Protein arginine deiminase 4 (PAD4): current understanding and future therapeutic potential

Justin Jones, Corey Causey, Bryan Knuckley, Jessica L. Slack-Noyes, and Paul R. Thompson
Department of Chemistry & Biochemistry, University of South Carolina, 631 Sumter Street, Columbia, SC 29208

Abstract
The Protein Arginine Deiminases (PAD), and in particular PAD4, have emerged over the last several years as potential therapeutic targets for the treatment of Rheumatoid Arthritis (RA). In this review, we discuss the current evidence linking dysregulated PAD activity to the onset and progression of RA, as well as its potential role in other human diseases, e.g. cancer and multiple sclerosis. We additionally describe the known physiological roles of the PADs, focusing on PAD4, as well as the current state of knowledge regarding PAD structure, catalysis, and inhibition.

Keywords
Protein Arginine Deiminase; Rheumatoid Arthritis; Multiple Sclerosis; Cancer; Treatment; Inhibitors

By expanding the relatively small number of proteinaceous amino acids, protein posttranslational modifications (PTMs) provide novel functional groups that can be used to regulate protein function. One increasingly prominent example, among the more than 100 known PTMs, is citrullination. Although first reported by Rogers and Simmonds in 1958 [1], a new appreciation for this modification began in 1998 when Schellekens et al demonstrated that Rheumatoid Arthritis (RA) patients produce autoantibodies that recognize citrullinated proteins [2]. Since then, the role of citrullination in the pathogenesis of RA, and many other disease states, has become increasingly apparent; as a result, the enzymes that catalyze this modification (Figure 1), i.e. the Protein Arginine Deiminases (PADs), and in particular PAD4, are now widely regarded as excellent therapeutic targets.

In humans, and other mammals, there are five highly related calcium-dependent PADs, which, for historical reasons, have been designated as PADs 1-4 and PAD6. While the primary structure of these enzymes is highly conserved (inter-isozyme conservation is ~50%), they exhibit tissue-specific expression patterns (Table 1) [3]. For example, PAD4 is primarily expressed in the cells of the immune system as well as several cancer cell lines and tumors (Table 1; [3-6]). The subcellular localization of the PADs has also been characterized; PADs 1-3 and PAD6 are principally cytoplasmic enzymes [3,7], whereas PAD4 localizes to both cytoplasmic granules and the nucleus [3,7]. Known substrates for the PADs are summarized in Table 1. While the substrate scope may appear limited, it is only beginning to be understood.
Role of PADs in Disease

Role of PADs in RA

Over the last 10 years, a significant body of evidence linking dysregulated PAD activity, most prominently PAD4, to the onset and progression of RA has accumulated – RA is a systemic autoimmune disorder that primarily causes synovial joint inflammation. More than 50 years ago it was recognized that RA patients produce a number of autoantibodies, including rheumatoid factor (RF), anti-perinuclear factor (APF), and anti-keratin autoantibodies (AKA). Although RF is present in 70-80% of RA patients [8], this autoantibody is also observed in a number of other inflammatory diseases, e.g. ankylosing spondylitis; thus, their specificity is quite low [8]. In contrast, APF and AKA are highly specific to RA [8]. It took, however, ~40 years to discover that these autoantibodies recognize citrulline-containing proteins (e.g., fibrin) [2]. It is now recognized that multiple intracellular and extracellular proteins, most of which are unknown, are deiminated within the RA synovium, and higher levels of these citrulline-containing proteins correlate with increased joint destruction [9]. This discovery led to the development of a novel diagnostic that detects the presence of these anti-citrulline autoantibodies by screening sera against a library of citrullinated cyclic peptides (CCP) [10]. This ELISA based assay, the anti-CCP test, provides a highly sensitive and specific method to diagnose RA [2,10]. Significantly, anti-CCP autoantibodies can be detected >10 years prior to clinical disease onset, are present in patients with RF negative RA, and are associated with a more severe and erosive form of the disease [2,10,11].

The link between abnormal protein citrullination and RA was strengthened when a genome-wide association study identified an RA-associated PAD4 haplotype that consists of 4 exonic single nucleotide polymorphisms, resulting in 3 amino acid substitutions: S55G, A82V, and A112G (the fourth polymorphism is silent) [12]. Although the odds ratio for this haplotype is quite low (OR = 1.14), it was associated with anti-CCP antibody levels and increased expression of PAD4 in the RA synovium. While the association between PAD4 mutations and RA is observed in Asian populations, its role in Caucasian populations is somewhat controversial because associations between PAD4 and RA have been observed in some (Canadian, French, and German), but not all (England, France, Hungary Spain) studies (see refs. [13,14] for a more thorough discussion of this controversy). The reasons for these discrepancies are unknown but may relate to the relatively low odds ratio.

Besides PAD4, additional RA-associated genetic loci include PTPN22, a protein tyrosine phosphatase that regulates T cell receptor signaling, and the HLA DR shared epitope, e.g. HLA-DRB1*0401. How these different loci contribute to RA is not clear in all cases, but it is noteworthy that the HLA-DRB1*0401 allele preferentially binds and presents citrullinated peptides to the immune system [15], providing a possible link between dysregulated PAD4 activity and this genetic locus. Also consistent with such a possibility is the observation that the association between anti-CCP antibodies and RA is strongest in patients who express the HLA “shared epitope” [16].

In total, these data suggest that increased PAD4 activity contributes to the onset and progression of RA. One hypothesis is that PAD4 expressing white blood cells are recruited to sites of inflammation, become activated, and release PAD4, through an unknown mechanism, to the surrounding tissue. The enzyme can then deiminate extracellular proteins, creating a neoantigen(s) for which an immune response is mounted in those who are genetically pre-disposed to RA. These neoantigens are likely expressed in articular tissues, thereby explaining why the clinical manifestations of RA principally affect the joints. The fact that PAD4 is expressed in immune cells [17] recruited to sites of inflammation further suggests that the recruitment and activation of these cells ultimately leads to a vicious circle
that perpetuates the disease. Consistent with such a mechanism, extracellular proteins in the RA synovium show increased levels of deimination [18]. Even though extracellular PAD activity has not been demonstrated, PAD4 has been detected in the sera of diseased patients [6,19]. Although the evidence discussed above favors a role for PAD4 in RA pathogenesis, PAD2 may also be involved as this isozyme is expressed in a subset of immune cells, e.g., macrophages, and is, like PAD4, detected in the RA synovium [20,21]. It should also be emphasized that it is not known whether one or both isoforms are responsible for generating the immunogenic citrullinated epitopes. Although the association between dysregulated PAD activity and RA is clear, the specific factors that cause the dysregulation remain unknown. Nonetheless, a number of environmental (e.g., smoking) and infectious agents could be possible triggers [22]. Uncovering these factors will undoubtedly be an exciting avenue for future research.

In addition to their direct role in creating neoantigens, PADs may also influence the immune response by triggering the release of Neutrophil Extracellular Traps (NETs) and regulating both antithrombin and chemokine activity. For example, in response to various stimuli, e.g., TNF α and lipopolysaccharide, a subset of neutrophils are known to rupture and form a net-like structure that is thought to physically immobilize pathogens [23,24]. PAD activity is critical for NET formation [23,24], and this process may provide a mechanism for PAD4 release into the extracellular milieu. With regard to antithrombin, once citrullinated, this protein no longer inhibits proteases involved in the blood coagulation cascade (e.g., thrombin) [25]. Significantly, the levels of citrullinated antithrombin are elevated in the synovial fluid of patients with RA [25]; in the joint, elevated thrombin activity would be expected to lead to abnormal fibrin deposition and angiogenesis, both of which are hallmarks of RA.

The deimination of several chemokines (e.g., CXCLs 5, 8, and 10-12) has also been reported, and this modification influences their ability to recruit and/or activate lymphocytes [26-28]. For example, CXCL8, purified from blood derived leukocytes, is deiminated at R5, and this modification reduces its ability to increase intracellular calcium concentrations and activate MAPK signaling cascades. In addition, CXCL8 signaling activity is normally potentiated by the thrombin or plasmin catalyzed proteolytic removal of its first 5 amino acids; deimination at R5 inhibits this process. These effects on cell signaling likely manifest themselves in vivo because citrullinated CXCL8 shows reduced chemotactic activity [27]. Similar effects on chemotaxis and cell signaling have been observed with CXCL10, 11, and 12 [26,28]. The modification of CXCL10 and 12 are particularly relevant to RA because the levels of these chemokines are elevated during inflammation [26]. Although these data suggest that PAD activity is anti-inflammatory, it is clear that PAD activity is required for NET formation, a pro-inflammatory response. Also, the deimination of anti-thrombin should also be considered pro-inflammatory. Therefore, the abundance of PAD activity in the joint may prevent the mounting of an effective immune response or prevent the proper resolution of that response.

Role of PADs in Multiple Sclerosis

Multiple sclerosis (MS) is a chronic and progressive demyelinating disease that reduces nerve cell communication, leading to a loss of motor function and impaired vision. Links between abnormal protein deimination and MS include the facts that in both MS patients and an animal model of MS (i.e., autoimmune encephalomyelitis (EAE)), myelin basic protein (MBP), a primary component of myelin, is deiminated at levels that are up to 3-fold higher than normal [29]. Significantly, in Marburg MS, a malignant form of the disease, the levels of citrullinated MBP are even higher [30]. Further bolstering the link between dysregulated PAD activity and MS is the observation that the levels of both PAD4 and PAD2 are elevated in MS brains [31]. While speculative, MBP deimination, by PAD2 and /
or PAD4, may alter the tertiary structure of MBP such that it is more susceptible to proteolysis and unable to properly form the multilamellar structure of the myelin sheath, thereby explaining the decreased nerve cell communication [29].

Although the links between abnormal protein deimination and MS appear strong, it is noteworthy that autoantibodies to citrullinated proteins have not been described in MS patients. It is also noteworthy that PAD2 knockout mice still develop EAE [32]. This animal model, however, does not faithfully recapitulate all aspects of human MS, e.g. less demyelination and deimination are observed in this model than in MS. Furthermore, these results do not rule out a role for PAD4 in MS as this enzyme is still present in the brains of PAD2 knockout mice and MBP is deiminated, albeit at lower levels [31]. Finally, transgenic mice designed to overexpress PAD2 develop MS-like symptoms and undergo nerve demyelination, at levels that are similar to those observed in other MS animal models [33]. Thus, a potential role for dysregulated PAD activity in MS remains strong, although the data are by no means definitive and warrant further research.

Role of PADs in Cancer

PAD4 is overexpressed in numerous malignant tumors, but not those that are benign (Table 1), thereby suggesting a role for dysregulated PAD4 activity in cancer progression [6]. Interestingly, PAD4 levels are also elevated in the blood of patients with malignant cancers and, significantly, these levels drop after tumor resection [6]. Mimicking RA, serum citrullinated antithrombin levels are also increased in patients with malignant cancers [6], offering a possible explanation for why increased thrombin activity is observed in numerous cancers [34]. This observation is relevant to oncogenesis because thrombin activity increases the expression of both VEGF and integrin $\beta_3$, thereby contributing to angiogenesis, hyperplasia, and metastasis [34]. We also note that PAD4 efficiently modifies RGD peptides [Knuckley B, Thompson PR: unpublished data], and postulate that this modification, if it occurs in vivo, could reduce cell-cell contacts and further contribute to tumor growth and metastasis. In addition to these effects, which are primarily due to extracellular PAD4 activity, PAD4 is known to act as a transcriptional corepressor for p53 within cells (see below). Thus increased PAD4 expression could contribute to tumorigenesis both intra- and extracellularly. Because inhibition of PAD4 would be expected to reverse all of these processes, examining the effect of PAD4 inhibitors on tumorigenesis is an interesting avenue for future research.

Role of PADs in other diseases

Dysregulated PAD activity has also been implicated, albeit tangentially, in a number of other diseases, including Ankylosing Spondylitis, Osteoarthritis, Ulcerative Colitis, Alzheimer’s Disease, Glaucoma, HIV/AIDS, and even Scrapie [21,26,35-38]. The evidence linking abnormal PAD activity to these diseases is summarized in Table 2. Although definitive roles for dysregulated PAD activity in the above described diseases has not been established, a direct link is plausible. Additionally, the fact that PAD activity appears to be upregulated in each of these diseases suggest that PAD inhibitors are likely to see great utility in treating multiple human diseases.

Physiological roles of PAD4

While our understanding of whether dysregulated PAD activity contributes to human disease has become clearer in recent years, our understanding of the physiological roles of these enzymes is only in its infancy. Because the physiological roles of PAD4 are the best characterized, we will focus our efforts on describing the in vivo roles of this isozyme. Putative roles for the remaining PADs are highlighted in Table 1.
PAD4 was initially discovered as a protein whose expression is increased during the terminal differentiation of HL-60 cells, a leukemic cell line, into granulocytes and monocytes [7], suggesting a possible role for this isozyme in cellular differentiation. Although this role has yet to be established [7], it is clear that PAD4 is expressed in the nucleus and cytoplasmic granules of differentiated neutrophils, and it plays a role in NET formation. In addition to differentiation, PAD4 has also been shown to influence gene expression via its ability to deiminate histones H2A, H3 and H4 [17]. This observation was highly significant because histone modifications (e.g. methylation, phosphorylation, and acetylation) are known to regulate gene expression, DNA repair, and replication by altering the local, and in some instances global, chromatin architecture [39].

Subsequent studies demonstrated that PAD4 is recruited to the promoters of Estrogen Receptor (ER) regulated genes, where it deiminites histones H3 and H4 at R17 and R3 respectively; deimination of these residues correlates with the decreased expression of ER regulated genes [4,5]. Interestingly, these arginines can also be methylated, and this modification is inversely correlated with citrullination. This data suggests the existence of enzymes that can reverse both arginine methylation and citrullination. Although PAD4 has been suggested to convert methylated arginines directly into citrulline [5], this activity is unlikely to be physiologically relevant [40-43]. Thus, these two modifications appear to be mutually exclusive. While an arginine demethylase has recently been described [44], a mechanism to reverse citrullination remains unknown.

PAD4 has more recently been found to negatively influence the expression of a subset of p53-target genes [45,46] – p53 is a sequence specific transcription factor that regulates cell fate by activating the transcription of genes involved in DNA repair, cell-cycle arrest, senescence, and apoptosis. Among the genes whose expression is influenced by PAD4 are p21, a cell cycle inhibitor, as well as PUMA and OKL38, which are inducers of apoptosis [45,46]. Importantly, the discovery of a role for PAD4 in p53 regulated transcription utilized a recently described PAD4 inhibitor that was developed in our lab [47]. This compound, denoted Cl-amidine, was also instrumental in identifying OKL38 (or pregnancy-induced growth inhibitor) as a novel p53 regulated gene [46]. In MCF7 cells, a breast cancer cell line, the expression of OKL38 is normally low, apparently held in check, at least in part, by the deimination of promoter associated nucleosomes [46]. Inhibition with Cl-amidine or siRNA knockdown of PAD4, increases OKL38 expression, which consequently induces apoptosis in a subset of cells [46]. Although PAD4 appears to act as a transcriptional corepressor for several transcription factors, this enzyme has also been shown to deiminate the GRIP1 binding domain of p300, facilitating its interaction with GRIP1 and activating transcription [48]. Whether PAD4 can act as both a repressor and activator of gene transcription more generally is unknown and requires further research.

The observation that PAD4 inhibition or knockdown induces apoptosis was somewhat surprising given reports that, in response to specific apoptotic inducers, i.e. calcium ionophores, PAD activity is increased and that overexpression of PAD4 can induce apoptosis [49,50]. However, camptothecin and staurosporine, conventional inducers of apoptosis, fail to stimulate PAD activity in HL-60 cells [23]. The data regarding the overexpression of PAD4 is also problematic because long times and high levels of expression were required to observe apoptosis. Overall, a role for PAD4 as an inducer of apoptosis seems unlikely, although it could be cell line dependent.

**Structure and mechanism of PAD4**

Due to its putative role in human disease, PAD4 has been the primary focus of most structure-function work on the PADs. The overall structure of the PAD4 monomer consists
of two immunoglobulin-like subdomains that are linked in series to a C-terminal catalytic domain (Figure 2)[51]. The catalytic domain consists of five \( \beta \beta \alpha \beta \) motifs that form a pseudo-five-fold structure that is termed an \( \alpha/\beta \) propeller. The immunoglobulin-like domains are likely important for regulating PAD4 activity as the first subdomain forms critical contacts with the catalytic domain of a second PAD4 monomer to form a head-to-tail dimer [51]. Although speculative, the two immunoglobulin-like domains may also regulate PAD4 activity via the formation of protein-protein interactions with unknown regulatory proteins. Because the RA associated mutations are located in subdomain 1 (Figure 2), interactions with these putative regulatory proteins and/or the stability of PAD4 itself could be altered and thereby contribute to the dysregulation of PAD4 activity. A direct effect on enzyme activity is unlikely because the mutated residues are not directly involved in dimer formation and are far from the PAD4 active site [51].

Based on the crystallographic data, there are five calcium binding sites; three of these sites are in subdomain 2 and two are in the catalytic domain (Figure 2). The residues that make up these sites are highly conserved among PAD family members, except PAD6, possibly explaining why this isozyme lacks detectable activity [42]. The calcium dependence of PAD4, and the other active PADs, has been characterized and the results indicate that high concentrations of calcium (\( \geq 100 \mu \text{M} \)) are required for activity [40,52]. Calcium activation is cooperative, and the Hill coefficients range from 1.5 to 3.0, indicating that multiple calcium ions are indeed required for enzyme activation [40]. Comparisons of the apoenzyme to calcium-bound PAD4 demonstrate that conformational changes at the sites of calcium binding are transmitted to the active site, where key catalytic residues are moved into positions that are competent for catalysis (Figure 3A)[51]. Interestingly, the in vitro concentration of calcium required for maximal activity is 100 to 1000 fold higher than that observed in activated cells. Although this dependence would have no effect on extracellular PAD4 activity, it is unclear how this enzyme is activated inside cells. It is likely that some type of PTM or interacting protein lowers the concentration of calcium required for intracellular enzyme activation.

Structural and mechanistic studies on PAD4, and other related enzymes, have provided insights into the mechanism of PAD4 catalysis [40,51,53]. Based on these studies, four residues are known to play key catalytic roles, i.e. D350, H471, D473, and C645. Although unproven, deimination likely proceeds via the nucleophilic attack by the active site cysteine, C645, on the guanidinium carbon to form a tetrahedral intermediate, which ultimately collapses to form an S-alkylthiouronium intermediate. Subsequent hydrolysis of this intermediate generates citrulline. Based on mechanistic data, H471 likely acts as both a general acid and general base in this mechanism while the two aspartates orient the guanidinium for nucleophilic attack (Figure 3B).

pH rate profiles and \( \text{p}K_a \) measurements of the active site thiol support the above described mechanism [53]. For example, the \( k_{cat}/K_m \) versus pH rate profiles show a bell shaped curve with a pH optimum of 7.6. The \( \text{p}K_a \)s for the ascending and descending limbs are 7.3 and 8.2, respectively, and based on \( \text{p}K_a \) measurements of the active site thiol, these \( \text{p}K_a \)s likely correspond to H471 and C645, respectively. Although the suggestion that the rate of the reaction would drop as the concentration of the thiolate increased is counterintuitive, this type of mechanism is often observed with thiolate reactive enzymes and is referred to as a ‘reverse protonation’ mechanism. Nevertheless, the thiolate is the reactive species and the loss in activity as the pH is increased is more a reflection of the fact that the \( \text{p}K_a \)s of the H471 and C645 are quite close; deprotonation of His471 would preclude it from acting as a general acid in the first step of the reaction.
Insights into substrate recognition have also been gained from structures of PAD4 bound to peptides whose sequence is based on the N-termini of histones H3 and H4 [54]. When bound to PAD4, these normally flexible histone-tail mimics adopt an ordered β-turn like configuration in which a 5 residue segment is discernable crystallographically. The majority of the contacts occur between the backbone carbonyl groups of the peptide and the side chains of residues surrounding the active site, e.g. Q346, W347, R372, and R374 (Figure 4) [54]. Significantly, R374 hydrogen bonds to the two carbonyl groups immediately surrounding the modified arginine residue, suggesting a critical role for this residue in substrate recognition, which was subsequently confirmed; mutation of this residue results in a ~20 to 50-fold reduction in enzyme activity [54]. The only contacts observed between the enzyme and the side chains of the substrate peptides occur between the side chain amide of Q346 and the side chain hydroxyl group of a serine or threonine at the R-2 position. The lack of significant contacts between the enzyme and functional groups on the substrate may explain why sequence specific substrate recognition elements have been difficult to identify for this enzyme.

**PAD Inhibitor Design and Development**

Due to their potential role in human disease, the interest in identifying PAD inhibitors is high. In recent years, several PAD inhibitors have been described, but most of these compounds are relatively weak (Figure 5). For example, taxol and benzoyl-NG-asymmetrical dimethyl-Arg are high μM to mM PAD4 inhibitors [41,55]. In contrast, our lab has successfully developed highly potent PAD4 inactivators (Figure 5) [47,56]. The best two compounds, i.e. F-amidine and Cl-amidine, incorporate a haloacetamidine warhead in place of a substrate guanidinium group. These compounds are analogues of benzoylarginine amide, the best small molecule substrate identified to date [40]. Note that Fast and colleagues independently reported the use of 2-chloroacetamidine, i.e. the warhead alone, as a PAD4 inactivator [57]. F- and Cl-amidine inactivate the enzyme via the alkylation of C645, which has been confirmed crystallographically [47]. Although the specific mechanism of inactivation has not been elucidated, inactivation likely proceeds through either: (i) direct displacement of the halide by the C645 thiolate, or (ii) initial attack of the thiolate on the iminium carbon, followed by displacement of the halide to form a three-membered sulfonium ring, ending with concomitant collapse of the tetrahedral intermediate and opening of the sulfonium ring. Although the latter mechanism was invoked to account for the poor leaving group potential of the fluoride, recent data from our lab suggests that the inactivation of PAD4 by both F- and Cl-amidine proceeds via this mechanism [Knuckley B, Causey, CP, Thompson, PR: unpublished data].

The inhibitory properties of these compounds have been explored in some detail. For example, the irreversibility of inhibition has been confirmed, and we have shown that these compounds preferentially inactivate the active, i.e. calcium bound, form of the enzyme [47,56]. The compounds are also bioavailable and, as described above, have been quite useful in discerning the roles of PAD4 in various physiological processes, including NET formation and transcriptional regulation [24,45-47,56]. Although the $K_I$ values are relatively high, the best indicator of an inactivator’s potency is $k_{\text{inact}}/K_I$. By this standard, F- and Cl-amidine are the two most potent PAD inhibitors described to date. Nevertheless, the $K_I$ term is instructive, as demonstrated by the fact that a comparison of the inactivation properties of the warhead alone, i.e. 2-chloroacetamidine, to those obtained for Cl-amidine reveal that the increase in $k_{\text{inact}}/K_I$ is driven mostly by a decrease in $K_I$ [47,53,57]. This result suggests that further improvements in inactivator potency can be achieved by enhancing interactions between the enzyme and the non-warhead portions of the molecule. Modifications at these sites will also likely lead to the development of inhibitors with greater selectivity. Increased
selectivity is important because F-amidine and Cl-amidine generally inhibit all of the active PAD isozymes with similar potency [Knuckley B, Thompson PR: unpublished data].

**Chemical Biology of PAD Inhibitors**

The fact that F- and Cl-amidine covalently modify PAD4, as well as other PAD isozymes, suggested the development of PAD targeted chemical probes (Figure 5) that can be used to both advance our understanding of both PAD biology and to identify novel PAD inhibitors. Towards this end, we recently described the synthesis of fluorescently tagged derivatives of F- and Cl-amidine [58]. These compounds, denoted rhodamine-conjugated F- and Cl-amidine (RFA and RCA, respectively) provide a method for the gel-based visualization of the active forms of the enzyme. Since their initial description, several different derivatives of these compounds have been synthesized, e.g. biotin conjugated versions, and they have shown great utility in identifying the active forms of PAD4 in cell extracts [Slack,-Noyes J Thompson PR: unpublished data]. In addition, RFA has been used to screen for PAD inhibitors – this is a competitive assay where the level of inhibition is measured as a function of fluorescent modification. Specifically, this assay was used to determine if Disease Modifying Anti-Rheumatic Drugs (DMARDs) derive their efficacy from PAD4 inhibition [59]. The use of this assay rapidly led to the identification of minocycline as a μM PAD4 inhibitor. Follow up studies further demonstrated that three other tetracycline derivatives inhibit PAD4 with similar potency [59]. Although these compounds are relatively weak mixed-type inhibitors, they represent potential chemical scaffolds that could be further elaborated to develop more potent PAD inhibitors. The potential use of RFA in a fluorescent polarization based high-throughput assay is also being explored, and the initial results look quite promising.

**Future Prospects**

The evidence linking dysregulated PAD activity to RA is strong and the development of inhibitors targeting these enzymes has drawn considerable interest, including our own. Although these links have not been proven, unpublished data from our lab suggests that PAD inhibition is a viable strategy for the treatment of RA. In the future, it will be interesting to explore the use of these inhibitors in other diseases where dysregulated PAD activity is implicated. Although a definitive role for dysregulated PAD activity in these diseases has not been established, a direct link is not only plausible, but likely. It will also be important to define whether or not the inhibition of only PAD4 or, alternatively, multiple PAD isozymes are required for *in vivo* efficacy in RA and other diseases. The development of PAD selective inhibitors should allow the testing of this possibility. Because dysregulated PAD activity is observed both intra- and extracellularly, it will also be of interest to determine whether PAD inhibition in either of these compartments, or even both, is required for efficacy. Finally, the fact that the substrate scope of the PADS and the Protein Arginine Methyltransferases is quite similar, leads to the question of whether diseases such as RA are simply due to dysregulated PAD activity or whether the inability to methylate these arginine residues also contributes to disease. Providing answers to these questions will undoubtedly be exciting avenues of research well into the future.

**Acknowledgments**

This work was supported in part by NIH grant GM079357 to PRT.

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAD</td>
<td>protein arginine deiminase</td>
</tr>
</tbody>
</table>

*Curr Opin Drug Discov Devel. Author manuscript; available in PMC 2013 September 12.*
RA  rheumatoid arthritis  
MS  multiple sclerosis  
PTM  post translational modifications  
RF  rheumatoid factor  
APF  anti-perinuclear factor  
AKA  anti-keratin autoantibodies  
CCP  citrullinated cyclic peptides  
OR  odds ratio  
BAA  benzoyl L-arginine amide  
NET  Neutrophil Extracellular Traps  
EAE  autoimmune encephalomyelitis  
MBP  myelin basic protein  
ER  estrogen receptor  

References


Curr Opin Drug Discov Devel. Author manuscript; available in PMC 2013 September 12.


Figure 1.
PADs catalyze the hydrolytic deimination of protein arginine to produce protein citrulline and ammonia.
Figure 2.
Crystal structure of the PAD4•Ca\(^{2+}\)•Benzoyl Arginine Amide (BAA) complex. The C-terminal catalytic domain is shown in blue. The N-terminal domain contains two immunoglobulin-like subdomains (red and green, respectively). There are five Ca\(^{2+}\) binding sites (black). Insert shows the location of the RA-associated single nucleotide polymorphisms (SNPs), S55G, A82V, and A112G.
**Figure 3.**
Active site of PAD4. A) Position of active site residues in the apoenzyme (dark grey) and calcium-bound form of the enzyme (light grey). B) Proposed catalytic mechanism. Adapted from [53].
Figure 4.
Diagram of PAD4 bound to a histone H3 tail mimic. Critical binding contacts are indicated by the dashed lines. Adapted from [54].
Figure 5.
PAD4 inhibitors. aEnzyme was incubated at 52 °C for 30 min with 50 mM HEPES (pH 7.6), CaCl$_2$ (5 mM), DTT (2 mM), and BAEE (5 mM). bAfter preincubating the enzyme with inhibitor, 100 mM Tris-HCl (pH 7.6), CaCl$_2$ (10 mM), DTT (2 mM), and NaCl (50 mM) at 37 °C for 15 min, BAEE (1 mM final) was added. cReactions were carried out in 100 mM Tris-HCl (pH 7.6), CaCl$_2$ (10 mM), DTT (1 mM), and Bz-Arg (0.1 mM) at 37 °C for 40 or 60 min. dReactions were carried out in 100 mM Tris-HCl (pH 7.6), CaCl$_2$ (10 mM), DTT (5 mM), and BAEE (5 mM) at 37 °C for 10 min. eReactions were carried out in 100 mM HEPES (pH 7.6), CaCl$_2$ (10 mM), TCEP (0.5 mM), NaCl (50 mM), and BAEE (10 mM final) at 37 °C for 15 min.
**Table 1**

PAD Enzymes: substrates, subcellular localization, cell and tissue-specific expression patterns, and physiological roles.

<table>
<thead>
<tr>
<th>PAD Isozyme</th>
<th>Substrates</th>
<th>Subcellular Localization, Cell and Tissue-specific Expression Patterns</th>
<th>Known and Putative Physiological Roles</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAD1</td>
<td>keratin K1</td>
<td>• cytoplasm • epidermis, uterus, keratinocytes</td>
<td>• skin differentiation • terminal differentiation of keratinocytes</td>
<td>[3]</td>
</tr>
<tr>
<td>PAD2</td>
<td>vimentin, myelin basic protein (MBP), glial fibrillary acidic protein (GFAP)</td>
<td>• cytoplasm • skeletal muscle, brain, pancreas, glial cells, macrophages, bone marrow, muscle, breast, colon, embryo, eye, kidney, epidermal, uterus, thymus</td>
<td>• may play a role in brain development</td>
<td>[3], [29], [30], [31]</td>
</tr>
<tr>
<td>PAD3</td>
<td>trichohyalin</td>
<td>• cytoplasm • hair follicles, keratinocytes</td>
<td>• skin differentiation • hair follicle formation • terminal differentiation of keratinocytes</td>
<td>[1], [3]</td>
</tr>
<tr>
<td>PAD4</td>
<td>histones H2A, H3, H4, vimentin, p300</td>
<td>• cytoplasmic granules and nucleus • normal † eosinophils ‡ neutrophils ‡ granulocytes ‡ macrophages ‡ cancerous † cancer cell lines (MCF7, A549, SKOV3, U937) ‡ cancerous tissues (breast carcinomas, lung adenocarcinomas, hepatocellular carcinomas, esophageal squamous cancer cells, colorectal adenocarcinomas, renal cancer cells, ovarian adenocarcinomas, endometrial carcinomas, uterine adenocarcinomas, bladder carcinomas, chondromas)</td>
<td>• cellular differentiation • transcriptional corepressor for the estrogen receptor and p53 • neutrophil extracellular traps (NETs) formation</td>
<td>[3], [4], [5], [6], [7]</td>
</tr>
<tr>
<td>PAD6</td>
<td>unknown</td>
<td>• egg, ovary, early embryo, thymus, oocyte</td>
<td>• embryonic development • oocyte cytoplasmic lattice formation • fertility</td>
<td>[3]</td>
</tr>
</tbody>
</table>
## Table 2
Evidence linking PADs and Disease

<table>
<thead>
<tr>
<th>Disease</th>
<th>Evidence</th>
<th>Refs.</th>
</tr>
</thead>
</table>
| Rheumatoid Arthritis (RA)     | • RA patients produce autoantibodies that recognize citrulline containing proteins.  
                                 | • A genome-wide association screen identified an RA-associated PAD4 haplotype.  
                                 | • PAD2 and PAD4 are overexpressed in the RA synovium.  
                                 | • PADs help trigger the release of Neutrophil Extracellular Traps (NETs) as part of the innate immune response.  
                                 | • The level of citrullinated antithrombin is elevated in synovial fluid of patients with RA.  
                                 | • Chemokine deimination alters chemokine function.                                                                                                                                                     | [2], [12], [21], [23], [24], [25], [26], [27], [28] |
| Multiple Sclerosis (MS)       | • Levels of deiminated myelin basic protein (MBP) are increased in patients with MS.  
                                 | • PAD2 and PAD4 are overexpressed in the brain of patients with MS.  
                                 | • The levels of citrullinated MBP are higher in the malignant form of MS, Marburg MS variant.  
                                 | • Transgenic mice designed to overexpress PAD2 develop MS-like symptoms and undergo nerve demyelination.                                                                                      | [29], [30], [31], [32], [33] |
| Cancer                        | • PAD4 is overexpressed in the blood of patients with malignant tumors (see Table 1).  
                                 | • Elevated serum levels of citrullinated antithrombin may alter thrombin activity.  
                                 | • PAD4 modifies RGD peptides, potentially altering extracellular cell-cell contacts.  
                                 | • PAD4 is a transcriptional corepressor for p53.                                                                                                                                                    | [6], [34], [45], [46] |
| Ankylosing Spondylitis (AS)   | • Levels of PAD2, PAD4, and citrullinated proteins are increased in the joints of patients with AS.                                                                                                        | [21], [36]                                                                                 |
| Osteoarthritis (OA)           | • PAD4 levels are elevated in the joints of patients with OA.                                                                                                                                               | [21], [36]                                                                                 |
| Ulcerative Colitis (UC)       | • The PAD4 RA associated mutations are also associated with UC.  
                                 | • PAD4 protein levels are increased in the inflammatory lesions that are characteristic of UC.                                                                                                          | [35]                                                                                       |
| Alzheimer's disease (AD)      | • Increased levels of PAD2 and deiminated proteins are found in hippocampal extracts of patients with AD.                                                                                                  | [21], [36]                                                                                 |
| Glaucoma                      | • Levels of protein deimination are elevated in the glaucomatous optic nerve.                                                                                                                              | [37]                                                                                       |
| HIV/AIDS                      | • Citrullinated CXCL12 does not bind to CXCR4, a coreceptor for HIV infection. This decreased binding fails to inhibit HIV infection of cells that are normally protected by this chemokine.                                        | [26]                                                                                       |
| Scrapie                       | • Citrullination of ovine Prion Protein (PrP) enhances its ability to form insoluble amyloid fibrils and plaques.                                                                                           | [38]                                                                                       |

*Curr Opin Drug Discov Devel. Author manuscript; available in PMC 2013 September 12.*