Title

Development of antibody-based tools to study PAD4 protein dynamics and cell surface interactions

Author

Kong, Sophie

Publication Date

2024

Peer reviewed|Thesis/dissertation

Antibody engineering to modulate PAD4 activity and investigate cell surface citrullination

by
Sophie Kong

THESIS
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Jim Wells
5F2F4D1A06164C2...

Jim Wells

Chair

DocuSigned by:

Arun Wiita

DocuSigned by:
JULIE ZIKHERMAN
D967968925C2437...

JULIE ZIKHERMAN

Committee Members

This thesis is dedicated to my parents, Jingyi Zhang and Lingmian Kong, for their endless support even thousands of miles away.

Acknowledgements

Without the endless support of family, friends, and mentors, I would not be the person or scientist that I am today. While reflecting on my scientific career up until this point, there are many people that I would like to thank.

First, I want to thank my parents, Jingyi and Lingmian, to which this work is dedicated. My parents continue to support me unconditionally and have always encouraged and pushed me to do whatever makes me happy. They have never missed a call from me, despite many late nights and a 3-hour time difference, and have always filled my carry-on with homemade food whenever I fly back to San Francisco. They knew whenever I need a picture of my pets, Pebbles and Artie, to make me feel better.

Next, I would like to thank my UCSF mentors. To Dr. Jim Wells, who accepted me into his lab right in the thick of the pandemic, and throughout my four years in his lab, never failed to create a boisterous lab environment a filled comradery and scientific creativity. To my committee members, Dr. Arun Wiita and Dr. Julie Zikherman, for asking insightful questions during committee meetings to strengthen my science and drive my project in the right direction. I would also like to thank my rotation mentors and the members of my qualifying exam committee, Danica Fujimori, Charly Craik, Balyn Zaro, Kole Roybal, and Michelle Arkin.

I would also like to thank the various members of the Wells lab that have been instrumental in my path during my PhD. First, Dr. Xin Zhou, for being an amazing mentor and role model that taught me so much of what I know about antibody engineering. To Dr. Kata Pance, Dr. Irene Lui, and Dr. Susanna Elledge, who all showed me what it was like to be a successful graduate student during my early years in the Wells lab. To Dr. Zi Yao, Dr. Fangzhu Zhao, and Dr. Kun Miao, who were always being down to play ping pong during working hours, explore good Chinese food, or trash talk each other's favorite sports teams. To my baymates, Dr. Ines Folger and Kaan Kumru, who never failed to make me laugh and brighten my day even when my science may not have been going so well. I would also like to thank other members of the Wells lab that I have

collaborated with on various projects, Dr. Corleone Delavaris, Dr. Soumya Remesh, Dr. Trent Peters-Clarke and Dr. Kevin Leung. I also need to thank my academic mentors before UCSF, Dr. Michael Krout, Dr. Timothy Strein, and Dr. David Rovnyak. Without my experience performing research in your labs at Bucknell and continued career guidance, I would have never pursued graduate studies.

Outside of the Wells lab, I was fortunate to make several life-long friends through the graduate program. To my first friend on the west coast, Minh Tran, for always pushing me to try new restaurants and different experiences; Zachary Gale-Day for making sure to help me pick up food when my chopsticks skills don't cut it; Michela Caffrey for moving to the bay and inspiring us to enjoy our time outside of lab; Martin Myslinski for endless movie recommendations and being the backbone of the apartment; Nick Young for being my partner in crime that accompanied me through every impulsive decision I've made; Elizabeth Sisko for being the best co-mom to our cat, Rhubarb, and the person that has shared a brain with me for the past five years. I know that making it through my PhD would not have been possible without you all. I would also like to thank my friends back on the east coast, Lauren O'Connor, Julia Burzynski, and Sruti Suryanarayanan, that have continued to stay in touch throughout the years and make the most of the time we spend together over the phone or when I'm home.

Contributions

Chapter 1: Adapted version of the following published manuscript: Zhou, X.*, Kong, S.*, Maker, A., Remesh, S.G., Leung, K.K., Verba, K.A., Wells, J.A. Antibody discovery identifies regulatory mechanisms of protein arginine deiminase 4. *Nature Chemical Biology* (2024). *Denotes equal contribution to this work.

Chapter 2: Unpublished work that has not yet been submitted for publication. Sophie Kong performed engineering and validation of protein arginine deiminase fusion proteins, cellular assays, and mass spectrometry experiments. Trent-Peters Clarke assisted with mass spectrometry methodology and data analysis. The UCSF preclinical therapeutics core (PTC) performed all *in vivo* experiments and Sophie Kong performed downstream analysis of *in vivo* data and harvested serum/cells. James A. Wells oversaw the project.

Chapter 3: Unpublished work that has not yet been submitted for publication. Katarina Pance and Jim Wells contributed to the conceptualization of the project. Arthur Tran performed synthesis of compounds and degradation experiments. Sophie Kong conceptualized and executed degradation experiments. James A. Wells and Ian B. Seiple oversaw the project.

Development of antibody-based tools to study PAD4 protein dynamics and cell surface interactions

Sophie Kong

Abstract

The development of engineered biologics as both tools and therapeutic molecules have transformed the scientific landscape tremendously. Due to their increased specificity, potency, bioavailability and decreased immunogenicity, antibodies have become a popular scaffold for drug development. In 2023, antibodies accounted for five of the top ten selling drugs in the US. In addition to their therapeutic purposes, the high specificity of antibodies allows them to be powerful tools for studying and modifying biological systems. Antibodies are now regularly used to modify cellular signaling pathways, block protein interactions, detect a specific antigen of interest, among many other applications.

Though our body naturally produces antibodies to fight off infection, the ability to generate synthetic antibodies against a chosen antigen of interest gives scientists the opportunity to expand the scope of druggable targets. A library consisting of up to 10^{11} fully synthetic antibodies can be displayed on systems such as phage, yeast, or mammalian cells, then screened for binding against an immobilized antigen. If binders are identified from the pool, they are isolated and sequenced to obtain information about the antibody's amino acid combination for further production and characterization. These antibodies can then be optimized and applied in various formats for downstream biological experiments.

Here, I describe three approaches for using antibody and protein engineering to uncover novel information on biological systems at the protein or cellular level. First in chapter 1, we use naïve antibody discovery to identify functional antibody modulators of protein arginine deiminase (PAD4), a protein highly implicated in the development of autoimmune diseases. In chapter 2, we

develop a PAD4 fusion protein to target the enzyme to HER2+ expressing cancer cells and identify the substrate scope of PAD4 at the cell surface. In the last chapter, we create an antibody-drug conjugate using site-specific bioconjugation for targeted degradation of challenging cell surface proteins.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A2AR – Adenosine receptor 2A

ACPA – Anti-citrullinated protein antibody

ACK – Ammonium-chloride-potassium

ADC – Antibody drug conjugate

ADC-Tac – Antibody drug conjugate PROTAC

BCA – Bicinchoninic acid assay

BLI – Biolayer interferometry

BSA – Buried surface area OR Bovine serum albumin

CDR – Complementarity determining region

DMEM - Dulbecco's Modified Eagle Medium

DNA – Deoxyribonucleic acid

DSF – Differential scanning fluorimetry

EDTA - Ethylenediaminetetraacetic acid

ELISA – Enzyme-linked immunosorbent assay

ETD – Electron-transfer dissociation

GPCR – G-protein coupled receptor

HCD – Higher-energy collisional dissociation

HER2 – Human epidermal growth factor receptor 2

HRP – Horseradish Peroxidase

IgG – Immunoglobulin G

K_D – Binding affinity

k_{off} – Dissociation constant

LOF – Loss of function

LYTAC – Lysosomal targeting chimera

Ni-NTA – Nickel-nitrilotriacetic acid

NSEM – Negative stain electron microscopy

PAD4 – Protein arginine deiminase 4

PBS – Phosphate buffer saline

PDB – Protein data bank

PROTAC – Proteolysis targeting chimera

PTM – Post-translational modification

PVDF - Polyvinylidene fluoride

RA – Rheumatoid arthritis

RAbD – Rosetta antibody design

RIPA – Radioimmunoprecipitation assay

SDS-PAGE – Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

SEC – Size exclusion chromatography

TB – Terrific broth

TBS – Tris buffer saline

TCEP - Tris(2-carboxyethyl)phosphine

TIL – Tumor infiltrating lymphocytes

Chapter 1

Antibody discovery identifies regulatory mechanisms of protein arginine deiminase 4

Abstract

Unlocking the potential of protein arginine deiminase 4 (PAD4) as a drug target for rheumatoid arthritis requires a deeper understanding of its regulation. In this study, we use unbiased antibody selections to identify functional antibodies capable of either activating or inhibiting PAD4 activity. Through cryo-electron microscopy, we characterized the structures of these antibodies in complex with PAD4 and revealed insights on their mechanisms of action. Remarkably, rather than steric occlusion of the substrate-binding catalytic pocket, the antibodies modulate PAD4 activity through interactions with allosteric binding sites adjacent to the catalytic pocket. These binding events lead to either alteration of the active site conformation or the enzyme oligomeric state, resulting in modulation of PAD4 activity. We then carried out extensive computational and library-based strategies to improve the affinity, solubility, and biophysical properties of our lead inhibitory antibody, hI365. Our study uses antibody engineering to reveal new mechanisms for enzyme regulation and highlights the potential of using PAD4 agonist and antagonist antibodies for studying PAD4-dependency in disease models and future therapeutic development.

Introduction

Protein arginine deiminase 4 (PAD4) is a calcium-dependent enzyme that catalyzes hydrolysis of peptidyl arginine sidechains to citrulline in proteins.[1,2,3] In neutrophils, PAD4-catalyzed chromatin decondensation plays a crucial yet ambiguous role in inducing an inflammatory form of neutrophil cell death called NETosis. Multiple reports have shown a correlation between PAD4 expression and histone citrullination leading to NETosis, but the link between citrullination and NET formation is still largely complicated. [4,5,6,7] During this process, PAD4 and other intracellular contents are released to the extracellular space where PAD4 can create citrullinated neoepitopes. These neoepitopes are recognized by the immune system and further trigger inflammatory diseases such as rheumatoid arthritis (RA).[8,9] Small-molecule PAD4 inhibitors were developed and found to be effective in alleviating RA phenotypes in mouse models, indicating the relevance of PAD4 in RA pathology.[10,11] While small-molecule PAD4 inhibitors show promise in preclinical studies of targeting RA, they often lack specificity and potency due to direct targeting of the enzyme's active site which is highly conserved across PAD isoforms.[12,13,14]

Antibodies are powerful tools for capturing the dynamic nature of proteins by binding to both active sites and allosteric sites of enzymes and interrogating different protein conformations.[15] For example, deliberately selecting for antibodies to natural on- or off-states in proteins can be achieved by first “trapping” those conformations and then using differential *in vitro* antibody selection to isolate state-specific binders.[16,17,18] Instead of trapping the enzymatic state prior to selection, we hypothesized that using unbiased antibody selections on a native enzyme sampling various conformations in solution will allow us to identify novel conformations that activate or inhibit the enzyme. PAD4 is ideally suited for such an approach to better understand the mechanisms and protein states associated with activation and inhibition of the enzyme.

Moreover, antibodies are an ideal modality for targeting extracellular PAD4 since, unlike small molecules, they do not penetrate cells and would not interfere with normal homeostatic

activities of intracellular PAD enzymes. Structural studies show each PAD4 binds five Ca^{2+} ions in two different pockets to become fully activated, and calcium-bound PAD4 is 10,000x as active as apo-PAD4.[19,20] PAD4's calcium-dependent activation may be relevant when released into extracellular milieu where calcium concentration reaches millimolar (mM) levels compared to micromolar (μM). However, it's extracellular regulation remains ambiguous given the high oxidizing environment outside of cells (https://www.sigmaaldrich.com/US/en/technical-documents/technical-article/cell-culture-and-cell-culture-analysis/mammalian-cell-culture/calcium-in-cell-culture). Also contributing to PAD4 extracellular regulation is the presence of autoantibodies that bind and further activate PAD4, which are found in 20-45% of RA patients. The presence of these autoantibodies is typically associated with a more severe and erosive disease phenotype.[21,22,23,24,25,26] The binding epitopes of several patient-derived anti-PAD4 autoantibodies have been characterized, but their mechanisms of action remain unclear.[25,26] The existence of endogenous anti-PAD4 antibodies lead us to believe that we can use *in vitro* methods to identify anti-PAD4 antibodies also capable of modulating extracellular PAD4 activity.

Here, we utilized an unbiased antibody selection strategy, coupled with functional screening and Cryo-EM structural analysis, to identify new conformations and mechanisms for inhibiting or activating both murine and human PAD4 in the presence of high Ca^{2+}. We discovered that PAD4 activity can be enhanced through antibody binding to an interface loop which promotes PAD4 dimerization while reducing disorder in the substrate-binding loop. We also discovered an inhibitory antibody that binds and re-structures a helix in the Ca^{+2} binding pocket that mediates a conformational change in the active site, preventing calcium ion and substrate binding. These engineered antibodies form a versatile toolkit for studying PAD4-dependence of disease states in both mouse models and patient samples. Through structural analysis of the antibody-PAD4 complexes, we shed light on previously unknown PAD4 regulatory mechanisms, providing new opportunities for pharmacological targeting of the enzyme.

Results

Identifying PAD4 antibodies by phage display

As a first step toward antibody generation, it was critical to produce highly purified native forms of PAD4 suitable for *in vitro* phage antibody selections. N-terminal biotinylated forms of human PAD4 (hPAD4) and mouse PAD4 (mPAD4) were constructed and expressed in BirA-expressing *E. coli* for facile one-step purification and immobilization on streptavidin magnetic beads for phage selection (**Fig. 1.1A**). Biotinylation was validated by a gel shift assay (**Fig. 1.1B**). Because PAD4 has free cysteines on the surface, 0.5 mM of Tris(2-carboxyethyl)phosphine (TCEP) was supplemented to the protein purification and phage selection buffers to prevent protein aggregation and loss of function (**Fig. 1.1C**). As antibodies are held together by a series of disulfide bonds, we confirmed that the Fab scaffold used in this phage library does not lose the antigen-binding capability even in the presence of the TCEP needed to stabilize PAD4 (**Fig. 1.1D,E**).

We carried out two phage selection campaigns with a synthetic antibody Fab-phage library either in the presence of 10 mM free calcium or EDTA to sample the active and inactive states of PAD4 as PAD4's activity is dependent on calcium (**Fig. 1.1F, Fig. 1.2A**).[1] We observed that the thermostability of the Ca^{2+} bound state is significantly higher than the apo-state, shown by a 20°C upward shift in the melting temperature, demonstrating the enzyme exhibits two distinct conformational states with or without calcium (**Fig. 1.1G,H**). Four rounds of selections were performed with either 10 mM Ca^{2+} or 1 mM EDTA, respectively. To increase the stringency of selection, we systematically decreased the PAD4 concentration in every subsequent round of selection. At the end of the fourth round of selection, 95 individual clones were screened for binding by Fab-phage enzyme-linked immunosorbent assay (ELISAs). The top 10 binders from both the Ca^{2+} and EDTA selections were sequenced and expressed. We then performed a second hPAD4 selection to identify PAD4 activators and inhibitors that targeted different epitopes than the ones discovered in our first selection campaign. To do this, we added hI281 in excess during

selection to block this previously identified epitope. This strategy allowed us to discover novel binders and create a toolkit of diverse PAD4 antibody modulators. The top binding clones from our second selection Fab-phage ELISAs were sequenced, affording 16 new anti-hPAD4 clones.

We were also interested in identifying antibodies against mouse PAD4, as these binders can be used as tools for studying the role of PAD4 in mouse models of RA and mimic the development of anti-PAD4 autoantibodies in RA patients.[27] We carried out an additional phage selection campaign against mPAD4 and after four rounds of selection, 12 antibodies that exclusively bind to mPAD4 and 16 antibodies that were hPAD4/mPAD4 cross-reactive were identified (**Fig. 1.3A**).

Characterization of PAD4 antibodies

We utilized two established PAD4 activity assays for initial characterization of the functional consequences of antibody binding. The first is an end-point immunoblot assay that detects citrullination of a large natural protein substrate, histone H3. The second is a spectrophotometric assay utilizing a small molecule trypsin-fluorogenic substrate pair in which the substrate citrullination by PAD4 prevents hydrolysis by trypsin, so that a reduced fluorescence readout is proportional to higher PAD4 activity (**Fig. 1.2B,C**).[28] As expected, PAD4 activity significantly increases with higher concentrations of Ca^{2+} as measured by each assay. We used the more rapid and higher-throughput trypsin/fluorogenic peptide assay to screen the function of antibodies identified through our phage selection strategy.

The naming of our antibodies was as follows: the first letter (h or m) stands for human or mouse PAD4 to which the antibody binds. If the antibody is cross-reactive to human and mouse, we designate it as hm. The second letter (I, A or N) means inhibitory, activating, or neutral, respectively. The letters are followed by the plasmid ID number. We found inhibitory, activating, and neutral binders towards both hPAD4 and mPAD4 and summarize their binding and functional properties in **Table 1.1**.

We identified five antibodies to hPAD4 (hI281, hA288, hA362, hI364 and hI365) which showed either antagonistic or agonistic effects (**Fig. 1.2C**). hI281 and hA288 were identified as the strongest inhibitors and activators from the first selection, respectively, and these binders were then used in a second selection against hPAD4 to mask their epitopes for explore others. This epitope blocking strategy allowed us to identify activating antibody hA362, and inhibitory antibodies hI364, and hI365 that bind to different epitopes than the masking hI281 and hA288 (**Fig. 1.2C**). Our PAD4 modulating binders from the second selection also inhibited or activated PAD4-mediated histone H3 citrullination, indicating that these PAD4-binders can modulate citrullination of protein substrates as well as small synthetic substrates (**Fig. 1.2D,E**).

Interestingly, some of our most potent inhibitors (hI364, hI365) exhibited selective binding to the Ca^{2+}-bound state of PAD4 and showed either weaker or no binding to the apo-enzyme as measured by biolayer interferometry (BLI) in the presence or absence of Ca^{2+} (**Fig. 1.3B)**. These results suggest that there are epitopes formed in only the Ca^{2+} bound form of PAD4 that are critical to enzyme activity. We believe that our calcium-dependent antibodies are targeting these critical epitopes, thus inhibiting activity.

Our top binders were also tested for cross-reactivity between mPAD4 and hPAD4 as this would be a useful feature for studying PAD4 between mouse and human. However, most binders identified from single species selections lacked cross-reactivity when terminally evolved for five rounds. This may be attributed to the fact that hPAD4 and mPAD4 share only 73% sequence homology and contain notable amino acid differences near important structural regions. For example, the N-terminal domain calcium-binding pocket contains several significant mutations from human to mouse PAD4 (D157E, E170K, and D388N). We hypothesize that these changes may prevent the identification of cross-reactive binders to this important of region of enzyme regulation. In a further effort to identify pan-binding clones we took the eluted phage pool from round three of selection on mPAD4, and continued to select an additional two rounds on hPAD4.

This selection on dual antigens significantly enriched for cross-reactive binders and we found one clone (hmI400) to inhibit both hPAD4 and mPAD4 (**Fig. 1.4A**).

Binding of antibodies to human and mouse PAD4

We characterized the binding affinity of several inhibiting and activating antibodies to human or mouse PAD4 using BLI. The binding affinity of each antibody was investigated under varying concentrations of calcium (0, 2, 10 mM). Of our inhibitory lead clones, hI364 and hI365 both exhibited no binding to PAD4 in calcium-free conditions but showed high affinity to PAD4 in 2 and 10 mM Ca^{2+} (K_D for hI364 = 0.64 nM; K_D for hI365 = 1.2 nM). The binding of other clones (hI281, hA362) remained unaffected by calcium concentration (**Fig. 1.4B**).

We similarly characterized the binding of our mPAD4 binders and identified mA342 to be a potent activator of mPAD4. Interestingly, this clone also exhibited calcium selectivity as its K_D improves from 200 uM to < 1 nM upon addition of 10 mM calcium (**Fig. 1.4B**). However, mA342 tended to aggregate in solution and suffered from poor biophysical properties characterized by a low intensity, widely dispersed peak on SEC (**Fig. 1.4C**). We hypothesized that the poor solubility was due to several hydrophobic residues in the complementarity determining regions (CDRs). To test this, we performed an alanine scan of the CDRs to identify hotspots of hydrophobicity to improve the solubility of mA342.

The alanine scan identified four hydrophobic CDR residues that were dispensable as they did not impact binding of mA342 to mPAD4. These four residues (light chain I97; heavy chain Y101, Y111, M114) were mutated to alanine to generate mA342-c3, which showed a clean, monodispersed peak on SEC, and improved binding and activation of mPAD4 measured by BLI (**Fig. 1.4C,D**). We believe that mA342-c3 could be a promising candidate for mouse models mimicking the pathology of activating anti-PAD4 autoantibodies.

Several antibodies modulate hPAD4 enzyme dimerization

We utilized negative stain EM (NSEM) to screen the overall shapes of PAD4 and the PAD4/antibody complexes which are schematically represented in **Fig. 1.5A**. As expected from X-ray structures of PAD4, the enzyme formed a C-shaped monomeric species and a globular, dimeric species of hPAD4 (**Fig. 1.5B**). Previous reports have shown that PAD4 exists as both low activity monomers and highly active dimeric states.[29] We next incubated hPAD4 with our functional antibody Fab fragments in equimolar amounts and characterized the stoichiometry of the protein complexes by NSEM. The antibody Fabs showed the typical donut-shaped structures in the EM micrographs. Interestingly, the 1:1 complex of hPAD4 to antibody was seen for hPAD4 in complex with the inhibitory antibody, hI281, while a 2:2 complex of hPAD4 to antibody was observed for hPAD4 in complex with both activating antibodies, hA288 and hA362 (**Fig. 1.5B**). Taken together, these data suggest that the inhibitor (hI281) may be blocking PAD4 dimerization as the antibody is only seen interacting with monomeric PAD4, while activators (hA288 and hA362) are promoting PAD4 dimerization and promoting activity.

We confirmed the stoichiometries for 1:1 inhibitor/PAD4 (120 kDa) and 2:2 activator/PAD4 (240 kDa) complexes by analytical size exclusion chromatography (SEC) (**Fig 1.5C**). The free hPAD4 shows a major dimeric peak and a minor monomeric peak on SEC reflecting its dimer to monomer equilibrium. The hPAD4/hA288 complex runs as a homogeneous species at its expected MWs of a 2:2 complex, whereas the hPAD4/hI281 complex runs at the MW of a 1:1 complex. These data confirm that hI281 stabilizes the hPAD4 monomer, forming a 1:1 complex with hPAD4, while hA288 promotes hPAD4 dimerization to a 2:2 complex.

PAD4 dimerization influences activity and antibody function

To further study how dimerization influences enzyme activity, we applied computation-guided mutational analysis to disrupt and test the PAD4 dimerization interface. Briefly, after relaxing and minimizing the structures in Rosetta (PDB: 1WDA), we used PyMOL to select

interacting residues within 5 Å from the two subunits. The interface energy function in Rosetta was applied to calculate an interaction score of different residues. Residues R8, Y435, F541, and W548 stood out as the top four residues contributing to interface association with large negative interaction scores (**Fig. 1.6A**). We chose R8 and Y435 located at the two ends of the interface for a mutational analysis (**Fig. 1.6B**). The interface energies of single mutant R8E, Y435A and double mutant R8E/Y435A were calculated to be significantly higher than the wildtype, indicating these mutations would be effective in breaking dimerization. R8 and Y435 were also reported in a previous mutational study to influence PAD4 dimerization constant.[29] Additionally, PAD1 is the only PAD isoform that exists as a monomer in solution. This finding is likely explained by an R8Q substitution in the N-terminal domain of PAD1, again supporting the importance of R8 in enzyme dimerization.[30]

Based on these analyses, we cloned and expressed the R8E, Y435A and double mutant R8E/Y435A of PAD4. We also expressed two additional mutants as negative controls, N438A and N438R, as residue N438 is predicted to minimally contribute to the interface energy. Indeed, we found that R8E, Y435A, and R8E/Y435A run as monomers on SEC while the WT, N438A, and N438R form dimers (**Fig. 1.6C**). All monomeric mutants were virtually inactive in the peptide citrullination assay while N438A and N438R had comparable activities to WT hPAD4 (**Fig. 1.5D**). These data confirmed that dimerization is important for PAD4 activity.[19,29] We also found the monomerizing mutations slightly destabilized the enzyme, reducing the melting temperature of apo-PAD4 from 45°C to 40-42°C, and Ca^{2+}-bound PAD4 from 65°C to 64°C (**Fig. 1.6D,E**).

We next studied how WT and mutant PAD4s bind to the activating antibody hA362 that promotes enzyme dimerization. The binding affinity of hA362 was decreased by about 3-, 10-, and 160-fold for R8E, Y435A and the R8E/Y435 double mutant, respectively; their ability to activate PAD4 decreased in rough proportion to their reduction in binding affinity. While the Y435A mutation directly affects the hA362 binding epitope, the R8E mutation alters a series of electrostatic interactions that also exist near the epitope, thus explaining the observed reduction

in binding affinity and antibody functionality. In contrast, the null mutants, N438A and N438R, did not significantly impact binding (**Fig. 1.5E, Fig. 1.6F**). These results further support that the binding and activation efficacy of hA362 is influenced by dimerization of PAD4.

hA362 activates hPAD4 by stabilizing dimerization

To further understand the molecular mechanism of hA362 activation of PAD4 we obtained the cryo-EM structure of hPAD4 in complex with hA362 at 3.5 Å resolution. The overall structure is a 2:2 complex, containing two copies of hA362 bound to a homodimeric PAD4 (**Fig. 1.7**). Consistent with previous structures of the enzyme alone, PAD4 forms an anti-parallel head-to-tail homodimer in C2 symmetry.[31] Each PAD4 monomer adopts an elongated fold with a N-terminal domain in immunoglobulin-like structure and a C-terminal domain in an α/β propeller structure, and it binds to a total of five Ca^{2+} ions in two pockets (**Fig. 1.8A,B**).

Although each hA362 Fab predominantly interacts with the N-terminal domain of each PAD4 (Buried Surface Area (BSA): 1196.3 $Å^2$), the Fab HC spans across the PAD4 monomer to make an additional interaction with the C-terminal domain of a second PAD4, close to the substrate binding site (BSA: 183.9 $Å^2$) (**Fig. 1.7B**). This readily explains why the binding affinity of hA362 is lower to the monomeric PAD4 mutants as monomerization eliminates a portion of the binding epitope. Importantly, this additional contact with the PAD4 in trans is with the interface loop (the I-loop, shown in red), a known regulatory motif (**Fig. 1.7C,D, Fig. 1.8C**). The specific interaction of hA362 with the I-loop mechanistically explains why PAD4 activity is enhanced, as the substrate binding loop (S-loop, shown in purple) neighbors the I-loop. The I-loop engages the S-loop through an electrostatic interaction between R441 and D465 and this interaction is important for catalytic activity.[31] Consistently, previous molecular dynamics simulations have found that the I-loop in a monomeric PAD4 W548 mutants exhibits high flexibility, accounting for the increased flexibility of the S loop and low activity of monomeric PAD4.[31] Our structural results further show that the I-loop and the associated S-loop can be stabilized through antibody binding.

Our data suggest that hA362 activates PAD4 by stabilizing the dimer and by directly facilitating organization of the active site.

hPAD4 inhibition by re-structuring binding pockets

Fabs hI364 and hI365 are two of the most potent inhibitors we identified. Both Fabs only recognize the Ca^{2+}-bound form of PAD4. We determined the cryo-EM structure of hPAD4 in complex with hI365 at 3.2 Å resolution to understand the antibody calcium dependency and inhibitory mechanism (**Fig. 1.9B**). Similar to hA362, hPAD4 adopts a homodimeric structure and forms a 2:2 complex with hI365 (**Fig. 1.9A,B**), but here each antibody only directly interacts with one PAD4 monomer (BSA: 1122.9 $Å^2$). The CDRH1 and H2 of hI365 binds to a region that is structured only upon binding of three Ca^{2+} ions in the N-terminal domain, explaining why the binding of hI365 is Ca^{2+}-dependent (**Fig. 1.9B**).

Interestingly, each hPAD4 monomer in the hI365/hPAD4 complex only contains three of the five Ca^{2+} ions that are typically bound in PAD4 structures and the hA362/PAD4 complex (**Fig. 1.8D**). A closer look of the structure reveals that binding of hI365 to hPAD4 leads to the occlusion of both Ca^{2+} ions in C-terminal domain while the ions in the N-terminal domain remain bound (**Fig 1.9C, Fig. 1.8E**). Additionally, the Y106 residue in CDRH3 loop of hI365 interacts with Trp347 in hPAD4 (**Fig. 1.9C, pink**), pulling the 340-352 loop out of the substrate binding site and disrupting a small helix usually formed in the active structures (residues 374-383).[32] In the original hPAD4 structure (PDB: 1WDA) with substrate benzoyl-L-arginine amide (BAA), the 340-352 loop contains key residues for interactions with both the C-terminal Ca^{2+} ions and the hPAD4 substrate.[32] The 347-350 region also stabilizes R374, which binds substrate, along with the surrounding Ca^{2+} binding region (purple). R374 and the nearby helical region are also no longer resolved with hI365 binding (**Fig 1.9C, right panel, Fig. 1.8E**). Therefore, structural alteration of these residues diminishes the ability of the protein to bind calcium in its C-terminal domain and the substrate in this pocket. Specifically, the negatively charged D350 residue is known to stabilize binding of the

arginine substrate, and mutation of D350 results in loss of enzyme activity.[32] Our structure shows that binding of hI365 to PAD4 results in D350 flipping away from the substrate pocket, thus leading to enzyme inhibition (**Fig. 1.9C**). Additionally, hI365 forms many hydrophobic interactions with one predominant PAD4 monomer though loop H3 also interacts with the second PAD4 monomer (**Fig. 1.9D**). These results reveal another mechanism for blocking PAD4 activity, namely pulling out loop 340-352 and preventing Ca^{2+} and substrate from binding.

Optimized hI365 shows potent inhibition of hPAD4 in vitro

While hI365 exhibited inhibitory properties against PAD4 in preliminary activity assays, the antibody performed inconsistently when attempting to obtain an IC_{50}. In addition, we observed late elution of the antibodies on SEC, suggesting presence of sticky, hydrophobic residues that may have contributed to assay variability by causing antibody aggregation (**Fig. 1.10B**). We thus used structural information to inform in silico and experimental methods to improve the solubility of hI365. First, using a Rosetta-based pairwise interaction analysis, we identified a G58D mutation that improved antibody solubility and SEC behavior (**Fig. 1.11A-C, Supplementary Note 2**). Next we used Rosetta Antibody Design (RAbD) algorithm to engineer light chain CDR loop 3 (L3). As seen in the cyro-EM structure, L3 is present at the binding interface but does not currently form any appreciable contacts to PAD4 (**Fig. 1.10A, Fig. 1.11D**). We predicted that optimizing the length of L3, which is composed of 9 amino acids in parental hI365, could promote additional contacts to PAD4. Computational screening of various randomized L3 grafts showed that the optimal length of L3 is 9 to 10 amino acids (**Fig. 1.11E-H**). However, though our L3 mutants bound PAD4 with similar affinity to WT hI365, all clones identified from our L3 engineering exhibited poor SEC profiles and did not approve the ability of the antibody to inhibit PAD4 (**Fig. 1.11G,H)**.

Lastly, a library-based randomization approach based on the knowledge gained by structure-informed *in silico* designs was carried out to simultaneously optimize binding and biophysical properties. We created two libraries. In Library 1, all CDR H3 residues were targeted

for soft randomization using a 70-10-10-10 formula.[33] In Library 2, an RVK degenerate codon strategy was employed at certain positions to bias mutagenesis to more hydrophilic residues (**Fig. 1.10C**). In both libraries, we incorporated information gained from our rAbD engineering of CDRL3 and randomized this loop to contain both 9 and 10 amino acids. After four rounds of panning, we identified six unique clones that improved binding to hPAD4 compared to the parent hI365 (**Fig. 1.12A-D**). Binding was initially measured via ELISA at 20 nM and 5 nM hPAD4 antigen, and results show that library-based clones outperformed the parental hI365, especially at 5 nM hPAD4 (**Fig. 1.12C**). To estimate antibody off-rates, we pre-formed an Ab/hPAD4 complex before adding the mixture to hPAD4-coated ELISA plates. While hI365 dissociated from the initial Ab/PAD4 complex and readily bound plate-bound hPAD4 antigen, the newly identified library clones remained associated to hPAD4 in the pre-formed complex, suggesting an improvement in affinity (**Fig. 1.12D**).

In particular, clones E3 and E6 showed significantly improved SEC profiles, binding kinetics measured via BLI, and a marked increase in PAD4 inhibition compared to hI365 (IC_{50} values of 94 nM and 13 nM, respectively; **Fig. 1.10D-F, Fig. 1.12E**). E3 and E6 both incorporate the G58D mutation previously shown to improve antibody solubility (**Fig. 1.10D**). In agreement with our computational CDR L3 studies, parental hI365 and E3 have 9 amino acid L3 loop length, while clone E6 has a 10 amino acid long loop length. Thus, the affinity and the developability profiles of these antibodies were successfully improved by structure-informed antibody optimization, making them more suitable probes of PAD4 function.

Antibody-PAD4 specificity and broader substrate profiling

While current small molecule inhibitors show poor specificity for PAD4 over other PAD isoforms, we believe that our antibodies may offer the advantage of improved specificity that is commonly noted with biologics. We tested binding of our functional human antibodies (hI281,

hA288, hA362, hI364, hI365, E3, and E6) against PAD4 alongside PAD2 and PAD3 via BLI, and all antibodies exhibited binding to PAD4 exclusively (**Fig. 1.13A**).

In addition to developing PAD4 specific binders, we wanted to ensure that our antibodies inhibited PAD4 in biologically relevant environments. As extracellular PAD4 activity is hypothesized to be pathogenic in RA, we performed a PAD4 activity assay against whole cell lysate containing cytosolic, nuclear, and membrane-bound proteins. While PAD4 alone citrullinated a variety of proteins present in whole cell lysate, our functional antibodies retained the ability to inhibit PAD4 (**Fig. 1.13B**).

Discussion

New therapeutic mechanisms for autoimmune diseases like RA are in significant demand as current approved treatments are limited in scope or susceptible to broad-spectrum resistance.[34] Upregulation of the expression and/or activity of PAD4, one of the five PAD isozymes expressed in the nucleus of bloodstream granulocytes, is associated with various autoimmune diseases including rheumatoid arthritis (RA), Alzheimer's disease, multiple sclerosis, lupus, Parkinson's disease, and cancer.[35,36] Genetic studies have also found an association between polymorphisms in PAD4 gene expression and RA disease risk, while PAD4 knockout or pharmacological inhibition through small molecules have proven to reduce disease severities in several RA mouse models.[37] These results have inspired a recent rise in interest for developing PAD4 inhibitors as a new generation of RA therapeutics.

Blocking PAD4-mediated extracellular citrullination is a promising strategy for autoimmune intervention, but further validation and mechanistic interrogation needs to be performed and requires new tools.[38] Antibodies are established protein tools and have been commonly employed to bind under specific conditions including pH and high calcium concentrations to target different protein conformations to interrogate their function.[39,40] Here, using an unbiased, Fab-phage display approach, we discovered conformation-selective antibodies and built a toolkit of both

inhibitory and activating antibodies to influence PAD4 activity. We believe these function-modulating antibodies can aid the study of PAD4's involvement in disease pathology and ultimately reveal insights for designing alternative and unique treatments of RA.

Our study reports the first structures of antibody binders in complex with PAD4, and our findings identify several allosteric mechanisms that may be harnessed for PAD4 activation or inhibition. The first mechanism is modulation of dimerization. Previous studies have discussed the impact of PAD4 oligomerization state on enzyme activity as structural changes along the dimerization interface heavily impact formation of the enzyme's substrate binding loop.[3,28,30] Though the monomeric form remains active, k_{cat} values of previously reported PAD4 monomerization mutants are decreased as much as 4-fold compared to the wild type enzyme. These results provide a rationale for developing PAD4 inhibitors that function as dimerization blockers. In our study we identified several binders, hI281, hA288, and hA362, that affect PAD4 dimerization and PAD4 function. In particular, hI281 blocks enzyme dimerization. Blocking enzyme dimerization with hI281 effectively inhibits enzyme activity, just as PAD4 monomerization mutants are also less active. On the other hand, the activating antibodies, hA288 and hA362, promoted dimerization and catalytic activity. Through cryo-EM analysis of the Fab/PAD4 complexes, hA362 was found to bind and stabilize the interface loop, which has previously been shown to stabilize the substrate binding loop upon PAD4 dimer formation.[30] These findings both reinforce previously known PAD4 regulation mechanisms afforded by enzyme oligomerization state and show that functional antibodies can be used to both activate and inhibit PAD4 activity through this mechanism.

Along with dimerization, calcium binding also plays an important role in PAD4 function. Ca^{2+} binding in the C-terminal domain is vital in structuring the enzyme catalytic pocket, while Ca^{2+} binding in the N-terminal domain plays an important role in proper protein folding.[1] As PAD4's catalytic pocket is partially disordered in the absence of C-terminal Ca^{2+} ions, this acidic pocket presents an allosteric site that directly affects PAD4's ability to catalyze citrullination. As

determined through cryo-EM, our antibody hI365 binds to a structural fragment in the C-terminal Ca^{2+} pocket. Once hI365 engages PAD4, several key residues in this pocket are rearranged, inhibiting Ca^{2+} binding and subsequently leading to disruption of substrate binding and loss of enzymatic activity. Interestingly, hI365 also recognizes the folded Ca^{2+} site in the N-terminal domain, thus exhibiting selective binding to the PAD4 only in the presence of Ca^{2+}. Specifically targeting calcium-bound PAD4 presents an intriguing strategy for RA treatment as a calcium-form selective binder will ignore inactive, apo-PAD4. This could hypothetically improve the therapeutic index of an anti-PAD4 drug, as we would only target protein actively contributing to disease pathogenesis in the joints.

Additionally, our functional antibodies show high specificity for PAD4 over other PAD isoforms. Of the 4 other PAD isoforms (-1, -2, -3, and -6) we thought it was most relevant to test binding of our antibodies against PAD2 and PAD3, both of which share significant structural similarities to PAD4.[41,42] PAD2 is also implicated in various autoimmune diseases and ubiquitously expressed across all tissue, so there is a need for PAD4-specific tools to address its role separately from PAD2 in disease pathology. As for PAD3, it is known that a subset of RA patients has PAD3/PAD4 cross-reactive autoantibodies.[21,22] Given the precedence for PAD3/PAD4 antibody cross-reactivity, we thought it would be important to show that our antibodies are still PAD4 specific. To address this, we tested the binding of our antibodies against commercially available PAD2 and PAD3 via BLI and found that none of our antibodies exhibit binding to PAD2 or PAD3. These PAD4 specific antibodies should be useful for isolating the functional effects of extracellular PAD4 from PAD2/3.

Taken together, we discovered functional antibody modulators of PAD4 through unbiased phage selection methods. These antibodies were key to reveal alternative mechanisms, direct and allosteric, for both PAD4 activation and inhibition. In the future, these highly specific and functional PAD4 binders may be used to investigate PAD4 activity in mouse arthritis models and human samples and shed light on how PAD4-mediated citrullination impacts RA disease

progression. Given the increasing evidence that PAD4 is a feasible anti-RA therapeutic target, using biologics to target extracellular PAD4 may provide a safer and more potent alternative for RA treatment.

Methods

Vector design and construction

We used a previously described vector for expression of Fabs in *E. coli*.[43] The pFUSE-hIgG1-Fc (InvivoGen) vector was used for expression of IgGs wherein the heavy chain was genetically fused to the hIgG1-Fc and the light chain was expressed on a separate copy of the vector without Fc. The vector used to express hPAD4 and mPAD4 was generated by Gibson cloning into the same vector for Fab expression. Each PAD4 was fused to an N-terminal His_6-AviTag-PreScission or TEV cleavage site.

Expression and purification of PAD4

C43 (DE3) Pro+ or BL21 Gold (DE3) *E. coli* containing PAD4 expression vectors were grown in 2xYT at 37 °C to an OD-600 of 0.4–0.8 and then protein expression was induced by the addition of 0.5–1.0 mM IPTG. Incubation temperature was subsequently reduced to 18°C and the cultures were allowed to shake for 16–20 h. Cells were harvested by centrifugation and lysed using sonication. The lysate was centrifuged to remove inclusion bodies. The enzymes were purified by Ni-NTA resin with 0.5 mM TCEP supplemented to all buffers to prevent PAD4 oxidation. The purified enzyme was buffer exchanged to 50 mM Tris (pH 8), 400 mM sodium chloride, and supplemented with 0.5 mM TCEP. Purification steps were performed on ice to maintain high PAD4 enzymatic activity. Purified enzyme was aliquoted and flash frozen.

Phage display selections

All phage selections were done according to previously established protocols. Briefly, selections with antibody phage Library E were performed using biotinylated antigens captured with streptavidin-coated magnetic beads (Promega).[43] Prior to each selection, the phage pool was incubated with streptavidin beads in order to deplete the library of any binders to the beads or sticky antibodies. In total, 3-5 rounds of selection were performed with decreasing amounts of PAD4 antigens (100 nM, 50 nM, 10 nM, 10 nM, 10 nM). 10 mM $CaCl_2$ or 1 mM EDTA was added for Ca^{2+} or calcium-free selection schemes. 0.5 mM TCEP was used in all buffers to keep thiols reduced.

Phage library-based affinity maturation

Kunkel mutagenesis was used to incorporate mutagenized oligo pools into CDR regions as previously described.[44,45,46] Briefly, each soft randomization library was first generated by producing ssDNA in *dut-/ung- E.* coli cells. Phosphorylated oligos were annealed to the ssDNA and subsequently amplified to generate cccDNA which was electroporated into SS320 *E. coli* cells, grown up, and infected with M13K07 helper phage to generate our Fab-phage library. Selections using this library were then performed against PAD4 as described previously.

Expression of Fabs

Fabs were expressed as previously described.[43] Briefly, C43 (DE3) Pro+ *E. coli* containing expression plasmids were grown in TB at 37°C in an autoinduction media for 6 hrs and incubation temperature was subsequently reduced to 30°C where the cultures were allowed to grow for an additional 16-18 h. Cells were harvested by centrifugation, lysed, and Fabs were purified by Ni-NTA resin. Fab purity and integrity were assessed by SDS-PAGE. Fab sequences are provided in the **Supplementary Note 1.**

Expression of IgGs

Expi293 (Life Technologies) cells were transiently co-transfected with two pFUSE (InvivoGen) vectors harboring either the IgG heavy chain and the IgG light chain at a mass ratio of 1:1. The ExpiFectamine 293 transfection kit (Life Technologies) was used for transfections as per manufacturer's instructions. Cells were incubated for 5 days at 37°C in an 8% CO_2 environment before the supernatants were harvested by centrifugation. IgGs were purified by Protein A affinity chromatography or Ni-NTA resin and assessed for quality and integrity by SDS-PAGE. IgG sequences are provided in the **Supplementary Note 1.**

Phage ELISAs

ELISAs were performed according to standard protocols. Briefly, 96-well Maxisorp plates were coated with NeutrAvidin (10 µg/ml) overnight at 4°C and subsequently blocked with BSA (2% w/v) for 1 h at 20°C. 20 nM of biotinylated PAD4 was captured on the NeutrAvidin-coated wells for 30 minutes followed by the addition phage supernatants diluted 1:5 in ELISA buffer (TBS, pH 7.4, 0.05% Tween-20, 0.2% BSA) for 30 minutes. 10 mM calcium or 1 mM EDTA were supplemented to buffers to determine binding with either the calcium bound form or the apo form of PAD4. The bound phage was then detected using a horseradish peroxidase (HRP)-conjugated anti-phage monoclonal antibody (GE Lifesciences 27-9421-01) and imaged on a Tecan i-control (v3.4.3.0) plate reader. For competition ELISAs, diluted phage supernatants and PAD4 were incubated for 30 minutes at 20°C before addition to the NeutrAvidin coated plates.

DSF antibody stability assay

Fab or IgG samples in PBS were mixed with Sypro Orange dye (20 X stock) to make a final antibody concentration of 2, 4, 8, or 16 µM and a 4X dye concentration. 10 µL of each mixture

was transferred to a Biorad 384-well PCR plate and covered by a qPCR Sealing Tape. The assay was preformed over a temperature range of 25°C to 95°C with a temperature ramping rate of approximately 0.5°C/30 seconds, and fluorescence was detected using a Roche LC480 Light Cycler. Due to instrumentation constraints, the ramp rate is set based on the number of data points acquired per °C. Currently this is set at 20 acquisitions per °C equating to a ramp rate of 0.03 °C/sec.

PAD4 293 Flp-in cells

To construct the PAD4 expressing HEK293 cell lines, Flp-In HEK293 (ThermoFisher) cells were co-transfected with the pOG44 vector (ThermoFisher) and a construct encoding PAD4, PAD4 D350A, or PAD4 with a green fluorescent protein Dronpa tag in the pcDNA5/FRT Mammalian Expression vector (ThermoFisher). Cells expressing PAD4 constructs were selected for in DMEM supplemented with 10% FBS, 100 µg/mL zeocin, and 100 µg/mL Hygromycin B (ThermoFisher). Protein expression was confirmed by fluorescence microscopy detection of the Dronpa tag or PAD4 activity assays. To prepare cell lysates for the PAD4 activity assays, cells were lysed in RIPA cell lysis buffer (50 million cells/mL) and supplemented with protease inhibitor cocktail (Sigma-Aldrich). After rotating at 4°C for 15 min, cell lysates were sonicated and spun at 16,000 g to remove cell debris.

Fluorescence-based PAD4 activity assay

PAD4 activities in the absence or presence of antibodies were assessed in a fluorescence-based assay with a pro-fluorescence substrate analog.[28] 1 µM of PAD4 was mixed with various concentrations of antibodies and calcium at 4°C for 45 min, followed by the addition of 25 µM substrate ZRcoum, an arginine mimetic that releases a fluorophore once cleaved by trypsin. The reaction was incubated at 37°C for 110 min. Upon addition of excess trypsin/EDTA to the solution,

the fluorophore was liberated from uncitrullinated ZRcoum, but remained quenched and unmodified when the substrate was citrullinated. The reaction was read on a Tecan i-control (v3.4.3.0) fluorescence plate reader with an excitation wavelength of 345 nm and an emission wavelength of 465 nm.

Citrullinated histone H3 PAD4 activity assay

PAD4 activities in the absence or presence of antibodies were also assessed using a citrullinated histone H3 Western assay. 10-100 nM recombinant PAD4 were mixed with antibodies with various concentrations of calcium at 4°C for 45 min, and then incubated with 760 nM recombinant histone H3.1 (New England Biolabs). The reaction was incubated at 37°C for 110 min, followed by western analysis using an anti-citrullinated H3 primary antibody (Abcam Ab5103) and an anti-rabbit HRP secondary antibody. For the PAD4 293-Flp-in cell lysate assay, 10 µL of cell lysate was mixed with varying concentrations of antibodies and calcium in a final volume of 19 µL at 4°C for 45 min. The reactions were then incubated at 37°C for 110 min and subsequently analyzed by western blot using an anti-citrullinated H3 antibody (Ab5103) and anti-rabbit HRP secondary antibody. Images were acquired in Image Lab (v5.0) and processed with Image Studio Software (v5.2). IC_{50} measurements were obtained with technical triplicates and quantified using Fiji.[47]

Modified citrulline western blot assay

Lysate was harvested from Expi293T cells using a 1x RIPA + protease inhibitor solution. Antibody-PAD4 complexes were pre-formed at 4°C for 45 min. The complexes were then incubated at 37°C for 110 min and subsequently analyzed by western blot using an anti-modified citrulline detection kit (EMD Millipore). PVDF membrane was incubated with anti-citrulline probe overnight, then blocked and stained with a primary anti-modified citrulline antibody and secondary HRP linked anti-IgG antibody. HRP signal was detected using a BioRad ChemiDoc imager.

Size exclusion chromatography

SEC analysis was performed using an Agilent HPLC 1260 Infinity II LC system equipped with an AdvanceBio SEC column (300 Å, 2.7 µm, Agilent). Each analyte was injected at 10 µM and run with a constant mobile phase of TBS high salt buffer (50 mM Tris, pH = 8, 400 mM NaCl) for 15 minutes. Fluorescence (excitation 285 nm, emission 340 nm) and absorbance were measured and analyzed with Agilent OpenLab CDS ChemStation software.

Biolayer interferometry

BLI measurements were made using an Octet RED384 (ForteBio) instrument. Biotinylated PAD4 was immobilized on optically transparent streptavidin biosensors (ForteBio) and loaded until a 1 nm signal was achieved. After blocking with 10 µM biotin, purified binders in solution were used as the analyte. TBSTB was used for all buffers. Data were analyzed using the ForteBio Octet analysis software, and kinetic parameters were determined using a 1:1 monovalent binding model (https://www.sartorius.com/en/products/protein-analysis/octet-bli-detection/octet-systems-software).

Negative-stain TEM

2.5 uL of PAD4 samples at 40 µg/mL were applied to a glow-discharged Cu grid covered by continuous carbon film, then stained with 0.75% (w/v) uranyl formate.[48] A Tecnai T12 microscope (ThermoFisher FEI Company) operated at 120 kV was used to analyze negatively stained grids. Images were recorded using an UltraScan 4000 camera (Gatan) at normal magnification of 52,000x corresponding to a pixel size of 2.21 Å on the specimen.

Cryo-EM studies of PAD4 in complex with hA362 and hI365

hPAD4/hA362: Sample concentrations used for cryo-EM were between 2-3 uM in 50 mM Tris pH 8.0, 150 mM NaCl, 10 mM CaCl2, 0.5 mM TCEP. Grids were frozen at 7 sec blot time, 5 ^{o}C, 95% humidity on a Vitrobot. Movies were collected on a Titan Krios at 300 kV and 0.835 Å /pix, 74 e^{-} total dose. 2214 images were collected untilted; 1620 movies were collected at 15º tilt; 1048 were collected at 30º tilt. CTF (constant transfer function) was estimated using Patch CTF in cryoSPARC2.[49] 1593955 particles were template picked after curation. 3D classification was performed in cryoSPARC2. One class at 3.41 Å resolution with 274910 particles was chosen for further refinement. These particles underwent 2 rounds of ab initio reconstruction into 2 models. After this point, C2 symmetry was imposed. The resulting stack of 134775 particles went into one round of ab initio (2 classes) and one round of heterogeneous refinement (2 classes). One class with 92424 particles underwent homogeneous refinement to 3.5 Å. These particles and corresponding map were imported into cisTEM and auto refined to 2 classes.[50] Four rounds of local refinement into 2 classes were performed at increasing starting resolutions (10 Å, 7 Å, 5 Å, 4.5 Å), leading to a final local refinement where 1 class with 59.78% (55251) of the particles at 3.3 Å was chosen. Generate3D was then used to generate half maps. Local and global resolution maps were generated in cryoSPARC2, and directional resolution was determined with the 3DFSC server.[51]

Model building: PDB 1WD9 was fit, SWISS-MODEL was used with model Fab 6OTC (light chain) and 1N8Z (heavy chain) to build a homology model from the hA362 Fab sequence. This model then underwent Phenix Real-Space Refinement,[52] followed by Rosetta relaxation in torsion space. Loops that were not visible in the density were deleted with Coot. Isolde was then used for local adjustments to fix Ramachandran outliers and manual placement into the map density.[53] A final Rosetta torsion relax with C2 symmetry imposed was used to create the final model. Validation

was performed in Phenix, and model-map correlation was performed with the MapQ plugin for UCSF Chimera.[54,55] Buried surface area was calculated with the PISA server (**Fig. 1.14**).

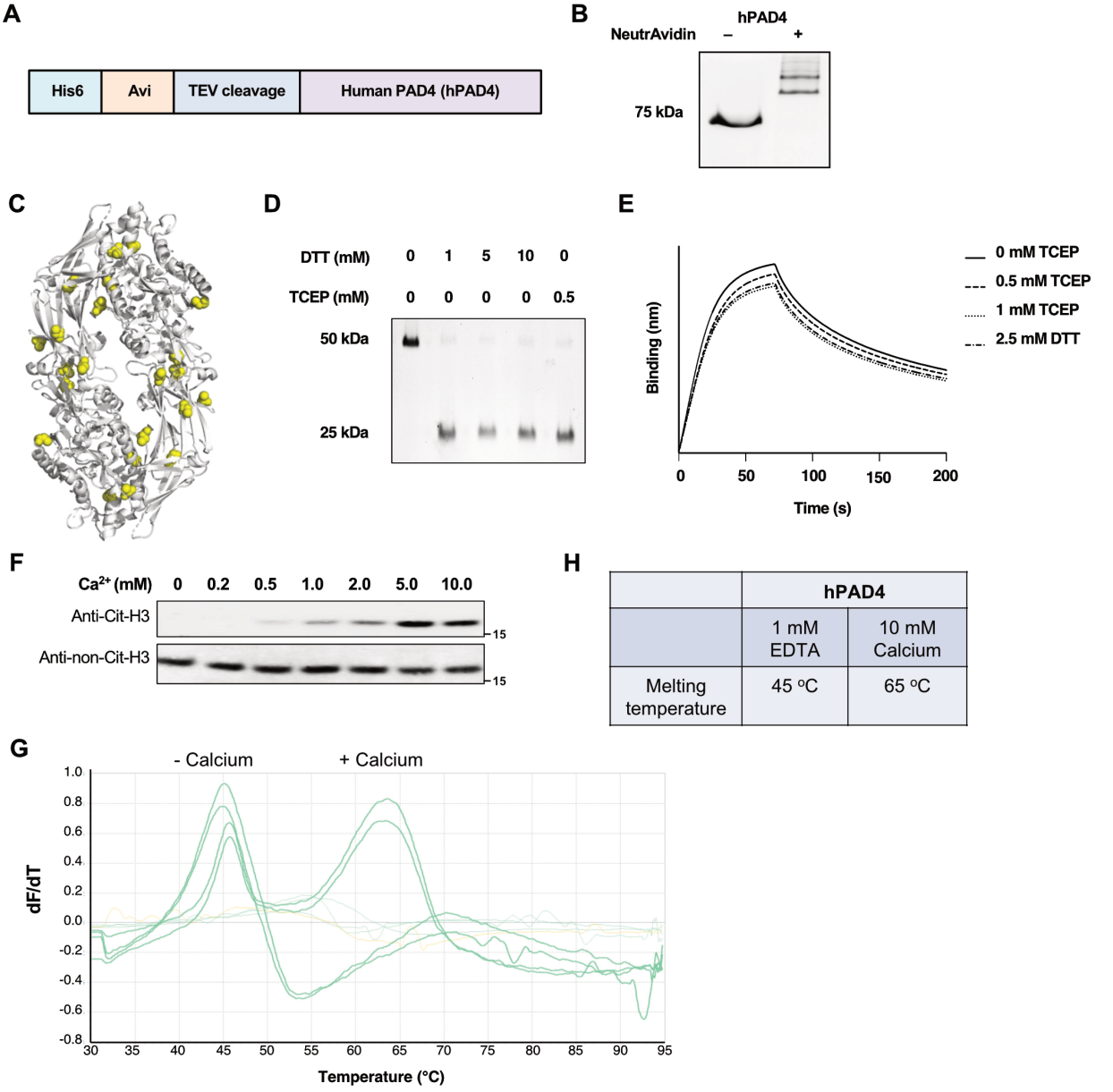
hPAD4/hI365: Sample concentrations used for cryo-EM were between 2-3 uM in 50 mM Tris pH 8.0, 150 mM NaCl, 10 mM CaCl2, 0.5 mM TCEP. Grids were frozen at 7 sec blot time, 5 ^{o}C, 95% humidity on a Vitrobot. As initial processing of untilited micrographs resulted in anisotropic resolution reconstruction, additional tilted data sets were collected. Movies were collected on a Titan Krios at 300 kV and 0.835 Å /pix, 70 e- total dose. 1773 images were collected untilted; 1844 movies were collected at 15º tilt; 842 were collected at 30º tilt. CTF was estimated with Patch CTF in cryoSPARC2. Bad micrographs were removed with the exposure curation tool, leading to 1225 untilted, 1366 at 15º tilt, and 648 at 30º tilt. These underwent template picking, and 1215874 particles were picked. These underwent ab Initio reconstruction into 3 classes; the best class with 780324 particles was chosen. This class underwent ab Initio into 2 classes; the best class had 391776 particles. These were re-extracted and particles were placed into 5 ab initio classes. 241 of these then went through 3 classes of heterogeneous refinement; the best class with 143413 particles at 6.3 Å was chosen. These particles then underwent 2 rounds of heterogeneous refinement with C2 symmetry imposed, leading to a stack of 73648 particles, which were then imported into cisTEM for further processing. These were classified in 2 classes with Auto Refinement; both classes resulted in 3.9 Å maps. All particles then underwent four rounds of local refinement into 2 classes at increasing starting resolutions (8 Å, 6 Å, 5 Å, 4.5 Å). In the final round, class 1 was chosen with 62.22% (45824) of the particles at 2.98 Å. Generate3D was then used to generate half maps. Local and global resolution maps were generated in cryoSPARC2, and directional resolution was determined with the 3DFSC server.

Model building: PDB 1WD9 was fit, SWISS-MODEL was used with model Fab 6OTC (light chain) and 1N8Z (heavy chain) to build a homology model with the 365 Fab sequence. Fab loops were deleted and RosettaES was used to rebuild the loops. This model was flexibly fit into the density

with Rosetta torsion relax. Loops and domains that were not visible in the density were deleted with Coot. Manual placement of resolved regions that did not match the homology models into map density was performed in Isolde. Finally, Rosetta torsion relax was performed again with C2 symmetry imposed to produce a final model. Validation was performed in Phenix, and model-map correlation was performed with the MapQ plugin for UCSF Chimera. Buried surface area was calculated with the PISA server (**Fig. 1.15**).

Statistics and Reproducibility

All representative SDS-PAGE protein gels and western blots shown were reproduced three times before inclusion in text.



	hPAD4	
	1 mM EDTA	10 mM Calcium
Melting temperature	45 °C	65 °C

Figure 1.1: Expression and biophysical characterization of hPAD4. (**A**) Map of hPAD4 gene expressed in C43 or BL21 *E. coli.* N-terminal 6x His tag and Avi tag are separated by a protease cleavage site to hPAD4 to remove tags. (**B**) Gel shift assay of full length hPAD4 (left, 74 kDa) and biotinylated hPAD4 bound to NeutrAvidin (right, 134 kDa). (**C**) PAD4 enzyme contains 19 free cysteines (colored in yellow), requiring storage with reducing agent to prevent oxidation of cysteines. (**D**) SDS-PAGE showing reduction of the inter-heavy and light chain disulfide of a $4G10^{4D5}$ antibody upon addition of reducing agents. The $4G10^{4D5}$ antibody adopts the antibody scaffold of the Fab-phage library used for PAD4 antibody selection. 50 kDa band represents disulfide-linked Fab while 25 kDa bands represent separated heavy and light chains following TCEP reduction. (**E**) Biolayer Interferometry (BLI) showing binding of Fab $4G10^{4D5}$ to its cognate antibody is not influenced by various reducing agents. (**F**) Immunoblotting detection of PAD4-mediated Histone 3 citrullination (Cit-H3) showing PAD4 activity is calcium dependent. (**G**) Differential scanning fluorimetry (DSF) plot showing melting temperatures of hPAD4 in the presence and absence of Ca^{2+}. (**H**) Table listing PAD4 melting temperatures measured by DSF. Representative gels and blots were reproduced three times before inclusion in main text.

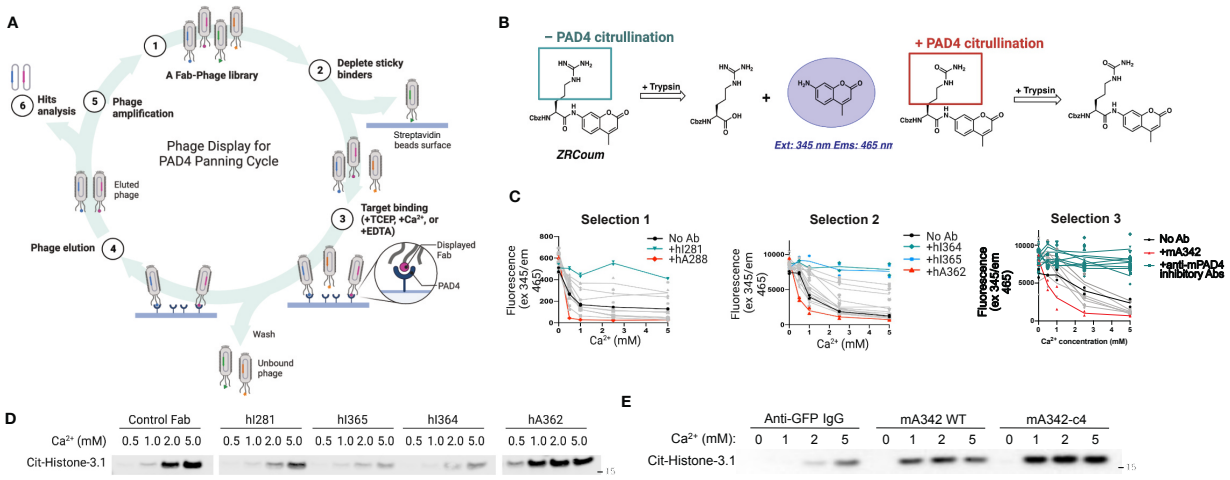


Figure 1.2: PAD4 antibody selection and hits characterization. (**A**) PAD4 phage display schematic. The Fab-phage library (1) was depleted of non-specific streptavidin (SA) binders (2); the remaining soluble Fab-phage were allowed to bind to PAD4 immobilized on SA in the presence of 1 mM Ca^{2+} or 1 mM EDTA with TCEP and washed (3); the bound Fab-phage were eluted (4), amplified (5) and subjected to additional selection rounds before final characterization of individual hits (6). **(B)** Schematic of fluorescent-substrate PAD4 activity assay (derived from Sabulski, et. al.). Higher fluorescent signal is indicative of lower PAD4 activity. **(C)** Characterization of Fab binding effect on PAD4 activity from hPAD4 selection 1 and 2 and mPAD4 selection 4 (see supplementary Fig. 3A) via fluorescent-substrate activity assay. Highlighted clones hI281, hA288, hA362, hI364, hI365, and mA342 are described in the manuscript text. **(D)** Activity of hPAD4 in the presence of control antibody, an inhibitory Fab to human PAD4 (hI281), or activating Fabs (hA288, and hA362) measured by citrullination of protein substrate H3. hI281 reduces the activity of PAD4 while hA288/hA362 increases PAD4 activity. **(E)** Activity of mPAD4 in the presence of control antibody, an activating IgG to murine PAD4 (mA342) and its variant mA342-c4 measured by citrullination of protein substrate H3. Representative blots were reproduced three times before inclusion in text.

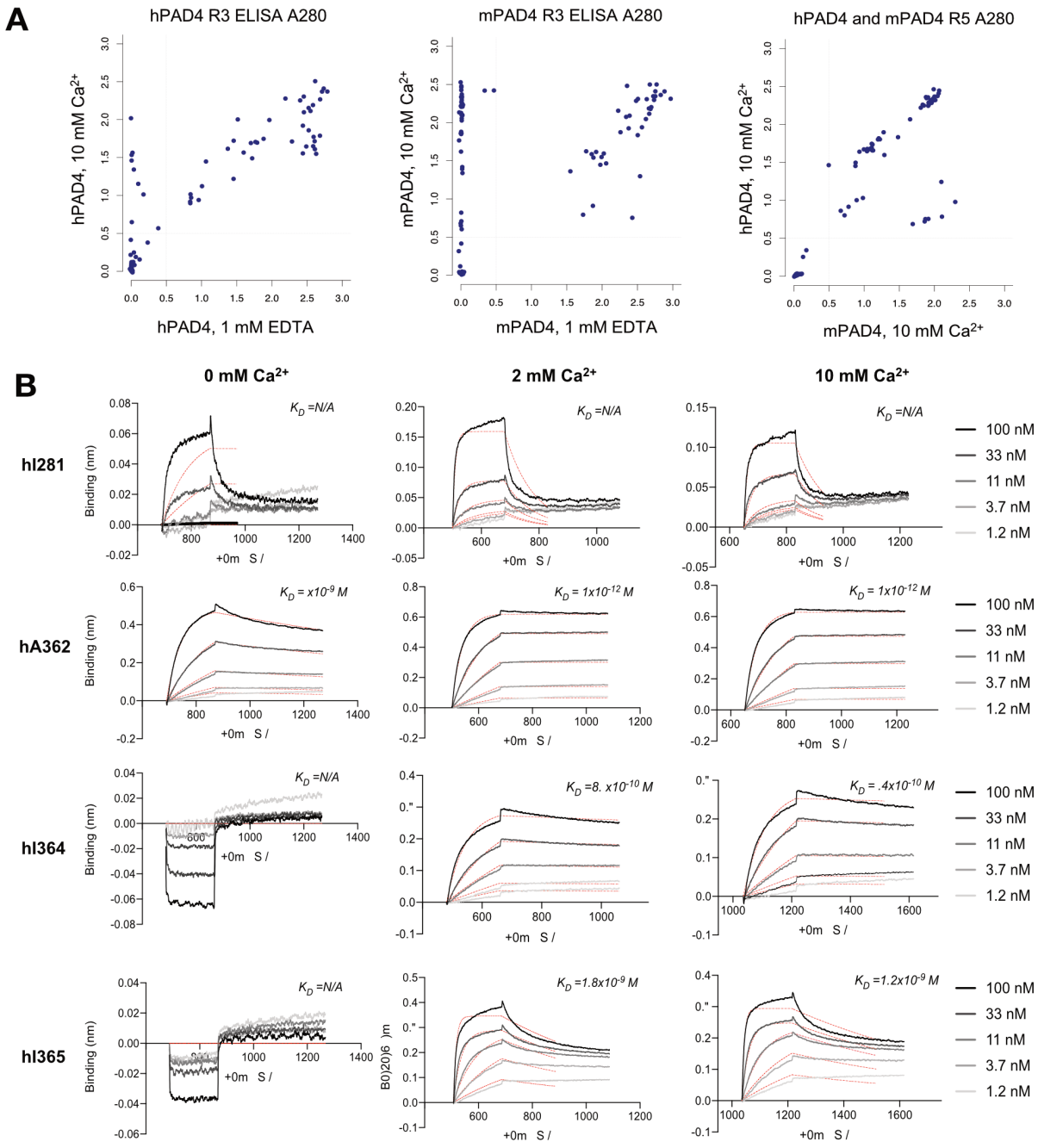


Figure 1.3: Phage ELISA for characterizing hPAD4/mPAD4 cross reactivity and multi-point biolayer interferometry (BLI) binding data of lead candidate IgGs to hPAD4. (**A**) Binding of Fab-phage to hPAD4/mPAD4 measured by ELISA A280 signal. Very few clones from hPAD4 or mPAD4 phage selections (left, middle) were cross reactive. By taking the round 3 phage pool and performing selections with the other antigen for two additional rounds, nearly all phage clones from round 5 are cross reactive with hPAD4 and mPAD4 (right). (**B**) Multi point BLI data show that hI364 and hI365 are dependent on calcium to bind hPAD4. hI281 and hA362 are calcium-independent binders although binding to the calcium-bound form of the enzyme is stronger. Binding of all clones to PAD4 at 2 mM Ca^{2+} is comparable to their binding at 10 mM Ca^{2+}. K_Ds reported as <1 nM are due to bottoming out of off-rate measurements on BLI.

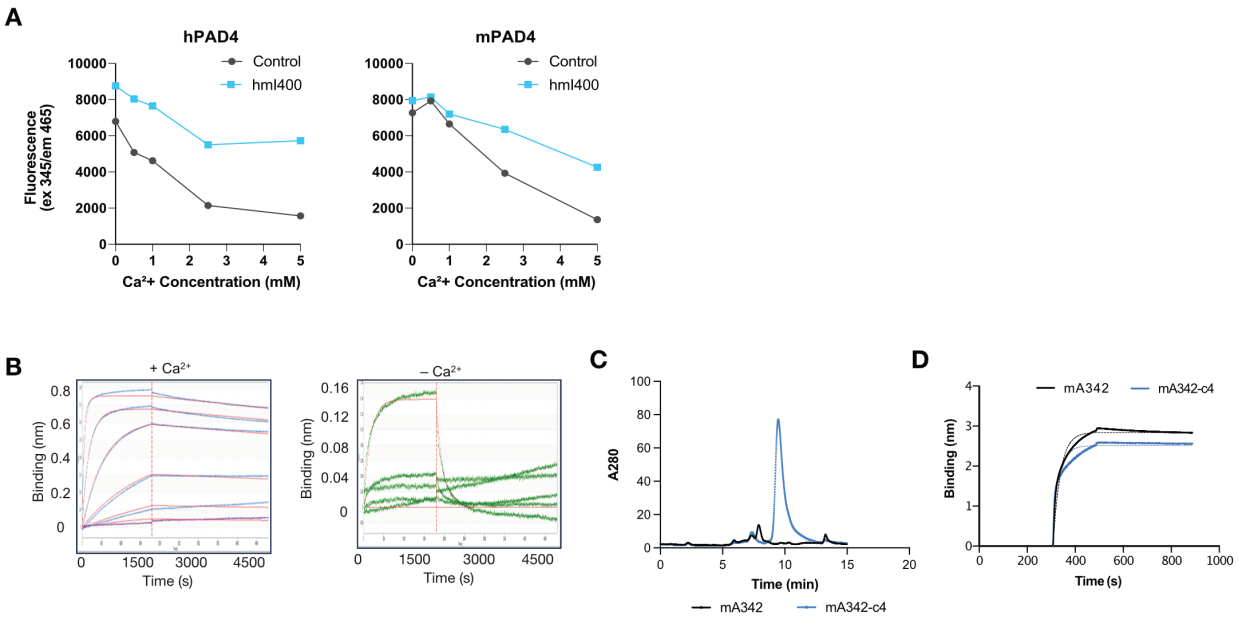
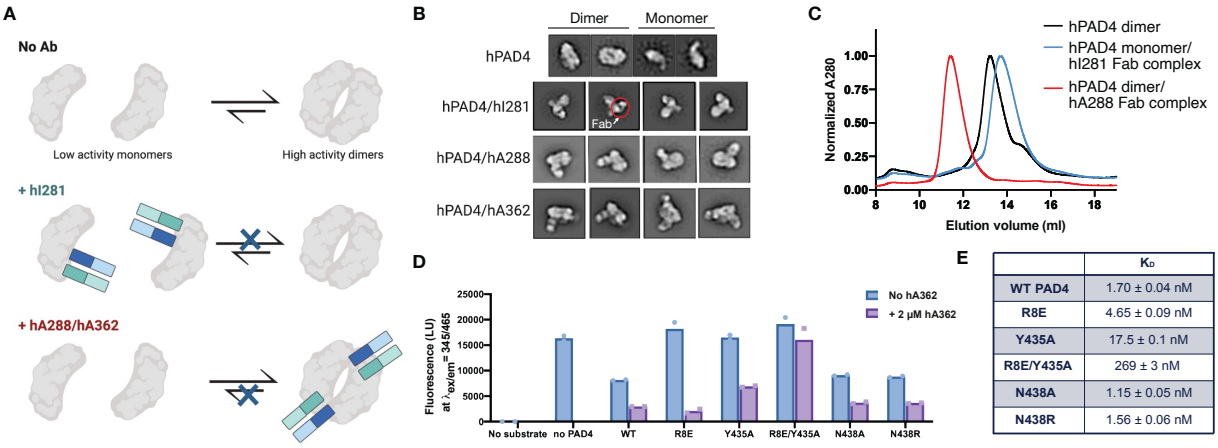
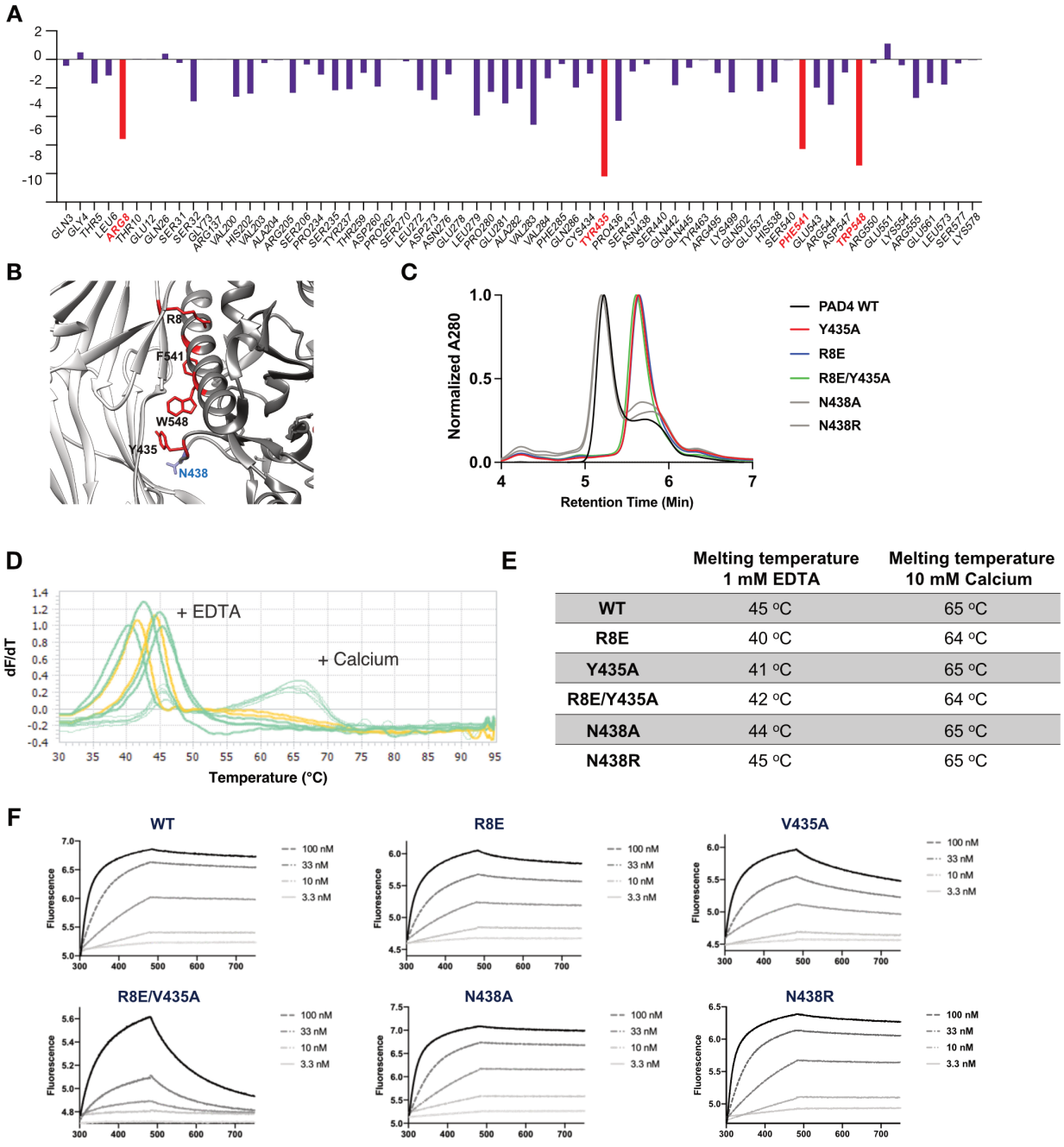


Figure 1.4: Functional characterization of mPAD4 and h/mPAD4 cross reactive antibodies. (**A**) hml400 was identified as a functional inhibitor of both hPAD4 and mPAD4. (**B**) BLI results show mA342 interacts with the calcium bound form of mPAD4 but binds minimally to the calcium-free form. (**C**) SEC trace showing improved biophysical properties of mA342-c4 as compared to WT mA342. (**D**) BLI comparison of WT mA342 and its improved variant mA342-c4 in 10 mM Ca^{2+}.



	K_D
WT PAD4	1.70 ± 0.04 nM
R8E	4.65 ± 0.09 nM
Y435A	17.5 ± 0.1 nM
R8E/Y435A	269 ± 3 nM
N438A	1.15 ± 0.05 nM
N438R	1.56 ± 0.06 nM

Figure 1.5: Antibodies that modulate hPAD4 dimerization influence enzymatic activity. (**A**) Summary diagram of antibodies influencing PAD4 activity through modulating its dimerization. Monomeric PAD4 is less active than dimeric PAD4, and hI281 blocks dimerization while hA288 and hA362 promote dimerization. (**B**) Negative stain-EM 2D micrographs showing monomeric state of PAD4 in complex with hI281 and dimeric state in complex with hA288 or hA362. (**C**) SEC traces of PAD4 alone, in complex with Fab-hI281, and in complex with Fab-hA362. PAD4 alone exists in both the monomer and dimer form, while addition of hI281 promotes the monomer form and hA362 promotes the dimer form as evidenced by elution time. (**D)** Binding is correlated with antibody function using a trypsin substrate assay. Biological duplicates are shown. Fabs binding to mutants with less affinity are less able to inhibit enzymatic activity (Right). (**E**) Table displaying measured binding affinities of activating Fab hA362 to PAD4 mutants (Left). Binding is drastically decreased upon mutation of several residues at the PAD4 dimer interface.



	Melting temperature 1 mM EDTA	Melting temperature 10 mM Calcium
WT	45 °C	65 °C
R8E	40 °C	64 °C
Y435A	41 °C	65 °C
R8E/Y435A	42 °C	64 °C
N438A	44 °C	65 °C
N438R	45 °C	65 °C

Figure 1.6: Biophysical characterization of PAD4 monomerization mutants and effect of PAD4 monomerization on antibody binding. (**A**) Computational modeling of the PAD4 dimer interface to identify key residues that contribute to dimerization (highlighted in red). (**B**) Key residues R8, Y435, and W548 highlighted at the dimer interface in red. The selected control residue N438 is highlighted in blue. (**C**) SEC trace showing earlier elution of dimeric, WT PAD4 compared to monomerization-promoting mutants R8E and Y435A. Control mutants, N438A and N438R (gray), elute with WT PAD4. (**D**). DSF plot of PAD4 mutants in the presence of 10 mM Ca^{2+} (green) or 1 mM EDTA (yellow). (**E**) Estimated T_ms of PAD4 mutants from DSF plots. (**F**) BLI measurements of hA362 binding to WT hPAD4 and monomerization mutants.

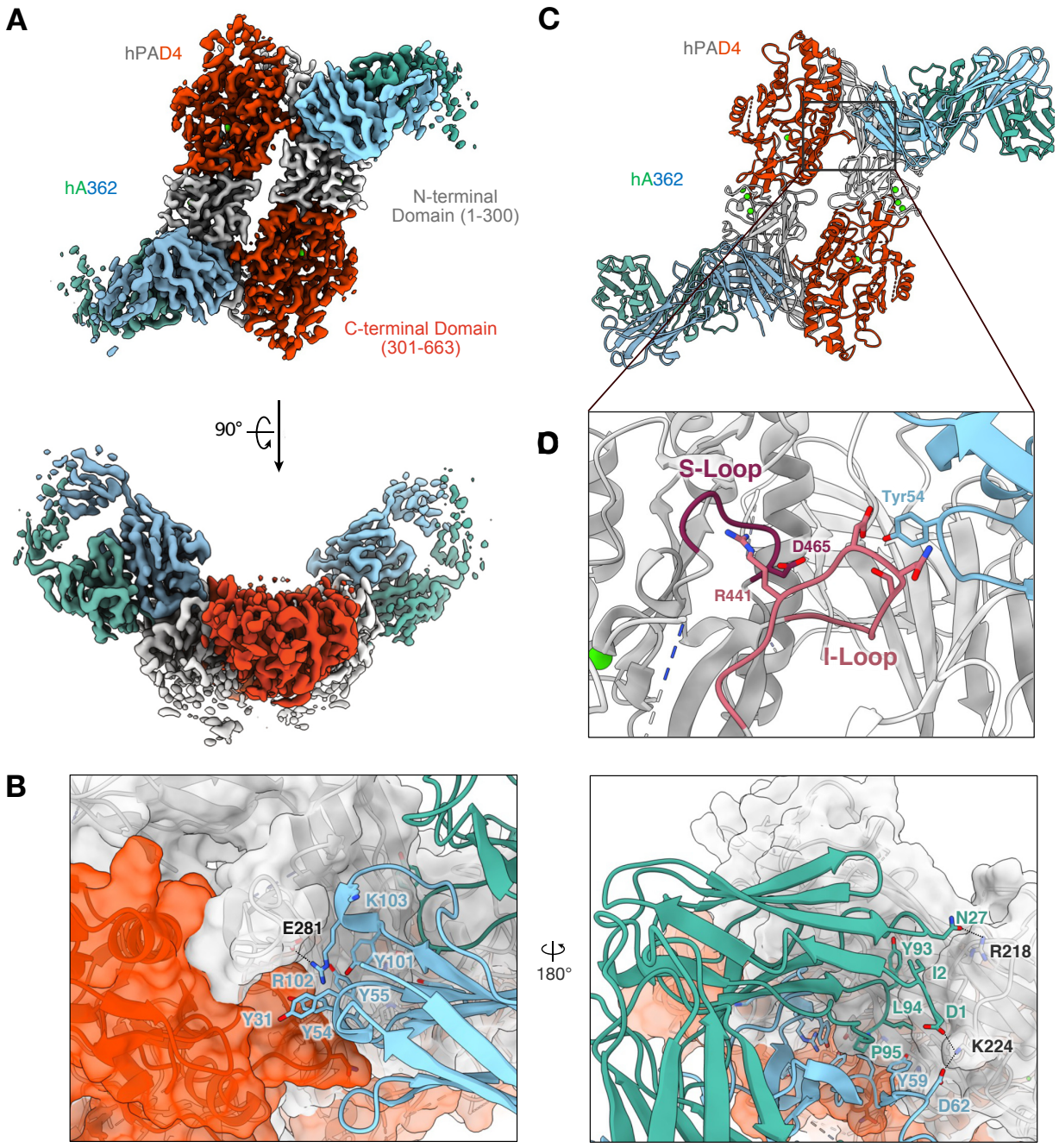


Figure 1.7: hA362 directly contributes to the PAD4 dimerization interface and helps order the substrate binding site. (**A)** Two views of the cryo-EM map of PAD4 in complex with Fab hA362 (PDB: 8SMK). PAD4 monomers are in shades of gray (NTD) and red (CTD), hA362 heavy and light chains are teal and blue. (**B**) Model of PAD4/hA362 derived from the cryo-EM map shown as ribbon. Boxed region delineates Fab CDR interaction with I- and S-loop on PAD4 (**C**) Zoomed-in view of the boxed region in (B) shows hA362 reaching across the PAD4 monomer to interact with the I loop on the other PAD4 monomer. This helps order the S loop via the R441-D465 salt bridge. (**D**) Detailed hA362-PAD4 interactions. Both chains of the Fab pack a large number of aromatics against both monomers of PAD4 dimer. Ion and hydrogen bonds marked with dashed lines.

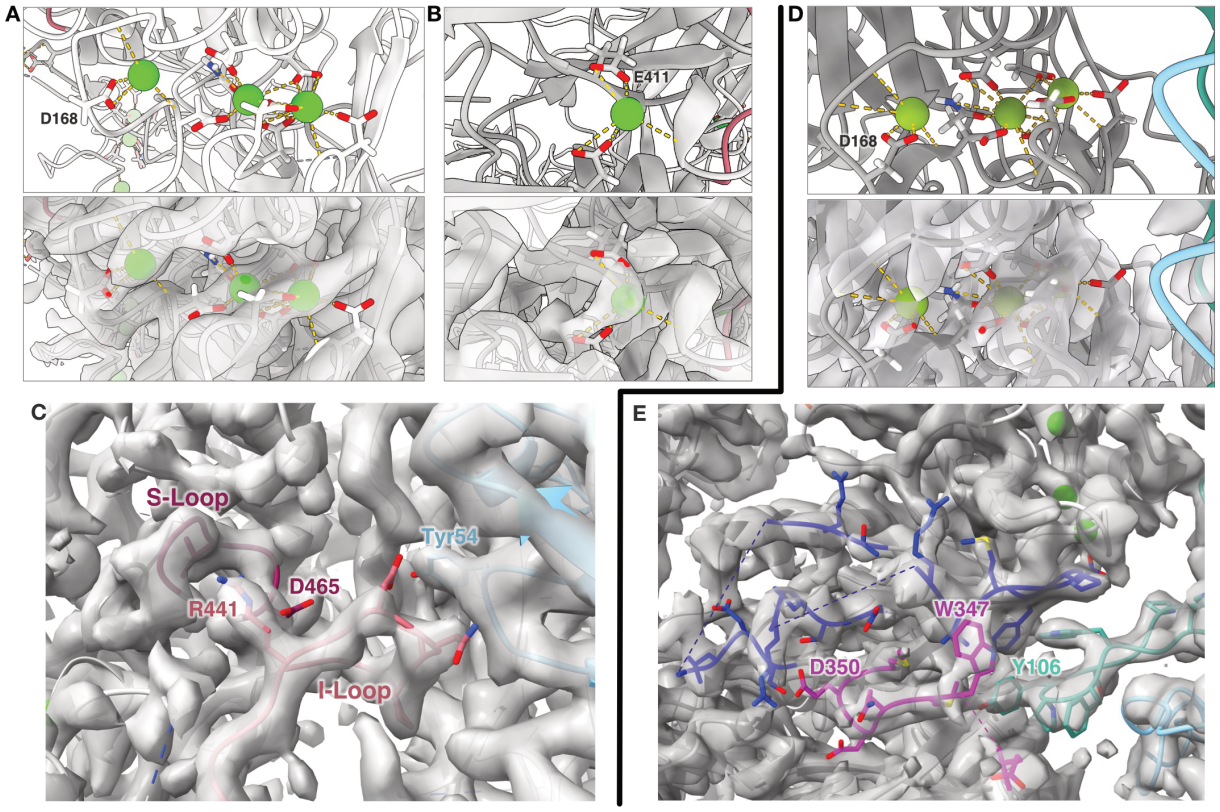


Figure 1.8: Detailed view of Ca^{2+} ions and Fab/hPAD4 interactions with cryo-EM map. (**A**) Three Ca^{2+} ions in the N terminal domain of hPAD4-hA362 structure. Model in ribbon and atoms on top, overlayed with cryo-EM map on the bottom. (**B**) Ca^{2+} ion in the C terminal domain of the hPAD4-hA362 structure. One Ca^{2+} was removed from analysis as the electron density was very weak. (**C**) Zoomed in view of the hA362 interactions with the hPAD4 I-loop with corresponding cryo-EM map. Adapted from **Fig. 3C**. (**D**) Ca^{2+} ions in the N terminal domain of the hPAD4-hA365 structure. (**E**) Zoomed in view of the hI365 interactions with the hPAD4 active site with corresponding cryo-EM map. Adapted from **Fig. .7C, right panel**.

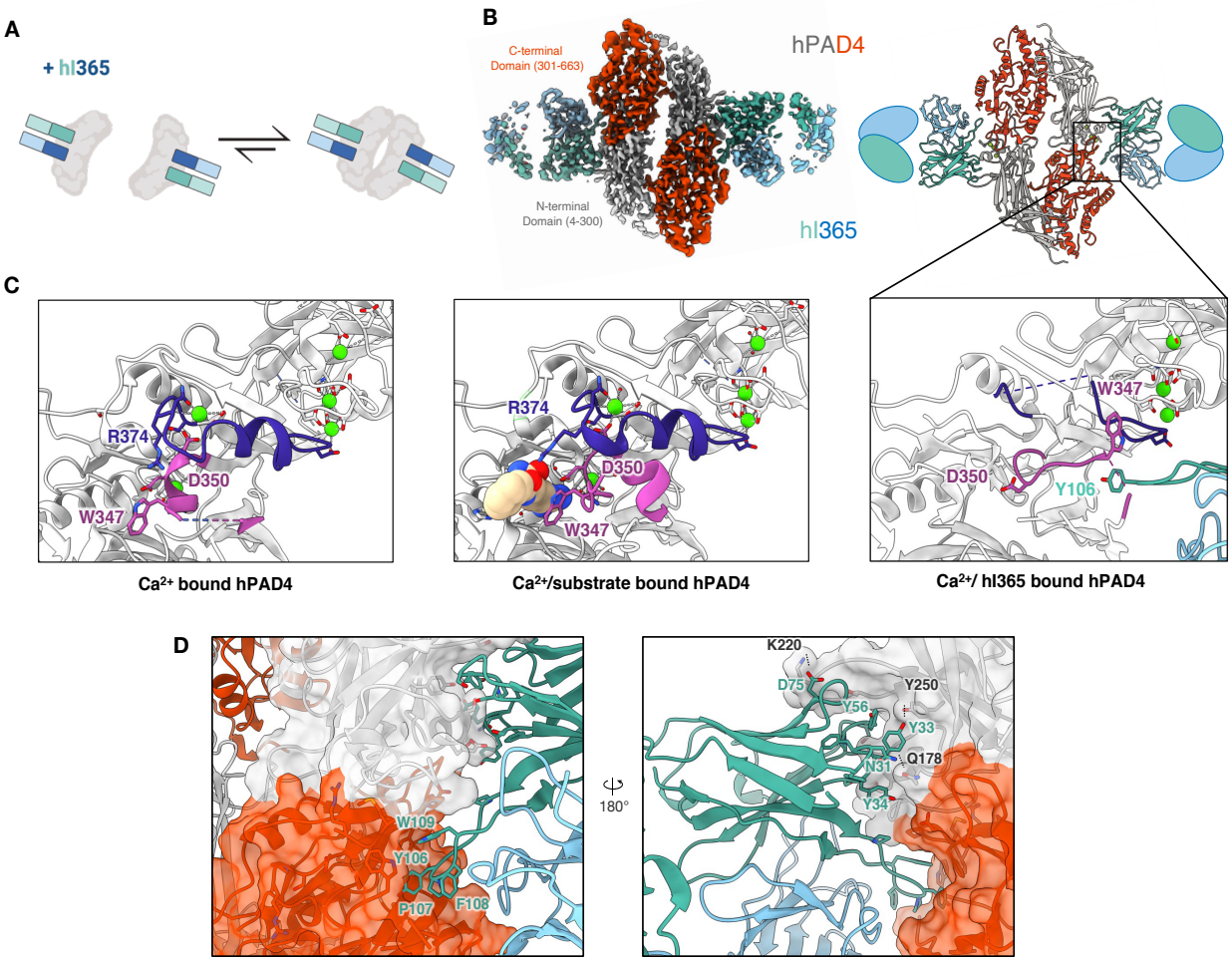


Figure 1.9: Cryo-EM structure illustrates the mechanism of calcium dependency and inhibitory function of hI365. (**A**) Diagram showing that inhibitory Fab hI365 is able to bind both the monomeric and dimeric form of hPAD4. (**B**) Cryo-EM map (left, PDB: 8SML) and the resulting model (right) of hPAD4 (NTD: grey, CTD: red) in complex with hI365 (HC: teal, LC: blue). The Ab primarily binds an N-terminal Ca^{2+}-coordinated region, but the H3 loop extends to the C-terminal pocket and induces conformational change. (**C**) Comparison of the Ca^{2+}-bound hPAD4 (PDB: 1WD9), Ca^{2+}/substrate bound hPAD4 (PDB: 1WDA), and the Ca^{2+}/ hI365 bound hPAD4. The interaction of hI365 with residues 340-352 in hPAD4 (magenta) alters the structure and orientation of this fragment disrupting the organization of several residues (D350, R374, W347) involved in calcium and substrate binding. (**D**) Detailed hA365-PAD4 interactions. The Fab interacts predominantly with one chain against one monomer of PAD4, burying considerable number of aromatics with some hydrogen bonds.

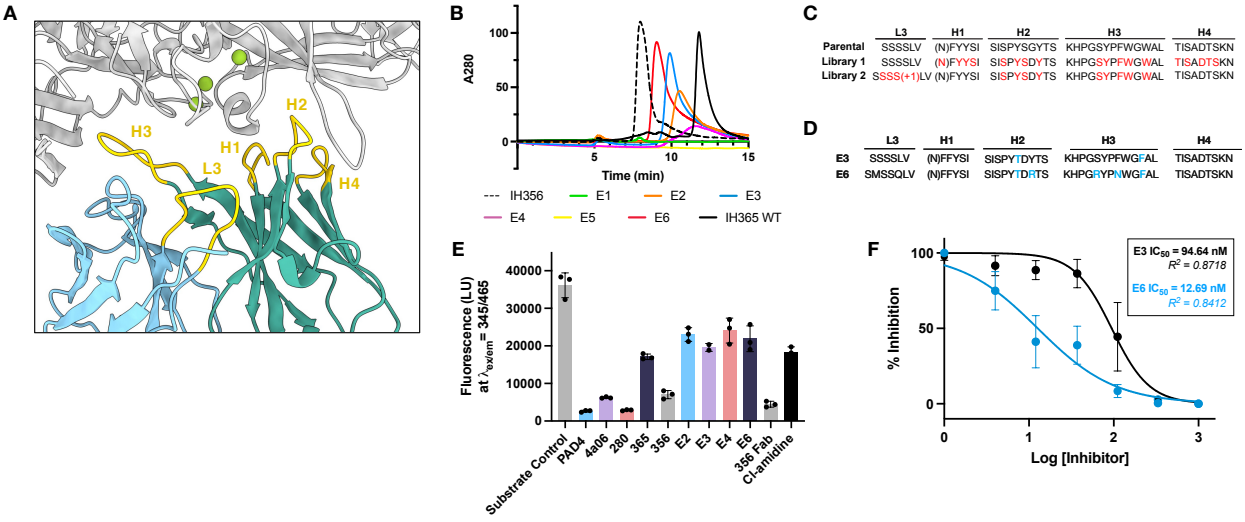


Figure 1.10: Antibody engineering strategies employed to improve affinity and inhibition activity of the calcium-dependent binder, hI365. (**A**) hPAD4-hI365 binding interface with CDRs L3, H1-H4 highlighted. CDR loops H1-H4 are forming contacts with PAD4, but L3 is too short to facilitate any contact with the enzyme. (**B**) SEC traces of WT hI365 and affinity matured clones. Selected hits show improved SEC profiles as indicated by earlier elution times. (**C**) CDR sequences of WT hI365 aligned with residues to be mutagenized via soft randomization; L3 is randomized to contain 9 or 10 amino acids. (**D**) CDR sequences of the two top engineered binders with good affinity and solubility. (**E**) PAD4 activity as measured by the fluorescent substrate assay shows that several Ab clones identified from soft randomization inhibit PAD4 more potently than WT hI365. 2 mM Ca^{2+} was used in this assay to mimic physiological extracellular conditions. Error bars represent mean ± standard deviation of three biological replicates. (**F**) IC50s of lead candidates E3 and E6 as measured by PAD4 citrullination H3. IC50s of E3 and E6 determined to be 95 nM and 13 nM, respectively and represent mean ± SEM of three biological replicates.

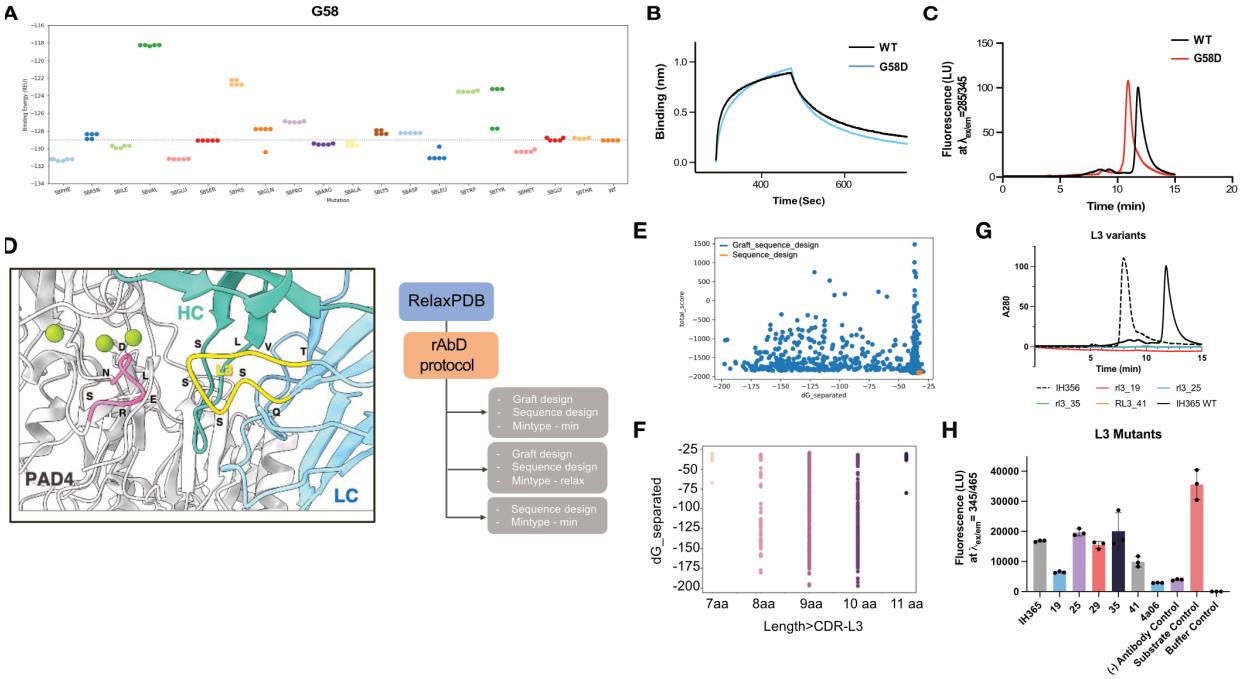


Figure 1.11: Rosetta antibody design (rAbD) guided optimization of hI365. (**A**) Several G58 mutants are predicted to reduce interface energy of hI365 and hPAD4. (**B**) BLI measurements showing that point mutant G58D binds PAD4 with similar affinity as WT hI365. (**C**) SEC trace showing improved biophysical properties of mutant G58D as compared to WT hI365, as evidenced by the earlier elution time. (**D**) Position of CDR L3 at the PAD4 binding interface and rAbD workflow for designing L3 variants. Though the loop sits at the interface, it is not forming any favorable interactions with PAD4 due to its short length. High resolution structure of PAD4/hI365 interface served as the input model for rAbD. CDR L3 of the relaxed model was allowed to "Graft" and "Sequence" design or only "Sequence" design with or without flexible backbone design. Neighboring CDRs L1 and H3 were repacked to avoid clashes and optimize interface interactions. (**E**) "Graft" and "Sequence" design generated models (blue) with low totalscore and interface energies shown by dG_separated (in REU). "Sequence" design alone did not improve the interface with all models (orange) having poor dG_separated score. The dG_separated score of the cryo_EM structure of PAD4/hI365 interface is ~-54.4 REU (**F**) Rosetta antibody design (rAbD) predicted CDR L3 at lengths of 9 and 10 aa to have the most optimal interface energies. Different CDR L3 lengths and corresponding dG separated scores are shown. (**G**) SEC chromatagram showing poor performance of L3 mutants compared to parent (hI365) and control antibodies (hI356). (**H**) Fluorescent substrate activity assay showing inhibition of PAD4 in the presence of WT hI365 and the predicted L3 variants by rAbD. Error bars represent mean ± standard deviation of three biological replicates.

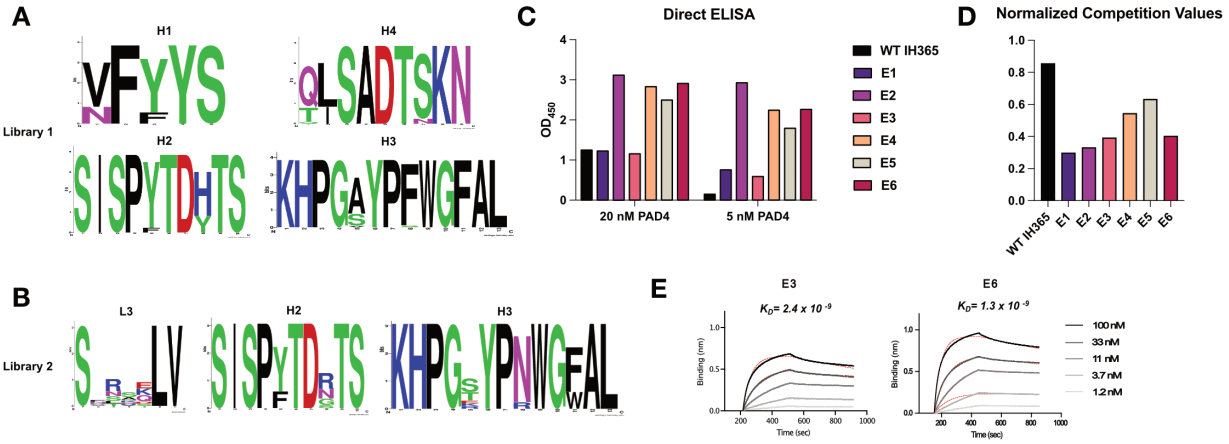


Figure 1.12: Soft randomization of hI365. (**A-B**) Sequence logos of library 1 and 2 showing amino acid preference for each CDR. (**C-D**) Binding of top Fab-phage clones to PAD4 as measured by direct and competition ELISAs. Top clones exhibit similar or improved binding to WT hI365 at both 20 nM and 5 nM PAD4. Top clones are less likely to be competed off PAD4, indicating a lower relative k_{off}.(**E**) Multi-point octet trace of leading clones, E3 and E6. Both appear to bind PAD4 at nanomolar affinity.

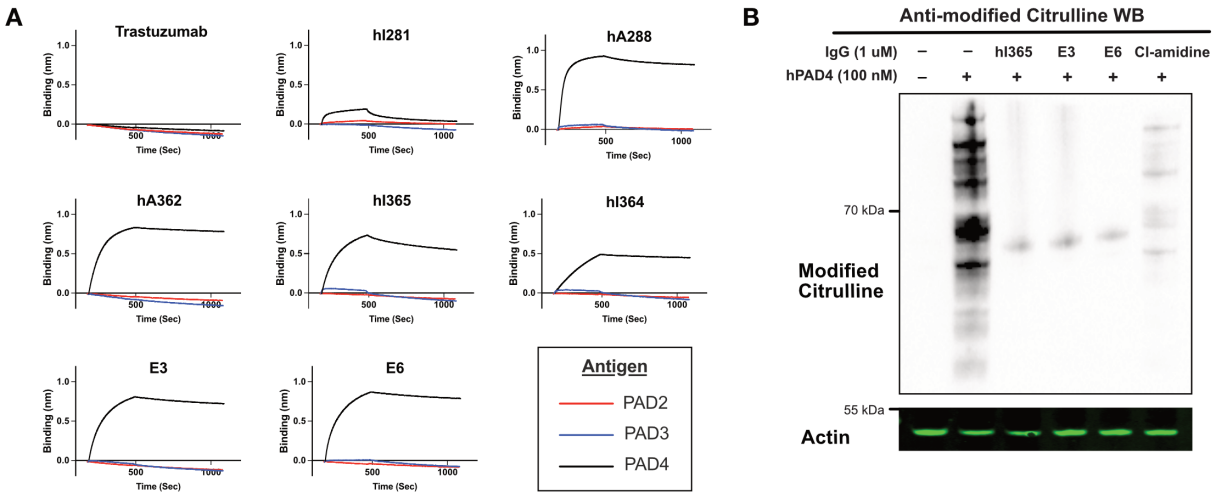


Figure 1.13: Anti-PAD4 antibody specificity and effect on PAD4 on whole cell lysate detect via anti-modified citrulline western blot. (A) Binding of hI365, E3, and E6 to PAD4, PAD2, PAD3, and Trastuzumab negative control measured by BLI. All PAD4 antibodies bind only to PAD4 and not PAD2 or PAD4. (**B**) Whole cell lysate treated with PAD4 +/- inhibitors. Functional antibodies, hI365, E3, and E6 inhibit PAD4-catalyzed citrullination on a variety of cytosolic, nuclear, and membrane-bound proteins. Trastuzumab IgG and pan-PAD inhibitor Cl-amidine included as negative and positive controls. An anti-actin loading control is included. Representative blots were reproduced three times before inclusion in main text.

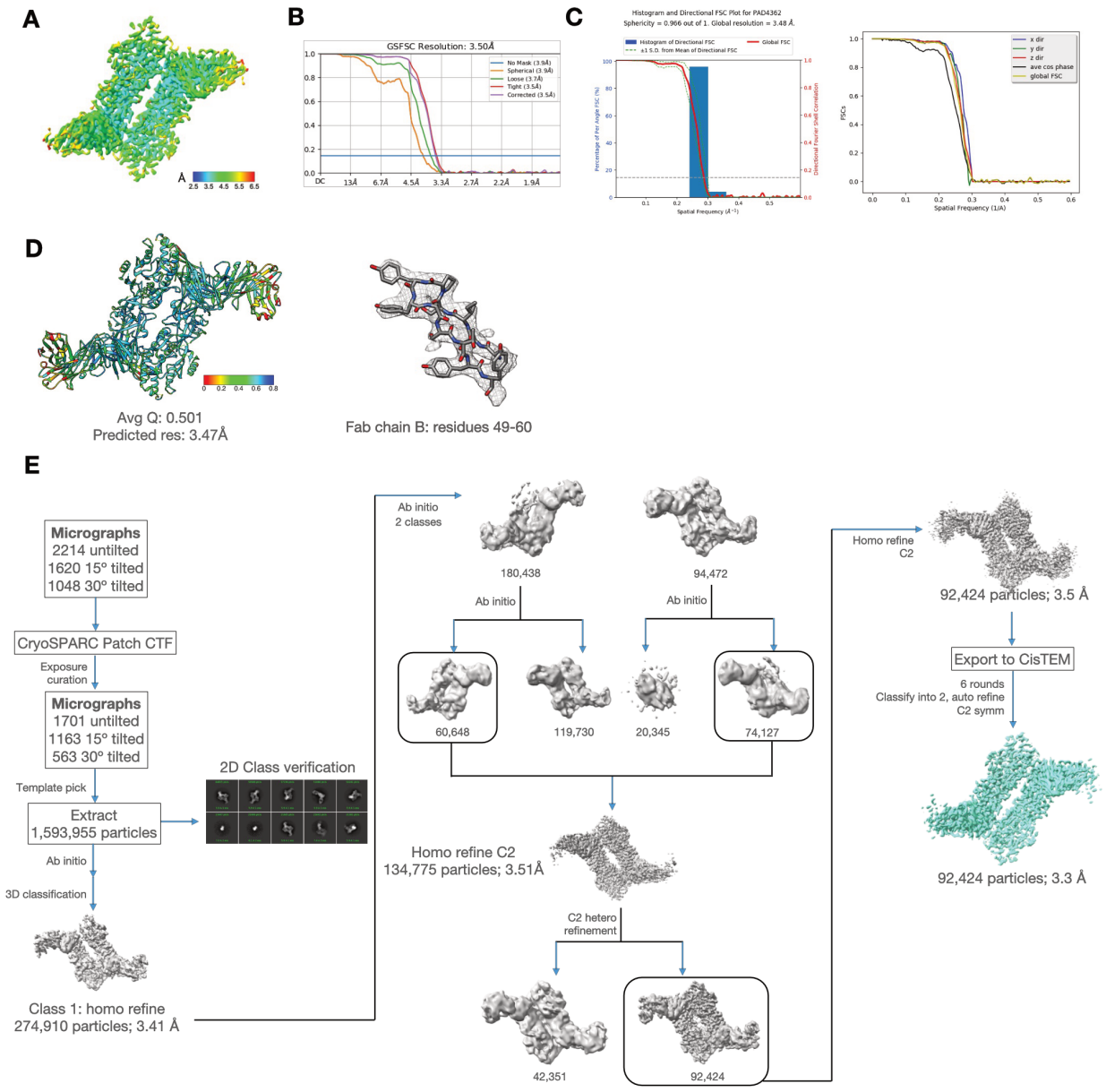


Figure 1.14: Cryo-EM map statistics for hPAD4/hA362. (A) Map colored by local resolution. **(B)** Global resolution FSC plots generated from half maps. Blue line indicates FSC=0.143 cutoff. **(C)** Directional FSC plots. Left: Histogram showing the percentage of per angle FSC vs resolution, showing agreement with global FSC and sphericity of 0.966. Right: Directional FSC curves. **(D)** Left: Model colored by per residue Q score; Right: Fab CDR loop showing model/map fit. PDB: 8SMK. (**E**) Cryo-EM data processing workflow for PAD4-hA362.

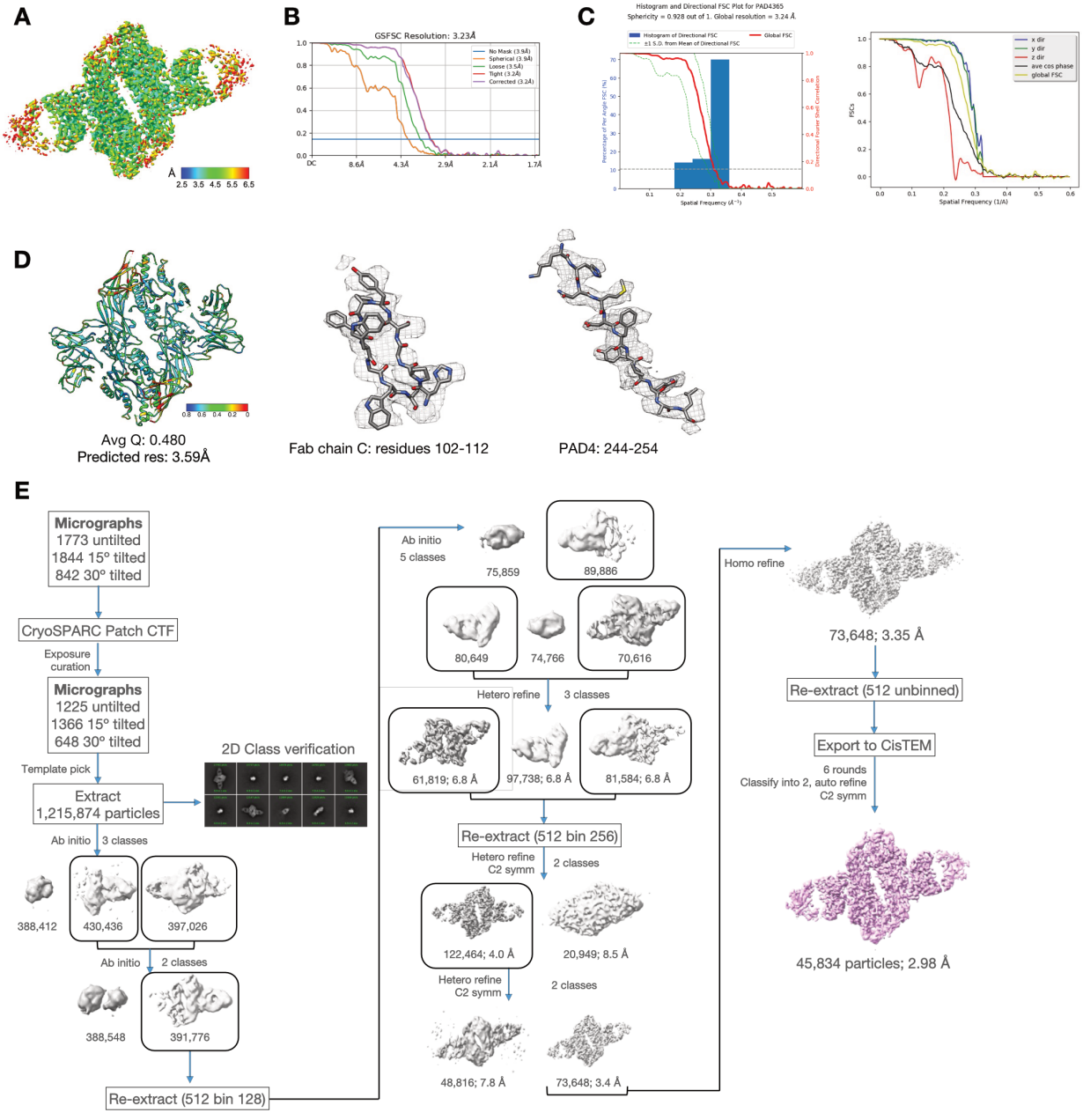


Figure 1.15: Cryo-EM map statistics for hPAD4/hI365. (A) Map colored by local resolution. **(B)** Global resolution FSC plots generated from half maps. Blue line indicates FSC=0.143 cutoff. **(C)** Directional FSC plots. Left: Histogram showing the percentage of per angle FSC vs resolution, showing some degree of anisotropy, but overall agreement with the global FSC and sphericity of 0.928. Right: Directional FSC curves demonstrating better than 3≈ resolution in two directions and ~4.5≈ in third direction. **(D)** Left: Model colored by per residue Q score; Right: Fab CDR loop showing model/map fit. PDB: 8SML. (**E**) Cryo-EM data processing workflow for PAD4-hI365.

Table 1.1 Summary table of all anti-human/mouse PAD4 antibodies detailing binding affinity, functional properties, mechanism of action, and associated PDB IDs.

Antibody	Target	Activity	Ca^{2+} dependency	Mechanism	Structural Analysis	Affinity (KD)	IC50	PDB ID
hI281	hPAD4	inhibitory	No	Promote monomerization	NSEM	-	-	-
hA288	hPAD4	activating	No	Promote dimerization	NSEM	-	-	-
hA362	hPAD4	activating	No	Promote dimerization	NSEM/ Cryo-EM	< 1 nM	-	PDB: 8SMK
hI364	hPAD4	inhibitory	Yes	-	-	< 1 nM	-	-
hI365	hPAD4	inhibitory	Yes	Disrupt Ca^{2+} and substrate binding	Cryo-EM	1.2 nM	Varied	PDB: 8SML
hI365-E3	hPAD4	inhibitory	Yes	Disrupt Ca^{2+} and substrate binding	-	2.3 nM	95 nM	-
hI365-E6	hPAD4	inhibitory	Yes	Disrupt Ca^{2+} and substrate binding	-	1.3 nM	13 nM	-
mA342	mPAD4	activating	Yes	-	-	< 1 nM	-	-
mA342-c4	mPAD4	activating	Yes	-	-	< 1 nM	-	-

Table 1.2: Cryo-EM data collection, processing, and refinement statistics.

	PAD4/hA362	PAD4/hI365
Data collection		
Grids	Quantifoil R1.2/1.3 Au 300 mesh	
Vitrification method	FEI Vitrobot	FEI Vitrobot
Microscope	Titan Krios	Titan Krios
Magnification	105,000x	105,000x
Voltage (kV)	300	300
Stage tilt (°)	0, (15, 30)	0, (15, 30)
Detector	K3	K3
Recording mode	Counting	Counting
Dose rate (e−/pix/sec)	8, (15)	8, (8)
Total electron exposure (e−/Å2)	70, (74)	77, (68)
Number of frames	140, (116)	140, (120)
Defocus range (μm)	-1 to -2	-1 to -2
Pixel size (Å)	0.835	0.835
Number of micrographs	4,882	4,559
Initial particle images (no.)	1,593,955	1,215,874
Data processing: C2 symmetry		
Final particle images (no.)	92,424	45,834
Symmetry	C2	C2
Map resolution (Å)	3.5	3.3
Refinement		
Initial model used (PDB code)	1WD9, 6OTC, 1N8Z	1WD9, 6OTC, 1N8Z
Symmetry	C2	C2
PDB code/EMDB code		
Model resolution (Å)	3.5	3.3
FSC threshold	0.143	0.143
Map sharpening B factor (Å2)	-67	-100
Model composition Non-hydrogen atoms	30762	24640
Protein residues	2000	1594
Ligands	10	6
B factors (Å2)		
Protein (min/max/mean)	0.00/578.92/95.83	37.44/357.03/79.88
Ligand	0.00/52.03/33.01	53.31/103.13/74.46
RMS deviations		
Bond lengths (Å)	0.011 (12)	0.010 (8)
Bond angles (°)	1.450 (8)	1.211 (4)
Validation		
MolProbity score	0.64	0.81
Clash score	0.39	0.85
Rotamer outliers (%)	0	0.14
Ramachandran plot		
Favored (%)	98.41	97.8
Allowed (%)	1.59	2.08
Disallowed (%)	0	0.13

Chapter 2

Targeted cell surface citrullination reveal novel PAD4 substrates

Abstract

PAD4-catalyzed citrullination of intracellular proteins has widespread implications in cellular activities, transcriptional regulation, and innate immunity. It's overexpression and hyperactivity are correlated with an increase of disease phenotypes, particularly in autoimmunity. Though PAD4 has primarily been studied either intracellularly or following inflammatory release into the extracellular space by neutrophils, there is increasing evidence of PAD4 localization at the cell surface. In this study, we build a tool to target PAD4 to HER2 expressing cells by genetically fusing PAD4 to ZHER2, an affibody with picomolar affinity for HER2. Using this tool, we study PAD4-catalyzed citrullination at the cell surface and describe a mass spectrometry proteomics method for confident identification of PAD4 substrates. Lastly, we investigated whether these pro-inflammatory post translational modifications may have an anti-tumor effect when installed on HER2-expressing breast cancer cells.

Introduction

Protein arginine deiminase 4 (PAD4) is a calcium-dependent enzyme that catalyzes the conversion of peptidyl arginine to peptidyl citrulline. This post translation modification changes the positively charged guanidium group on arginine to a neutrally charged urea group. Though citrullination results in only a 0.984 Da mass difference, the loss of a positive charge can have both significant structural and immunogenicity implications on substrate proteins.[56,57,58] The generation of citrullinated peptides and neo-epitopes is a hallmark in inflammation and a variety of autoimmune diseases. About two-thirds of the patient population in rheumatoid arthritis (RA) develop anti-citrullinated protein antibodies (ACPAs), and these patients tend to exhibit more severe disease phenotypes, suggesting the inflammatory nature of citrullinated neo-epitopes.[59]

The human proteome contains five PAD isoforms, PAD1-PAD4 and PAD6. Of all five, PAD4 is the only isoform that contains a nuclear localization signal (Stadler).[60] PAD2 has previously been reported to localize to the nucleus though it's mechanism of transport remains unclear.[61,62] PAD4 is also expressed exclusively by immune cells, specifically by monocytes and granulocytes. In a process of inflammatory neutrophil cell death called NETosis, neutrophils release their intracellular and nuclear contents, allowing PAD4 to enter the extracellular space and citrullinate extracellular targets.[4,5,6,7] These substrates are recognized by the immune system, which develops ACPAs to the neo-antigens and chronic inflammation.

Recent work from the Durrant group have shown that citrullinated peptides as immunogenic from RA studies can be leveraged to improve immune responses to cancer.[63,64,65,66,67] Citrullinated peptides have been shown to bind certain human leukocyte antigen (HLA) alleles with higher affinity than their non-citrullinated peptide counterparts, and recognition of these presented peptides by CD4+ T cells elicits an anti-tumor response. When injecting a select combination of citrullinated peptides into tumor-bearing mice, these vaccines have caused tumor regression and production of pro-inflammatory cytokines. These peptide vaccines are composed of citrullinated peptides from vimentin and enolase that are known targets of ACPAs in RA, but we believe that

there are a variety of PAD4 substrates at the cell surface that can be identified and leveraged for similar anti-tumor purposes.[68,69]

Here, we develop a tool to target and tether PAD4 to tumor cell surfaces, then use mass spectrometry proteomics to identify novel PAD4 substrates. We hypothesize that these substrates may be pro-inflammatory and follow up substrate identification with *in vivo* vaccination or tumor studies to investigate the effect of citrullination on immune response to tumor cells. These results show the impact of using protein tools to reveal novel cellular processes.

Results

Visualization of PAD4 on the cell surface

Though PAD4 intracellular activity has been well characterized, we sought out the implications of PAD4 activity at the cell surface. To a culture of murine EMT6 cells, we added 1 uM of soluble, recombinant PAD4 and imaged every 15 min for 4 hours, then hourly for anther 20 hours. Media was supplemented with 2 mM Ca^{2+} and reducing agent, TCEP, to promote PAD4 activity and stability (**Fig. 2.1A**). Compared to vehicle or catalytically inactive PAD4 D350A conditions, cells treated with active PAD4 enzyme exhibited decreased cellular eccentricity and delayed proliferation for the first 10 hours before returning to its normal morphology (**Fig. 2.1B**). Eccentricity is a measure of how much the cell deviates from being circular, meaning that PAD4 treated cells morphologically appear more circular than control cells. Though these two characteristics changed, cell viability remained similar at all time points during the experiment.

Lastly, we detected molecular events of citrullination using an anti-modified citrulline western blot (**Fig. 2.1C**). Since peptidyl arginine and peptidyl citrulline are structurally similar, a citrulline-specific probe is reacted with the sample on-membrane at low pH before detecting sing an anti-modified citrulline primary antibody.

Migration of citrullinated cells

PAD4 has been implicated in promoting cancer metastasis. Incucyte data showed that the morphology of PAD4-treated cancer cells mimic the morphology of cells treated with versene, a reagent commonly used in cell culture to lift adherent cells from cell culture flasks. To test the hypothesis that PAD4 treatment of cells leads to disruption of cell adherence, we used a transwell migration assay to show increased migration of citrullinated cells compared to non-citrullinated cells. T24 cells were plated in serum-free media and migration of into complete media, plus or minus PAD4, was measured.

Engineering of a PAD4 targeted to HER2 expressing cells

Modern targeted HER2 therapies typically involve an anti-HER2 antibody, namely Trastuzumab or Pertuzumab. These antibody scaffolds have been used to direct small molecule drugs, peptides, or even enzymes specifically to HER2 expressing cells.[70] Several attempts were made at forming a Trastuzumab IgG or Fab conjugate with PAD4, but genetic fusions suffered from low expression yields and chemical conjugates led to protein instability and aggregation. Additionally, PAD4 contains several free, surface-exposed cysteines and requires storage with reducing agent to prevent aggregation. At concentrations of 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), the inter-chain antibody disulfide bonds are broken. Though most antibodies still retain binding to their targets when disulfides bonds are broken, their stability drastically decreases. Due to this consideration, we decided against using antibodies to target PAD4 to the cell surface.

ZHER2 is a HER2 binding affibody protein derived from the Z domain of bacterial Protein A that exhibits picomolar binding affinity to HER2.[71] This three-helical bundle contains no disulfide bonds and has been used clinically as a PET imaging diagnostic with minimal immunogenicity considerations.[72,73] We engineered a ZHER2-PAD4 fusion by genetically linking ZHER2 to the N-terminus of PAD4, separated by a 10 amino acid glycine/serine linker (**Fig. 2.2A**). The fusion protein can be recombinantly expressed in *E.coli* and retains both binding to HER2 ectodomain

as measured by biolayer interferometry (BLI), and on-cell binding to HER2 by flow cytometry (**Fig. 2.2B,C**). PAD4 appears to inherently stick to cell surfaces at high concentrations (1 uM), but only ZHER2-PAD4 is targeted to HER2 expressing cells at concentrations as low as 10 nM. In non-HER2 expressing cells, PAD4 and ZHER2-PAD4 stick to cells in a similar fashion (**Fig. 2.3**).

ZHER2-PAD4 catalytic activity is also conserved compared to wild-type PAD4, as measured by detecting citrullination of histone H3, a canonical substrate of PAD4. Using the anti-modified citrulline western blot, we find that citrullination of HER2+ cell surface substrates is detected at lower concentrations (100 nM) of ZHER2-PAD4 compared to untargeted WT PAD4. On non-HER2 expressing cells, ZHER2-PAD4 and WT PAD4 only exhibit cell surface citrullination at high enzyme concentrations (1 uM) (**Fig. 2.1C**).

Confidently identifying sites of citrullination in a recombinant system using MS proteomics

We sought out to optimize a method to confidently identify citrullination events using a recombinant system containing only PAD4 and histone H3. PAD4 and histone H3 were incubated together at 37C for 90 minutes in the presence of 0.5 mM TCEP and 2 mM Ca^{2+} before the sample was trypsin-digested for proteomic analysis. PBS and PAD4 D350A control conditions were subjected to the same workflow.

Citrullination is a difficult post-translational modification to detect via mass spectrometry (MS). Several methods have been developed to mine for citrullination, including LC separations and specific anti-citrulline handles for enrichment of citrullinated peptides.[74] However, these methods are time consuming and require harsh conditions for peptide labelling.[75] An alternative strategy involves a combination of high energy collisional dissociation (HCD) and electron-transfer dissociation (ETD) to confidently identify sites of citrullination, as ETD can fully fragment peptide backbones for greater resolution, but comes at a time-cost.[76,77] HCD fragmentation of peptidyl citrulline results in spontaneous release of a 43 Da isocyanate group (**Fig. 2.3A**). When detected, this 43 Da fragment can trigger ETD scans of the corresponding peptide.[78] By manually inspecting

peptide spectra from ETD scans, b and y ions can only be observed when citrullination is defined as a modification (**Fig. 2.3B**). We used this method to map sites of citrullination onto H3 and obtained similar results to reports from literature.[77]

PAD4 is known to autocitrullinate, so we also mapped sites of citrullination onto PAD4 (**Fig. 2.3C**). These results are corroborated by far fewer citrullination events being detected in PBS or PAD4 D350A conditions.

Mining whole cell lysate for novel PAD4 targets

Following optimization of a method to confidently detect citrullination events, we moved to a more complex sample of whole cell lysate treated with PAD4. A hHER2+ expressing cell line was generated by transducing EMT6s murine breast cancer cells with a hHER2 expressing lentiviral vector. Cells were then lysed and treated with PBS, PAD4, PAD4 D350A or ZHER2-PAD4, then digested and fractionated offline by HPLC to ensure greater separation of peptides. Samples were then prepared for downstream proteomic analysis using a method called high-energy collision dissociation product dependent electron-transfer/higher-energy collision dissociation (HCD-pd-EThcD). In each sample, about 6,000 proteins were identified with close to 50,000 total peptides. In PAD4 active samples, we identified >3,000 unique citrullinated peptides, but in control samples, about 1,000 citrullinated peptides were still identified (**Fig. 2.4**).

False positives in the dataset were further filtered out by eliminating peptides that end in a terminal citrulline (**Fig. 2.5**). Though there is evidence that trypsin may still be able to cleave after citrulline, enzymatic efficiency is decreased >100 fold so we decided to eliminate all terminal citrullines.[79,80,81] Through this filter, we decreased the number of citrullinated peptides detected in negative control samples while minimally affecting our active PAD4 datasets. Of the top 150 proteins identified in each dataset, several membrane or membrane-associated proteins stood out to us, including HER2 and transferrin receptor (**Table 2.1**).

Humoral response to citrullinated cell surfaces

Since citrullination is known to be an immunogenic PTM, we investigated if installing citrulline onto tumor cell surfaces would improve immune responses to cancer *in vivo.* We chose EMT6s, a syngeneic mouse cell line in the middle of the spectrum between immunogenically hot and cold tumors.[82] This cell line has previously shown moderate immune cell infiltration, so we believe that generating immunogenic neo-epitopes could take advantage of already present tumor infiltration lymphocytes (TILs) to improve immune response.

As a proof-of-concept experiment, we targeted PAD4 to EMT6 cells *in vitro* to generate cell surface citrullination, then heat-shock killed the cells. Mild heat shock of cells has been shown to act as an adjuvant for T cell infiltration of tumors, so we decided to use this in combination with cell surface citrullination to create a whole cell vaccine and probe immune response.[83,84] BALB/c mice were vaccinated with citrullinated and heat-shocked whole cells on a weekly basis for four weeks on alternating flanks. Five mice were used for each condition, and at the end of four weeks, mice were sacrificed, and serum was isolated from the blood.

To test for serum response, cells were citrullinated *in vitro* and incubated with serum from vaccinated mice. Binding of mouse IgGs to citrullinated cells was detected via flow cytometry using a fluorescent anti-mouse IgG secondary (**Fig. 2.6**). In the group of mice vaccinated with citrullinated cells, several mice showed serum binding to citrullinated cells, while mice vaccinated with wild-type (WT) EMT6 or EMT6 cells treated with inactive PAD4 D350A did not exhibit serum binding to citrullinated cells. This indicates that mice vaccinated with citrullinated cells generated a humoral response specifically to the citrullinated neo-epitopes (**Fig. 2.7A**).

We then tested serum responses using an in-house developed ELISA, where plates were coated with BSA or citrullinated BSA, then we added serum and detected mouse IgG binding. Serum from mice immunized with citrullinated EMT6 seemed to bind stronger to citrullinated BSA than other conditions **(Fig. 2.7B**).

Anti-tumor response to citrullinated cell surfaces

Given the humoral response built towards cell surface citrullinated cells, we wanted to see if this humoral response translates to an anti-tumor effect. To test this, we conducted two experiments: (1) a tumor regression study and (2) a vaccinate-then-rechallenge study. For tumor regression, EMT6 hHER2+ cells were injected subcutaneously into BALB/c mice, then treated weekly with ZHER2-PAD4 for four weeks. Tumor volume and mice weight were tracked until the end of the study. Unfortunately, no significant response in tumor regression was observed between ZHER2-PAD4 treated mice and vehicle treated mice (**Fig. 2.8**).

For our vaccinate-then-rechallenge study, we again injected heat shocked and citrullinated whole cells into alternating flanks of mice on a weekly basis for four weeks to generate an immune response (**Fig. 2.9A**). Following vaccination, live, citrullinated EMT6 cells were implanted orthotopically into the mammary fat pad of mice and tumor growth rate was tracked. We hypothesized that mice that have been vaccinated with citrullinated cells may experience a slower tumor growth rate since their immune systems have been trained to attack citrullinated cells. However, no difference in tumor take rate was observed (**Fig. 2.9B**). This could be due to many extenuating factors, one being that syngeneic cells grow extremely aggressively when implanted orthotopically. It is possible that the fast-growing nature of the cells completely overshadowed the immune response developed against citrullinated cells.

Discussion

In this study, we develop a tool to investigate citrullination of cell surface substrates. Though canonically an intracellular protein, emerging evidence is beginning to show the significance of extracellular, cell surface PAD4. By targeting PAD4 to HER2+ expressing cells using an affibody-PAD4 fusion protein, we can efficiently install citrullination on the surface of cells. Following tethering of PAD4 to the cell surface, we used a combination of collisional and electron transfer dissociation to confidently identify substrates of PAD4, then further triaged for false positives and

cell surface substrates. Currently, we have generated a citrullinated protein data set form mouse whole cell lysate, but the ultimate goal would be to generate a dataset using surface proteins from a human cancer cell line. Substrates identified from our dataset can be further investigated for use as anti-cancer treatments or tools.

Though no anti-tumor effect was observed in mice vaccinated with citrullinated cancer cells, a mild humoral response developed. Serum analysis shows a specific response towards cell surface citrullination instead of PAD4 or non-citrullinated tumor cells. Citrullination is known to stimulate CD4+ responses, not CD8+ responses. Perhaps in this scenario, CD8+ cells need to be activated to see tumor regression and further investigation is needed on citrullination-induced lymphocyte activation. Furthermore, combination therapies including immune checkpoint inhibitors or B cell activators can be explored to further drive an anti-tumor effect against citrullinated cells.

Methods

Expression and purification of PAD4

C43 (DE3) Pro+, BL21 Gold (DE3), or BL21 ClearColi *E. coli* containing PAD4 expression vectors were grown in 2xYT at 37 °C to an OD-600 of 0.4–0.8 and then protein expression was induced by the addition of 0.5–1.0 mM IPTG. Incubation temperature was subsequently reduced to 18°C and the cultures were allowed to shake for 16–20 h. Cells were harvested by centrifugation and lysed using sonication. The lysate was centrifuged to remove inclusion bodies. The enzymes were purified by Ni-NTA resin with 0.5 mM TCEP supplemented to all buffers to prevent PAD4 oxidation. The purified enzyme was buffer exchanged to 50 mM Tris (pH 8), 400 mM sodium chloride, and supplemented with 0.5 mM TCEP. Purification steps were performed on ice to maintain high PAD4 enzymatic activity. Purified enzyme was aliquoted and flash frozen.

Citrullinated histone H3 PAD4 activity assay

PAD4 activities in the absence or presence of antibodies were also assessed using a citrullinated histone H3 Western assay. 10-100 nM recombinant PAD4 were mixed with antibodies with various concentrations of calcium at 4°C for 45 min, and then incubated with 760 nM recombinant histone H3.1 (New England Biolabs). The reaction was incubated at 37°C for 110 min, followed by western analysis using an anti-citrullinated H3 primary antibody (Abcam Ab5103) and an anti-rabbit HRP secondary antibody. The reactions were then incubated at 37°C for 110 min and subsequently analyzed by western blot using an anti-citrullinated H3 antibody (Abcam, Ab5103) and anti-rabbit HRP secondary antibody. Images were acquired in Image Lab (v5.0) and processed with Image Studio Software (v5.2). IC_{50} measurements were obtained with technical triplicates and quantified using Fiji.[47]

Modified citrulline western blot assay

Lysate was harvested from Expi293T cells using a 1x RIPA + protease inhibitor solution. Antibody-PAD4 complexes were pre-formed at 4°C for 45 min. The complexes were then incubated at 37°C for 110 min and subsequently analyzed by western blot using an anti-modified citrulline detection kit (EMD Millipore). PVDF membrane was incubated with anti-citrulline probe overnight, then blocked and stained with a primary anti-modified citrulline antibody and secondary HRP linked anti-IgG antibody. HRP signal was detected using a BioRad ChemiDoc imager.

Biolayer interferometry

BLI measurements were made using an Octet RED384 (ForteBio) instrument. Biotinylated PAD4 was immobilized on optically transparent streptavidin biosensors (ForteBio) and loaded until a 1 nm signal was achieved. After blocking with 10 μM biotin, purified binders in solution were used as the analyte. TBSTB was used for all buffers. Data were analyzed using the ForteBio Octet

analysis software, and kinetic parameters were determined using a 1:1 monovalent binding model (https://www.sartorius.com/en/products/protein-analysis/octet-bli-detection/octet-systems-software).

On-cell binding of PAD4 by flow cytometry

Adherent cells were lifted from culture with Versene (Fisher Scientific) at 37 °C, harvested by centrifugation (500 rcf), and resuspended in PBS in the presence of 1 uM PAD4. Cells were incubated at 4°C for 30 min, washed three times with cold 1% BSA in PBS (500 rcf, 5 min), and incubated on ice for 30 min with an anti-PAD4 primary antibody (Abcam) at a 1:500 dilution in 1% BSA in PBS. Cells were washed three times with cold 1% BSA in PBS (500 rcf, 5 min) then incubated with anti-rabbit IgG AlexaFluor 647 conjugated secondary antibody (Abcam) and incubated for 30 min at 4°C. Cells were washed three times with cold 1% BSA in PBS then incubated for 15 min with Propidium Iodide Ready Flow (Thermo Scientific) live/dead cell stain. Cells were analyzed on a Beckman Coulter CytoFLEX flow cytometer.

On-cell activity of PAD4 and ZHER2-PAD4 fusion proteins

EMT-6 hHER2+ cells were harvested with Versene and harvested by centrifugation (500 rcf) and resuspended in DMEM with 2 mM Ca^{2+} and 0.5 mM TCEP supplemented. Cells were incubated for 60 min at 37 °C, rotating, in the presence of PAD4 fusion proteins, inactive enzyme controls, or vehicle. Cells were then washed three times with PBS and fractionated and lysed using a subcellular fractionation kit. Protease inhibitor (Sigma) and benzonase nuclease (Sigma Aldrich) were added to the lysate buffer. Membrane fractions were isolated and clarified by centrifugation (21000 rcf, 15 min) and protein concentration was quantitated by RapidGold BCA (Thermo Scientific). Normalized concentrations of lysates were separated by SDS-PAGE and transferred to PVDF membranes, then blotted for citrullination using the modified citrulline western blot kit (described above).

Mouse immunization experiments

EMT-6 hHER2+ cells were harvested with Versene and harvested by centrifugation (500 rcf) and resuspended in DMEM with 2 mM Ca^{2+} and 0.5 mM TCEP supplemented. Cells were incubated for 60 min at 37 °C, rotating, in the presence of PAD4 fusion proteins, inactive enzyme controls, or vehicle. Cells were then washed three times with PBS and then resuspended at 1e7 cell per mL in PBS and killed by heat-shock at 47 °C for 1 hour with shaking at 900 rpm. Heat-killed lysates were immediately flash-frozen until use.

BALB/c mice received subcutaneous injections of 100 µL heat-killed lysate in alternating flanks weekly for four weeks. At the end of the four weeks, mice were euthanized according to standard protocols per the Institutional Animal Care and Use Committee at the University of California, San Francisco. Blood and spleens were harvested for later use.

Serum was separated from whole blood by centrifugation. Serum was then aliquoted and flash-frozen in liquid nitrogen. Frozen aliquots were stored at -80 °C.

Splenocytes were harvested by homogenizing spleens using a 45 µm nylon mesh filter and the rubber end of a 3 mL syringe plunger in 2 mL PBS. The dissociated cells were harvested by centrifugation (600 rcf, 5 min). Pellets were resuspended in 2 mL ACK lysis buffer (Fisher Scientific) and incubated at room temperature for 3 min, at which point 10 mL PBS was added to neutralize the lysis. Cells were pelleted by centrifugation (600 rcf, 5 min), washed with 10 mL PBS again, and resuspended in 1 mL cryopreservation medium (Biolife) for gradual cooling and long-term storage in liquid nitrogen vapor phase.

Serum flow cytometry

EMT-6 hHER2+ cells were harvested with Versene and harvested by centrifugation (500 rcf) and resuspended in DMEM with 2 mM Ca^{2+} and 0.5 mM TCEP supplemented. Cells were incubated for 60 min at 37 °C, rotating, in the presence of PAD4 fusion proteins, inactive enzyme controls, or vehicle. Cells were washed three times with cold 1% BSA in PBS (500 rcf, 5 min) and incubated

on ice for 30 min with diluted serum (1:50) from immunized animals. Cells were washed three times with cold 1% BSA in PBS (500 rcf, 5 min) and incubated on ice for 30 min with goat anti-mouse IgG AlexaFluor 647 conjugate (Thermo) at a 1:200 dilution in 1% BSA in PBS. Cells were washed two times with cold 1% BSA in PBS (500 rcf, 5 min), once with PBS (500 rcf, 5 min), and incubated for 15 min in PBS with Propidium Iodide Ready Flow (Thermo Scientific) live/dead cell stain. Cells were analyzed on a Beckman Coulter CytoFLEX flow cytometer.

Mouse immunization then challenge experiments

Mice were immunized as before. EMT-6 hHER2+ cells were harvested with Versene and harvested by centrifugation (500 rcf) and resuspended in DMEM with 2 mM Ca^{2+} and 0.5 mM TCEP supplemented. Cells were incubated for 60 min at 37 °C, rotating, in the presence of PAD4 fusion proteins, inactive enzyme controls, or vehicle. Live cells were then resuspended at 1e7 cell per mL in PBS and injected subcutaneously into the flanks of immunized or naïve BALB/c mice. Tumor growth and take rate was monitored periodically.

Preparation of samples for mass spectrometry proteomics

Recombinant histone H3.1 protein and hPAD4 were incubated at 37 °C with TCEP and Ca^{2+} to generate citrullinated protein. Samples were then alkylated and digested on-bead using Preomics iST 96x kits (Preomics) and purified according to manufacturer's protocols. Peptides were dried *in vacuo* and resuspended with 0.2% formic acid in before quantification with a Pierce Colorimetric Peptide Quantification Kit (Thermo). For cell lysate experiments, EMT6 and EMT6 hHER2+ cells were grown to confluency and lifted with Versene. Cells were washed once with ice cold PBS, then lysed with RIPA buffer supplemented with protease inhibitor. Lysates were spun down at max speed for 10

minutes at 4 °C and finally treated with PAD4 to generate citrullinated proteins in the lysate. Lysates were prepared for mass spectrometry analysis as described above.

Mass spectrometry proteomics of citrullinated samples

Samples were analyzed using a Vanquish HPLC system coupled to an Orbitrap Eclipse mass spectrometer (Thermo). Resuspended peptide samples were loaded onto a 60 cm long 75 μm inner diameter column. Mobile phase A was composed of water and 0.2% formic acid (FA) while mobile phase B was composed of 80% ACN and 0.2% FA. Separation was performed using a gradient elution of 7-48% mobile phase B over 65 minutes followed by 48-60% mobile phase B over 7 minutes and a flus of 99% mobile phase B for 24 minutes. Flow rate was kept at 0.3 μL/min throughout. Survey scans of peptide precursors from 350 – 2000 *m/z* were performed at a resolving power of 60k with an AGC target of 1×10^6, and maximum injection time of 50 ms. For the HCD-pd-EThcD method, HCD product ions of *m/z* 43.0058 were listed to trigger the subsequent EThcD fragmentation. The HCD scans were set to an ASG target of 1.25×10^5, a resolving power of 30K, and maximum injection time of 100 ms. The ETD reaction time was set to an ASG Target of 1.25×10^5, a resolving power of 60K, and a maximum injection time of 118 ms.

Mass spectrometry data analysis

Raw sample files were analyzed in FragPipe (v21.1) against a murine proteome plus human PAD4 and HER2 with trypsin selected as the enzyme and 4 missed cleavages allowed. Cysteine carbamidomethylation (+57.02146) was chosen as a fixed modification, while methionine oxidation residues (+15.9949 Da), N-terminal acetylation (+42.0106 Da), and arginine citrullination (+0.984 Da) were chosen as variable modifications. Peptides identified to end in a C-terminal citrulline residue were filtered out of the dataset. All other parameters were set as default.

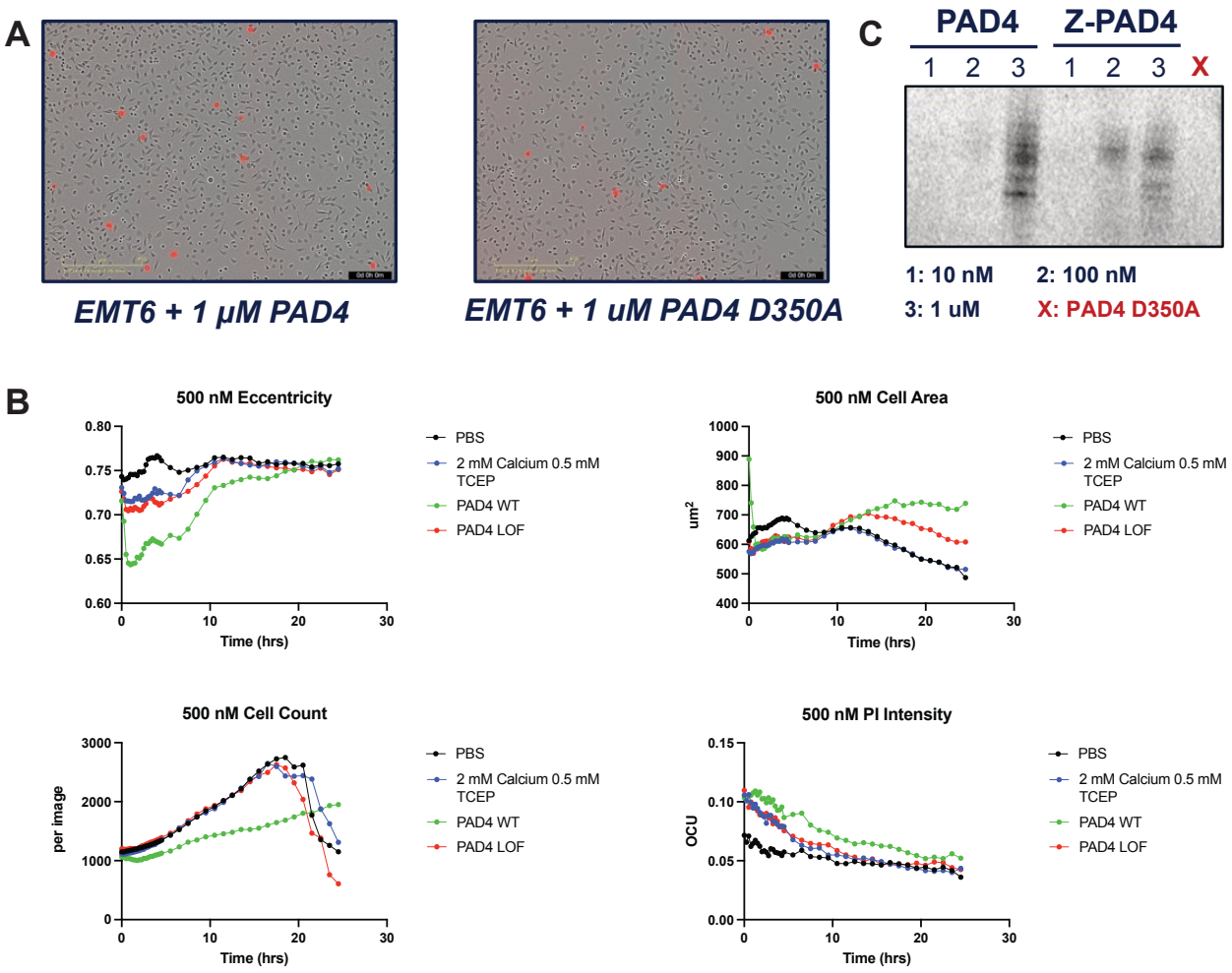


Figure 2.1: Treatment of EMT6 cells with active PAD4 changes cell morphology while retaining viability. (**A**) Time-lapse videos showing 24 hour treatment of cells with active or inactive PAD4. (**B**) Incucyte-generated plots of cell eccentricity, cell area, cell count, and live-dead stain. More eccentricity is defined as a less spherical/circular cell shape. (**C**) A cell-tethered variant of PAD4 is more active than parental enzyme on cell surfaces measured by anti-modified citrulline western blot.

Figures and Tables

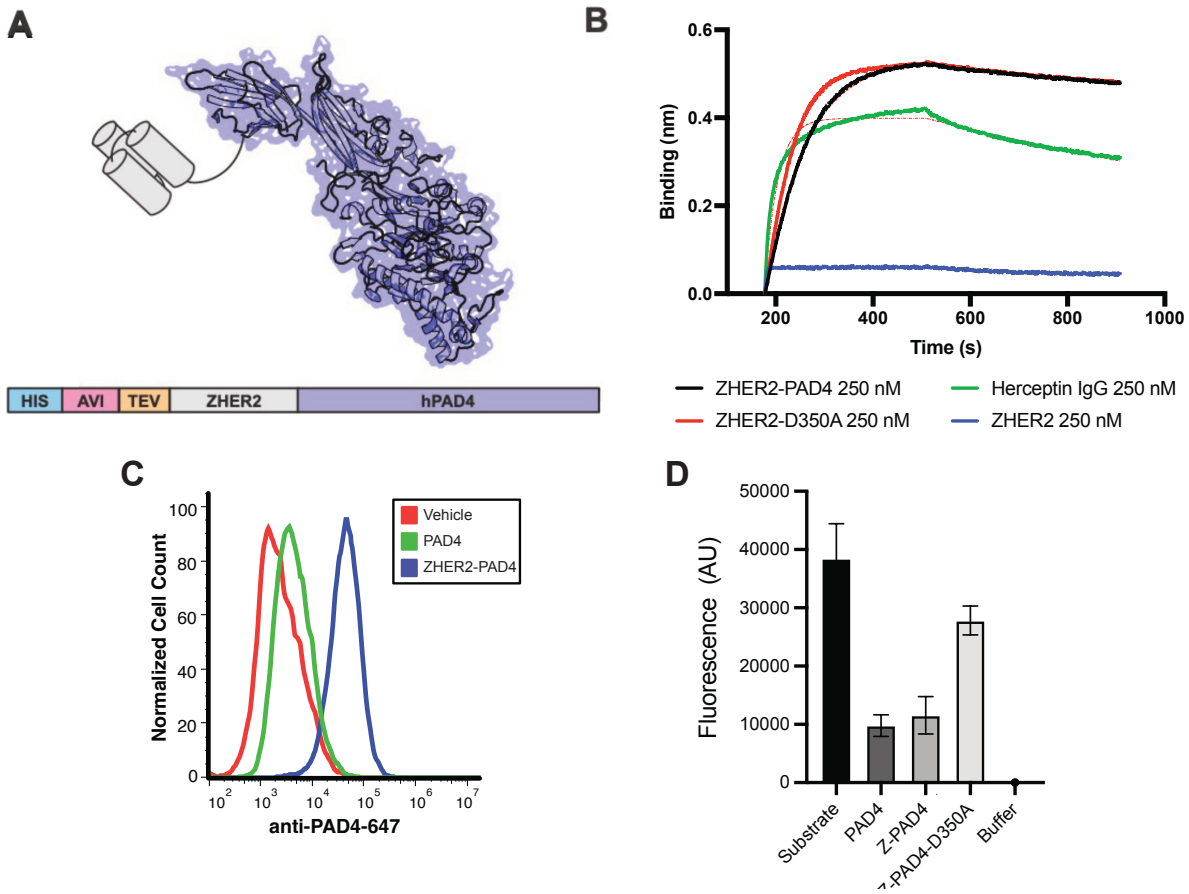


Figure 2.2: An engineered form of PAD4 is active on the cell surface. (**A**) Schematic representation of ZHER2-PAD4 conjugate with N-terminal His and Avi tag separated by a TEV cleavage site. (**B**) Binding of ZHER-PAD4 to HER2 ectodomain measured via BLI. (**C**) and (**D**) Binding of ZHER2-PAD4 to EMT6 HER2+ cell surfaces at 100 nM measured via flow cytometry.

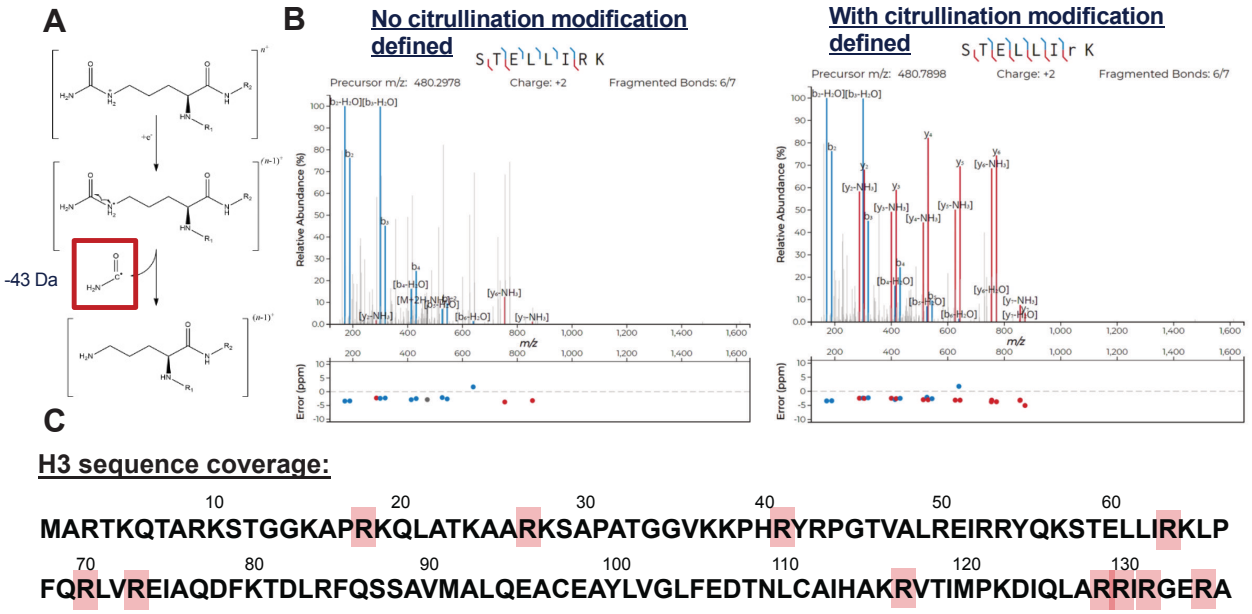


Figure 2.3: Piloting HCD-pd-EThcD workflow using a recombinant PAD4 and histone H3 system. (**A**) Neutral loss of isocyanate ion during HCD triggers product dependent EthcD for peptide backbone fragmentation. (**B**) Spectral searches performed on a histone H3 peptide (amino acids 58-65) with no citrullination modification defined (left) and with citrullination defined (right). B and y ions only clearly observed with citrullination defined. (**C**) Sequence map of histone H3 with arginine residues where citrullination is detected highlighted.

A

Sample	Total Proteins	Total Peptides	Unique cit. peptides
PBS	6670	51832	1690
PAD4 WT	6566	50415	3833
PAD4 D350A	6531	49089	1466
ZHER2-PAD4	5326	32401	3181

B

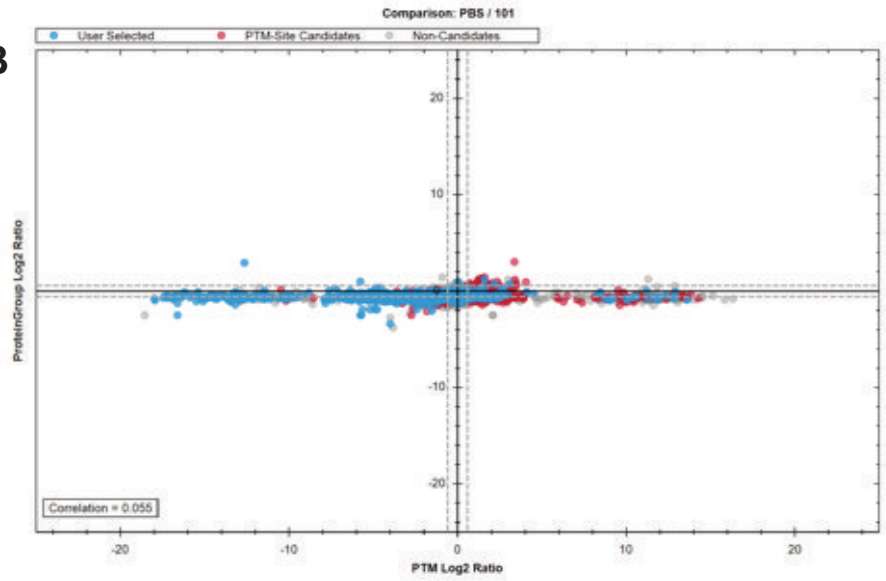


Figure 2.4: Results from off-line fractionation and LC/MS/MS of whole cell lysate samples treated with PAD4. (**A**) Number of unique proteins and peptides detected in mass spectrometry data set following off-line high pH reverse-phase separation. (**B**) Change in protein level (y-axis) vs. citrulline PTM level (x-axis) between PBS and Z-PAD4 treated cells.

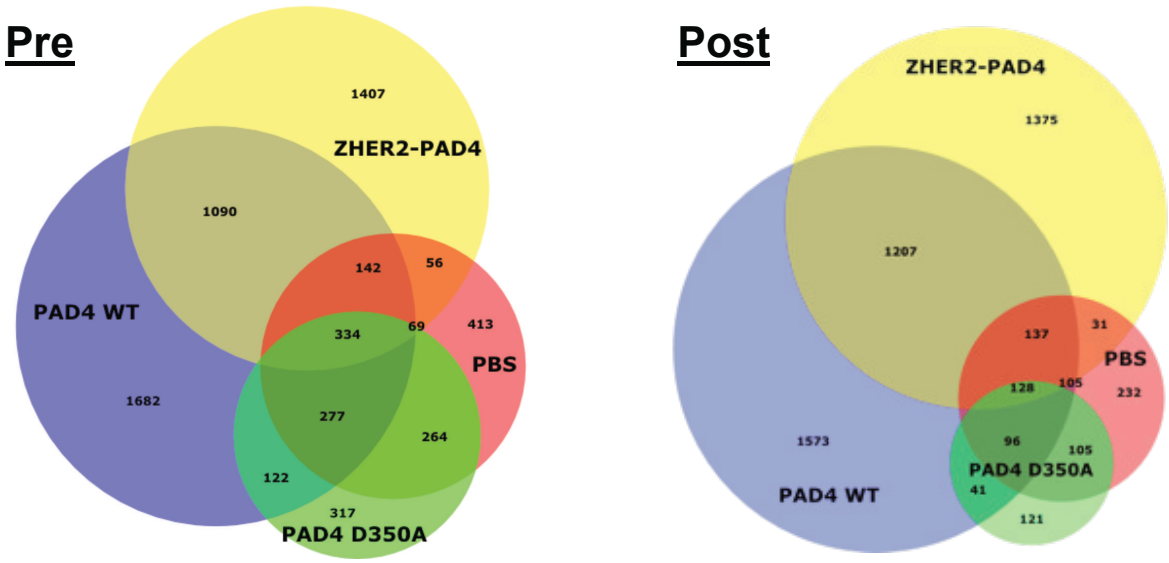


Figure 2.5: Pre (left) and post (right) filtering of peptide citrullination hits. Peptides ending in citrulline were discarded from data set.

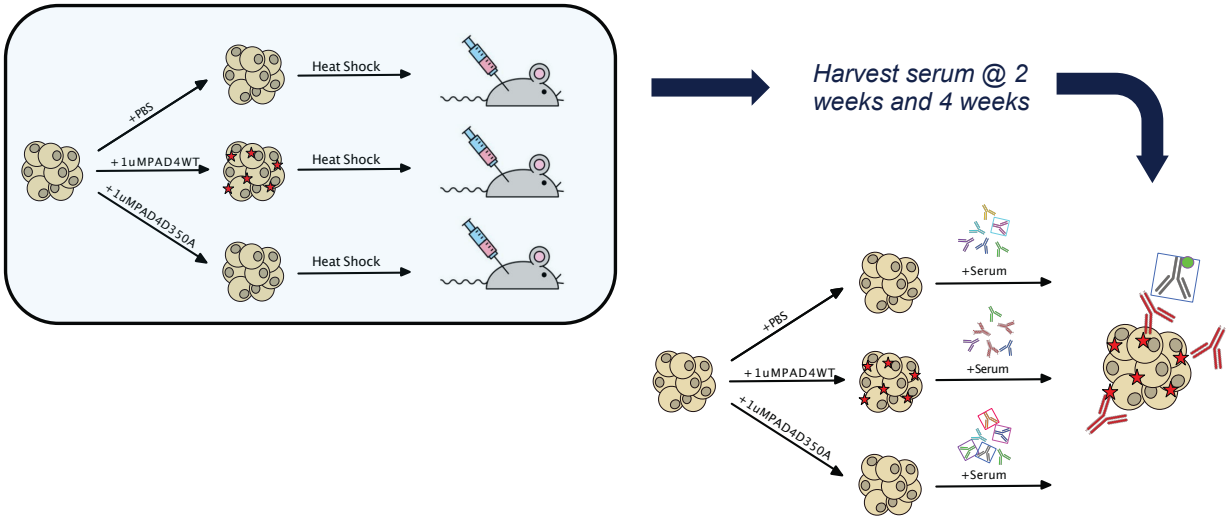


Figure 2.6: Schematic of mouse immunization studies and downstream serum flow analysis. EMT6 cells were pooled into 3 groups and treated with PBS, PAD4, or PAD4 LOF, then heat shocked and injected into mice for four weeks on a weekly basis. Serum was harvested after 4 weeks and subjected to flow cytometry over PBS/PAD4/PAD4 D350A treated cells to detect binding of mouse IgGs.

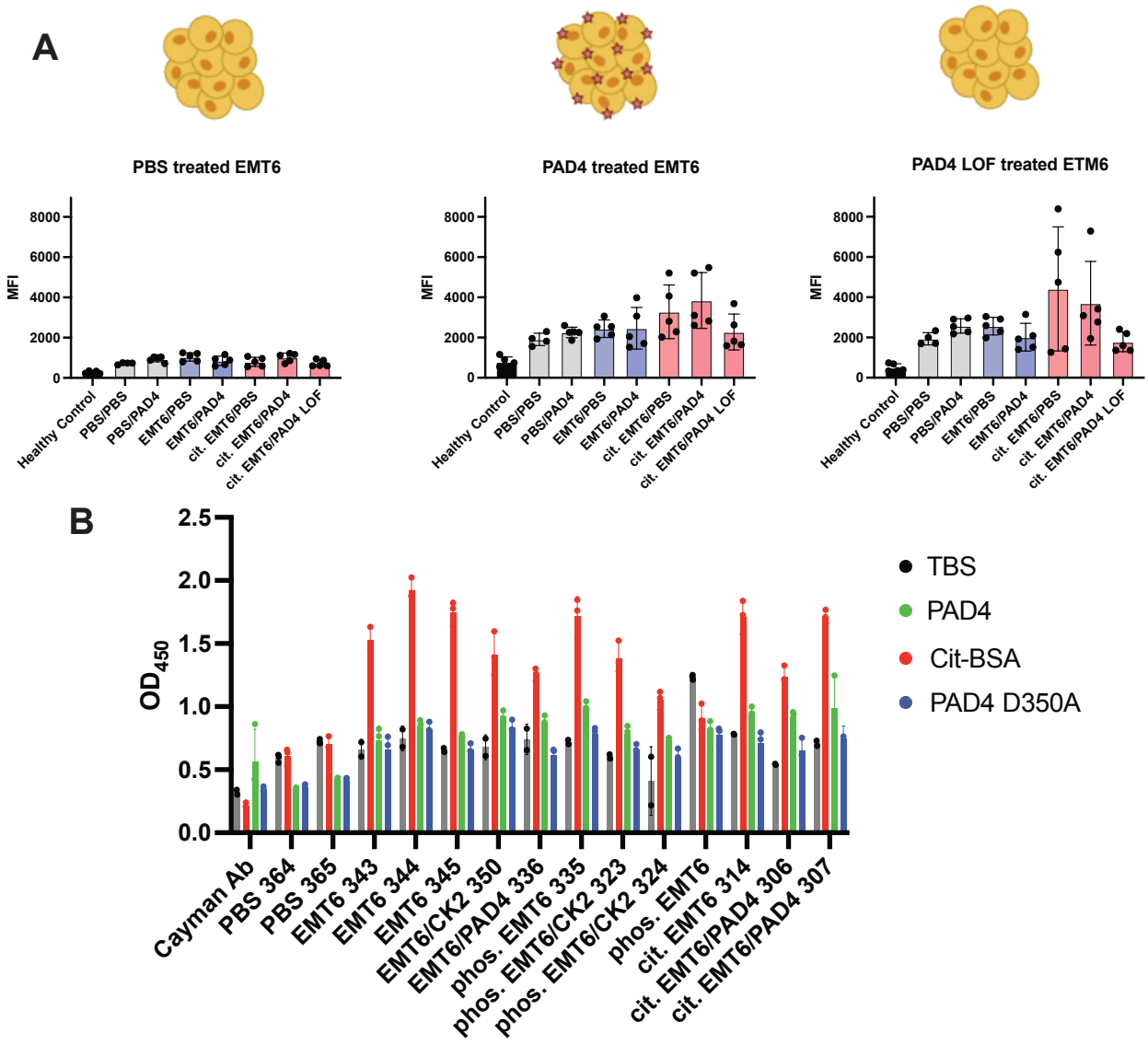


Figure 2.7: Serum harvested from immunized mice show modest response to citrullinated EMT6 cells. (**A**) Flow cytometry analysis of mouse IgG from serum binding to citrullinated EMT6 cells. (**B**) Serum binding ELISA shows elevated response to citrullinated BSA over non-citrullinated BSA in all mouse-immunized conditions.

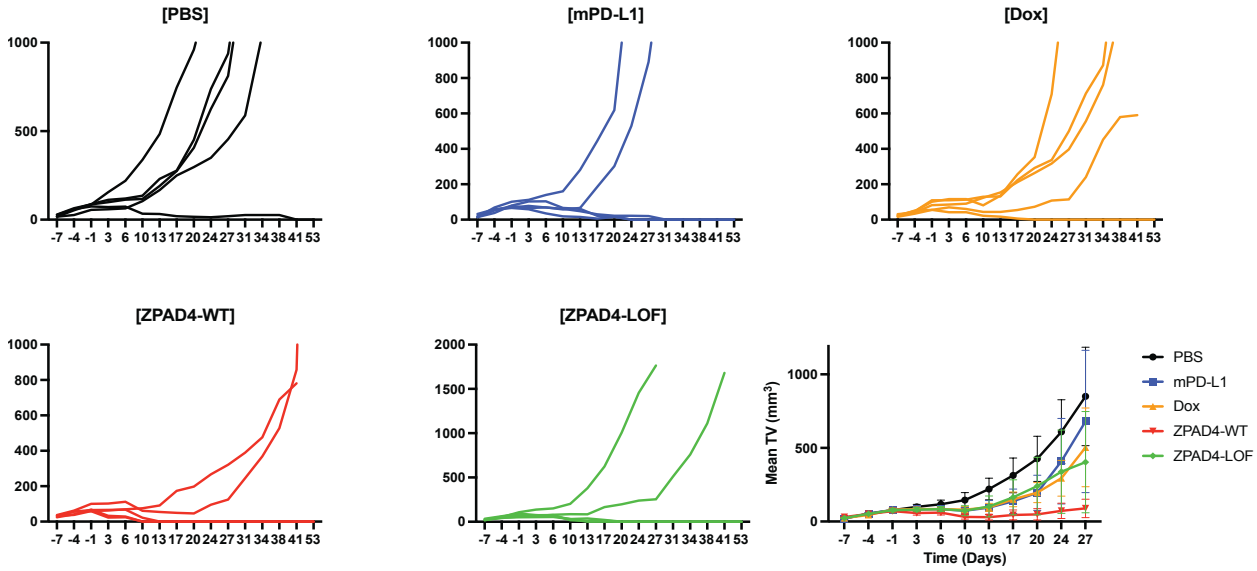


Figure 2.8: Tumor regression studies of mice implanted with EMT6 hHER2+ cells. Mice subcutaneously implanted with EMT6 hER2+ cells were treated with PBS, mPD-L1, Dox, ZPAD4-WT, or ZPAD4-LOF. No significant difference in tumor regression was observed across conditions.

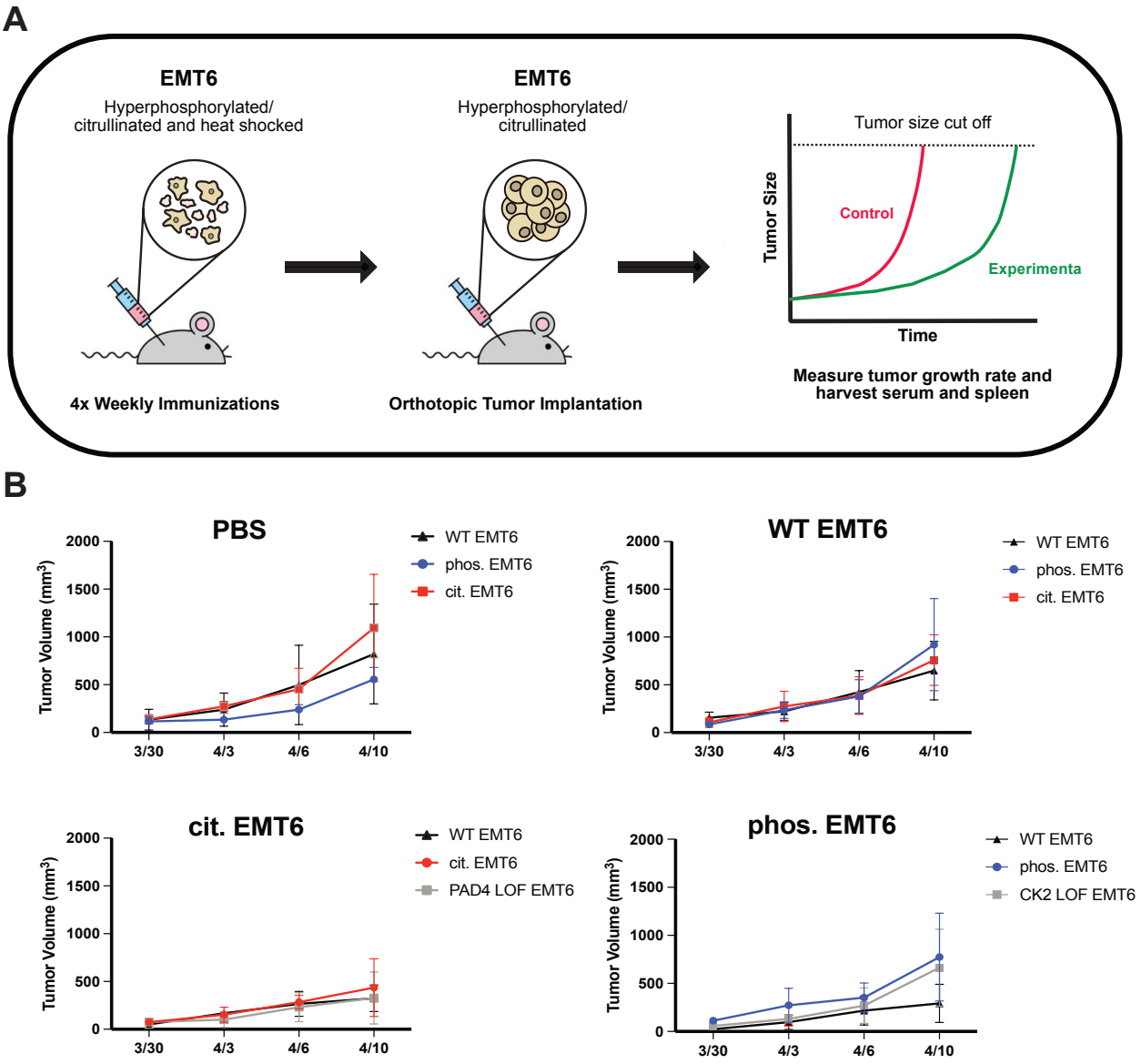


Figure 2.9: No anti-tumor effect observed in mice immunized and implanted with citrullinated EMT6 cells. (**A**) Schematic of mouse vaccinate then re-challenge studies. (**B**) Tumor growth rate of mice treated with various treated EMT6 cells. No difference in growth rate observed.

Table 2.1: Top citrullinated protein hits identified from proteomic analysis and their canonical cellular location.

PAD4

ERBB2	Erb-b2 receptor tyrosine kinase 2	Cytosol/membrane
Tnks1bp1	Tankyrase 1 binding protein 1	Cytosol/membrane
Pdap1	PDGFA interacting protein 1	Cytosol/membrane
Rcc2	Regulator of chromosome condensation 2	Nuclear
Ccdc124	Coiled-coil domain containing 124	Cytosol/membrane
Dpysl2	Dihydropyrimidinase like 2	Cytosol/membrane
Cavin1	Caveolae associated protein 1	Cytosol/membrane
Rpsa	Ribosomal protein SA	Cytosol/membrane
Cttn	Cortactin	Cytosol/membrane

ZHER2-PAD4

ERBB2	Erb-b2 receptor tyrosine kinase 2	Cytosol/membrane
Aip	Aryl hydrocarbon receptor interacting protein	Cytosol
Nedd4	NEDD4 E3 ubiquitin protein ligase	Cytosol
Tnks1bp1	Tankyrase 1 binding protein 1	Cytosol/membrane
Pdap1	PDGFA interacting protein 1	Cytosol/membrane
Atp2a2	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 2	Cytosol
Cttn	Cortactin	Cytosol/membrane
Tfrc	Transferrin receptor	Cytosol/membrane/endosomes/lysosomes
Nisch	Nischarin	Nucleoplasm/cytosol
Eps8l2	EPS8 like 2	Cytosol
Ccdc124	Coiled-coil domain containing 124	Cytosol/membrane

Chapter 3

Development of an antibody-drug conjugate for targeted degradation of challenging extracellular targets

Abstract

Targeted protein degradation has emerged as a powerful strategy for eliminating problematic protein targets and offers several advantages over traditional protein inhibition. The first class of degraders used heterobifunctional small molecules that could pass the cell membrane to target cytoplasmic or nuclear proteins. These molecules would recruit intracellular E3 ligases to target a protein of interest to the proteosome. Recently, various approaches have been described to include extracellular targets into the "degradome."[85] These technologies include LYTACs, AbTACs, sweeping antibodies, KineTACs, and more. All current approaches use antibodies to recruit the extracellular protein of interest, but there are several classes of proteins that the field has struggled to develop antibodies against. Multi-pass membrane proteins like GPCRs typically lack a large, structured ectodomain, and have traditionally been difficult to target with antibody binders. Despite the lack of antibody therapies, there have been a wealth of small molecule binders described for various multi-pass membrane proteins. Here, we describe a novel approach, termed antibody-drug conjugate PROTACs (ADC-TACs), that utilizes a small molecule binder targeting challenging extracellular proteins conjugated to a Fab targeting an E3 ligase or internalizing receptor. This strategy enables the targeted degradation of multi-pass membrane proteins. Simultaneous engagement of the membrane protein and the E3 ligase leads to lysosomal degradation of the target protein. We demonstrate proof-of-concept experiments for ADC-TACs targeting the G protein-coupled receptor adenosine 2a receptor (A2AR) and explore alternative applications of our ADC-TACs, including in a bifunctional T cell engager system. Overall, this strategy is generalizable to targeting various classes of multi-pass membrane proteins and other membrane proteins lacking antibody binders.

Introduction

Targeted protein degradation (TPD) has emerged as an effective strategy to remove problematic proteins from cells in an efficient and sustained manner. Proteolysis targeting chimeras (PROTACs), the term for molecules targeting proteins for degradation, simultaneously engage a protein of interest (POI) and E3 ligase to recruit the cell's native degradation machinery towards problematic proteins.[86] This strategy offers several advantages compared to drugs that act on targets through traditional protein inhibition. While traditional inhibitors rely on occupying the active site of an enzyme, degraders can bind any epitope on the POI for trafficking to the proteasome for degradation. These molecules also work catalytically, meaning each PROTAC can target multiple POIs for degradation, allowing for a lower dose to achieve the same therapeutic effect as traditional inhibitors and first class PROTACs have now progressed to the clinic.[87]

While the proteosome is responsible for degrading most intracellular proteins, extracellular proteins can undergo either proteosome or lysosome-based degradation. In recent years, both these pathways have been exploited for extracellular TPD: AbTACs and REULR recruit cell surface E3 ligases for proteasomal degradation of POIs, while LYTACs, KineTACs, and TransTACs employ internalizing receptors for lysosomal degradation.[88,89,90,91,92] However, while there have been major advances in the field of extracellular TPD in terms of identifying new "degrader" arms, the POI targeting arm has always employed an antibody or antibody-based fragment for protein engagement. Though antibodies offer high specificity, their binding is reliant on the antigen containing a large, ordered, extracellular epitope for antibody recognition. This presents a challenge when targeting POIs like G-protein coupled receptors (GPCRs) that notably have a small ectodomain.[93]

While antibodies struggle to recognize GPCRs, many of these targets have well validating, existing, even FDA approved small molecule binders. By chemically linking these small molecules onto existing antibody degraders, we can use existing chemistry to expand the scope of

extracellular degradation targets. Here, we describe the engineering and optimization of antibody-drug conjugate PROTACs (ADC-TACs) and show preliminary data for degradation of adenosine receptor 2A (A2AR). We follow this up by formatting our small molecules into a bispecific T cell engager to recruit T cells to tumors bearing A2AR. This general strategy allows for the conversion of any small molecule against an extracellular target to be converted into an ADC-TAC degrader, expanding the range of proteins available for degradation at the cell surface.

Results

A2AR is a GPCR that responds to and is activated by adenosine in the extracellular space. Overactivity of A2AR is a common feature of various types of cancer due to high levels of ATP and adenosine in the tumor microenvironment (TME).[94] Downstream A2AR signaling may contribute to cancer by (1) activating the MAPK/ERK/JNK cell proliferation pathway and (2) leading to suppression of immune cells through buildup of cAMP. A2AR has been a well-established target in oncology and has been targeted by small molecule in the past.[95,96] One molecule, CGS21680, is an agonist of A2AR and has previously been used for bioconjugation onto an antibody Fc domain.[97] When bound to A2AR, CGS21680 contains a solvent-exposed site where a carboxylic acid reactive handle could be installed for downstream bioconjugation without affecting molecule binding.[98]

A two-step bioconjugation strategy to link CGS21680 to the scaffold of a Fab targeting RNF43 E3 ligase was devised using methionine-specific chemistry. RNF43 Fab were conjugated with oxaziridine-azide, a methionine-specific redox reagent developed in the Chang lab.[99] To ensure single site labelling, methionine residues were introduced at sites according to previously validated studies of the Fab scaffold that have demonstrated high labelling efficiency and labelling stability.[100] With an azide handle installed, copper-free click chemistry was performed to attach a DBCO-PEG-CGS21680 molecule onto the Fab (**Fig. 3.1**). CGS21680 was synthesized with PEG linkers of various lengths to allow for flexibility when recruiting the E3 ligase, which is a known

factor for consideration when designing PROTACs. CGS21680-conjugated RNF43 Fab retaining binding to RNF43 as measured by BLI to RNF43-Fc, and on-cell binding was measured by flow using a GFP fab conjugated to CGS21680.

Initial degradation experiments were carried out in MOLT-4 cells that endogenously express both A2AR and RNF43, and degradation levels were investigated through western blotting. We tested 4 different sites on the Fab for methionine conjugation and tested 3 linker lengths at each site, but unfortunately saw no degradation for each condition (**Fig. 3.2**). Though we checked for RNF43 expression via flow, we hypothesized this may be due to very low amounts of RNF43 expressed by MOLT-4 cells (0.1 transcripts per million (TPM) according to human protein atlas). In general, it was challenging to find cell lines endogenously expressing both A2AR and RNF43, so we began to look for alternative methods for A2AR degradation.

We are currently testing out three other ideas: (1) degradation of A2AR using ZNRF3, another cell surface E3 ligase that is generally has higher endogenous expression levels; (2) internalization of A2AR using CGS21680 conjugated to Kine-TACs; (3) recruitment of cytotoxic T cells to A2AR-expressing cells using CGS21680 conjugated to an anti-CD3 Fab (**Fig. 3.3**).

First, methionine mutations were introduced onto the scaffold of the ZNRF3 Fab, and conjugation with oxaziridine-azide was tested. We saw significant double labelling of the Fab, most likely due to the presence of a methionine residue in the heavy chain (HC) framework 4 regions of the Fab. This methionine was converted to a leucine residue (M83L), and oxaziridine labelling of the methionine-scrubbed Fab now showed single labelling. Fabs were then conjugated with CGS21680 and tested for retained binding to ZNRF3-Fc via BLI. However, when testing for A2AR degradation in Jurkat cells, once again no degradation was observed.

Next, we sought to degrade A2AR using KineTACs. While degradation via recruitment of an E3 ligase is geometry-dependent and may not form an interaction that leads to transfer of ubiquitin to the POI, KineTACs offer an advantage by targeting POIs to receptors that are already constantly undergoing internalization. The KineTAC developed by Pance et al. used a bispecific

IgG scaffold with one arm targeting a POI and the other harboring recombinant CXCL12, a cytokine that targets atypical chemokine receptor CXCR7, a decoy receptor that undergoes constitutive internalization and recycling.

Instead of using a bispecific format, we conjugated CGS21680 to methionine scrubbed CXCL12-Fc harboring a V262M mutation for site specific conjugation. We are currently working to confirm that CGS21680 conjugated CXCL12-Fc retains binding to cells via either the CGS21680 or CXCL12 arm and testing for degradation via flow cytometry. Lastly, we have conjugated CGS21680 to OKT3, a Fab targeting CD3 on T cells for use in a bispecific T cell engager (BiTE) format and are also exploring co-culture assays to look for BiTE-induced T cell activation and target cell death.

Methods

Cell lines

Cell lines were grown and maintained in T75 (Thermo Fisher Scientific) flasks at 37ºC and 5% CO2. MOLT-4 CCR5+ cells were grown in RPMI-1640 supplemented with 10% fetal bovine serum (FBS) and 2% geneticin. MOLT-4 CCR5+ cells were obtained from the NIH AIDS Reagent Program.

Antibody cloning, expression, and purification

Anti-RNF43 Fab LC S7M single mutation was introduced using Gibson Assembly. Fabs were expressed in E. coli C43(DE3) Pro+ using an optimized autoinduction media and purified by Protein A affinity chromatography. Purity and integrity of Fabs were assessed by SDS-PAGE and intact LC/MS mass spectrometry (Waters).

Conjugation of engineered anti-RNF43 Fab with oxaziridine and DBCO-CGS21680

For conjugation with oxaziridine, 50 μM Fab was incubated with 5 molar equivalents of oxaziridine azide for 30 min at room temperature in phosphate-buffer saline (PBS). The reaction was quenched with 10 molar equivalents methionine. The antibody was buffer exchanged into PBS and desalted using a 0.5-mL Zeba 7-kDa desalting column (Thermo Fisher Scientific). Then, 10 molar equivalents of DBCO-CGS21680 was added and incubated at room temperature overnight. The agonist-labeled conjugate was desalted using the 0.5-mL Zeba 7-kDa desalting column to remove excess DBCO-CGS21680. Full conjugation at each step was monitored by intact mass spectrometry using a Xevo G2-XS Mass Spectrometer (Waters).

Degradation assays

Cells at 1 million cells/mL were treated with antibody-drug conjugate, agonist, or antagonist in complete growth medium. After 24 hrs, cells were pelleted by centrifugation (300xg, 5 min, 4ºC). Cell pellets were lysed with RIPA buffer containing cOmplete mini protease inhibitor cocktail on ice for 40 min. Lysates were spun at 16,000xg for 10 min at 4ºC and protein concentrations were normalized using BCA assay. 4x NuPAGE LDS sample buffer and 2-mercaptoethanol (BME) was added to the lysates. Equal amounts of lysates were loaded onto a 4-12% Bis-Tris gel and ran at 200V for 37 min. The gel was incubated in 20% ethanol for 10 min and then transferred onto a polyvinylidene difluoride (PVDF) membrane. The membrane was blocked in PBS with 0.1% Tween20 + 5% bovine serum albumin (BSA) for 30 min at room temperature with gentle shaking. Membranes were co-incubated overnight with rabbit-anti-A2aR (Abcam, ab3461, 1:1000) and mouse-anti-tubulin (Cell Signaling Technologies, DM1A, 1:1600) at 4ºC with gentle shaking in PBS + 0.2% Tween20 + 5% BSA. Membranes were washed four times with tris-buffered saline (TBS) + 0.1% Tween20 and then co-incubated with HRP-anti-rabbit IgG (Cell Signaling Technologies, 7074S, 1:2000) and 680RD goat anti-mouse IgG (LI-COR, 926-68070, 1:10000) in PBS + 0.2% Tween20 + 5% BSA for 1 hr at room temperature. Membranes were washed four times with TBS + 0.1% Tween20, then washed with PBS. Membranes were first imaged using an

OdysseyCLxImager (LI-COR). SuperSignal West Pico PLUS Chemiluminescent Substrate was then added and image using a ChemiDoc Imager (BioRad). Band intensities were quantified using Image Studio Software (LI-COR).

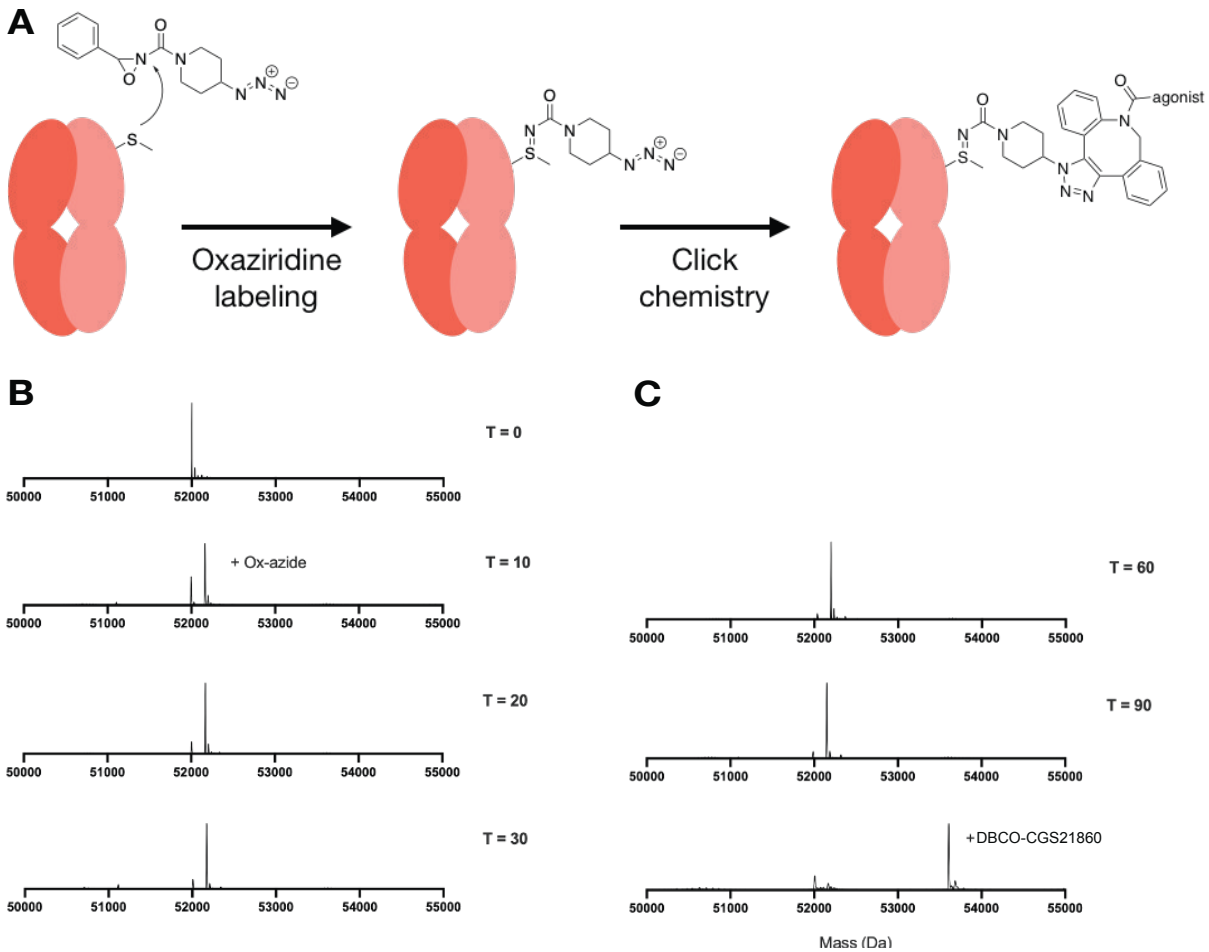


Figure 3.1: ADC-TAC bioconjugation strategy. (**A**) Schematic showing two-step conjugation involving (1) methionine labelling using oxaziridine azie and (2) click chemistry using DBCO-agonist. (**B**) Example LC/MS trace of oxaziridine labelling of Fabs. (**C**) Example LC/MS trace of DBCO-agonist labelling of Fabs.

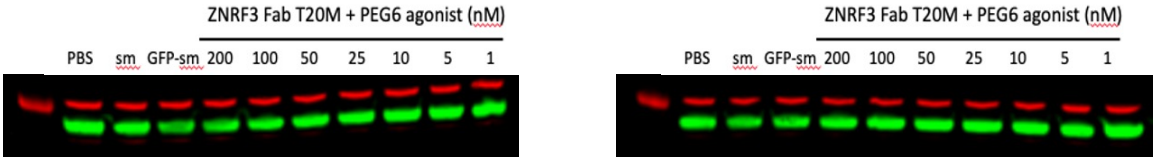


Figure 3.2: Representative western blots show no A2AR degradation observed in MOLT-4 cells. Left and right blots show biological duplicates.

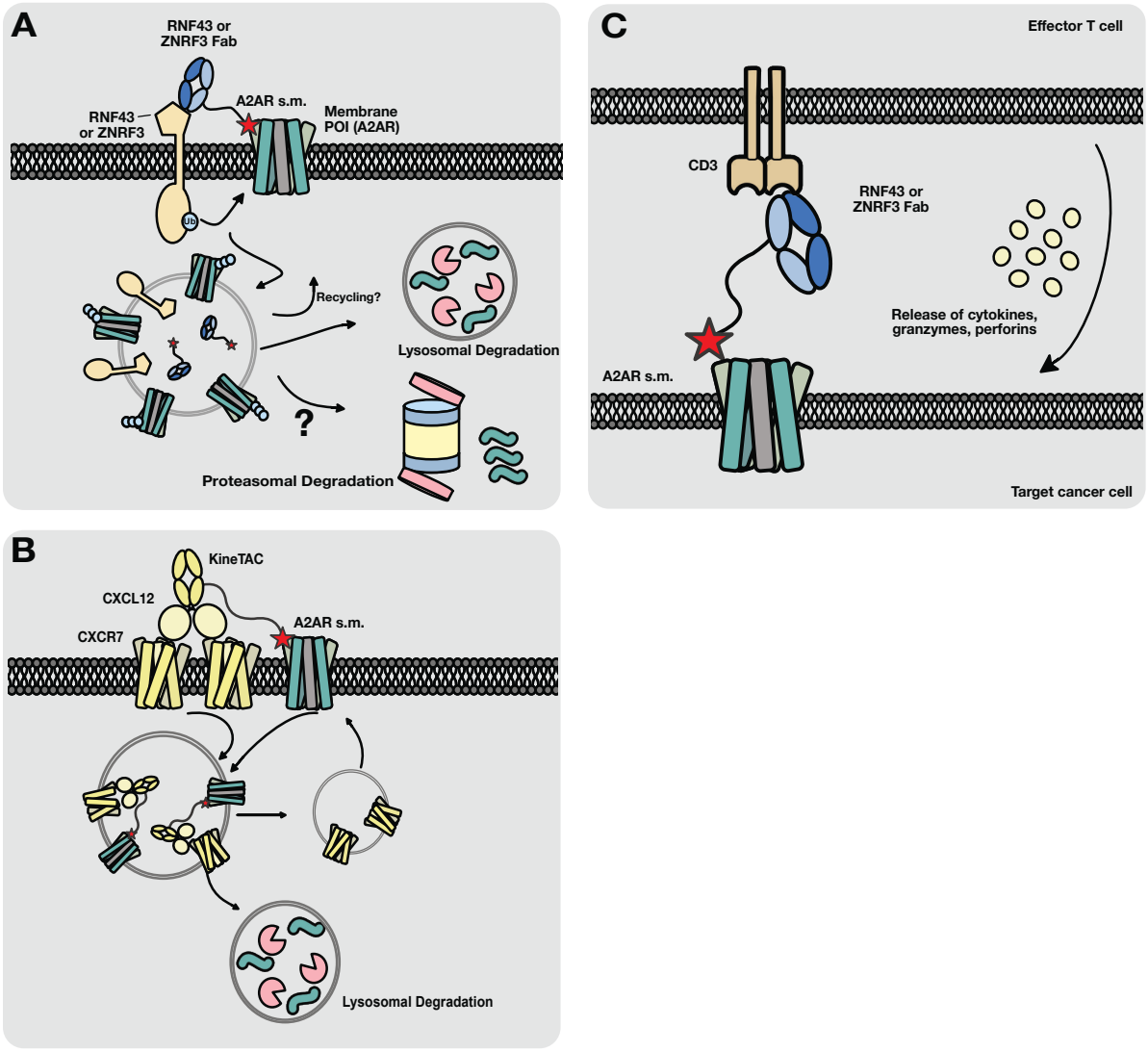


Figure 3.3: Schematic of various applications for ADC-TACs to be used in targeted protein degradation or cell killing. (**A**) Degradation using ADC-TAC to recruit A2AR to E3 ligase RNF43 or ZNRF3. (**B**) Degradation using ADC-TAC to recruit A2AR to internalizing receptor CXCR7. (**C**) Cell killing in a co-culture setup showing recruitment of CD3+ T cells to A2AR expressing target cells via ADC-TAC.

References

1. Witalison, E. E., Thompson, P. R., & Hofseth, L. J. Protein Arginine Deiminases and Associated Citrullination: Physiological Functions and Diseases Associated with Dysregulation. *Curr. Drug Targets* **16**, 700-710 (2015).
2. Fujisaki, M., & Sugawara, K. Properties of peptidylarginine deiminase from the epidermis of newborn rats. *J. Biochem.* **89,** 257-63 (1981).
3. Watanabe, K., & Shenshu, T. Isolation and characterization of cDNA clones encoding rat skeletal muscle peptidylarginine deiminase. *JBC.* **26,** 15255-15260 (1989).
4. Leshner, M., et al. PAD4 mediated histone hypercitrullination induces heterochromatin decondensation and chromatin unfolding to form neutrophil extracellular trap-like structures. *Front. Immunol*. **3**, 307 (2012).
5. Wong, S. L., & Wagner, D. D. Peptidylarginine deiminase 4: a nuclear button triggering neutrophil extracellular traps in inflammatory diseases and aging. *FASEB J*. **32**, 6358-6370 (2018).
6. Rohrbach, A. S., Slade, D. J., Thompson, P. R., & Mowen, K. A. Activation of PAD4 in NET formation. *Front. Immunol*. **3**, 360 (2012).
7. Curran, A. M., Naik, P., Giles, J. T., & Darrah, E. PAD enzymes in rheumatoid arthritis: pathogenic effectors and autoimmune targets. *Nat. Rev. Rheumatol*. **16**, 301-315 (2020).
8. Bicker, K. L., & Thompson, P. R. The protein arginine deiminases: Structure, function, inhibition, and disease. *Biopolymers* **99**, 155-163 (2013).
9. Darrah, E., & Andrade, F. Rheumatoid arthritis and citrullination. *Curr. Opin. Rheumatol*. **30**, 72-78 (2018).
10. Mondal, S., & Thompson, P. R. Chemical biology of protein citrullination by the protein A arginine deiminases. *Curr. Opin. Chem. Biol*. **63**, 19-27 (2021).

11. Mondal, S., & Thompson, P. R. Protein Arginine Deiminases (PADs): Biochemistry and Chemical Biology of Protein Citrullination. *Acc. Chem. Res*. **52**, 818-832 (2019).
12. Kouk, S., et al. PAD4: pathophysiology, current therapeutics and future perspective in rheumatoid arthritis. *Expert Opin. Ther. Targets* **21**, 433-447 (2017).
13. Aliko, A., et al. Discovery of Novel Potential Reversible Peptidyl Arginine Deiminase Inhibitor. *Int. J. Mol. Sci*. **20**, 2174 (2019).
14. Liu, X., et al. PAD4 takes charge during neutrophil activation: Impact of PAD4 mediated NET formation on immune-mediated disease. *J. Thromb. Haemost.* **19**, 1607-1617 (2021).
15. Paduch, M., et al. Generating conformation-specific synthetic antibodies to trap proteins in selected functional states. *Methods* **60**, 3-14 (2013).
16. Gao, J., Sidhu, S. S., & Wells, J. A. Two-state selection of conformation-specific antibodies. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3071-3076 (2009).
17. Mukherjee, S., et al. Engineered synthetic antibodies as probes to quantify the energetic contributions of ligand binding to conformational changes in proteins. *J. Biol. Chem*. **293**, 2815-2828 (2018).
18. Rizk, S. S., et al. Allosteric control of ligand-binding affinity using engineered conformation-specific effector proteins. *Nat. Struct. Mol. Biol.*, **18**, 437-442 (2011).
19. Liu, Y. L., et al. Probing the Roles of Calcium-Binding Sites during the Folding of Human Peptidylarginine Deiminase 4. *Sci. Rep*. **7**, 2429 (2017).
20. Knuckley, B., et al. Substrate specificity and kinetic studies of PADs 1, 3, and 4 identify potent and selective inhibitors of protein arginine deiminase 3. *Biochemistry* **49**, 4852-4863 (2010).
21. Darrah, E., et al. Erosive rheumatoid arthritis is associated with antibodies that activate PAD4 by increasing calcium sensitivity. *Sci. Transl. Med.* **5**, (2013).

22. Shi, J., et al. Affinity maturation shapes the function of agonistic antibodies to peptidylarginine deiminase type 4 in rheumatoid arthritis. *Ann. Rheum. Dis.* **77**, 141-148 (2018).
23. Darrah, E., et al. Association of Baseline Peptidylarginine Deiminase 4 Autoantibodies With Favorable Response to Treatment Escalation in Rheumatoid Arthritis. *Arthritis Rheum*. **71**, 696-702 (2019).
24. Reyes-Castillo, Z., Francisco Muñoz-Valle, & J., Llamas-Covarrubias, M. Clinical and immunological aspects of anti-peptidylarginine deiminase type 4 (anti-PAD4) autoantibodies in rheumatoid arthritis. *Autoimmun. Rev*. **17,** 94-102 (2018).
25. Harris, M. L., et al. Association of autoimmunity to Peptidyle Arginine Deiminase Type 4 with Genotype and Disease Severity in Rheumatoid Arthritis. *Arthritis Rheum*. **58**, 1958-1967 (2008).
26. Auger, I., Martin, M., Balandraud, N., & Roudier, J. Rheumatoid arthritis-specific autoantibodies to peptidyl arginine deiminase type 4 inhibit citrullination of fibrinogen. *Arthritis Rheum.* **62**, 126-131 (2010).
27. Martinez-Prat, L., et al. Autoantibodies to protein-arginine deiminase (PAD) 4 in rheumatoid arthritis: immunological and clinical significance, and potential for precision medicine. *Expert Rev. Clin. Immunol*. **15**, 1073-1087 (2019).
28. Sabulski, M. J., Fura, J. M., & Pires, M. M. Fluorescence-based monitoring of PAD4 activity via a pro-fluorescence substrate analog. *J. Vis. Exp.* **93**, e52114 (2014).
29. Liu, Y. L., Chiang, Y. H., Liu, G. Y., & Hung, H. C. Functional role of dimerization of human peptidylarginine deiminase 4 (PAD4). *PLoS One* **6**, e21314 (2011).
30. Saijo, S., et al. Monomeric Form of Peptidylarginine Deiminase Type I Revealed by X-ray Crystallography and Small-Angle X-ray Scattering. *J. Mol. Biol.* **425**, 3058-73 (2016).
31. Lee, C. Y., et al. Molecular Interplay between the Dimer Interface and the Substrate-Binding Site of Human Peptidylarginine Deiminase 4. *Sci. Rep*. **7**, 42662 (2017).

32. Arita, K., et al. Structural basis for Ca(2+)-induced activation of human PAD4. *Nat. Struct. Mol. Biol*. **11**, 777-783 (2004).

33. Kim, H. Y., Stojadinovic, A., & Izadjoo, M. J. (2014). Affinity Maturation of Monocolonal Anitbodies by Multi-Site-Directed Mutagenesis. In *Monoclonal Antibodies,* pp. 407-420.

34. Jansen, G., Scheper, R. J., & Dijkmans, B. A. Multidrug resistance proteins in rheumatoid arthritis, role in disease-modifying antirheumatic drug efficacy and inflammatory processes: an overview. *Scand. J. Rheumatol*. **32**, 325-336 (2003).

35. Zhu, D., Lu, Y., Wang, Y., & Wang, Y. PAD4 and Its Inhibitors in Cancer Progression and Prognosis. *Pharmaceutics* **14**, (2022).

36. Jones, J. E., Causey, C. P., Knuckley, B., Slack-Noyes, J. L., & Thompson, P. R. Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Curr. Opin. Drug Discov. Devel*. **12**, 616-627 (2009).

37. Lewis, H. D., et al. Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation. *Nat. Chem. Biol*. **11**, 189-191 (2015).

38. Smolen, J. S., Aletaha, D., & McInnes, I. B. Rheumatoid arthritis. *Lancet* **388**, 2023-2038 (2016).

39. Igawa, T., et al. Antibody recycling by engineered pH-dependent antigen binding improves the duration of antigen neutralization. *Nat. Biotechnol*. **28**, 1203-1207 (2010).

40. Glasgow, J. E., et al. Identifying and antagonizing the interactions between layilin and glycosylated collagens. *Cell. Chem. Biol*. **29**, 597-604 (2022).

41. Slade, D.J., et al. Protein arginine deiminase 2 binds calcium in an ordered fashion: implications for inhibitor design. *ACS Chem. Biol.* **10**, 1043-1053 (2015).

42. Funabashi, K., et al. Structures of human peptidylarginine deiminase type III provide insights into substrate recognition and inhibitor design. *Arch. Biochem. Biophys.* **708**, 102911 (2021).

43. Hornsby, M., et al. A High Through-put Platform for Recombinant Antibodies to Folded Proteins. *Mol Cell Proteomics* **14**, 2833-2847 (2015).
44. Chen, G., & Sidhu, S. S. Design and generation of synthetic antibody libraries for phage display. *Methods Mol. Biol*. **1131**, 113-131 (2014).
45. Tonikian, R., Zhang, Y., Boone, C., & Sidhu, S. S. Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nat. Protoc*. **2**, 1368-1386 (2007).
46. Huang, R., Fang, P., & Kay, B. K. Improvements to the Kunkel mutagenesis protocol for constructing primary and secondary phage-display libraries. *Methods* **58**, 10-17 (2012).
47. Schindelin, J., et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682 (2012).
48. Ohi, M., Li, Y., Cheng, Y., Walz, T. Negative Staining and Image Classification – Powerful Tools in Modern Electron Microscopy. *Biol. Proced. Online, 6*, 23-24 (2004).
49. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods*, **14**, 290-296 (2017).
50. Grant, T., Rohou, A., & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *Elife*, **7** (2018).
51. Tan, Y. Z., et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* **14**, 793-796 (2017).
52. Liebschner, D., et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol*. **75**, 861-877 (2019).
53. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol*. **74**, 519-530 (2018).
54. Pettersen, E. F., et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*. **30**, 70-82 (2021).

55. Goddard, T. D., et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14-25 (2018).
56. Jang, B., et al. Peptidylarginine deiminase and protein citrullination in prion diseases: strong evidence of neurodegeneration. *Prion.* **7**, 42-46 (2013).
57. Zhu, G., et al. Proteomics of post-translational modifications in colorectal cancer: Discovery of new biomarkers. *BBA – Reviews on Cancer.* **1877**(2022).
58. Yuzhalin, A.E., et al. Colorectal cancer liver metastatic growth depends on PAD4-driven citrullination of the extracellular matrix. *Nat Comm.* **9**, 4783 (2018).
59. Li, K., et al. ACPA-negative rheumatoid arthritis: From immune mechanisms to clinical translation. *The Lancet.* **83** (2022).
60. Wang, S., Wang Y. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. *BBA*. **1829** (2013).
61. Zhang, X., et al. Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor α target gene activation. *PNAS.* **109,** 13331-13336 (2012).
62. Zheng, L., et al. Calcium Regulates the Nuclear Localization of Protein Arginine Deiminase 2. *Biochem.* **58**, 3042-3056 (2019).
63. Durrant, L., Autophagy, citrullination and cancer. *Autophagic Puncta.* 1055-1056 (2016).
64. Brentville, V., et al. Citrullinated Vimentin Presented on MHC-II in Tumor Cells Is a Target for CD4+ T-cell-Mediated Antitumor Immunity. **76**, 548-560 (2016).
65. Cook, K., et al. Citrullinated α-enolase is an effective target for anti-cancer immunity. *OncoImmunology.* **7** (2017).
66. Brentville, V., et al. T cell repertoire to citrullinated self-peptides in healthy humans is not confined ot the HLA-DR SE alleles; Targeting of citrullinated self-peptides presented by HLA-DP4 for tumor therapy. *OncoImmunology.* **8** (2019).

67. Brentville, V., et al. Combination vaccine based on citrullinated vimentin and enolase peptides induces potent CD4-mediated anti-tumor responses. *J Immunother Cancer.* **8** (2020).
68. Snir, O., et al. Identification and Functional Characterization of T Cells Reactive to Citrullinated Vimentin in HLA-DRB1*0401-Positive Humanized Mice and Rheumatoid Arthritis Patients. *A&R.* **63**, 2873-2883 (2011).
69. Tan, Y., et al. Barcode-Enabled Sequencing of Plasmablast Antibody Repertoires in Rheumatoid Arthritis. *Arthritis Rheumatol.* **66**, 2706-2715 (2014).
70. Swain, S.M., Shastry, M., Hamilton, E., Targeting HER2-positive breast cancer: advances and future directions. *Nat Rev Drug Disc.* **22**, 101-126 (2023).
71. Eigenbrot, C., et al. Structural basis for high-affinity HER2 receptor binding by an engineering protein. *PNAS.* **107**, 15039-15044 (2010).
72. Ekerljung, L., Lennartsson, J., Gedda, L. The HER2-Binding Affibody Molecule $(ZHER_{2:342})_2$ Increases Radiosensitivity in SKBR-3 Cells. *PLoS One.* **7** (2012).
73. Nazari, M., Radmanesh, R. ZHER2 Affibody as a Good Candidate for Detection of Metastatic Prostate Cancer. *Avicenna J Med Biotechnol.* **13** (2021).
74. Lewallen, D.M., et al. Chemical Proteomic Platform To Identify Citrullinated Proteins. *ACS Chem Biol.* **10**, 2520-2528 (2015).
75. Rebak, A.S., Hendriks, I.A., Nielsen, M.L. Characterizing citrullination by mass spectrometry-based proteomics. *Phil Trans R Soc B*. **378** (2023).
76. Shi, Y., et al. Enabling Global Analysis of Protein Citrullination via Biotin Thiol Tag-Assisted Mass Spectrometry. *Anal Chem.* **94 (**2022).
77. Wang, B., et al. Mass Spectrometry-Based Precise Identification of Citrullinated Histone via Limited Digestion and Biotin Derivative Tag Enrichment. *Anal Chem.* **96** (2024).

78. Hao, G., et al. Neutral loss of isocyanic acid in peptide CID spectra: a novel diagnostic marker for mass spectrometric identification of protein citrullination. *J Am Soc Mass Spectrom.* **20**, 723-727 (2009).
79. Hensen, S.M.M., Pruijn, G.J.M. Methods for the Detectrion of Peptidylarginine Deiminase (PAD) Activity and Protein Citrullination. *MCP.* **13** (2014).
80. Bennike, T., et al. Optimizing the Identification of Citrullinated Peptides by Mass Spectrometry: Utilizing the Inability of Trypsin to Cleave after Citrullinated Amino Acids. *J Proteomics Bioinf* (2013)jj.
81. Jin, Z., et al. Identification and characterization of citrulline-modifed brain proteins by combining HCD and CID fragmentation. *Proteomics*. **13**, 2682-2691 (2013).
82. Jin, Y., et al. Different syngeneic tumors show distinctive intrinsic tumor-immunity and mechanisms of actions (MOA) of anti-PD-1 treatment. *Sci Rep.* **12** (2022).
83. Adkins, I., et al. Sever, but not mild heat-shock treatment induces immunmogenic cell death in cancer cells. *OncoImmunology*. **6** (2017).
84. Kooreman, N.G. Autoogous iPSC-Based Vaccines Elicit Anti-tumor Responses *In Vivo. Cell Stem Cell.* **22**, 501-513 (2018).
85. Wells, J.A., Kumru, K. Extracellular targeted protein degradation: an emerging modality for drug discovery. *Nat Rev Drug Discov.* **23**, 126-140 (2024).
86. Pettersson, M., Crews, C.M. PROteolysis TArgeting Chimeras (PROTACs) – Past, present and future. *Drug Discov Today Technol.* **31**, 15-27 (2019).
87. Mullard, A. First targeted protein degrader hits the clinic. *Nat Rev Drug Discov.* (2019).
88. Banik, S.M., et al. Lysosome-targeting chimeras for degreadation of extracellular proteins. *Nature.* **584**, 291-297 (2020).
89. Cotton, A.D., et al. Development of Antibody-Based PROTACs for the Degradation of Cell-Surface Immune Checkpoint Protein PD-L1. *JACS*. **143**, 593-598 (2021).

90. Pance, K., et al. Modular cytokine receptor-targeting chimeras for targeted degradation of cell surface and extracellular proteins. *Nat Biotech.* **41**, 273-281 (2023).
91. Zhang, D. et al. Transferrin Receptor Targeting Chimeras (TransTACs) for Membrane Protein Degradation. *bioRxiv* (2023).
92. Siepe, D.H., Picton, L.K., Garcia, C.K. Receptor Elimination by E3 Ubiquitin Ligase Recruitment (REULR): A Targeted Protein Degradation Toolbox. *ACS Synth Biol.* **12**, 1081-1093 (2023).
93. Schlimgen, R.R. Structural basis for selectivity and antagonism in extracellular GPCR-nanobodies. *Nat Comm.* **15** (2024).
94. Vijayan, D., et al. Targeting immunosuppressive adenosine in cancer. *Nat Rev Cancer* **17**, 709-724 (2017).
95. Sek, K., et al. Targeting Adenosine Receptor Signaling in Cancer Immunotherapy. *Int J Mol Sci.* **19** (2018).
96. Sun, C., Wang, B., Hao, S. Adenosine-A2A Receptor Pathway in Cancer Immunotherapy. *Front Immunol.* **13** (2022).
97. Chiang, M.J., et al. An Fc Domain Protein–Small Molecule Conjugate as an Enhanced Immunomodulator. *JACS.* **136**, 3370–3373 (2014).
98. Lebon, G., et al. Molecular Determinants of CGS21680 Binding to the Human Adenosine A2A Receptor. *Mol Pharmacol*. **87**, 907–915 (2015).
99. Lin, S. et al. Redox-based reagents for chemoselective methionine bioconjugation. *Science*. **355**, 597–602 (2017).
100. Elledge, S. K. et al. Systematic identification of engineered methionines and oxaziridines for efficient, stable, and site-specific antibody bioconjugation. *PNAS*. **117**, 5733– 5740 (2020).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
0A51554BDEF946F...
Author Signature

8/27/2024
Date