

SCIENTIFIC REPORTS

OPEN

Probing the Roles of Calcium-Binding Sites during the Folding of Human Peptidylarginine Deiminase 4

Yi-Liang Liu^{1,2}, Chien-Yun Lee^{1,3,4}, Yu-Ni Huang¹, Hui-Yi Chen⁵, Guang-Yaw Liu² & Hui-Chih Hung^{1,6,7}

Our recent studies of peptidylarginine deiminase 4 (PAD4) demonstrate that its non-catalytic Ca²⁺-binding sites play a crucial role in the assembly of the correct geometry of the enzyme. Here, we examined the folding mechanism of PAD4 and the role of Ca²⁺ ions in the folding pathway. Multiple mutations were introduced into the calcium-binding sites, and these mutants were termed the Ca1_site, Ca2_site, Ca3_site, Ca4_site and Ca5_site mutants. Our data indicate that during the unfolding process, the PAD4 dimer first dissociates into monomers, and the monomers then undergo a three-state denaturation process via an intermediate state formation. In addition, Ca²⁺ ions assist in stabilizing the folding intermediate, particularly through binding to the Ca3_site and Ca4_site to ensure the correct and active conformation of PAD4. The binding of calcium ions to the Ca1_site and Ca2_site is directly involved in the catalytic action of the enzyme. Finally, this study proposes a model for the folding of PAD4. The nascent polypeptide chains of PAD4 are first folded into monomeric intermediate states, then continue to fold into monomers, and ultimately assemble into a functional and dimeric PAD4 enzyme, and cellular Ca²⁺ ions may be the critical factor governing the interchange.

The peptidylarginine deiminase (PAD; protein-arginine deiminase, EC 3.5.3.15) enzyme family catalyzes the Ca²⁺-dependent deimination of arginine to citrulline in proteins, concurrently producing ammonia^{1,2}. The citrullination catalyzed by this deiminase family is a type of post-translational modification³⁻⁵ that may have significant effects on the physiological functions of the target proteins and may play essential roles in cell differentiation⁶, nerve growth⁷, embryonic development⁸, cell apoptosis and gene regulation⁹⁻¹³.

PAD has various tissue distributions¹⁴⁻¹⁹. Five isoforms of PAD (PAD1-4 and PAD6) have been identified. PAD1 is found in the skin epidermis, where it citrullinates keratins and filaggrins^{6,20}. PAD2 is found in the brain, the nervous system and muscle tissues¹⁵. PAD3 is found in hair follicles, where it citrullinates trichohyalin for hair follicle hardening^{16,21}. PAD4 is found in granulocytes, monocytes and macrophages; it citrullinates histones H2A, H3 and H4 and nucleophosmin/B23^{12,13,17,22}. Finally, PAD6 is found in embryonic stem cells and oocytes¹⁹. PAD has broad substrate specificity. Filaggrin and histones H3 and H4 are the most extensively studied of the known PAD protein substrates^{12,13,20,23}. The citrullination sites of these proteins have been identified; thus, synthetic peptides derived from these proteins have been used to determine the sequence specificity of PAD protein substrates^{24,25}. The structures of PAD4 in a complex with various histone H3 and H4 peptides have been resolved, suggesting that PAD4 may recognize a structural motif on the protein surface rather than a specific consensus sequence²⁶.

¹Department of Life Sciences, National Chung Hsing University (NCHU), Taichung, Taiwan. ²Institute of Biochemistry, Microbiology & Immunology, Chung Shan Medical University, and Division of Allergy, Immunology, and Rheumatology, Chung Shan Medical University Hospital, Taichung, Taiwan. ³Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan. ⁴Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan. ⁵Biotechnology Center, NCHU, Taichung, Taiwan. ⁶Institute of Genomics and Bioinformatics, NCHU, Taichung, Taiwan. ⁷Agricultural Biotechnology Center, NCHU, Taichung, Taiwan. Correspondence and requests for materials should be addressed to G.-Y.L. (email: liugy@csmu.edu.tw) or H.-C.H. (email: hchung@dragon.nchu.edu.tw)

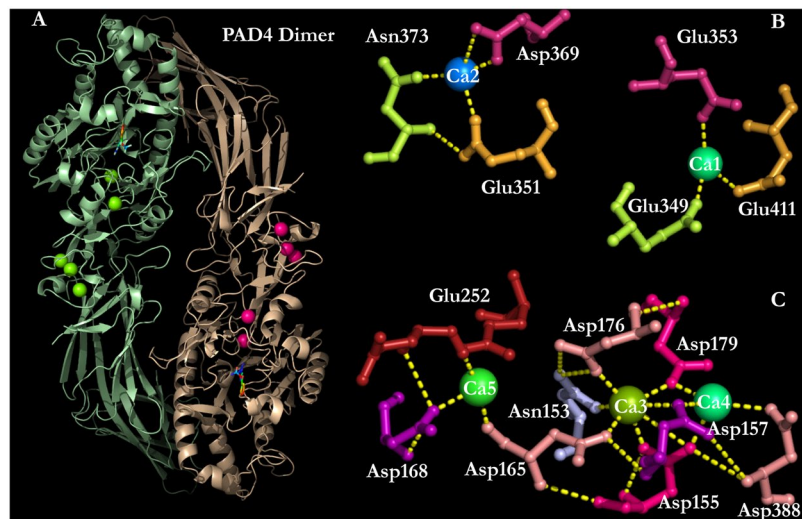


Figure 1. Five Ca^{2+} -binding sites in the human PAD4 enzyme. Binding ligands for the five Ca^{2+} -binding sites, which include the C-terminal calcium-binding sites, i.e., the Ca1_site and Ca2_site, and the N-terminal calcium-binding sites, i.e., the Ca3_site, Ca4_site and Ca5_site. This figure was generated using PyMOL⁵⁰.

During the past ten years, studies of the PAD enzyme and citrullination have attracted much attention. First, high PAD4 activity and high levels of citrullinated proteins are highly related to the pathogenesis of an autoimmune disease known as rheumatoid arthritis (RA)²⁷. An excess of autoantibodies against citrullinated proteins is often discovered in the blood of RA patients^{28,29}. A case control study by a Japanese group revealed that the *PAD4* haplotype that is associated with susceptibility to RA increases production of deiminated peptides that act as autoantigens^{27,30}. In particular, PAD4 is autocitrullinated *in vitro* and *in vivo*, and this modification inactivates its enzymatic function and enhances its recognition by human autoantibodies³¹. As a result, antibodies against these proteins can be used as diagnostic markers for RA. Furthermore, PAD inhibitors are drug development targets. The Cl- and F-amidine inhibitors have been synthesized to effectively inhibit PAD activity^{32–34}.

Second, PAD4 is involved in histone demethylation (citrullination) and causes a decrease in gene expression. Therefore, PAD4 is considered a transcriptional co-repressor³⁵. Indeed, PAD4 is involved in the repression of p53 target genes by its interactions with the C-terminus of p53 and has other regulatory effects on p53 target genes^{36,37}. Because PAD4 has histone methylarginine deiminase activity, the enzyme results in the negative regulation of downstream p53 target genes, including the p21 protein. PAD4 functions as a p53 corepressor; therefore, inhibitors of this enzyme are considered to be potential treatments for cancer. The synthetic Cl- and F-amidine inhibitors have also been examined for their inhibitory effects on PAD4 activity and used to evaluate cancer cell survival rates^{32,37–39}.

Multiple X-ray structures of PAD4 complexed with ligands and calcium ions have been solved and are available in the Protein Data Bank^{26,40}. PAD4 is a homodimeric enzyme with an elongated rubber boot structure. Each monomer contains a separate active site with two calcium ions, Ca1 and Ca2, in the C-terminal domain and a distinct binding region with three additional calcium ions, Ca3, Ca4 and Ca5, situated at the N-terminal domain. The calcium-free, calcium-bound and substrate-bound PAD4 structures indicate that the binding of Ca^{2+} ions to the acidic concave enzyme surface induces a conformational change that creates the active site cleft⁴⁰. The sequence identities among the five PAD isoforms are approximately 50–55%, and the catalytic and calcium-binding residues are mostly conserved, suggesting that the isoforms may have similar tertiary structures and catalytic mechanisms.

Our recent studies of PAD4 demonstrate that the dimerization of the enzyme is essential for full enzymatic activity and calcium-binding cooperativity⁴¹. In addition, we also provided evidence that the binding of Ca^{2+} ions to the Ca3-, Ca4- and Ca5-binding sites plays a crucial role in achieving the correct geometry for full activation of the enzyme⁴². However, the exact role of these calcium-binding sites during the folding of PAD4 is not yet clear. To address these questions, multiple mutations were made at the calcium-binding sites, and these mutations were termed the Ca1_site, Ca2_site, Ca3_site, Ca4_site and Ca5_site mutants. Based on biophysical data from several types of measurements, we propose that an intermediate state is formed during the folding process of PAD4 and that this intermediate is mainly stabilized by the binding of Ca^{2+} ions to the N-terminal Ca^{2+} -binding site.

Results and Discussion

To explore the folding pathway of human dimeric PAD4 and to identify the specific roles of each of its calcium-binding sites (Ca1, Ca2, Ca3, Ca4 and Ca5; Fig. 1) during folding, PAD4 was subjected to urea-induced denaturation and renaturation, and multiple mutagenesis experiments were performed on its calcium-binding sites (Table 1).

Characteristics of human PAD4 and the Ca^{2+} -binding-site mutant enzymes. The C-terminal Ca^{2+} -binding sites, namely the Ca1_site and Ca2_site, are part of the active site of the enzyme (Fig. 1)⁴⁰. The

PAD4	Variants
Ca1_site mutant	Q349A/E353A/E411A
Ca2_site mutant	E351A/D369A/N373A
Ca3_site mutant	N153A/D155A/D157A/D165A/D176A/D179A
Ca4_site mutant	D155A/D157A/D179A/D388A
Ca5_site mutant	D168A/E252A

Table 1. Mutations of the PAD4 enzyme at the Ca1-, Ca2-, Ca3-, Ca4- and Ca5-binding sites.

PAD4	$K_{m,BAEE}^a$ (mM)	k_{cat}^a (s ⁻¹)	$k_{cat}/K_{m,BAEE}$ (s ⁻¹ mM ⁻¹)	$K_{0.5,Ca}$ (mM)	h coefficient
WT	0.57 ± 0.33	14.5 ± 0.06	25.4	0.28 ± 0.05	1.7 ± 0.2
Ca1_site mutant	1.22 ± 0.40	0.02 ± 0.44	0.02	0.58 ± 0.27	1.0 ± 0.5
Ca2_site mutant	1.30 ± 0.15	1.45 ± 0.02	1.12	1.23 ± 0.44	1.0 ± 0.2
Ca3_site mutant	3.28 ± 0.81	1.84 ± 0.01	0.56	1.20 ± 0.37	1.0 ± 0.2
Ca4_site mutant	6.52 ± 2.40	0.07 ± 0.01	0.01	2.21 ± 0.38	1.0 ± 0.2
Ca5_site mutant	1.49 ± 0.08	4.54 ± 0.07	3.05	3.25 ± 0.09	1.8 ± 0.1

Table 2. Kinetic parameters of PAD4 and the calcium-binding-site mutant enzymes.

N-terminal Ca²⁺-binding sites, namely, the Ca3_site, Ca4_site and Ca5_site, are also required for enzymatic activity⁴². The kinetic parameters of the wild-type (WT) PAD4 and the calcium-binding-site mutant enzymes are presented in Table 2. The Michaelis constant of the *in vitro* substrate BAEE ($K_{m,BAEE}$), the catalytic constant (k_{cat}) and the specificity constant ($k_{cat}/K_{m,BAEE}$) of the WT enzyme were 0.57 mM, 14.5 s⁻¹ and 25.4 mM⁻¹ s⁻¹, respectively. The binding of Ca²⁺ ions to the enzyme was cooperative with sigmoidal kinetics³⁸. The half-saturation constant for Ca²⁺ ($K_{0.5,Ca}$) and the degree of cooperativity of calcium binding (h) for the WT enzyme were 0.28 mM and 1.7, respectively.

All the calcium-binding-site mutant enzymes displayed poor catalytic efficiency (Table 2), which is not surprising for the Ca1_site and Ca2_site mutants because Ca²⁺ ions were bound to these sites for catalysis. The $K_{m,BAEE}$, k_{cat} and $k_{cat}/K_{m,BAEE}$ values of the Ca1_site mutant were 1.22 mM, 0.02 s⁻¹ and 0.02 mM⁻¹ s⁻¹, respectively, and these values for the Ca2_site mutant were 1.3 mM, 1.45 s⁻¹ and 1.12 mM⁻¹ s⁻¹, respectively (Table 2). Compared with that of the WT, the $K_{m,BAEE}$ values of the Ca1_site and Ca2_site mutants increased approximately 2-fold, but the $k_{cat}/K_{m,BAEE}$ values of these two mutants severely decreased (Table 2), clearly showing the catalytic role of these two calcium ions. This role is especially obvious for the Ca1_site mutant. The overall catalytic efficiency of this mutant was just 0.1% of the WT efficiency (Table 2). The cooperative binding of calcium ions to the Ca1_site and Ca2_site mutants was completely abolished with an h value of 1.0.

For the Ca3_site, Ca4_site and Ca5_site mutants, the $K_{m,BAEE}$ values were elevated with significantly decreased k_{cat} and $k_{cat}/K_{m,BAEE}$ values. For the Ca3_site mutant, the $K_{m,BAEE}$, k_{cat} and $k_{cat}/K_{m,BAEE}$ values were 3.28 mM, 1.84 s⁻¹ and 0.56 mM⁻¹ s⁻¹, respectively; for the Ca4_site mutant, these values were 6.52 mM, 0.07 s⁻¹ and 0.01 mM⁻¹ s⁻¹, respectively, and for the Ca5_site mutant, they were 1.49 mM, 4.57 s⁻¹ and 3.05 mM⁻¹ s⁻¹, respectively (Table 2). The cooperative binding of calcium ions to the Ca3_site and Ca4_site mutants was also abolished with an h value of 1.0. The Ca5_site mutant, however, is the only mutant that retained a slightly level of catalytic activity and cooperativity with an h value of 1.8, similar to that of the WT.

Although Ca3_site, Ca4_site and Ca5_site are not thought to be catalytic sites, mutations abolishing these binding sites severely affected the enzyme catalysis and increased the $K_{m,BAEE}$ values by more than 6-fold (Table 2), indicating that the binding of calcium ions to these N-terminal Ca²⁺-binding sites may have profound effects on the precise conformation of the catalytic site. This effect is noticeable for the Ca4_site mutant, for which the $K_{m,BAEE}$ value increased more than 10-fold over that of the WT (Table 2). Based on these findings, the Ca3_site and Ca4_site are crucial for the catalytic function of the enzyme by binding Ca²⁺ to ensure the correct conformation of the active site.

The PAD4 WT exhibited a stable dimer with a dissociation constant between monomer and dimer (K_d) of 0.25 μM, and these calcium-binding-site mutants were also present as dimers with K_d values similar to that of the WT enzyme (Fig. S1 and Table S1). Although mutations in these calcium-binding sites did not affect the dimerization of PAD4, the calcium-binding-site mutants exhibited inactive dimers (Table 2), suggesting that the folding pathway of these mutants may be different from that of the WT enzyme.

Unfolding and refolding of the WT PAD4 enzyme. To investigate the folding pathway of PAD4, the urea-induced denaturation and renaturation of the WT enzyme was first examined. The conformational changes of the PAD4 enzyme were monitored by CD and intrinsic fluorescence (Fig. 2), and the thermodynamic parameters are shown in Table 3. The unfolding curve measured from the far-UV CD signals (molar ellipticity at 222 nm) was biphasic (Fig. 2A, open circles), and the urea concentrations at half-maximal denaturation, [Urea]_{0.5}, for the first and second phases ([Urea]_{0.5,N→I} and [Urea]_{0.5,I→U}, respectively) were approximately 3.0 and 5.3 M, respectively (Table 3). The intermediate state was achieved with approximately 3.5–4.0 M urea (Fig. 2A). Similar results

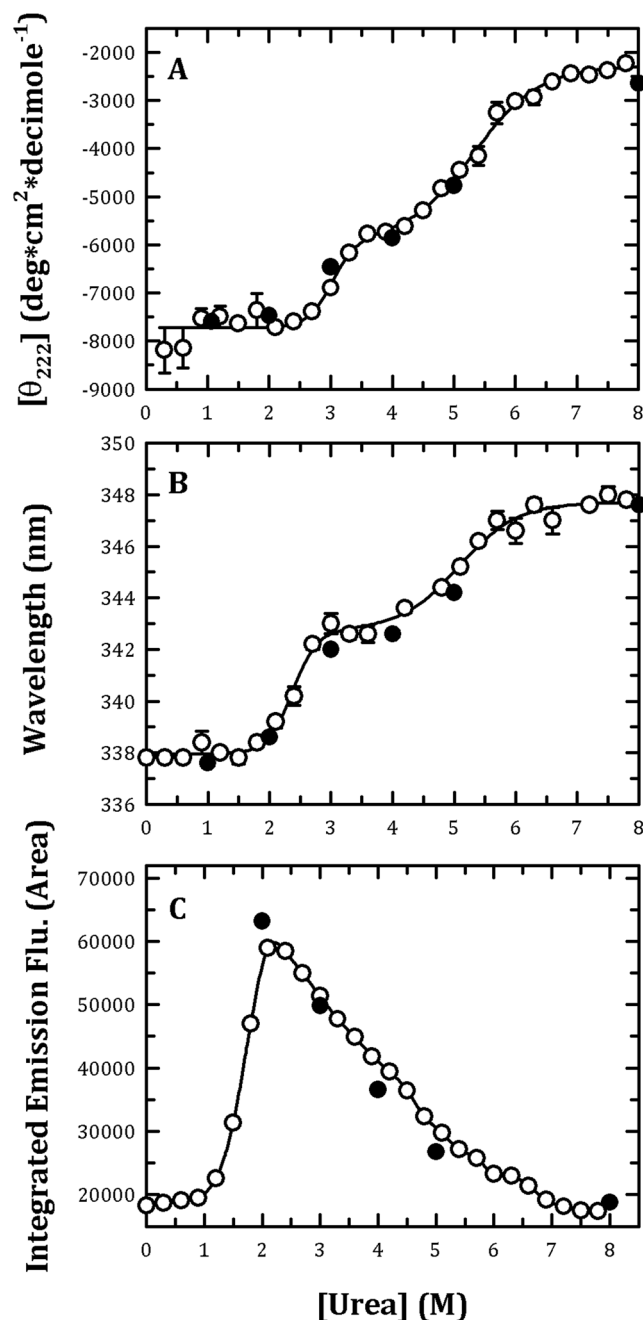


Figure 2. Monitoring of the urea-induced unfolding and refolding of human PAD4 WT enzyme by CD spectrometry and the intrinsic protein fluorescence. The PAD4 WT enzyme in the presence of 10 mM Ca²⁺ was treated with various concentrations of urea in 50 mM Tris-HCl buffer (pH 7.4) at 25 °C for 16 h and then monitored through CD spectrometry (A), fluorescence (B) or ANS fluorescence (C). Open circles: the PAD4 enzyme was denatured with different concentrations of urea. Closed circles: the PAD4 enzyme was completely denatured with 8 M urea and then renatured by diluting the urea concentration to 5, 4, 3, 2 and 1 M, as indicated in the figures. The experimental data in (A,B) were fitted by either a two-state or three-state model. The fit results and residues are shown as a solid line with error bars.

were observed in the fluorescence experiments. The unfolding curve measured from the average fluorescence emission wavelengths was biphasic (Fig. 2B, open circles), with $[\text{Urea}]_{0.5}$ values of 2.4 M for the first phase and 6.0 M for the second phase (Table 3). The intermediate state appeared in the presence of approximately 3–4 M urea (Fig. 2B).

The PAD4 enzyme was completely denatured with 8 M urea and then renatured by diluting the urea concentration to 5, 4, 3, 2 and 1 M (Fig. 2, closed circles). The refolding CD spectra for PAD4 WT enzyme in the native (0 M urea, black line), partially unfolded (1 M urea, red line), unfolded (8 M urea, blue line) and refolded (denatured with 8 M urea then renatured with 1 M urea, green line) states are shown in Fig. S2. Basically, the CD profile

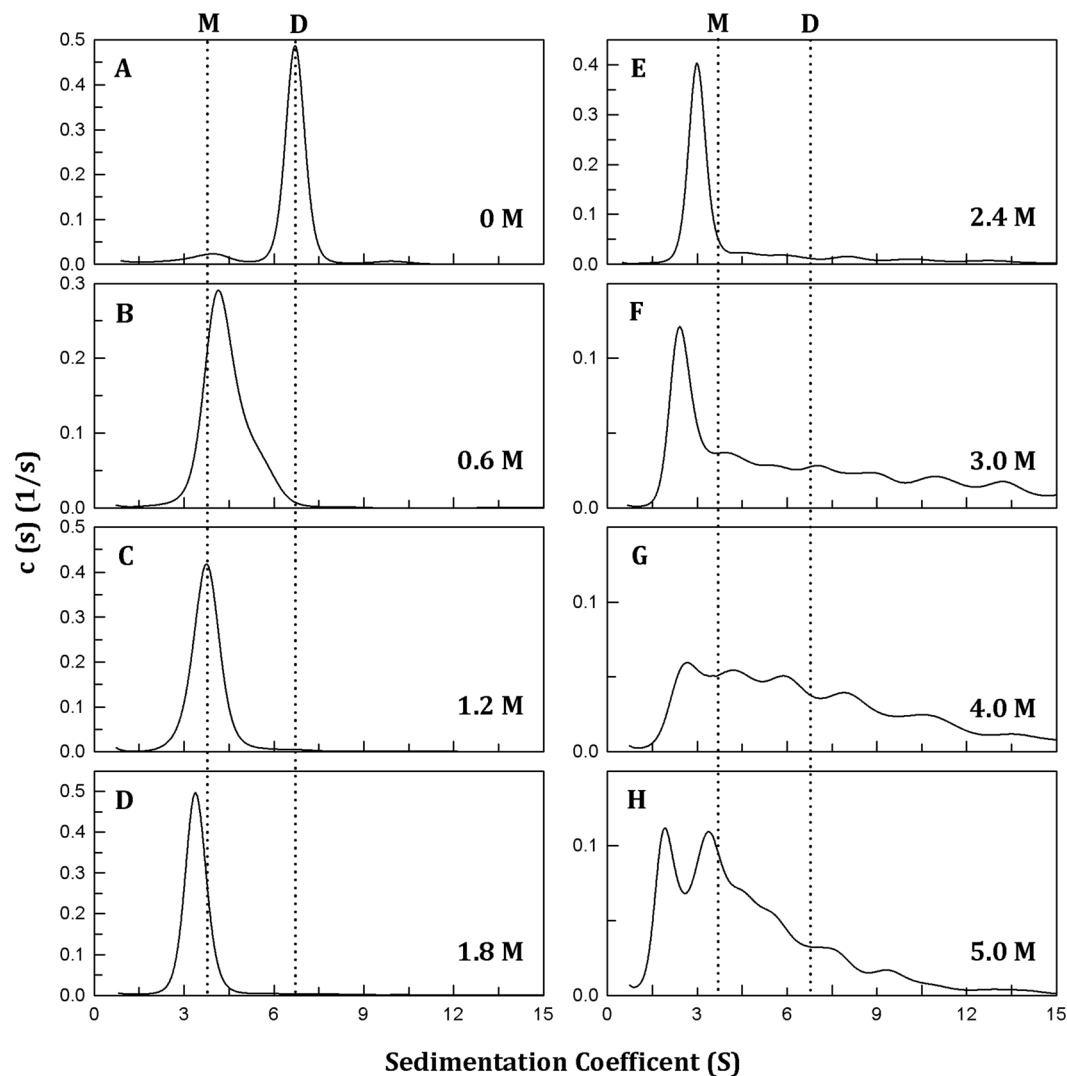


Figure 3. Continuous sedimentation coefficient distributions of human PAD4 WT enzyme during urea denaturation. The PAD4 WT enzyme (0.3 mg/ml) in the presence of 10 mM Ca^{2+} was treated with various concentrations of urea in 50 mM Tris-HCl buffer (pH 7.4) at 25 °C for 16 h: (A) 0 M urea, (B) 0.6 M urea, (C) 1.2 M urea, (D) 1.8 M urea, (E) 2.4 M urea, (F) 3.0 M urea, (G) 4.0 M urea and (H) 5.0 M urea.

	Midpoint of Urea-induced Denaturation					
	CD measurement			Fluorescence measurement		
PAD4	[Urea] _{0.5,N→I} (M)	[Urea] _{0.5,I→U} (M)	[Urea] _{0.5,N→U} (M)	[Urea] _{0.5,N→I} (M)	[Urea] _{0.5,I→U} (M)	[Urea] _{0.5,N→U} (M)
WT	3.0 ± 1.6	5.3 ± 1.5	—	2.4 ± 0.8	6.0 ± 2.8	—
Ca1_site mutant	2.6 ± 1.0	5.7 ± 2.1	—	2.0 ± 1.5	5.3 ± 2.7	—
Ca2_site mutant	2.9 ± 1.0	6.3 ± 3.4	—	1.3 ± 1.3	4.1 ± 1.0	—
Ca3_site mutant	—	—	4.7 ± 1.1	—	—	4.1 ± 0.9
Ca4_site mutant	—	—	5.1 ± 1.4	—	—	4.9 ± 2.5
Ca5_site mutant	1.6 ± 1.2	5.3 ± 1.2	—	1.1 ± 0.8	4.6 ± 1.9	—

Table 3. Thermodynamic parameters derived from the circular dichroism (CD) and fluorescence measurements of PAD4 and the calcium-binding-site mutant enzymes. These values were derived from the circular dichroism data presented in Fig. 5A–E. These values were derived from the fluorescence data presented in Fig. 5F–J.

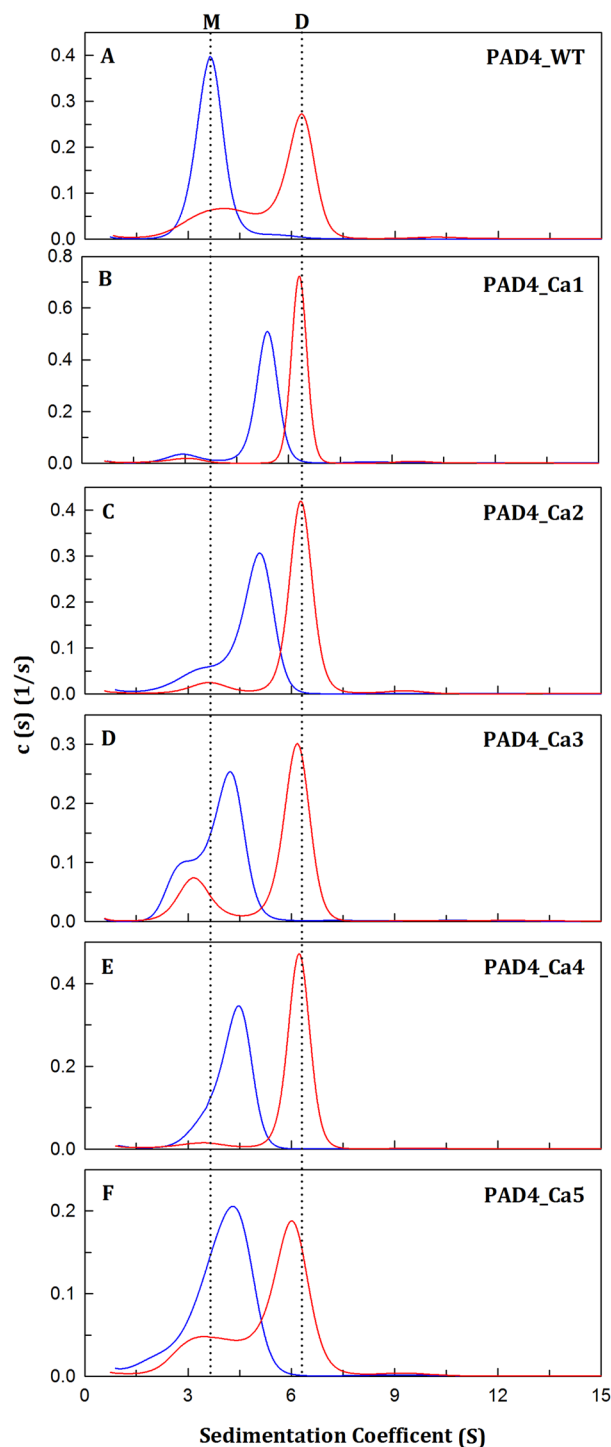


Figure 4. Dissociation-reassociation of human PAD4 WT and the calcium-binding-site mutant enzymes. The PAD4 WT and the calcium-binding-site mutant enzymes were treated with 1.2 M urea in 50 mM Tris-HCl buffer (pH 7.4) at 25 °C for 16 h, and the protein samples were then diluted 10-fold to reduce the urea concentration to 0.12 M. (A) PAD4_WT. (B) PAD4_Ca1. (C) PAD4_Ca2. (D) PAD4_Ca3. (E) PAD4_Ca4. (F) PAD4_Ca5. Blue line: the enzyme treated with 1.2 M urea. Red line: 10-fold dilution of the 1.2 M urea-treated enzyme (0.12 M urea).

of the refolding curve (green line) was similar to the native (black line and partially unfolded states (red line). Thus, the value at 222 nm was used to present the unfolding/refolding status. As determined through monitoring the molar ellipticity at 222 nm or the fluorescence emission wavelength, the signals of the PAD4 enzyme reached the respective denatured condition (Fig. 2A,B, closed circles, respectively), indicating that the unfolding-refolding of PAD4 is a reversible process.

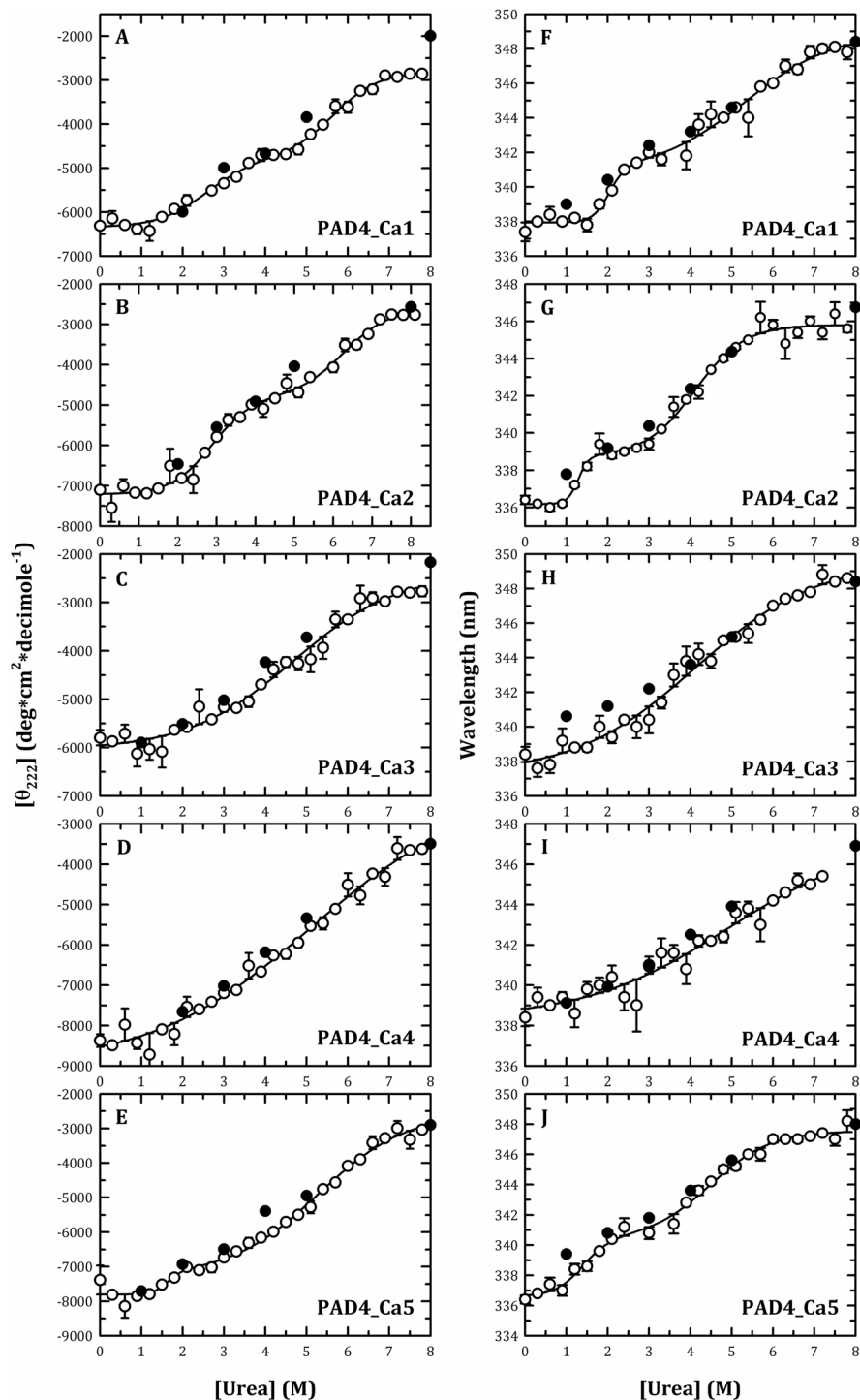


Figure 5. Monitoring of the urea-induced unfolding and refolding of the human PAD4 calcium-binding-site mutant enzymes by CD spectrometry and the intrinsic protein fluorescence. The PAD4 calcium-binding-site mutant enzymes in the presence of 10 mM Ca^{2+} were treated with various concentrations of urea in 50 mM Tris-HCl buffer (pH 7.4) at 25 °C for 16 h and then monitored through CD spectrometry (Panels (A–E)) or fluorescence (Panels (F–J)). Panels (A and F) PAD4_Ca1. Panels (B and G) PAD4_Ca2. Panels (C and H) PAD4_Ca3. Panels (D and I) PAD4_Ca4. Panels (E and J) PAD4_Ca5. Open circles: the PAD4 enzyme was denatured with different concentrations of urea. Closed circles: the PAD4 enzyme was completely denatured with 8 M urea and then renatured by diluting the urea concentration as indicated in the figures. All the data were fitted by either a two-state or three-state model. The fit results and residues are shown as a solid line with error bars.

The ANS fluorescence of the urea-induced denaturation of PAD4 was also determined. ANS is generally employed as a sort of hydrophobic probe on protein surfaces. The changes in the ANS fluorescence for the PAD4-WT enzyme with increasing concentrations of urea were measured (Fig. 2C). A bell-shaped curve with a single peak that occurred at approximately 2 M urea and was followed by a shoulder was observed. The ANS fluorescence of PAD4 increased to a maximal value at approximately 2 M urea, indicating that some hydrophobic regions were largely exposed in the presence of 2 M urea, although no significant changes were detected using CD and fluorescence (Fig. 2A,B, respectively).

We examined the quaternary structure changes of PAD4 during urea-induced denaturation by analytical ultracentrifugation (AUC). Figure 3 shows the continuous sedimentation coefficient distribution of PAD4 during urea denaturation. The sedimentation coefficients of 3.7 S and 6.8 S correspond to monomers and dimers, with molecular weights of 37 and 74 kDa, respectively. The quaternary structure of the WT enzyme displayed a dimer with a sedimentation coefficient of 6.8 S (Fig. 3A). However, the PAD4 dimer is sensitive to urea denaturation because the enzyme was significantly dissociated with 0.6 M urea (Fig. 3B). The enzyme completely dissociated into monomers with 1.2–1.8 M urea (Fig. 3C,D). At higher concentrations of urea, the size distribution plots of the enzyme became broad and irregular, indicating the unfolding of the monomeric enzyme (Fig. 3E–H). The 1.2 M urea-treated PAD4 enzyme, as previously mentioned, was found to exist as a monomer (Fig. 4A, blue line), and diluting the urea concentration from 1.2 M to 0.12 M caused the enzyme to re-associate into dimers (Fig. 4A, red line), indicating that the urea-induced dissociation of PAD4 is reversible. Furthermore, the dissociation process of PAD4 may correspond to the increased ANS fluorescence observed with 2 M urea (Fig. 2C) because the dimeric enzyme was completely dissociated into monomers in the presence of 2 M urea (Fig. 3), indicating that some hydrophobic regions in the dimer interface of PAD4 were exposed gradually during its dissociation from a dimer to monomers. In fact, a significant hydrophobic interface is present in the PAD4 dimer. Several hydrophobic amino acid residues are located at the dimer interface, including Y435, F541, W548 and F576 in the A subunit (colored in green; Fig. S3) and L6, L279, V283, V284 and F285 in the B subunit (colored in purple; Fig. S3). We have identified the hydrophobic amino acid residues at the dimer interface that significantly affected dimerization, particularly L6, L279 and V283⁴³. Therefore, when the enzyme dissociated during urea-induced denaturation, the hydrophobic interface was gradually exposed, and then the hydrophobic probe ANS bound to this region, allowing us to monitor the dissociation of the enzyme.

Based on the denaturation data from the CD (Fig. 2A), fluorescence (Fig. 2B) and AUC measurements (Fig. 3), we suggest that the dissociation and denaturation processes of PAD4 are separate: the enzyme first dissociates from a dimer to a monomer at <2 M urea, and the monomer then begins to denature at higher urea concentrations (Fig. 2). The ANS fluorescence clearly increased to reach a maximal value at approximately 2 M urea; in the presence of this concentration of urea, the PAD4 dimer completely dissociated into monomers in the AUC experiments (Fig. 3), although dissociation was not detected in CD and fluorescence titration experiments (Fig. 2). This finding suggests that the first transition reflects the denaturation of the enzyme rather than the dissociation of the dimers and that dissociation of the dimers to monomers exposes the buried region in the dimer interface. Furthermore, the midpoints detected by CD and fluorescence, either between N and I or I and U, are close to each other (Table 3), suggesting that the enzyme truly follows a three-state unfolding process involving an intermediate that appeared at a urea concentration of 3–4 M (Fig. 2). In addition, the renaturation data from the CD (Fig. 2A, closed circles), fluorescence (Fig. 2B, closed circles) and AUC measurements (Fig. 4A) revealed that the urea-induced denaturation and dissociation of PAD4 is a reversible process; hence, the folding pathway of PAD4 can be characterized by tracing the unfolding of PAD4.

Unfolding and refolding of the catalytic Ca²⁺-binding-site mutant enzymes (Ca1_site and Ca2_site mutants). To probe the role of the calcium-binding sites during the folding of PAD4, the urea-induced denaturation and renaturation of the calcium-binding-site mutants were examined. The conformational changes of the Ca1_site and Ca2_site mutants during urea-induced denaturation was monitored by far-UV CD (open circles in Fig. 5A,B, respectively) and fluorescence signals (open circles in Fig. 5E,G, respectively), and the thermodynamic parameters were determined (Table 3). The urea denaturation curves of the Ca1_site and Ca2_site mutants were also biphasic, similar to that of the WT enzyme (Fig. 2, open circles). The monitoring of the unfolding process by far-UV CD signals revealed that [Urea]_{0.5} values of the Ca1_site and Ca2_site mutants were 2.6 and 2.9 M, respectively, for the first phase (3.0 M for WT) and 5.7 M and 6.3 M, respectively, for the second phase (5.3 M for WT; Table 3). The intermediate state of the Ca1_site and Ca2_site mutants appeared in the presence of approximately 4 M urea (3.5–4.0 M for WT). The monitoring of the unfolding process by fluorescence signals showed that the [Urea]_{0.5} values of the Ca1_site and Ca2_site mutants were 2.0 and 1.3 M, respectively, for the first phase (2.4 M for WT) and 5.3 M and 4.1 M, respectively, for the second phase (6.0 M for WT; Table 3). The intermediate state of the Ca1_site and Ca2_site mutants appeared in the presence of approximately 2.5–3 M urea (3.0–4.0 M for WT). In summary, the unfolding patterns of the Ca1_site and Ca2_site mutants are similar to those of the WT protein, which followed a three-state unfolding pathway, although the intermediate states of these two mutants appeared earlier than the WT protein, as determined using fluorescence monitoring.

The Ca1_site and Ca2_site mutants were refolded through a reverse of unfolding path. After denaturation with 8 M urea, the mutant enzymes could be renatured by diluting the urea concentrations. The signals for protein renaturation, monitored through either the molar ellipticity at 222 nm (closed circles in Fig. 5A,B, respectively) or the fluorescence emission wavelength (closed circles in Fig. 5E,G, respectively), returned to values similar to those obtained for the respective denatured condition, indicating that these two mutants could refold to a native state similar to the WT enzyme. The 1.2 M urea-treated Ca1_site and Ca2_site mutants were dissociated (Fig. 4B,C, blue lines, respectively), and diluting the urea concentration from 1.2 M to 0.12 M caused the enzyme to re-associate into dimers (Fig. 4B,C, red lines, respectively), similar to the pattern found for the WT enzyme.

(Fig. 4A, red line). According to the abovementioned data, the folding of the tertiary structures and the assembly of the quaternary structures of the Ca1_site and Ca2_site mutants are similar to those of the WT protein, which follows a three-state pathway with an intermediate. The destruction of the catalytic Ca²⁺-binding sites of PAD4 or the diminished binding of Ca²⁺ ions to the Ca1_site and Ca2_site severely affected enzyme catalysis but not the formation of the intermediate, which is critical for the correct folding of PAD4. Therefore, the binding of calcium ions to the Ca1 and Ca2 sites mainly contribute to catalysis and are less important for structural stability.

Unfolding and refolding of the non-catalytic Ca²⁺-binding-site mutant enzymes (Ca3_site, Ca4_site and Ca5_site mutants). In contrast to the WT protein and the Ca1_site and Ca2_site mutants, the unfolding pattern of the Ca3_site and Ca4_site mutants displayed a monophasic denaturation curve with a broad shoulder, implying that an unstable intermediate may exist in equilibrium. The monitoring of the unfolding process by far-UV CD signals (open circles in Fig. 5C,D, respectively) revealed that the [Urea]_{0.5} values for the Ca3_site and Ca4_site mutants were 4.7 M and 5.1 M, respectively, and the monitoring of the unfolding process by fluorescence signals (open circles in Fig. 5H,I, respectively) showed [Urea]_{0.5} values for the Ca3_site and Ca4_site mutants of 4.1 and 4.9 M, respectively (Table 3).

The unfolding process of the Ca3_site and Ca4_site mutant enzymes was reversible. Most protein signals of renaturation monitored by CD (closed circles in Fig. 5C,D, respectively) or fluorescence (closed circles in Fig. 5H,I, respectively) returned to the values of the respective denatured condition. The Ca3_site and Ca4_site mutant enzymes in the presence of 1.2 M urea were dissociated into monomers (Fig. 4D,E, blue lines, respectively), and diluting the urea concentration from 1.2 M to 0.12 M caused the enzyme to re-associate from monomers into dimers (Fig. 4D,E, red lines, respectively), similar to the results found for the WT enzyme (Fig. 4A, red line). Based on these aforementioned data, the folding of the tertiary structures of the Ca3_site and Ca4_site mutants follows a two-state pathway with the formation of an unstable intermediate, and the assembly of the quaternary structures of these two mutants is similar to the WT protein. The destruction of the non-catalytic Ca²⁺-binding sites in PAD4 or the diminished binding of Ca²⁺ ions to the Ca3_site and Ca4_site severely affect the stability of the intermediate, which is critical for the correct folding of PAD4. Therefore, calcium binding to the Ca3 and Ca4 sites is crucial for tertiary structure folding.

The Ca5_site mutant displayed a biphasic denaturation curve. The monitoring of the unfolding process by far-UV CD signals (open circles in Fig. 5E) revealed [Urea]_{0.5} values for the Ca5_site mutant of 1.6 M for the first phase (3.0 M for WT) and 5.3 M for the second phase (5.3 M for WT; Table 3). The monitoring of the unfolding process by fluorescence signals (open circles in Fig. 5J) showed that the [Urea]_{0.5} values of the Ca5_site mutant was 1.1 M for the first phase (2.4 M for WT) and 4.6 M for the second phase (6.0 M for WT; Table 3). Although the unfolding pattern of the Ca5_site mutants was similar to that of the WT, the intermediate states were observed earlier than with the WT. Similar to the Ca1_site and Ca2_site of PAD4, the destruction of the Ca5_site did not severely influence the formation of an intermediate during the folding process.

The unfolding of the Ca5_site mutant enzyme is a reversible process. The signals of protein renaturation monitored by the molar ellipticity at 222 nm and the fluorescence emission wavelength returned to values similar to those of the respective denatured condition, indicating that the Ca5_site mutant could refold to a native state similar to the WT enzyme (closed circles in Fig. 5E,J, respectively). The Ca5_site mutant enzyme treated with 1.2 M urea was dissociated into monomers (Fig. 4F, blue line), and diluting the urea concentration from 1.2 M to 0.12 M caused the enzyme to re-associate into dimers (Fig. 4F, red line), similar to the pattern observed for the WT enzyme (Fig. 4A, red line). Thus, the folding of the tertiary structure and the assembly of the quaternary structure of the Ca5_site mutant are similar to the WT protein, suggesting that the binding of calcium to the Ca5_site is less important for the folding of the PAD4 protein.

The folding pathway of the human PAD4 enzyme. Ca²⁺ ions are essential for the enzymatic activity and structural stability of the human PAD4 enzyme^{40,42}. Structural evidence suggests that the binding of Ca²⁺ to the enzyme induces significant conformational changes that generate the active enzyme⁴⁰. The C-terminal Ca²⁺-binding sites, namely, the Ca1_site and Ca2_site, are believed to form part of the catalytic calcium-binding site of the enzyme⁴⁰. Furthermore, the N-terminal Ca²⁺-binding sites, specifically the Ca3_site, Ca4_site and Ca5_site, have been shown to be required for enzymatic activity because they stabilize the conformational geometry of the enzyme⁴². This analysis of the unfolding-refolding process of PAD4 led us to propose a folding process for an active and dimeric PAD4 enzyme (Fig. 6).

During the unfolding process, the PAD4 dimer first dissociates into monomers, and the monomers subsequently follow a three-state denaturation process via the formation of an intermediate state. In addition, the binding of calcium ions to the Ca3_site and Ca4_site primarily affects the formation of a stable intermediate. Calcium ions are important for the folding of PAD4; in particular, the binding of calcium ions to the Ca3_ and Ca4_ sites stabilizes the folding intermediate to ensure the correct folding of PAD4. The binding of calcium ions to the Ca1_site and Ca2_site may directly affect enzymatic catalysis and may be partially involved in achieving the correct active site conformation. Therefore, we propose a potential model for the folding pathway of PAD4 (Fig. 6). When a nascent polypeptide chain of PAD4 is synthesized, the unfolded PAD4 monomer is first folded into a monomeric intermediate state. The intermediate state continues to fold into a monomer. Two monomers then assemble into a dimeric PAD4 enzyme, which is fully functional for enzyme catalysis (Fig. 6, left). The binding of calcium ions to Ca3_site and Ca4_site is important for the formation of the intermediate state, which plays a crucial role during the folding of PAD4. If the Ca3_site or Ca4_site is mutated, the enzyme cannot fold properly, leading to an inactive enzyme (Fig. 6, right). The binding of calcium ions to the Ca1_site and Ca2_site is crucial for enzyme catalysis, and mutations that prevent these two sites from binding Ca²⁺ ions truly abolished the catalytic activity of PAD4 but did not severely affect the formation of an intermediate during the folding process (Fig. 6, left).

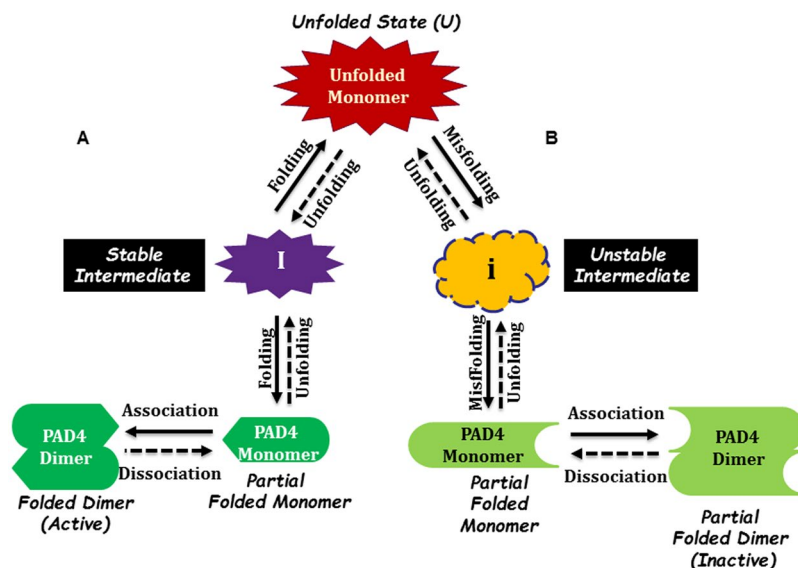


Figure 6. Proposed model of the PAD4 folding pathway. (A) Three-state folding pathway with the formation of a stable intermediate and assembly of dimers of the PAD4 WT, PAD4_Ca1, PAD4_Ca2 and PAD4_Ca5 mutant enzymes. (B) Two-state folding pathway with the formation of an unstable intermediate and assembly of dimers of PAD4_Ca3 and PAD4_Ca4 mutant enzymes.

The role of calcium ions in the folding pathway of PAD4. It appears much of the PAD4 enzyme is folded either in the absence or presence of Ca^{2+} . Actually, the urea-induced denaturation of PAD4 holo-form (with 10 mM CaCl_2) and apo-form (without 10 mM CaCl_2) was examined through intrinsic fluorescence experiments, revealing that the two forms exhibited dissimilar conformational stabilities (Fig. S4). Although the urea denaturation curves of the holo-form and apo-form of PAD4 are apparently biphasic (Fig. S4, open and closed circles, respectively), the thermodynamic parameters and the curve transitions are quite different. The midpoints of the holo-form and apo-form of PAD4 were 2.4 and 0.4 M, respectively, for the first phase and 6.0 M and 5.2 M, respectively, for the second phase. The first transition of the apo-form emerged much earlier, and the intermediate state of the apo-form was much less obvious than those of the holo-form of PAD4, revealing that the holo-form and apo-form of PAD4 represent different conformations. We also performed the titration experiments of Ca^{2+} ions into the PAD4 WT enzyme which were monitored by the CD and fluorescence (Fig. S5). There were just slight changes in CD and fluorescence experiments during Ca^{2+} titration. Although Ca^{2+} ions enhance the PAD4 enzyme to produce a stable and active fold, the Ca^{2+} -induced conformational changes of PAD4 could not be traced by CD or fluorescence. The native PAD4 enzyme without Ca^{2+} ions has CD or fluorescence spectra similar to the PAD4 enzyme with saturated Ca^{2+} ions.

These data coincide with the structural analysis of PAD4 in its Ca^{2+} -free and Ca^{2+} -bound form⁴⁰. PAD4 can fold even the Ca^{2+} ions are absent⁴⁰. However, an r.m.s. deviation is 1.1 Å observed by superimposing the backbone $\text{C}\alpha$ atoms of Ca^{2+} -free and Ca^{2+} -bound PAD4. A number of loop regions are missing in the Ca^{2+} -free structure but are visible in the Ca^{2+} -bound structure. Binding of Ca^{2+} to Ca1 and Ca2 induces marked conformational changes that generate the active site cleft in the C-terminal domain. In addition, Ca^{2+} binding shows a marked difference in conformation in the region between Asn158 and Val171, which is disordered in Ca^{2+} -free PAD4 but becomes ordered to form an α 1-helix when Ca^{2+} binds to Ca3, Ca4, Ca5 and two water molecules in Ca^{2+} -bound PAD4. As described above, PAD4 folds in the absence of Ca^{2+} but the folded state of the apo-form is different from the holo-form. Based on our data and structural analysis, we suggest that Ca^{2+} binding assists in stabilizing the intermediate state and protein structure. The pattern of the unfolding curve of the Ca1_site and Ca2_site mutants is similar to that of the holo-form of PAD4, and the pattern of the unfolding curve of the Ca3_site and Ca4_site mutants is similar to that of the apo-form of PAD4 (Fig. 5).

Finally, we suggest that the protein folding of PAD4 would occur at the intracellular concentrations of calcium ions (100 nM) found within a typical cell, where an equilibrium between unfolded monomers, partially folded intermediates, folded monomers and folded dimers exist (Fig. 6, left), and the enzyme displays very low activity under such low calcium ion concentrations. However, 10- to 100-fold increases in the calcium ion concentration, which can be observed during various cellular functions, causes the PAD4 enzymes to become activated, a process that may be attributed to the folding of most partially folded intermediates into monomers and their subsequent assembly into fully active dimers.

Materials and Methods

Expression and purification of recombinant WT and mutant PAD4. Human PAD4 cDNA was cloned into a pQE30 vector with an N-terminal His tag for purifying the overexpressed PAD4 enzyme. This vector contained an ampicillin-resistant gene and was transfected into the JM109 strain of *E. coli*. The expression of the PAD4 enzyme was induced with 1.0 mM isopropyl-1-thio- β -D-galactoside (IPTG), and the cells were harvested

after overnight-incubation at 25 °C. Ni-NTA sepharose gel (Sigma, St. Louis, MO, USA), which was used to purify the target protein, was equilibrated with binding buffer (5 mM imidazole in 500 mM NaCl, 2 mM β -mercaptoethanol and 30 mM Tris-HCl, pH 7.6). The lysate-Ni-NTA mixture was loaded onto a column and washed with a stepwise procedure (5 and 10 mM imidazole in 500 mM NaCl, 2 mM β -mercaptoethanol and 30 mM Tris-HCl, pH 7.6) to remove undesired proteins. Lastly, PAD4 enzymes were washed out with elution buffer (250 mM imidazole, 500 mM NaCl, 2 mM β -mercaptoethanol and 30 mM Tris-HCl, pH 7.6). The purified enzyme was then subjected to buffer exchange with dialysis buffer (500 mM NaCl, 2 mM β -mercaptoethanol, and 50 mM Tris-HCl, pH 7.4) and concentrated using a centrifugal filter device (Amicon Ultra-15, Millipore) with a molecular weight cutoff of 50 kDa. The purity of the enzymes was examined by SDS-PAGE, and the protein concentrations were determined using the Bradford method⁴⁴.

Site-directed mutagenesis. Site-directed mutagenesis was conducted using a QuikChange™ kit (Stratagene, La Jolla, CA USA). This mutagenesis method was accomplished using *Pfu* DNA polymerase, an enzyme with high fidelity for DNA replication. The specific primers for mutagenesis were 25- to 45-mer oligonucleotides that bind specifically to the template DNA. Multiple mutagenic primers were used to make the calcium-binding-site mutants. For the Ca1_site, Ca2_site and Ca5_site mutants, three sets of primers for each were used; six and four sets of primers were used for the Ca3_site and Ca4_site mutants, respectively. The synthetic oligonucleotides used as mutagenic primers were the following:

N153A 5'-GCCATCCTGCTGGTGGCTTGTGACAGAGACAATC-3',
 D155A 5'-CCTGCTGGTGAAGTGTGCTAGAGACAATCTCG-3',
 D157A 5'-GGTGAAGTGTGACAGAGCTAATCTCGAATCTTCTGCC-3',
 D165A 5'-GAATCTTCTGCCATGGCTTGCAGGATGATG-3',
 D168A 5'-GCCATGGACTGCGAGGCTGATGAAGTCTTGAC-3',
 D176A 5'-GTGCTTGACAGCGAAGCTCTGCAGGACATGTGC-3',
 D179A 5'-GACAGCGAAGACCTGCAGGCTATGTCGCTGATGACCC-3',
 E252A 5'-CATGGACTTCTACGTGGCTGCCCTCGCTTTCCCG-3',
 Q349A 5'-GGATGACCAGTGGATGGCTGATGAAATGGAGATCGGC-3',
 E351A 5'-CCAGTGGATGCAGGATGCTATGGAGATCGGCTACATCC-3',
 E353A 5'-TGCAGGATGAAATGGCTATCGGCTACATCCAAGCCCC-3',
 D369A 5'-GCCCGTGGTCTTCGCTTCTCCAAGGAACAGAGGC-3',
 N373A 5'-GGTCTTCGACTCTCCAAGGGCTAGAGGCCTGAAGGAG-3',
 D388A 5'-GAGTGATGGGTCCAGCTTTTGGCTATGTAAC-3', and
 E411A 5'-CCTTTGGGAACCTGGCTGTGAGCCCCCAGTCACAGTC-3'.

The PCR used 16–18 temperature cycles, and the desired mutant plasmids that included staggered nicks were produced. After the PCR reactions, the products were treated with DpnI to digest the PAD4 WT templates, and the nicked DNA with the anticipated mutations was transformed into the XL-1 strain of *Escherichia coli*.

Enzyme assay and kinetic data analysis. The continuous assay method for measuring PAD4 activity has been reported earlier⁴⁵. The regular reaction mixture for the spectrophotometric assay of PAD4 contained 10 mM benzoyl-L-arginine ethyl ester (BAEE) as an artificial substrate, 10 mM CaCl₂, 2.5 mM dithiothreitol (DTT), 8.5 mM α -ketoglutarate (α -KG), 0.22 mM NADH and 8.4 U of glutamate dehydrogenase (GDH) in 100 mM Tris-HCl (pH 7.6) in a 1-ml cuvette at 25 °C. The appropriate amount of PAD4 was added to the assay mixture to start the reaction. After the enzyme was added to the reaction, the decline in absorbance at 340 nm was monitored continuously using a Perkin-Elmer Lambda-25 spectrophotometer. An enzyme unit was defined as the amount of enzyme that catalyzes the consumption of 1 μ mol of NADH per min. An extinction coefficient of 6.22 cm⁻¹ mM⁻¹ at 340 nm for NADH was used in calculations. The K_m value of the BAEE substrate was determined by varying the BAEE concentration while maintaining constant concentrations of other components. The experimental data were analyzed using Prism 4.0. The sigmoidal curves of the [Ca²⁺] versus the initial velocities were fitted to the Hill equation, and the data were further analyzed to calculate the $K_{0.5, Ca}$ value (i.e., the calcium concentration at half-maximal velocity) and the Hill coefficient (h), which were employed to assess the degree of cooperativity.

$$v = \frac{V_{\max} [Ca^{2+}]^h}{K_{0.5, Ca}^h + [Ca^{2+}]^h} \quad (1)$$

All data fitting was performed using Sigma Plot 10.0 (Jandel, San Rafael, CA, USA).

Equilibrium denaturation and renaturation of the PAD4 enzyme. Urea was used as the chemical denaturant for the denaturation and renaturation of PAD4. The PAD4 enzymes were preincubated with various concentrations of urea in 50 mM Tris-HCl buffer (pH 7.4) for 16 h at 25 °C to enable the unfolding reactions to reach equilibrium. Before urea denaturation, the enzyme was incubated with 10 mM Ca²⁺ to prevent the precipitation of PAD4. The urea-induced denaturation of PAD4 was performed in the presence of 10 mM CaCl₂. For the renaturation experiments, the enzyme was first denatured with 8 M urea and subsequently renatured by diluting the urea concentration. The final concentrations of PAD4 and CaCl₂ in the denaturation and renaturation experiments were maintained equal for further measurement of the conformational changes of the enzyme by equilibrium denaturation and renaturation in the presence of urea.

Monitoring the urea-induced denaturation and renaturation of PAD4 and the calcium-binding-site mutant enzymes using circular dichroism spectrometry and fluorescence spectroscopy. Two techniques, circular dichroism (CD) and fluorescence spectroscopy, were employed to monitor the unfolding and refolding process of the enzymes.

CD signals was measured using a Jasco J-815 spectropolarimeter with 0.1-cm quartz cuvettes and a 1-mm slit width. The ellipticity (222 nm) of all samples was recorded to analyze the protein conformational changes during the urea denaturation. The mean residue ellipticity (Θ) at 222 nm was calculated using the following equation:

$$\Theta = \frac{\text{MRW} \times \theta_{\lambda}}{10 \times l \times c} \quad (2)$$

where MRW is the mean residue weight, θ_{λ} is the measured ellipticity in degrees at wavelength λ , l is the cuvette pathlength (0.1 cm), and c is the protein concentration in g/ml.

Fluorescence spectroscopy was performed using a Hitachi F-4500 FL luminescence spectrometer at 25 °C. The excitation wavelength was set to 290 nm, and the fluorescence emission spectra were recorded by scanning the emission from 300 to 400 nm. All spectra were corrected for buffer absorption and the Raman spectrum of water. The average emission wavelength ($\langle \lambda \rangle$) was analyzed using the average emission wavelength method⁴⁶ and calculated according to the following equation:

$$\langle \lambda \rangle = \frac{\sum F_i \lambda_i}{\sum F_i} \quad (3)$$

in which F_i is the fluorescence intensity at a specific emission wavelength (λ_i).

Monitoring the surface hydrophobicity of PAD4 denaturation through ANS fluorescence. An extrinsic fluorescent probe, ANS (8-anilino-1-naphthalene sulfonate), is used to detect the surface hydrophobic properties of a protein molecule. Therefore, the change in surface hydrophobicity of the WT human PAD4 during urea-induced denaturation was further measured through ANS fluorescence. The enzyme was denatured using various concentrations of urea for 16 h at 25 °C. ANS was then added to the protein sample, and the ANS fluorescence of the protein was then measured using a Hitachi F-4500 FL luminescence spectrometer at 25 °C. All spectra were corrected for buffer absorption. The excitation wavelength was set to 370 nm to monitor the changes in the surface hydrophobicity of the enzyme during the urea-induced unfolding process. The ANS fluorescence spectra were scanned from 400 to 600 nm, and the area of ANS fluorescence was integrated from 425 to 575 nm.

Analysis of the urea-induced denaturation curve of PAD4 and the calcium-binding-site mutant enzymes. The analysis of the unfolding curves of the enzymes was performed as described by Pace⁴⁷, assuming a two-state or three-state unfolding mechanism. For a two-state model, the $\Delta G_{(\text{H}_2\text{O})}$ and m values were estimated by fitting the data to the following equation:

$$y_{\text{obs}} = \frac{y_{\text{N}} + y_{\text{U}} \exp\{-(\Delta G_{(\text{H}_2\text{O})} - m[\text{D}])/RT\}}{1 + \exp\{-(\Delta G_{(\text{H}_2\text{O})} - m[\text{D}])/RT\}} \quad (4)$$

where y_{obs} denotes the observed signal change, and y_{N} and y_{U} represent the signals of the folded and unfolded states, respectively. $\Delta G_{(\text{H}_2\text{O})}$ denotes the intrinsic free energy change in the absence of denaturant, and m represents the dependence of ΔG on the denaturant. $[\text{D}]$ denotes the denaturant concentration, T is the absolute temperature in degrees Kelvin, and R is the gas constant.

The denaturation curve was also analyzed using a three-state model. The $\Delta G_{(\text{H}_2\text{O})}$ and m values at each step were estimated by fitting the overall data to the following equation:

$$y_{\text{obs}} = \frac{y_{\text{N}} + y_{\text{I}} \exp\{-(\Delta G_{(\text{H}_2\text{O}),\text{N} \rightarrow \text{I}} - m_{\text{N} \rightarrow \text{I}}[\text{D}])/RT\} + y_{\text{U}} \exp\{-(\Delta G_{(\text{H}_2\text{O}),\text{N} \rightarrow \text{I}} - m_{\text{N} \rightarrow \text{I}}[\text{D}])/RT\} * \exp\{-(\Delta G_{(\text{H}_2\text{O}),\text{I} \rightarrow \text{U}} - m_{\text{I} \rightarrow \text{U}}[\text{D}])/RT\}}{1 + \exp\{-(\Delta G_{(\text{H}_2\text{O}),\text{N} \rightarrow \text{I}} - m_{\text{N} \rightarrow \text{I}}[\text{D}])/RT\} + \exp\{-(\Delta G_{(\text{H}_2\text{O}),\text{N} \rightarrow \text{I}} - m_{\text{N} \rightarrow \text{I}}[\text{D}])/RT\} * \exp\{-(\Delta G_{(\text{H}_2\text{O}),\text{I} \rightarrow \text{U}} - m_{\text{I} \rightarrow \text{U}}[\text{D}])/RT\}} \quad (5)$$

where y_{I} represent the signals of the intermediate states, $\Delta G_{(\text{H}_2\text{O}),\text{N} \rightarrow \text{I}}$ and $\Delta G_{(\text{H}_2\text{O}),\text{I} \rightarrow \text{U}}$ denote the intrinsic free energy change for the native to intermediate ($\text{N} \rightarrow \text{I}$) and for the intermediate to denatured ($\text{I} \rightarrow \text{U}$) processes, respectively, and $m_{\text{N} \rightarrow \text{I}}$ and $m_{\text{I} \rightarrow \text{U}}$ are the m values for the corresponding processes. The urea concentration for half-denaturation of the protein, $[\text{Urea}]_{0.5}$, was estimated by dividing ΔG by m ⁴⁷.

Quaternary structure analysis of PAD4 and the calcium-binding-site mutant enzymes by analytical ultracentrifugation. To examine the quaternary structural changes of the dimeric PAD4 enzyme during urea denaturation, sedimentation velocity experiments were performed using a Beckman Optima XL-A analytical ultracentrifuge (Beckman Coulter, CA, USA). Before ultracentrifugation, the protein sample was preincubated with various concentrations of urea (0–5.4 M) in the presence of 10 mM CaCl_2 at 25 °C for 16 h. The sample (380 μl) and buffer (400 μl) solutions were separately loaded into the double sector centerpiece and placed in a Beckman An-50 Ti rotor. The experiments were performed at 20 °C and at a rotor speed of 42,000 rpm. The protein samples were monitored by the UV absorbance at 280 nm in continuous mode with a time interval of

480 s and a step size of 0.002 cm. Multiple scans at different time points were fitted to a continuous size distribution model by the program SEDFIT^{48,49}. All size distributions were solved at a confidence level of $p = 0.95$, a best fitted average anhydrous frictional ratio (f/f_0), and a resolution N of 250 sedimentation coefficients between 0.1 and 20.0 S.

References

- Lu, X., Galkin, A., Herzberg, O. & Dunaway-Mariano, D. Arginine deiminase uses an active-site cysteine in nucleophilic catalysis of L-arginine hydrolysis. *J. Am. Chem. Soc.* **126**, 5374–5375 (2004).
- Knuckley, B., Bhatia, M. & Thompson, P. R. Protein arginine deiminase 4: evidence for a reverse protonation mechanism. *Biochemistry* **46**, 6578–6587 (2007).
- Anzilotti, C., Pratesi, F., Tommasi, C. & Migliorini, P. Peptidylarginine deiminase 4 and citrullination in health and disease. *Autoimmun. Rev.* **9**, 158–160 (2010).
- Gyorgy, B., Toth, E., Tarcsa, E., Falus, A. & Buzas, E. I. Citrullination: a posttranslational modification in health and disease. *Int. J. Biochem. Cell Biol.* **38**, 1662–1677 (2006).
- Ishigami, A. & Maruyama, N. Importance of research on peptidylarginine deiminase and citrullinated proteins in age-related disease. *Geriatr. Gerontol. Int.* **10**(Suppl 1), S53–58 (2010).
- Senshu, T., Akiyama, K., Ishigami, A. & Nomura, K. Studies on specificity of peptidylarginine deiminase reactions using an immunochemical probe that recognizes an enzymatically deiminated partial sequence of mouse keratin K1. *J. Dermatol. Sci.* **21**, 113–126 (1999).
- Lange, S. *et al.* Protein deiminases: New players in the developmentally regulated loss of neural regenerative ability. *Dev. Biol.* **355**, 205–214 (2011).
- Wright, P. W., Bolling, L. C., Calvert, M. E., Sarmiento, O. F. & Berkeley, E. V. ePAD, an oocyte and early embryo-abundant peptidylarginine deiminase-like protein that localizes to egg cytoplasmic sheets. *Dev. Biol.* **256**, 73–88 (2003).
- Hung, H. C. *et al.* The functional haplotype of peptidylarginine deiminase IV (S55G, A82V and A112G) associated with susceptibility to rheumatoid arthritis dominates apoptosis of acute T leukemia Jurkat cells. *Apoptosis* **12**, 475–487 (2007).
- Liu, G. Y. *et al.* Overexpression of peptidylarginine deiminase IV features in apoptosis of haematopoietic cells. *Apoptosis* **11**, 183–196 (2006).
- Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylination. *Science* **306**, 279–283 (2004).
- Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H. & Hagiwara, T. Histone deimination antagonizes arginine methylation. *Cell* **118**, 545–553 (2004).
- Wang, S. & Wang, Y. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. *Biochim. Biophys. Acta. (BBA) - Gene Regulatory Mechanisms* **1829**, 1126–1135 (2013).
- Guerrin, M., Ishigami, A., Mechin, M. C., Nachat, R. & Valmary, S. cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type I. *Biochem. J.* **370**, 167–174 (2003).
- Ishigami, A., Ohsawa, T., Asaga, H., Akiyama, K. & Kuramoto, M. Human peptidylarginine deiminase type II: molecular cloning, gene organization, and expression in human skin. *Arch. Biochem. Biophys.* **407**, 25–31 (2002).
- Kanno, T., Kawada, A., Yamanouchi, J., Yosida-Noro, C. & Yoshiki, A. Human peptidylarginine deiminase type III: molecular cloning and nucleotide sequence of the cDNA, properties of the recombinant enzyme, and immunohistochemical localization in human skin. *J. Invest. Dermatol.* **115**, 813–823 (2000).
- Nakashima, K., Hagiwara, T. & Yamada, M. Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *J. Biol. Chem.* **277**, 49562–49568 (2002).
- Zhang, J., Dai, J., Zhao, E., Lin, Y. & Zeng, L. cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type VI. *Acta. Biochim. Pol.* **51**, 1051–1058 (2004).
- Chavanas, S., Mechin, M. C., Takahara, H., Kawada, A. & Nachat, R. Comparative analysis of the mouse and human peptidylarginine deiminase gene clusters reveals highly conserved non-coding segments and a new human gene, PADI6. *Gene* **330**, 19–27 (2004).
- Ishida-Yamamoto, A., Senshu, T., Eady, R. A., Takahashi, H. & Shimizu, H. Sequential reorganization of cornified cell keratin filaments involving filaggrin-mediated compaction and keratin 1 deimination. *J. Invest. Dermatol.* **118**, 282–287 (2002).
- Tarcsa, E., Marekov, L. N., Andreoli, J., Idler, W. W. & Candi, E. The fate of trichohyalin. Sequential post-translational modifications by peptidyl-arginine deiminase and transglutaminases. *J. Biol. Chem.* **272**, 27893–27901 (1997).
- Hagiwara, T., Nakashima, K., Hirano, H., Senshu, T. & Yamada, M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem. Biophys. Res. Commun.* **290**, 979–983 (2002).
- Denis, H. *et al.* Functional connection between deimination and deacetylation of histones. *Mol. Cell. Biol.* **29**, 4982–4993 (2009).
- Knuckley, B. *et al.* Substrate specificity and kinetic studies of PADs 1, 3, and 4 identify potent and selective inhibitors of protein arginine deiminase 3. *Biochemistry* **49**, 4852–4863 (2010).
- Stensland, M. E. *et al.* Primary sequence, together with other factors, influence peptide deimination by peptidylarginine deiminase-4. *Biol. Chem.* **390**, 99–107 (2009).
- Arita, K. *et al.* Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4. *Proc. Natl. Acad. Sci. USA* **103**, 5291–5296 (2006).
- Suzuki, A., Yamada, R., Chang, X., Tokuihiro, S. & Sawada, T. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat. Genet.* **34**, 395 (2003).
- Nijenhuis, S., Zendman, A. J., Vossenaar, E. R., Pruijn, G. J. & van Venrooij, W. J. Autoantibodies to citrullinated proteins in rheumatoid arthritis: clinical performance and biochemical aspects of an RA-specific marker. *Clin. Chim. Acta.* **350**, 17–34 (2004).
- Van Boekel, M. A., Vossenaar, E. R., van den Hoogen, F. H. & van Venrooij, W. J. Autoantibody systems in rheumatoid arthritis: specificity, sensitivity and diagnostic value. *Arthritis Res.* **4**, 87–93 (2002).
- Chang, X., Yamada, R., Suzuki, A., Sawada, T. & Yoshino, S. Localization of peptidylarginine deiminase 4 (PADI4) and citrullinated protein in synovial tissue of rheumatoid arthritis. *Rheumatology (Oxford)* **44**, 40–50 (2005).
- Andrade, F., Darrah, E., Gucek, M., Cole, R. N. & Rosen, A. Autocitrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation. *Arthritis Rheum.* **62**, 1630–1640 (2010).
- Jones, J. E., Causey, C. P., Knuckley, B., Slack-Noyes, J. L. & Thompson, P. R. Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Curr. Opin. Drug Discov. Devel.* **12**, 616–627 (2009).
- Jones, J. E. *et al.* Synthesis and screening of a haloacetamide containing library to identify PAD4 selective inhibitors. *ACS Chem. Biol.* **7**, 160–165 (2012).
- Willis, V. *et al.* Efficacy of Cl-amidine in the collagen induced model of rheumatoid arthritis. *J. Immuno.* **186**, 4396–4404, PMID: PMC3085980 (2011).
- Klose, R. J. & Zhang, Y. Regulation of histone methylation by demethylination and demethylation. *Nat. Rev. Mol. Cell Biol.* **8**, 307–318 (2007).
- Li, P., Wang, D., Yao, H., Doret, P. & Hao, G. Coordination of PAD4 and HDAC2 in the regulation of p53-target gene expression. *Oncogene* **29**, 3153–3162 (2010).

37. Li, P., Yao, H., Zhang, Z., Li, M. & Luo, Y. Regulation of p53 target gene expression by peptidylarginine deiminase 4. *Mol. Cell Biol.* **28**, 4745–4758 (2008).
38. Slack, J. L., Causey, C. P. & Thompson, P. R. Protein arginine deiminase 4: a target for an epigenetic cancer therapy. *Cell Mol. Life Sci.* **68**, 709–720 (2011).
39. Causey, C. P. *et al.* The development of *N*- α -(2-carboxyl)benzoyl-*N*⁵-(2-fluoro-1-iminoethyl)-L-ornithine amide (*o*-F-amidine) and *N*- α -(2-carboxyl)benzoyl-*N*⁵-(2-Chloro-1-iminoethyl)-L-ornithine amide (*o*-Cl-amidine) as second generation Protein Arginine Deiminase (PAD) inhibitors. *J. Med. Chem.* **54**, 6919–6935 (2011).
40. Arita, K., Hashimoto, H., Shimizu, T., Nakashima, K. & Yamada, M. Structural basis for Ca²⁺-induced activation of human PAD4. *Nat. Struct. Mol. Biol.* **11**, 777–783 (2004).
41. Liu, Y. L., Chiang, Y. H., Liu, G. Y. & Hung, H. C. Functional role of dimerization of human peptidylarginine deiminase 4 (PAD4). *PLoS One* **6**(6), e21314, doi:10.1371/journal.pone.0021314 (2011).
42. Liu, Y. L. *et al.* Functional Roles of the Non-Catalytic Calcium-Binding Sites in the N-Terminal Domain of Human Peptidylarginine Deiminase 4. *PLoS One* **8**(1), e51660, doi:10.1371/journal.pone.0051660 (2013).
43. Lee, C. Y. *et al.* Molecular interplay between the dimer interface and the substrate-binding site of human peptidylarginine deiminase 4. *Sci Rep.* 2017 Feb 17, 7, 42662. doi:10.1038/srep42662 (2017).
44. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
45. Liao, Y. F., Hsieh, H. C., Liu, G. Y. & Hung, H. C. A continuous spectrophotometric assay method for peptidylarginine deiminase type 4 activity. *Anal. Biochem.* **347**, 176–181 (2005).
46. Sánchez del Pino, M. M. & Fersht, A. R. Nonsequential unfolding of the α/β barrel protein indole-3-glycerol-phosphate synthase. *Biochemistry* **36**, 5560–5565 (1997).
47. Pace, C. N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280 (1986).
48. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. & Schubert, D. Size distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys. J.* **82**, 1096–1111 (2002).
49. Schuck, P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal. Biochem.* **320**, 104–124 (2003).
50. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. URL <http://www.pymol.org/>.

Acknowledgements

This work was supported by the Ministry of Science and Technology, ROC (MOST 104-2311-B-005-009-MY3 to HCH; 102-2320-B-040-018-MY3 to GYL) and in part by the Ministry of Education, Taiwan, R.O.C., under the ATU plan.

Author Contributions

H.C.H. and G.Y.L. conceived and coordinated the study and wrote the manuscript. Y.L.L., C.Y.L., and Y.N.H. designed, performed and analyzed the experiments. H.Y.C. provided technical assistance. All the authors reviewed the results and approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-02677-1

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017