Journal of Medicinal Chemistry

Anatolian University Libraries Consortium a·n·k·o·s

Subscriber access provided by HACETTEPE UNIVERSITESI KUTUPHANESI

Article

Development of a selective inhibitor of Protein Arginine Deiminase 2

Aaron Muth, Venkataraman Subramanian, Edward Beaumont, Mitesh Nagar, Philip Kerry, Paul McEwan, Hema Srinath, Kathleen Wanda Clancy, Sangram S. Parelkar, and Paul R. Thompson

J. Med. Chem., **Just Accepted Manuscript** • Publication Date (Web): 22 Mar 2017

Downloaded from http://pubs.acs.org on March 24, 2017

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

Journal of Medicinal Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Development of a selective inhibitor of Protein Arginine Deiminase 2

Aaron Muth[1,2], Venkataraman Subramanian[1], Edward Beaumont,[3] Mitesh Nagar[1], Philip Kerry,[3] Paul McEwan,[3] Hema Srinath[1], Kathleen Clancy[1], Sangram Parelkar[1], Paul R. Thompson[1,4*]

[1]Department of Biochemistry and Molecular Pharmacology, UMass Medical School, 364 Plantation Street, Worcester, MA, 01605, USA.

[2]Department of Pharmaceutical Sciences, College of Pharmacy and Health Sciences, St. John's University, Queens, NY, 11439, USA.

[3]Evotec, Milton Park, Abingdon, Oxfordshire OX14 4RZ, UK.

[4]Program in Chemical Biology, UMass Medical School, 364 Plantation Street, Worcester, MA 01605, USA

[*]Author to whom correspondence should be addressed: Paul R. Thompson

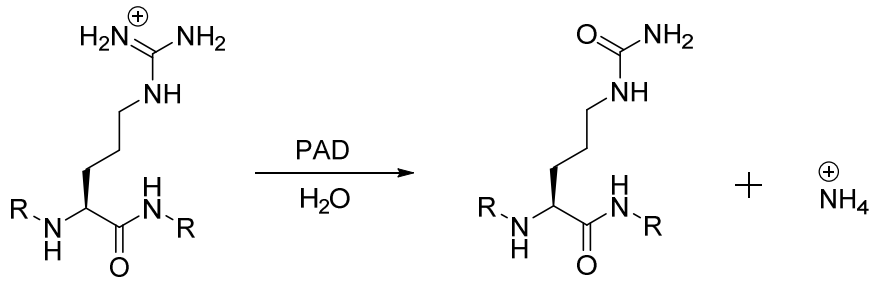
paul.thompson@umassmed.edu

Abstract

Protein arginine deiminase 2 (PAD2) plays a key role in the onset and progression of multiple sclerosis, rheumatoid arthritis and breast cancer. To date, no PAD2-selective inhibitor has been developed. Such a compound will be critical for elucidating the biological roles of this isozyme and may ultimately be useful for treating specific diseases in which PAD2 activity is dysregulated. To achieve this goal, we synthesized a series of benzimidazole-based derivatives of Cl-amidine, hypothesizing that this scaffold would allow access to a series of PAD2-selective inhibitors with enhanced cellular efficacy. Herein, we demonstrate that substitutions at both the N-terminus and C-terminus of Cl-amidine result in >100-fold increases in PAD2 potency and selectivity (**30a**, **41a**, and **49a**) as well as cellular efficacy **30a.** Notably, these compounds use the far less reactive fluoroacetamidine warhead. In total, we predict that **30a** will be a critical tool for understanding cellular PAD2 function and sets the stage for treating diseases in which PAD2 activity is dysregulated.

Introduction.

The protein arginine deiminase (PAD) family of enzymes, which catalyze protein citrullination (Figure 1), have garnered significant recent interest because aberrantly upregulated protein citrullination is associated with a variety of autoimmune diseases (e.g., rheumatoid arthritis (RA), multiple sclerosis, lupus, and ulcerative colitis) as well as certain cancers.[1, 2] In addition to playing an important role in human pathology, the deiminase activity of the PADs regulates a diverse array of cellular processes including NET formation, the epigenetic control of gene transcription, differentiation, and the maintenance of pluripotency.[1, 3-7] There are five PAD isozymes (PADs 1-4 and 6) that are uniquely distributed in various tissues.[8] PADs 1-4 all possess deiminase activity, whereas PAD6 has a number of mutations that render it inactive.[1, 9] The active PAD enzymes (i.e., PADs 1-4) are all regulated by calcium, wherein calcium binding triggers a conformational change that moves a nucleophilic cysteine residue into the active site, resulting in a >10,000-fold increase in PAD activity.[10-12]

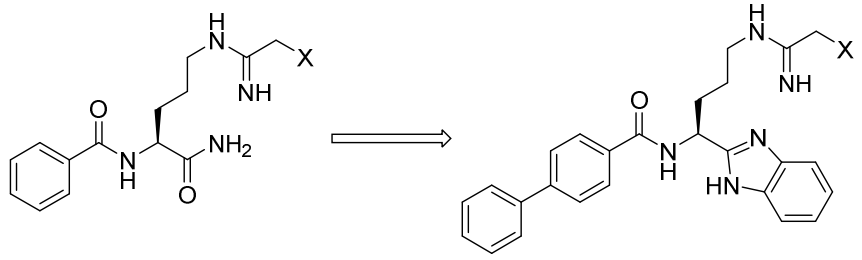


1: peptidyl-arginine **2:** peptidyl-citrulline

R=corresponding amino acids in protein sequence

Figure 1. The PAD catalyzed hydrolysis of peptidyl-arginine (**1**) to generate peptidyl-citrulline (**2**).

PAD2, which is expressed in most tissue and cell types,[8] contributes to the development of certain inflammatory diseases and cancers.[1, 2] Specifically, PAD2 is upregulated in multiple sclerosis (MS) where it citrullinates myelin basic protein (MBP) leading to demyelination.[13, 14] PAD2 is also released into the synovial fluid of arthritic joints where it citrullinates a number of extracellular targets, thereby promoting the development of anti-citrullinated antibodies (ACPA), which is a characteristic feature of RA.[8, 15-18] Additionally, PAD2 is highly expressed in luminal breast cancers and its expression correlates with the levels of the HER2 protooncogene.[19-21] When PAD2 activity is inhibited in MCF10-DCIS cells, a breast cancer cell line, proliferation is decreased with a corresponding increase in apoptosis.[19] This appears to be an epigenetic phenomenon as PAD2 is recruited to ERα promoters where it citrullinates histone H3 at R26 to trigger the localized decondensation of chromatin to facilitate ERα binding to its promoters to drive gene transcription.[19, 20]



3a: X=F; **F-amidine**	**3b:** X=Cl; **Cl-amidine**	**4a:** X=F; **BB-F-amidine**	**4b:** X=Cl; **BB-Cl-amidine**
k_{inact}/K_I (M^{-1} min^{-1})	k_{inact}/K_I (M^{-1} min^{-1})	k_{inact}/K_I (M^{-1} min^{-1})	k_{inact}/K_I (M^{-1} min^{-1})
PAD1=2800	PAD1=37000	PAD1=900	PAD1=16100
PAD2=380	PAD2=1200	PAD2=1200	PAD2=4100
PAD3=170	PAD3=2000	PAD3=3400	PAD3=6800
PAD4=3000	PAD4=13000	PAD4=3750	PAD4=13300

Figure 2. Structures and k_{inact}/K_I values for the 1st-Generation PAD inhibitors F-amidine (**3a**) and Cl-amidine (**3b**) and the 2nd-Generation PAD inhibitors BB-F-amidine (**4a**) and BB-Cl-amidine (**4b**).

Given the strong evidence linking dysregulated PAD2 activity to MS, RA and breast cancer, we and others have focused on developing inhibitors targeting this enzyme, as well as PAD4, whose activity is also upregulated in a variety of inflammatory diseases including RA[2, 22-24] and lupus.[25-27] Notably, our first-generation *pan*-PAD inhibitor Cl-amidine (**3b**, Figure 2) demonstrates efficacy in animal models of RA,[24] lupus,[27] ulcerative colitis,[28] spinal cord injury,[29] breast cancer,[19] and atherosclerosis.[30] Despite these successes, many peptide-based inhibitors similar to Cl-amidine (**3b**) have displayed less than optimal potency *in cellulo*, which has been attributed to their poor metabolic stability and cell membrane permeability. In an effort to address these issues, the C-terminal carboxamide was replaced with a benzimidazole moiety to yield BB-Cl-amidine (**4b**, Figure 2).[27] This second-generation PAD inhibitor is at least 10-fold more potent than Cl-amidine (**3b**) in cells and animal models of lupus and RA, which represents a significant advance over Cl-amidine (**3b**).[22, 23, 27]

Interestingly, the fluoroacetamidine counterparts of Cl-amidine (**3a**, F-amidine) and BB-Cl-amidine (**4a**, BB-F-amidine) have yet to find much utility in animal models as they display a marked decrease in PAD inhibition as compared to their chloroacetamidine analogs. This has been attributed to the less electrophilic nature of the fluoroacetamidine warhead, which reacts with an active site thiol conserved in all PAD isozymes.[31, 32] It is also important to note that while BB-Cl-amidine (**4b**) brought a significant advance to PAD inhibitor development, as a pan PAD inhibitor,[27] it does not provide utility in investigating the roles of specific isozymes.

Prior to the development of BB-Cl-amidine (**4b**) advances had been made in improving isozyme-specific inhibition. Specifically, TDFA provided PAD4-selective inhibition[33], D-Cl-amidine provided PAD1-selective inhibition[34], while Cl4-amidine[10, 31] and a series of hydantoin-

based inhibitors[35, 36] provided PAD3-selective inhibition. While these inhibitors have brought about significant advances in PAD inhibitor development, two main issues remain: 1) most of these inhibitors exhibit poor efficacy in cell-based assays (potentially due to metabolic instability and poor membrane permeability) and 2) a PAD2-selective inhibitor has yet to be developed. The deficiencies of current PAD inhibitors underscore a pressing need for PAD2-selective inhibitors with enhanced cellular efficacy. Herein, we report the development of benzimidazole-based PAD2-selective inhibitors that inhibit cellular PAD2 activity.

Results and Discussion.

Inhibitor Design

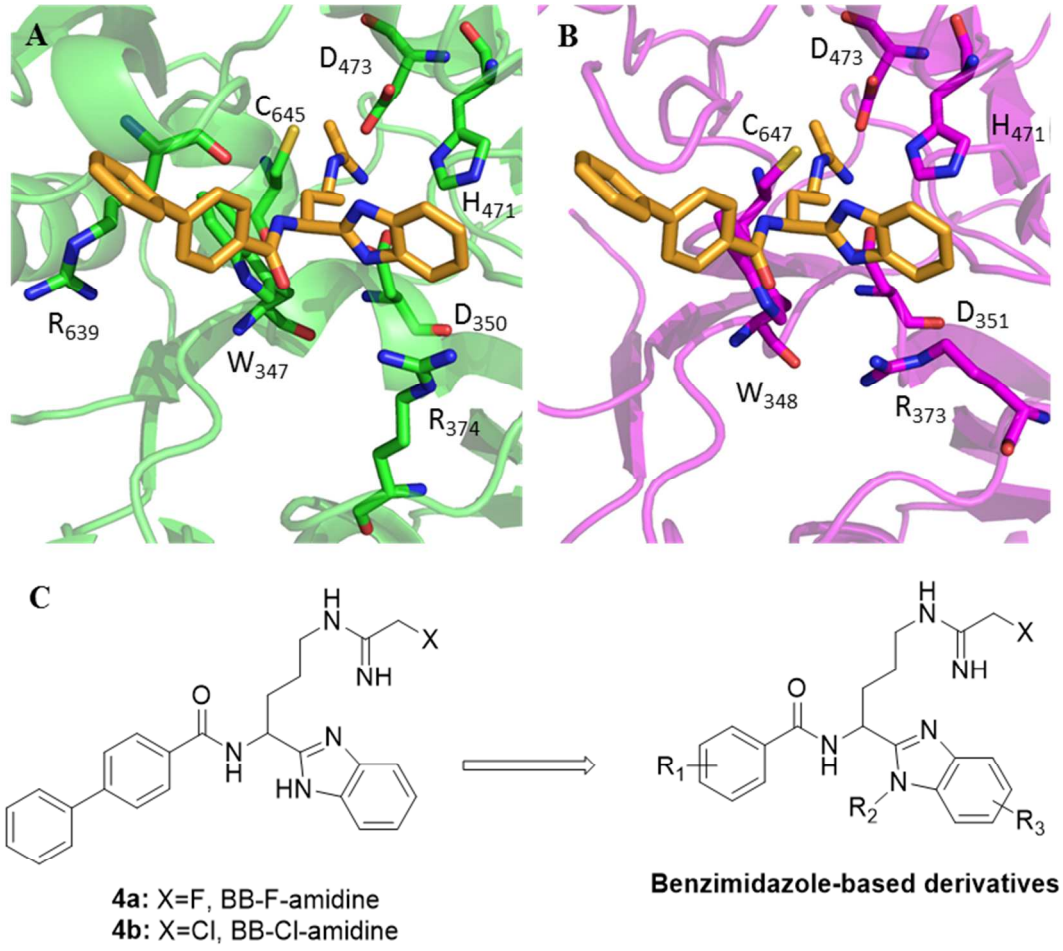


Figure 3. (**A**) Co-crystal structure of BB-F-amidine (**4a**) bound to PAD4 (PDB ID: 5N0M). (**B**) PAD2 (PDB: 4N2C)[12] was superimposed on the PAD4·BB-F-amidine (**4a**) co-crystal structure and then PAD4 was removed leaving BB-F-amidine. (**C**) Evolution of 2nd-Generation PAD inhibitors BB-F-amidine (**4a**) and BB-Cl-amidine (**4b**) to 3rd-Generation benzimidazole-based PAD inhibitors.

Given the favorable cellular efficacy observed with BB-Cl-amidine (**4b**), we hypothesized that further elaboration of this scaffold would identify compounds with enhanced potency, selectivity and cellular efficacy. To aid these efforts we first generated the fluoro-containing analog, BB-F-amidine (Figure 2, **4a**). We then determined the crystal structure of BB-F-amidine (**4a**) bound to PAD4 (Figure 3A). Similar to crystal structures of F-amidine,[31] TDFA[33], and *o*-F-amidine[37] bound to PAD4, the amidine forms an intricate hydrogen bonding network with D350, H471, and D473. Interestingly, while R374 forms a hydrogen bond with the C-terminal carboxamide of the 1st-Generation PAD inhibitors, it instead forms a cation-π interaction with the benzimidazole. Overlays of the PAD2 structure[12] onto the BB-F-amidine•PAD4 co-crystal structure (Figure 3B) shows a similar network of favorable interactions with D473, D351, and H471. These interactions properly orient the warhead for covalent modification by C647. By contrast to the PAD4 structure, R373, which corresponds to R374 in PAD4, orients away from the benzimidazole moiety such that it does not form the favorable cation-π interaction observed with PAD4.

Ortho-carboxylate Series

Using these structures as a guide, we hypothesized that further elaboration of the benzimidazole scaffold could lead to compounds with enhanced potency. Given our previous demonstration that the installation of an ortho-carboxylate (*o*-F-amidine and *o*-Cl-amidine)[37] enhanced potency by ≥10-fold, due to the formation of a key hydrogen bond with W348 in PAD2 or W347 in PAD4,[37] we sought to combine the benzimidazole scaffold with the ortho-carboxylate functionality, expecting that the ortho-carboxylate would enhance *in vitro* potency and the benzimidazole would promote cellular uptake. This approach had previously been

highly successful in developing highly potent tetrazole-based inhibitors.[38] Combining these two features (compounds **9a** and **9b**, Figure 4) increased PAD1 (18- and 1.8-fold, respectively) and PAD4 (6.5- and 2.3-fold, respectively) inhibition as compared to BB-F-amidine (**4a**) and BB-Cl-amidine (**4b**), however, PAD2 potency was relatively unaffected.

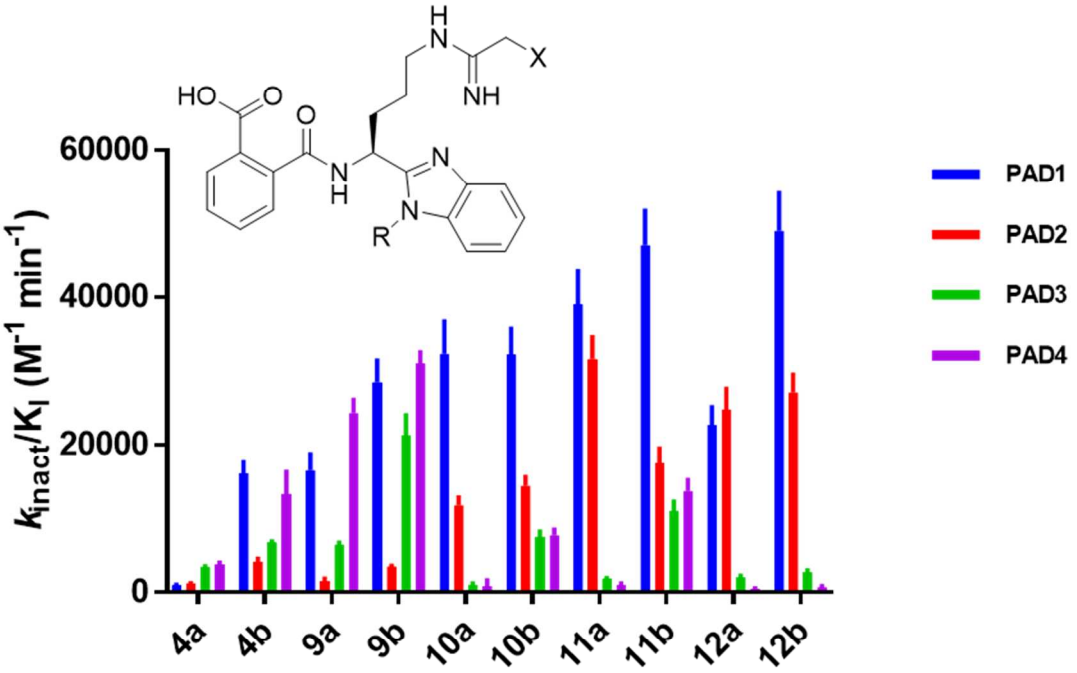
N-alkyl Series

We next explored other aspects of the benzimidazole scaffold to improve PAD2 potency and selectivity. Focusing on the nitrogen atoms of the benzimidazole scaffold, we noted that when BB-F-amidine (**4a**) is bound to PAD4, each nitrogen atom is involved in a water-mediated hydrogen bonding network; similar interactions are also observed with F-amidine (**3a**). While no crystal structure of BB-F-amidine (**4a**) bound to PAD1 exists, these water molecules are conserved in its structure. However, this hydrogen bonding network is absent in PAD2 and is replaced with a unique hydrophobic binding region created by V469, L467, L642, and F573 (Figure S1). This insight was intriguing and we hypothesized that *N*-alkylation of the benzimidazole could target this binding region for potent and selective PAD2 inhibition. *N*-alkylation would also increase the overall hydrophobicity of the inhibitor, which we expected to enhance cell permeability.

As discussed above, the ortho-carboxylic acid containing series (compounds **9a** and **9b**) improved PAD1 and PAD4 inhibition as compared to BB-Cl-amidine (**4b**) and BB-F-amidine (**4a**), but surprisingly did not improve PAD2 inhibition. By contrast, *N*-alkylation of the benzimidazole exhibited a marked increase in PAD2 inhibition and selectivity (Figure 4). Specifically, the *N*-ethyl derivative (**11a**) showed a 26-fold increase in PAD2 inhibition as compared to BB-F-amidine (**4a**) while also exhibiting 31-fold selectivity for PAD2 over PAD4

and 18-fold selectivity for PAD2 over PAD3. Also noteworthy is the fact that the fluoroacetamidine warhead (**9a-12a**) demonstrated similar or better PAD2 inhibition than its chloroacetamidine counterpart (**9b-12b**).

A



B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
9a (R=H, X=F)	11	1.0	4.3	16
9b (R=H, X=Cl)	8.3	1.0	6.2	9.0
10a (R=Me, X=F)	40	15	1.3	1.0
10b (R=Me, X=Cl)	4.3	1.9	1.0	1.0
11a (R=Et, X=F)	39	31	1.7	1.0
11b (R=Et, X=Cl)	4.3	1.6	1.0	1.2
12a (R=^{i}Pr, X=F)	49	54	4.4	1.0
12b (R_1^{i}Pr, X=Cl)	68	38	3.8	1.0

Figure 4. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 1 compounds **9-12**. (**B**) Summary of isozyme selectivity for series 1 compounds **9-12** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

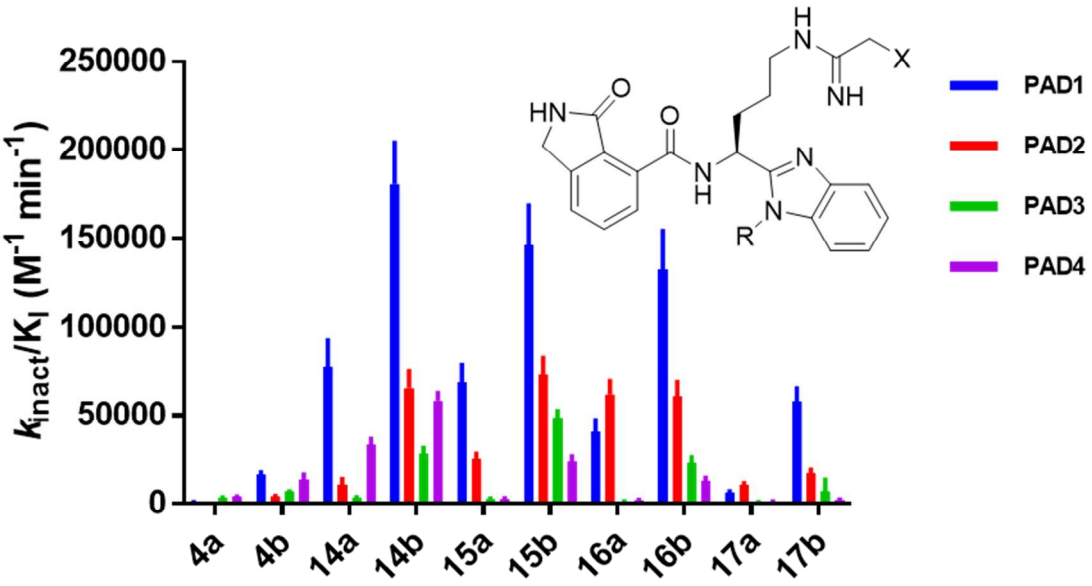
While this improvement in PAD2 potency and selectivity is not completely understood, the co-crystal structure of BB-F-amidine (**4a**) with PAD4 and the subsequent PAD2 overlay do provide some insight (Figure 3). These structures suggest that there is a key water-mediated hydrogen bond that exists for PAD4 and not PAD2 with the nitrogen atom of the benzimidazole moiety. Introduction of an *N*-alkyl group may disrupt this hydrogen bond causing a decrease in PAD4 inhibition. The PAD2 overlay also shows a hydrophobic binding region unique to PAD2 which the *N*-alkyl group could be oriented toward. This gained hydrophobic interaction could explain the improved PAD2 inhibition of compounds **10-12** with the *N*-ethyl substitution (**11a**) proving to be optimal for this binding region. The fluoroacetamidine warhead demonstrates similar (and in some instances improved) PAD2 inhibition as compared to the chloroacetamidine warhead potentially due to a smaller binding region near C647, as fluorine is smaller than chlorine. The differences in PAD2 selectivity, may also be attributed to the less electrophilic nature of the fluoroacetamidine warhead as compared to the chloroacetamidine warhead leading to a higher degree of specificity for the intended C647 target.

Lactam Series

Hypothesizing that the hydrophilic nature of the ortho-carboxylate might limit cellular uptake, we next replaced this moiety with a lactam, yielding series 2 compounds (**14-17**). We further hypothesized that the enhanced conformational constraint of this amide would further improve PAD2 inhibition. Rewardingly, introduction of the lactam moiety resulted in a ~7-fold

increase in PAD2 inhibition for **14a** as compared to both **9a** and ~9-fold compared to BB-F-amidine (**4a**). By contrast, there was only a minor increase in PAD4 inhibition observed for **14a** as compared to **9a** (~1.4-fold). Similar to the series 1 compounds, *N*-alkylation of the benzimidazole caused a marked increase in PAD2 inhibition and selectivity. Specifically, the *N*-ethyl benzimidazole (**16a**) displayed the most potent PAD2 inhibition (k_{inact}/K_I = 61,600 M^{-1} min^{-1}) utilizing the fluoroacetamidine warhead, representing a ~41-fold increase in potency compared to **9a** and ~51-fold increase compared to BB-F-amidine (**4a**). Furthermore, **16a** has the greatest degree of PAD2 selectivity (56-fold over PAD3 and ~33-fold over PAD4). The fluoroacetamidine warhead again proved superior with an optimal combination of PAD2 potency and selectivity.

A



B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
14a (R=H, X=F)	24	3.3	1.0	10
14b (R=H, X=Cl)	6.5	2.3	1.0	2.1
15a (R=Me, X=F)	25	8.9	1.0	1.0
15b (R=Me, X=Cl)	61	3.1	2.0	1.0
16a (R=Et, X=F)	37	56	1.0	1.7
16b (R=Et, X=Cl)	10	4.8	1.8	1.0
17a (R=^{i}Pr, X=F)	12	20	1.0	2.4
17b (R=^{i}Pr, X=Cl)	27	8.0	3.2	1.0

Figure 5. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 2 compounds **14-17**. (**B**) Summary of isozyme selectivity for series 2 compounds **14-17** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

A co-crystal structure of **14a** bound to PAD4 was then determined (Figure 6A); we used PAD4 because efforts to co-crystallize PAD2 with **14a** as well as other inhibitors proved unsuccessful. The PAD4•**14a** complex is highly similar to the PAD4•*o*-F-amidine structure (Figure 6B). Notably, these structures show that the carbonyl of either the *ortho*-carboxylic acid or the lactam form favorable hydrogen bonds with the indole nitrogen of W347. This interaction is likely a key determinant of the enhanced potency observed with this series. A potential hydrogen bond (either direct or water mediated) between R639 and the *ortho*-carboxylate or the nitrogen of the lactam may also contribute to the observed potency. Overlays of these PAD4 structures with PAD1[39] (Figures 6C-D) suggest similar hydrogen bonding networks with W349 and R348 also led to the marked increase in PAD1 inhibition observed for these compounds. However, modeling studies with PAD2[12] suggest that the increased potency for PAD2 is due solely to a hydrogen bond with W348 as an arginine residue corresponding to R639 in PAD4 is far removed from this binding region. The improved PAD2 inhibition for series 2 compounds (**14-17**) as compared to series 1 compounds (**9-12**) may be attributed to the replacement of the carboxylic acid with an amide functionality, the introduction of conformational constraint, or a synergistic effect of the two changes. Efforts to fully understand this are underway through the pursuit of a PAD2 co-crystal structure.

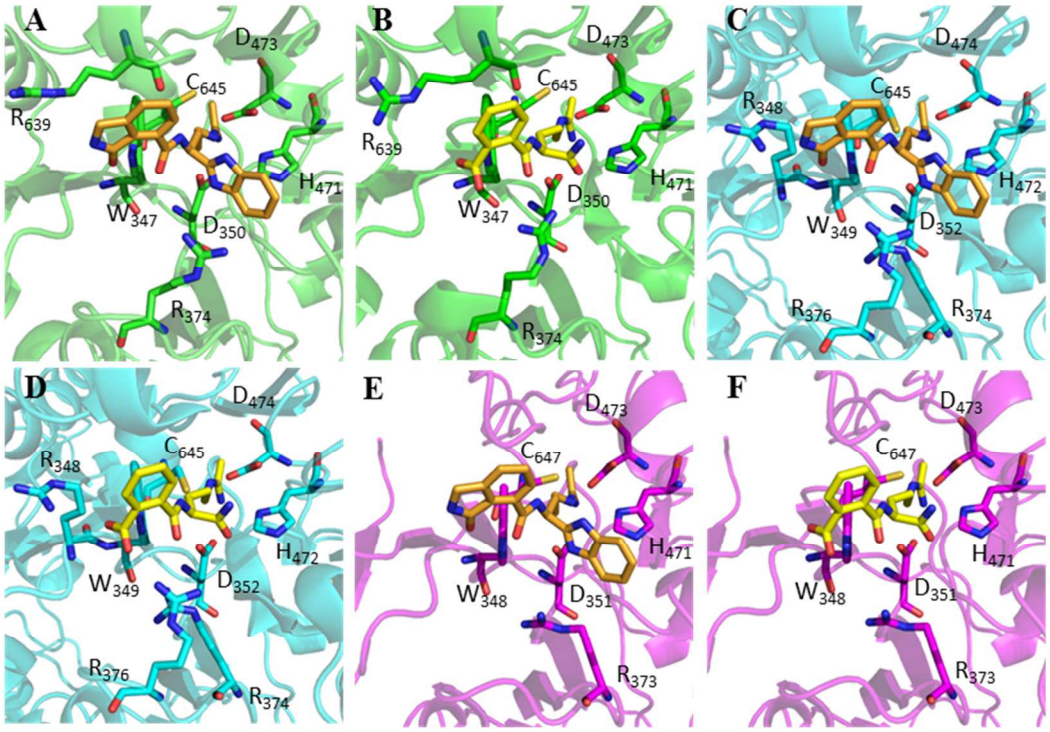


Figure 6. (**A**) Co-crystal structure of **AFM-14a** bound to PAD4 (PDB: 5N1B). (**B**) Co-crystal structure of *o*-F-amidine bound to PAD4[37] (PDB: 3B1T). (**C**) PAD1[39] (PDB: 5HP5) overlay with PAD4•**AFM-14a** co-crystal structure in PyMol. (**D**) PAD1[39] (PDB: 5HP5) overlay with PAD4•*o*-F-amidine co-crystal structure in PyMol. (**E**) PAD2[12] (PDB: 4N2C) overlay with PAD4•**AFM-14a** co-crystal structure in PyMol. (**F**) PAD2[12] (PDB: 4N2C) overlay with PAD4•o-F-amidine co-crystal structure in PyMol.

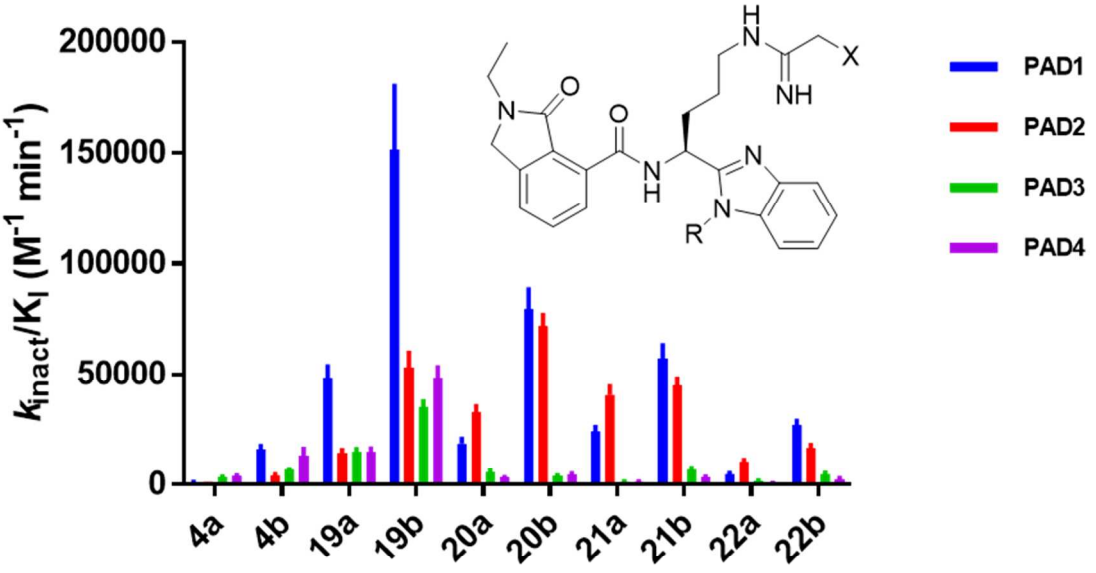
N-alkyl Lactams

To further optimize this scaffold, we next focused on *N*-alkylation of the lactam moiety (Figure 7) because the PAD4•**14a** co-crystal, as well as our PAD1 modeling studies, suggested that modification of this position would disrupt the hydrogen bonding network with R639 and with R348 in PAD1,[39] neither of which appear to aid in PAD2[12] binding. Furthermore, this

substitution would further increase the overall hydrophobicity (**19a** ClogP = 1.61 and **14a** ClogP = 1.00), which we expected to further improve the cell permeability.

Interestingly, this series of compounds (**19-22**) demonstrated similar PAD2 potencies and PAD2-selectivity as compared to series 2 compounds (**14-17**). Specifically, compounds **21a** and **16a** exhibited identical PAD2 vs PAD4 selectivity (33-fold), PAD2 vs PAD1 selectivity (~2-fold) and similar PAD2 vs PAD3 selectivity (>33-fold). While the *N*-ethyl lactam does not influence PAD2 potency or selectivity, it does improve the overall hydrophobicity of the compounds.

A



B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
19a (R=H, X=F)	3.4	1.0	1.0	1.0
19b (R=H, X=Cl)	4.3	1.5	1.0	1.4
20a (R=Me, X=F)	6.1	11	1.9	1.0
20b (R=Me, X=Cl)	21	19	1.0	1.2
21a (R=Et, X=F)	19	33	1.0	1.0
21b (R=Et, X=Cl)	18	14	2.1	1.0
22a (R=^{i}Pr, X=F)	10	22	3.4	1.0
22b (R=^{i}Pr, X=Cl)	11	6.8	2.0	1.0

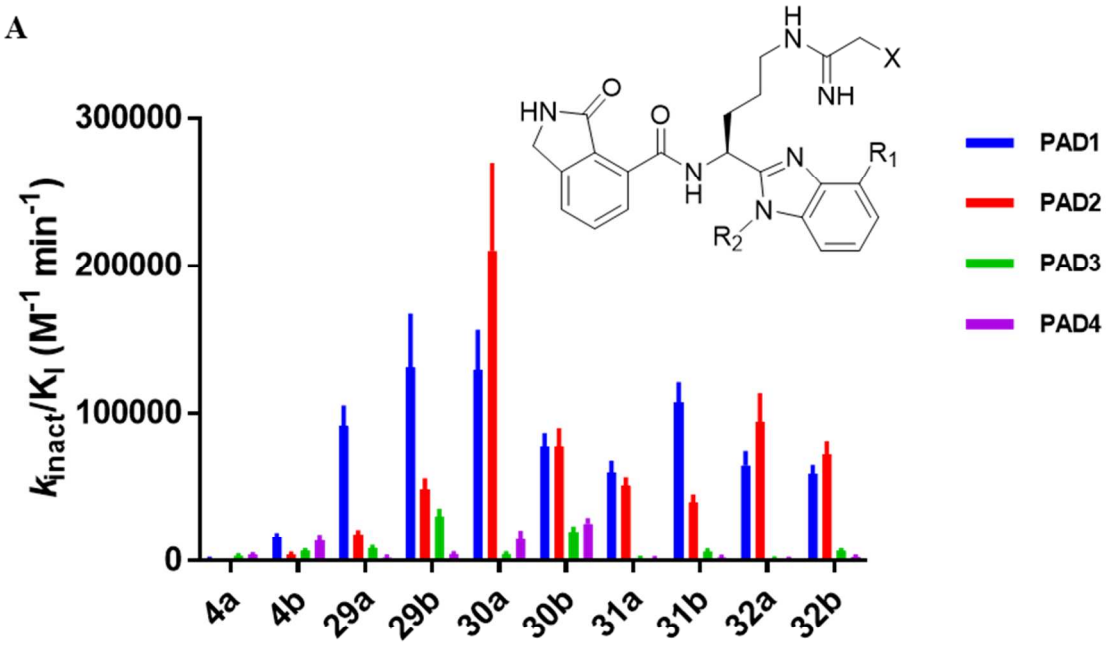
Figure 7. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 3 compounds **19-22**. (**B**) Summary of isozyme selectivity for series 3 compounds **19-22** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Alkoxy-substituted Benzimidazoles

Having demonstrated that the lactam enhanced potency and the *N*-ethyl benzimidazole modification enhanced PAD2 selectivity, we next explored substitutions of the 6-membered ring of the benzimidazole. Specifically, alkoxy substitutions were introduced to probe potential interactions with D351, R373 and H471 in PAD2, all of which surround the 6-membered ring of the benzimidazole.

This investigation began with a methoxy substitution at the 4-position giving series 4 compounds **29a** and **29b** which both showed improvements in PAD2 inhibition and selectivity over PAD4 as compared to their parent compounds **14a** and **14b**. This improvement was then coupled with *N*-alkylation of the benzimidazole to give series 4 compounds **30-31**. Rewardingly, compound **30a** (*N*-methyl benzimidazole) showed a dramatic increase in PAD2 inhibition (k_{inact}/K_I = 210,300 M^{-1} min^{-1}) while also displaying very good selectivity for PAD2 (1.6-fold over PAD1, 47-fold over PAD3 and ~15-fold over PAD4). Interestingly, the *N*-ethyl benzimidazole (compound **31a**) displayed a 4-fold decrease in PAD2 potency (compared to **30a**) while improving PAD2 selectivity over PAD4 to 32-fold. This result suggests that introduction of the 4-methoxy group shifts the position of the benzimidazole so that the smaller *N*-methyl group would better access the unique hydrophobic binding region of PAD2 than the larger *N*-ethyl group. This is in contrast to series 1, 2, and 3 compounds which showed that the *N*-ethyl group was optimal at this position. Utilization of a 4-ethoxy group (**32a** and **32b**) also showed excellent PAD2 inhibition, albeit ~2-fold less than **30a**. However, **32a** did demonstrate superior selectivity for PAD2 over PAD4 (95-fold) and PAD2 over PAD3 (79-fold).

A



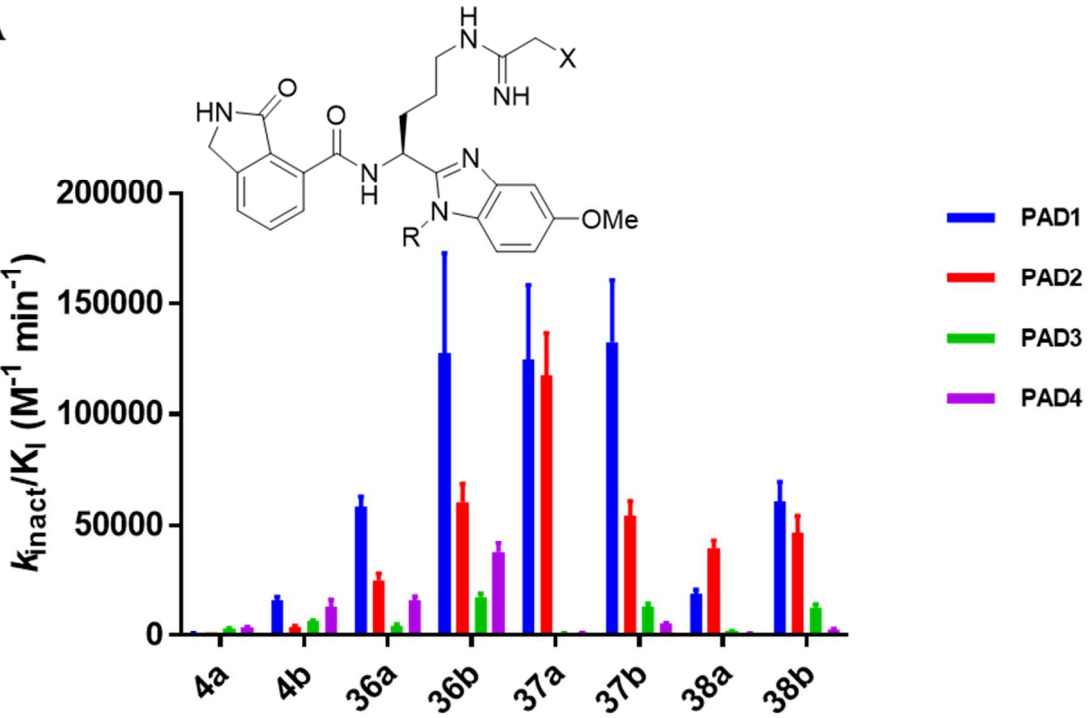
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
29a (R_1=OMe, R_2=H, X=F)	42	8.0	3.8	1.0
29b (R_1=OMe, R_2=H, X=Cl)	30	11	6.9	1.0
30a (R_1=OMe, R_2=Me, X=F)	29	47	1.0	3.2
30b (R_1=OMe, R_2=Me, X=Cl)	3.9	3.9	1.0	1.3
31a (R_1=OMe, R_2=Et, X=F)	38	32	1.0	1.0
31b (R_1=OMe, R_2=Et, X=Cl)	50	19	2.9	1.0
32a (R_1=OEt, R_2=Me, X=F)	65	95	1.2	1.0
32b (R_1=OEt, R_2=Me, X=Cl)	24	30	2.7	1.0

Figure 8. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 4 compounds **29-32**. (**B**) Summary of isozyme selectivity for series 4 compounds **29-32** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Further SAR of the benzimidazole moiety then explored a 5-methoxy substitution. As seen in Figure 9, compounds **36a** and **36b** showed an increase in PAD2 inhibition compared to series 2 compounds **14a** and **14b**. Similar to series 4 compounds, *N*-methyl alkylation of the benzimidazole resulted in a significant increase in PAD2 inhibition (albeit ~2-fold lower than **30a**) as well as excellent PAD2 selectivity over PAD4 (95-fold) and PAD3 (114-fold). The *N*-methyl substitution again displayed a greater degree of PAD2 inhibition and PAD2 selectivity than the *N*-ethyl substitution (**38a** and **38b**).

A



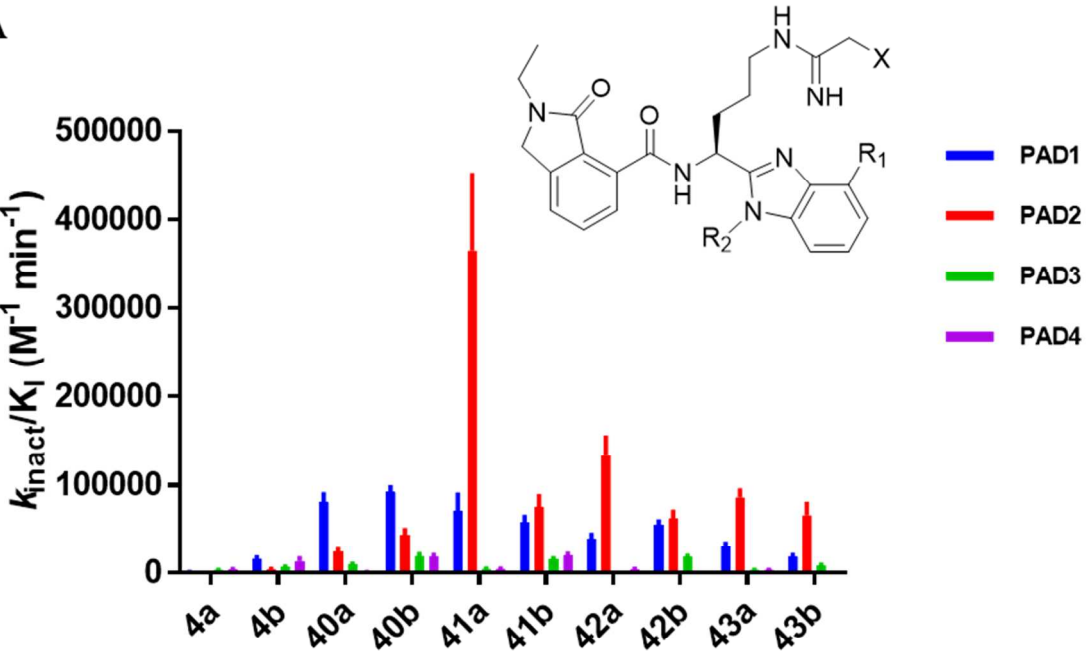
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
36a (R=H, X=F)	13	5.4	1.0	3.5
36b (R=H, X=Cl)	7.3	3.5	1.0	2.2
37a (R=Me, X=F)	121	114	1.0	1.2
37b (R=Me, X=Cl)	25	10	2.4	1.0
38a (R=Et, X=F)	24	50	2.5	1.0
38b (R=Et, X=Cl)	21	16	4.4	1.0

Figure 9. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 5 compounds **36-38**. (**B**) Summary of isozyme selectivity for series 5 compounds **36-38** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

In order to improve the hydrophobic nature of these compounds, we then reintroduced the *N*-ethyl lactam moiety in conjunction with the 4-alkoxy substitutions on the benzimidazole (**30a** ClogP=1.26 and **41a**=1.62) to give series 6 compounds. Similar to **29a**, compound **40a** showed enhanced PAD2 potency and selectivity as compared to **14a** (Figure 10). Series 6 compounds also demonstrated that the *N*-methyl benzimidazole substitution (**41a**) was optimal for PAD2 inhibition resulting in the most potent PAD2 inhibitor to date. Notably, the k_{inact}/K_I for this compound, which is 365,400 M^{-1} min^{-1}, represents a >300-fold improvement over F-amidine (**3a**) and BB-F-amidine (**4a**). Compound **41a** also demonstrated excellent selectivity for PAD2 over PAD4 (85-fold) and PAD3 (85-fold), while improving selectivity for PAD2 over PAD1 (~5-fold). The 4-methoxy benzimidazole substitution also proved superior over the 4-ethoxy and 5-methoxy substitutions utilized by series 6 compounds **43a** and **43b** (Figure 10) and series 7 compounds **45-47**.

A

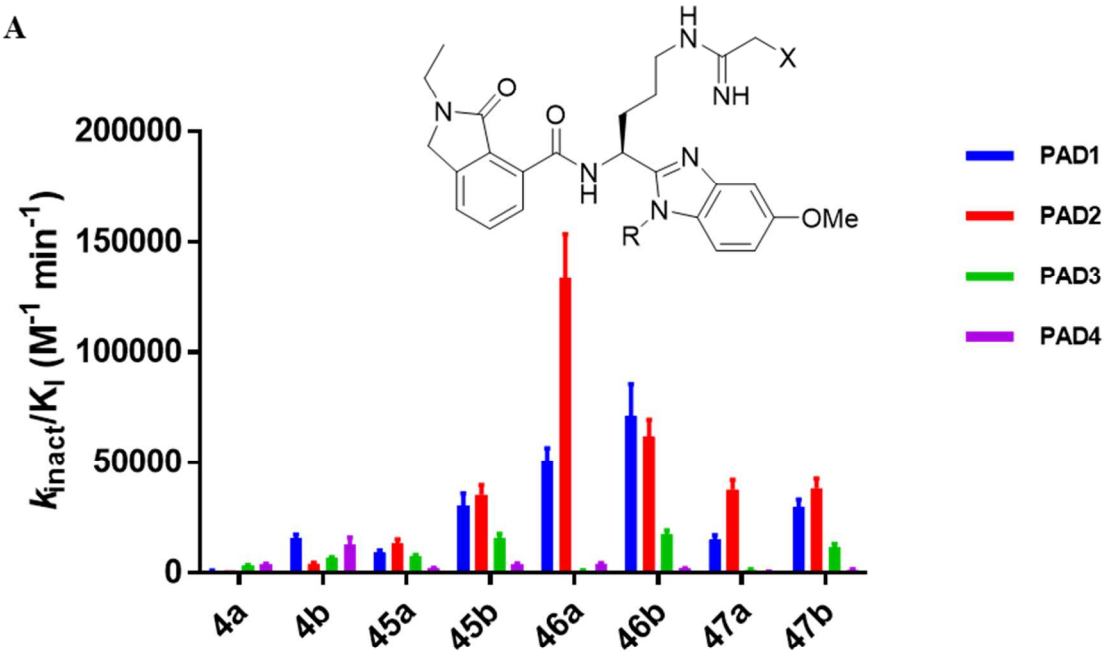


B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
40a (R_1=OMe, R_2=H, X=F)	107	32	12	1.0
40b (R_1=OMe, R_2=H, X=Cl)	5.0	2.3	1.0	1.0
41a (R_1=OMe, R_2=Me, X=F)	16	85	1.0	1.0
41b (R_1=OMe, R_2=Me, X=Cl)	3.9	5.0	1.0	1.3
42a (R_1=OMe, R_2=Et, X=F)	28	40	2.0	1.0
42b (R_1=OMe, R_2=Et, X=Cl)	31	27	3.4	1.0
43a (R_1=OEt, R_2=Me, X=F)	8.9	26	1.0	1.0
43b (R_1=OEt, R_2=Me, X=Cl)	12	43	5.5	1.0

Figure 10. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 6 compounds **40-43**. (**B**) Summary of isozyme selectivity for series 6 compounds **40-43** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

A



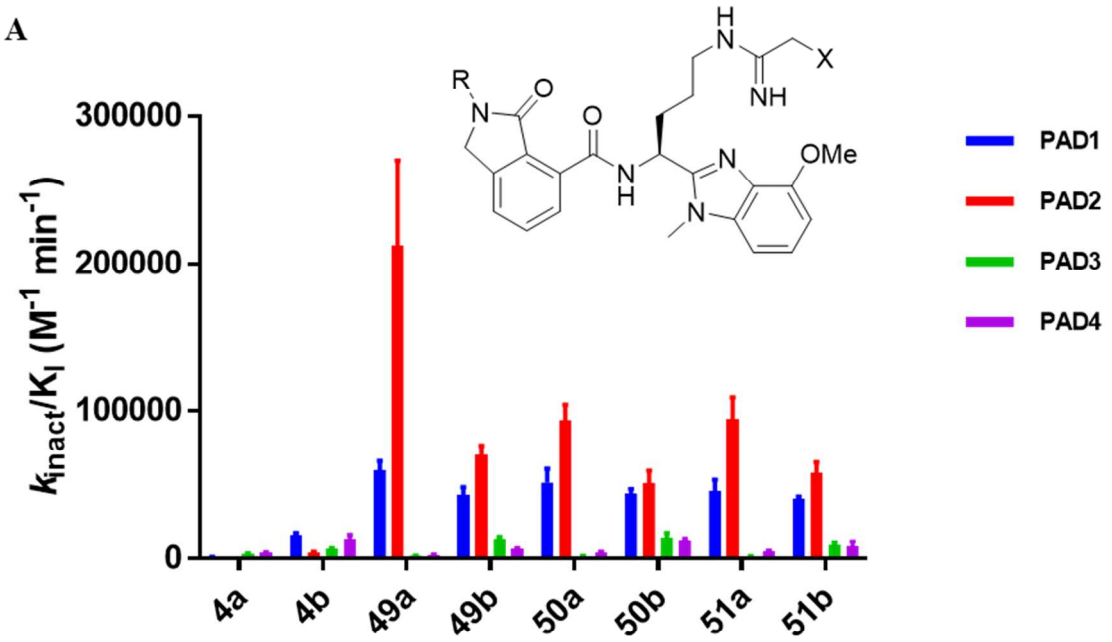
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
45a (R=H, X=F)	4.4	6.3	3.6	1.0
45b (R=H, X=Cl)	7.8	9.1	4.1	1.0
46a (R=Me, X=F)	47	125	1.0	3.8
46b (R=Me, X=Cl)	37	32	9.3	1.0
47a (R=Et, X=F)	28	70	2.7	1.0
47b (R=Et, X=Cl)	21	27	8.4	1.0

Figure 11. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 7 compounds **45-47**. (**B**) Summary of isozyme selectivity for series 1 compounds **45-47** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Series 6 compound **41a** demonstrated that a combination of a 4-methoxy *N*-methyl benzimidazole with an *N*-ethyl lactam presented an ideal combination for potent and selective PAD2-inhibition. However, further investigation of lactam *N*-alkylation was needed to fully understand this phenomenon as well as finding a good balance between optimal potency and hydrophobicity. Therefore, series 8 compounds **49-51** were synthesized to explore *N*-alkylation of the lactam moiety. These data show that an *N*-methyl substitution of **49a** was similar in PAD2 potency to the unsubstituted lactam (**30a**) while also exhibiting similar PAD2-selectivity to the *N*-ethyl lactam, **41a**. *N*-isopropyl and *N*-cyclopropyl lactams (**50a** and **51a**, respectively) also proved to be highly potent and selective PAD2 inhibitors, although neither were as potent and selective as **30a**, **41a** or **49a**. Despite this, **50a** and **51a** are more hydrophobic than other lactam *N*-alkyl substitutions.

A



B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
49a (R=Me, X=F)	30	106	1.0	1.1
49b (R=Me, X=Cl)	6.8	11	2.2	1.0
50a (R=^{i}Pr, X=F)	32	58	1.0	2.6
50b (R=^{i}Pr, X=Cl)	3.5	4.1	1.2	1.0
51a (R=cyclopropyl, X=F)	38	77	1.0	3.8
51b (R=cyclopropyl, X=Cl)	4.7	6.7	1.1	1.0

Figure 12. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 8 compounds **45-47**. (**B**) Summary of isozyme selectivity for series 1 compounds **45-47** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

*PAD1, PAD2 and PAD4 binding of **30a** and **41a***

Co-crystal structures of **30a** and **41a** bound to PAD4 were determined to shed light on how these compounds bind PAD4 (Figure 13). Overlays of these structures with those determined for PADs 1 and 2 further highlight common and unique interactions. For example, these compounds interact with PAD1 (D474, H472, W349, D352 and C645), PAD2 (D473, H471, D351, W348 and C647) and PAD4 (D473, H471, W347, D350 and C645) in a similar fashion to **14a**. Close inspection of these structures reveal three main reasons for the PAD2-selectivity versus both PAD1 and PAD4. First, the *N*-methyl benzimidazole appears to orient into a unique hydrophobic binding region of PAD2, while simultaneously eliminating a water mediated hydrogen bonding network that is important for PAD4 binding. Second, the 4-methoxy group of the benzimidazole appears to be ideal for this binding region in PAD2 and tolerated for PAD1. However, the PAD4 cation-π with R374 is no longer optimal as the methoxy group forces the benzimidazole moiety to move and it is not able to properly interact with R374. Third, the *N*-ethyl lactam appears to cause R639 in PAD4 to orient away from the lactam functionality. Since the corresponding residue in PAD2 is not an arginine, this likely explains the higher selectivity for PAD2. Similarly, the enhanced PAD2-selectivity over PAD1 is likely explained by the fact that PAD1 does place an arginine residue close to this position. Despite these findings, selective inhibition of PAD2 over PAD1 proved difficult and will be a focus for future inhibitor development.

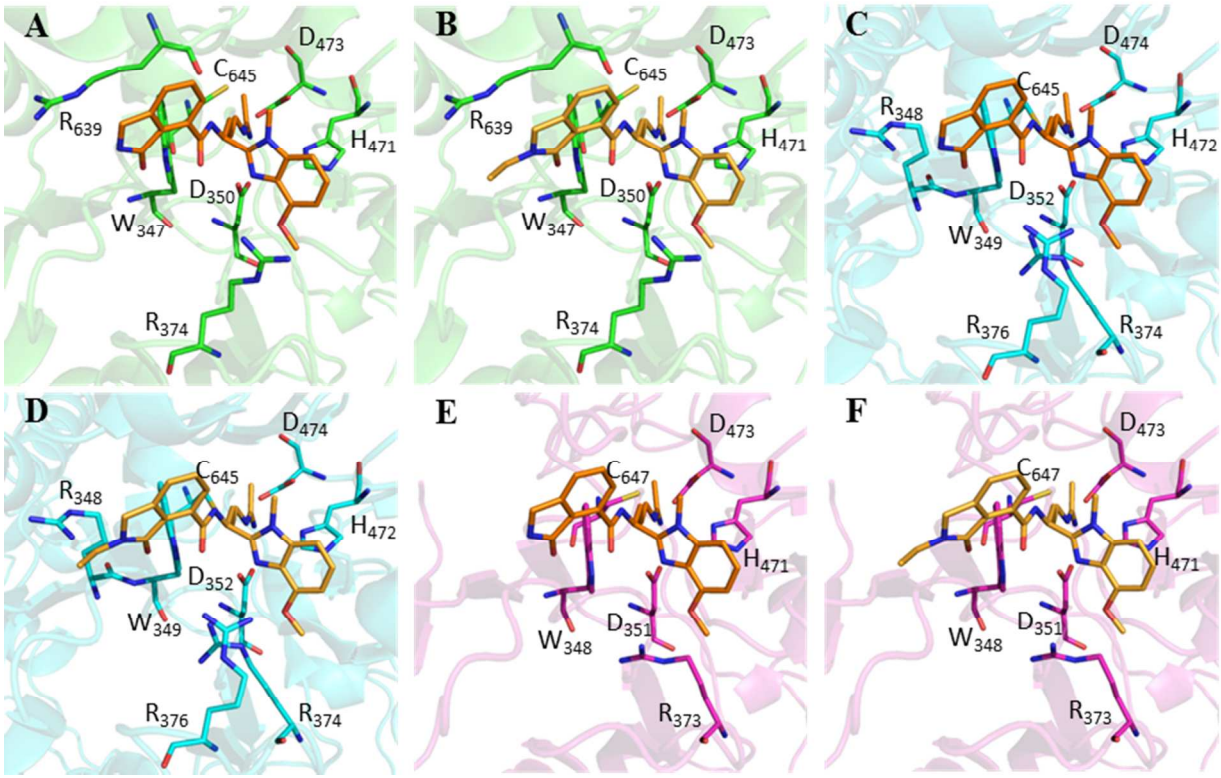


Figure 13. (**A**) Co-crystal structure of **AFM-30a** bound to PAD4 (PDB: 5N0Y). (**B**) Co-crystal structure of **AFM-41a** bound to PAD4 (PDB: 5N0Z). (**C**) PAD1[39] (PDB: 5HP5) overlay with PAD4•**AFM-30a** co-crystal structure in PyMol. (**D**) PAD1[39] (PDB: 5HP5) overlay with PAD4•**AFM-41a** co-crystal structure in PyMol. (**E**) PAD2[12] (PDB: 4N2C) overlay with PAD4•**AFM-30a** co-crystal structure in PyMol. (**F**) PAD2[12] (PDB: 4N2C) overlay with PAD4•**AFM-41a** co-crystal structure in PyMol.

Target Engagement Assays

To measure target engagement, we used a stably transfected HEK293T cell line that overexpresses PAD2 (HEK293T/PAD2). Briefly, cells were incubated with an inhibitor of interest for 15 min, ionomycin (a calcium ionophore) for an additional 15 min, and calcium chloride was then added and the cells incubated for 1 h. Cells were then harvested, lysed, and

cell lysates incubated with RIBFA, **55**, a fluorescently tagged PAD probe (see Supporting Information) and calcium chloride for 1 h. Proteins were then resolved by SDS-PAGE and fluorescently tagged proteins imaged. For proof of concept, BB-Cl-amidine (**4b**) was first evaluated in this assay due to its high cell permeability and efficacy in other cell-based assays. As seen in Figure 14, BB-Cl-amidine (**4b**) was effective at entering the cell and covalently modifying PAD2, giving an EC_{50} of 7.5 μM in this target engagement assay. Figure 14 also shows that BB-F-amidine (**4a**) was ~3-fold less potent in the target engagement assay than BB-Cl-amidine (**4b**), which was expected as it was less potent *in vitro* as well as less hydrophobic.

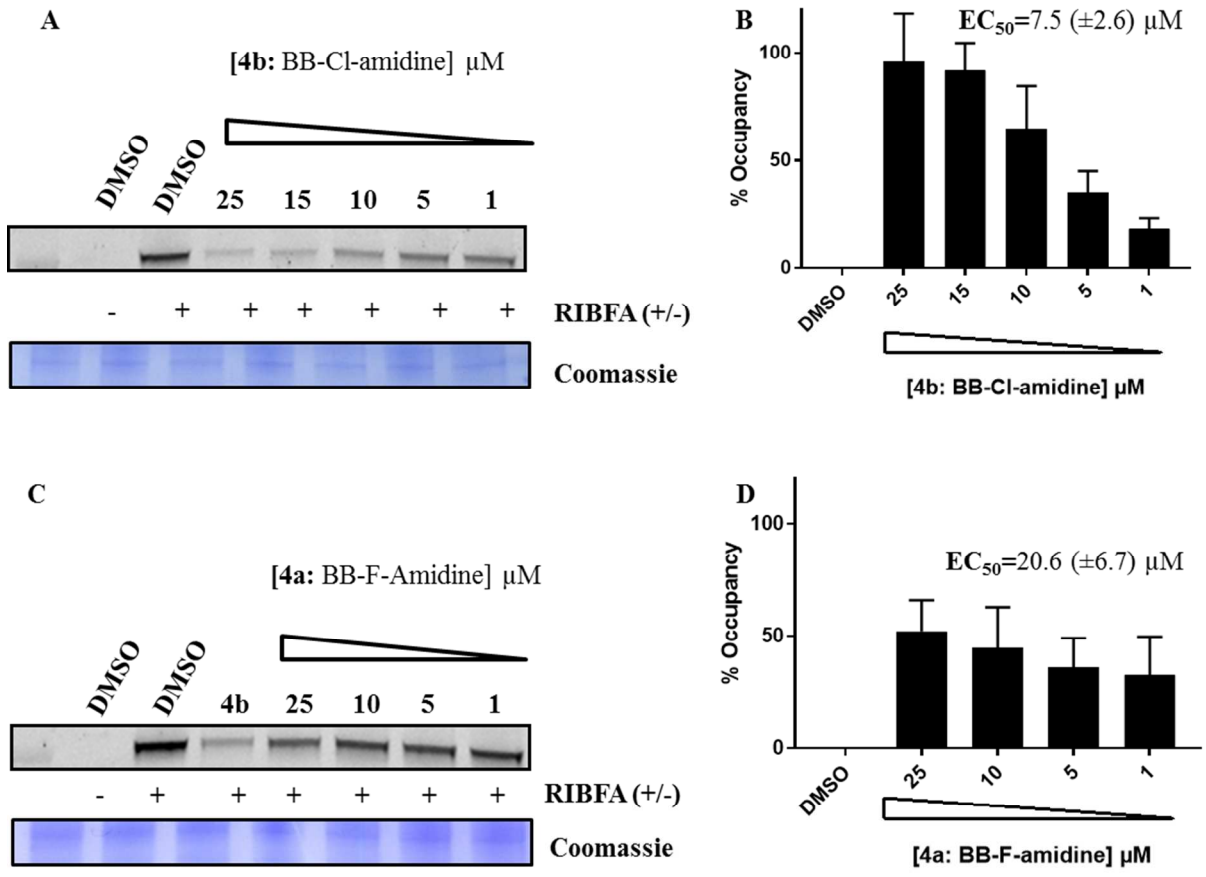


Figure 14. (**A**) Target engagement assay in HEK293/PAD2 cells for BB-Cl-amidine (**4b**). (**B**) Quantified results of target engagement assay plotted as % occupancy of PAD2 by BB-Cl-

amidine (**4b**). (**C**) Target engagement assay in HEK293/PAD2 cells for BB-F-amidine (**4a**). (**D**) Quantified results of target engagement assay plotted as % occupancy of PAD2 by BB-F-amidine (**4a**). Note: Panel C, BB-Cl-amidine (**4b**) was evaluated at 25 μM.

This target engagement assay was then used a primary screen to evaluate the cellular efficacy of compounds that showed a k_{inact}/K_I for PAD2 of ≥50,000 M^{-1} min^{-1}. Specifically, the most potent PAD2 inhibitors (**16a**, **30a-32a**, **37a**, **41a**, and **49a-51a**) were first tested at 25 μM for their ability to enter HEK293T/PAD2 cells and covalent modify PAD2 (Figure S2). The unsubstituted lactams **30a**, **31a**, **32a**, and **37a** exhibited good potency in this assay with >60% occupancy at 25 μM, while **16a** was roughly half as potent. Interestingly, *N*-alkyl lactams **41a**, **49a**, **50a**, and **51a** were all ~2-fold less potent than their unsubstituted counterparts (<37% occupancy). This was unexpected as the *N*-alkyl lactams displayed a similar or greater ability to inhibit PAD2 *in vitro* while also possessing a higher ClogP, or higher degree of hydrophobicity. These attributes were expected to enable better cell penetration and inactivation of PAD2. Mass spectrometry analysis of compound entering the cell showed that ~6-fold less of **41a** was found internalized by the cell than unsubstituted lactams **30a**, **31a**, **32a**, and **37a** (see Supporting Information, Table S1). The exact reason for this difference is not completely understood but is an active area of investigation.

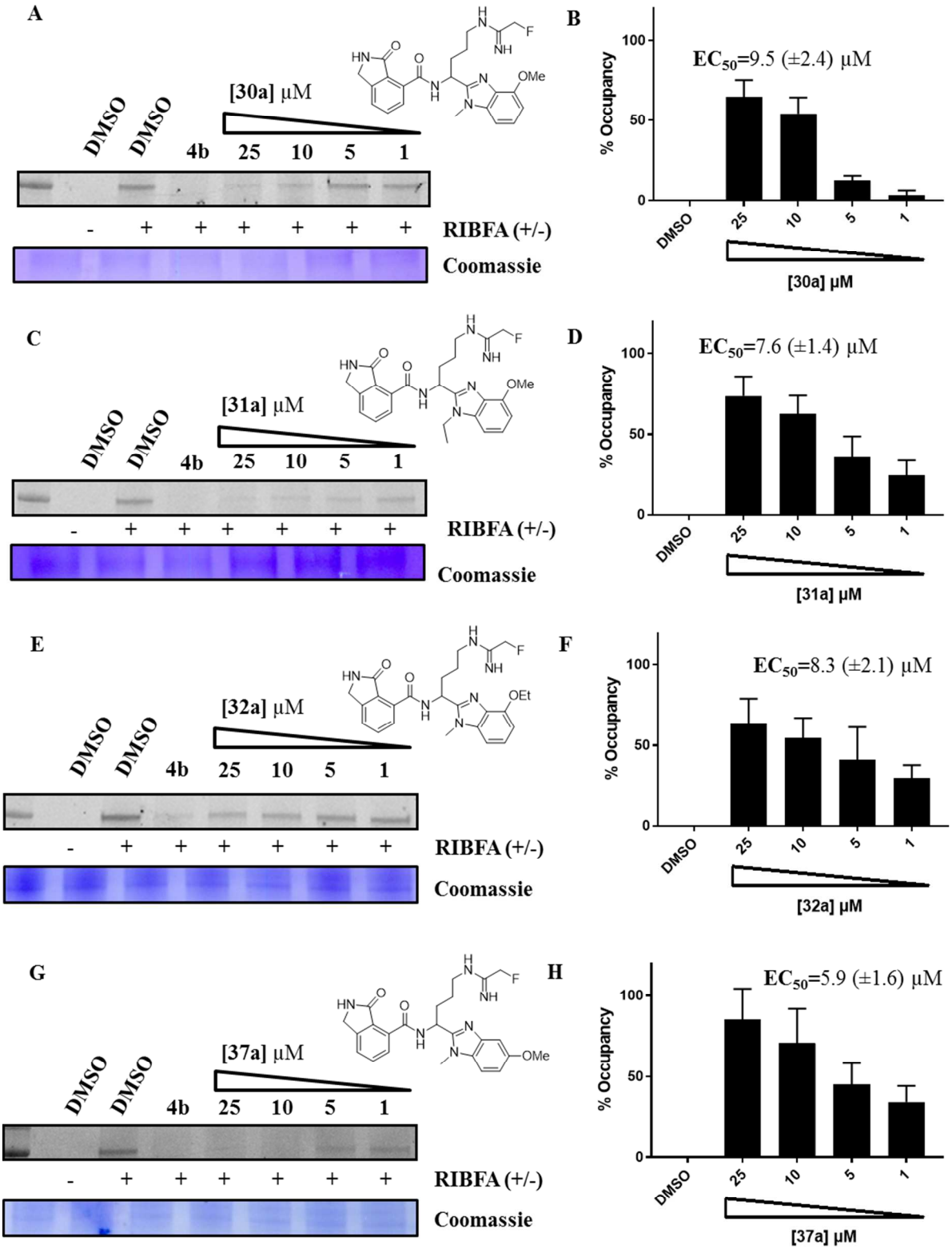

A
[30a] µM
DMSO
DMSO
4b
25
10
5
1
RIBFA (+/-)
Coomassie
B
EC50=9.5 (±2.4) µM
% Occupancy
[30a] µM
C
[31a] µM
EC50=7.6 (±1.4) µM
[31a] µM
E
[32a] µM
EC50=8.3 (±2.1) µM
[32a] µM
G
[37a] µM
H
EC50=5.9 (±1.6) µM
[37a] µM

Figure 15. (**A**) Target engagement assay in HEK293/PAD2 cells for **30a**. (**B**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **30a**. (**C**) Target engagement assay in HEK293/PAD2 cells for **31a**. (**D**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **31a**. (**E**) Target engagement assay in HEK293/PAD2 cells for **32a**. (**F**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **32a**. (**G**) Target engagement assay in HEK293/PAD2 cells for **37a**. (**H**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **37a**. Note: BB-Cl-amidine (**4b**) was evaluated at 25 μM.

For the best compounds (**30a**, **31a**, **32a**, and **37a**) we next obtained a 5-point dose response using our target engagement assay (Figure 15). Excitingly, these compounds all proved to be more potent than BB-F-amidine (**4a**) while also exhibiting similar potencies to BB-Cl-amidine (**4b**); **37a** proved to be the most potent. Notably, however, BB-Cl-amidine (**4b**) was 25-35-fold better at gaining access to the cell than these four unsubstituted lactams (Supporting Information, Table S1). This apparent discrepancy is readily explained by the fact that **30a**, **31a**, **32a**, and **37a** are ~12-50-fold more potent PAD2 inhibitors than BB-Cl-amidine (**4b**) *in vitro*. Thus, the enhanced potency of these compounds overcomes their relatively poor ability to enter cells.

Histone H3 Citrullination in HEK293T/PAD2 Cells

Based on the promising results of our target engagement assay, BB-Cl-amidine (**4b**, Figure S3), **30a**, **31a**, **32a**, and **37a** were then evaluated for their ability to inhibit histone H3 citrullination in HEK293T/PAD2 cells (Figure 16). For these studies, we used a stably transfected HEK293T cell line that overexpresses PAD2 (HEK293T/PAD2). Cells were

incubated with the inhibitor of interest, ionomycin and calcium chloride for 3 h. Cells were then lysed and proteins were resolved by SDS-PAGE and transferred to PVDF membrane. The membranes were then incubated with antibodies for histone H3 or citrullinated histone H3 (Cit 2,8,17), followed by incubation with the appropriate secondary antibodies and imaged by Licor analysis. Similar to the target engagement assay, BB-Cl-amidine (**4b**) strongly inhibited histone H3 citrullination with an EC_{50} of 1.2 μM. Rewardingly, compounds **30a**, **31a**, **32a**, and **37a** showed similar efficacy to BB-Cl-amidine (**4b**) with **30a** demonstrating the greatest ability to inhibit H3 citrullination (Figure 16A, EC_{50}=0.4 μM). Notably, while BB-Cl-amidine (**4b**) demonstrated similar efficacies to **30a**, **31a**, **32a**, and **37a** in both the target engagement and histone citrullination assays, these compounds were >30-fold less cytotoxic than BB-Cl-amidine (Table S1).

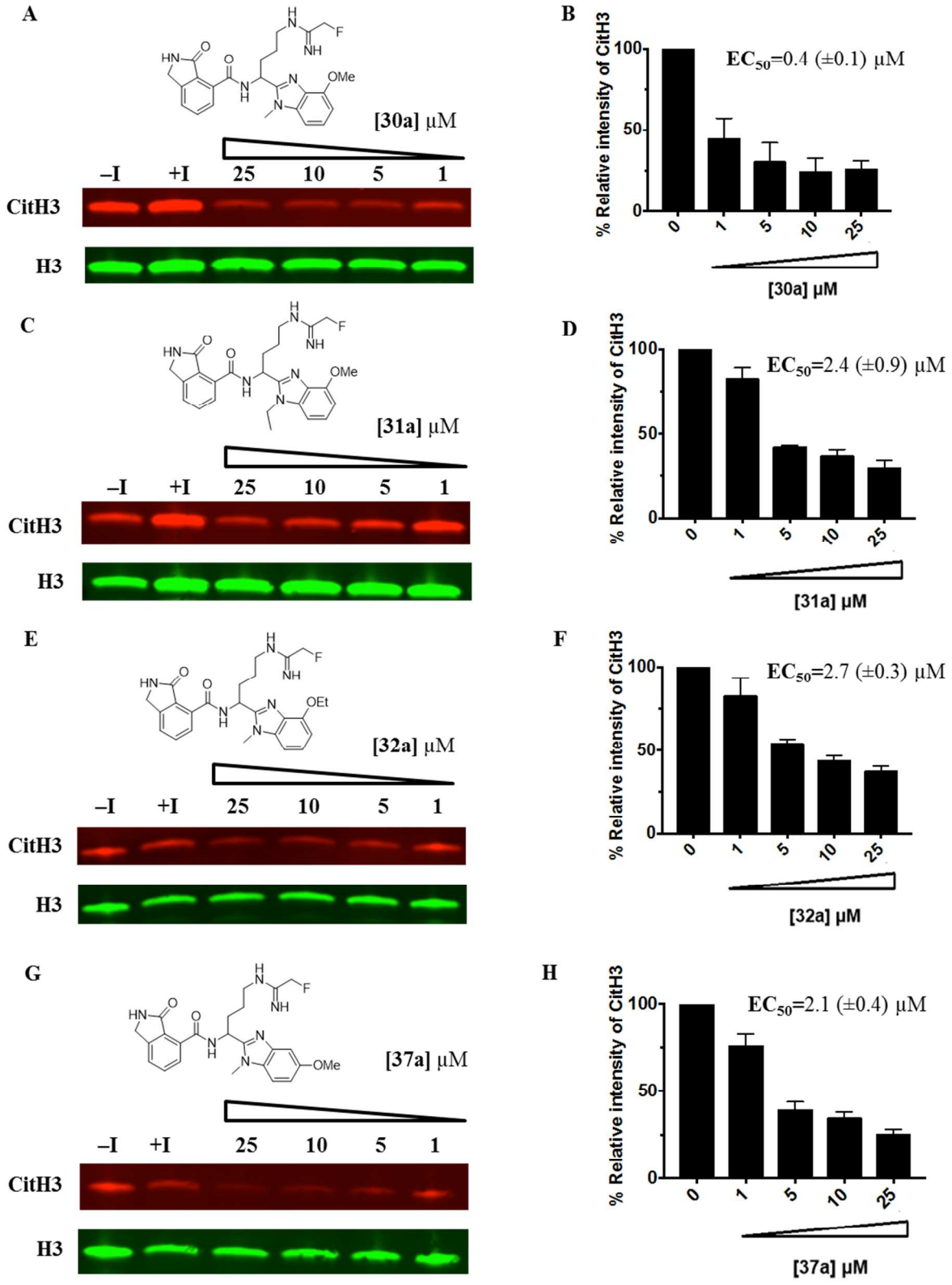

A
[30a] µM
–I
+I
25
10
5
1
CitH3
H3
B
% Relative intensity of CitH3
EC50=0.4 (±0.1) µM
[30a] µM
C
[31a] µM
D
EC50=2.4 (±0.9) µM
[31a] µM
E
[32a] µM
F
EC50=2.7 (±0.3) µM
[32a] µM
G
[37a] µM
H
EC50=2.1 (±0.4) µM
[37a] µM

Figure 16. (**A**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **30a**. (**B**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **30a**. (**C**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **31a**. (**D**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **31a**. (**E**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **32a**. (**F**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **32a**. (**G**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **37a**. (**H**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **37a**. Note: Lane 1 is HEK293T/PAD2 cells treated with DMSO only, lane 2 is HEK293T/PAD2 cells treated with DMSO and 10 μM ionomycin, and lanes 3-6 show HEK293T/PAD2 cells treated with increasing amounts of compound along with 10 μM ionomycin.

Conclusions

Motivated by the pathological roles of PAD2 in numerous inflammatory diseases and cancer, coupled with the dearth of highly potent and selective PAD2 inhibitors, we generated a detailed SAR of a benzimidazole-based scaffold to explore several key features for PAD2 inhibition, including: replacement of the N-terminal ortho-carboxylate with a lactam, *N*-alkylation of the lactam, *N*-alkylation of the benzimidazole, 4- and 5-alkoxy substitution of the benzimdazole, and the electrophilic warhead. While **30a**, **41a** and **49a** were the three most potent PAD2 inhibitors that also exhibit excellent PAD2-selectivity (up to 106-fold), only **30a** proved to be potent in multiple cell-based assays. This was surprising since **41a** and **49a** are more hydrophobic than **30a**, which we expected to enhance cellular uptake. This discrepancy is

partially explained by uptake studies showing that 6-fold more **30a** was able to enter HEK293T cells than **41a**.

These compounds represent a significant advance over BB-Cl-amidine (**4b**) due to their enhanced selectivity for a single PAD isozyme and their use of the less reactive fluoroacetamidine warhead, which significantly limits the number of off targets.[32] Although PAD inhibitors have been developed to inhibit specific isozymes, most of these demonstrate poor efficacy in cell-based assays and none have shown specific and potent inhibition of PAD2. By contrast, **30a** is both highly potent and selective for PAD2 while also demonstrating excellent efficacy in multiple PAD2 cell-based assays. Furthermore, **30a** possesses a large therapeutic window as it is equipotent to BB-Cl-amidine (**4b**) in multiple cell-based assays while exhibiting >30-fold less non-specific toxicity. Ultimately, we expect that **30a** will prove to be a powerful tool for validating the therapeutic relevance of inhibiting PAD2 in multiple disease states, i.e., breast cancer, multiple sclerosis and rheumatoid arthritis. Furthermore, this compound will enable studies to probe the role of this enzyme in numerous cellular pathways where PAD2 is presumed to play a role, including the epigenetic control of gene transcription.

EXPERIMENTAL SECTION

Chemistry. ^{1}H NMR were recorded at either 400 MHz using a Bruker DRX-400 with H/C/P/F QNP gradient probe spectrometer or 500 MHz using a Bruker BioSpin 500MHz Avance AV-III Digital NMR Spectrometer and ^{13}C NMR spectra were recorded at either 100 MHz or 125 MHz: Chemical shifts are reported in δ (ppm) relative to the internal chloroform-*d* ($CDCl_3$, 7.26 ppm) or methanol-*d* (CD_3OD, 3.31 ppm). ESI-HRMS signals were recorded with a Micromass Q-TQF I at the Mass Spectrometry Center at the University of South Carolina (Columbia, SC). The purity of all compounds was determined to be >95% by ^{1}H NMR and ^{13}C NMR spectra. TLC was performed on glass backed silica get plates (Uniplate) with spots visualized by UV light. All solvents were reagent grade and, when necessary, were purified and dried by standard methods. Concentrations of solutions after reactions and extractions involved the use of a rotary evaporator operating at reduced pressure. Synthetic procedures and spectral characterizations of intermediates and final compounds are provided in the supporting information.

Proteins and Cell Culture. PADs 1, 2, 3 and 4 were purified as previously reported.[10, 11, 40] HEK293T and HEK293T cells stably expressing human PAD2 (HEK293T/PAD2) were cultured as previously described.[41, 42] For protein crystallography, recombinant GST-PAD4[43] was expressed in LEMO21 (DE3) cells overnight at 16 °C and cells were lysed in buffer (20mM Tris pH 8.1, 400 mM NaCl, 2 mM DTT, 5 mM EDTA, 20% glycerol, protease inhibitor, 0.5 mg/mL lysozyme, 5 U/mL benzonase) by sonication. Protein was incubated with glutathione sepharose resin and eluted with 10 mM glutathione. The GST tag was removed by cleavage with HRV 3C protease (PreScission Protease) and ion exchange with HiTrapQ FF column (protein eluted on gradient of 0.1-2 M NaCl). Pooled fractions were applied to a Superdex S200 26/60 column,

equilibrated with buffer (20 mM Tris pH 8.1, 400 mM NaCl, 0.5 mM TCEP) before concentration to 4 mg/mL.[43]

Crystallography. Native PAD4 crystals were grown by vapour diffusion at 10 °C by mixing equal volume of protein with precipitant (0.1 M imidazole pH 7.5, 0.2 M Li_2SO_4, 8-11% PEG3350). For soaking, crystals were transferred to soaking solution (0.1 M Imidazole pH 7.5, 0.2 M Li_2SO_4, 10% PEG3350, 1 mM compound (prepared at 100 mM stock in DMSO), 5 mM $CaCl_2$, 0.5 mM TCEP) and incubated overnight. Prior to freezing in liquid nitrogen, crystals were cryoprotected by brief transfer to a solution of mother liquor supplemented with 25% glycerol.

Data Collection. Data were collected on beamlines at the Diamond Light Source (Oxford, UK) and ESRF (Grenoble, France). Automated data processing was performed using Xia2[44-47] or GrenADES.[46, 48, 49] Molecular replacement was performed using DIMPLE[50] (using PDB ID: 3B1U as a reference model) and refinement was performed with REFMAC5[51] and coot.[52] Compound dictionaries were generated using AFITT.[53] Refinement statistics are provided in Table S2.

Inactivation Kinetics. Inactivation kinetic parameters were obtained via established methods.[31, 54] Briefly, values were determined by incubating PAD1, PAD2, PAD4 (2.0 μM) or PAD3 (5.0 μM) in a pre-warmed (10 min; 37 °C) inactivation mixture (50 mM HEPES, 10 mM $CaCl_2$, and 2 mM DTT, pH 7.6, with a final volume of 60 μL) containing various concentrations of inhibitor. Aliquots were removed at various time points and added to a pre-warmed (10 min, 37

°C) reaction mixture (50 mM HEPES, 10 mM $CaCl_2$, and 2 mM DTT, and 10 mM BAEE or 10 mM BAA in the case of PAD3; pH 7.6). After 15 min, reactions were quenched in liquid nitrogen and citrulline production was quantified using the COLDER assay.[11, 55] Data were plotted as a function of time and fit to **eq 1**,

$$v = v_o e^{-kt} \qquad \textbf{eq 1,}$$

using Grafit version 5.0.11, where v is velocity, v_o is initial velocity, k (or k_{obs}) is the pseudo-first order rate constant of inactivation, and t is time. When saturation was reached upon plotting k_{obs} versus inactivator concentration, the data were fit to **eq 2**,

$$k_{obs}=k_{inact}[I]/(K_I+[I]) \qquad \textbf{eq 2,}$$

using GraFit version 5.0.11, where k_{inact} corresponds to that maximal rate of inactivation and K_I is the concentration of inhibitor that gives half-maximal inactivation. If the plot of k_{obs} versus [I] was linear and did not saturate, then the value for k_{inact}/K_I equaled the slope of the line. All experiments were performed in at least duplicate. k_{inact}/K_I values are depicted graphically in the main text and summarized in Tables S3-S10 in the Supporting Information.

Target engagement assay with HEK293T/PAD2 cells. HEK293T cells stably expressing human PAD2 (HEK293T/PAD2) were cultured as previously described.[41, 42] Cells were grown to ~80% confluence in a Nunc Cell Culture Treated EasYFlask™ T175 flask, trypsinized, and trypsin activity quenched with complete media. The cells were harvested by centrifugation at 1000 x g for 3.5 min, and resuspended in complete media. Cells were then added to two 6-well plates (2 mL/well) and incubated overnight. After overnight incubation, cells reached 80-90% confluence and were ready for the target engagement assay.

The complete media was removed from each well of the 6-well plate and carefully replaced with 1x HBS (2 mL/well) containing 2 mM $CaCl_2$ and incubated at 37 °C with 5% CO_2 for 15 min. Cells were then treated with either DMSO (1 µL) or varying concentrations of inhibitor (1 µL in DMSO) and incubated at 37 °C with 5% CO_2 for another 15 min. Cells were then treated with 1 µL of ionomycin (0.5 µM final concentration) and incubated at 37 °C with 5% CO_2 for 1 h.

A cell scraper was then used to harvest the cells, followed by centrifugation at 1000 x g for 3.5 min. The cell pellets were then washed 4x with 1x HBS, transferred to 1.6 mL tubes, and the HBS was removed. The cell pellets were then resuspended in 1x HBS with 1% Triton X-100 and incubated for 1 h at 0 °C. Lysates were cleared by centrifugation at 21,000 x g for 15 min. The soluble protein fraction was isolated and quantified by the DC assay (Bio-Rad). Lysates (2 µg/µL, 50 µL) were then treated with $CaCl_2$ (2 mM final concentration) and RIBFA (2 µM final concentration). The tubes were then incubated at 37 °C for 1 h followed by quenching with 6x SDS loading buffer and separated by SDS-PAGE (4-15% gradient gel). The bands were visualized by scanning the gel in a typhoon scanner (excitation/emission maxima ~546/579, respectively).

Histone H3 Citrullination in HEK293T/PAD2 cells. HEK293T cells stably expressing human PAD2 (HEK293T/PAD2) were cultured as previously described.[41, 42] Cells were grown to ~80% confluence (8 x 10^6 cells), trypsinized, and quenched with complete media. The cells were harvested by centrifugation at 1000 x g for 3.5 min and washed 4x with 1x HBS. Cells were resuspended in 1x HBS at 8 x 10^6 cells/mL, and 4 x 10^5 cells were added to 0.65 mL tubes for the subsequent assay. Ionomycin (10 µM), $CaCl_2$ (1 mM) and either DMSO or varying

concentrations of inhibitor were added to the cells and incubated for 3 h. The final concentration of DMSO was 1% in each sample.

Triton X-100 was then added to the cell suspensions (1% final concentration) and sonicated at 0 °C for 1 h. Lysates were cleared by centrifugation at 21,000 x g for 15 min, and soluble proteins were removed and quantified by DC assay (Bio-Rad). Lysates were normalized and 1 µg of protein was separated (4-15% gradient gel) and transferred to PVDF membranes (Bio-Rad) at 80 V for 60 min. Membranes were then blocked with TBST and BSA (5%) for 1 h at 23 °C. Blocked membranes were incubated with antibodies for histone H3 (1:2000) or histone H3 Cit 2,8,17 (1:1000) in TBST with 5% BSA for 12 h at 4 °C. Membranes were then washed with TBST (9x) and incubated with anti-Rabbit IgG Licor conjugate (1:5000) for 1 h at 23 °C. Membranes were washed with TBST (9x) and imaged by Licor analysis.

Cellular Uptake Assays. HEK293T/PAD2 cells stably expressing PAD2 were cultured and treated with various inhibitors similarly to the target engagement assay described above. After the cells were treated with $CaCl_2$ (2 mM), ionomycin (0.5 µM) and various inhibitors (25 µM), a cell scraper was then used to harvest the cells, followed by centrifugation at 1000 x g for 3.5 min. The cell pellets were then washed 4x with 1x HBS, transferred to 1.6 mL tubes, and the HBS was removed. The cell pellets were then resuspended in acetonitrile and sonicated for 1 h at 0 °C. Lysates were cleared by centrifugation at 21,000 x g for 15 min. The soluble protein fraction was isolated, filtered, and injected onto the LCMS. The area under the curve was then taken to be the 'area of analyte' as described in Table S1. The retention time and 'area of standard' were both determined by injecting 25 µM of inhibitor directly onto the LCMS and integrating the area under the curve.

Cytotoxicity Studies. Human HEK293T cells were plated (2.5 x 10^6 cells/well) in a 96-well plate. The next day the cells were treated with DMSO (1 μL), Triton X-100 (1 μL), or various concentrations of BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), **30a**, **31a**, **32a**, **37a** or **41a** and incubated for 24 h. Cell viability was measured using the XTT reagent kit (ATCC) by reading the absorption at 475 nm using a Spectramax plate reader. EC_{50} values for cell growth inhibition were determined by fitting an eight-point dose-response curve to eq 3,

$$Y = \text{Range}/(1 + ([I]/EC_{50}))^s + \text{background} \qquad \textbf{eq 3,}$$

using GraFit 5.0.11, where range is the uninhibited value minus the background and *s* is the slope factor.

ASSOCIATED CONTENT

Supporting Information

Synthetic procedures, experimental details, Figures S1-S3, and Tables S1-S10. This material is available free of charge on the ACS Publications websites at DOI:

AUTHOR INFORMATION

Corresponding Author

*E-mail: paul.thompson@umassmed.edu. Tel: 508-856-8492

ORCID

Paul R. Thompson: 0000-0002-1621-3372

Notes

The authors declare the following competing financial interest(s): P.R.T. is a consultant to Bristol-Myers Squibb. The remaining authors have no competing interest to declare.

ACKNOWLEDGMENTS

This work was supported in part by NIH grants GM079357, GM110394, and GM118112 to P.R.T.

ABBREVIATIONS USED

RA, Rheumatoid arthritis; PAD, Protein arginine deiminase; ACPA, Anti-citrullinated protein antibodies; NET, Neutrophil extracellular trap; RFA, Rhodamine-conjugated F-amidine; TAMRA-N_3, Tetramethylrhodamine azide; Fmoc, Fluorenylmethyloxycarbonyl; Orn, Ornithine; Boc, *tert*-butyloxycarbonyl; HBTU, (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, Hydroxybenzotriazole; AcOH, Acetic acid; DIPEA, N,N-Diisopropylethylamine; DMF, Dimethylformamide; HCl, Hydrochloric acid; Et_2O, Diethyl ether; MeOH, Methanol; TEA, Triethylamine; SDS-PAGE, Sodium dodecyl sulfate polyacrylamide gel electrophoresis; kDa, Kilodalton; MBP, myelin basic protein; MS, multiple sclerosis; TDFA, threonine-aspartic acid-F-amidine; HEK293T, human embryonic kidney 293 cells; PBS, phosphate buffered saline; HBS, HEPES buffered saline; PVDF, polyvinylidene fluorine; DMSO, dimethyl sulfoxide; TBTA, Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine

Accession Codes

The coordinates for the PAD4 bound to BB-F-amidine, AFM-14a, AFM-30a, and AFM-41a have been deposited in the PDB under accession codes 5N0M, 5N1B, 5N0Y, and 5N0Z, respectively. Authors will release the atomic coordinates and experimental data upon article publication.

REFERENCES

(1) Fuhrmann, J.; Clancy, K. W.; Thompson, P. R. Chemical biology of protein arginine modifications in epigenetic regulation. *Chem. Rev.* **2015,** *115*, 5413-5461.

(2) 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.* **2009,** *12*, 616-627.

(3) Slade, D. J.; Subramanian, V.; Thompson, P. R. Pluripotency: Citrullination unravels stem cells. *Nat. Chem. Biol.* **2014,** *10*, 327-328.

(4) Wang, Y.; Wysocka, J.; Sayegh, J.; Lee, Y. H.; Perlin, J. R.; Leonelli, L.; Sonbuchner, L. S.; McDonald, C. H.; Cook, R. G.; Dou, Y.; Roeder, R. G.; Clarke, S.; Stallcup, M. R.; Allis, C. D.; Coonrod, S. A. Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **2004,** *306*, 279-283.

(5) Wang, Y.; Li, M.; Stadler, S.; Correll, S.; Li, P.; Wang, D.; Hayama, R.; Leonelli, L.; Han, H.; Grigoryev, S. A.; Allis, C. D.; Coonrod, S. A. Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J. Cell Biol.* **2009,** *184*, 205-213.

(6) Christophorou, M. A.; Castelo-Branco, G.; Halley-Stott, R. P.; Oliveira, C. S.; Loos, R.; Radzisheuskaya, A.; Mowen, K. A.; Bertone, P.; Silva, J. C.; Zernicka-Goetz, M.; Nielsen, M. L.; Gurdon, J. B.; Kouzarides, T. Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **2014,** *507*, 104-108.

(7) Cuthbert, G. L.; Daujat, S.; Snowden, A. W.; Erdjument-Bromage, H.; Hagiwara, T.; Yamada, M.; Schneider, R.; Gregory, P. D.; Tempst, P.; Bannister, A. J.; Kouzarides, T. Histone deimination antagonizes arginine methylation. *Cell* **2004,** *118*, 545-553.

(8) Vossenaar, E. R.; Zendman, A. J.; van Venrooij, W. J.; Pruijn, G. J. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* **2003,** *25*, 1106-1118.

(9) Raijmakers, R.; Zendman, A. J.; Egberts, W. V.; Vossenaar, E. R.; Raats, J.; Soede-Huijbregts, C.; Rutjes, F. P.; van Veelen, P. A.; Drijfhout, J. W.; Pruijn, G. J. Methylation of arginine residues interferes with citrullination by peptidylarginine deiminases in vitro. *J. Mol. Biol.* **2007,** *367*, 1118-1129.

(10) Knuckley, B.; Causey, C. P.; Jones, J. E.; Bhatia, M.; Dreyton, C. J.; Osborne, T. C.; Takahara, H.; Thompson, P. R. Substrate specificity and kinetic studies of PADs 1, 3, and 4 identify potent and selective inhibitors of protein arginine deiminase 3. *Biochemistry* **2010,** *49*, 4852-4863.

(11) Kearney, P. L.; Bhatia, M.; Jones, N. G.; Luo, Y.; Glascock, M. C.; Catchings, K. L.; Yamada, M.; Thompson, P. R. Kinetic characterization of protein arginine deiminase 4: A transcriptional corepressor implicated in the onset and progression of rheumatoid arthritis. *Biochemistry* **2005,** *44*, 10570-10582.

(12) Slade, D. J.; Fang, P.; Dreyton, C. J.; Zhang, Y.; Fuhrmann, J.; Rempel, D.; Bax, B. D.; Coonrod, S. A.; Lewis, H. D.; Guo, M.; Gross, M. L.; Thompson, P. R. Protein arginine deiminase 2 binds calcium in an ordered fashion: implications for inhibitor design. *ACS Chem. Biol.* **2015,** *10*, 1043-1053.

(13) Lamensa, J. W.; Moscarello, M. A. Deimination of human myelin basic protein by a peptidylarginine deiminase from bovine brain. *J. Neurochem.* **1993,** *61*, 987-996.

(14) Moscarello, M. A.; Pritzker, L.; Mastronardi, F. G.; Wood, D. D. Peptidylarginine deiminase: a candidate factor in demyelinating disease. *J. Neurochem.* **2002,** *81*, 335-343.

(15) Damgaard, D.; Senolt, L.; Nielsen, M. F.; Pruijn, G. J.; Nielsen, C. H. Demonstration of extracellular peptidylarginine deiminase (PAD) activity in synovial fluid of patients with rheumatoid arthritis using a novel assay for citrullination of fibrinogen. *Arthritis Res. Ther.* **2014,** *16*, 498.

(16) Masson-Bessiere, C.; Sebbag, M.; Girbal-Neuhauser, E.; Nogueira, L.; Vincent, C.; Senshu, T.; Serre, G. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. *J. Immunol.* **2001,** *166*, 4177-4184.

(17) Burska, A. N.; Hunt, L.; Boissinot, M.; Strollo, R.; Ryan, B. J.; Vital, E.; Nissim, A.; Winyard, P. G.; Emery, P.; Ponchel, F. Autoantibodies to posttranslational modifications in rheumatoid arthritis. *Mediators Inflamm.* **2014,** *2014*, 492873.

(18) Van Steendam, K.; Tilleman, K.; Deforce, D. The relevance of citrullinated vimentin in the production of antibodies against citrullinated proteins and the pathogenesis of rheumatoid arthritis. *Rheumatology (Oxford)* **2011,** *50*, 830-837.

(19) McElwee, J. L.; Mohanan, S.; Griffith, O. L.; Breuer, H. C.; Anguish, L. J.; Cherrington, B. D.; Palmer, A. M.; Howe, L. R.; Subramanian, V.; Causey, C. P.; Thompson, P. R.; Gray, J. W.; Coonrod, S. A. Identification of PADI2 as a potential breast cancer biomarker and therapeutic target. *BMC Cancer* **2012,** *12*, 500.

(20) Zhang, X.; Bolt, M.; Guertin, M. J.; Chen, W.; Zhang, S.; Cherrington, B. D.; Slade, D. J.; Dreyton, C. J.; Subramanian, V.; Bicker, K. L.; Thompson, P. R.; Mancini, M. A.; Lis, J. T.; Coonrod, S. A. Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor alpha target gene activation. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109*, 13331-13336.

(21) Mohanan, S.; Cherrington, B. D.; Horibata, S.; McElwee, J. L.; Thompson, P. R.; Coonrod, S. A. Potential role of peptidylarginine deiminase enzymes and protein citrullination in cancer pathogenesis. *Biochem. Res. Int.* **2012,** *2012*, 895343.

(22) Ghari, F.; Quirke, A. M.; Munro, S.; Kawalkowska, J.; Picaud, S.; McGouran, J.; Subramanian, V.; Muth, A.; Williams, R. O.; Kessler, B.; Thompson, P. R.; Fillipakopoulos, P.; Knapp, S.; Venables, P. J.; La Thangue, N. B. Citrullination-acetylation interplay guides E2F-1 activity during the inflammatory response. *Sci. Adv.* **2016,** *2*, e1501257.

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

(24) Willis, V. C.; Gizinski, A. M.; Banda, N. K.; Causey, C. P.; Knuckley, B.; Cordova, K. N.; Luo, Y.; Levitt, B.; Glogowska, M.; Chandra, P.; Kulik, L.; Robinson, W. H.; Arend, W. P.; Thompson, P. R.; Holers, V. M. N-alpha;-benzoyl-N5-(2-chloro-1-iminoethyl)-L-ornithine amide, a protein arginine deiminase inhibitor, reduces the severity of murine collagen-induced arthritis. *J. Immunol.* **2011,** *186*, 4396-4404.

(25) Kaplan, M. J. Neutrophils in the pathogenesis and manifestations of SLE. *Nat. Rev. Rheumatol.* **2011,** *7*, 691-699.

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

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

(28) Chumanevich, A. A.; Causey, C. P.; Knuckley, B. A.; Jones, J. E.; Poudyal, D.; Chumanevich, A. P.; Davis, T.; Matesic, L. E.; Thompson, P. R.; Hofseth, L. J. Suppression of colitis in mice by Cl-amidine: A novel peptidylarginine deiminase (PAD) inhibitor. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2011,** *300*, G929-G938.

(29) Lange, S.; Gogel, S.; Leung, K. Y.; Vernay, B.; Nicholas, A. P.; Causey, C. P.; Thompson, P. R.; Greene, N. D.; Ferretti, P. Protein deiminases: New players in the developmentally regulated loss of neural regenerative ability. *Dev. Biol.* **2011,** *355*, 205-214.

(30) Knight, J. S.; Luo, W.; O'Dell, A. A.; Yalavarthi, S.; Zhao, W.; Subramanian, V.; Guo, C.; Grenn, R. C.; Thompson, P. R.; Eitzman, D. T.; Kaplan, M. J. Peptidylarginine deiminase inhibition reduces vascular damage and modulates innate immune responses in murine models of atherosclerosis. *Circ. Res.* **2014**, *114*, 947-956.

(31) Luo, Y.; Arita, K.; Bhatia, M.; Knuckley, B.; Lee, Y. H.; Stallcup, M. R.; Sato, M.; Thompson, P. R. Inhibitors and inactivators of protein arginine deiminase 4: Functional and structural characterization. *Biochemistry* **2006,** *45*, 11727-11736.

(32) Luo, Y.; Knuckley, B.; Lee, Y. H.; Stallcup, M. R.; Thompson, P. R. A fluoro-acetamidine based inactivator of protein arginine deiminase 4 (PAD4): Design, synthesis, and in vitro and in vivo evaluation. *J. Am. Chem. Soc.* **2006,** *128*, 1092-1093.

(33) Jones, J. E.; Slack, J. L.; Fang, P.; Zhang, X.; Subramanian, V.; Causey, C. P.; Coonrod, S. A.; Guo, M.; Thompson, P. R. Synthesis and screening of a haloacetamidine containing library to identify PAD4 selective inhibitors. *ACS Chem. Biol.* **2012,** *7*, 160-165.

(34) Bicker, K. L.; Anguish, L.; Chumanevich, A. A.; Cameron, M. D.; Cui, X.; Witalison, E.; Subramanian, V.; Zhang, X.; Chumanevich, A. P.; Hofseth, L. J.; Coonrod, S. A.; Thompson, P. R. D-amino acid based protein arginine deiminase inhibitors: Synthesis, pharmacokinetics, and in cellulo efficacy. *ACS Med. Chem. Lett.* **2012,** *3*, 1081-1085.

(35) Jamali, H.; Khan, H. A.; Stringer, J. R.; Chowdhury, S.; Ellman, J. A. Identification of multiple structurally distinct, nonpeptidic small molecule inhibitors of protein arginine deiminase 3 using a substrate-based fragment method. *J. Am. Chem. Soc.* **2015,** *137*, 3616-3621.

(36) Jamali, H.; Khan, H. A.; Tjin, C. C.; Ellman, J. A. Cellular activity of new small molecule protein arginine deiminase 3 (PAD3) inhibitors. *ACS Med. Chem. Lett.* **2016,** *7*, 847-851.

(37) Causey, C. P.; Jones, J. E.; Slack, J. L.; Kamei, D.; Jones, L. E.; Subramanian, V.; Knuckley, B.; Ebrahimi, P.; Chumanevich, A. A.; Luo, Y.; Hashimoto, H.; Sato, M.; Hofseth, L. J.; Thompson, P. R. The development of N-alpha-(2-carboxyl)benzoyl-N(5)-(2-fluoro-1-iminoethyl)-l-ornithine amide (o-F-amidine) and N-alpha-(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.* **2011,** *54*, 6919-6935.

(38) Subramanian, V.; Knight, J. S.; Parelkar, S.; Anguish, L.; Coonrod, S. A.; Kaplan, M. J.; Thompson, P. R. Design, synthesis, and biological evaluation of tetrazole analogs of Cl-amidine as protein arginine deiminase inhibitors. *J. Med. Chem.* **2015,** *58*, 1337-1344.

(39) Saijo, S.; Nagai, A.; Kinjo, S.; Mashimo, R.; Akimoto, M.; Kizawa, K.; Yabe-Wada, T.; Shimizu, N.; Takahara, H.; Unno, M. Monomeric form of peptidylarginine deiminase type I revealed by x-ray crystallography and small-angle x-ray scattering. *J. Mol. Biol.* **2016,** *428*, 3058-3073.

(40) Dreyton, C. J.; Knuckley, B.; Jones, J. E.; Lewallen, D. M.; Thompson, P. R. Mechanistic studies of protein arginine deiminase 2: evidence for a substrate-assisted mechanism. *Biochemistry* **2014,** *53*, 4426-4433.

(41) Lewallen, D. M.; Bicker, K. L.; Madoux, F.; Chase, P.; Anguish, L.; Coonrod, S. A.; Hodder, P.; Thompson, P. R. A FluoPol-ABPP PAD2 high-throughput screen identifies the first calcium site inhibitor targeting the PADs. *ACS Chem. Biol.* **2014,** *9*, 913-921.

(42) Lewallen, D. M.; Bicker, K. L.; Subramanian, V.; Clancy, K. W.; Slade, D. J.; Martell, J.; Dreyton, C. J.; Sokolove, J.; Weerapana, E.; Thompson, P. R. Chemical proteomic platform to identify citrullinated proteins. *ACS Chem. Biol.* **2015,** *10*, 2520-2528. .

(43) Arita, K.; Hashimoto, H.; Shimizu, T.; Nakashima, K.; Yamada, M.; Sato, M. Structural basis for Ca^{2+}-induced activation of human PAD4. *Nat. Struct. Mol. Biol.* **2004,** *11*, 777-783.

(44) Winter, G. XIA2: an expert system for macromolecular crystallography data reduction. *J Appl. Cryst.* **2010,** *43*, 186-190.

(45) Winter, G.; Lobley, C. M.; Prince, S. M. Decision making in xia2. *Acta. Crystallogr. D Biol. Crystallogr.* **2013,** *69*, 1260-1273.

(46) Kabsch, W. Xds. *Acta. Crystallogr. D Biol. Crystallogr.* **2010,** *66*, 125-132.

(47) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta. Crystallogr. D Biol. Crystallogr.* **2013,** *69*, 1204-1214.

(48) Monaco, S.; Gordon, E.; Bowler, M. W.; Delageniere, S.; Guijarro, M.; Spruce, D.; Svensson, O.; McSweeney, S. M.; McCarthy, A. A.; Leonard, G.; Nanao, M. H. Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. *J. Appl. Crystallogr.* **2013,** *46*, 804-810.

(49) Evans, P. R. Scaling and asessment of data quality. *Acta. Crystallogr. D Biol. Crystallogr.* **2006,** *62*, 72-82.

(50) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta. Crystallogr. D Biol. Crystallogr.* **2011,** *67*, 235-242.

(51) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D Biol. Crystallogr.* **1997,** *53*, 240-255.

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

(53) Wlodek, S.; Skillman, A. G.; Nicholls, A. Automated ligand placement and refinement with a combined force field and shape potential. *Acta. Crystallogr. D Biol. Crystallogr.* **2006,** *62*, 741-749.

(54) Knuckley, B.; Causey, C. P.; Pellechia, P. J.; Cook, P. F.; Thompson, P. R. Haloacetamidine-based inactivators of protein arginine deiminase 4 (PAD4): Evidence that general acid catalysis promotes efficient inactivation. *Chembiochem* **2010,** *11*, 161-165.

(55) Knipp, M.; Vasak, M. A colorimetric 96-well microtiter plate assay for the determination of enzymatically formed citrulline. *Anal. Biochem.* **2000,** *286*, 257-264.

Figure Legends

Figure 1. Peptidyl-arginine (**1**) to peptidyl-citrulline (**2**) hydrolysis reaction catalyzed by the PADs.

Figure 2. Structures and k_{inact}/K_I values for the 1st-Generation PAD inhibitors F-amidine (**3a**) and Cl-amidine (**3b**) and the 2nd-Generation PAD inhibitors BB-F-amidine (**4a**) and BB-Cl-amidine (**4b**).

Figure 3. (**A**) Co-crystal structure of BB-F-amidine (**4a**) bound to PAD4 (PDB ID: 5N0M). (**B**) PAD2 (PDB ID: 4N2C)[12] was superimposed on the PAD4·BB-F-amidine (**4a**) co-crystal structure and then PAD4 was removed leaving BB-F-amidine. (**C**) Evolution of 2nd-Generation PAD inhibitors BB-F-amidine (**4a**) and BB-Cl-amidine (**4b**) to 3rd-Generation benzimidazole-based PAD inhibitors.

Figure 4. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 1 compounds **9-12**. (**B**) Summary of isozyme selectivity for series 1 compounds **9-12** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 5. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 2 compounds **14-17**. (**B**) Summary of isozyme selectivity for series 2 compounds **14-17** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 6. (**A**) Co-crystal structure of **AFM-14a** bound to PAD4. (**B**) Co-crystal structure of *o*-F-amidine bound to PAD4[37] (PDB: 3B1T). (**C**) PAD1[39] (PDB: 5HP5) overlay with PAD4•**AFM-14a** co-crystal structure in PyMol. (**D**) PAD1[39] (PDB: 5HP5) overlay with PAD4•*o*-F-amidine

co-crystal structure in PyMol. (**E**) PAD2[12] (PDB: 4N2C) overlay with PAD4•**AFM-14a** co-crystal structure in PyMol. (**F**) PAD2[12] (PDB: 4N2C) overlay with PAD4•o-F-amidine co-crystal structure in PyMol.

Figure 7. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 3 compounds **19-22**. (**B**) Summary of isozyme selectivity for series 3 compounds **19-22** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 8. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 4 compounds **29-32**. (**B**) Summary of isozyme selectivity for series 4 compounds **29-32** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 9. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 5 compounds **36-38**. (**B**) Summary of isozyme selectivity for series 5 compounds **36-38** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 10. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 6 compounds **40-43**. (**B**) Summary of isozyme selectivity for series 6 compounds **40-43** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 11. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 7 compounds **45-47**. (**B**) Summary of isozyme selectivity for series 1 compounds **45-47** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 12. (**A**) k_{inact}/K_I values for BB-F-amidine (**4a**), BB-Cl-amidine (**4b**), and series 8 compounds **45-47**. (**B**) Summary of isozyme selectivity for series 1 compounds **45-47** based on their respective k_{inact}/K_I ratios setting the enzyme with the lowest value as 1.

Figure 13. (**A**) Co-crystal structure of **AFM-30a** bound to PAD4 (PDB: 5N0Y). (**B**) Co-crystal structure of **AFM-41a** bound to PAD4 (PDB: 5N0Z). (**C**) PAD1[39] (PDB: 5HP5) overlay with PAD4•**AFM-30a** co-crystal structure in PyMol. (**D**) PAD1[39] (PDB: 5HP5) overlay with PAD4•**AFM-41a** co-crystal structure in PyMol. (**E**) PAD2[12] (PDB: 4N2C) overlay with PAD4•**AFM-30a** co-crystal structure in PyMol. (**F**) PAD2[12] (PDB: 4N2C) overlay with PAD4•**AFM-41a** co-crystal structure in PyMol.

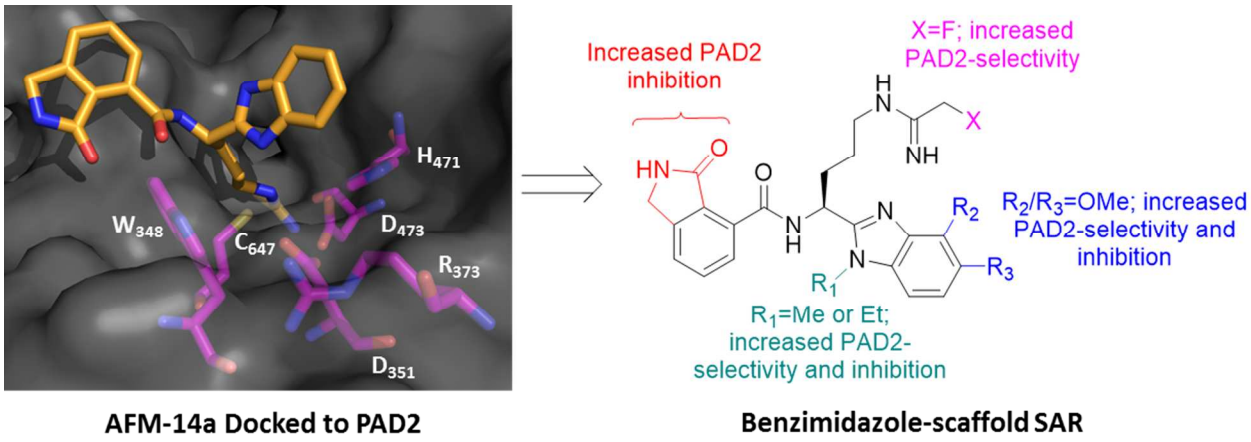
Figure 14. (**A**) Target engagement assay in HEK293/PAD2 cells for BB-Cl-amidine (**4b**). (**B**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by BB-Cl-amidine (**4b**). (**C**) Target engagement assay in HEK293/PAD2 cells for BB-F-amidine (**4a**). (**D**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by BB-F-amidine (**4a**). Note: Panel C, BB-Cl-amidine (**4b**) was evaluated at 25 μM.

Figure 15. (**A**) Target engagement assay in HEK293/PAD2 cells for **30a**. (**B**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **30a**. (**C**) Target engagement assay in HEK293/PAD2 cells for **31a**. (**D**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **31a**. (**E**) Target engagement assay in HEK293/PAD2 cells for **32a**. (**F**) Quantified results of target engagement assay plotted as

percent occupancy of PAD2 by **32a**. (**G**) Target engagement assay in HEK293/PAD2 cells for **37a**. (**H**) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by **37a**. Note: BB-Cl-amidine (**4b**) was evaluated at 25 μM.

Figure 16. (**A**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **30a**. (**B**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **30a**. (**C**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **31a**. (**D**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **31a**. (**E**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **32a**. (**F**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **32a**. (**G**) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for **37a**. (**H**) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for **37a**. Note: Lane 1 is HEK293T/PAD2 cells treated with DMSO only, lane 2 is HEK293T/PAD2 cells treated with DMSO and 10 μM ionomycin, and lanes 3-6 show HEK293T/PAD2 cells treated with increasing amounts of compound along with 10 μM ionomycin.

Graphical Abstract



AFM-14a Docked to PAD2 **Benzimidazole-scaffold SAR**

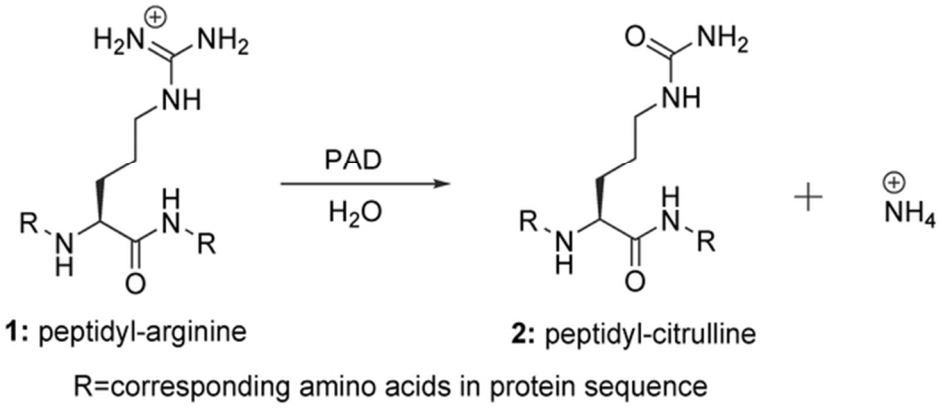
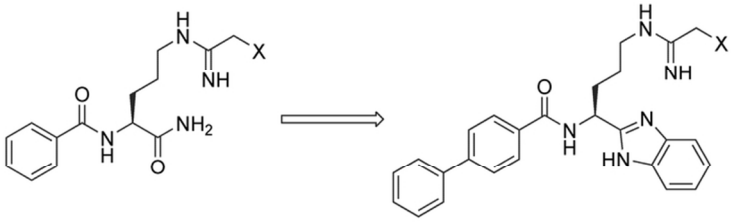


Figure 1. Peptidyl-arginine (1) to peptidyl-citrulline (2) hydrolysis reaction catalyzed by the PADs.

59x28mm (300 x 300 DPI)



3a: X=F; **F-amidine**	**3b:** X=Cl; **Cl-amidine**	**4a:** X=F; **BB-F-amidine**	**4b:** X=Cl; **BB-Cl-amidine**
k_{inact}/K_I (M^{-1} min^{-1})	k_{inact}/K_I (M^{-1} min^{-1})	k_{inact}/K_I (M^{-1} min^{-1})	k_{inact}/K_I (M^{-1} min^{-1})
PAD1=2800	PAD1=37000	PAD1=900	PAD1=16100
PAD2=380	PAD2=1200	PAD2=1200	PAD2=4100
PAD3=170	PAD3=2000	PAD3=3400	PAD3=6800
PAD4=3000	PAD4=13000	PAD4=3750	PAD4=13300

Figure 2. Structures and kinact/KI values for the 1st-Generation PAD inhibitors F-amidine (3a) and Cl-amidine (3b) and the 2nd-Generation PAD inhibitors BB-F-amidine (4a) and BB-Cl-amidine (4b).

83x40mm (300 x 300 DPI)

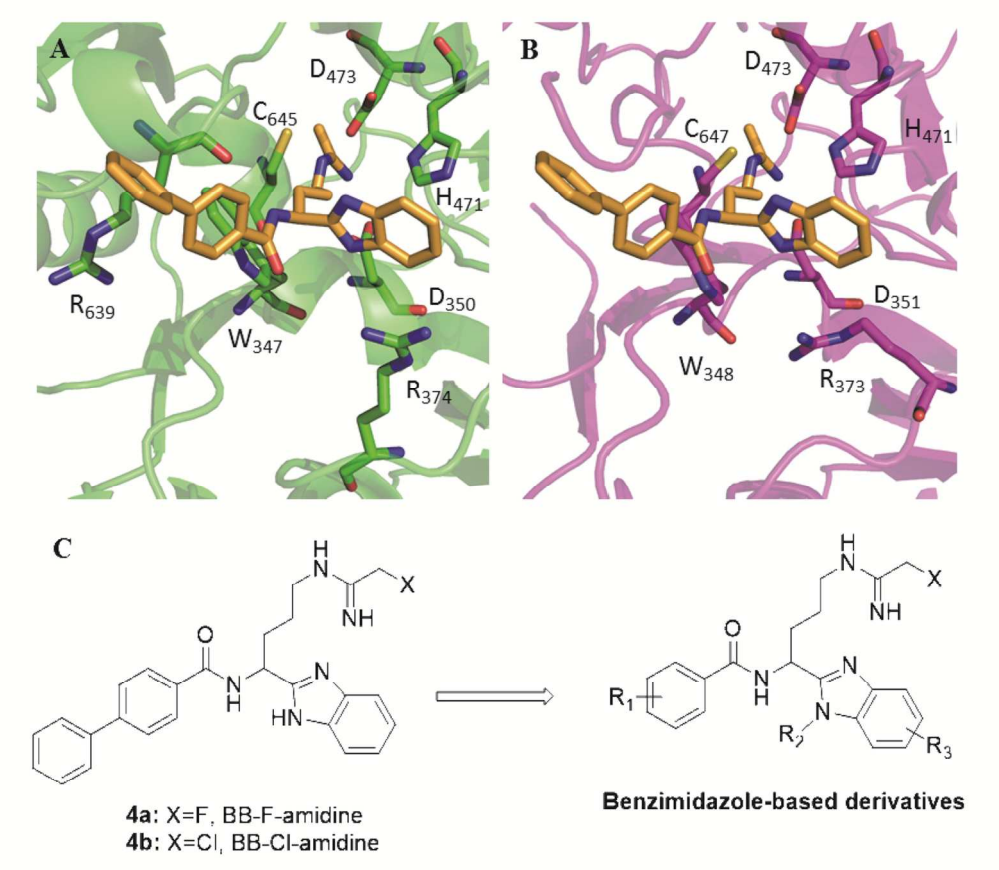
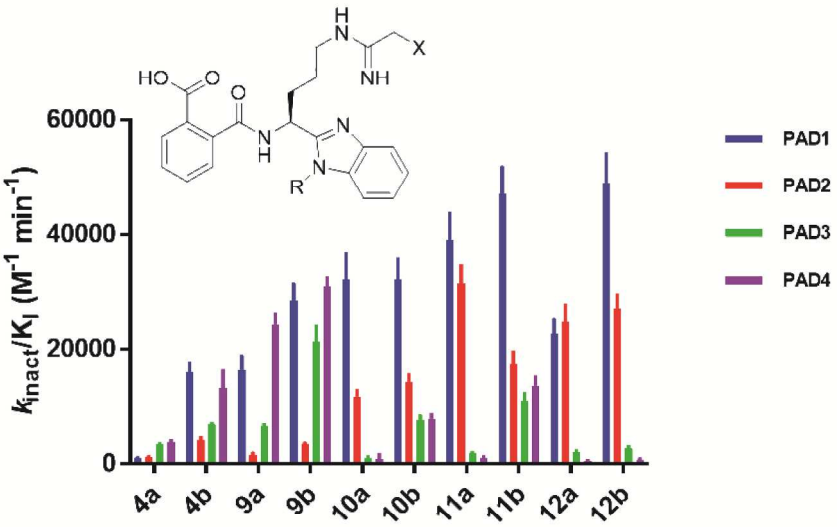


Figure 3. (A) Co-crystal structure of BB-F-amidine (4a) bound to PAD4. (B) PAD2 (PDB: 4N2C)12 overlay with the PAD4•BB-F-amidine (4a) co-crystal structure in PyMol. (C) Evolution of 2nd-Generation PAD inhibitors BB-F-amidine (4a) and BB-Cl-amidine (4b) to 3rd-Generation benzimidazole-based PAD inhibitors.

131x115mm (300 x 300 DPI)

A



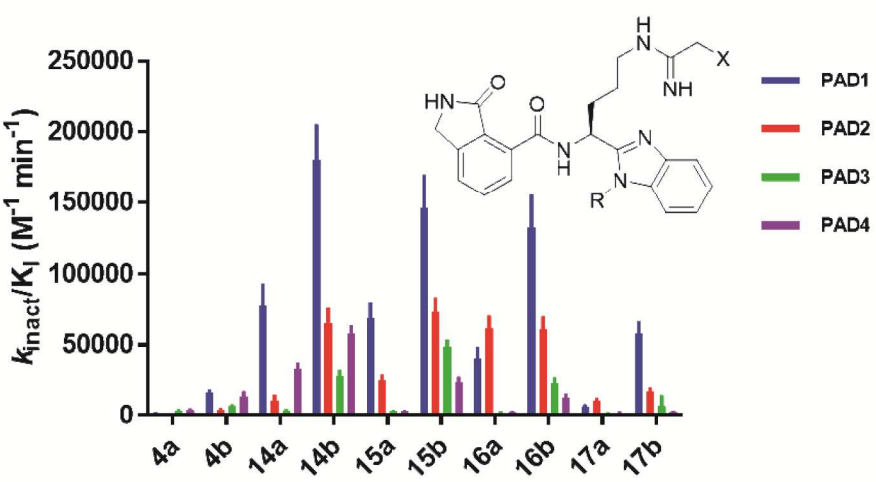
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
9a (R=H, X=F)	11	1.0	4.3	16
9b (R=H, X=Cl)	8.3	1.0	6.2	9.0
10a (R=Me, X=F)	40	15	1.3	1.0
10b (R=Me, X=Cl)	4.3	1.9	1.0	1.0
11a (R=Et, X=F)	39	31	1.7	1.0
11b (R=Et, X=Cl)	4.3	1.6	1.0	1.2
12a (R=^{i}Pr, X=F)	49	54	4.4	1.0
12b (R_1^{i}Pr, X=Cl)	68	38	3.8	1.0

Figure 4. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 1 compounds 9-12. (B) Summary of isozyme selectivity for series 1 compounds 9-12 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

203x260mm (300 x 300 DPI)

A



B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
14a (R=H, X=F)	24	3.3	1.0	10
14b (R=H, X=Cl)	6.5	2.3	1.0	2.1
15a (R=Me, X=F)	25	8.9	1.0	1.0
15b (R=Me, X=Cl)	61	3.1	2.0	1.0
16a (R=Et, X=F)	37	56	1.0	1.7
16b (R=Et, X=Cl)	10	4.8	1.8	1.0
17a (R=^{i}Pr, X=F)	12	20	1.0	2.4
17b (R=^{i}Pr, X=Cl)	27	8.0	3.2	1.0

Figure 5. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 2 compounds 14-17. (B) Summary of isozyme selectivity for series 2 compounds 14-17 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

197x252mm (300 x 300 DPI)

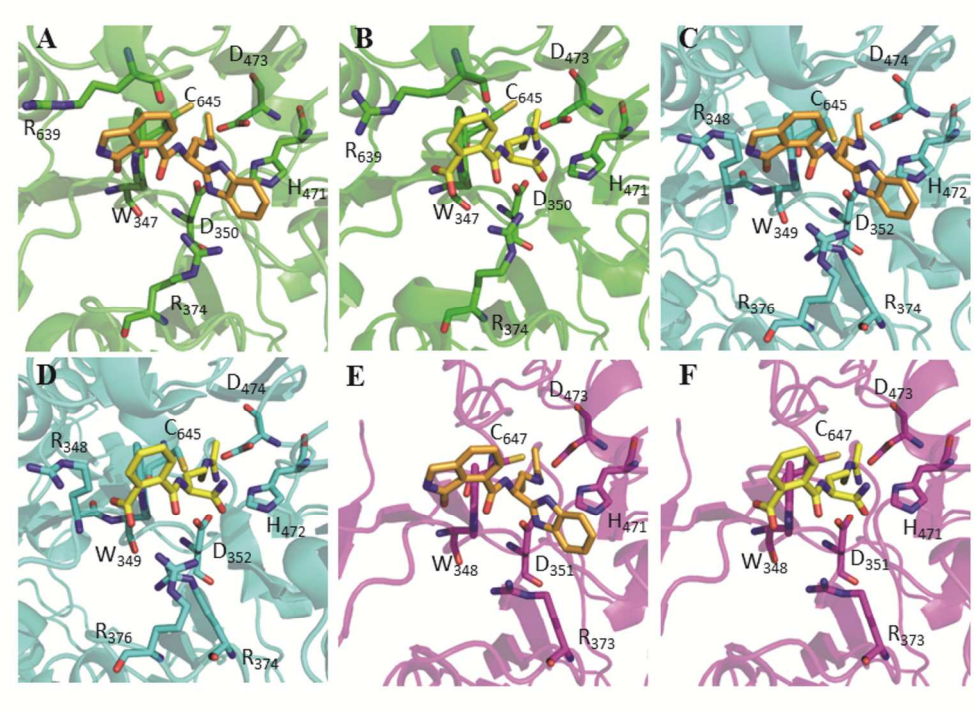
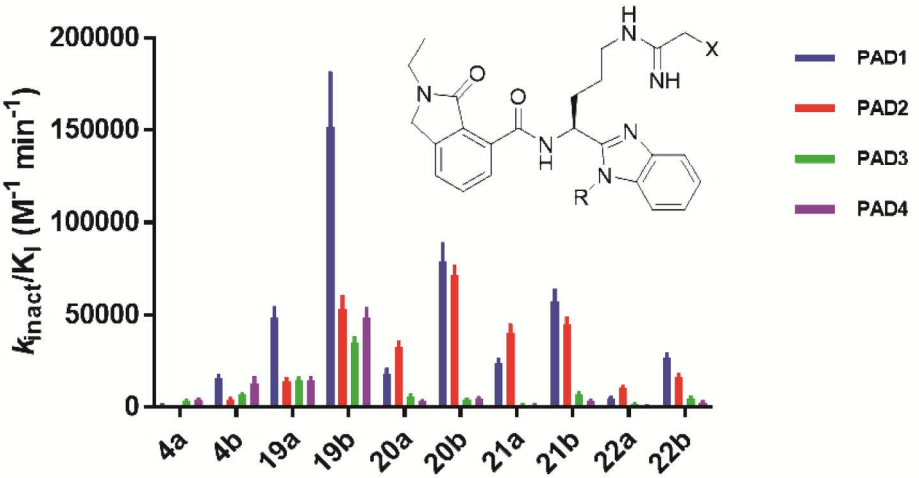


Figure 6. (A) Co-crystal structure of AFM-14a bound to PAD4. (B) Co-crystal structure of o-F-amidine bound to PAD437 (PDB: 3B1T). (C) PAD139 (PDB: 5HP5) overlay with PAD4•AFM-14a co-crystal structure in PyMol. (D) PAD139 (PDB: 5HP5) overlay with PAD4•o-F-amidine co-crystal structure in PyMol. (E) PAD212 (PDB: 4N2C) overlay with PAD4•AFM-14a co-crystal structure in PyMol. (F) PAD212 (PDB: 4N2C) overlay with PAD4•o-F-amidine co-crystal structure in PyMol.

104x75mm (300 x 300 DPI)

A



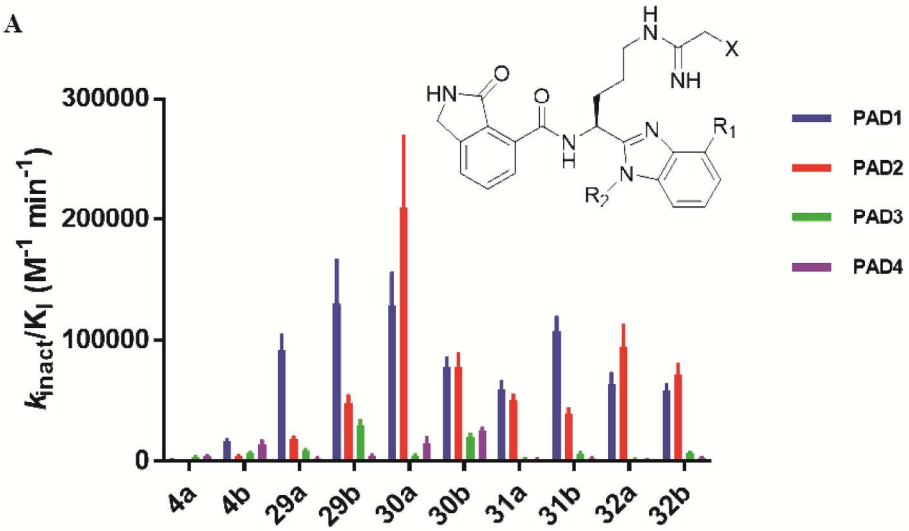
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
19a (R=H, X=F)	3.4	1.0	1.0	1.0
19b (R=H, X=Cl)	4.3	1.5	1.0	1.4
20a (R=Me, X=F)	6.1	11	1.9	1.0
20b (R=Me, X=Cl)	21	19	1.0	1.2
21a (R=Et, X=F)	19	33	1.0	1.0
21b (R=Et, X=Cl)	18	14	2.1	1.0
22a (R=^{i}Pr, X=F)	10	22	3.4	1.0
22b (R=^{i}Pr, X=Cl)	11	6.8	2.0	1.0

Figure 7. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 3 compounds 19-22. (B) Summary of isozyme selectivity for series 3 compounds 19-22 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

189x235mm (300 x 300 DPI)

A



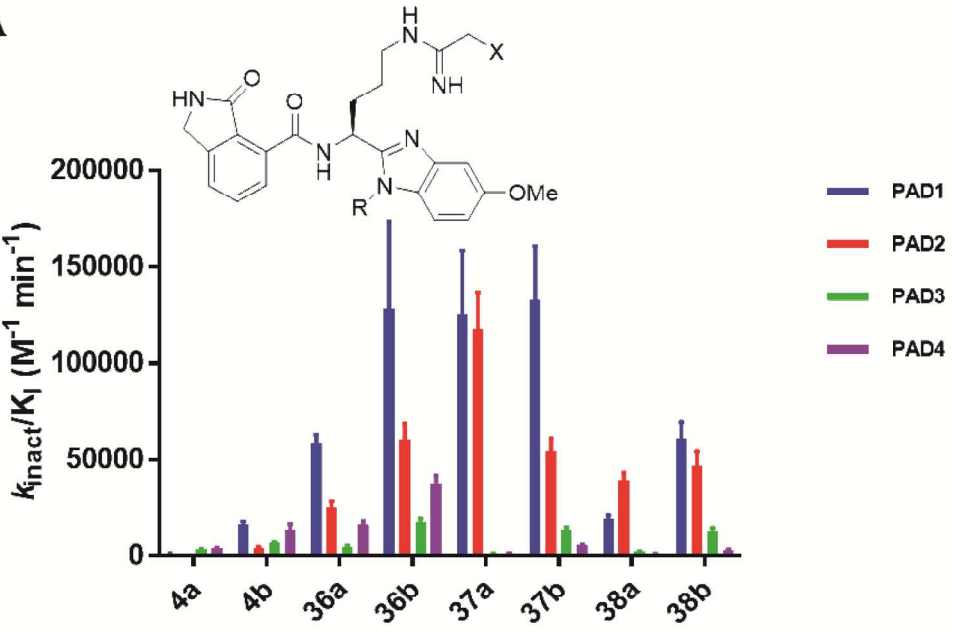
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
29a (R_1=OMe, R_2=H, X=F)	42	8.0	3.8	1.0
29b (R_1=OMe, R_2=H, X=Cl)	30	11	6.9	1.0
30a (R_1=OMe, R_2=Me, X=F)	29	47	1.0	3.2
30b (R_1=OMe, R_2=Me, X=Cl)	3.9	3.9	1.0	1.3
31a (R_1=OMe, R_2=Et, X=F)	38	32	1.0	1.0
31b (R_1=OMe, R_2=Et, X=Cl)	50	19	2.9	1.0
32a (R_1=OEt, R_2=Me, X=F)	65	95	1.2	1.0
32b (R_1=OEt, R_2=Me, X=Cl)	24	30	2.7	1.0

Figure 8. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 4 compounds 29-32. (B) Summary of isozyme selectivity for series 4 compounds 29-32 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

193x242mm (300 x 300 DPI)

A



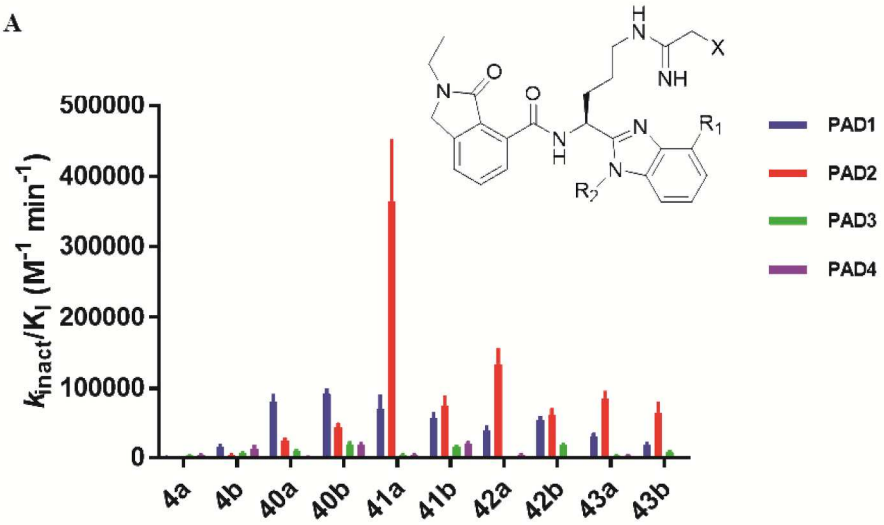
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
36a (R=H, X=F)	13	5.4	1.0	3.5
36b (R=H, X=Cl)	7.3	3.5	1.0	2.2
37a (R=Me, X=F)	121	114	1.0	1.2
37b (R=Me, X=Cl)	25	10	2.4	1.0
38a (R=Et, X=F)	24	50	2.5	1.0
38b (R=Et, X=Cl)	21	16	4.4	1.0

Figure 9. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 5 compounds 36-38. (B) Summary of isozyme selectivity for series 5 compounds 36-38based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

181x218mm (300 x 300 DPI)

A



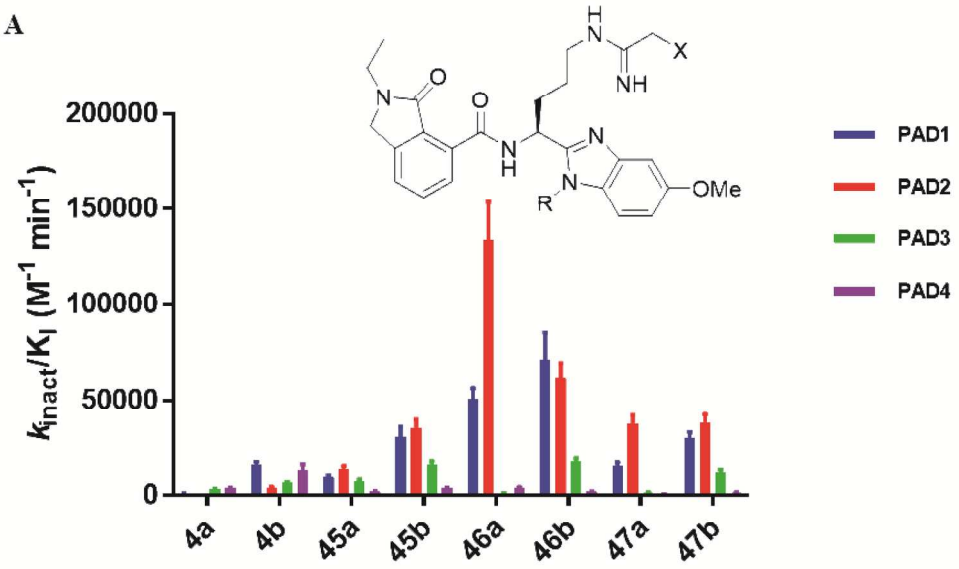
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
40a (R_1=OMe, R_2=H, X=F)	107	32	12	1.0
40b (R_1=OMe, R_2=H, X=Cl)	5.0	2.3	1.0	1.0
41a (R_1=OMe, R_2=Me, X=F)	16	85	1.0	1.0
41b (R_1=OMe, R_2=Me, X=Cl)	3.9	5.0	1.0	1.3
42a (R_1=OMe, R_2=Et, X=F)	28	40	2.0	1.0
42b (R_1=OMe, R_2=Et, X=Cl)	31	27	3.4	1.0
43a (R_1=OEt, R_2=Me, X=F)	8.9	26	1.0	1.0
43b (R_1=OEt, R_2=Me, X=Cl)	12	43	5.5	1.0

Figure 10. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 6 compounds 40-43. (B) Summary of isozyme selectivity for series 6 compounds 40-43 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

199x257mm (300 x 300 DPI)

A



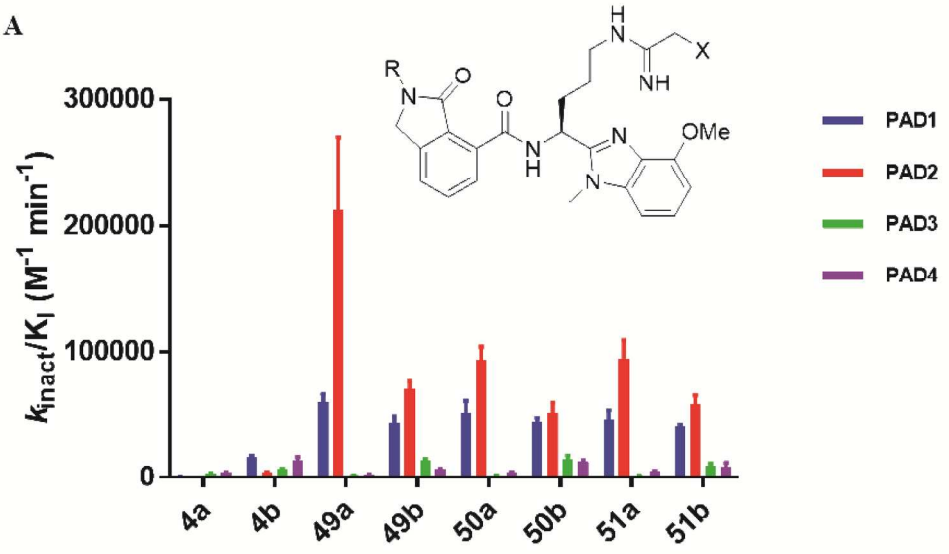
B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
45a (R=H, X=F)	4.4	6.3	3.6	1.0
45b (R=H, X=Cl)	7.8	9.1	4.1	1.0
46a (R=Me, X=F)	47	125	1.0	3.8
46b (R=Me, X=Cl)	37	32	9.3	1.0
47a (R=Et, X=F)	28	70	2.7	1.0
47b (R=Et, X=Cl)	21	27	8.4	1.0

Figure 11. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 7 compounds 45-47. (B) Summary of isozyme selectivity for series 1 compounds 45-47 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

176x202mm (300 x 300 DPI)

A



B

Compound	Fold PAD1 Selectivity	Fold PAD2 Selectivity	Fold PAD3 Selectivity	Fold PAD4 Selectivity
49a (R=Me, X=F)	30	106	1.0	1.1
49b (R=Me, X=Cl)	6.8	11	2.2	1.0
50a (R=^{i}Pr, X=F)	32	58	1.0	2.6
50b (R=^{i}Pr, X=Cl)	3.5	4.1	1.2	1.0
51a (R=cyclopropyl, X=F)	38	77	1.0	3.8
51b (R=cyclopropyl, X=Cl)	4.7	6.7	1.1	1.0

Figure 12. (A) kinact/KI values for BB-F-amidine (4a), BB-Cl-amidine (4b), and series 8 compounds 45-47. (B) Summary of isozyme selectivity for series 1 compounds 45-47 based on their respective kinact/KI ratios setting the enzyme with the lowest value as 1.

185x226mm (300 x 300 DPI)

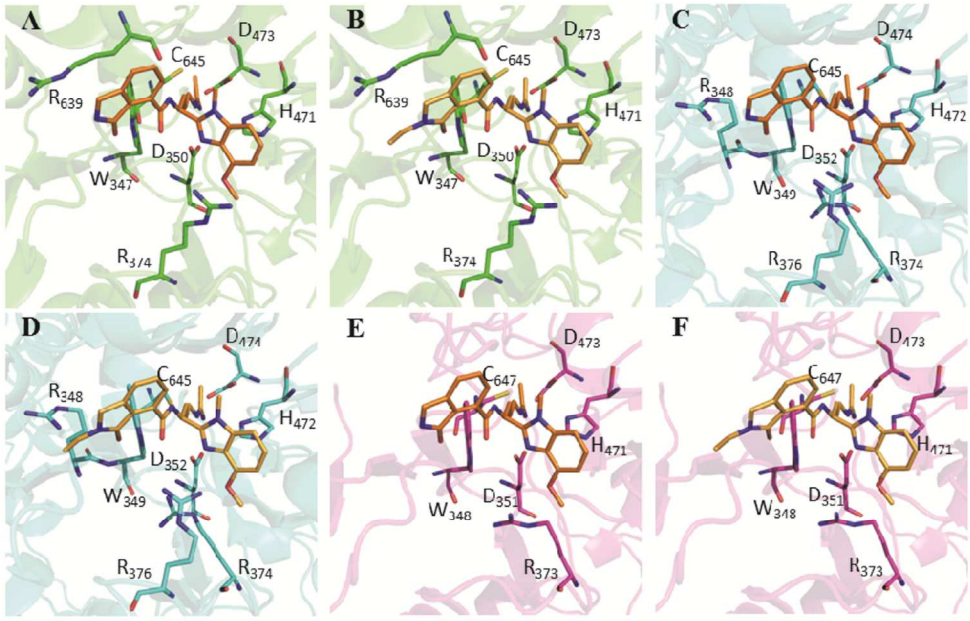


Figure 13. (A) Co-crystal structure of AFM-30a bound to PAD4. (B) Co-crystal structure of AFM-41a bound to PAD4. (C) PAD139 (PDB: 5HP5) overlay with PAD4•AFM-30a co-crystal structure in PyMol. (D) PAD139 (PDB: 5HP5) overlay with PAD4•AFM-41a co-crystal structure in PyMol. (E) PAD212 (PDB: 4N2C) overlay with PAD4•AFM-30a co-crystal structure in PyMol. (F) PAD212 (PDB: 4N2C) overlay with PAD4•AFM-41a co-crystal structure in PyMol.

105x66mm (300 x 300 DPI)

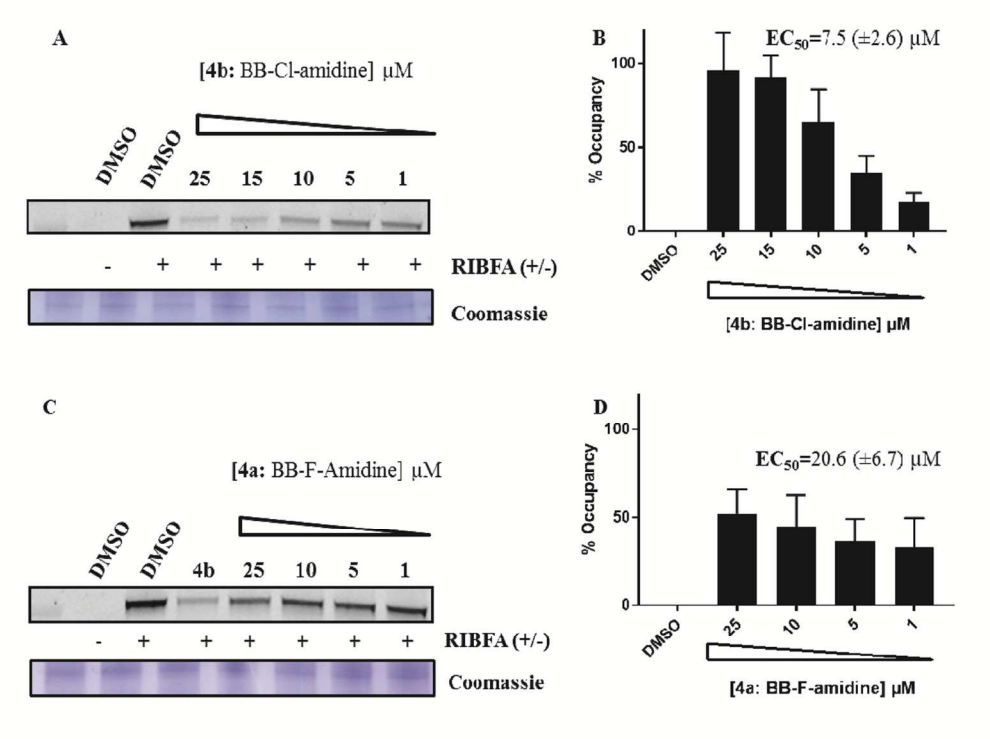


Figure 14. (A) Target engagement assay in HEK293/PAD2 cells for BB-Cl-amidine (4b). (B) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by BB-Cl-amidine (4b). (C) Target engagement assay in HEK293/PAD2 cells for BB-F-amidine (4a). (D) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by BB-F-amidine (4a). Note: Panel C, BB-Cl-amidine (4b) was evaluated at 25 µM.

126x93mm (300 x 300 DPI)

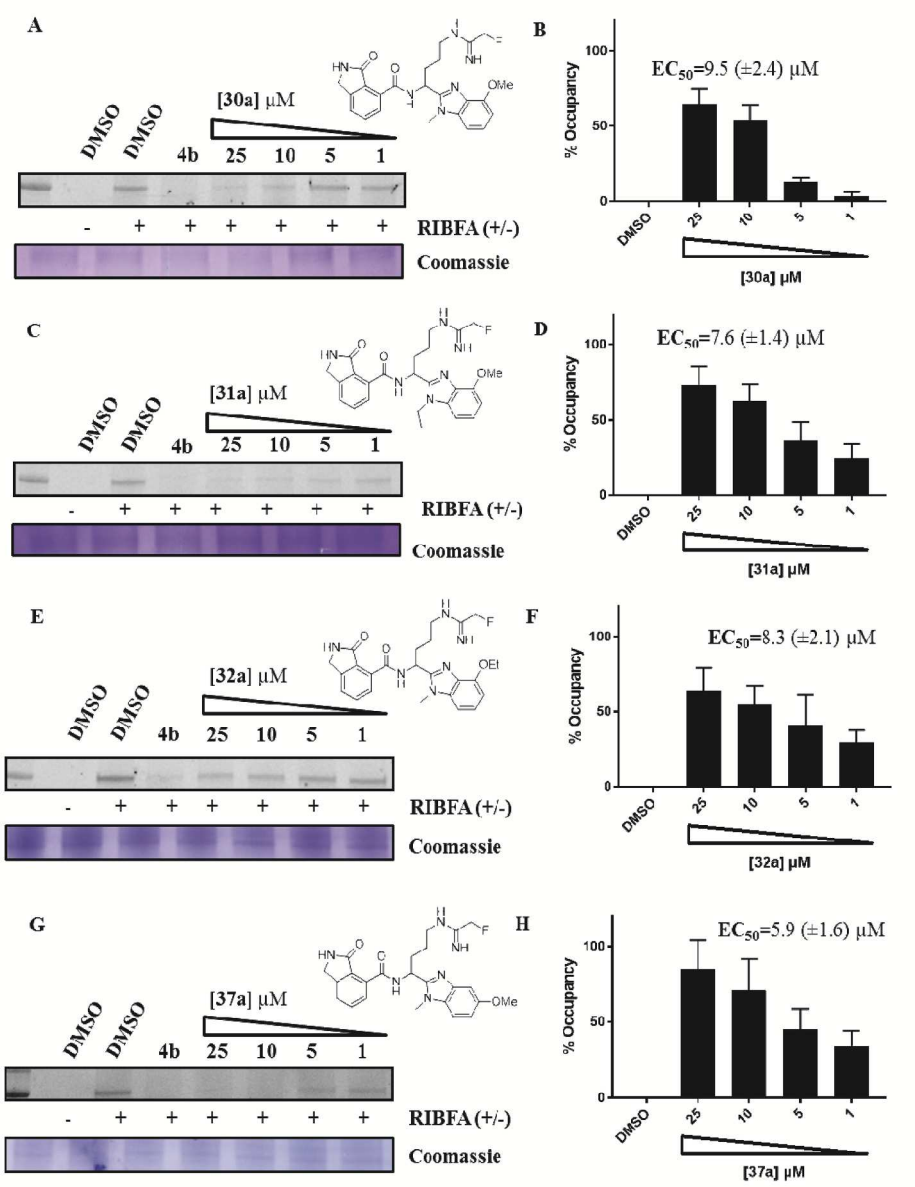


Figure 15. (A) Target engagement assay in HEK293/PAD2 cells for 30a. (B) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by 30a. (C) Target engagement assay in HEK293/PAD2 cells for 31a. (D) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by 31a. (E) Target engagement assay in HEK293/PAD2 cells for 32a. (F) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by 32a. (G) Target engagement assay in HEK293/PAD2 cells for 37a. (H) Quantified results of target engagement assay plotted as percent occupancy of PAD2 by 37a. Note: BB-Cl-amidine (4b) was evaluated at 25 µM.

212x277mm (300 x 300 DPI)

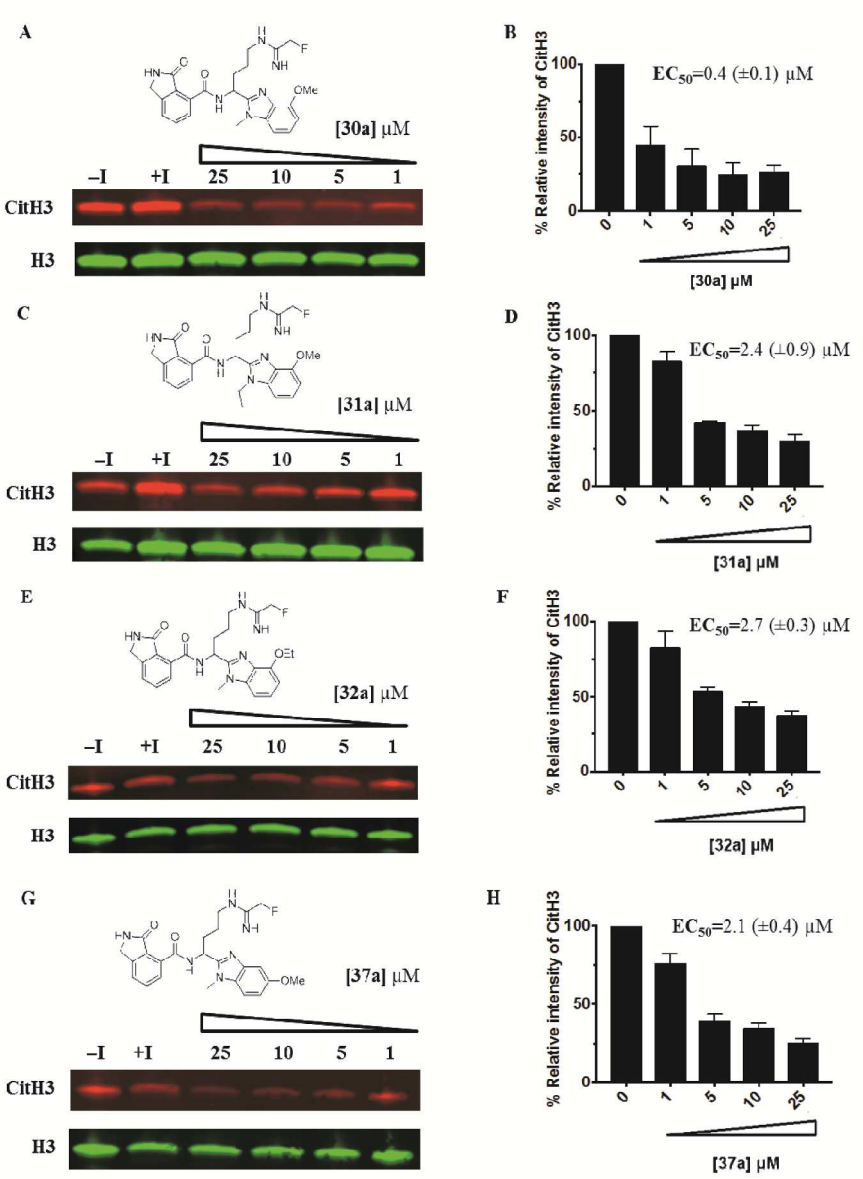


Figure 16. (A) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for 30a. (B) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for 30a. (C) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for 31a. (D) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for 31a. (E) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for 32a. (F) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for 32a. (G) Histone H3 citrullination (CitH3) assay in HEK293/PAD2 cells for 37a. (H) Quantified results of target engagement assay plotted as percent relative intensity of CitH3 for 37a. Note: Lane 1 is HEK293T/PAD2 cells treated with DMSO only, lane 2 is HEK293T/PAD2 cells treated with DMSO and 10 μM ionomycin, and lanes 3-6 show HEK293T/PAD2 cells treated with increasing amounts of compound along with 10 μM ionomycin.

217x299mm (300 x 300 DPI)