

# Functional Role of Dimerization of Human Peptidylarginine Deiminase 4 (PAD4)

Yi-Liang Liu<sup>1,2</sup>, Yu-Hsiu Chiang<sup>1</sup>, Guang-Yaw Liu<sup>2\*</sup>, Hui-Chih Hung<sup>1\*</sup>

**1** Department of Life Sciences and Institute of Genomics and Bioinformatics, National Chung-Hsing University, Taichung, Taiwan, **2** Institute of Microbiology & Immunology, Chung Shan Medical University, and Division of Allergy, Immunology, and Rheumatology, Chung Shan Medical University Hospital, Taichung, Taiwan

## Abstract

Peptidylarginine deiminase 4 (PAD4) is a homodimeric enzyme that catalyzes Ca<sup>2+</sup>-dependent protein citrullination, which results in the conversion of arginine to citrulline. This paper demonstrates the functional role of dimerization in the regulation of PAD4 activity. To address this question, we created a series of dimer interface mutants of PAD4. The residues Arg8, Tyr237, Asp273, Glu281, Tyr435, Arg544 and Asp547, which are located at the dimer interface, were mutated to disturb the dimer organization of PAD4. Sedimentation velocity experiments were performed to investigate the changes in the quaternary structures and the dissociation constants ( $K_d$ ) between wild-type and mutant PAD4 monomers and dimers. The kinetic data indicated that disrupting the dimer interface of the enzyme decreases its enzymatic activity and calcium-binding cooperativity. The  $K_d$  values of some PAD4 mutants were much higher than that of the wild-type (WT) protein (0.45  $\mu$ M) and were concomitant with lower  $k_{cat}$  values than that of WT (13.4 s<sup>-1</sup>). The  $K_d$  values of the monomeric PAD4 mutants ranged from 16.8 to 45.6  $\mu$ M, and the  $k_{cat}$  values of the monomeric mutants ranged from 3.3 to 7.3 s<sup>-1</sup>. The  $k_{cat}$  values of these interface mutants decreased as the  $K_d$  values increased, which suggests that the dissociation of dimers to monomers considerably influences the activity of the enzyme. Although dissociation of the enzyme reduces the activity of the enzyme, monomeric PAD4 is still active but does not display cooperative calcium binding. The ionic interaction between Arg8 and Asp547 and the Tyr435-mediated hydrophobic interaction are determinants of PAD4 dimer formation.

**Citation:** Liu Y-L, Chiang Y-H, Liu G-Y, Hung H-C (2011) Functional Role of Dimerization of Human Peptidylarginine Deiminase 4 (PAD4). PLoS ONE 6(6): e21314. doi:10.1371/journal.pone.0021314

**Editor:** Paul Proost, University of Leuven, Rega Institute, Belgium

**Received:** April 8, 2011; **Accepted:** May 25, 2011; **Published:** June 22, 2011

**Copyright:** © 2011 Liu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Science Council, Republic of China (NSC-99-2313-M-005-010 and NSC-99-2314-B-040-002-MY3) and in part, by the Ministry of Education, Taiwan, Republic of China under the ATU plan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: hchung@dragon.nchu.edu.tw (H-CH); liugy@csmu.edu.tw (G-YL)

## Introduction

Peptidylarginine deiminase (PAD; protein-arginine deiminase; EC 3.5.3.15) is a Ca<sup>2+</sup>-dependent enzyme that catalyzes protein citrullination (deimination) in the presence of Ca<sup>2+</sup>. Protein citrullination, which is catalyzed by the PAD enzyme, is a post-translational modification process that converts arginine to citrulline. The change from a positively charged arginine to a neutral citrulline is thought to induce protein unfolding [1–3]. PADs play essential roles in epithelial terminal differentiation [4], nervous growth [5,6], embryonic development [7], apoptosis [8,9], and transcriptional regulation of gene expression [10,11]. Also, some human diseases, such as psoriasis, multiple sclerosis, rheumatoid arthritis Alzheimer's disease, and various types of cancers are associated with the PAD enzymes and their citrullinated targets [12–16].

Human peptidylarginine deiminase 4 (PAD4) is one of the PAD isoforms. To date, five isoforms of human PADs, PAD1–4 (PAD5 is the same as PAD4) and PAD6, have been identified, and they display tissue-specific expression patterns. PAD1 is found in the skin epidermis and citrullinates keratins and filaggrins. PAD2 is found in the brain and muscle tissues and citrullinates myelin basic proteins. PAD3 is found in hair follicles and citrullinates trichohyalin. PAD4 is found in granulocytes, monocytes and macrophages and citrullinates histones (H2A, H3 and H4) and nucleophosmin/B23. Finally, PAD6 has been identified in embryonic stem

cells and oocytes [3,17–21]. In addition to those listed above, the proteins that have been identified as PAD substrates comprise vimentin [22], fibrin [12], fibrinogen [23], alpha-enolase and collagen type I [13–16]. Although some protein substrates of PADs have been identified, the protein substrate specificity of PADs remains unknown. The sequence identities among the five PAD isoforms are approximately 50–55%. The catalytic and calcium binding residues are almost all conserved. PAD4 is the only isoform that is located in the nucleus because PAD4 contains a nuclear localization signal (NLS). Nuclear proteins, histones and nucleophosmin/B23, which are citrullinated by PAD4, may be associated with apoptosis [24]. Some studies have shown that citrullination correlates with specific biological events, such as inflammation, apoptosis, trauma, aging and histone-related gene expression and regulation [10,11,25,26]; however, little is known about the physiological roles of PAD4 and other PAD isoforms *in vivo*.

Rheumatoid arthritis (RA) is a chronic, autoimmune disease that is thought to be associated with PAD4 [27]. This disease produces pain, swelling, and stiffness of the synovial joints, joint inflammation and damage to bone and cartilage. Large numbers of autoantibodies against citrullinated proteins are often detected in the blood of patients with RA [28], including factors that are strongly associated with RA, indicating that the citrullinated proteins in RA patients are major autoantigenic epitopes of anti-citrullinated protein antibodies. A number of antibodies in sera of

RA patients, such as anti-perinuclear factor, anti-keratin antibody, anti-Sa antibody (Sa antigen has been suggested to be identical to citrullinated vimentin), anti-cyclic citrullinated peptide (CCP) antibodies and anti-citrullinated fibrinogen (CitFib) antibody, are produced early in the progression of the disease and are more specific for RA than rheumatoid factor [14,15,29,30]. As a result, antibodies against these proteins can be used as diagnostic markers for RA. A genome-wide, single-nucleotide polymorphism (SNP)-based analysis of RA patients identified a specific haplotype of the PAD4 gene, which suggests that this enzyme seems to increase the susceptibility for RA in Japanese and Korean populations [27, 31,32]. These SNPs in the PAD4 gene are thought to make the PAD4 mRNA more stable [27]. In addition, the PAD4 SNPs increase the PAD4 enzyme activity to enhance apoptosis through the mitochondrial pathway [33]. Furthermore, PAD inhibitors are drug development targets. The Cl- and F-amidine inhibitors are synthesized to effectively inhibit PAD1, 3 and 4 activity [34–36].

Overexpression of PAD4 induces apoptosis in hematopoietic cells. PAD4 also mediates cell growth arrest and apoptosis, both of which are associated with expression of the tumor suppressor p53 [37]. PAD4 is involved in the repression of p53 target gene expression which interacts with the C-terminus of p53 and further regulates p53 target genes [26,38]. Since PAD4 functions as a p53 corepressor, inhibitors of this enzyme are considered potential treatments for cancer therapy. The synthetic Cl- and F-amidine inhibitors have also been used to examine their inhibitory effects on PAD4 activity to evaluate cancer cell survival rates [34,36,38].

A recent report has shown that PAD4 is autocitrullinated *in vitro* and *in vivo*, and this modification inactivates enzyme function and augments its recognition by human autoantibodies [39]. As a result, antibodies against these proteins can be used as diagnostic markers for RA. In contrast, Thompson's group reports that autocitrullination of PAD4 and PAD4 SNPs does not alter the enzyme activities, substrate specificity and calcium dependence of the enzyme but alters the ability to bind the PAD4-interacting proteins such as histone deacetylase 1, citrullinated histone H3, and protein methyltransferase 1 [40].

Multiple X-ray structures of PAD4 in complex with ligands and calcium ions have been solved and are available in the Protein Data Bank [41,42]. PAD4 is a homodimeric enzyme with an elongated rubber boot structure (Figure 1A). Each monomer contains a separate active site with two calcium ions (Ca1, 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 (Figure 1A). Calcium-free, calcium-bound and substrate-bound PAD4 structures indicate that binding of Ca<sup>2+</sup> to the acidic concave surface of the enzyme induces a conformational change that subsequently creates the active site cleft [41].

The overall quaternary structure of PAD4 shows that PAD4 is a dimer of two monomers that make head-to-tail contact (Figure 1A). Although the dimeric form seems to be the functional unit for PAD4, it is still unclear whether the dimeric form is unique to PAD4 or common to all PADs. Furthermore, the relationship between PAD4 enzyme activity and dimer formation is also unclear. To address these questions, we created a series of PAD4 mutants of residues Arg8, Tyr237, Glu281, Tyr435, and Asp547, which are located on the dimer interface (Figures 1B and 1C). The subunit-subunit interactions at the dimer interface were disrupted by site-directed mutagenesis, which resulted in the dissociation of the dimers into monomers. The quaternary structures of these dimer interface mutants were analyzed by analytical ultracentrifugation. Our data indicate that the PAD4 monomer is less active than the dimer, which suggests that dimerization is required for the full enzyme activity of PAD4.

## Results

The amino acid residues at the dimer interface, including Arg8, Tyr237, Asp273, Glu281, Tyr435, Arg544 and Asp547 (Figure 1), may be involved in subunit-subunit interactions. The following mutants were constructed: R8A, R8E, R8H, R8K, R8L and R8Q; D547A, D547E and D547N; Y237A; D273A; E281A; R544A; and Y435A and Y435N.

### Quaternary Structures of Human WT and Mutant PAD4s

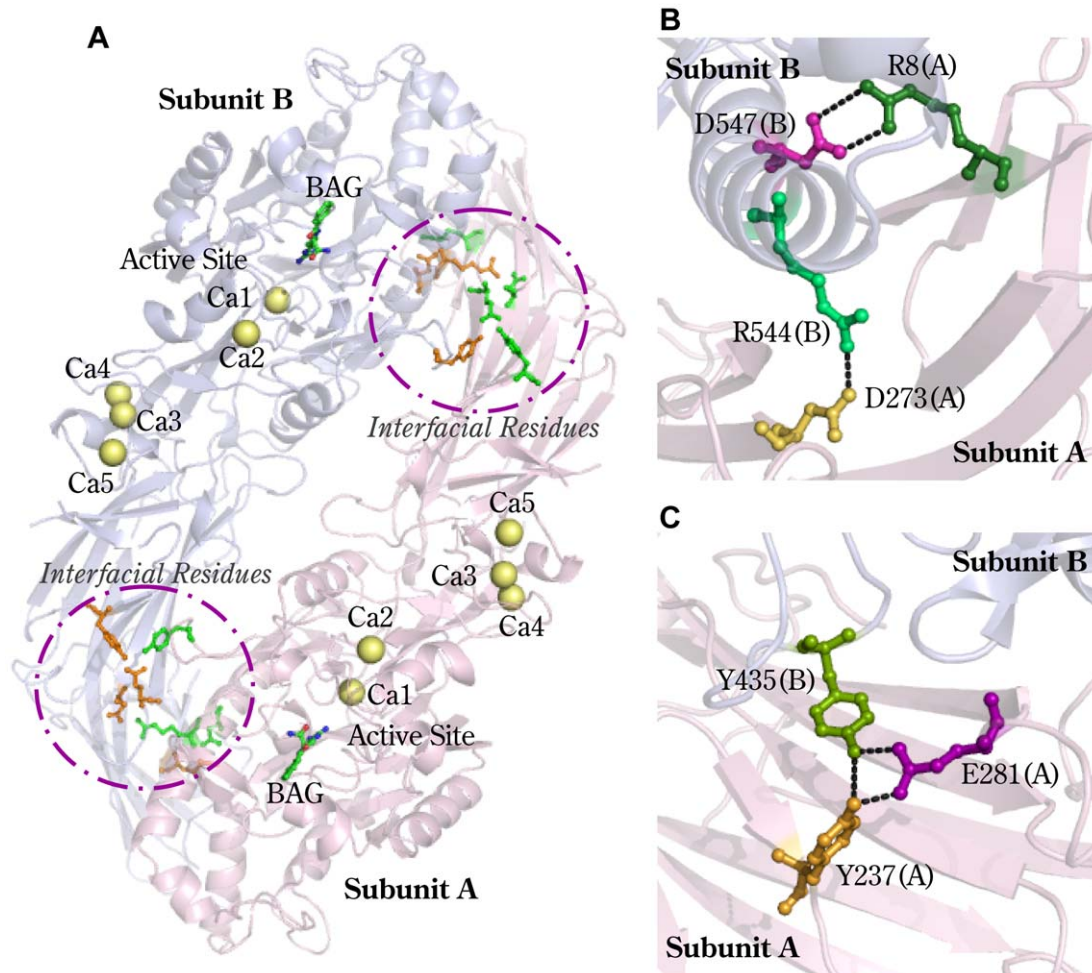
The quaternary structures of the WT and mutant PAD4 proteins were examined by analytical ultracentrifugation (Figures 2, 3 and 4), and the dissociation constants ( $K_d$ ) between the dimers and monomers were determined (Table 1). Differences in the size distributions among WT and mutant enzymes were analyzed by sedimentation velocity experiments. The WT enzyme principally exists as a dimer in solution, with a  $K_d$  value of 0.45  $\mu$ M (Figure 2A). At the dimer interface, Arg8 of subunit A is ion-paired with Asp547 of subunit B, and Asp273 of subunit A is ion-paired with Arg544 of subunit B (Figure 1B). In addition, Tyr237 and Glu281 of subunit A form a hydrogen-bonding network with Tyr435 of subunit B (Figure 1C). The single mutants, Y237A and E281A, existed as dimers, with  $K_d$  values of 0.29  $\mu$ M and 0.1  $\mu$ M, respectively (Figure 2, B and C). The double mutants Y237A/E281A and D273A/R544A also dimerized and had  $K_d$  values of 0.1  $\mu$ M and 0.68  $\mu$ M, respectively (Figure 2, D and E), which suggests that the salt bridge between Asp273 and Arg544 and the Tyr237 and Glu281-mediated hydrogen-bonding network are insignificant for PAD4 dimer formation.

The mutant with a single substitution of Arg8 with a histidine (R8H) did not dissociate into monomers, and the  $K_d$  value of this mutant was 0.47  $\mu$ M (Figure 2E); however, substitution of Arg8 with Ala, Lys or Gln caused the dimeric enzyme to dissociate (Figure 3, B, C and D, respectively). The  $K_d$  values for the R8A, R8K and R8Q single mutants were 9.3  $\mu$ M, 10.2  $\mu$ M and 15.7  $\mu$ M, respectively. Asp547, the binding partner of Arg8, was also mutated into Ala, Glu and Asn. The  $K_d$  values for D547A, D547E and D547N mutant enzymes were 6.4  $\mu$ M, 11.2  $\mu$ M and 4.9  $\mu$ M, respectively (Figures 3, E, F and G, respectively). These results suggest that the electrostatic interactions between Arg8 (subunit A) and Asp547 (subunit B) at the dimer interface are important for stabilization of the dimer; however, removal of the ion pair formed by Arg8 and Asp547 did not cause the enzyme to completely dissociate into monomers. The  $K_d$  value of the R8A/D547A double mutant was 3.9  $\mu$ M (Figure 3H), suggesting that additional factors are involved in PAD4 dimer formation.

Mutation of Arg8 to leucine and glutamate caused the enzyme to predominantly dissociate into monomers, with  $K_d$  values of 16.8  $\mu$ M for R8L and 45.6  $\mu$ M for R8E (Figure 4, B and C). The R8E/D547E double mutant was also monomeric (Figure 4D) and had a  $K_d$  value of 24  $\mu$ M, which also suggests that the electrostatic interactions between Arg8 and Asp547 at the dimer interface are significant.

The Y435A and Y435N mutants were also monomeric (Figure 4, E and F, respectively) and had  $K_d$  values of 30.3  $\mu$ M and 33.8  $\mu$ M, respectively. Tyr435 hydrogen bonds to Tyr237 and Glu281 (Figure 1C). In the Y237A/E281A double mutant, the Tyr435-mediated hydrogen-bonding network was not present. The Y237A/E281A double mutant formed a dimer with a similar  $K_d$  value to the WT, which suggests that the monomeric Y435A and Y435N enzymes did not form dimers due to the abolishment of the hydrogen-bonding network but because they lacked the hydrophobicity or geometry related to Tyr435.

We also examined the effect of Ca<sup>2+</sup> on the quaternary structure of the WT and mutant PAD4 enzymes (Figure 5). Generally, the



**Figure 1. Homodimeric PAD4 and residues located at the dimer interface.** Homodimer of human PAD4 (PDB code: 1WDA, **Panel A**). The active site and calcium-binding site are indicated. Five calcium ions (Ca1–Ca5) are indicated as yellow balls. The substrate analog, benzoyl-L-arginine amide (BAG), is shown as a stick model. Panels B and C display the amino acid residues in the dimer interface of PAD4, which are represented by ball-and-stick models. **Panel B:** Arg8(A) is ion-paired with Asp547(B), and Asp273(A) is ion-paired with Arg544(B). **Panel C:** The hydrogen-bonding network formed by Tyr435(B), Tyr237(A) and Glu281(A). The figures were generated with PYMOL (DeLano Scientific LLC, San Carlos, CA, USA). doi:10.1371/journal.pone.0021314.g001

monomer-dimer equilibrium of PAD4 is not influenced by  $\text{Ca}^{2+}$ . The dissociation constants between the dimeric and monomeric forms of WT and interface mutant PAD4 enzymes in the presence of  $\text{Ca}^{2+}$  ( $K_{d,\text{Ca}}$ ) are listed in Table 1. For the dimeric WT enzyme, the  $K_{d,\text{Ca}}$  was 0.51  $\mu\text{M}$ , which is similar to the value in the absence of  $\text{Ca}^{2+}$  (0.45  $\mu\text{M}$ ). For the R8K mutant, which existed in a monomer-dimer equilibrium, the  $K_{d,\text{Ca}}$  was 7.8  $\mu\text{M}$ , which is slightly lower than the  $K_d$  without  $\text{Ca}^{2+}$  (10.2  $\mu\text{M}$ ; Table 1); however, the monomer peak of the R8K mutant did not shift to form more dimers (Figure 5, C and D). Similar results were observed with the Y435N and R8E mutant enzymes. The  $K_{d,\text{Ca}}$  of Y435N was 19.8  $\mu\text{M}$ , which is lower than the value without  $\text{Ca}^{2+}$  (33.8  $\mu\text{M}$ ). For R8E, the  $K_{d,\text{Ca}}$  was 22.3  $\mu\text{M}$ , which is lower than the value without  $\text{Ca}^{2+}$  (45.6  $\mu\text{M}$ ). Although the  $K_{d,\text{Ca}}$  values of these two mutant enzymes were lower than the values without  $\text{Ca}^{2+}$ , the monomeric Y435N and R8E enzymes did not significantly form dimers (Figure 5, E–H). Therefore, the effect of  $\text{Ca}^{2+}$  on these mutant enzymes may contribute to the conformational stability of the monomeric structure but not to dimer formation.

#### Kinetic properties of WT and interface mutant PAD4

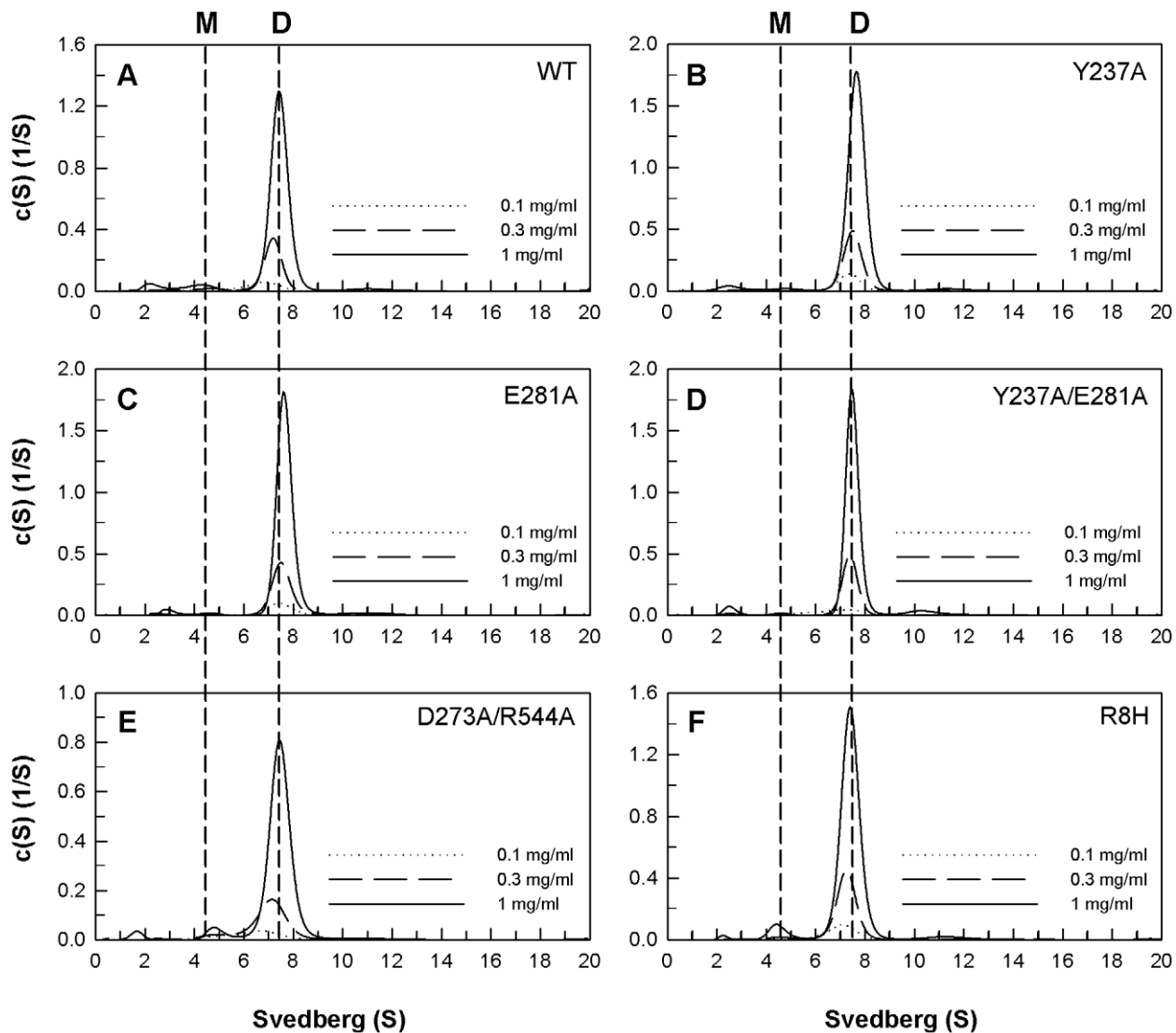
For the dimeric PAD4 enzyme, the active site of each monomer is independent and located far away from the dimer interface.

Based on this, we assumed that the changes in the enzyme kinetics of these mutant enzymes were more significantly affected by dissociation of the enzyme than by mutation of a specific residue.

The kinetic parameters of the WT and mutant PAD4s are shown in Table 1. The PAD4 enzyme binds  $\text{Ca}^{2+}$  ions cooperatively. The initial velocities of PAD4, which were measured at various concentrations of  $\text{Ca}^{2+}$ , exhibited sigmoidal kinetics. Table 1 summarizes the results obtained by fitting these sigmoidal curves to the Hill equation. The half-saturation constant for  $\text{Ca}^{2+}$  ( $K_{0.5,\text{Ca}}$ ) and the degree of cooperativity of calcium binding ( $h$ ) were estimated.

The Michaelis constant of *in vitro* substrate BAEE ( $K_{m,\text{BAEE}}$ ) and  $K_{0.5,\text{Ca}}$  values of the PAD4 interface mutant were roughly similar to those of the WT enzyme, indicating that the active site was not significantly perturbed by the mutations; however, the  $k_{\text{cat}}$  values of some of the interface mutants were less than that of the WT enzyme. For dimeric WT PAD4, the  $k_{\text{cat}}$  value was about 13.4  $\text{s}^{-1}$ . For the R8H, Y237A, Y281A, Y237A/Y281A and D273A/R544A mutants, which existed predominantly as dimers, the  $k_{\text{cat}}$  values were almost the same as that of the WT enzyme.

For the R8A, R8K, R8Q, D547A, D547E, D547N and R8A/D547A mutant enzymes, which existed in monomer-dimer equilibria, the  $k_{\text{cat}}$  values ranged from 8.9 to 12.2  $\text{s}^{-1}$ , which



**Figure 2. Continuous sedimentation coefficient distributions of the dimeric PAD4 WT and interface mutants.** The enzyme concentrations used in the experiments were 0.1, 0.3 and 1 mg/mL in 50 mM Tris-HCl and 250 mM NaCl (pH 7.6) at 20°C. (A) WT; (B) Y237A mutant; (C) E281A mutant; (D) Y237A/E281A double mutant; (E) D273A/R544A double mutant; (F) R8H mutant.  
doi:10.1371/journal.pone.0021314.g002

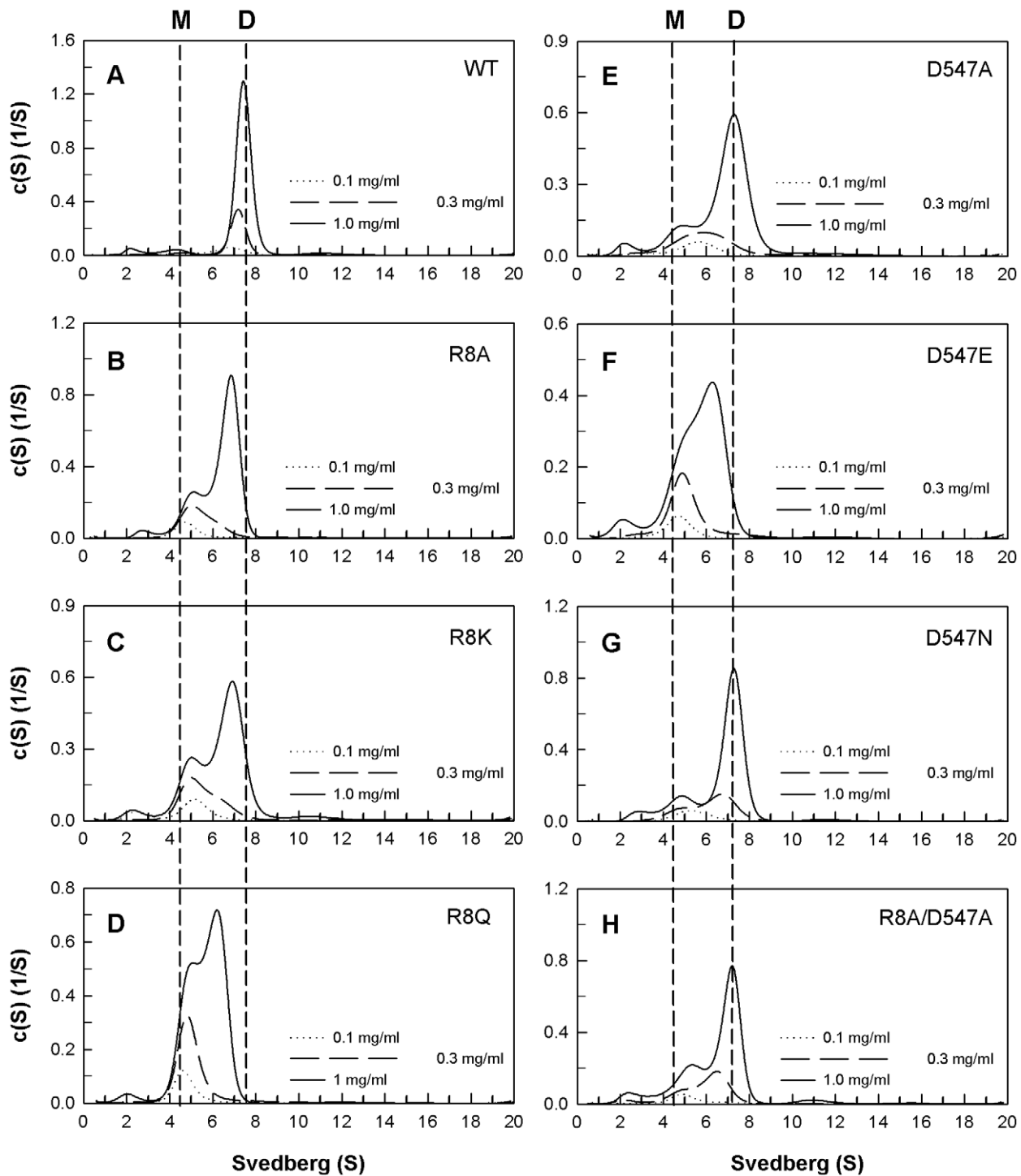
represents a slight decrease in the  $k_{cat}$  and implies that the enzyme activity may be reduced by disrupting the interactions at the subunit interface. For the monomeric PAD4 interface mutants R8L, R8E, R8E/D547E, Y435A and Y435N, the  $k_{cat}$  values ranged from 3.3 to 7.3  $s^{-1}$ . The significant decrease in the  $k_{cat}$  values of these mutants indicates that the enzymatic activity was reduced due to the dissociation of the dimeric enzyme into monomers. The catalytic activity of monomeric PAD4 was approximately half or less than half of the WT enzyme, which suggests that the monomeric PAD4 enzyme is still active but less active than the dimeric enzyme. In addition, the nine-fold increase in the  $K_{m,BAEE}$  and  $K_{0.5,Ca}$  values and the greater than three-fold decrease in the  $k_{cat}$  values for the R8E/D547E and Y435N mutants may have partially resulted from the instability of the monomeric form of the enzyme, suggesting that dimer formation is advantageous for PAD4 stability.

The  $h$  value for WT PAD4 was greater than two; however, the  $h$  values of the interface mutants were significantly reduced (Table 1). The dimeric interface mutant enzymes (Figure 2) had lower  $h$

values of around 1.5 to 2.0, which indicates that calcium-binding cooperativity is very sensitive to changes in the dimer interface, even though the dimeric structure was maintained. The mutant enzymes that showed a monomer-dimer equilibrium (Figure 3) had  $h$  values of 1.3 and 1.6, which also indicates that cooperativity was partially lost for these mutant enzymes when the dimer interface was disrupted. For the monomeric interface mutant enzymes (Figure 4), the  $h$  value was reduced to 1, indicating that the calcium-binding cooperativity for these mutant enzymes was totally lost. The PAD4 monomers were less active and more non-cooperative enzymes.

## Discussion

PAD4 is a homodimer with a fully functional active site in each subunit. The correlation between the catalytic efficiency and dimer formation of this enzyme is unknown. This paper establishes the significance of the dimeric structure of PAD4 for catalysis, regulation and stability.

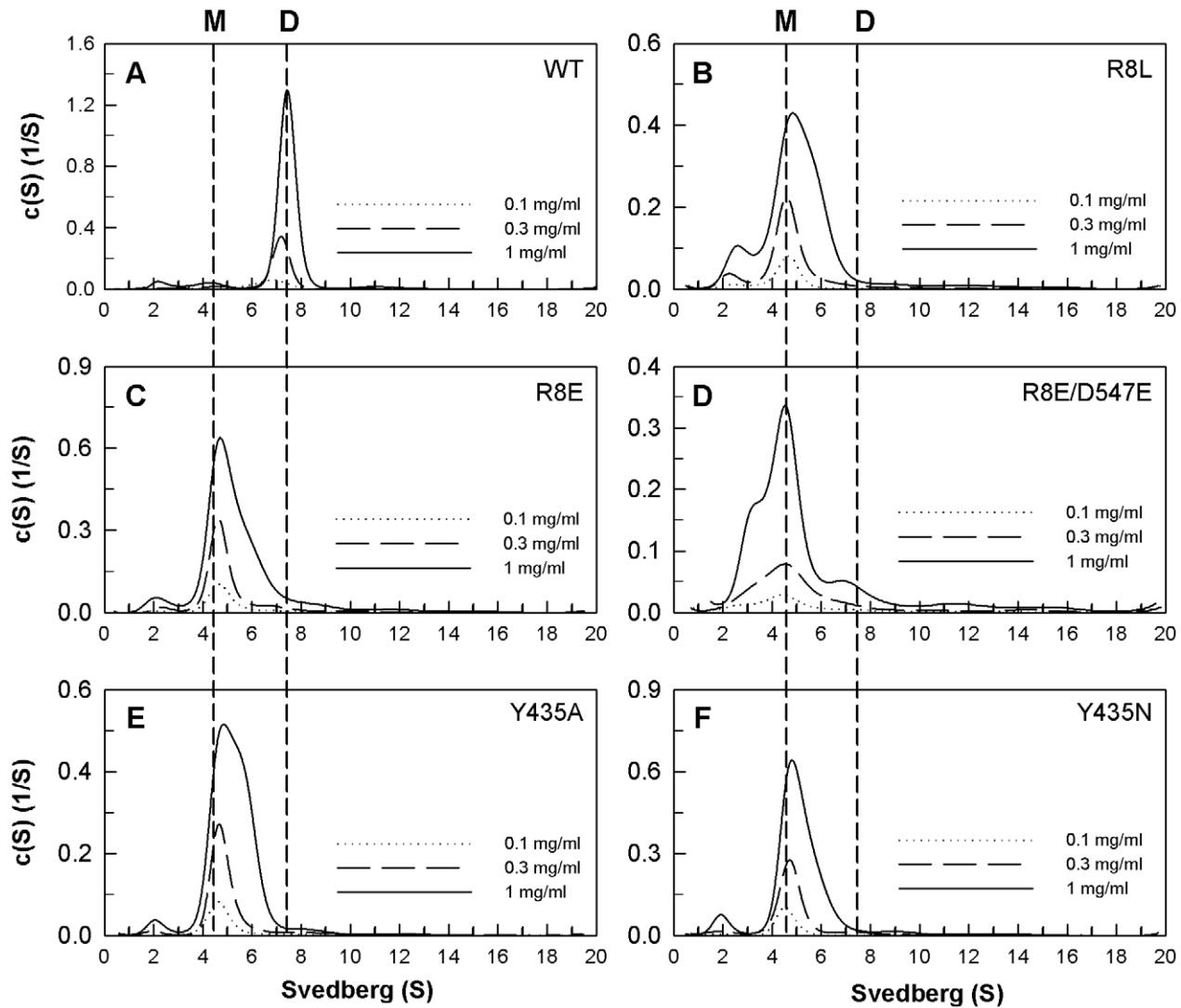


**Figure 3. Continuous sedimentation coefficient distributions of the PAD4 interface mutants in monomer-dimer equilibria.** The enzyme concentrations used in the experiments were 0.1, 0.3 and 1 mg/mL in 50 mM Tris-HCl and 250 mM NaCl (pH 7.6) at 20°C. (A) WT; (B) R8A mutant; (C) R8K mutant; (D) R8Q mutant; (E) D547A mutant; (F) D547E mutant; (G) D547N mutant; (H) R8A/D547A double mutant. doi:10.1371/journal.pone.0021314.g003

#### Enzyme regulation and subunit-subunit interactions of PAD4

The analytical ultracentrifugation and enzyme kinetics analyses clearly revealed that disruption of the dimer interfaces of PAD4 causes the mutants to be less active than the WT protein. Figure 6 shows the correlation between the  $K_d$  and  $k_{cat}$  values, the  $K_d$  and  $h$  values and the  $k_{cat}$  and  $h$  values. The  $k_{cat}$  values of these interface

mutants decreased with increasing  $K_d$  values (Figure 6A), suggesting that the dissociation of dimers into monomers considerably influences the activity of the enzyme. Furthermore, the  $h$  values of these interface mutants also decreased with increasing  $K_d$  values, and the  $K_d$  value was closest to 1 when the enzyme dissociated into monomers (Figure 6B), suggesting that the organization at the dimer



**Figure 4. Continuous sedimentation coefficient distributions of the monomeric PAD4 interface mutants.** The enzyme concentrations used in the experiments were 0.1, 0.3 and 1 mg/mL in 50 mM Tris-HCl and 250 mM NaCl (pH 7.6) at 20°C. (A) WT; (B) R8L mutant; (C) R8E mutant; (D) R8E/D547E double mutant; (E) Y435A mutant; (F) Y435N mutant.  
doi:10.1371/journal.pone.0021314.g004

interface is critical for proper calcium-binding cooperativity. It is not surprising that the increased  $k_{cat}$  values of these interface mutants corresponded to greater  $h$  values (Figure 6C). We suggest that the dimer is the fully functional form of human PAD4 because dimer formation correlated with catalytic activity and cooperativity.

#### Ionic and hydrophobic interactions stabilize the dimer interface of PAD4

Size distribution analyses demonstrated that the electrostatic interaction between Arg8 and Asp547 and the Tyr435-mediated interactions determine the dimer formation of PAD4. In contrast, the ionic interaction between Asp273 and Arg544 did not seem to be the determinant for PAD4 dimer formation.

Our data also suggested that steric effects are important for the subunit-subunit organization of the PAD4 dimer. The R8H enzyme existed as a dimer while the R8K enzyme was in a monomer-dimer equilibrium (Figures 2 and 3). Similarly, D547E formed a greater proportion of monomers than D547A and D547N (Figure 3). The R8K and D547E enzymes cannot form

salt bridges at the dimer interface of the enzyme as well as WT. The improper geometry of the mutant enzyme causes it to dissociate even though the charges of its side chains are preserved.

#### The dimer interface residues of peptidylarginine deiminase isoforms

Figure 7 displays a multiple sequence alignment of the PAD isoforms. Asp547 is highly conserved among the PAD enzymes. Arg8 has a conservative substitution for His, and residue 435 is conserved as a hydrophobic amino acid (Figure 7B). For the PAD4 enzymes, residue 8 is Arg or His, residue 547 is Asp and residue 435 is Tyr. We hypothesize that the PAD4 enzyme in different species should be dimers because the human R8H enzyme is as stable as a dimer as the WT protein. Residues 8 and 547 in PAD2 and PAD3 are Arg and Asp, respectively. Residue 435 is Phe in PAD2 and Leu in PAD3, both of which are hydrophobic residues similar to Tyr. We postulate that Tyr435 in human PAD4 is involved in the hydrophobic interactions in the dimer interface of the enzyme and that Phe435 in PAD2 and Leu

**Table 1.** Kinetic Parameters of the Human WT and Mutant PAD4s.

PAD4	$K_{m,BAEE}^a$ (mM)	$K_{0.5,Ca}^b$ (mM)	$k_{cat}^c$ ( $s^{-1}$ )	$h^d$	Quaternary Structure <sup>e</sup>	$K_d^f$ ( $\mu$ M)	$K_{d,ca}^f$ ( $\mu$ M)
WT	0.31±0.03	0.71±0.02	13.4±0.19	2.30±0.13	D	0.45±0.003	0.51±0.004
R8A	0.47±0.03	0.71±0.04	12.2±0.24	1.50±0.12	M/D	9.3±0.08	ND <sup>g</sup>
R8E	1.06±0.08	1.65±0.41	7.3±0.71	1.01±0.12	M	45.6±0.37	22.3±0.12
R8H	0.47±0.06	0.58±0.04	13.3±0.32	1.71±0.17	D	0.47±0.005	ND <sup>g</sup>
R8K	0.50±0.08	0.82±0.04	10.8±0.19	1.54±0.10	M/D	10.2±0.11	7.8±0.08
R8L	1.36±0.17	1.72±0.52	5.3±0.59	1.06±0.18	M	16.8±0.14	ND <sup>g</sup>
R8Q	0.60±0.07	1.07±0.20	12.2±0.86	1.35±0.29	M/D	15.7±0.14	ND <sup>g</sup>
D547A	0.36±0.05	0.77±0.09	11.6±0.48	1.26±0.16	M/D	6.4±0.07	ND <sup>g</sup>
D547E	0.72±0.10	1.24±0.13	8.9±0.35	1.26±0.11	M/D	11.2±0.11	ND <sup>g</sup>
D547N	0.86±0.10	1.02±0.05	12.2±0.25	1.81±0.12	M/D	4.9±0.05	ND <sup>g</sup>
R8A/D547A	0.40±0.07	0.51±0.03	10.7±0.26	1.57±0.17	M/D	3.9±0.04	ND <sup>g</sup>
R8E/D547E	2.77±0.67	2.88±0.51	3.3±0.23	1.11±0.13	M	24.0±0.23	ND <sup>g</sup>
Y237A	0.36±0.04	0.38±0.02	13.9±0.28	1.79±0.19	D	0.29±0.002	ND <sup>g</sup>
E281A	0.45±0.05	0.39±0.01	12.1±0.19	1.95±0.14	D	0.10±0.001	ND <sup>g</sup>
Y435A	2.33±0.43	1.80±0.21	5.9±0.26	1.08±0.07	M	30.3±0.26	ND <sup>g</sup>
Y435N	2.73±0.45	2.16±0.24	4.0±0.19	1.21±0.08	M	33.8±0.24	19.8±0.19
Y237A/E281A	0.38±0.05	0.82±0.06	13.2±0.41	1.60±0.16	D	0.10±0.001	ND <sup>g</sup>
D273A/R544A	0.35±0.04	0.97±0.08	13.5±0.33	1.49±0.09	D	0.68±0.006	ND <sup>g</sup>

<sup>a</sup> $K_{m,BAEE}$ : Michaelis constant for benzoyl-L-arginine ethyl ester (BAEE) as the *in vitro* substrate for PAD4.

<sup>b</sup> $K_{0.5,Ca}$ : half-saturation constant for  $Ca^{2+}$ .

<sup>c</sup> $k_{cat}$ : catalytic constant.

<sup>d</sup> $h$ : the Hill coefficient.

<sup>e</sup>M: monomer; D: dimer.

<sup>f</sup> $K_d$ : dissociation constant between the monomer and dimer without  $Ca^{2+}$ ;  $K_{d,ca}$ : dissociation constant between the monomer and dimer with 10 mM  $Ca^{2+}$ .

<sup>g</sup>ND: not determined.

doi:10.1371/journal.pone.0021314.t001

in PAD3 should have similar hydrophobic effects in the interface of the enzyme. Therefore, we propose that the PAD2 and PAD3 enzymes may also be stable dimers. For some PAD1 enzymes in different species, residue 8 is Gln, and the other two residues are conserved. It is possible that the PAD1 enzymes might be unstable dimers because the human R8Q enzyme displayed a monomer-dimer equilibrium. For PAD6, residue 435 is Tyr, but the ion pair between residues 8 and 547 is not conserved. Our data indicate that the human R8A/D547A enzyme is in equilibrium between the monomeric and dimeric forms, with a nine-fold larger  $K_d$  value than the WT protein. The PAD6 enzyme, in which Tyr435 is conserved, may sufficiently dimerize; however, the quaternary structures of the PAD isoforms need further investigation in the future.

## Conclusion

In summary, we suggest that the dimerization of PAD4 is essential for full enzyme activity and calcium-binding cooperativity. Monomeric PAD4 maintains 25–50% enzyme activity compared to the dimeric enzyme, suggesting that the PAD4 monomer is still active. Although PAD4 is the only isoform of the PAD family for which a crystal structure has been determined [42], our data suggest that other PAD isoforms may also exist as dimers with different association affinities. In addition, excess PAD4 expression seems to be associated with rheumatoid arthritis. Our data may provide a rationale for the development of PAD-inhibiting drugs by disrupting the dimerization of PAD4 to decrease its activity.

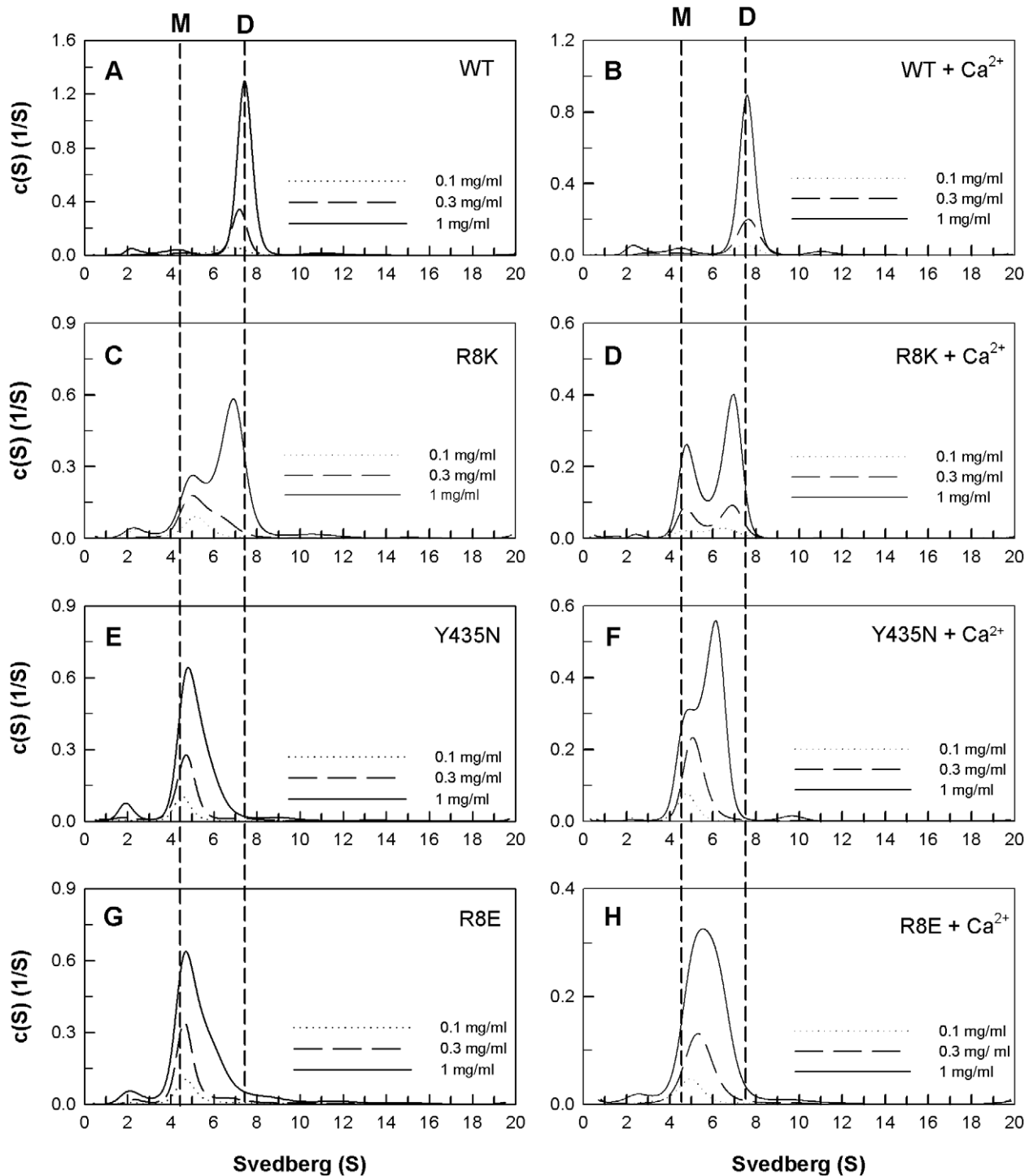
## Materials and Methods

### Site-Directed Mutagenesis

Site-directed mutagenesis was performed using the QuikChange<sup>TM</sup> mutagenesis kit (Stratagene, USA) to create the mutant human PAD4 constructs. The human PAD4 DNA was purchased from RZPD Deutsches Ressourcenzentrum für Genomforschung. The purified PAD4 DNA was used as template, and primers with the desired codon changes were used for PCR (polymerase chain reaction) with high fidelity *Pfu* DNA polymerase. The primers with the desired mutations ranged from 25 to 45 bases in length, which were required for specific binding to the template DNA. After 16–18 temperature cycles, mutated plasmids that included staggered nicks were produced. The PCR products were treated with DpnI to digest the wild-type human PAD4 templates, and the nicked DNA with the desired mutations was transformed into the XL-1 strain of *Escherichia coli*. The sequences of the mutants were verified by dideoxy-based sequencing.

### Expression and Purification of Recombinant PAD4

PAD4 was sub-cloned into the pQE30 vector, which carries an N-terminal 6His tag. This ampicillin-resistant vector was transformed into the JM109 strain of *Escherichia coli*. PAD4 expression was induced with 1.0 mM isopropyl-1-thio- $\beta$ -D-galactoside (IPTG) at 25°C overnight. The over-expressed enzyme was purified using Ni-NTA Sepharose (Sigma). The lysate-Ni-NTA mixture was washed in a stepwise fashion (5, 10 and 20 mM imidazole in 500 mM sodium chloride, 2.5 mM dithiothreitol and



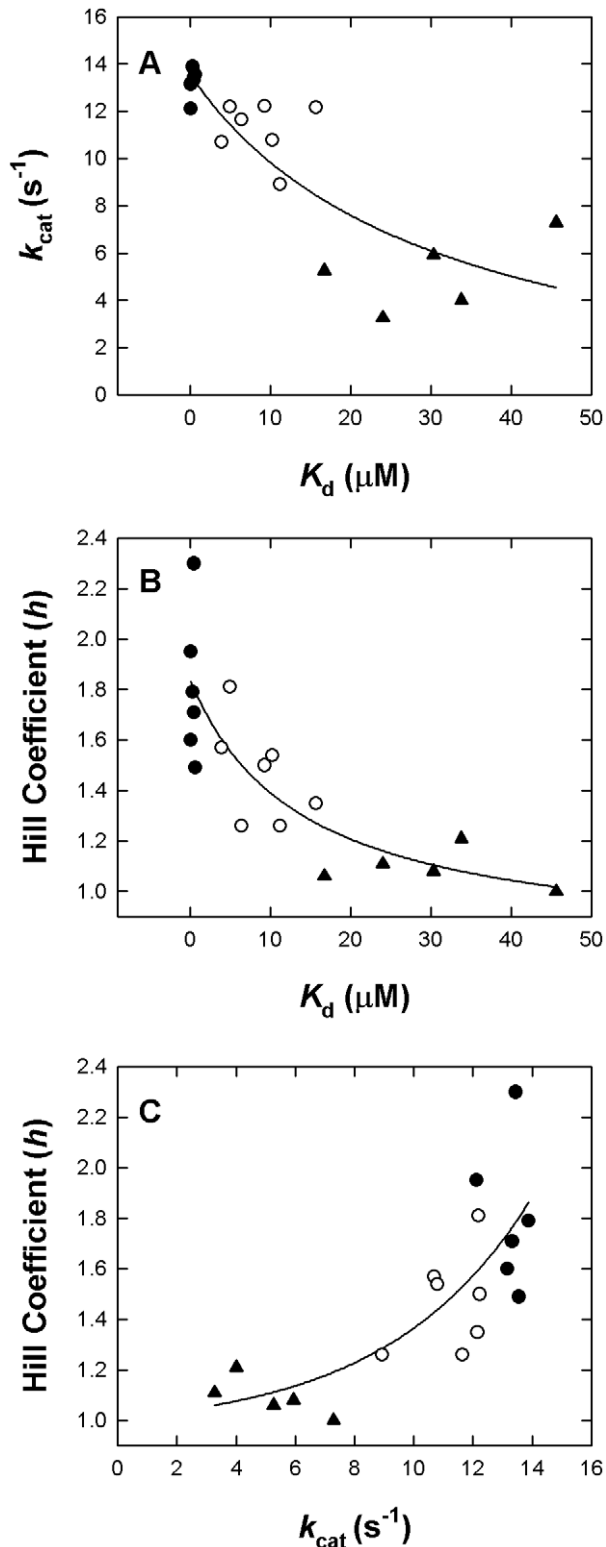
**Figure 5. Continuous sedimentation coefficient distributions of the WT PAD4 and dimer interface mutants in the presence of calcium ions.** The enzyme concentrations were 0.1, 0.3 and 1.0 mg/mL in 50 mM Tris-HCl and 250 mM NaCl (pH 7.6) at 20°C. The concentration of the calcium ion was 10 mM. (A) WT; (B) WT with Ca<sup>2+</sup>; (C) R8K mutant; (D) R8K with Ca<sup>2+</sup>; (E) Y435N mutant; (F) Y435N with Ca<sup>2+</sup>; (G) R8E mutant; (H) R8E with Ca<sup>2+</sup>.

doi:10.1371/journal.pone.0021314.g005

30 mM Tris-HCl, pH 7.6) to remove unwanted proteins, and then PAD4 was eluted with elution buffer (80 mM imidazole, 500 mM sodium chloride, 2.5 mM dithiothreitol and 30 mM Tris-HCl, pH 7.6). The purified enzyme was then buffer exchanged into 50 mM Tris-HCl (pH 7.6), 500 mM NaCl and 2 mM  $\beta$ -mercaptoethanol and concentrated using a centrifugal

filter device (Amicon Ultra-15, Millipore) with a molecular weight cutoff of 30 kDa. The enzyme purity was verified by resolving the eluted protein with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and staining with Coomassie, and protein concentrations were approximately calculated with the Bradford method [43].





**Figure 6. Correlation plots for the dissociation constant ( $K_d$ ), Hill coefficient ( $h$ ), and catalytic constant ( $k_{cat}$ ).** The data points were derived from Table 1. Closed circle: dimeric PAD4 (WT, R8H, Y237A, E281A, Y237A/E281A, and D273A/R544A); open circle: PAD4 in dimer-monomer equilibrium (R8A, R8K, R8Q, D547A, D547E, D547N, and R8A/D547A); closed triangle: monomeric PAD4 (R8E, R8L, R8E/D547E, Y435A, and Y435N). **Panel A:**  $K_d$  versus  $h$ . **Panel B:**  $K_d$  versus  $k_{cat}$ . **Panel C:**  $k_{cat}$  versus  $h$ .  
doi:10.1371/journal.pone.0021314.g006

### Enzyme assay of PAD4 and analysis of kinetic data

A protocol for the continuous measurement of PAD4 enzyme activity, which is coupled to a glutamate dehydrogenase-catalyzed reaction, has been reported [44]. The reaction mixture for the spectrophotometric assay of PAD4 activity contained 10 mM benzoyl-L-arginine ethyl ester (BAEE) as the *in vitro* substrate for PAD4, 10 mM CaCl<sub>2</sub>, 2.5 mM dithiothreitol, 8.5 mM  $\alpha$ -ketoglutarate ( $\alpha$ -KG), 0.22 mM NADH and 8.4 U of glutamate dehydrogenase (GDH) in 100 mM Tris-HCl (pH 7.5) in a total volume of 1 ml at 25°C. The reaction was initiated by the addition of the appropriate amount of enzyme to the reaction mixture, and the decrease in absorbance at 340 nm was continuously monitored using a Perkin-Elmer Lambda-25 spectrophotometer. An enzyme unit is defined as the amount of enzyme that catalyzed the decomposition of 1  $\mu$ mol of NADH per min. An extinction coefficient of 6220 M<sup>-1</sup> for NADH was employed in the calculations. The apparent Michaelis constants were determined for the BAEE substrate by varying its concentration near its  $K_m$  value with invariable and saturated concentrations of the other constituents.

Furthermore, the  $k_{cat}$  values of PAD4 were calculated with the following equation:

$$k_{cat}(s^{-1}) = \{[v/6.22]/\mu g \text{ protein}\} \times \text{Molecular weight}\}/60$$

where  $v$  represents  $\Delta A_{340}/\text{min}$ , 6.22 is the millimolar absorption coefficient of NAD(P)H, 148,000 is the molecular weight of the human PAD4 dimer and 60 is the number of seconds per minute.

The sigmoidal curves of  $[Ca^{2+}]$  versus the initial velocities were input into the Hill equation, and the data were further analyzed to calculate the  $K_{0.5, Ca}$  value, which is the calcium concentration at half-maximal velocity, and the Hill coefficient ( $h$ ), which were utilized to evaluate the degree of cooperativity using the following equation:

$$v = V_{max}[Ca^{2+}]^h / (K_{0.5, Ca}^h + [Ca^{2+}]^h)$$

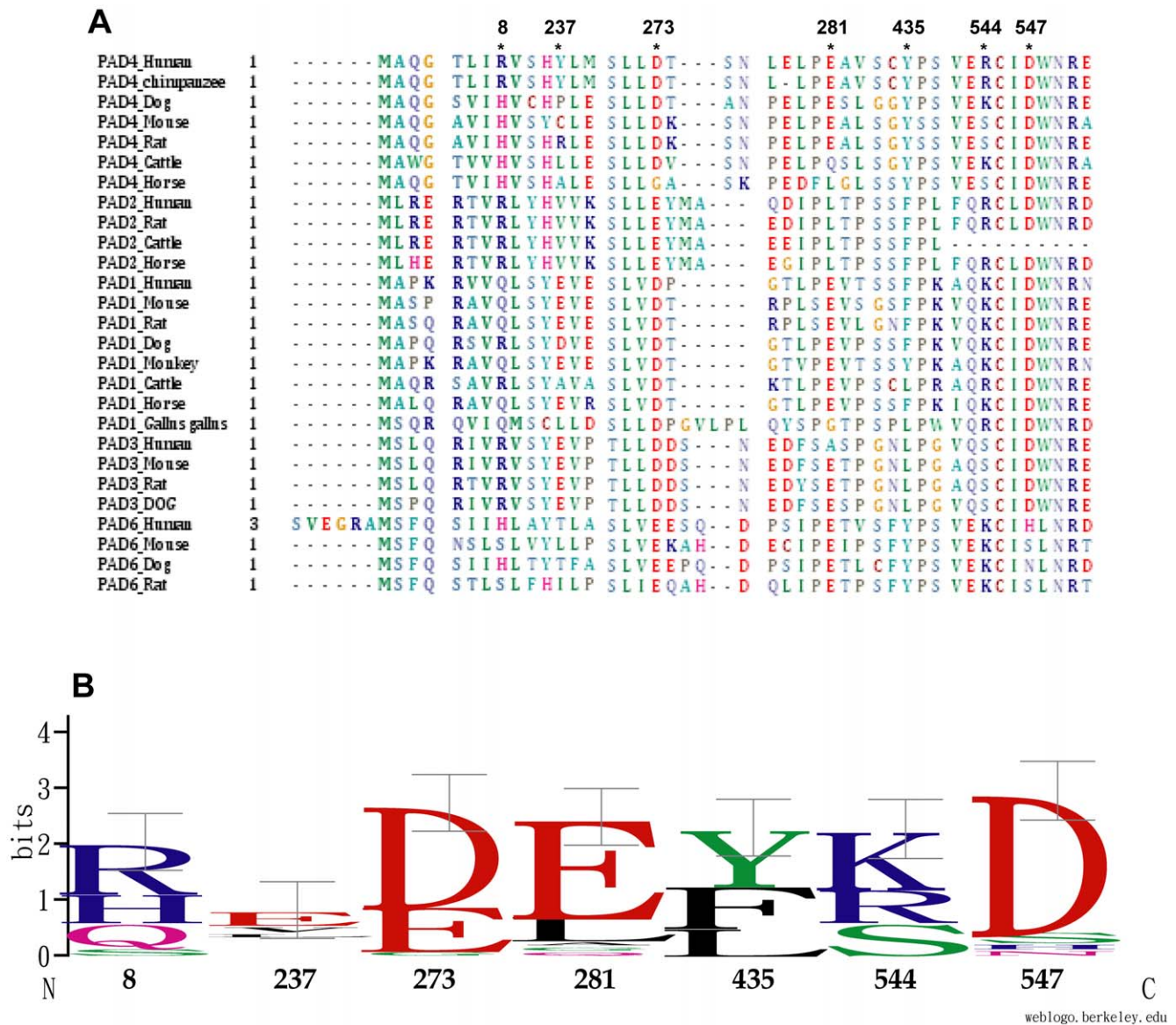
Graphical analysis of the data was performed with the Sigma Plot 8.0 program (Jandel, San Rafael, CA).

### Analytical ultracentrifugation

The quaternary structure of PAD4 was analyzed with an Optima XL-A Analytical Ultracentrifuge (Beckman, USA). The concentrations of the PAD4 enzyme in sedimentation velocity (SV) experiments were fixed at 0.1, 0.3, and 1 mg/ml. The sample cell was loaded with 380  $\mu$ l of sample, and the reference cell was loaded with 400  $\mu$ l of 50 mM Tris-HCl and 250 mM NaCl (pH 7.6). After the protein samples were loaded, the cells were transferred into an An-50 Ti analytical rotor. The enzyme was detected at an absorbance of 280 nm in continuous mode with high-speed centrifugation (42000 rpm) for 4 hours at 20°C, with a time interval of 480 seconds and a step size of 0.002 cm. Numerous scans of the sedimentation velocity data were collected and analyzed globally with the SEDFIT 9.4c software [45]. All size distributions were determined with a confidence level of  $p = 0.95$ , a best-fitted average anhydrous frictional ratio ( $f/f_0$ ), and a resolution of 200 sedimentation coefficients between 0.1 and 20.0 S.

### Self-association of the PAD4 enzyme

The dissociation constant ( $K_d$ ) of PAD4 was analyzed using the SEDPHAT program with a monomer-dimer equilibrium model



**Figure 7. Multiple sequence alignments of the PAD family.** The amino acid sequences of PADs were identified using the BLAST [48], and the alignments were generated with ClustalW [49]. **Panel A:** Multiple sequence alignments of 27 isoforms of peptidylarginine deiminase. This figure was generated using the BioEdit sequence alignment editor [50]. **Panel B:** Sequence conservation with error bars for the dimer interface residues. This figure shows the frequency of conservation for the respective amino acid residues at a given position [51]. doi:10.1371/journal.pone.0021314.g007

[45,46]. The sedimentation velocity data, which were collected with three different enzyme concentrations, were globally fitted with SEDPHAT, and the partial, specific volume of the enzyme, solvent density, and viscosity were calculated by the SEDNTERP program [47].

## References

- Rogers GE, Harding HW, Llewellyn-Smith IJ (1977) The origin of citrulline-containing proteins in the hair follicle and the chemical nature of trichohyalin, an intracellular precursor. *Biochim Biophys Acta* 495: 159–175.
- Tarcsa E, Marekov LN, Mei G, Melino G, Lee SC, et al. (1996) Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. *J Biol Chem* 271: 30709–30716.
- Vossenaar ER, Zendman AJ, van Venrooij WJ, Pruijn GJ (2003) PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* 25: 1106–1118.
- Senshu T, Akiyama K, Ishigami, A, Nomura K (1999) 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.
- Moscarello MA, Wood DD, Ackerley C, Boulias C (1994) Myelin in multiple sclerosis is developmentally immature. *J Clin Invest* 94: 146–154.
- Wood DD, Bilbao JM, O'Connors P, Moscarello MA (1996) Acute multiple sclerosis (Marburg type) is associated with developmentally immature myelin basic protein. *Ann Neurol* 40: 18–24.
- Wright PW, Bolling LC, Calvert ME, Sarmiento OF, Berkeley EV, et al. (2003) ePAD, an oocyte and early embryo-abundant peptidylarginine deiminase-like protein that localizes to egg cytoplasmic sheets. *Dev Biol* 256: 73–88.

8. Asaga H, Yamada M, Senshu T (1998) Selective deimination of vimentin in calcium ionophore-induced apoptosis of mouse peritoneal macrophages. *Biochem Biophys Res Commun* 243: 641–646.
9. Mizoguchi M, Manabe M, Kawamura Y, Kondo Y, Ishidoh K, et al. (1998) Deimination of 70-kD nuclear protein during epidermal apoptotic events in vitro. *J Histochem Cytochem* 46: 1303–1309.
10. Wang Y, Wyssocka J, Sayegh J, Lee YH, Perlin JR, et al. (2004) Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* 306: 279–283.
11. Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, et al. (2004) Histone deimination antagonizes arginine methylation. *Cell* 118: 545–553.
12. Senshu T, Serre G (2001) The major synovial targets of the rheumatoid arthritis-specific anti-filagrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. *J Immunol* 166: 4177–4184.
13. Suzuki A, Yamada R, Ohtake-Yamanaka M, Okazaki Y, Sawada T, et al. (2005) Anti-citrullinated collagen type I antibody is a target of autoimmunity in rheumatoid arthritis. *Biochem Biophys Res Commun* 333: 418–426.
14. Anzilotti C, Pratesi F, Tommasi C, Migliorini P (2010) Peptidylarginine deiminase 4 and citrullination in health and disease. *Autoimmun Rev* 9: 158–160.
15. Gyorgy B, Toth E, Tarcsa E, Falus A, Buzas EI (2006) Citrullination: a posttranslational modification in health and disease. *Int J Biochem Cell Biol* 38: 1662–1677.
16. Ishigami A, Maruyama N (2010) Importance of research on peptidylarginine deiminase and citrullinated proteins in age-related disease. *Geriatr Gerontol Int* 10(Suppl 1): S53–58.
17. Guerrin M, Ishigami A, Mechin MC, Nachat R, Valmary S, et al. (2003) cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type I. *Biochem J* 370: 167–174.
18. Ishigami A, Ohsawa T, Asaga H, Akiyama K, Kuramoto M, et al. (2002) Human peptidylarginine deiminase type II: molecular cloning, gene organization, and expression in human skin. *Arch. Biochem Biophys* 407: 25–31.
19. Kanno T, Kawada A, Yamanouchi J, Yosida-Noro C, Yoshiki A, et al. (2000) 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.
20. Nakashima K, Hagiwara T, Ishigami A, Nagata S, Asaga H, et al. (1999) Molecular characterization of peptidylarginine deiminase in HL-60 cells induced by retinoic acid and lalpha,25-dihydroxyvitamin D(3). *J Biol Chem* 274: 27786–27792.
21. Chavanas S, Mechin MC, Takahara H, Kawada A, Nachat R, et al. (2004) 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.
22. Vossenaar ER, Despres N, Lapointe E, van der Heijden A, Lora M, et al. (2004) Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin. *Arthritis Res Ther* 6: R142–150.
23. Vossenaar ER, Nijenhuis S, Helsen MM, van der Heijden A, Senshu T, et al. (2003) Citrullination of synovial proteins in murine models of rheumatoid arthritis. *Arthritis Rheum* 48: 2489–2500.
24. Nakashima K, Hagiwara T, Yamada M (2002) Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. *J Biol Chem* 277: 49562–49568.
25. Klose R, Zhang Y (2007) Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell Biol* 8: 307–318.
26. Li P, Wang D, Yao H, Doret P, Hao G, et al. (2010) Coordination of PAD4 and HDAC2 in the regulation of p53-target gene expression. *Oncogene* 29: 3153–3162.
27. Suzuki A, Yamada R, Chang X, Tokuhira S, Sawada T, et al. (2003) Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 34: 395–402.
28. van Boeckel MA, Vossenaar ER, van den Hoogen FH, van Venrooij WJ (2002) Autoantibody systems in rheumatoid arthritis: specificity, sensitivity and diagnostic value. *Arthritis Res* 4: 87–93.
29. Reparón-Schuijt CC, van Esch WJ, van Kooten C, Schellekens GA, de Jong BA, et al. (2001) Secretion of anti-citrulline-containing peptide antibody by B lymphocytes in rheumatoid arthritis. *Arthritis Rheum* 44: 41–47.
30. Vossenaar ER, Zendman AJ, Van Venrooij WJ (2004) Citrullination, a possible functional link between susceptibility genes and rheumatoid arthritis. *Arthritis Res Ther* 6: 1–5.
31. Barton A, Bowes J, Eyre S, Spreckley K, Hinks A, et al. (2004) A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. *Arthritis Rheum* 50: 1117–1121.
32. Kang CP, Lee HS, Ju H, Cho H, Kang C, et al. (2006) A functional haplotype of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Koreans. *Arthritis Rheum* 54: 90–96.
33. Hung HC, Lin CY, Liao YF, Hsu PC, Tsay GJ, et al. (2007) 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.
34. Jones JE, Causey CP, Knuckley B, Slack-Noyes JL, Thompson PR (2009) Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Curr Opin Drug Discov Devel* 12: 616–627.
35. Knuckley B, Causey CP, Jones JE, Bhatia M, Dreyton CJ, et al. (2010) 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.
36. Slack JL, Causey CP, Thompson PR (2011) Protein arginine deiminase 4: a target for an epigenetic cancer therapy. *Cell Mol Life Sci* 68: 709–720.
37. Liu GY, Liao YF, Chang WH, Liu CC, Hsieh MC, et al. (2006) Overexpression of peptidylarginine deiminase IV features in apoptosis of haematopoietic cells. *Apoptosis* 11: 183–96.
38. Li P, Yao H, Zhang Z, Li M, Luo Y, et al. (2008) Regulation of p53 target gene expression by peptidylarginine deiminase 4. *Mol Cell Biol* 28: 4745–4758.
39. Andrade F, Darrah E, Gucek M, Cole RN, Rosen A, et al. (2010) Autocitrullination of human peptidyl arginine deiminase type 4 regulates protein citrullination during cell activation. *Arthritis Rheum* 62: 1630–1640.
40. Slack JL, Jones LE, Bhatia MM, Thompson PR (2011) Autodeimination of Protein Arginine Deiminase 4 Alters Protein-Protein Interactions but Not Activity. *Biochemistry In press*.
41. Arita K, Hashimoto H, Shimizu T, Nakashima K, Yamada M, et al. (2004) Structural basis for Ca(2+)-induced activation of human PAD4. *Nat Struct Mol Biol* 11: 777–783.
42. Arita K, Shimizu T, Hashimoto H, Hidaka Y, Yamada M, et al. (2006) Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4. *Proc Natl Acad Sci USA* 103: 5291–5296.
43. Bradford MM (1976) 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.
44. Liao YF, Hsieh HC, Liu GY, Hung HC (2005) A continuous spectrophotometric assay method for peptidylarginine deiminase type 4 activity. *Anal Biochem* 347: 176–181.
45. Schuck P (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* 320: 104–124.
46. Schuck P (2010) Sedimentation patterns of rapidly reversible protein interactions. *Biophys J* 98: 2005–2013.
47. Laue TM, Shah BD, Ridgeway TM, Pelleter SL (1992) Analytical Ultracentrifugation in Biochemistry and Polymer Science, The Royal Society of Chemistry, Cambridge, UK.
48. Altschul SF, Boguski MS, Gish W, Wootton JC (1994) *Nat Genet* 6: 119–129.
49. Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, et al. (1994) ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res* 22: 4673–4680.
50. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95–98.
51. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14: 1188–1190.