S., Spreckley, K., Hinks, A., John, S. & Worthington, J. 2004. A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. *Arthritis Rheum,* 50**,** 1117-21.

Batista, F. D. & Harwood, N. E. 2009. The who, how and where of antigen presentation to B cells. *Nat Rev Immunol,* 9**,** 15-27.

Bhattacharya, S. K., Crabb, J. S., Bonilha, V. L., Gu, X., Takahara, H. & Crabb, J. W. 2006. Proteomics implicates peptidyl arginine deiminase 2 and optic nerve citrullination in glaucoma pathogenesis. *Invest Ophthalmol Vis Sci,* 47**,** 2508-14.

Bielecka, E., Scavenius, C., Kantyka, T., Jusko, M., Mizgalska, D., Szmigielski, B., Potempa, B., Enghild, J. J., Prossnitz, E. R., Blom, A. M. & Potempa, J. 2014. Peptidyl Arginine Deiminase from Porphyromonas gingivalis Abolishes Anaphylatoxin C5a Activity. *J Biol Chem,* 289**,** 32481-7.

Billings, F. 1912. CHronic focal infections and their etiologic relations to arthritis and nephritis. *Archives of Internal Medicine,* IX**,** 484-498.

Binstadt, B. A., Hebert, J. L., Ortiz-Lopez, A., Bronson, R., Benoist, C. & Mathis, D. 2009. The same systemic autoimmune disease provokes arthritis and endocarditis via distinct mechanisms. *Proc Natl Acad Sci U S A,* 106**,** 16758-63.

Biyikoglu, B., Buduneli, N., Aksu, K., Nalbantsoy, A., Lappin, D. F., Evrenosoglu, E. & Kinane, D. F. 2013. Periodontal therapy in chronic periodontitis lowers

gingival crevicular fluid interleukin-1beta and DAS28 in rheumatoid arthritis patients. *Rheumatol Int,* 33**,** 2607-16.

Bonfil, J. J., Dillier, F. L., Mercier, P., Reviron, D., Foti, B., Sambuc, R., Brodeur, J. M. & Sedarat, C. 1999. A "case control" study on the role of HLA DR4 in severe periodontitis and rapidly progressive periodontitis. Identification of types and subtypes using molecular biology (PCR.SSO). *J Clin Periodontol,* 26**,** 77-84.

Brand, D. D., Kang, A. H. & Rosloniec, E. F. 2003. Immunopathogenesis of collagen arthritis. *Springer Semin Immunopathol,* 25**,** 3-18.

Brink, M., Hansson, M., Mathsson, L., Jakobsson, P. J., Holmdahl, R., Hallmans, G., Stenlund, H., Ronnelid, J., Klareskog, L. & Rantapaa-Dahlqvist, S. 2013. Multiplex analyses of antibodies against citrullinated peptides in individuals prior to development of rheumatoid arthritis. *Arthritis Rheum,* 65**,** 899-910.

Brinkmann, V., Reichard, U., Goosmann, C., Fauler, B., Uhlemann, Y., Weiss, D. S., Weinrauch, Y. & Zychlinsky, A. 2004. Neutrophil extracellular traps kill bacteria. *Science,* 303**,** 1532-5.

Brown, L. J. & Loe, H. 1993. Prevalence, extent, severity and progression of periodontal disease. *Periodontol 2000,* 2**,** 57-71.

Burr, M. L., Naseem, H., Hinks, A., Eyre, S., Gibbons, L. J., Bowes, J., Wilson, A. G., Maxwell, J., Morgan, A. W., Emery, P., Steer, S., Hocking, L., Reid, D. M., Wordsworth, P., Harrison, P., Thomson, W., Worthington, J., Consortium, B., Consortium, Y. & Barton, A. 2010. PADI4 genotype is not associated with rheumatoid arthritis in a large UK Caucasian population. *Ann Rheum Dis,* 69**,** 666-70.

Byrne, S. J., Dashper, S. G., Darby, I. B., Adams, G. G., Hoffmann, B. & Reynolds, E. C. 2009. Progression of chronic periodontitis can be predicted by the levels of Porphyromonas gingivalis and Treponema denticola in subgingival plaque. *Oral Microbiol Immunol,* 24**,** 469-77.

Calsina, G., Ramon, J. M. & Echeverria, J. J. 2002. Effects of smoking on periodontal tissues. *J Clin Periodontol,* 29**,** 771-6.

Cambridge, G., Acharya, J., Cooper, J. A., Edwards, J. C. & Humphries, S. E. 2013. Antibodies to citrullinated peptides and risk of coronary heart disease. *Atherosclerosis,* 228**,** 243-6.

Camelo-Castillo, A., Novoa, L., Balsa-Castro, C., Blanco, J., Mira, A. & Tomas, I. 2015. Relationship between periodontitis-associated subgingival microbiota and clinical inflammation by 16S pyrosequencing. *J Clin Periodontol*.

Cao, L., Goodin, R., Wood, D., Moscarello, M. A. & Whitaker, J. N. 1999. Rapid release and unusual stability of immunodominant peptide 45-89 from citrullinated myelin basic protein. *Biochemistry,* 38**,** 6157-63.

Caplan, A. 1953. Certain unusual radiological appearances in the chest of coal-miners suffering from rheumatoid arthritis. *Thorax,* 8**,** 29-37.

Caponi, L., Petit-Teixeira, E., Sebbag, M., Bongiorni, F., Moscato, S., Pratesi, F., Pierlot, C., Osorio, J., Chapuy-Regaud, S., Guerrin, M., Cornelis, F., Serre, G., Migliorini, P. & Ecraf 2005. A family based study shows no association between rheumatoid arthritis and the PADI4 gene in a white French population. *Ann Rheum Dis,* 64**,** 587-93.

Casiano-Colon, A. & Marquis, R. E. 1988. Role of the arginine deiminase system in protecting oral bacteria and an enzymatic basis for acid tolerance. *Appl Environ Microbiol,* 54**,** 1318-24.

Castellarin, M., Warren, R. L., Freeman, J. D., Dreolini, L., Krzywinski, M., Strauss, J., Barnes, R., Watson, P., Allen-Vercoe, E., Moore, R. A. & Holt, R. A. 2012. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res,* 22**,** 299-306.

Chang, X., Han, J., Pang, L., Zhao, Y., Yang, Y. & Shen, Z. 2009. Increased PADI4 expression in blood and tissues of patients with malignant tumors. *BMC Cancer,* 9**,** 40.

Chang, X., Xia, Y., Pan, J., Meng, Q., Zhao, Y. & Yan, X. 2013. PADI2 Is Significantly Associated with Rheumatoid Arthritis. *PLoS One,* 8**,** e81259.

Chemin, K., Pollastro, S., James, E., Ge, C., Albrecht, I., Herrath, J., Gerstner, C., Tandre, K., Sampaio Rizzi, T., Ronnblom, L., Catrina, A., Holmdahl, R., Klareskog, L., De Vries, N. & Malmstrom, V. 2016. A Novel HLA-DRB1*10:01-Restricted T Cell Epitope From Citrullinated Type II Collagen Relevant to Rheumatoid Arthritis. *Arthritis Rheumatol,* 68**,** 1124-35.

Chen, C. C., Isomoto, H., Narumi, Y., Sato, K., Oishi, Y., Kobayashi, T., Yanagihara, K., Mizuta, Y., Kohno, S. & Tsukamoto, K. 2008. Haplotypes of PADI4 susceptible to rheumatoid arthritis are also associated with ulcerative colitis in the Japanese population. *Clin Immunol,* 126**,** 165-71.

Chen, H. H., Huang, N., Chen, Y. M., Chen, T. J., Chou, P., Lee, Y. L., Chou, Y. J., Lan, J. L., Lai, K. L., Lin, C. H. & Chen, D. Y. 2013. Association between a history of periodontitis and the risk of rheumatoid arthritis: a nationwide, population-based, case-control study. *Ann Rheum Dis,* 72**,** 1206-11.

Chen, R., Wei, Y., Cai, Q., Duan, S., Ren, D., Shen, J., He, D., Fang, M., Lv, K., Cheng, N. & Sun, S. 2011. The PADI4 gene does not contribute to genetic susceptibility to rheumatoid arthritis in Chinese Han population. *Rheumatol Int,* 31**,** 1631-4.

Christophorou, M. A., Castelo-Branco, G., Halley-Stott, R. P., Oliveira, C. S., Loos, R., Radzisheuskaya, A., Mowen, K. A., Bertone, P., Silva, J. C., Zernicka-Goetz, M., Nielsen, M. L., Gurdon, J. B. & Kouzarides, T. 2014. Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature,* 507**,** 104-8.

Chumanevich, A. A., Causey, C. P., Knuckley, B. A., Jones, J. E., Poudyal, D., Chumanevich, A. P., Davis, T., Matesic, L. E., Thompson, P. R. & Hofseth, L. J. 2011. Suppression of colitis in mice by Cl-amidine: a novel peptidylarginine deiminase inhibitor. *Am J Physiol Gastrointest Liver Physiol,* 300**,** G929-38.

Collaborative Computational Project, N. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr,* 50**,** 760-3.

Cornelis, F., Faure, S., Martinez, M., Prud'homme, J. F., Fritz, P., Dib, C., Alves, H., Barrera, P., De Vries, N., Balsa, A., Pascual-Salcedo, D., Maenaut, K., Westhovens, R., Migliorini, P., Tran, T. H., Delaye, A., Prince, N., Lefevre, C., Thomas, G., Poirier, M., Soubigou, S., Alibert, O., Lasbleiz, S., Fouix, S., Bouchier, C., Liote, F., Loste, M. N., Lepage, V., Charron, D., Gyapay, G., Lopes-Vaz, A., Kuntz, D., Bardin, T., Weissenbach, J. & Ecraf 1998. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc Natl Acad Sci U S A,* 95**,** 10746-50.

Courtenay, J. S., Dallman, M. J., Dayan, A. D., Martin, A. & Mosedale, B. 1980. Immunisation against heterologous type II collagen induces arthritis in mice. *Nature,* 283**,** 666-8.

Cunningham, M. W. 2000. Pathogenesis of group A streptococcal infections. *Clin Microbiol Rev,* 13**,** 470-511.

Cunningham, M. W. 2016. Post-Streptococcal Autoimmune Sequelae: Rheumatic Fever and Beyond. *In:* FERRETTI, J. J., STEVENS, D. L. & FISCHETTI, V. A. (eds.) *Streptococcus pyogenes: Basic Biology to Clinical Manifestations.* Oklahoma City (OK).

Curtis, M. A., Slaney, J. M., Carman, R. J. & Johnson, N. W. 1991. Identification of the major surface protein antigens of Porphyromonas gingivalis using IgG antibody reactivity of periodontal case-control serum. *Oral Microbiol Immunol,* 6**,** 321-6.

Dadoun, S., Zeboulon-Ktorza, N., Combescure, C., Elhai, M., Rozenberg, S., Gossec, L. & Fautrel, B. 2013. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. *Joint Bone Spine,* 80**,** 29-33.

Damgaard, D., Senolt, L. & Nielsen, C. H. 2016. Increased levels of peptidylarginine deiminase 2 in synovial fluid from anti-CCP-positive rheumatoid arthritis patients: Association with disease activity and inflammatory markers. *Rheumatology (Oxford)*.

Damgaard, D., Senolt, L., Nielsen, M. F., Pruijn, G. J. & Nielsen, C. H. 2014. 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.

Darrah, E., Giles, J. T., Ols, M. L., Bull, H. G., Andrade, F. & Rosen, A. 2013. Erosive rheumatoid arthritis is associated with antibodies that activate PAD4 by increasing calcium sensitivity. *Sci Transl Med,* 5**,** 186ra65.

Darrah, E., Rosen, A., Giles, J. T. & Andrade, F. 2012. Peptidylarginine deiminase 2, 3 and 4 have distinct specificities against cellular substrates: novel insights into autoantigen selection in rheumatoid arthritis. *Ann Rheum Dis,* 71**,** 92-8.

Darveau, R. P. 2010. Periodontitis: a polymicrobial disruption of host homeostasis. *Nat Rev Microbiol,* 8**,** 481-90.

De Pablo, P., Dietrich, T., Chapple, I. L., Milward, M., Chowdhury, M., Charles, P. J., Buckley, C. D. & Venables, P. J. 2013. The autoantibody repertoire in periodontitis: a role in the induction of autoimmunity to citrullinated proteins in rheumatoid arthritis? *Ann Rheum Dis*.

De Pablo, P., Dietrich, T., Chapple, I. L., Milward, M., Chowdhury, M., Charles, P. J., Buckley, C. D. & Venables, P. J. 2014. The autoantibody repertoire in periodontitis: a role in the induction of autoimmunity to citrullinated proteins in rheumatoid arthritis? *Ann Rheum Dis,* 73**,** 580-6.

De Pablo, P., Dietrich, T. & Mcalindon, T. E. 2008. Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population. *J Rheumatol,* 35**,** 70-6.

De Rycke, L., Nicholas, A. P., Cantaert, T., Kruithof, E., Echols, J. D., Vandekerckhove, B., Veys, E. M., De Keyser, F. & Baeten, D. 2005. Synovial intracellular citrullinated proteins colocalizing with peptidyl arginine deiminase as pathophysiologically relevant antigenic determinants of

rheumatoid arthritis-specific humoral autoimmunity. *Arthritis Rheum,* 52**,** 2323-30.

De Rycke, L., Peene, I., Hoffman, I. E., Kruithof, E., Union, A., Meheus, L., Lebeer, K., Wyns, B., Vincent, C., Mielants, H., Boullart, L., Serre, G., Veys, E. M. & De Keyser, F. 2004. Rheumatoid factor and anticitrullinated protein antibodies in rheumatoid arthritis: diagnostic value, associations with radiological progression rate, and extra-articular manifestations. *Ann Rheum Dis,* 63**,** 1587-93.

De Smit, M. J., Brouwer, E., Westra, J., Nesse, W., Vissink, A. & Van Winkelhoff, A. J. 2012. [Effect of periodontal treatment on rheumatoid arthritis and vice versa]. *Ned Tijdschr Tandheelkd,* 119**,** 191-7.

Deeg, J., Axmann, M., Matic, J., Liapis, A., Depoil, D., Afrose, J., Curado, S., Dustin, M. L. & Spatz, J. P. 2013. T cell activation is determined by the number of presented antigens. *Nano Lett,* 13**,** 5619-26.

Delbosc, S., Alsac, J. M., Journe, C., Louedec, L., Castier, Y., Bonnaure-Mallet, M., Ruimy, R., Rossignol, P., Bouchard, P., Michel, J. B. & Meilhac, O. 2011. Porphyromonas gingivalis participates in pathogenesis of human abdominal aortic aneurysm by neutrophil activation. Proof of concept in rats. *PLoS One,* 6**,** e18679.

Dessen, A., Lawrence, C. M., Cupo, S., Zaller, D. M. & Wiley, D. C. 1997. X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. *Immunity,* 7**,** 473-81.

Di Giuseppe, D., Alfredsson, L., Bottai, M., Askling, J. & Wolk, A. 2012. Long term alcohol intake and risk of rheumatoid arthritis in women: a population based cohort study. *BMJ,* 345**,** e4230.

Dias, I. H., Marshall, L., Lambert, P. A., Chapple, I. L., Matthews, J. B. & Griffiths, H. R. 2008. Gingipains from Porphyromonas gingivalis increase the chemotactic and respiratory burst-priming properties of the 77-amino-acid interleukin-8 variant. *Infect Immun,* 76**,** 317-23.

Dipaolo, R. J. & Unanue, E. R. 2001. The level of peptide-MHC complex determines the susceptibility to autoimmune diabetes: studies in HEL transgenic mice. *Eur J Immunol,* 31**,** 3453-9.

Dissick, A., Redman, R. S., Jones, M., Rangan, B. V., Reimold, A., Griffiths, G. R., Mikuls, T. R., Amdur, R. L., Richards, J. S. & Kerr, G. S. 2010. Association of periodontitis with rheumatoid arthritis: a pilot study. *J Periodontol,* 81**,** 223-30.

Dorn, B. R., Burks, J. N., Seifert, K. N. & Progulske-Fox, A. 2000. Invasion of endothelial and epithelial cells by strains of Porphyromonas gingivalis. *FEMS Microbiol Lett,* 187**,** 139-44.

Dumont, N., Aubin, E., Proulx, D. P., Lemieux, R. & Bazin, R. 2009. Increased secretion of hyperimmune antibodies following lipopolysaccharide stimulation of CD40-activated human B cells in vitro. *Immunology,* 126**,** 588-95.

Dusad, A., Duryee, M. J., Shaw, A. T., Klassen, L. W., Anderson, D. R., Wang, D., Ren, K., Gravallese, E. M., O'dell, J. R., Mikuls, T. R. & Thiele, G. M. 2013. Induction of bone loss in DBA/1J mice immunized with citrullinated autologous mouse type II collagen in the absence of adjuvant. *Immunol Res*.

Dwivedi, N. & Radic, M. 2014. Citrullination of autoantigens implicates NETosis in the induction of autoimmunity. *Ann Rheum Dis,* 73**,** 483-91.

Eichinger, A., Beisel, H. G., Jacob, U., Huber, R., Medrano, F. J., Banbula, A., Potempa, J., Travis, J. & Bode, W. 1999. Crystal structure of gingipain R: an Arg-specific bacterial cysteine proteinase with a caspase-like fold. *EMBO J,* 18**,** 5453-62.

Emsley, P. & Cowtan, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

Esposito, G., Vitale, A. M., Leijten, F. P., Strik, A. M., Koonen-Reemst, A. M., Yurttas, P., Robben, T. J., Coonrod, S. & Gossen, J. A. 2007. Peptidylarginine deiminase (PAD) 6 is essential for oocyte cytoskeletal sheet formation and female fertility. *Mol Cell Endocrinol,* 273**,** 25-31.

Evans, P. 2006. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr,* 62**,** 72-82.

Evans, P. R. & Murshudov, G. N. 2013. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr,* 69**,** 1204-14.

Farquharson, D., Butcher, J. P. & Culshaw, S. 2012. Periodontitis, Porphyromonas, and the pathogenesis of rheumatoid arthritis. *Mucosal Immunol,* 5**,** 112-20.

Feitsma, A. L., Toes, R. E., Begovich, A. B., Chokkalingam, A. P., De Vries, R. R., Huizinga, T. W. & Van Der Helm-Van Mil, A. H. 2007. Risk of progression from undifferentiated arthritis to rheumatoid arthritis: the effect of the PTPN22 1858T-allele in anti-citrullinated peptide antibody positive patients. *Rheumatology (Oxford),* 46**,** 1092-5.

Feitsma, A. L., Van Der Voort, E. I., Franken, K. L., El Bannoudi, H., Elferink, B. G., Drijfhout, J. W., Huizinga, T. W., De Vries, R. R., Toes, R. E. & Ioan-Facsinay, A. 2010. Identification of citrullinated vimentin peptides as T cell epitopes in HLA-DR4-positive patients with rheumatoid arthritis. *Arthritis Rheum,* 62**,** 117-25.

Finckh, A. & Turesson, C. 2014. The impact of obesity on the development and progression of rheumatoid arthritis. *Ann Rheum Dis,* 73**,** 1911-3.

Fisher, B. A., Bang, S. Y., Chowdhury, M., Lee, H. S., Kim, J. H., Charles, P., Venables, P. & Bae, S. C. 2013. Smoking, the HLA-DRB1 shared epitope and ACPA fine-specificity in Koreans with rheumatoid arthritis: evidence for more than one pathogenic pathway linking smoking to disease. *Ann Rheum Dis*.

Fisher, B. A., Cartwright, A. J., Quirke, A. M., De Pablo, P., Romaguera, D., Panico, S., Mattiello, A., Gavrila, D., Navarro, C., Sacerdote, C., Vineis, P., Tumino, R., Lappin, D. F., Apatzidou, D., Culshaw, S., Potempa, J., Michaud, D. S., Riboli, E. & Venables, P. J. 2015. Smoking, Porphyromonas gingivalis and the immune response to citrullinated autoantigens before the clinical onset of rheumatoid arthritis in a Southern European nested case-control study. *BMC Musculoskelet Disord,* 16**,** 331.

Fisher, B. A., Plant, D., Brode, M., Van Vollenhoven, R. F., Mathsson, L., Symmons, D., Lundberg, K., Ronnelid, J. & Venables, P. J. 2011. Antibodies to citrullinated alpha-enolase peptide 1 and clinical and radiological outcomes in rheumatoid arthritis. *Ann Rheum Dis,* 70**,** 1095-8.

Fremont, D. H., Monnaie, D., Nelson, C. A., Hendrickson, W. A. & Unanue, E. R. 1998. Crystal structure of I-Ak in complex with a dominant epitope of lysozyme. *Immunity,* 8**,** 305-17.

Freund, J. 1956. The mode of action of immunologic adjuvants. *Bibl Tuberc***,** 130-48.

Freund, J. & Lipton, M. M. 1955. Experimental allergic encephalomyelitis after the excision of the injection site of antigen-adjuvant emulsion. *J Immunol,* 75**,** 454-9.

Fujinami, R. S., Oldstone, M. B., Wroblewska, Z., Frankel, M. E. & Koprowski, H. 1983. Molecular mimicry in virus infection: crossreaction of measles virus phosphoprotein or of herpes simplex virus protein with human intermediate filaments. *Proc Natl Acad Sci U S A,* 80**,** 2346-50.

Gabarrini, G., De Smit, M., Westra, J., Brouwer, E., Vissink, A., Zhou, K., Rossen, J. W., Stobernack, T., Van Dijl, J. M. & Van Winkelhoff, A. J. 2015. The peptidylarginine deiminase gene is a conserved feature of Porphyromonas gingivalis. *Sci Rep,* 5**,** 13936.

Garzelli, C., Taub, F. E., Scharff, J. E., Prabhakar, B. S., Ginsberg-Fellner, F. & Notkins, A. L. 1984. Epstein-Barr virus-transformed lymphocytes produce monoclonal autoantibodies that react with antigens in multiple organs. *J Virol,* 52**,** 722-5.

Genco, C. A., Cutler, C. W., Kapczynski, D., Maloney, K. & Arnold, R. R. 1991. A novel mouse model to study the virulence of and host response to Porphyromonas (Bacteroides) gingivalis. *Infect Immun,* 59**,** 1255-63.

Gibson, F. C., 3rd & Genco, C. A. 2001. Prevention of Porphyromonas gingivalis-induced oral bone loss following immunization with gingipain R1. *Infect Immun,* 69**,** 7959-63.

Giles, J. T., Danoff, S. K., Sokolove, J., Wagner, C. A., Winchester, R., Pappas, D. A., Siegelman, S., Connors, G., Robinson, W. H. & Bathon, J. M. 2013. Association of fine specificity and repertoire expansion of anticitrullinated peptide antibodies with rheumatoid arthritis associated interstitial lung disease. *Ann Rheum Dis*.

Glew, M. D., Veith, P. D., Chen, D., Seers, C. A., Chen, Y. Y. & Reynolds, E. C. 2014. Blue native-PAGE analysis of membrane protein complexes in Porphyromonas gingivalis. *J Proteomics,* 110**,** 72-92.

Glew, M. D., Veith, P. D., Peng, B., Chen, Y. Y., Gorasia, D. G., Yang, Q., Slakeski, N., Chen, D., Moore, C., Crawford, S. & Reynolds, E. C. 2012. PG0026 is the C-terminal signal peptidase of a novel secretion system of Porphyromonas gingivalis. *J Biol Chem,* 287**,** 24605-17.

Gonzalez, D., Tzianabos, A. O., Genco, C. A. & Gibson, F. C., 3rd 2003. Immunization with Porphyromonas gingivalis capsular polysaccharide prevents P. gingivalis-elicited oral bone loss in a murine model. *Infect Immun,* 71**,** 2283-7.

Gonzalez, S. M., Payne, J. B., Yu, F., Thiele, G. M., Erickson, A. R., Johnson, P. G., Schmid, M. J., Cannon, G. W., Kerr, G. S., Reimold, A. M., Sokolove, J., Robinson, W. H. & Mikuls, T. R. 2014. Alveolar Bone Loss is Associated with Circulating Anti-Citrullinated Protein Antibody (ACPA) in Rheumatoid Arthritis Patients. *J Periodontol***,** 1-16.

Gonzalez-Gay, M. A., Hajeer, A. H., Garcia-Porrua, C., Dababneh, A., Thomson, W., Ollier, W. E. & Mattey, D. L. 2002. Patients chosen for treatment with

cyclosporine because of severe rheumatoid arthritis are more likely to carry HLA-DRB1 shared epitope alleles, and have earlier disease onset. *J Rheumatol,* 29**,** 271-5.

Gorasia, D. G., Veith, P. D., Chen, D., Seers, C. A., Mitchell, H. A., Chen, Y. Y., Glew, M. D., Dashper, S. G. & Reynolds, E. C. 2015. Porphyromonas gingivalis Type IX Secretion Substrates Are Cleaved and Modified by a Sortase-Like Mechanism. *PLoS Pathog,* 11**,** e1005152.

Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R. & Strominger, J. L. 1987. Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line. *J Biol Chem,* 262**,** 16087-94.

Goulas, T., Mizgalska, D., Garcia-Ferrer, I., Kantyka, T., Guevara, T., Szmigielski, B., Sroka, A., Millan, C., Uson, I., Veillard, F., Potempa, B., Mydel, P., Sola, M., Potempa, J. & Gomis-Ruth, F. X. 2015. Structure and mechanism of a bacterial host-protein citrullinating virulence factor, Porphyromonas gingivalis peptidylarginine deiminase. *Sci Rep,* 5**,** 11969.

Gregersen, P. K. & Batliwalla, F. 2005. PTPN22 and rheumatoid arthritis: gratifying replication. *Arthritis Rheum,* 52**,** 1952-5.

Gregersen, P. K., Silver, J. & Winchester, R. J. 1987. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum,* 30**,** 1205-13.

Griffen, A. L., Beall, C. J., Campbell, J. H., Firestone, N. D., Kumar, P. S., Yang, Z. K., Podar, M. & Leys, E. J. 2012. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J,* 6**,** 1176-85.

Gully, N., Bright, R., Marino, V., Marchant, C., Cantley, M., Haynes, D., Butler, C., Dashper, S., Reynolds, E. & Bartold, M. 2014. Porphyromonas gingivalis peptidylarginine deiminase, a key contributor in the pathogenesis of experimental periodontal disease and experimental arthritis. *PLoS One,* 9**,** e100838.

Guo, Y., Nguyen, K. A. & Potempa, J. 2010. Dichotomy of gingipains action as virulence factors: from cleaving substrates with the precision of a surgeon's knife to a meat chopper-like brutal degradation of proteins. *Periodontol 2000,* 54**,** 15-44.

Haffajee, A. D., Cugini, M. A., Tanner, A., Pollack, R. P., Smith, C., Kent, R. L., Jr. & Socransky, S. S. 1998. Subgingival microbiota in healthy, well-maintained elder and periodontitis subjects. *J Clin Periodontol,* 25**,** 346-53.

Hagiwara, T., Nakashima, K., Hirano, H., Senshu, T. & Yamada, M. 2002. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem Biophys Res Commun,* 290**,** 979-83.

Hajishengallis, G., Darveau, R. P. & Curtis, M. A. 2012. The keystone-pathogen hypothesis. *Nat Rev Microbiol,* 10**,** 717-25.

Hajishengallis, G. & Lamont, R. J. 2012. Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. *Mol Oral Microbiol,* 27**,** 409-19.

Halperin, R. O., Gaziano, J. M. & Sesso, H. D. 2008. Smoking and the risk of incident hypertension in middle-aged and older men. *Am J Hypertens,* 21**,** 148-52.

Halvorsen, E. H., Haavardsholm, E. A., Pollmann, S., Boonen, A., Van Der Heijde, D., Kvien, T. K. & Molberg, O. 2009. Serum IgG antibodies to peptidylarginine

deiminase 4 predict radiographic progression in patients with rheumatoid arthritis treated with tumour necrosis factor-alpha blocking agents. *Ann Rheum Dis,* 68**,** 249-52.

Hara, Y., Kaneko, T., Yoshimura, A. & Kato, I. 1996. Serum rheumatoid factor induced by intraperitoneal administration of periodontopathic bacterial lipopolysaccharide in mice. *J Periodontal Res,* 31**,** 502-7.

Harlow, L., Rosas, I. O., Gochuico, B. R., Mikuls, T. R., Dellaripa, P. F., Oddis, C. V. & Ascherman, D. P. 2013. Identification of citrullinated hsp90 isoforms as novel autoantigens in rheumatoid arthritis-associated interstitial lung disease. *Arthritis Rheum,* 65**,** 869-79.

Harre, U., Georgess, D., Bang, H., Bozec, A., Axmann, R., Ossipova, E., Jakobsson, P. J., Baum, W., Nimmerjahn, F., Szarka, E., Sarmay, G., Krumbholz, G., Neumann, E., Toes, R., Scherer, H. U., Catrina, A. I., Klareskog, L., Jurdic, P. & Schett, G. 2012. Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. *J Clin Invest,* 122**,** 1791-802.

Harris, M. L., Darrah, E., Lam, G. K., Bartlett, S. J., Giles, J. T., Grant, A. V., Gao, P., Scott, W. W., Jr., El-Gabalawy, H., Casciola-Rosen, L., Barnes, K. C., Bathon, J. M. & Rosen, A. 2008. Association of autoimmunity to peptidyl arginine deiminase type 4 with genotype and disease severity in rheumatoid arthritis. *Arthritis Rheum,* 58**,** 1958-67.

Harvey, G. P., Fitzsimmons, T. R., Dhamarpatni, A. A., Marchant, C., Haynes, D. R. & Bartold, P. M. 2013. Expression of peptidylarginine deiminase-2 and -4, citrullinated proteins and anti-citrullinated protein antibodies in human gingiva. *J Periodontal Res,* 48**,** 252-61.

Heidari, B. 2011. Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I. *Caspian J Intern Med,* 2**,** 205-12.

Heitz-Mayfield, L. J. & Lang, N. P. 2013. Surgical and nonsurgical periodontal therapy. Learned and unlearned concepts. *Periodontol 2000,* 62**,** 218-31.

Hensen, S. M., Boelens, W. C., Bonger, K. M., Van Cruchten, R. T., Van Delft, F. L. & Pruijn, G. J. 2015. Phenylglyoxal-based visualization of citrullinated proteins on Western blots. *Molecules,* 20**,** 6592-600.

Hidaka, Y., Hagiwara, T. & Yamada, M. 2005. Methylation of the guanidino group of arginine residues prevents citrullination by peptidylarginine deiminase IV. *FEBS Lett,* 579**,** 4088-92.

Hill, J., Cairns, E. & Bell, D. A. 2004. The joy of citrulline: new insights into the diagnosis, pathogenesis, and treatment of rheumatoid arthritis. *J Rheumatol,* 31**,** 1471-3.

Hill, J. A., Bell, D. A., Brintnell, W., Yue, D., Wehrli, B., Jevnikar, A. M., Lee, D. M., Hueber, W., Robinson, W. H. & Cairns, E. 2008. Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. *J Exp Med,* 205**,** 967-79.

Hill, J. A., Southwood, S., Sette, A., Jevnikar, A. M., Bell, D. A. & Cairns, E. 2003. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol,* 171**,** 538-41.

Hitchon, C. A., Chandad, F., Ferucci, E. D., Willemze, A., Ioan-Facsinay, A., Van Der Woude, D., Markland, J., Robinson, D., Elias, B., Newkirk, M., Toes, R. M., Huizinga, T. W. & El-Gabalawy, H. S. 2010. Antibodies to porphyromonas gingivalis are associated with anticitrullinated protein antibodies in

patients with rheumatoid arthritis and their relatives. *J Rheumatol,* 37**,** 1105-12.

Holmdahl, R., Andersson, M. E., Goldschmidt, T. J., Jansson, L., Karlsson, M., Malmstrom, V. & Mo, J. 1989. Collagen induced arthritis as an experimental model for rheumatoid arthritis. Immunogenetics, pathogenesis and autoimmunity. *APMIS,* 97**,** 575-84.

Holmdahl, R., Jansson, L., Andersson, M. & Jonsson, R. 1992. Genetic, hormonal and behavioural influence on spontaneously developing arthritis in normal mice. *Clin Exp Immunol,* 88**,** 467-72.

Holmdahl, R., Rubin, K., Klareskog, L., Larsson, E. & Wigzell, H. 1986. Characterization of the antibody response in mice with type II collagen-induced arthritis, using monoclonal anti-type II collagen antibodies. *Arthritis Rheum,* 29**,** 400-10.

Holoshitz, J. 2010. The rheumatoid arthritis HLA-DRB1 shared epitope. *Curr Opin Rheumatol,* 22**,** 293-8.

Holt, S. C., Ebersole, J., Felton, J., Brunsvold, M. & Kornman, K. S. 1988. Implantation of Bacteroides gingivalis in nonhuman primates initiates progression of periodontitis. *Science,* 239**,** 55-7.

Holt, S. C., Kesavalu, L., Walker, S. & Genco, C. A. 1999. Virulence factors of Porphyromonas gingivalis. *Periodontol 2000,* 20**,** 168-238.

Huizinga, T. W., Amos, C. I., Van Der Helm-Van Mil, A. H., Chen, W., Van Gaalen, F. A., Jawaheer, D., Schreuder, G. M., Wener, M., Breedveld, F. C., Ahmad, N., Lum, R. F., De Vries, R. R., Gregersen, P. K., Toes, R. E. & Criswell, L. A. 2005. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. *Arthritis Rheum,* 52**,** 3433-8.

Hviid, A., Svanstrom, H. & Frisch, M. 2011. Antibiotic use and inflammatory bowel diseases in childhood. *Gut,* 60**,** 49-54.

Igboin, C. O., Griffen, A. L. & Leys, E. J. 2009. Porphyromonas gingivalis strain diversity. *J Clin Microbiol,* 47**,** 3073-81.

Iida, A. & Nakamura, Y. 2004. Identification of 45 novel SNPs in the 83-kb region containing peptidylarginine deiminase types 1 and 3 loci on chromosomal band 1p36.13. *J Hum Genet,* 49**,** 387-90.

Imamura, T. 2003. The role of gingipains in the pathogenesis of periodontal disease. *J Periodontol,* 74**,** 111-8.

Imamura, T., Potempa, J., Pike, R. N. & Travis, J. 1995. Dependence of vascular permeability enhancement on cysteine proteinases in vesicles of Porphyromonas gingivalis. *Infect Immun,* 63**,** 1999-2003.

Imamura, T., Potempa, J., Tanase, S. & Travis, J. 1997. Activation of blood coagulation factor X by arginine-specific cysteine proteinases (gingipain-Rs) from Porphyromonas gingivalis. *J Biol Chem,* 272**,** 16062-7.

Inglis, J. J., Criado, G., Medghalchi, M., Andrews, M., Sandison, A., Feldmann, M. & Williams, R. O. 2007. Collagen-induced arthritis in C57BL/6 mice is associated with a robust and sustained T-cell response to type II collagen. *Arthritis Res Ther,* 9**,** R113.

Ioan-Facsinay, A., El-Bannoudi, H., Scherer, H. U., Van Der Woude, D., Ménard, H. A., Lora, M., Trouw, L. A., Huizinga, T. W. & Toes, R. E. 2011. Anti-cyclic citrullinated peptide antibodies are a collection of anti-citrullinated

protein antibodies and contain overlapping and non-overlapping reactivities. *Ann Rheum Dis,* 70**,** 188-93.

Ippolito, A., Wallace, D. J., Gladman, D., Fortin, P. R., Urowitz, M., Werth, V., Costner, M., Gordon, C., Alarcon, G. S., Ramsey-Goldman, R., Maddison, P., Clarke, A., Bernatsky, S., Manzi, S., Bae, S. C., Merrill, J. T., Ginzler, E., Hanly, J. G., Nived, O., Sturfelt, G., Sanchez-Guerrero, J., Bruce, I., Aranow, C., Isenberg, D., Zoma, A., Magder, L. S., Buyon, J., Kalunian, K., Dooley, M. A., Steinsson, K., Van Vollenhoven, R. F., Stoll, T., Weisman, M. & Petri, M. 2011. Autoantibodies in systemic lupus erythematosus: comparison of historical and current assessment of seropositivity. *Lupus,* 20**,** 250-5.

Ireland, J., Herzog, J. & Unanue, E. R. 2006. Cutting edge: unique T cells that recognize citrullinated peptides are a feature of protein immunization. *J Immunol,* 177**,** 1421-5.

Ireland, J. M. & Unanue, E. R. 2011. Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells. *J Exp Med,* 208**,** 2625-32.

Ishida-Yamamoto, A., Senshu, T., Takahashi, H., Akiyama, K., Nomura, K. & Iizuka, H. 2000. Decreased deiminated keratin K1 in psoriatic hyperproliferative epidermis. *J Invest Dermatol,* 114**,** 701-5.

Ishigami, A., Ohsawa, T., Hiratsuka, M., Taguchi, H., Kobayashi, S., Saito, Y., Murayama, S., Asaga, H., Toda, T., Kimura, N. & Maruyama, N. 2005. Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. *J Neurosci Res,* 80**,** 120-8.

Ivanov, Ii, Atarashi, K., Manel, N., Brodie, E. L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K. C., Santee, C. A., Lynch, S. V., Tanoue, T., Imaoka, A., Itoh, K., Takeda, K., Umesaki, Y., Honda, K. & Littman, D. R. 2009. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell,* 139**,** 485-98.

James, E. A., Moustakas, A. K., Bui, J., Papadopoulos, G. K., Bondinas, G., Buckner, J. H. & Kwok, W. W. 2010. HLA-DR1001 presents "altered-self" peptides derived from joint-associated proteins by accepting citrulline in three of its binding pockets. *Arthritis Rheum,* 62**,** 2909-18.

James, E. A., Rieck, M., Pieper, J., Gebe, J. A., Yue, B. B., Tatum, M., Peda, M., Sandin, C., Klareskog, L., Malmstrom, V. & Buckner, J. H. 2014. Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis Rheumatol,* 66**,** 1712-22.

Jang, B., Ishigami, A., Maruyama, N., Carp, R. I., Kim, Y. S. & Choi, E. K. 2013. Peptidylarginine deiminase and protein citrullination in prion diseases: strong evidence of neurodegeneration. *Prion,* 7**,** 42-6.

Jones, J. E., Causey, C. P., Knuckley, B., Slack-Noyes, J. L. & Thompson, P. R. 2009. Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Curr Opin Drug Discov Devel,* 12**,** 616-27.

Joshi, S., Rawat, K., Yadav, N. K., Kumar, V., Siddiqi, M. I. & Dube, A. 2014. Visceral Leishmaniasis: Advancements in Vaccine Development via Classical and Molecular Approaches. *Front Immunol,* 5**,** 380.

Kabsch, W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr,* 66**,** 125-32.

Kallberg, H., Ding, B., Padyukov, L., Bengtsson, C., Ronnelid, J., Klareskog, L., Alfredsson, L. & Group, E. S. 2011. Smoking is a major preventable risk

factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Ann Rheum Dis,* 70**,** 508-11.

Kallberg, H., Jacobsen, S., Bengtsson, C., Pedersen, M., Padyukov, L., Garred, P., Frisch, M., Karlson, E. W., Klareskog, L. & Alfredsson, L. 2009. Alcohol consumption is associated with decreased risk of rheumatoid arthritis: results from two Scandinavian case-control studies. *Ann Rheum Dis,* 68**,** 222-7.

Kallberg, H., Padyukov, L., Plenge, R. M., Ronnelid, J., Gregersen, P. K., Van Der Helm-Van Mil, A. H., Toes, R. E., Huizinga, T. W., Klareskog, L., Alfredsson, L. & Group, E. I. O. R. a. S. 2007. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. *Am J Hum Genet,* 80**,** 867-75.

Kang, C. P., Lee, H. S., Ju, H., Cho, H., Kang, C. & Bae, S. C. 2006. A functional haplotype of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Koreans. *Arthritis Rheum,* 54**,** 90-6.

Kaplan, M. J. & Radic, M. 2012. Neutrophil extracellular traps: double-edged swords of innate immunity. *J Immunol,* 189**,** 2689-95.

Kasser, U. R., Gleissner, C., Dehne, F., Michel, A., Willershausen-Zonnchen, B. & Bolten, W. W. 1997. Risk for periodontal disease in patients with longstanding rheumatoid arthritis. *Arthritis Rheum,* 40**,** 2248-51.

Katz, J., Black, K. P. & Michalek, S. M. 1999. Host responses to recombinant hemagglutinin B of Porphyromonas gingivalis in an experimental rat model. *Infect Immun,* 67**,** 4352-9.

Kawalkowska, J., Quirke, A. M., Ghari, F., Davis, S., Subramanian, V., Thompson, P. R., Williams, R. O., Fischer, R., La Thangue, N. B. & Venables, P. J. 2016. Abrogation of collagen-induced arthritis by a peptidyl arginine deiminase inhibitor is associated with modulation of T cell-mediated immune responses. *Sci Rep,* 6**,** 26430.

Kearney, P. L., Bhatia, M., Jones, N. G., Yuan, L., Glascock, M. C., Catchings, K. L., Yamada, M. & Thompson, P. R. 2005. Kinetic characterization of protein arginine deiminase 4: a transcriptional corepressor implicated in the onset and progression of rheumatoid arthritis. *Biochemistry,* 44**,** 10570-82.

Keffer, J., Probert, L., Cazlaris, H., Georgopoulos, S., Kaslaris, E., Kioussis, D. & Kollias, G. 1991. Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. *EMBO J,* 10**,** 4025-31.

Keijser, B. J., Zaura, E., Huse, S. M., Van Der Vossen, J. M., Schuren, F. H., Montijn, R. C., Ten Cate, J. M. & Crielaard, W. 2008. Pyrosequencing analysis of the oral microflora of healthy adults. *J Dent Res,* 87**,** 1016-20.

Khachigian, L. M. 2006. Collagen antibody-induced arthritis. *Nat Protoc,* 1**,** 2512-6.

Khandpur, R., Carmona-Rivera, C., Vivekanandan-Giri, A., Gizinski, A., Yalavarthi, S., Knight, J. S., Friday, S., Li, S., Patel, R. M., Subramanian, V., Thompson, P., Chen, P., Fox, D. A., Pennathur, S. & Kaplan, M. J. 2013. NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. *Sci Transl Med,* 5**,** 178ra40.

Kharlamova, N., Jiang, X., Sherina, N., Potempa, B., Israelsson, L., Quirke, A. M., Eriksson, K., Yucel-Lindberg, T., Venables, P. J., Potempa, J., Alfredsson, L. & Lundberg, K. 2016. Antibodies to Porphyromonas gingivalis Indicate

Interaction Between Oral Infection, Smoking, and Risk Genes in Rheumatoid Arthritis Etiology. *Arthritis Rheumatol,* 68**,** 604-13.

Kim, Y. C., Ko, Y., Hong, S. D., Kim, K. Y., Lee, Y. H., Chae, C. & Choi, Y. 2010. Presence of Porphyromonas gingivalis and plasma cell dominance in gingival tissues with periodontitis. *Oral Dis,* 16**,** 375-81.

Kinloch, A., Lundberg, K., Wait, R., Wegner, N., Lim, N. H., Zendman, A. J., Saxne, T., Malmstrom, V. & Venables, P. J. 2008. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. *Arthritis Rheum,* 58**,** 2287-95.

Kinloch, A., Tatzer, V., Wait, R., Peston, D., Lundberg, K., Donatien, P., Moyes, D., Taylor, P. C. & Venables, P. J. 2005. Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis. *Arthritis Res Ther,* 7**,** R1421-9.

Kinloch, A. J., Alzabin, S., Brintnell, W., Wilson, E., Barra, L., Wegner, N., Bell, D. A., Cairns, E. & Venables, P. J. 2011. Immunization with Porphyromonas gingivalis enolase induces autoimmunity to mammalian α-enolase and arthritis in DR4-IE-transgenic mice. *Arthritis Rheum,* 63**,** 3818-23.

Kitamura, Y., Matono, S., Aida, Y., Hirofuji, T. & Maeda, K. 2002. Gingipains in the culture supernatant of Porphyromonas gingivalis cleave CD4 and CD8 on human T cells. *J Periodontal Res,* 37**,** 464-8.

Klareskog, L., Malmström, V., Lundberg, K., Padyukov, L. & Alfredsson, L. 2011. Smoking, citrullination and genetic variability in the immunopathogenesis of rheumatoid arthritis. *Semin Immunol,* 23**,** 92-8.

Klareskog, L., Stolt, P., Lundberg, K., Kallberg, H., Bengtsson, C., Grunewald, J., Ronnelid, J., Harris, H. E., Ulfgren, A. K., Rantapaa-Dahlqvist, S., Eklund, A., Padyukov, L. & Alfredsson, L. 2006. 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.

Knight, J. S., Luo, W., O'dell, A. A., Yalavarthi, S., Zhao, W., Subramanian, V., Guo, C., Grenn, R. C., Thompson, P. R., Eitzman, D. T. & Kaplan, M. J. 2014. Peptidylarginine Deiminase Inhibition Reduces Vascular Damage and Modulates Innate Immune Responses in Murine Models of Atherosclerosis. *Circ Res*.

Knight, J. S., Subramanian, V., O'dell, A. A., Yalavarthi, S., Zhao, W., Smith, C. K., Hodgin, J. B., Thompson, P. R. & Kaplan, M. J. 2015. Peptidylarginine deiminase inhibition disrupts NET formation and protects against kidney, skin and vascular disease in lupus-prone MRL/lpr mice. *Ann Rheum Dis,* 74**,** 2199-206.

Knight, J. S., Zhao, W., Luo, W., Subramanian, V., O'dell, A. A., Yalavarthi, S., Hodgin, J. B., Eitzman, D. T., Thompson, P. R. & Kaplan, M. J. 2013. Peptidylarginine deiminase inhibition is immunomodulatory and vasculoprotective in murine lupus. *J Clin Invest,* 123**,** 2981-93.

Kokkonen, H., Mullazehi, M., Berglin, E., Hallmans, G., Wadell, G., Ronnelid, J. & Rantapaa-Dahlqvist, S. 2011. Antibodies of IgG, IgA and IgM isotypes against cyclic citrullinated peptide precede the development of rheumatoid arthritis. *Arthritis Res Ther,* 13**,** R13.

Kolfenbach, J. R., Deane, K. D., Derber, L. A., O'donnell, C. I., Gilliland, W. R., Edison, J. D., Rosen, A., Darrah, E., Norris, J. M. & Holers, V. M. 2010. Autoimmunity

to peptidyl arginine deiminase type 4 precedes clinical onset of rheumatoid arthritis. *Arthritis Rheum,* 62**,** 2633-9.

Konig, M. F., Paracha, A. S., Moni, M., Bingham, C. O., 3rd & Andrade, F. 2014. Defining the role of Porphyromonas gingivalis peptidylarginine deiminase (PPAD) in rheumatoid arthritis through the study of PPAD biology. *Ann Rheum Dis*.

Korganow, A. S., Ji, H., Mangialaio, S., Duchatelle, V., Pelanda, R., Martin, T., Degott, C., Kikutani, H., Rajewsky, K., Pasquali, J. L., Benoist, C. & Mathis, D. 1999. From systemic T cell self-reactivity to organ-specific autoimmune disease via immunoglobulins. *Immunity,* 10**,** 451-61.

Kostic, A. D., Gevers, D., Pedamallu, C. S., Michaud, M., Duke, F., Earl, A. M., Ojesina, A. I., Jung, J., Bass, A. J., Tabernero, J., Baselga, J., Liu, C., Shivdasani, R. A., Ogino, S., Birren, B. W., Huttenhower, C., Garrett, W. S. & Meyerson, M. 2012. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res,* 22**,** 292-8.

Kouskoff, V., Korganow, A. S., Duchatelle, V., Degott, C., Benoist, C. & Mathis, D. 1996. Organ-specific disease provoked by systemic autoimmunity. *Cell,* 87**,** 811-22.

Kuhn, K. A., Kulik, L., Tomooka, B., Braschler, K. J., Arend, W. P., Robinson, W. H. & Holers, V. M. 2006. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. *J Clin Invest,* 116**,** 961-73.

Kyburz, D. & Corr, M. 2003. The KRN mouse model of inflammatory arthritis. *Springer Semin Immunopathol,* 25**,** 79-90.

Lappin, D. F., Apatzidou, D., Quirke, A. M., Oliver-Bell, J., Butcher, J. P., Kinane, D. F., Riggio, M. P., Venables, P., Mcinnes, I. B. & Culshaw, S. 2013. Influence of periodontal disease, Porphyromonas gingivalis and cigarette smoking on systemic anti-citrullinated peptide antibody titres. *J Clin Periodontol*.

Lee, A. T., Li, W., Liew, A., Bombardier, C., Weisman, M., Massarotti, E. M., Kent, J., Wolfe, F., Begovich, A. B. & Gregersen, P. K. 2005. The PTPN22 R620W polymorphism associates with RF positive rheumatoid arthritis in a dose-dependent manner but not with HLA-SE status. *Genes Immun,* 6**,** 129-33.

Lepage, P., Hasler, R., Spehlmann, M. E., Rehman, A., Zvirbliene, A., Begun, A., Ott, S., Kupcinskas, L., Dore, J., Raedler, A. & Schreiber, S. 2011. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology,* 141**,** 227-36.

Lewis, H. D., Liddle, J., Coote, J. E., Atkinson, S. J., Barker, M. D., Bax, B. D., Bicker, K. L., Bingham, R. P., Campbell, M., Chen, Y. H., Chung, C. W., Craggs, P. D., Davis, R. P., Eberhard, D., Joberty, G., Lind, K. E., Locke, K., Maller, C., Martinod, K., Patten, C., Polyakova, O., Rise, C. E., Rudiger, M., Sheppard, R. J., Slade, D. J., Thomas, P., Thorpe, J., Yao, G., Drewes, G., Wagner, D. D., Thompson, P. R., Prinjha, R. K. & Wilson, D. M. 2015. Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation. *Nat Chem Biol,* 11**,** 189-91.

Li, P., Li, M., Lindberg, M. R., Kennett, M. J., Xiong, N. & Wang, Y. 2010a. PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps. *J Exp Med,* 207**,** 1853-62.

Li, P., Wang, D., Yao, H., Doret, P., Hao, G., Shen, Q., Qiu, H., Zhang, X., Wang, Y., Chen, G. & Wang, Y. 2010b. Coordination of PAD4 and HDAC2 in the regulation of p53-target gene expression. *Oncogene,* 29**,** 3153-62.

Li, P., Yao, H., Zhang, Z., Li, M., Luo, Y., Thompson, P. R., Gilmour, D. S. & Wang, Y. 2008. Regulation of p53 target gene expression by peptidylarginine deiminase 4. *Mol Cell Biol,* 28**,** 4745-58.

Li, S., Yu, Y., Yue, Y., Liao, H., Xie, W., Thai, J., Mikuls, T. R., Thiele, G. M., Duryee, M. J., Sayles, H., Payne, J. B., Klassen, L. W., O'dell, J. R., Zhang, Z. & Su, K. 2016. Autoantibodies From Single Circulating Plasmablasts React With Citrullinated Antigens and Porphyromonas gingivalis in Rheumatoid Arthritis. *Arthritis Rheumatol,* 68**,** 614-26.

Li, W., Hu, X., Zhang, X., Ge, Y., Zhao, S., Hu, Y. & Ashman, R. B. 2011. Immunisation with the glycolytic enzyme enolase confers effective protection against Candida albicans infection in mice. *Vaccine,* 29**,** 5526-33.

Li, Y., He, J., He, Z., Zhou, Y., Yuan, M., Xu, X., Sun, F., Liu, C., Li, J., Xie, W., Deng, Y., Qin, Y., Vannostrand, J. D., Xiao, L., Wu, L., Zhou, J., Shi, W. & Zhou, X. 2014. Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. *ISME J,* 8**,** 1879-91.

Linsky, T. & Fast, W. 2010. Mechanistic similarity and diversity among the guanidine-modifying members of the pentein superfamily. *Biochim Biophys Acta,* 1804**,** 1943-53.

Listgarten, M. A. 1976. Structure of the microbial flora associated with periodontal health and disease in man. A light and electron microscopic study. *J Periodontol,* 47**,** 1-18.

Listgarten, M. A. 1994. The structure of dental plaque. *Periodontol 2000,* 5**,** 52-65.

Liu, Y. L., Chiang, Y. H., Liu, G. Y. & Hung, H. C. 2011. Functional role of dimerization of human peptidylarginine deiminase 4 (PAD4). *PLoS One,* 6**,** e21314.

Lugli, E. B., Correia, R., Fischer, R., Lundberg, K., Bracke, K. R., Montgomery, A. B., Kessler, B. M., Brusselle, G. G. & Venables, P. J. 2015. Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: implications for autoimmunity in rheumatoid arthritis. *Arthritis Res Ther,* 17**,** 9.

Lundberg, K., Bengtsson, C., Kharlamova, N., Reed, E., Jiang, X., Kallberg, H., Pollak-Dorocic, I., Israelsson, L., Kessel, C., Padyukov, L., Holmdahl, R., Alfredsson, L. & Klareskog, L. 2012. Genetic and environmental determinants for disease risk in subsets of rheumatoid arthritis defined by the anticitrullinated protein/peptide antibody fine specificity profile. *Ann Rheum Dis*.

Lundberg, K., Kinloch, A., Fisher, B. A., Wegner, N., Wait, R., Charles, P., Mikuls, T. R. & Venables, P. J. 2008. Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase. *Arthritis Rheum,* 58**,** 3009-19.

Luo, Y., Arita, K., Bhatia, M., Knuckley, B., Lee, Y. H., Stallcup, M. R., Sato, M. & Thompson, P. R. 2006a. Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization. *Biochemistry,* 45**,** 11727-36.

Luo, Y., Knuckley, B., Lee, Y. H., Stallcup, M. R. & Thompson, P. R. 2006b. A fluoroacetamidine-based inactivator of protein arginine deiminase 4:

design, synthesis, and in vitro and in vivo evaluation. *J Am Chem Soc,* 128**,** 1092-3.

Macgregor, A. J., Snieder, H., Rigby, A. S., Koskenvuo, M., Kaprio, J., Aho, K. & Silman, A. J. 2000. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum,* 43**,** 30-7.

Mahdi, H., Fisher, B. A., Kallberg, H., Plant, D., Malmstrom, V., Ronnelid, J., Charles, P., Ding, B., Alfredsson, L., Padyukov, L., Symmons, D. P., Venables, P. J., Klareskog, L. & Lundberg, K. 2009. Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis. *Nat Genet,* 41**,** 1319-24.

Majumder, A., Kirabo, A., Karrupiah, K., Tsuda, S., Caldwell-Busby, J., Cardounel, A. J., Keseru, G. M. & Sayeski, P. P. 2011. Cell death induced by the Jak2 inhibitor, G6, correlates with cleavage of vimentin filaments. *Biochemistry,* 50**,** 7774-86.

Makrygiannakis, D., Hermansson, M., Ulfgren, A. K., Nicholas, A. P., Zendman, A. J., Eklund, A., Grunewald, J., Skold, C. M., Klareskog, L. & Catrina, A. I. 2008. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. *Ann Rheum Dis,* 67**,** 1488-92.

Mangialaio, S., Ji, H., Korganow, A. S., Kouskoff, V., Benoist, C. & Mathis, D. 1999. The arthritogenic T cell receptor and its ligand in a model of spontaneous arthritis. *Arthritis Rheum,* 42**,** 2517-23.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P., Roca, J. & Dore, J. 2006. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut,* 55**,** 205-11.

Maradit-Kremers, H., Crowson, C. S., Nicola, P. J., Ballman, K. V., Roger, V. L., Jacobsen, S. J. & Gabriel, S. E. 2005. Increased unrecognized coronary heart disease and sudden deaths in rheumatoid arthritis: a population-based cohort study. *Arthritis Rheum,* 52**,** 402-11.

Marchesan, J. T., Gerow, E. A., Schaff, R., Taut, A. D., Shin, S. Y., Sugai, J., Brand, D., Burberry, A., Jorns, J., Lundy, S. K., Nunez, G., Fox, D. A. & Giannobile, W. V. 2013. Porphyromonas gingivalis oral infection exacerbates the development and severity of collagen-induced arthritis. *Arthritis Res Ther,* 15**,** R186.

Maresz, K. J., Hellvard, A., Sroka, A., Adamowicz, K., Bielecka, E., Koziel, J., Gawron, K., Mizgalska, D., Marcinska, K. A., Benedyk, M., Pyrc, K., Quirke, A. M., Jonsson, R., Alzabin, S., Venables, P. J., Nguyen, K. A., Mydel, P. & Potempa, J. 2013. Porphyromonas gingivalis facilitates the development and progression of destructive arthritis through its unique bacterial peptidylarginine deiminase (PAD). *PLoS Pathog,* 9**,** e1003627.

Martin, F. & Kearney, J. F. 2002. Marginal-zone B cells. *Nat Rev Immunol,* 2**,** 323-35.

Martinez, A., Valdivia, A., Pascual-Salcedo, D., Lamas, J. R., Fernandez-Arquero, M., Balsa, A., Fernandez-Gutierrez, B., De La Concha, E. G. & Urcelay, E. 2005. PADI4 polymorphisms are not associated with rheumatoid arthritis in the Spanish population. *Rheumatology (Oxford),* 44**,** 1263-6.

Masson-Bessiere, C., Sebbag, M., Girbal-Neuhauser, E., Nogueira, L., Vincent, C., Senshu, T. & Serre, G. 2001. The major synovial targets of the rheumatoid

arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. *J Immunol,* 166**,** 4177-84.

Matsumoto, I., Staub, A., Benoist, C. & Mathis, D. 1999. Arthritis provoked by linked T and B cell recognition of a glycolytic enzyme. *Science,* 286**,** 1732-5.

Mcaleer, J. P. & Vella, A. T. 2008. Understanding how lipopolysaccharide impacts CD4 T-cell immunity. *Crit Rev Immunol,* 28**,** 281-99.

Mccoll, K. E. 2010. Clinical practice. Helicobacter pylori infection. *N Engl J Med,* 362**,** 1597-604.

Mcelwee, J. L., Mohanan, S., Griffith, O. L., Breuer, H. C., Anguish, L. J., Cherrington, B. D., Palmer, A. M., Howe, L. R., Subramanian, V., Causey, C. P., Thompson, P. R., Gray, J. W. & Coonrod, S. A. 2012. Identification of PADI2 as a potential breast cancer biomarker and therapeutic target. *BMC Cancer,* 12**,** 500.

Mcgraw, W. T., Potempa, J., Farley, D. & Travis, J. 1999. Purification, characterization, and sequence analysis of a potential virulence factor from Porphyromonas gingivalis, peptidylarginine deiminase. *Infect Immun,* 67**,** 3248-56.

Mellors, R. C., Heimer, R., Corcos, J. & Korngold, L. 1959. Cellular Origin of Rheumatoid Factor. *J Exp Med,* 110**,** 875-86.

Menard, L., Saadoun, D., Isnardi, I., Ng, Y. S., Meyers, G., Massad, C., Price, C., Abraham, C., Motaghedi, R., Buckner, J. H., Gregersen, P. K. & Meffre, E. 2011. The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. *J Clin Invest,* 121**,** 3635-44.

Mercado, F. B., Marshall, R. I., Klestov, A. C. & Bartold, P. M. 2001. Relationship between rheumatoid arthritis and periodontitis. *J Periodontol,* 72**,** 779-87.

Mezyk-Kopec, R., Bzowska, M., Potempa, J., Bzowska, M., Jura, N., Sroka, A., Black, R. A. & Bereta, J. 2005. Inactivation of membrane tumor necrosis factor alpha by gingipains from Porphyromonas gingivalis. *Infect Immun,* 73**,** 1506-14.

Miall, W. E., Caplan, A., Cochrane, A. L., Kilpatrick, G. S. & Oldham, P. D. 1953. An epidemiological study of rheumatoid arthritis associated with characteristic chest x-ray appearances in coal-workers. *Br Med J,* 2**,** 1231-6.

Mikuls, T. R., Payne, J. B., Reinhardt, R. A., Thiele, G. M., Maziarz, E., Cannella, A. C., Holers, V. M., Kuhn, K. A. & O'dell, J. R. 2009. Antibody responses to Porphyromonas gingivalis (P. gingivalis) in subjects with rheumatoid arthritis and periodontitis. *Int Immunopharmacol,* 9**,** 38-42.

Mikuls, T. R., Payne, J. B., Yu, F., Thiele, G. M., Reynolds, R. J., Cannon, G. W., Markt, J., Mcgowan, D., Kerr, G. S., Redman, R. S., Reimold, A., Griffiths, G., Beatty, M., Gonzalez, S. M., Bergman, D. A., Hamilton, B. C., 3rd, Erickson, A. R., Sokolove, J., Robinson, W. H., Walker, C., Chandad, F. & O'dell, J. R. 2014. Periodontitis and Porphyromonas gingivalis in patients with rheumatoid arthritis. *Arthritis Rheumatol,* 66**,** 1090-100.

Mikuls, T. R., Thiele, G. M., Deane, K. D., Payne, J. B., O'dell, J. R., Yu, F., Sayles, H., Weisman, M. H., Gregersen, P. K., Buckner, J. H., Keating, R. M., Derber, L. A., Robinson, W. H., Holers, V. M. & Norris, J. M. 2012. Porphyromonas

gingivalis and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis. *Arthritis Rheum,* 64**,** 3522-30.

Mita, Y., Dobashi, K., Endou, K., Kawata, T., Shimizu, Y., Nakazawa, T. & Mori, M. 2002. Toll-like receptor 4 surface expression on human monocytes and B cells is modulated by IL-2 and IL-4. *Immunol Lett,* 81**,** 71-5.

Moen, K., Brun, J. G., Valen, M., Skartveit, L., Eribe, E. K., Olsen, I. & Jonsson, R. 2006. Synovial inflammation in active rheumatoid arthritis and psoriatic arthritis facilitates trapping of a variety of oral bacterial DNAs. *Clin Exp Rheumatol,* 24**,** 656-63.

Monach, P. A., Mathis, D. & Benoist, C. 2008. The K/BxN arthritis model. *Curr Protoc Immunol,* Chapter 15**,** Unit 15 22.

Montes, A., Dieguez-Gonzalez, R., Perez-Pampin, E., Calaza, M., Mera-Varela, A., Gomez-Reino, J. J. & Gonzalez, A. 2011. Particular association of clinical and genetic features with autoimmunity to citrullinated α-enolase in rheumatoid arthritis. *Arthritis Rheum,* 63**,** 654-61.

Montgomery, A. B., Lugli, E. B. & Venables, P. 2015. Is Citrullination the Missing Link between Periodontal Disease and Rheumatoid Arthritis? *Current Oral Health Reports* 2**,** 30-36.

Montgomery, A. B., Venables, P. J. & Fisher, B. A. 2013. The case for measuring antibodies to specific citrullinated antigens. *Expert Rev Clin Immunol,* 9**,** 1185-92.

Moscarello, M. A., Lei, H., Mastronardi, F. G., Winer, S., Tsui, H., Li, Z., Ackerley, C., Zhang, L., Raijmakers, R. & Wood, D. D. 2013. Inhibition of peptidyl-arginine deiminases reverses protein-hypercitrullination and disease in mouse models of multiple sclerosis. *Dis Model Mech,* 6**,** 467-78.

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr,* 67**,** 355-67.

Nabozny, G. H., Bull, M. J., Hanson, J., Griffiths, M. M., Luthra, H. S. & David, C. S. 1994. Collagen-induced arthritis in T cell receptor V beta congenic B10.Q mice. *J Exp Med,* 180**,** 517-24.

Nachat, R., Mechin, M. C., Takahara, H., Chavanas, S., Charveron, M., Serre, G. & Simon, M. 2005. Peptidylarginine deiminase isoforms 1-3 are expressed in the epidermis and involved in the deimination of K1 and filaggrin. *J Invest Dermatol,* 124**,** 384-93.

Nakashima, K., Hagiwara, T., Ishigami, A., Nagata, S., Asaga, H., Kuramoto, M., Senshu, T. & Yamada, M. 1999. Molecular characterization of peptidylarginine deiminase in HL-60 cells induced by retinoic acid and 1alpha,25-dihydroxyvitamin D(3). *J Biol Chem,* 274**,** 27786-92.

Nakayama-Hamada, M., Suzuki, A., Furukawa, H., Yamada, R. & Yamamoto, K. 2008. Citrullinated fibrinogen inhibits thrombin-catalysed fibrin polymerization. *J Biochem,* 144**,** 393-8.

Nesse, W., Westra, J., Van Der Wal, J. E., Abbas, F., Nicholas, A. P., Vissink, A. & Brouwer, E. 2012. The periodontium of periodontitis patients contains citrullinated proteins which may play a role in ACPA (anti-citrullinated protein antibody) formation. *J Clin Periodontol,* 39**,** 599-607.

Newman, J. 2004. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr D Biol Crystallogr,* 60**,** 610-2.

Niederman, R., Brunkhorst, B., Smith, S., Weinreb, R. N. & Ryder, M. I. 1990. Ammonia as a potential mediator of adult human periodontal infection: inhibition of neutrophil function. *Arch Oral Biol,* 35 Suppl**,** 205S-209S.

Nielen, M. M., Van Der Horst, A. R., Van Schaardenburg, D., Van Der Horst-Bruinsma, I. E., Van De Stadt, R. J., Aarden, L., Dijkmans, B. A. & Hamann, D. 2005. Antibodies to citrullinated human fibrinogen (ACF) have diagnostic and prognostic value in early arthritis. *Ann Rheum Dis,* 64**,** 1199-204.

Nienhuis, R. L. & Mandema, E. 1964. A New Serum Factor in Patients with Rheumatoid Arthritis; the Antiperinuclear Factor. *Ann Rheum Dis,* 23**,** 302-5.

Niesen, F. H., Berglund, H. & Vedadi, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc,* 2**,** 2212-21.

Nijenhuis, S., Zendman, A. J., Vossenaar, E. R., Pruijn, G. J. & Vanvenrooij, W. J. 2004. Autoantibodies to citrullinated proteins in rheumatoid arthritis: clinical performance and biochemical aspects of an RA-specific marker. *Clin Chim Acta,* 350**,** 17-34.

O'dell, J. R., Elliott, J. R., Mallek, J. A., Mikuls, T. R., Weaver, C. A., Glickstein, S., Blakely, K. M., Hausch, R. & Leff, R. D. 2006. Treatment of early seropositive rheumatoid arthritis: doxycycline plus methotrexate versus methotrexate alone. *Arthritis Rheum,* 54**,** 621-7.

O'dwyer, D. N., Armstrong, M. E., Cooke, G., Dodd, J. D., Veale, D. J. & Donnelly, S. C. 2013. Rheumatoid Arthritis (RA) associated interstitial lung disease (ILD). *Eur J Intern Med,* 24**,** 597-603.

Obukhanych, T. V. & Nussenzweig, M. C. 2006. T-independent type II immune responses generate memory B cells. *J Exp Med,* 203**,** 305-10.

Okumura, N., Haneishi, A. & Terasawa, F. 2009. Citrullinated fibrinogen shows defects in FPA and FPB release and fibrin polymerization catalyzed by thrombin. *Clin Chim Acta,* 401**,** 119-23.

Oldstone, M. B. 1989. Molecular mimicry as a mechanism for the cause and a probe uncovering etiologic agent(s) of autoimmune disease. *Curr Top Microbiol Immunol,* 145**,** 127-35.

Oleksy, A., Banbula, A., Bugno, M., Travis, J. & Potempa, J. 2002. Proteolysis of interleukin-6 receptor (IL-6R) by Porphyromonas gingivalis cysteine proteinases (gingipains) inhibits interleukin-6-mediated cell activation. *Microb Pathog,* 32**,** 173-81.

Ortiz, P., Bissada, N. F., Palomo, L., Han, Y. W., Al-Zahrani, M. S., Panneerselvam, A. & Askari, A. 2009. Periodontal therapy reduces the severity of active rheumatoid arthritis in patients treated with or without tumor necrosis factor inhibitors. *J Periodontol,* 80**,** 535-40.

Otten, H. A. & Boerma, F. W. 1959. Significance of the Waaler-Rose test, streptococcal agagglutination, and antistreptolysin titre in the prognosis of rheumatoid arthritis. *Ann Rheum Dis,* 18**,** 24-8.

Ouhara, K., Komatsuzawa, H., Yamada, S., Shiba, H., Fujiwara, T., Ohara, M., Sayama, K., Hashimoto, K., Kurihara, H. & Sugai, M. 2005. Susceptibilities of periodontopathogenic and cariogenic bacteria to antibacterial peptides, {beta}-defensins and LL37, produced by human epithelial cells. *J Antimicrob Chemother,* 55**,** 888-96.

Panati, K., Pal, S., Rao, K. V. & Reddy, V. D. 2012. Association of single nucleotide polymorphisms (SNPs) of PADI4 gene with rheumatoid arthritis (RA) in Indian population. *Genes Genet Syst,* 87**,** 191-6.

Parseghian, M. H. & Luhrs, K. A. 2006. Beyond the walls of the nucleus: the role of histones in cellular signaling and innate immunity. *Biochem Cell Biol,* 84**,** 589-604.

Pearson, C. M. 1956. Development of arthritis, periarthritis and periostitis in rats given adjuvants. *Proc Soc Exp Biol Med,* 91**,** 95-101.

Pelegri, C., Morante, M. P., Castellote, C., Franch, A. & Castell, M. 1996. Treatment with an anti-CD4 monoclonal antibody strongly ameliorates established rat adjuvant arthritis. *Clin Exp Immunol,* 103**,** 273-8.

Pietrogrande, M., Corona, M., Milani, S., Rosti, A., Ramella, M. & Tordato, G. 1995. Relationship between rheumatoid factor and the immune response against hepatitis C virus in essential mixed cryoglobulinemia. *Clin Exp Rheumatol,* 13 Suppl 13**,** S109-13.

Pike, R. M., Sulkin, S. E. & Coggeshall, H. C. 1949. Concerning the nature of the factor in rheumatoid-arthritis serum responsible for increased agglutination of sensitized sheep erythrocytes. *J Immunol,* 63**,** 447-63.

Pischon, N., Pischon, T., Kroger, J., Gulmez, E., Kleber, B. M., Bernimoulin, J. P., Landau, H., Brinkmann, P. G., Schlattmann, P., Zernicke, J., Buttgereit, F. & Detert, J. 2008. Association among rheumatoid arthritis, oral hygiene, and periodontitis. *J Periodontol,* 79**,** 979-86.

Polak, D., Wilensky, A., Shapira, L., Halabi, A., Goldstein, D., Weiss, E. I. & Houri-Haddad, Y. 2009. Mouse model of experimental periodontitis induced by Porphyromonas gingivalis/Fusobacterium nucleatum infection: bone loss and host response. *J Clin Periodontol,* 36**,** 406-10.

Polak, D., Wilensky, A., Shapira, L., Weiss, E. I. & Houri-Haddad, Y. 2010. Vaccination of mice with Porphyromonas gingivalis or Fusobacterium nucleatum modulates the inflammatory response, but fails to prevent experimental periodontitis. *J Clin Periodontol,* 37**,** 812-7.

Popadiak, K., Potempa, J., Riesbeck, K. & Blom, A. M. 2007. Biphasic effect of gingipains from Porphyromonas gingivalis on the human complement system. *J Immunol,* 178**,** 7242-50.

Potempa, J., Banbula, A. & Travis, J. 2000. Role of bacterial proteinases in matrix destruction and modulation of host responses. *Periodontol 2000,* 24**,** 153-92.

Potempa, J. & Nguyen, K. A. 2007. Purification and characterization of gingipains. *Curr Protoc Protein Sci,* Chapter 21**,** Unit 21 20.

Potempa, J., Pike, R. & Travis, J. 1997. Titration and mapping of the active site of cysteine proteinases from Porphyromonas gingivalis (gingipains) using peptidyl chloromethanes. *Biol Chem,* 378**,** 223-30.

Potempa, J., Sroka, A., Imamura, T. & Travis, J. 2003. Gingipains, the major cysteine proteinases and virulence factors of Porphyromonas gingivalis: structure, function and assembly of multidomain protein complexes. *Curr Protein Pept Sci,* 4**,** 397-407.

Prakasam, A., Elavarasu, S. S. & Natarajan, R. K. 2012. Antibiotics in the management of aggressive periodontitis. *J Pharm Bioallied Sci,* 4**,** S252-5.

Pratesi, F., Dioni, I., Tommasi, C., Alcaro, M. C., Paolini, I., Barbetti, F., Boscaro, F., Panza, F., Puxeddu, I., Rovero, P. & Migliorini, P. 2013. Antibodies from

patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps. *Ann Rheum Dis*.

Pritzker, L. B., Joshi, S., Harauz, G. & Moscarello, M. A. 2000. Deimination of myelin basic protein. 2. Effect of methylation of MBP on its deimination by peptidylarginine deiminase. *Biochemistry,* 39**,** 5382-8.

Pritzker, L. B. & Moscarello, M. A. 1998. A novel microtubule independent effect of paclitaxel: the inhibition of peptidylarginine deiminase from bovine brain. *Biochim Biophys Acta,* 1388**,** 154-60.

Probert, L., Akassoglou, K., Alexopoulou, L., Douni, E., Haralambous, S., Hill, S., Kassiotis, G., Kontoyiannis, D., Pasparakis, M., Plows, D. & Kollias, G. 1996. Dissection of the pathologies induced by transmembrane and wild-type tumor necrosis factor in transgenic mice. *J Leukoc Biol,* 59**,** 518-25.

Pyrc, K., Milewska, A., Kantyka, T., Sroka, A., Maresz, K., Koziel, J., Nguyen, K. A., Enghild, J. J., Knudsen, A. D. & Potempa, J. 2013. Inactivation of epidermal growth factor by Porphyromonas gingivalis as a potential mechanism for periodontal tissue damage. *Infect Immun,* 81**,** 55-64.

Quirke, A. M., Lugli, E. B., Wegner, N., Hamilton, B. C., Charles, P., Chowdhury, M., Ytterberg, A. J., Zubarev, R. A., Potempa, J., Culshaw, S., Guo, Y., Fisher, B. A., Thiele, G., Mikuls, T. R. & Venables, P. J. 2014a. Heightened immune response to auto-citrullinated Porphyromonas gingivalis peptidylarginine deiminase: a potential mechanism for breaching immunologic tolerance in rheumatoid arthritis. *Ann Rheum Dis,* 73**,** 263-9.

Quirke, A. M., Lundberg, K., Potempa, J., Mikuls, T. R. & Venables, P. J. 2014b. PPAD remains a credible candidate for inducing autoimmunity in rheumatoid arthritis: comment on the article by Konig et al. *Ann Rheum Dis*.

Quirke, A. M., Perry, E., Cartwright, A., Kelly, C., De Soyza, A., Eggleton, P., Hutchinson, D. & Venables, P. J. 2015. Bronchiectasis is a Model for Chronic Bacterial Infection Inducing Autoimmunity in Rheumatoid Arthritis. *Arthritis Rheumatol,* 67**,** 2335-42.

Ragan, C. 1961. The history of the rheumatoid factor. *Arthritis Rheum,* 4**,** 571-3.

Rajapakse, P. S., O'brien-Simpson, N. M., Slakeski, N., Hoffmann, B. & Reynolds, E. C. 2002. Immunization with the RgpA-Kgp proteinase-adhesin complexes of Porphyromonas gingivalis protects against periodontal bone loss in the rat periodontitis model. *Infect Immun,* 70**,** 2480-6.

Rantapaa-Dahlqvist, S., De Jong, B. A., Berglin, E., Hallmans, G., Wadell, G., Stenlund, H., Sundin, U. & Van Venrooij, W. J. 2003. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum,* 48**,** 2741-9.

Rawlings, A. V. & Harding, C. R. 2004. Moisturization and skin barrier function. *Dermatol Ther,* 17 Suppl 1**,** 43-8.

Reichert, S., Haffner, M., Keyßer, G., Schäfer, C., Stein, J. M., Schaller, H. G., Wienke, A., Strauss, H., Heide, S. & Schulz, S. 2013. Detection of oral bacterial DNA in synovial fluid. *J Clin Periodontol,* 40**,** 591-8.

Reynisdottir, G., Karimi, R., Joshua, V., Olsen, H., Hensvold, A. H., Harju, A., Engstrom, M., Grunewald, J., Nyren, S., Eklund, A., Klareskog, L., Skold, C. M. & Catrina, A. I. 2014. Structural changes and antibody enrichment in the lungs are early features of anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheumatol,* 66**,** 31-9.

Reynisdottir, G., Olsen, H., Joshua, V., Engstrom, M., Forsslund, H., Karimi, R., Skold, C. M., Nyren, S., Eklund, A., Grunewald, J. & Catrina, A. I. 2015. Signs of immune activation and local inflammation are present in the bronchial tissue of patients with untreated early rheumatoid arthritis. *Ann Rheum Dis*.

Rodriguez, S. B., Stitt, B. L. & Ash, D. E. 2009. Expression of peptidylarginine deiminase from Porphyromonas gingivalis in Escherichia coli: enzyme purification and characterization. *Arch Biochem Biophys,* 488**,** 14-22.

Rodriguez, S. B., Stitt, B. L. & Ash, D. E. 2010. Cysteine 351 is an essential nucleophile in catalysis by Porphyromonas gingivalis peptidylarginine deiminase. *Arch Biochem Biophys,* 504**,** 190-6.

Rohrbach, A. S., Hemmers, S., Arandjelovic, S., Corr, M. & Mowen, K. A. 2012. PAD4 is not essential for disease in the K/BxN murine autoantibody-mediated model of arthritis. *Arthritis Res Ther,* 14**,** R104.

Rosamma, J., Sreeraj, R., Sameera, N. S. G. & Binoy, P. J. 2013. Association between chronic periodontitis and rheumatoid arthritis: a hospital-based case-control study. *Rheumatol Int,* 33**,** 103-9.

Rose, H. M., Ragan, C. & Et Al. 1948. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. *Proc Soc Exp Biol Med,* 68**,** 1-6.

Rosenstein, E. D., Greenwald, R. A., Kushner, L. J. & Weissmann, G. 2004. Hypothesis: the humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis. *Inflammation,* 28**,** 311-8.

Sasso, E. H. 2000. The rheumatoid factor response in the etiology of mixed cryoglobulins associated with hepatitis C virus infection. *Ann Med Interne (Paris),* 151**,** 30-40.

Sato, K., Naito, M., Yukitake, H., Hirakawa, H., Shoji, M., Mcbride, M. J., Rhodes, R. G. & Nakayama, K. 2010. A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proc Natl Acad Sci U S A,* 107**,** 276-81.

Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. 2004. Robust indexing for automatic data collection. *J Appl Crystallogr,* 37**,** 399-409.

Scally, S. W., Petersen, J., Law, S. C., Dudek, N. L., Nel, H. J., Loh, K. L., Wijeyewickrema, L. C., Eckle, S. B., Van Heemst, J., Pike, R. N., Mccluskey, J., Toes, R. E., La Gruta, N. L., Purcell, A. W., Reid, H. H., Thomas, R. & Rossjohn, J. 2013. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med,* 210**,** 2569-82.

Schaller, M., Burton, D. R. & Ditzel, H. J. 2001. Autoantibodies to GPI in rheumatoid arthritis: linkage between an animal model and human disease. *Nat Immunol,* 2**,** 746-53.

Schellekens, G. A., De Jong, B. A., Van Den Hoogen, F. H., Van De Putte, L. B. & Van Venrooij, W. J. 1998. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J Clin Invest,* 101**,** 273-81.

Scher, J. U., Ubeda, C., Equinda, M., Khanin, R., Buischi, Y., Viale, A., Lipuma, L., Attur, M., Pillinger, M. H., Weissmann, G., Littman, D. R., Pamer, E. G., Bretz, W. A. & Abramson, S. B. 2012. Periodontal disease and the oral microbiota in new-onset rheumatoid arthritis. *Arthritis Rheum,* 64**,** 3083-94.

Scherer, H. U., Van Der Woude, D., Willemze, A., Trouw, L. A., Knevel, R., Syversen, S. W., Van Der Linden, M. P., Lie, B., Huizinga, T. W., Van Der Heijde, D. M., Van Der Helm-Van Mil, A. H., Kvien, T. K. & Toes, R. E. 2011. Distinct ACPA fine specificities, formed under the influence of HLA shared epitope alleles, have no effect on radiographic joint damage in rheumatoid arthritis. *Ann Rheum Dis,* 70**,** 1461-4.

Schulte, S., Unger, C., Mo, J. A., Wendler, O., Bauer, E., Frischholz, S., Von Der Mark, K., Kalden, J. R., Holmdahl, R. & Burkhardt, H. 1998. Arthritis-related B cell epitopes in collagen II are conformation-dependent and sterically privileged in accessible sites of cartilage collagen fibrils. *J Biol Chem,* 273**,** 1551-61.

Schwenzer, A., Jiang, X., Mikuls, T. R., Payne, J. B., Sayles, H. R., Quirke, A. M., Kessler, B. M., Fischer, R., Venables, P. J., Lundberg, K. & Midwood, K. S. 2015. Identification of an immunodominant peptide from citrullinated tenascin-C as a major target for autoantibodies in rheumatoid arthritis. *Ann Rheum Dis*.

Seaman, A., Darrah, E., Infantino, M., Meacci, F., Manfredi, M., Benucci, M. & Mahler, M. 2016. Anti-peptidyl-arginine deaminase 3 (PAD3) antibodies as a promising marker to measure joint damage in patients with rheumatoid arthritis. *Autoimmun Rev,* 15**,** 776-80.

Sebbag, M., Moinard, N., Auger, I., Clavel, C., Arnaud, J., Nogueira, L., Roudier, J. & Serre, G. 2006. Epitopes of human fibrin recognized by the rheumatoid arthritis-specific autoantibodies to citrullinated proteins. *Eur J Immunol,* 36**,** 2250-63.

Sebbag, M., Simon, M., Vincent, C., Masson-Bessière, C., Girbal, E., Durieux, J. J. & Serre, G. 1995. The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies. *J Clin Invest,* 95**,** 2672-9.

Seers, C. A., Slakeski, N., Veith, P. D., Nikolof, T., Chen, Y. Y., Dashper, S. G. & Reynolds, E. C. 2006. The RgpB C-terminal domain has a role in attachment of RgpB to the outer membrane and belongs to a novel C-terminal-domain family found in Porphyromonas gingivalis. *J Bacteriol,* 188**,** 6376-86.

Sellam, J., Hendel-Chavez, H., Rouanet, S., Abbed, K., Combe, B., Le Loët, X., Tebib, J., Sibilia, J., Taoufik, Y., Dougados, M. & Mariette, X. 2011. B cell activation biomarkers as predictive factors for the response to rituximab in rheumatoid arthritis: a six-month, national, multicenter, open-label study. *Arthritis Rheum,* 63**,** 933-8.

Senshu, T., Kan, S., Ogawa, H., Manabe, M. & Asaga, H. 1996. Preferential deimination of keratin K1 and filaggrin during the terminal differentiation of human epidermis. *Biochem Biophys Res Commun,* 225**,** 712-9.

Seri, Y., Shoda, H., Suzuki, A., Matsumoto, I., Sumida, T., Fujio, K. & Yamamoto, K. 2015. Peptidylarginine deiminase type 4 deficiency reduced arthritis severity in a glucose-6-phosphate isomerase-induced arthritis model. *Sci Rep,* 5**,** 13041.

Shibasaki, S., Karasaki, M., Tafuku, S., Aoki, W., Sewaki, T. & Ueda, M. 2014. Oral Immunization Against Candidiasis Using Lactobacillus casei Displaying Enolase 1 from Candida albicans. *Sci Pharm,* 82**,** 697-708.

Shiozawa, S., Hayashi, S., Tsukamoto, Y., Goko, H., Kawasaki, H., Wada, T., Shimizu, K., Yasuda, N., Kamatani, N., Takasugi, K., Tanaka, Y., Shiozawa, K. & Imura, S. 1998. Identification of the gene loci that predispose to rheumatoid arthritis. *Int Immunol,* 10**,** 1891-5.

Shirai, H., Mokrab, Y. & Mizuguchi, K. 2006. The guanidino-group modifying enzymes: structural basis for their diversity and commonality. *Proteins,* 64**,** 1010-23.

Shoji, M., Sato, K., Yukitake, H., Kondo, Y., Narita, Y., Kadowaki, T., Naito, M. & Nakayama, K. 2011. Por secretion system-dependent secretion and glycosylation of Porphyromonas gingivalis hemin-binding protein 35. *PLoS One,* 6**,** e21372.

Silman, A. J., Macgregor, A. J., Thomson, W., Holligan, S., Carthy, D., Farhan, A. & Ollier, W. E. 1993. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol,* 32**,** 903-7.

Simon, M., Girbal, E., Sebbag, M., Gomes-Daudrix, V., Vincent, C., Salama, G. & Serre, G. 1993. The cytokeratin filament-aggregating protein filaggrin is the target of the so-called "antikeratin antibodies," autoantibodies specific for rheumatoid arthritis. *J Clin Invest,* 92**,** 1387-93.

Slade, D. J., Fang, P., Dreyton, C. J., Zhang, Y., Fuhrmann, J., Rempel, D., Bax, B. D., Coonrod, S. A., Lewis, H. D., Guo, M., Gross, M. L. & Thompson, P. R. 2015. Protein Arginine Deiminase 2 Binds Calcium in an Ordered Fashion: Implications for Inhibitor Design. *ACS Chem Biol*.

Slakeski, N., Seers, C. A., Ng, K., Moore, C., Cleal, S. M., Veith, P. D., Lo, A. W. & Reynolds, E. C. 2011. C-terminal domain residues important for secretion and attachment of RgpB in Porphyromonas gingivalis. *J Bacteriol,* 193**,** 132-42.

Slaney, J. M. & Curtis, M. A. 2008. Mechanisms of evasion of complement by Porphyromonas gingivalis. *Front Biosci,* 13**,** 188-96.

Smith, C. J., Sayles, H., Mikuls, T. R. & Michaud, K. 2011. Minocycline and doxycycline therapy in community patients with rheumatoid arthritis: prescribing patterns, patient-level determinants of use, and patient-reported side effects. *Arthritis Res Ther,* 13**,** R168.

Snir, O., Rieck, M., Gebe, J. A., Yue, B. B., Rawlings, C. A., Nepom, G., Malmstrom, V. & Buckner, J. H. 2011. Identification and functional characterization of T cells reactive to citrullinated vimentin in HLA-DRB1*0401-positive humanized mice and rheumatoid arthritis patients. *Arthritis Rheum,* 63**,** 2873-83.

Snir, O., Widhe, M., Hermansson, M., Von Spee, C., Lindberg, J., Hensen, S., Lundberg, K., Engström, A., Venables, P. J., Toes, R. E., Holmdahl, R., Klareskog, L. & Malmström, V. 2010. Antibodies to several citrullinated antigens are enriched in the joints of rheumatoid arthritis patients. *Arthritis Rheum,* 62**,** 44-52.

Snir, O., Widhe, M., Von Spee, C., Lindberg, J., Padyukov, L., Lundberg, K., Engström, A., Venables, P. J., Lundeberg, J., Holmdahl, R., Klareskog, L. & Malmström, V. 2009. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. *Ann Rheum Dis,* 68**,** 736-43.

Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C. & Kent, R. L., Jr. 1998. Microbial complexes in subgingival plaque. *J Clin Periodontol,* 25**,** 134-44.

Socransky, S. S., Smith, C. & Haffajee, A. D. 2002. Subgingival microbial profiles in refractory periodontal disease. *J Clin Periodontol,* 29**,** 260-8.

Sokolove, J., Bromberg, R., Deane, K. D., Lahey, L. J., Derber, L. A., Chandra, P. E., Edison, J. D., Gilliland, W. R., Tibshirani, R. J., Norris, J. M., Holers, V. M. & Robinson, W. H. 2012. Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. *PLoS One,* 7**,** e35296.

Sorice, M., Iannuccelli, C., Manganelli, V., Capozzi, A., Alessandri, C., Lococo, E., Garofalo, T., Di Franco, M., Bombardieri, M., Nerviani, A., Misasi, R. & Valesini, G. 2016. Autophagy generates citrullinated peptides in human synoviocytes: a possible trigger for anti-citrullinated peptide antibodies. *Rheumatology (Oxford)*.

Spengler, J., Lugonja, B., Ytterberg, A. J., Zubarev, R. A., Creese, A. J., Pearson, M. J., Grant, M. M., Milward, M., Lundberg, K., Buckley, C. D., Filer, A., Raza, K., Cooper, P. R., Chapple, I. L. & Scheel-Toellner, D. 2015. Release of Active Peptidyl Arginine Deiminases by Neutrophils Can Explain Production of Extracellular Citrullinated Autoantigens in Rheumatoid Arthritis Synovial Fluid. *Arthritis Rheumatol,* 67**,** 3135-45.

Stastny, P. 1976. Mixed lymphocyte cultures in rheumatoid arthritis. *J Clin Invest,* 57**,** 1148-57.

Stastny, P. 1978. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. *N Engl J Med,* 298**,** 869-71.

Steadman, F. S. 1914. Rheumatoid Arthritis twice Cured by the Removal of Septic Teeth. *Proc R Soc Med,* 7**,** 21-8.

Stein, J. M., Machulla, H. K., Smeets, R., Lampert, F. & Reichert, S. 2008. Human leukocyte antigen polymorphism in chronic and aggressive periodontitis among Caucasians: a meta-analysis. *J Clin Periodontol,* 35**,** 183-92.

Steinert, P. M., Parry, D. A. & Marekov, L. N. 2003. Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root sheath. *J Biol Chem,* 278**,** 41409-19.

Stoerk, H. C., Knowlton, A. I. & Loeb, E. N. 1954. Failure of liver feeding to counteract cortisone effects other than growth-inhibition. *Proc Soc Exp Biol Med,* 85**,** 627-31.

Stolt, P., Yahya, A., Bengtsson, C., Kallberg, H., Ronnelid, J., Lundberg, I., Klareskog, L., Alfredsson, L. & Group, E. S. 2010. Silica exposure among male current smokers is associated with a high risk of developing ACPA-positive rheumatoid arthritis. *Ann Rheum Dis,* 69**,** 1072-6.

Stone, E. M., Schaller, T. H., Bianchi, H., Person, M. D. & Fast, W. 2005. Inactivation of two diverse enzymes in the amidinotransferase superfamily by 2-chloroacetamidine: dimethylargininase and peptidylarginine deiminase. *Biochemistry,* 44**,** 13744-52.

Stuart, J. M. & Dixon, F. J. 1983. Serum transfer of collagen-induced arthritis in mice. *J Exp Med,* 158**,** 378-92.

Sur Chowdhury, C., Giaglis, S., Walker, U. A., Buser, A., Hahn, S. & Hasler, P. 2014. Enhanced neutrophil extracellular trap generation in rheumatoid arthritis: analysis of underlying signal transduction pathways and potential diagnostic utility. *Arthritis Res Ther,* 16**,** R122.

Susanto, H., Nesse, W., Kertia, N., Soeroso, J., Huijser Van Reenen, Y., Hoedemaker, E., Agustina, D., Vissink, A., Abbas, F. & Dijkstra, P. U. 2013. Prevalence and

severity of periodontitis in Indonesian patients with rheumatoid arthritis. *J Periodontol,* 84**,** 1067-74.

Suwannalai, P., Britsemmer, K., Knevel, R., Scherer, H. U., Levarht, E. W., Van Der Helm-Van Mil, A. H., Van Schaardenburg, D., Huizinga, T. W., Toes, R. E. & Trouw, L. A. 2013. Low-avidity anticitrullinated protein antibodies (ACPA) are associated with a higher rate of joint destruction in rheumatoid arthritis. *Ann Rheum Dis*.

Suwannalai, P., Van De Stadt, L. A., Radner, H., Steiner, G., El-Gabalawy, H. S., Zijde, C. M., Van Tol, M. J., Van Schaardenburg, D., Huizinga, T. W., Toes, R. E. & Trouw, L. A. 2012. Avidity maturation of anti-citrullinated protein antibodies in rheumatoid arthritis. *Arthritis Rheum,* 64**,** 1323-8.

Suzuki, A., Kochi, Y., Shoda, H., Seri, Y., Fujio, K., Sawada, T., Yamada, R. & Yamamoto, K. 2016. Decreased severity of experimental autoimmune arthritis in peptidylarginine deiminase type 4 knockout mice. *BMC Musculoskelet Disord,* 17**,** 205.

Suzuki, A., Yamada, R., Chang, X., Tokuhiro, S., Sawada, T., Suzuki, M., Nagasaki, M., Nakayama-Hamada, M., Kawaida, R., Ono, M., Ohtsuki, M., Furukawa, H., Yoshino, S., Yukioka, M., Tohma, S., Matsubara, T., Wakitani, S., Teshima, R., Nishioka, Y., Sekine, A., Iida, A., Takahashi, A., Tsunoda, T., Nakamura, Y. & Yamamoto, K. 2003. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet,* 34**,** 395-402.

Suzuki, T., Ikari, K., Yano, K., Inoue, E., Toyama, Y., Taniguchi, A., Yamanaka, H. & Momohara, S. 2013. PADI4 and HLA-DRB1 are genetic risks for radiographic progression in RA patients, independent of ACPA status: results from the IORRA cohort study. *PLoS One,* 8**,** e61045.

Syversen, S. W., Gaarder, P. I., Goll, G. L., Odegard, S., Haavardsholm, E. A., Mowinckel, P., Van Der Heijde, D., Landewe, R. & Kvien, T. K. 2008. High anti-cyclic citrullinated peptide levels and an algorithm of four variables predict radiographic progression in patients with rheumatoid arthritis: results from a 10-year longitudinal study. *Ann Rheum Dis,* 67**,** 212-7.

Takizawa, Y., Sawada, T., Suzuki, A., Yamada, R., Inoue, T. & Yamamoto, K. 2005. Peptidylarginine deiminase 4 (PADI4) identified as a conformation-dependent autoantigen in rheumatoid arthritis. *Scand J Rheumatol,* 34**,** 212-5.

Takizawa, Y., Suzuki, A., Sawada, T., Ohsaka, M., Inoue, T., Yamada, R. & Yamamoto, K. 2006. Citrullinated fibrinogen detected as a soluble citrullinated autoantigen in rheumatoid arthritis synovial fluids. *Ann Rheum Dis,* 65**,** 1013-20.

Tanikawa, C., Espinosa, M., Suzuki, A., Masuda, K., Yamamoto, K., Tsuchiya, E., Ueda, K., Daigo, Y., Nakamura, Y. & Matsuda, K. 2012. Regulation of histone modification and chromatin structure by the p53-PADI4 pathway. *Nat Commun,* 3**,** 676.

Taurog, J. D., Sandberg, G. P. & Mahowald, M. L. 1983. The cellular basis of adjuvant arthritis. II. Characterization of the cells mediating passive transfer. *Cell Immunol,* 80**,** 198-204.

Tedeschi, S. K., Bermas, B. & Costenbader, K. H. 2013. Sexual disparities in the incidence and course of SLE and RA. *Clin Immunol,* 149**,** 211-8.

Thiele, G. M., Duryee, M. J., Dusad, A., Hunter, C. D., Lacy, J. P., Anderson, D. R., Wang, D., O'dell, J. R., Mikuls, T. R. & Klassen, L. W. 2012. Citrullinated mouse collagen administered to DBA/1J mice in the absence of adjuvant initiates arthritis. *Int Immunopharmacol,* 13**,** 424-31.

Tolo, K. & Jorkjend, L. 1990. Serum antibodies and loss of periodontal bone in patients with rheumatoid arthritis. *J Clin Periodontol,* 17**,** 288-91.

Tonetti, M. S. 1998. Cigarette smoking and periodontal diseases: etiology and management of disease. *Ann Periodontol,* 3**,** 88-101.

Travis, J., Pike, R., Imamura, T. & Potempa, J. 1997. Porphyromonas gingivalis proteinases as virulence factors in the development of periodontitis. *J Periodontal Res,* 32**,** 120-5.

Trentham, D. E., Townes, A. S. & Kang, A. H. 1977. Autoimmunity to type II collagen an experimental model of arthritis. *J Exp Med,* 146**,** 857-68.

Turesson, C., O'fallon, W. M., Crowson, C. S., Gabriel, S. E. & Matteson, E. L. 2003. Extra-articular disease manifestations in rheumatoid arthritis: incidence trends and risk factors over 46 years. *Ann Rheum Dis,* 62**,** 722-7.

Turesson, C., Schaid, D. J., Weyand, C. M., Jacobsson, L. T., Goronzy, J. J., Petersson, I. F., Sturfelt, G., Nyhall-Wahlin, B. M., Truedsson, L., Dechant, S. A. & Matteson, E. L. 2005. The impact of HLA-DRB1 genes on extra-articular disease manifestations in rheumatoid arthritis. *Arthritis Res Ther,* 7**,** R1386-93.

Tutturen, A. E., Fleckenstein, B. & De Souza, G. A. 2014. Assessing the citrullinome in rheumatoid arthritis synovial fluid with and without enrichment of citrullinated peptides. *J Proteome Res,* 13**,** 2867-73.

Van Beers, J. J., Raijmakers, R., Alexander, L. E., Stammen-Vogelzangs, J., Lokate, A. M., Heck, A. J., Schasfoort, R. B. & Pruijn, G. J. 2010. Mapping of citrullinated fibrinogen B-cell epitopes in rheumatoid arthritis by imaging surface plasmon resonance. *Arthritis Res Ther,* 12**,** R219.

Van De Stadt, L. A., De Koning, M. H., Van De Stadt, R. J., Wolbink, G., Dijkmans, B. A., Hamann, D. & Van Schaardenburg, D. 2011a. Development of the anti-citrullinated protein antibody repertoire prior to the onset of rheumatoid arthritis. *Arthritis Rheum,* 63**,** 3226-33.

Van De Stadt, L. A., Van Der Horst, A. R., De Koning, M. H., Bos, W. H., Wolbink, G. J., Van De Stadt, R. J., Pruijn, G. J., Dijkmans, B. A., Van Schaardenburg, D. & Hamann, D. 2011b. The extent of the anti-citrullinated protein antibody repertoire is associated with arthritis development in patients with seropositive arthralgia. *Ann Rheum Dis,* 70**,** 128-33.

Van De Stadt, L. A., Witte, B. I., Bos, W. H. & Van Schaardenburg, D. 2012. A prediction rule for the development of arthritis in seropositive arthralgia patients. *Ann Rheum Dis*.

Van Der Woude, D., Houwing-Duistermaat, J. J., Toes, R. E., Huizinga, T. W., Thomson, W., Worthington, J., Van Der Helm-Van Mil, A. H. & De Vries, R. R. 2009. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. *Arthritis Rheum,* 60**,** 916-23.

Van Der Woude, D., Rantapää-Dahlqvist, S., Ioan-Facsinay, A., Onnekink, C., Schwarte, C. M., Verpoort, K. N., Drijfhout, J. W., Huizinga, T. W., Toes, R. E. & Pruijn, G. J. 2010a. Epitope spreading of the anti-citrullinated protein

antibody response occurs before disease onset and is associated with the disease course of early arthritis. *Ann Rheum Dis,* 69**,** 1554-61.

Van Der Woude, D., Syversen, S. W., Van Der Voort, E. I., Verpoort, K. N., Goll, G. L., Van Der Linden, M. P., Van Der Helm-Van Mil, A. H., Van Der Heijde, D. M., Huizinga, T. W., Kvien, T. K. & Toes, R. E. 2010b. The ACPA isotype profile reflects long-term radiographic progression in rheumatoid arthritis. *Ann Rheum Dis,* 69**,** 1110-6.

Van Dongen, H., Van Aken, J., Lard, L. R., Visser, K., Ronday, H. K., Hulsmans, H. M., Speyer, I., Westedt, M. L., Peeters, A. J., Allaart, C. F., Toes, R. E., Breedveld, F. C. & Huizinga, T. W. 2007. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. *Arthritis Rheum,* 56**,** 1424-32.

Van Oene, M., Wintle, R. F., Liu, X., Yazdanpanah, M., Gu, X., Newman, B., Kwan, A., Johnson, B., Owen, J., Greer, W., Mosher, D., Maksymowych, W., Keystone, E., Rubin, L. A., Amos, C. I. & Siminovitch, K. A. 2005. Association of the lymphoid tyrosine phosphatase R620W variant with rheumatoid arthritis, but not Crohn's disease, in Canadian populations. *Arthritis Rheum,* 52**,** 1993-8.

Van Oosterhout, M., Bajema, I., Levarht, E. W., Toes, R. E., Huizinga, T. W. & Van Laar, J. M. 2008. Differences in synovial tissue infiltrates between anti-cyclic citrullinated peptide-positive rheumatoid arthritis and anti-cyclic citrullinated peptide-negative rheumatoid arthritis. *Arthritis Rheum,* 58**,** 53-60.

Van Venrooij, W. J., Hazes, J. M. & Visser, H. 2002. Anticitrullinated protein/peptide antibody and its role in the diagnosis and prognosis of early rheumatoid arthritis. *Neth J Med,* 60**,** 383-8.

Veillard, F., Sztukowska, M., Mizgalska, D., Ksiazek, M., Houston, J., Potempa, B., Enghild, J. J., Thogersen, I. B., Gomis-Ruth, F. X., Nguyen, K. A. & Potempa, J. 2013. Inhibition of gingipains by their profragments as the mechanism protecting Porphyromonas gingivalis against premature activation of secreted proteases. *Biochim Biophys Acta,* 1830**,** 4218-28.

Verlinde, C. L. & Hol, W. G. 1994. Structure-based drug design: progress, results and challenges. *Structure,* 2**,** 577-87.

Verpoort, K. N., Cheung, K., Ioan-Facsinay, A., Van Der Helm-Van Mil, A. H., De Vries-Bouwstra, J. K., Allaart, C. F., Drijfhout, J. W., De Vries, R. R., Breedveld, F. C., Huizinga, T. W., Pruijn, G. J. & Toes, R. E. 2007. Fine specificity of the anti-citrullinated protein antibody response is influenced by the shared epitope alleles. *Arthritis Rheum,* 56**,** 3949-52.

Vignali, D. A., Moreno, J., Schiller, D. & Hammerling, G. J. 1992. Species-specific binding of CD4 to the beta 2 domain of major histocompatibility complex class II molecules. *J Exp Med,* 175**,** 925-32.

Vincent, C., Simon, M., Sebbag, M., Girbal-Neuhauser, E., Durieux, J. J., Cantagrel, A., Fournie, B., Mazieres, B. & Serre, G. 1998. Immunoblotting detection of autoantibodies to human epidermis filaggrin: a new diagnostic test for rheumatoid arthritis. *J Rheumatol,* 25**,** 838-46.

Vincents, B., Guentsch, A., Kostolowska, D., Von Pawel-Rammingen, U., Eick, S., Potempa, J. & Abrahamson, M. 2011. Cleavage of IgG1 and IgG3 by gingipain K from Porphyromonas gingivalis may compromise host defense in progressive periodontitis. *FASEB J,* 25**,** 3741-50.

Vitkov, L., Klappacher, M., Hannig, M. & Krautgartner, W. D. 2010. Neutrophil fate in gingival crevicular fluid. *Ultrastruct Pathol,* 34**,** 25-30.

Vossenaar, E. R., Nijenhuis, S., Helsen, M. M., Van Der Heijden, A., Senshu, T., Van Den Berg, W. B., Van Venrooij, W. J. & Joosten, L. A. 2003. Citrullination of synovial proteins in murine models of rheumatoid arthritis. *Arthritis Rheum,* 48**,** 2489-500.

Vossenaar, E. R., Radstake, T. R., Van Der Heijden, A., Van Mansum, M. A., Dieteren, C., De Rooij, D. J., Barrera, P., Zendman, A. J. & Van Venrooij, W. J. 2004. Expression and activity of citrullinating peptidylarginine deiminase enzymes in monocytes and macrophages. *Ann Rheum Dis,* 63**,** 373-81.

Waaler, E. 2007. On the occurrence of a factor in human serum activating the specific agglutintion of sheep blood corpuscles. 1939. *APMIS,* 115**,** 422-38; discussion 439.

Walport, M. J. 2001a. Complement. First of two parts. *N Engl J Med,* 344**,** 1058-66.

Walport, M. J. 2001b. Complement. Second of two parts. *N Engl J Med,* 344**,** 1140-4.

Wang, F., Chen, F. F., Gao, W. B., Wang, H. Y., Zhao, N. W., Xu, M., Gao, D. Y., Yu, W., Yan, X. L., Zhao, J. N. & Li, X. J. 2016. Identification of citrullinated peptides in the synovial fluid of patients with rheumatoid arthritis using LC-MALDI-TOF/TOF. *Clin Rheumatol*.

Wang, L., Guan, N., Jin, Y., Lin, X. & Gao, H. 2015. Subcutaneous vaccination with Porphyromonas gingivalis ameliorates periodontitis by modulating Th17/Treg imbalance in a murine model. *Int Immunopharmacol,* 25**,** 65-73.

Wang, Y., Li, M., Stadler, S., Correll, S., Li, P., Wang, D., Hayama, R., Leonelli, L., Han, H., Grigoryev, S. A., Allis, C. D. & Coonrod, S. A. 2009. Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J Cell Biol,* 184**,** 205-13.

Warnock, M. G. & Goodacre, J. A. 1997. Cryptic T-cell epitopes and their role in the pathogenesis of autoimmune diseases. *Br J Rheumatol,* 36**,** 1144-50.

Wegner, N., Wait, R., Sroka, A., Eick, S., Nguyen, K. A., Lundberg, K., Kinloch, A., Culshaw, S., Potempa, J. & Venables, P. J. 2010. Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and α-enolase: implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum,* 62**,** 2662-72.

Weinberg, A., Belton, C. M., Park, Y. & Lamont, R. J. 1997. Role of fimbriae in Porphyromonas gingivalis invasion of gingival epithelial cells. *Infect Immun,* 65**,** 313-6.

Wevers-De Boer, K., Visser, K., Heimans, L., Ronday, H. K., Molenaar, E., Groenendael, J. H., Peeters, A. J., Westedt, M. L., Collee, G., De Sonnaville, P. B., Grillet, B. A., Huizinga, T. W. & Allaart, C. F. 2012. Remission induction therapy with methotrexate and prednisone in patients with early rheumatoid and undifferentiated arthritis (the IMPROVED study). *Ann Rheum Dis,* 71**,** 1472-7.

Wilensky, A., Potempa, J., Houri-Haddad, Y. & Shapira, L. 2016. Vaccination with recombinant RgpA peptide protects against Porphyromonas gingivalis-induced bone loss. *J Periodontal Res*.

Willemze, A., Böhringer, S., Knevel, R., Levarht, E. W., Stoeken-Rijsbergen, G., Houwing-Duistermaat, J. J., Van Der Helm-Van Mil, A. H., Huizinga, T. W.,

Toes, R. E. & Trouw, L. A. 2012. The ACPA recognition profile and subgrouping of ACPA-positive RA patients. *Ann Rheum Dis,* 71**,** 268-74.

Williams, R. O., Feldmann, M. & Maini, R. N. 1992. Anti-tumor necrosis factor ameliorates joint disease in murine collagen-induced arthritis. *Proc Natl Acad Sci U S A,* 89**,** 9784-8.

Willis, V. C., Gizinski, A. M., Banda, N. K., Causey, C. P., Knuckley, B., Cordova, K. N., Luo, Y., Levitt, B., Glogowska, M., Chandra, P., Kulik, L., Robinson, W. H., Arend, W. P., Thompson, P. R. & Holers, V. M. 2011. N-alpha-benzoyl-N5-(2-chloro-1-iminoethyl)-L-ornithine amide, a protein arginine deiminase inhibitor, reduces the severity of murine collagen-induced arthritis. *J Immunol,* 186**,** 4396-404.

Winter, G., Lobley, C. M. & Prince, S. M. 2013. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr,* 69**,** 1260-73.

Woods, A., Chen, H. Y., Trumbauer, M. E., Sirotina, A., Cummings, R. & Zaller, D. M. 1994. Human major histocompatibility complex class II-restricted T cell responses in transgenic mice. *J Exp Med,* 180**,** 173-81.

Wucherpfennig, K. W., Yu, B., Bhol, K., Monos, D. S., Argyris, E., Karr, R. W., Ahmed, A. R. & Strominger, J. L. 1995. Structural basis for major histocompatibility complex (MHC)-linked susceptibility to autoimmunity: charged residues of a single MHC binding pocket confer selective presentation of self-peptides in pemphigus vulgaris. *Proc Natl Acad Sci U S A,* 92**,** 11935-9.

Ximenez-Fyvie, L. A., Haffajee, A. D. & Socransky, S. S. 2000. Comparison of the microbiota of supra- and subgingival plaque in health and periodontitis. *J Clin Periodontol,* 27**,** 648-57.

Xu, H., Liew, L. N., Kuo, I. C., Huang, C. H., Goh, D. L. & Chua, K. Y. 2008. The modulatory effects of lipopolysaccharide-stimulated B cells on differential T-cell polarization. *Immunology,* 125**,** 218-28.

Yao, H., Li, P., Venters, B. J., Zheng, S., Thompson, P. R., Pugh, B. F. & Wang, Y. 2008. Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis. *J Biol Chem,* 283**,** 20060-8.

Yilmaz, O., Watanabe, K. & Lamont, R. J. 2002. Involvement of integrins in fimbriae-mediated binding and invasion by Porphyromonas gingivalis. *Cell Microbiol,* 4**,** 305-14.

Young, A. & Koduri, G. 2007. Extra-articular manifestations and complications of rheumatoid arthritis. *Best Pract Res Clin Rheumatol,* 21**,** 907-27.

Young, B. J., Mallya, R. K., Leslie, R. D., Clark, C. J. & Hamblin, T. J. 1979. Anti-keratin antibodies in rheumatoid arthritis. *Br Med J,* 2**,** 97-9.

Ytterberg, A. J., Joshua, V., Reynisdottir, G., Tarasova, N. K., Rutishauser, D., Ossipova, E., Haj Hensvold, A., Eklund, A., Skold, C. M., Grunewald, J., Malmstrom, V., Jakobsson, P. J., Ronnelid, J., Padyukov, L., Zubarev, R. A., Klareskog, L. & Catrina, A. I. 2015. Shared immunological targets in the lungs and joints of patients with rheumatoid arthritis: identification and validation. *Ann Rheum Dis,* 74**,** 1772-7.

Yun, L. W., Decarlo, A. A., Collyer, C. & Hunter, N. 2003. Enhancement of Th2 pathways and direct activation of B cells by the gingipains of Porphyromonas gingivalis. *Clin Exp Immunol,* 134**,** 295-302.

Yun, P. L., Decarlo, A. A. & Hunter, N. 1999. Modulation of major histocompatibility complex protein expression by human gamma

interferon mediated by cysteine proteinase-adhesin polyproteins of Porphyromonas gingivalis. *Infect Immun,* 67**,** 2986-95.

Zhang, J., Dai, J., Zhao, E., Lin, Y., Zeng, L., Chen, J., Zheng, H., Wang, Y., Li, X., Ying, K., Xie, Y. & Mao, Y. 2004. cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type VI. *Acta Biochim Pol,* 51**,** 1051-8.

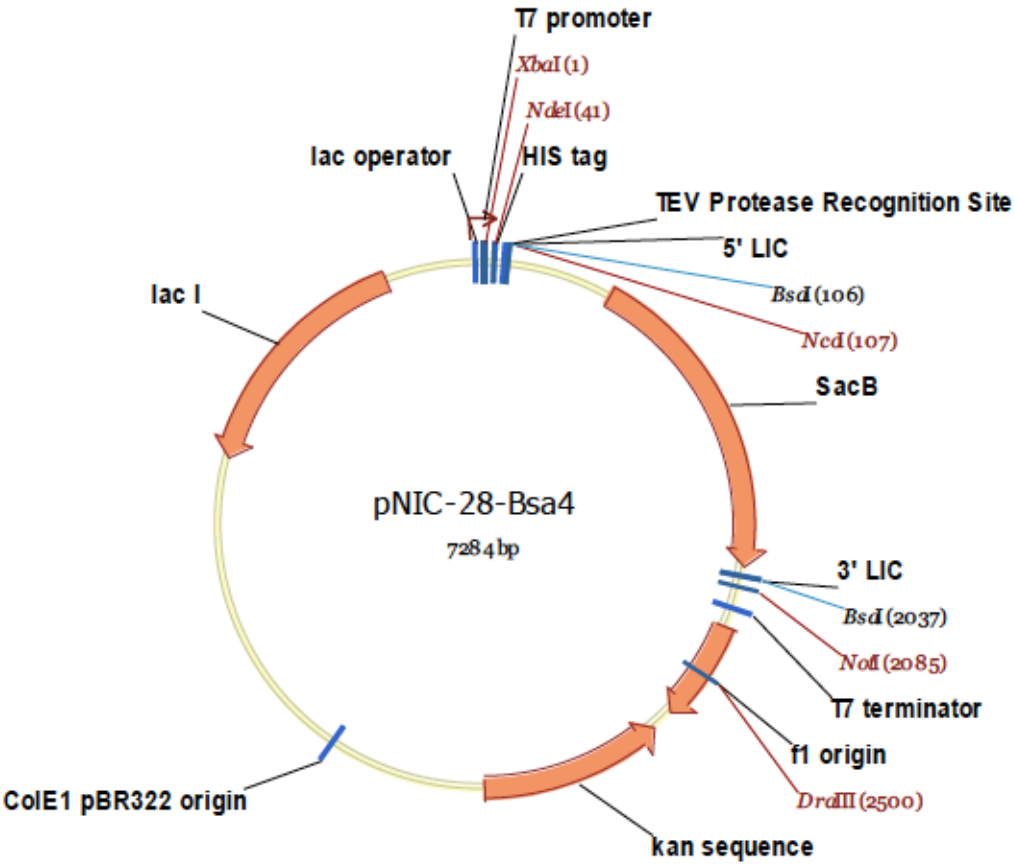
Zhao, J., Zhao, Y., He, J., Jia, R. & Li, Z. 2008. Prevalence and significance of anti-peptidylarginine deiminase 4 antibodies in rheumatoid arthritis. *J Rheumatol,* 35**,** 969-74.

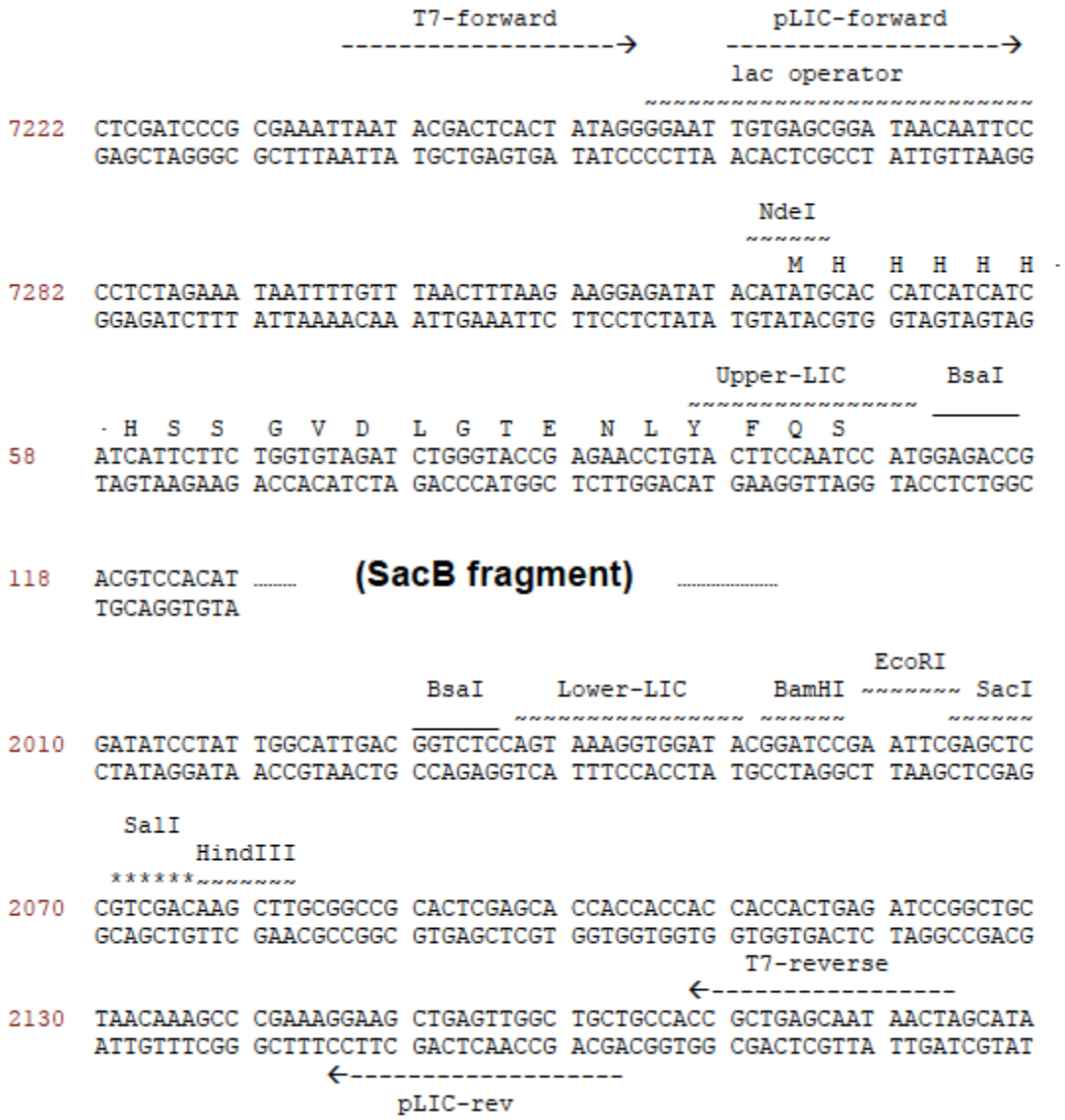
Zhou, X. Y., Gao, J. L., Hunter, N., Potempa, J. & Nguyen, K. A. 2013. Sequence-independent processing site of the C-terminal domain (CTD) influences maturation of the RgpB protease from Porphyromonas gingivalis. *Mol Microbiol,* 89**,** 903-17.

Appendix

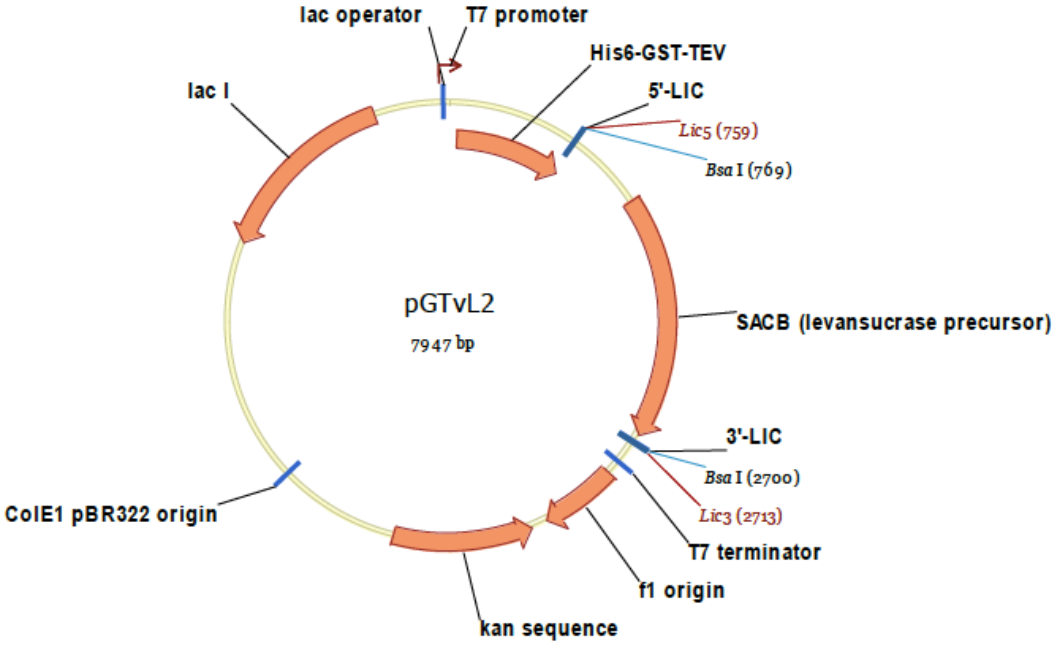
Vectors

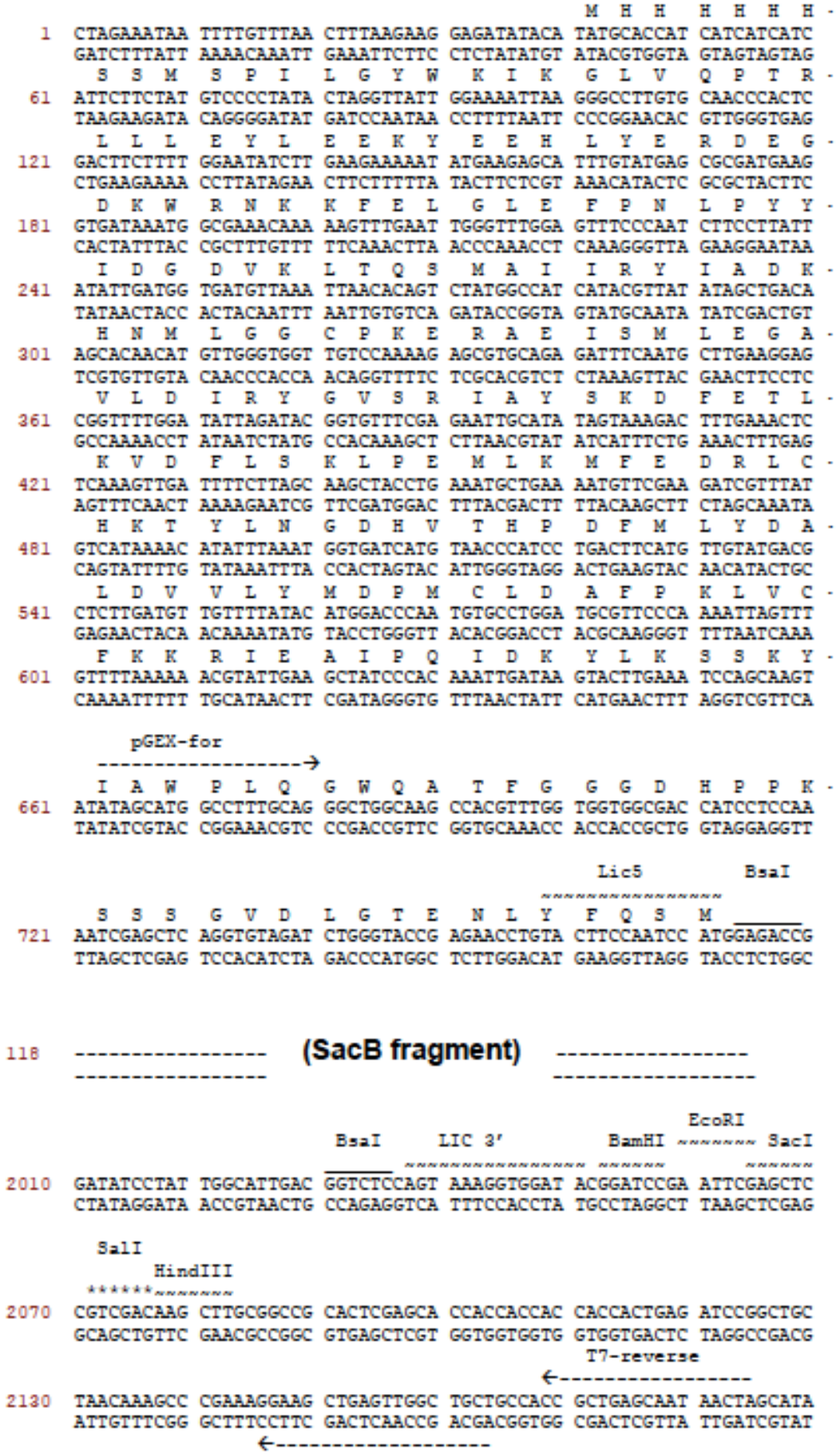
Name	pNIC28-Bsa4
Source	Opher Gileadi (SGC)
Sequence accession/link	Genebank EF198106
Description	pET expression vector with his_6 tag in 22-aa N-terminal fusion peptide, with TEV protease cleavage site. Includes sites for LIC and a 'stuffer' fragment that includes the SacB gene for negative selection on 5% sucrose
Antibiotic resistance	Kanamycin, 50µg/ml
Promoter	T7- LacO
Cloning	LIC
Initiation Codon	Supplied in PCR primer
N-terminal fusion sequence	MHHHHHHSSGVDLGTENLFQ*SM * = TEV cleavage site
N-terminal fusion MW	2684.1 Da including Met (2465.8 Da removed by TEV)
Termination codons	supplied in PCR primer
Protease cleavage	TEV
Preferred host	DE3 hosts: BL21, Rosetta, etc. Must express T7 RNA polymerase
5' sequencing primer	pLIC- for. TGTGAGCGGATAACAATTCC
3' sequencing primer	pLIC- rev. AGCAGCCAACTCAGCTTCC
Upstream LIC primer	add TACTTCCAATCCATG to the 5' end (ATG in-frame with the desired coding sequence)
Downstream LIC primer	add TATCCACCTTTACTG to the 5' end of downstream primer (add termination codon if necessary)


T7 promoter
XbaI (1)
NdeI (41)
lac operator
HIS tag
TEV Protease Recognition Site
5' LIC
BsaI (106)
NcoI (107)
lac I
SacB
pNIC-28-Bsa4
7284 bp
3' LIC
BsaI (2037)
NotI (2085)
T7 terminator
f1 origin
DraIII (2500)
ColE1 pBR322 origin
kan sequence


T7-forward
pLIC-forward
lac operator
NdeI
M H H H H H
Upper-LIC
BsaI
H S S G V D L G T E N L Y F Q S
(SacB fragment)
EcoRI
BsaI
Lower-LIC
BamHI
SacI
SalI
HindIII
T7-reverse
pLIC-rev
7222 CTCGATCCCG CGAAATTAAT ACGACTCACT ATAGGGGAAT TGTGAGCGGA TAACAATTCC
GAGCTAGGGC GCTTTAATTA TGCTGAGTGA TATCCCCTTA ACACTCGCCT ATTGTTAAGG
7282 CCTCTAGAAA TAATTTTGTT TAACTTTAAG AAGGAGATAT ACATATGCAC CATCATCATC
GGAGATCTTT ATTAAAACAA ATTGAAATTC TTCCTCTATA TGTATACGTG GTAGTAGTAG
58 ATCATTCTTC TGGTGTAGAT CTGGGTACCG AGAACCTGTA CTTCCAATCC ATGGAGACCG
TAGTAAGAAG ACCACATCTA GACCCATGGC TCTTGGACAT GAAGGTTAGG TACCTCTGGC
118 ACGTCCACAT
TGCAGGTGTA
2010 GATATCCTAT TGGCATTGAC GGTCTCCAGT AAAGGTGGAT ACGGATCCGA ATTCGAGCTC
CTATAGGATA ACCGTAACTG CCAGAGGTCA TTTCCACCTA TGCCTAGGCT TAAGCTCGAG
2070 CGTCGACAAG CTTGCGGCCG CACTCGAGCA CCACCACCAC CACCACTGAG ATCCGGCTGC
GCAGCTGTTC GAACGCCGGC GTGAGCTCGT GGTGGTGGTG GTGGTGACTC TAGGCCGACG
2130 TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA
ATTGTTTCGG GCTTTCCTTC GACTCAACCG ACGACGGTGG CGACTCGTTA TTGATCGTAT

Name	pGTVL2
Source	Pavel Savitsky (SGC)
Sequence accession/link	N/A
Description	pET expression vector with His_6 and GST tag at the N-terminal fusion, followed by TEC protease cleavage site
Antibiotic resistance	Kanamycin 50µg/ml
Promoter	T7 LacO
Cloning	LIC
Initiation Codon	Supplied in PCR primer
N-terminal fusion sequence	MHHHHHHSSMSPILGYWKIKGLVQPTRLLEYLEEK YEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD VKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGA VLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDP MCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPL QGWQATFGGGDHPPKSSSGVDLGTENLYFQ*SM *= TEV cleavage site
N-terminal fusion MW	28441 Da including Met
Termination codons	Supplied in PCR primer
Protease cleavage	TEV
Preferred host	DE3: BL21, Rosetta etc. Must express T7 RNA polymerase
5' sequencing primer	pGEX for ATAGCATGGCCTTGCAGG
3' sequencing primer	pLIC rev AGCAGCCAACTCAGCTTCC or T7R
Upstream LIC primer	add TACTTCCAATCCATG to the 5' end (ATG in-frame with the desired coding sequence)
Downstream LIC primer	add TATCCACCTTTACTG to the 5' end of downstream primer (add termination codon if necessary)


lac operator
T7 promoter
His6-GST-TEV
5'-LIC
Lic5 (759)
Bsa I (769)
lac I
pGTvL2
7947 bp
SACB (levansucrase precursor)
3'-LIC
Bsa I (2700)
Lic3 (2713)
T7 terminator
f1 origin
kan sequence
ColE1 pBR322 origin


M H H H H H H
1 CTAGAAATAA TTTTGTTTAA CTTTAAGAAG GAGATATACA TATGCACCAT CATCATCATC
GATCTTTATT AAAACAAATT GAAATTCTTC CTCTATATGT ATACGTGGTA GTAGTAGTAG
S S M S P I L G Y W K I K G L V Q P T R
61 ATTCTTCTAT GTCCCCTATA CTAGGTTATT GGAAAATTAA GGGCCTTGTG CAACCCACTC
TAAGAAGATA CAGGGGATAT GATCCAATAA CCTTTTAATT CCCGGAACAC GTTGGGTGAG
L L L E Y L E E K Y E E H L Y E R D E G
121 GACTTCTTTT GGAATATCTT GAAGAAAAAT ATGAAGAGCA TTTGTATGAG CGCGATGAAG
CTGAAGAAAA CCTTATAGAA CTTCTTTTTA TACTTCTCGT AAACATACTC GCGCTACTTC
D K W R N K K F E L G L E F P N L P Y Y
181 GTGATAAATG GCGAAACAAA AAGTTTGAAT TGGGTTTGGA GTTTCCCAAT CTTCCTTATT
CACTATTTAC CGCTTTGTTT TTCAAACTTA ACCCAAACCT CAAAGGGTTA GAAGGAATAA
I D G D V K L T Q S M A I I R Y I A D K
241 ATATTGATGG TGATGTTAAA TTAACACAGT CTATGGCCAT CATACGTTAT ATAGCTGACA
TATAACTACC ACTACAATTT AATTGTGTCA GATACCGGTA GTATGCAATA TATCGACTGT
H N M L G G C P K E R A E I S M L E G A
301 AGCACAACAT GTTGGGTGGT TGTCCAAAAG AGCGTGCAGA GATTTCAATG CTTGAAGGAG
TCGTGTTGTA CAACCCACCA ACAGGTTTTC TCGCACGTCT CTAAAGTTAC GAACTTCCTC
V L D I R Y G V S R I A Y S K D F E T L
361 CGGTTTTGGA TATTAGATAC GGTGTTTCGA GAATTGCATA TAGTAAAGAC TTTGAAACTC
GCCAAAACCT ATAATCTATG CCACAAAGCT CTTAACGTAT ATCATTTCTG AAACTTTGAG
K V D F L S K L P E M L K M F E D R L C
421 TCAAAGTTGA TTTTCTTAGC AAGCTACCTG AAATGCTGAA AATGTTCGAA GATCGTTTAT
AGTTTCAACT AAAAGAATCG TTCGATGGAC TTTACGACTT TTACAAGCTT CTAGCAAATA
H K T Y L N G D H V T H P D F M L Y D A
481 GTCATAAAAC ATATTTAAAT GGTGATCATG TAACCCATCC TGACTTCATG TTGTATGACG
CAGTATTTTG TATAAATTTA CCACTAGTAC ATTGGGTAGG ACTGAAGTAC AACATACTGC
L D V V L Y M D P M C L D A F P K L V C
541 CTCTTGATGT TGTTTTATAC ATGGACCCAA TGTGCCTGGA TGCGTTCCCA AAATTAGTTT
GAGAACTACA ACAAAATATG TACCTGGGTT ACACGGACCT ACGCAAGGGT TTTAATCAAA
F K K R I E A I P Q I D K Y L K S S K Y
601 GTTTTAAAAA ACGTATTGAA GCTATCCCAC AAATTGATAA GTACTTGAAA TCCAGCAAGT
CAAAATTTTT TGCATAACTT CGATAGGGTG TTTAACTATT CATGAACTTT AGGTCGTTCA
pGEX-for
I A W P L Q G W Q A T F G G G D H P P K
661 ATATAGCATG GCCTTTGCAG GGCTGGCAAG CCACGTTTGG TGGTGGCGAC CATCCTCCAA
TATATCGTAC CGGAAACGTC CCGACCGTTC GGTGCAAACC ACCACCGCTG GTAGGAGGTT
Lic5 BsaI
S S S G V D L G T E N L Y F Q S M
721 AATCGAGCTC AGGTGTAGAT CTGGGTACCG AGAACCTGTA CTTCCAATCC ATGGAGACCG
TTAGCTCGAG TCCACATCTA GACCCATGGC TCTTGGACAT GAAGGTTAGG TACCTCTGGC
118 ---------- (SacB fragment) ----------
---------- ----------
EcoRI
BsaI LIC 3' BamHI SacI
2010 GATATCCTAT TGGCATTGAC GGTCTCCAGT AAAGGTGGAT ACGGATCCGA ATTCGAGCTC
CTATAGGATA ACCGTAACTG CCAGAGGTCA TTTCCACCTA TGCCTAGGCT TAAGCTCGAG
SalI
HindIII
2070 CGTCGACAAG CTTGCGGCCG CACTCGAGCA CCACCACCAC CACCACTGAG ATCCGGCTGC
GCAGCTGTTC GAACGCCGGC GTGAGCTCGT GGTGGTGGTG GTGGTGACTC TAGGCCGACG
T7-reverse
2130 TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA
ATTGTTTCGG GCTTTCCTTC GACTCAACCG ACGACGGTGG CGACTCGTTA TTGATCGTAT

Primers

Wild-type constructs	**LIC primers**	
Forward	1	TACTTCCAATCCATGAAAAAGCTTTTACAGGCTAAAGCCTTG
	2	TACTTCCAATCCATGCAACGAGCATTCCAGGAAAC
	3	TACTTCCAATCCATGAATCCCCCTGCAGGTCCTG
	4	TACTTCCAATCCATGATCGCTGAGTACGAACGCTC
Reverse	1	TATCCACCTTTACTGTCATTTGAGAATTTTCATTGTCTCACGGA
	2	TATCCACCTTTACTGTCAGTGCTTGATATAGAGATAGCCCTTATC
	3	TATCCACCTTTACTGTCAGCCCTTATCCGCTACCTC
	4	TATCCACCTTTACTGTCAGGCGTTGAACCATGCACGAA
	5	TATCCACCTTTACTGTCAAGTACATGTATTGGTTTCGT
	6	TATCCACCTTTACTGTCAGTGAGTACGACAATGCAGGG
	7	TATCCACCTTTACTGTCATGTTCCTGAAGCCCCTTTGA
Double mutation	**Mega-primers**	
PPAD$^{R152/4Q}$	F	TGGACTTTATTTATAACcagCCTcagCCTAACGATGATGAATTCCC
	R	GGGAATTCATCATCGTTAGGCTGAGGCTGGTTATAAATAAAGTCCA
PPAD$^{R152/4A}$	F	CTTTATTTATAACgccCCTgccCCTAACGATGATGAATTCCC
	R	GGGAATTCATCATCGTTAGGGGCAGGGGCGTTATAAATAAAG
Single mutation	**Mega-primers**	
PPAD^{R152A}	F	CTTTATTTATAAC**GCC**CCTCGACCTAACGATGATGAATTCCC
	R	GGGAATTCATCATCGTTAGGGGCAGGGGCGTTATAAATAAAG
PPAD^{R154A}	F	CTTTATTTATAACCGACCT**GCC**CCTAACGATGATGAATTCCC
	R	GGGAATTCATCATCGTTAGGGGCAGGGGCGTTATAAATAAAG

Published papers

1. **Montgomery AB**, Kopec J, Shrestha L, Thezenas ML, Burgess-Brown NA, Fischer R, Yue W, Venables PJ. **Crystal structure of Porphyromonas gingivalis peptidylarginine deiminase: implications for autoimmunity in rheumatoid arthritis**. Ann Rheum Dis 2015

Research article detailing the 3D crystal structure of PPAD. As a main component of my thesis project, this was entirely my own work with the exception of structural determination, which was carried out by an experienced crystallographer. The main findings are presented in Chapter 2 and Chapter 3 of this thesis.

2. Lugli EB, Correia R, Fischer R, Lundberg K, Bracke KR, **Montgomery AB**, Kessler BM, Brusselle GG, Venables PJ. **Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: implications for the autoimmunity in rheumatoid arthritis.** Arthritis Res Ther. 2015;17(1):9

Research article investigating the expression of citrullinated proteins and PAD enzymes in lung tissue of smokers and non-smokers. My contribution to this work was western blotting of tissues for the presence of citrullinated proteins, and determining the detection levels for the antibodies used. The results are not used in this thesis, as findings were not directly related to my project.

3. **Montgomery AB**, Lugli EB, Venables PJ. **Is citrullination the missing link between periodontal disease and rheumatoid arthritis?** Current Oral Health Reports. 2015;2(1)30-6

Review article investigating citrullination as the link between PD and RA. As lead author I composed the majority of the article, with assistance from EB Lugli- who prepared the main table of results, and editing by PJ Venables. Findings from this research article have been discussed in the introduction of this thesis, and the table by EB Lugli has been used with permission.

4. **Montgomery AB,** Venables PJ, Fisher BA. **The case for measuring antibodies to specific citrullinated antigens.** Expert Rev Clin Immunol. 2013;9(12):1185-92

Review article investigating ACPA fine specificity. As lead author I composed a major proportion of the article with BA Fisher, and editing by PJ Venables. Aspects of these findings have been used in the introduction of this thesis, and a figure drafted by myself has been used (Figure 1.2).

Permission was obtained from all relevant publishing bodies for use of article contents in this thesis.