Title
Expanding the Antibody Toolbox

Author
Cotton, Adam David

Publication Date
2021

Peer reviewed|Thesis/dissertation

Expanding the Antibody Toolbox

by
Adam D. Cotton

DISSERTATION
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:
James Wells
5F2F4D1A06164C2...

James Wells
Chair

DocuSigned by:
Ian Seiple
DocuSigned by:
Jason Gestwicki
DocuSigned by:
David Toczyski
B39DA64A208E4CB...

Ian Seiple

Jason Gestwicki

David Toczyski

Committee Members

Acknowledgements

I am incredibly grateful to the group of people who have helped me through my time at UCSF. Firstly, I need to thank both my PI's. One thing I learned from years of running was you should find **who** you want to run (or work) for, and after that everything will fall into place. It was a week into my summer rotation when I knew I wanted to work for Ian. Your enthusiasm for both science and mentorship has been unwavering, and I have loved watching the lab grow into the bustling atmosphere of collaborative science it is today. As much as you have been a mentor to me, you have been a great friend, and I'll be eternally grateful that you've always been there for me throughout the highs and the lows. Next, Jim, I still vividly remember sitting in the breakroom as you pitched the AbTAC idea to me; your excitement for creative science was just as contagious then as it is today. Beyond your creativity, I think a major reason you have been so successful is that you truly care about your trainees, and I feel lucky that I've had you in my corner these last four years. You have both been so supportive in letting me work on a range of projects I've found interesting, and it's not lost on me how rare and really cool that is.

Leaving both my labs will be hard; there are so many colleagues, now friends, who have helped me grow personally and scientifically over the years. Every member of both labs has had a positive impact on my personal experience and thus projects. Whether that's brainstorming crazy ideas over a beer, helping me analyze data, answering my stupid questions, or being there for me when times got tough. Thank you, and I'm excited to see you all achieve your scientific and career goals.

Outside of the lab, I am indebted to my family and friends. Thanks to Mom and Dad for always encouraging me to do what I love; while we are thousands of miles and many time zones apart thank you for the weekly reminder of home. Thanks to my big sisters, Emma, and Abi, for their continued support over the last 29 years. Thanks to great friends Carl, Dan & Amy, and Fred &

Cate, your board game, Fortnite, and wine tasting nights have been the perfect get-away that have kept me sane here in the Bay Area. Thanks to Cole and Jamie for being great friends in and out of the lab; I've loved talking science with you as much as I loved our group runs and adventures in Yosemite.

Finally, Courtney. My best friend, life partner, and greatest supporter. Thank you for your constant push to keep pursuing my dreams and for being there for me through the highs and lows. Our educational paths have been long and windy, but we are finally there. The light at the end of the long-distance tunnel is finally beaming on our faces, and I couldn't be more excited to spend the rest of our lives together.

Contributions

Chapter 1 of this thesis is a reprint of the material as it appears in:

Cotton, A; Nguyen, D; Gramespacher, J; Seiple, I; Wells, J. Development of Antibody-based PROTACs (AbTACs) for the degradation of the cell-surface immune checkpoint protein PD-L1. *Journal Of The American Chemical Society,* 2021. DOI: 10.1021/jacs.0c10008

Chapter 2 of this thesis has been submitted for publication, and the work was completed in collaboration with Josef A. Gramspacher Ph.D.

Chapter 3 of this thesis is a reprint of the material as it appears in:

Cotton, A; Seiple, I. Examining gender imbalance in chemistry authorship. *ACS Chemical Biology*, 2021. DOI: 10.1021/acschembio.1c00142

Chapter 5 of this thesis is a reprint of the material as it appears in:

Cotton, A; Wells, J; Seiple, I. Biotin as a Reactive Handle to Selectively Label Proteins and DNA with Small Molecules. *ACS Chemical Biology*, 2021. DOI: 10.1021/acschembio.1c00252

Expanding the Antibody Toolbox

By

Adam D. Cotton

Abstract

The field of antibody engineering focuses on improving the antibody scaffold for their desired needs, whether that's delivering cytotoxic payloads as an antibody-drug-conjugate (ADC) or generating bispecific antibodies (bsIgGs) to gain tissue specificity. To date, the field has created top-selling therapies on the market. Antibodies can elicit a therapeutic effect in a variety of ways from, direct inhibition to T-cell recruitment. This thesis project seeks to both improve and expand the current toolbox of antibody modalities.

Chapters 1 and 2 describe the development and optimization of a technology termed Antibody-Based PROTACs (AbTACs) for the degradation of cell surface proteins. We demonstrate that the cell-surface E3 ligases RNF43 and ZNRF3 can be co-opted via a bsIgG to induce the targeted degradation of a range of clinically relevant proteins, including PD-L1 and EGFR.

Chapter 3 expands on redox-reactive oxaziridine reagents to label biotinylated biomolecules, including antibodies and DNA. Despite being solvent-exposed, the resulting conjugates are exceedingly stable, and we were able to generate both antibody-drug conjugates and flow cytometry reagents.

Chapter 4 models a hypothetical antibody therapeutic to treat COVID-19 infections. We look at the required in-vitro neutralization values and link that to patient dosing and production capacities for a large pharmaceutical company trying to deliver these molecules to patients.

Chapter 5 describes a sociological meta-analysis of gender diversity in chemistry publishing. We generated a dataset of name-predicted genders of both first and corresponding authors in a range

of journals since 2005. The resulting analysis demonstrated the disparity in gender diversity in scientific publishing that has shown little to no signs of improvement over the last 16 years.

Table of Contents

List of Figures

List of Tables

1. Development of Antibody-based PROTACs (AbTACs) for the degradation of the cell-surface immune checkpoint protein PD-L1

1.1 Abstract

Targeted protein degradation has emerged as a new paradigm to manipulate cellular proteostasis. Proteolysis-targeting chimeras (PROTACs) are bifunctional small molecules that recruit an E3 Ligase to a target protein of interest, promoting its ubiquitination and subsequent degradation. Here we report the development of Antibody-Based PROTACs (AbTACs), fully recombinant bispecific antibodies that recruit membrane-bound E3 ligases for the degradation of cell-surface proteins. We show that an AbTAC can induce the lysosomal degradation of programmed death-ligand 1 (PD-L1) by recruitment of the membrane-bound E3 ligase RNF43. AbTACs represent a new archetype within the PROTAC field to target cell-surface proteins with fully recombinant biological molecules.

1.2 Introduction

Proteolysis targeting chimeras (PROTACs)[1] and related degradation technologies, such as LYTACs[2,3], dTAGs[4], Trim-Away[5] and SNIPERs[6] have arisen as novel therapeutic modalities to target traditionally "undruggable" proteins. PROTACs are heterobifunctional molecules consisting of an E3 ligase binder and a substrate targeting ligand that exploit the cellular protein degradation machinery to selectively degrade target proteins[1]. Unlike occupancy-based inhibition,

PROTACs act catalytically[7]. This enables them to target previously intractable proteins and to be effective even when resistance to inhibitors develops[8,9]. Despite great promise for PROTACs, they require targets with cytosolic binding domains for small-molecule ligands, leaving many membrane proteins un-targetable. Membrane proteins comprise ~23% of encoded genes[10], and ~70% of FDA approved drugs target this important class[11]. Therefore, a new strategy to degrade cell-surface proteins has the potential to be transformative to the field.

IgGs have long serum half-lives[12] and can be rapidly generated as high-affinity binders to target proteins through phage display[13,14]. Bispecific IgGs can bind to two proteins simultaneously, colocalizing them[15]. We hypothesized that this biological construct could mimic PROTACs by recruiting membrane-bound E3 ligases to proteins of interest, inducing degradation. We have termed these Antibody-based-PROTACs (AbTACs). If successful, this approach would have several advantages over previous technologies. First, AbTACs are fully recombinant bispecific IgG's allowing for their rapid and renewable generation. Next, we utilize standard phage display to generate multiple recombinant antibody binders, resulting in high affinity and high specificity. The AbTACs are thus recombinant by nature allowing for simple genetic modification of binding properties that govern the technologies efficacy. Finally, AbTACs expand the PROTAC fields attempts to target challenging membrane proteins.

The most commonly used E3 ligases by the PROTAC field are von Hippel-Lindau disease tumor suppressor (VHL)[16] and Cereblon (CRBN)[17]; there have been numerous recent accounts of successfully recruiting different ligases for degradation[18–21], although the use of a transmembrane E3 ligase (as required for our approach) has not been reported. We sought a single-pass protein with a structured ectodomain to facilitate phage display antibody generation. Ideally, it would be widely expressed across cell types to enable generalizability and would naturally serve as a

degrader of its putative target protein. RNF43 is a single-pass E3 ligase comprising a structured ectodomain and an intracellular RING domain that meets these requirements[22]. It negatively regulates the Wnt signaling pathway by ubiquitinating Frizzled, a Wnt co-receptor, causing its endocytosis and degradation[23,24]. While broadly expressed, in certain diseases RNF43 acts as a tumor suppressor and is mutated or silenced[25]; this notion would limit AbTAC use in these indications.

Here we report AbTACs, targeted biologic-based degraders to recruit membrane-bound E3 ligases for the degradation of cell-surface proteins. Generating AbTACs in the bispecific IgG format, we show that our AbTAC molecule AC-1 induces the lysosomal degradation of programmed death-ligand 1 (PD-L1) by recruitment of the membrane-bound E3 ligase, RNF43.

1.3 Results

We hypothesized that recruitment of RNF43 to a target protein of interest would lead to its internalization and lysosomal degradation (**Fig. 1.1a**). To test this hypothesis, we engineered constructs to evaluate whether RNF43 can cause the internalization of a non-cognate target protein. We fused GFP via a transmembrane domain to a NanoLuc domain[26] to serve as a reporter; this allows for orthogonal visual and biochemical readouts. To induce dimerization, we fused an anti-GFP single chain Fab (scFab) to the N-terminus of RNF43[27]; for an isotype-control, we used an anti-GCN4 scFab. Upon expression of these constructs in HeLa cells, the confocal micrograph shows GFP localized to the cell surface in the isotype control and to reporter cells treated with an anti-GFP Fab (**Fig. 1.1b,c, 1.S1a,b**). Only the anti-GFP-RNF43 fusion caused the internalization of the GFP with co-localization in the lysosome (**Fig. 1.1d, 1.S1c**). Further, we performed a Nanoluciferase assay to quantify the amount of reporter – this showed a modest ~20% reduction

(**Fig. 1.S2**). This value is comparable to degradation levels seen for other over-expression systems[28]. These data suggested that RNF43 could be used to induce protein degradation of endogenous proteins.

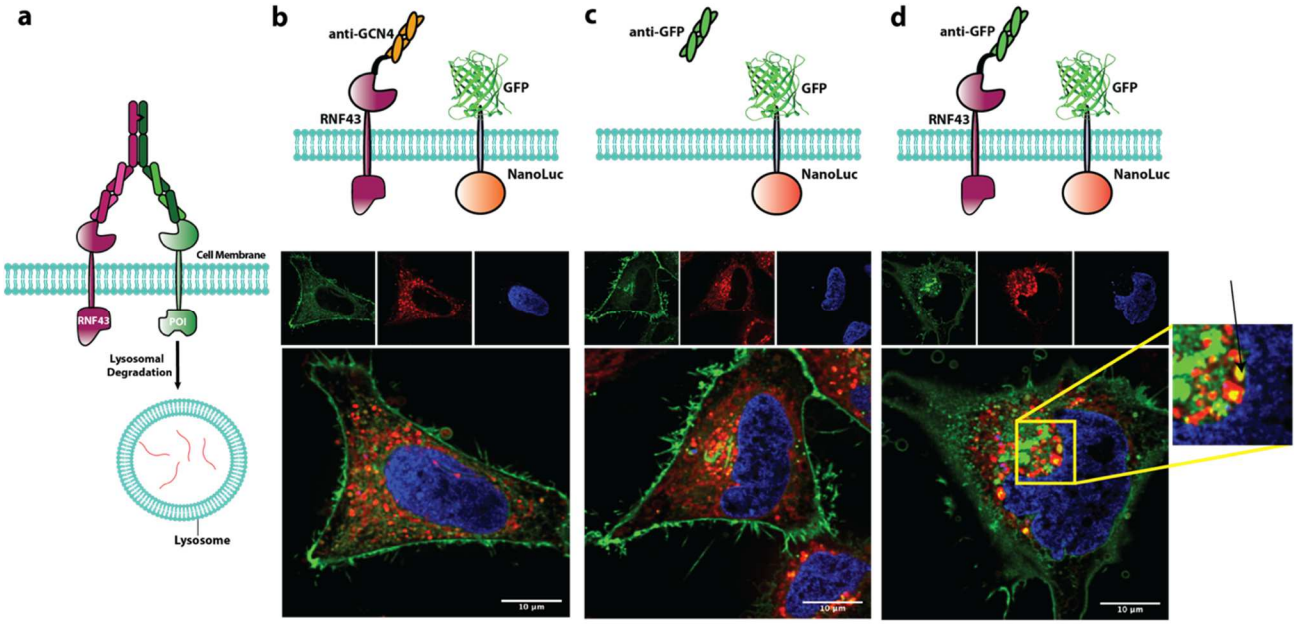


Figure 1.1: Antibody-Based-PROTACs (AbTACs) recruit RNF43 to internalize cell surface proteins. a) Graphical representation of the AbTAC mode of action. b, c, d) Engineered RNF43 constructs and the GFP-Nanoluciferase reporter. With their corresponding confocal microscopy images showing GFP localization for each experimental condition. b) Anti-GCN4-RNF43 fusion as an isotype control. c) Soluble Anti-GFP Fab to control fobr Fab binding effects. d) Anti-GFP-RNF43. (Green - GFP reporter protein, Red - Lysosome Tracker, Blue – DAPI)

We next sought to generate a recombinant antibody for the ectodomain of RNF43. For this, we utilized phage display; Fig. S3 outlines our selection strategy. After four rounds, we used a Fab-phage ELISA to triage clones with predicted affinities greater than 20 nM (**Fig. 1.2a**). Following the sequencing of passing clones, we had four unique Fab-phage. One of these clones, R3 bound well to RNF43 in vitro and on cells. Biolayer interferometry confirms that the Fab derived from

R3 has an in vitro K_D of 12.5 nM (**Fig. 1.2b**). To assess on-cell binding we overexpressed full-length RNF43 in HEK293T cells and used flow cytometry to verify Fab binding (**Fig. 1.2c**).

We chose programmed death-ligand 1 (PD-L1) for an initial degradation target. PD-L1 is a protein over-expressed in numerous cancers, causing the modulation of the T-cell response via binding to the inhibitory receptor PD-1 on T cells[29]. PD-L1 has a small, 31 amino acids long cytoplasmic domain, with no known small-molecule ligands, making it challenging to target with conventional small-molecule PROTAC approaches[30]. PD-L1 has previously been degraded using the LYTAC approach[2], and there are glyco-form specific antibodies that induce its degradation[31]. The AbTAC method introduces another mechanism for degrading cell surface proteins based on E3 ligases.

With an RNF43 binder in hand, we chose to use the published sequence for Tecentriq (atezolizumab) as our PD-L1 binder[32] to generate PD-L1-specific AbTACs. The validated knobs into holes Fc ensures correct heavy chain pairing when making bispecific IgGs (**Fig. 1.2d**)[33]. To prevent light chain mismatch pairing we expressed the antibodies as half IgGs followed by their in vitro assembly to construct AC-1 (**Fig. 1.2d**)[34]. We appropriated a His-tag on the knob half IgG to purify away unwanted hole-hole homodimer.

can bind to both antigens simultaneously, and they exhibit Fab binding kinetics due to lack of avidity effect. To test whether concurrent binding was occurring, we immobilized PD-L1 on a BLI tip, followed by subsequent addition of AbTAC and RNF43. The two sequential increases in signal indicate that AC-1 can bind to both RNF43 and PD-L1 in parallel (**Fig. 1.2e**). Due to the purification strategy, the RNF43 binding homo-dimer is the most likely contaminant. To assess purity, we compared the in vitro binding affinities for AC-1 to its Fab components, AC-1 binds to

RNF43 with the same 12.5 nM Kd as the Fab (**Fig. 1.S4**) These BLI experiments indicate AC-1 is a BsIgG that can bind to both antigens concurrently.

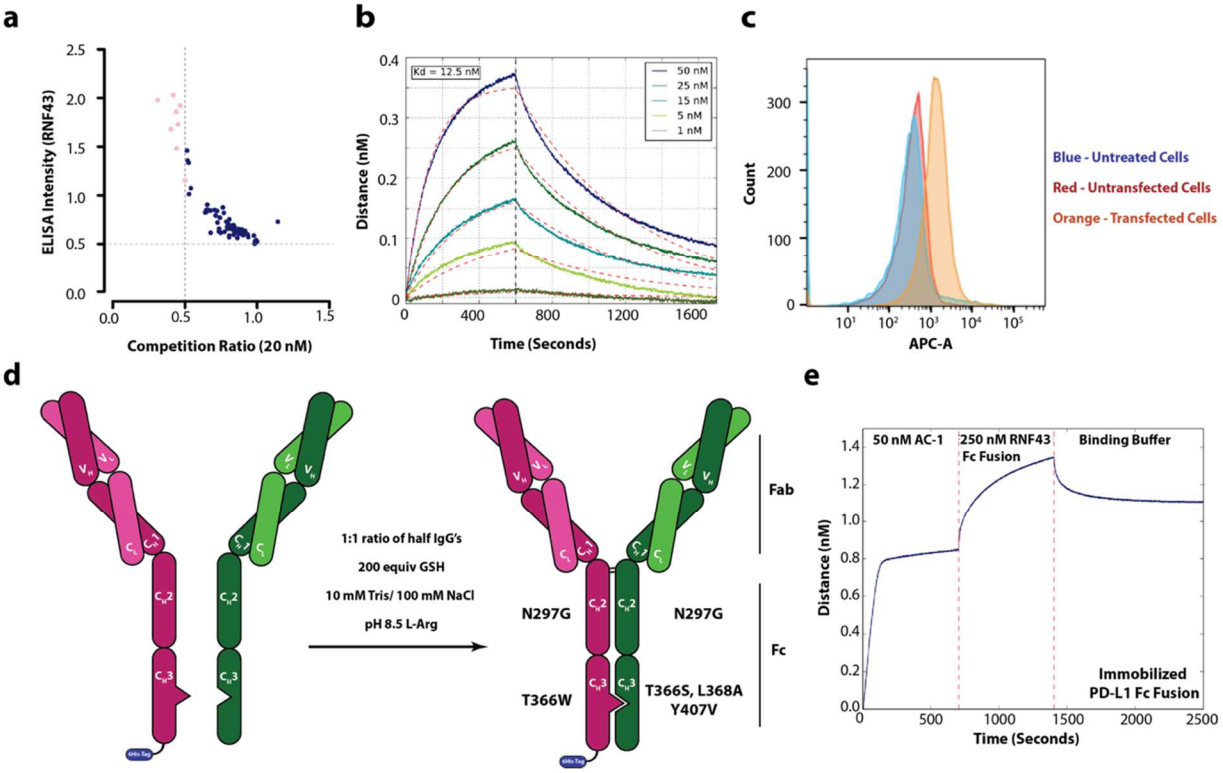


Figure 1.2: AbTACs are bispecific IgG's that can bind to RNF43 and PD-L1 concurrently. a) Fab-phage ELISA data showing binders from a phage display campaign with RNF43 Fc fusion as the target antigen. b) In vitro BLI, showing our RNF43 antibody (R3) binds to RNF43-Fc-fusion with a Kd of 12.5 nM c) Flow cytometry showing Fab R3 binds to RNF43 on cells. d) Conditions for in vitro assembly of separately expressed half IgG's to form a bispecific IgG. e) Two-step BLI experiment showing that AC-1 can simultaneously bind to both purified ecto-domains of RNF43 and PD-L1.

After verifying that AC-1 can engage both antigens in vitro, we sought to test if it could degrade PD-L1 on cells. We first chose to assess this in the triple-negative breast cancer cell line MDA-MB-231 due to its high expression of PD-L1. Flow-cytometry confirmed the presence of RNF43 on these cells (**Fig. 1.S5**). Due to PD-L1's heterogeneously glycosylated extracellular domain, it stains as a doublet in western blot analysis[35]. We found that AC-1 can induce the degradation of PD-L1 with a DC_{50} = 3.4 nM and maximal percent degradation, D_{max} = 63% at 24 hours (**Fig.**

1.3b). Treatment with either component individually or simultaneously at 10 nM did not affect PD-L1 levels, indicating both targets must be brought together to cause degradation (**Fig. 1.3c, 1.S6**). Treatment of MDA-MB-231 cells with 10 nM AC-1 induced degradation within 12 hours with maximal degradation at 24 hours (**Fig. 1.3a**).

To measure proteomic changes that might result from loss of PD-L1, we ran whole-cell proteomics, which showed no large cellular perturbations (**Fig. 1.S7**). Due to the relative low abundance of cell-surface proteins in a cell, PD-L1 peptides were not observed in either sample. We also found that PD-L1 levels recovered within 24 hours after wash-out (**Fig. 1.S8**), consistent with our findings that AC-1 does not produce irreversible cellular perturbations.

To test if PD-L1 degradation by AC-1 is dependent on RNF43, we knocked it down in MDA-MB-231 cells using CRISPR interference (CRISPRi) and confirmed knock down via RT-qPCR (**Fig. 1.S9**). AC-1 did not induce degradation of PD-L1 in these RNF43 knock down cells after 24 hours of 10 nM treatment (**Fig. 1.3d**). To evaluate the degradation pathway, we pre-treated with either Bafilomycin (a lysosome acidification inhibitor) or MG-132 (Proteasome inhibitor); Bafilomycin mitigated the degradation of PD-L1, whereas MG-132 did not (**Fig. 1.3e**). These data are consistent with the hypothesis that AC-1 causes degradation of PD-L1 in an RNF43- and lysosomal-dependent manner.

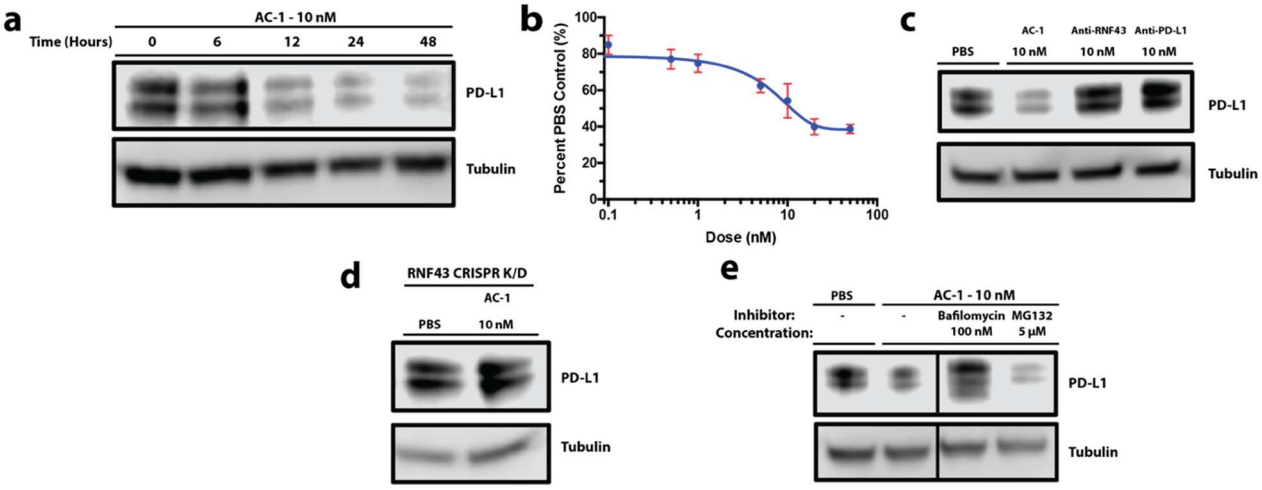


Figure 1.3: AC-1 causes the degradation of PD-L1 in MDA-MB-241 in an RNF43 and lysosomal dependent manner. a) Time course experiment showing degradation of PD-L1 after 12 hours, with maximum degradation after 24 hours. b) Dose escalation experiment showing maximum degradation of PD-L1 at 20 nM bispecific IgG treatment. Densitometry was used to quantify degradation levels. Error bars represents SD of 3 biological replicates. c) Western-blot analysis showing AC-1 can degrade PD-L1 at 10 nM after 24 hours, whereas each component of the bispecific has no effect on protein levels. d) CRISPRi knockdown of RNF43 rescues degradation, demonstrating degradation is RNF43 dependent. e) Pre-treatment with either Bafilomycin or MG132 indicates that AC-1 degrades PD-L1 in a lysosomal dependent manner. The presented data is representative of three independent replicates.

Next, we wanted to test if AC-1 could degrade PD-L1 on other cell lines. Tecentriq is approved for combination therapy in three indications, triple-negative breast cancer, non-small-cell lung cancer, and advanced bladder cancer. We chose a cell line from each of these indications that are known to have high PD-L1 levels. In each cell line, MDA-MB-231, HCC827, and T24, treatment with AC-1 at 10 nM induces degradation of PD-L1 after 24 hours, suggesting broad cellular applicability (**Fig. 1.4**).

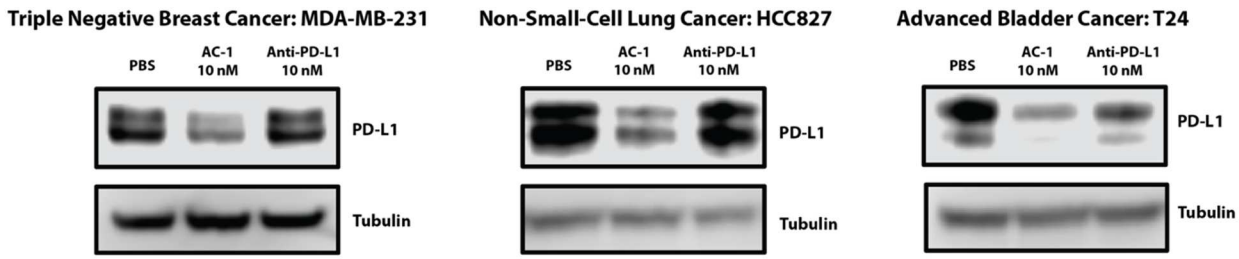


Figure 1.4: AC-1 degrades PD-L1 on clinically relevant cell-lines. To demonstrate the applicability of AC-1 to degrade PD-L1 we chose three cancer cell-lines that correspond to indications that Anti-PD-L1 checkpoint inhibitors have been approved for. AC-1 is able to degrade PD-L1 in each of these cell-lines at 10 nM after 24 hours. The presented data is representative of three independent replicates.

1.4 Discussion

PROTACs have provided an important alternative to traditional small-molecule inhibitors by their ability to deplete a protein of interest. Despite the great potential of PROTACs for targeted degradation, the examples so far have only been applied to cytosolic E3 ligases. Here, using the AbTAC technology we expand targeted degradation to transmembrane E3 ligases, allowing depletion of a cell-surface protein without a known small-molecule ligand.

As this work was underway, the Bertozzi laboratory reported a clever method, termed LYTACs[2], to degrade cell surface proteins by hijacking lysosomal targeting receptors. This elegant platform technology focused on degrading surface proteins via chemical conjugation of glycans to an antibody directed to the victim protein. In the first LYTAC iteration the glycan targets the mannose-6-phosphate receptor, which upon binding drags the cargo to the lysosome via a receptor internalization mechanism. We envision that the AbTACs can act as a complementary approach by providing fully recombinant and renewable parts to induce membrane protein degradation using membrane bound E3 ligases.

AC-1 recruits RNF43, a cell-surface E3 ligase, to degrade PD-L1, an important cell-surface check point protein with a small-intracellular domain and no reported small molecule ligands, which is a

requirement for PROTACs. The lysosomal degradation is caused by the hetero-dimerization of the two proteins, and there are no large cellular proteomic perturbations that occur as was similarly reported by LYTACs. AC-1 is a fully recombinant bispecific IgG, making it renewable and due to being built of human parts, is less likely to illicit an immune response. It is possible that engaging RNF43 for degrading PD-L1 could compromise it's function in the Wnt signaling pathways. However, it has been shown that inhibiting RNF43 or ZNRF3 (a close homolog that's involved in redundancy for the Wnt pathway) alone has very little effect on Wnt activity[36].

AC-1 achieves a D_{Max} of 63% which represents a steady state between synthesis and degradation rates. AbTACs have three main components, an E3 ligase binding domain, a protein of interest binding domain, and the antibody construct. We envision that multiple factors including binding properties, cell-surface levels, E3-target stoichiometry, endocytosis kinetics upon antibody binding, turnover rate of the protein of interest, and E3 ligase play important roles in this technology. With this deeper understanding of these factors we hope to expand the AbTAC toolbox to degrade a range of cell-surface proteins and to rationally optimize hits to test their phenotypic effects in vivo. The modular nature and genetic tractability of AbTACs offer promise for academic and translational applications.

1.5 Materials and Methods

Cell lines

Cells were grown and maintained in T75 flasks (Thermo Fisher) at 37ºC and 5% CO_2. MDA-MB-231 and HeLa cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. T24 cells were grown in in McCoy's 5a supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. HCC827 cells were grown in in RPMI supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. HeLa, MDA-MB-231 and T24 cells were obtained from the UCSF Cell Culture Facility. HCC827 cells were obtained from American Type Culture Collection (ATCC). Cell lines were authenticated by the supplier.

Bio-layer interferometry (BLI) Experiments

Bio-layer interferometry data (BLI) was measured using an Octet RED384 (ForteBio) instrument. RNF43-Fc-Fusion was immobilized on a streptavidin biosensor and loaded until 1.0 nm signal was achieved. After blocking with 10 μM biotin, purified antibodies in solution was used as the analyte. PBSTB 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.

Whole cell proteomics

MDA-MB-231 cells in a T-75 flask (Corning) were treated with either 10 nM AC-1 or PBS. Following 24 hour, cells were lifted using Versene (Gibco) and pelleted by centrifugation. Cell pellets (~7 million cells) were then lysed in 6 M guanidine hydrochloride (Chem-Impex), 5 mM tris(2-carboxyethyl)phosphine (Fisher Scientific), 10 mM chloroacetamide (Sigma Aldrich), with 50 mM Tris (Thermo Fisher) for 10 minutes at 95°C with periodic vortexing. DNA in the lysate was sheared using a probe sonicator. Lysates were then diluted three-fold with 50 mM Tris, pH

8.5 and digested at room temperature overnight with 1:100 (enzyme mass:protein mass) proteomics-grade trypsin (Promega). Following digestion, peptides were desalted using SOLA HRP columns (Thermo Fisher), dried using a GeneVac system, and resuspended in 2% acetonitrile (ACN, Fisher Scientific) + 0.1% formic acid (FA, Sigma) for subsequent liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

For LC-MS/MS, 1 μg of peptides were separated using an UltiMate 3000 UHPLC system (Thermo) with a pre-packed 0.75mm x 150mm Acclaim Pepmap C18 reversed phase column (2μm pore size, Thermo) and analyzed with a Q Exactive Plus (Thermo Fisher Scientific) mass spectrometer. Separation was performed using a linear gradient of 3-35% solvent B (Solvent A: 0.1% FA, solvent B: 80% ACN, 0.1% FA) over 230 mins at 300 μL/min. A top 20 method for data-dependent acquisition was performed during analysis (dynamic exclusion 35 seconds; selection of peptides with a charge of 2, 3, or 4). A resolution of 140,000 (at 200 m/z) was used to gather full spectra in MS1 using an AGC target of 3e6, maximum injection time of 120 ms, and scan range of 400 - 1800 m/z. Centroided data from MS2 scans were collected at a resolution of 17,500 (at 200 m/z) and an AGC target of 5e4. The maximum injection time was set at 60 milliseconds. For MS2, a collision energy of 27 was used, with an isolation window of 1.5 m/z and an isolation offset of 0.5 m/z.

Raw mass spectrometry data from two independent experiments were searched and quantified using Label-free Quantitation (LFQ) in MaxQuant (Version 1.6.7). The Uniprot Human Reference Proteome was used to generate the search database (downloaded July 2019). Cysteine carbamidomethyl was set as the only fixed modification with methionine oxidation and N-terminal glutamate to pyroglutamate as variable modifications. The MaxQuant output was then filtered and data visualized using Perseus (Version 1.6.7). Briefly, contaminants and decoys were removed and

only proteins identified with more than two unique peptides were carried forward. LFQ data were log_2 transformed and missing data were imputed using a normal distribution. A volcano plot was then generated using default Perseus statistical settings to visualize proteome changes following AbTAC treatment.

Generation of CRISPR knockdown cell line

RNF43 targeting sgRNAs, identified in a previously published genome-wide CRISPRi dataset[34], were cloned into a pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro lentiviral vector (Addgene 71236). Lentivirus were produced by transfecting HEK293T cells with standard packaging vectors. The RNF43-knockdown MDA-MB-231 cell line was generated by transducing cells with the all-in-one CRISPRi lentivirus[35]. The stable cell line expressing dCas9-KRAB and sgRNA was selected with puromycin (1 μg/mL) and validated by quantitative reverse transcription PCR (RT-qPCR).

Degradation experiments

Cells at 70% confluency were treated with bispecific antibody or control antibody in complete growth medium. At the experiment end point, cells were washed with cold phosphate-buffered saline (PBS), lifted with Versene solution and pelleted by centrifugation (500g, 5 min, 4ºC) and samples tested by either 'western blotting' or 'degradation flow cytometry'.

Western blotting

Cells pellets were lysed with RIPA buffer containing cOmplete mini protease inhibitor cocktail (Sigma) on ice for 45 min. The lysates were spun at 21,000g for 10 min at 4ºC and protein concentrations were normalized using BCA assay (Pierce). 4x NuPAGE LDS sample buffer (Invitrogen) and 2-Mercaptoethanol (BME) was added to the lysates and boiled for 10 min. 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 to a polyvinylidene difluoride (PVDF) membrane. The membrane was then blocked in Odyssey Blocking Buffer (TBS) (LI-COR) for 45 min at room temperature with gentle shaking. Membranes were incubated overnight with rabbit-anti-PDL1 (CST: E1L3N, 1:1,000) and mouse-anti-tubulin (CST: DM1A, 1:2,000) at 4ºC with gentle shaking in Odyssey Blocking Buffer (TBS) + 0.2% Tween-20, then washed five times with TBS + 0.1% Tween-20. The membrane was then incubated with HRP-anti-rabbit IgG (1:2000) for 1 h at room temperature in Odyssey Blocking Buffer (TBS) + 0.2% Tween-20. Membranes were washed five times with TBS + 0.1% Tween-20 and SuperSignal West Pico PLUS Chemiluminescent Substrate was added before imaging using a LICOR C-DiGit blot scanner. The membrane was then washed three times in TBS + 0.1% Tween-20, and 800CW goat anti-mouse IgG (1:10,000) in Odyssey Blocking Buffer (TBS) + 0.2% Tween-20 + 0.01% SDS for 1 h at room temperature. Membranes were washed five times with TBS + 0.1% Tween-20, then imaged using an OdysseyCLxImager (LI-COR). Band intensities were quantified using Image Studio Software (LI-COR)

Surface panning flow cytometry

Cells were washed with cold phosphate-buffered saline (PBS), lifted with Versene solution and pelleted by centrifugation (500g, 5 min, 4ºC). Cell pellets were washed with cold PBS, before pelleting (500g, 5 min, 4ºC). Cells were blocked with cold PBS + 3% BSA at 4ºC for 10 min. Cells were washed three times with cold PBS + 3% BSA and 20 µg/mL biotinylated anti-RNF43 clone R3 IgG was added and incubated at 4ºC for 30 min. Cells were washed three times with cold PBS + 3% BSA and neutravidin-647 (1:1,000) was added and incubated at 4ºC for 30 min. Cells were washed three times with cold PBS + 3% BSA and finally resuspended in cold PBS. Flow cytometry was performed on a CytoFLEX cytometer (Beckman Coulter), and gating was performed on single

cells and live cells, before acquisition of 10,000 cells. Analysis was performed using the FlowJo software package.

IgG Expression

IgGs were expressed and purified from Expi293 BirA cells according to established protocol from the manufacturer. Briefly, 30 µg of pFUSE (InvivoGen) vector was transiently transfected into 75 million Expi293 BirA cells using the Expifectamine kit. Enhancer was added 20 h after transfection. Cells were incubated for a total of 6 d at 37ºC in a 8% CO_2 environment before the supernatants were harvested by centrifugation.

Fab Expression

Fabs were expressed and purified by an optimized autoinduction protocol previously described[14]. In brief, C43 (DE3) Pro + *E. coli* containing expression plasmids were grown in TB autoinduction media at 37ºC for 6 h, then cooled to 30ºC for 18 h. Fabs were purified by Protein A affinity chromatography. Purity was assessed by SDS/PAGE and intact protein mass spectrometry.

Bispecific IgG Expression and Assembly

Half IgG's with either knob or hole Fc domains were expressed using described IgG expression protocol. Hole half IgG's were purified using Protein A affinity chromatography. Knob half IgG's were purified using Nickel affinity chromatography, appropriating a C-terminal his tag. Half IgG's were mixed together in a 1:1 ratio in 10 mM Tris pH 7.5, 100 mM NaCl. The pH was adjusted to 8.5 with addition of 20% 800 mM L-Arg pH 10. 200-fold excess reduced glutathione in 800 mM L-Arg pH 10 was added and the mixture incubated overnight at 35ºC with 150 rpm shaking. Reaction mixture was buffer exchanged into PBS pH 7.0 and purified by Nickel affinity chromatography, appropriating a C-terminal his tag on the knob half IgG. Purity was assessed by SDS/PAGE and BLI.

RT-qPCR

CRISPR knockdown cell lines were verified using RNAseq. Total RNA was extracted using the Qiagen RNAeasy kit. Reverse transcription was performed on 1µg of RNA using a Quantitect Reverse Transcription Kit. qPCR was performed on a BioRad CFX Connect, using primers: RNF43 - Hs.PT.58.19324722 – 52ºC annealing temperature. And Actin - Hs.PT.39a.22214847 - 60ºC annealing temperature. Experiments were performed with three biological and three technical replicates.

Phage display selections

Phage display was performed as previously described[34]. In brief selections with antibody phage Library E or Library UCSF were performed using biotinylated RNF43 Fc-fusion as the positive antigen or Biotinylated Fc-fusion for the negative selection. A 'catch and release' strategy was used with streptavidin-coated magnetic beads (Promega) and TEV protease. Four rounds of selections were performed with decreasing RNF43 Fc-fusion (100 nM, 50 nM, 10 nM, 10 nM). Individual clones from the fourth round of selection were carried onto fab-phage ELISA.

Fab-phage ELISA

384 Maxisorp plates were coated with Neutravidin (10 µg/mL) overnight at 4ºC with gentle shaking and subsequently blocked with PBS + 0.05% Tween20 + 0.2% BSA (Blocking buffer) for 1 h at RT. Blocking buffer was removed and 20 nM biotinylated antigen or BSA was added and incubated at RT for 20 min. Antigen solution was removed and plates blocked for 10 min with 1 µM biotin in blocking buffer. The plate was washed three times with PBS + 0.05% Tween20. Phage diluted 1:5 in blocking buffer or blocking buffer with 20 nM soluble competitor was added and incubated for 20 min at RT. Plates were washed three times with PBS + 0.05% Tween20. Horseradish peroxidase (HRP)-conjugated anti-phage monoclonal antibody (GE LiveScience's

27-9421-01) diluted 1:5000 in blocking buffer was added and incubated for 30 min at RT. Plates were washed three times with PBS + 0.05% Tween20. HRP TMB substrate was added and incubated at RT until signal appeared. Reaction was quenched with 1 M phosphoric acid and plates analyzed at OD_{450nM} on a SpectraMax plate reader.

Confocal Microscopy

HeLa cells were transiently transfected with each described construct using FuGENE transfection reagent. After 24 h in complete growth medium cells were transferred to Mat-Tek 35 mm glass bottom petri dishes pre-treated with poly-d-lysine. After a further 12 h in complete growth medium cells were stained using standard protocols using DAPI (Cell Signaling Technologies) and LysoTracker (Invitrogen). Samples were imaged using a Nikon Ti Microscope with Yokogawa CSU-22 spinning disk confocal and a 100x objective lens. 405, 488 and 640 nM lasers were used to image DAPI, GFP, LysoTracker respectively with a step size of 0.1 μM. Images were deconvoluted and processed using NIS-Element and FIJI software packages.

1.6 Supplemental Figures

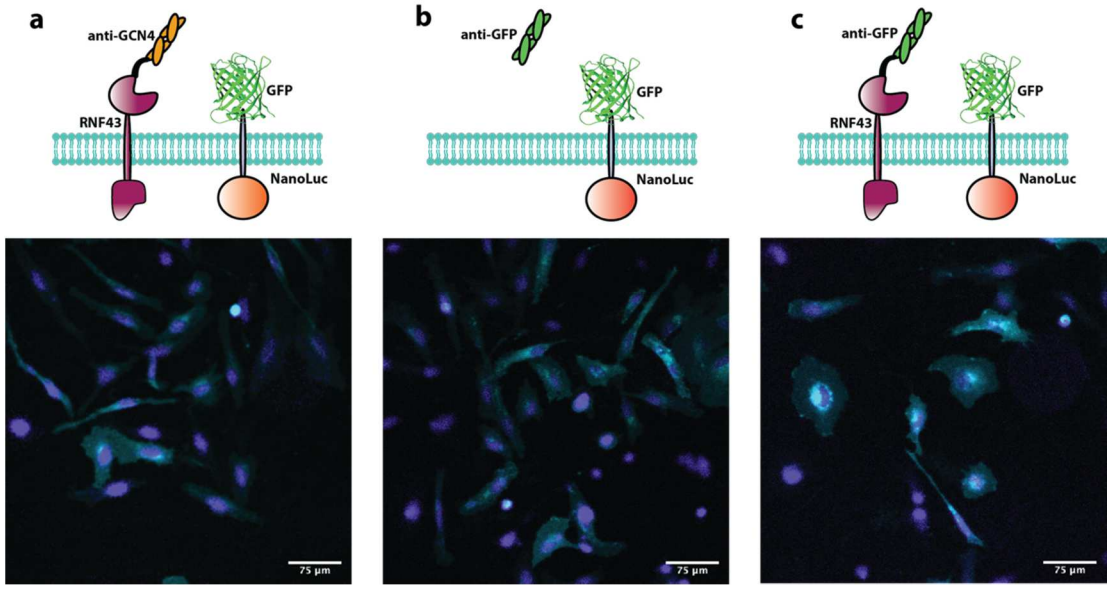


Figure 1.S1: Antibody-Based-PROTACs (AbTACs) recruit RNF43 to internalize cell surface proteins. a, b, c) Engineered RNF43 constructs and the GFP-Nanoluciferase reporter. With their corresponding wide-field confocal microscopy images showing GFP localization for each experimental condition. a) Anti-GCN4-RNF43 fusion as an isotype control. b) Soluble Anti-GFP Fab to control for Fab binding effects. c) Anti-GFP-RNF43. (Green - GFP reporter protein, Blue – DAPI)

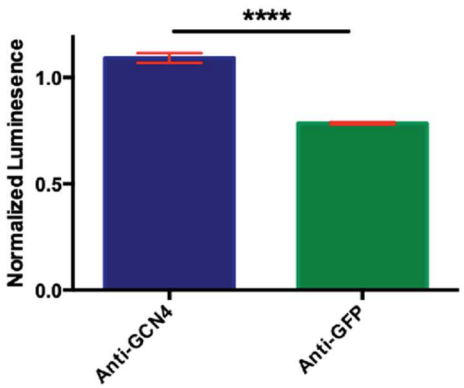


Figure 1.S2: NanoLuciferase assay showing decrease in reporter signal after co-expression of either Anti-GFP or Anti-GCN4 scFab-RNF43 fusions. Error bars represents SD of 3 biological replicates (P value = 0.00002).

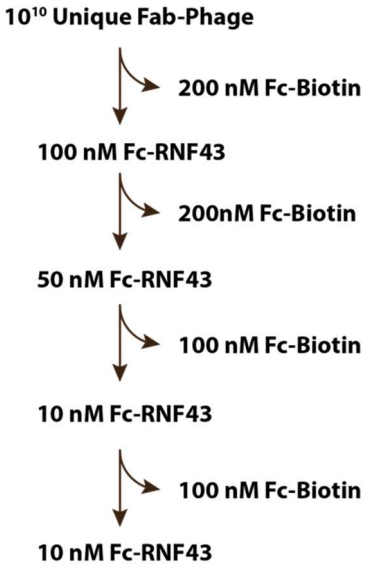


Figure 1.S3: Phage display selection strategy utilized against RNF43 Fc-fusion.

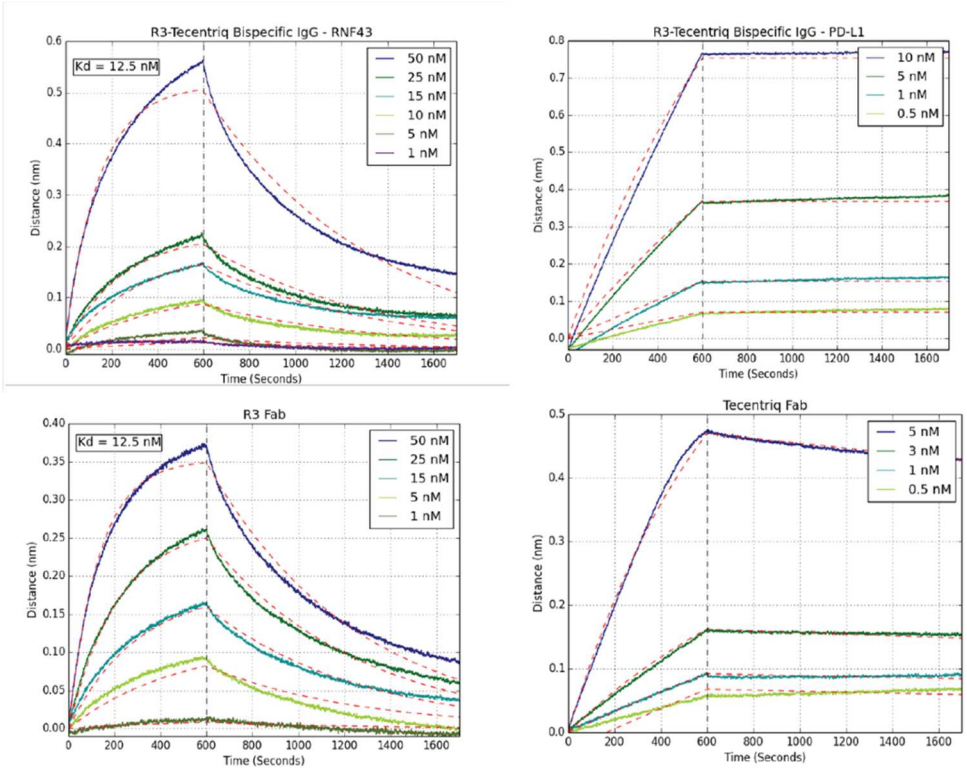


Figure 1.S4: Dose escalation BLI experiment showing binding of AC-1 either in Fab or BsIgG format to either RNF43 (left) or PD-L1 (right).

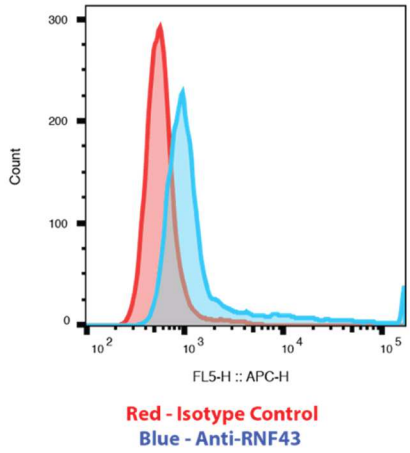


Figure 1.S5: Flow cytometry showing presence of RNF43 on MDA-MB-231 cells. Cells were stained with biotinylated IgG clone R3 and Neutravidin-Alexa Fluor 647.

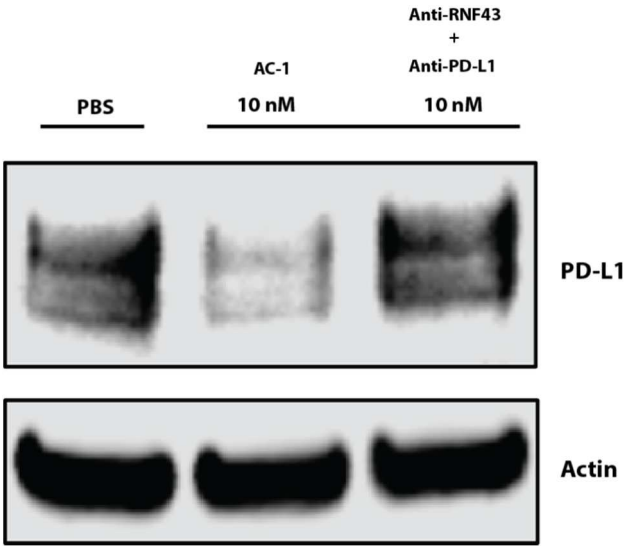


Figure 1.S6: Degradation experiment showing PD-L1 degradation is dependent on the dimerization of RNF43 and PD-L1, where simultaneous targeting has no effect of PD-L1 levels.

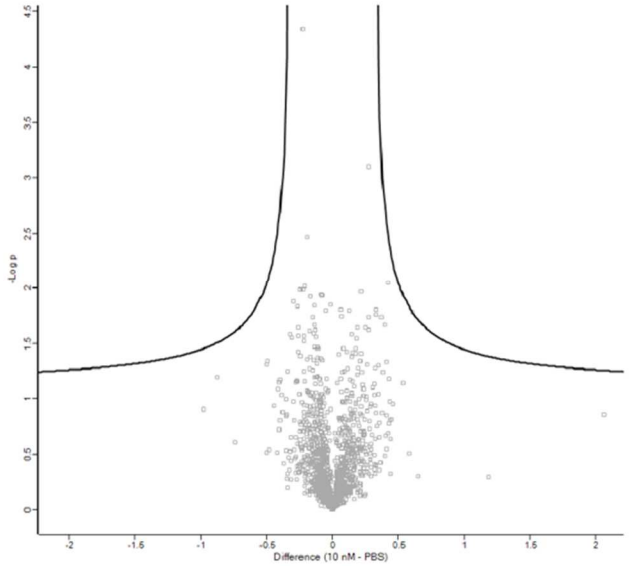


Figure 1.S7: Whole cell proteomics showing no significant change in protein expression in MDA-MB-231 cells after 24 hours treatment with 10 nM AC-1.

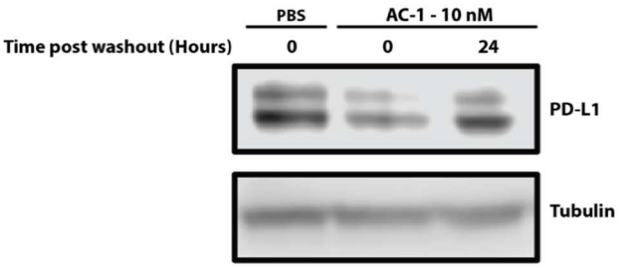


Figure 1.S8: Wash-out experiment showing PD-L1 levels largely recover 24 hours post treatment with 10 nM AC-1.

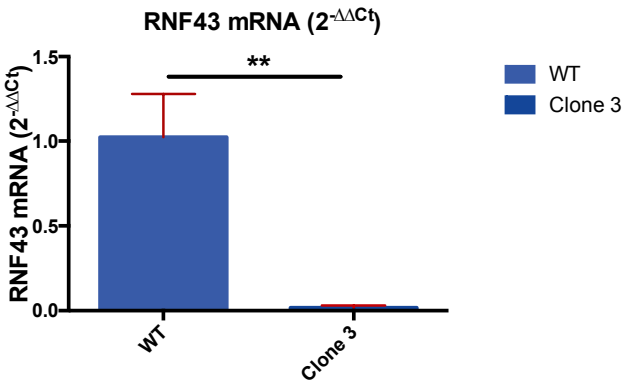


Figure 1.S9: qPCR confirms RNF43 knockdown in the engineered MDA-MB-231 cells. Error bars represents SD of 3 biological replicates and 2 technical replicates (P value = 0.0009).

Table 1.S1: Primers used for sgRNA cloning into the all-in-one CRISPR constructs.

Sequence Name	Sequence
sgRNA-RNF43-A-fwd	caccGGGCCCACTGGAATCCACGG
sgRNA-RNF43-A-rev	aaacCCGTGGATTCCAGTGGGCCC
sgRNA-RNF43-B-fwd	caccGTGGTTGCAGAGTAAGAAGG
sgRNA-RNF43-B-rev	aaacCCTTCTTACTCTGCAACCAC
sgRNA-RNF43-C-fwd	caccGTCGGGCCCACTGGAATCCA
sgRNA-RNF43-C-rev	aaacTGGATTCCAGTGGGCCCGAC
sgRNA-RNF43-D-fwd	caccGGGCACCTACCTGTAGTATG
sgRNA-RNF43-D-rev	aaacCATACTACAGGTAGGTGCCC
sgRNA-RNF43-E-fwd	caccGAGGCAGTATCTCTGAATCA
sgRNA-RNF43-E-rev	aaacTGATTCAGAGATACTGCCTC
sgRNA-RNF43-A-fwd	caccGGGCCCACTGGAATCCACGG

1.7 References

(1) Sakamoto, K. M.; Kim, K. B.; Kumagai, A.; Mercurio, F.; Crews, C. M.; Deshaies, R. J. Protacs: Chimeric Molecules That Target Proteins to the Skp1-Cullin-F Box Complex for Ubiquitination and Degradation. *Proceedings of the National Academy of Sciences* **2001**, *98* (15), 8554–8559. https://doi.org/10.1073/pnas.141230798.

(2) Banik, S. M.; Pedram, K.; Wisnovsky, S.; Ahn, G.; Riley, N. M.; Bertozzi, C. R. Lysosome-Targeting Chimaeras for Degradation of Extracellular Proteins. *Nature* **2020**, *584* (7820), 291–297. https://doi.org/10.1038/s41586-020-2545-9.

(3) Ahn, G.; Banik, S.; Miller, C. L.; Riley, N.; Cochran, J. R.; Bertozzi, C. Lysosome Targeting Chimeras (LYTACs) That Engage a Liver-Specific Asialoglycoprotein Receptor for Targeted Protein Degradation. *Chemrxiv* **2020**. https://doi.org/10.26434/chemrxiv.12736778.v1.

(4) Nabet, B.; Roberts, J. M.; Buckley, D. L.; Paulk, J.; Dastjerdi, S.; Yang, A.; Leggett, A. L.; Erb, M. A.; Lawlor, M. A.; Souza, A.; Scott, T. G.; Vittori, S.; Perry, J. A.; Qi, J.; Winter, G. E.; Wong, K.-K.; Gray, N. S.; Bradner, J. E. The DTAG System for Immediate and Target-Specific Protein Degradation. *Nature Chemical Biology* **2018**, *14* (5), 431–441. https://doi.org/10.1038/s41589-018-0021-8.

(5) Clift, D.; McEwan, W. A.; Labzin, L. I.; Konieczny, V.; Mogessie, B.; James, L. C.; Schuh, M. A Method for the Acute and Rapid Degradation of Endogenous Proteins. *Cell* **2017**, *171* (7), 1692-1706.e18. https://doi.org/10.1016/j.cell.2017.10.033.

(6) Naito, M.; Ohoka, N.; Shibata, N. SNIPERs—Hijacking IAP Activity to Induce Protein Degradation. *Drug Discovery Today: Technologies* **2019**, *31*, 35–42. https://doi.org/10.1016/j.ddtec.2018.12.002.

(7) Bondeson, D. P.; Mares, A.; Smith, I. E. D.; Ko, E.; Campos, S.; Miah, A. H.; Mulholland, K. E.; Routly, N.; Buckley, D. L.; Gustafson, J. L.; Zinn, N.; Grandi, P.; Shimamura, S.; Bergamini, G.; Faelth-Savitski, M.; Bantscheff, M.; Cox, C.; Gordon, D. A.; Willard, R. R.; Flanagan, J. J.; Casillas, L. N.; Votta, B. J.; den Besten, W.; Famm, K.; Kruidenier, L.; Carter, P. S.; Harling, J. D.; Churcher, I.; Crews, C. M. Catalytic in Vivo Protein Knockdown by Small-Molecule PROTACs. *Nat Chem Biol* **2015**, *11* (8), 611–617. https://doi.org/10.1038/nchembio.1858.

(8) Salami, J.; Alabi, S.; Willard, R. R.; Vitale, N. J.; Wang, J.; Dong, H.; Jin, M.; McDonnell, D. P.; Crew, A. P.; Neklesa, T. K.; Crews, C. M. Androgen Receptor Degradation by the Proteolysis-Targeting Chimera ARCC-4 Outperforms Enzalutamide in Cellular Models of Prostate Cancer Drug Resistance. *Communications Biology* **2018**, *1* (1). https://doi.org/10.1038/s42003-018-0105-8.

(9) Burslem, G. M.; Smith, B. E.; Lai, A. C.; Jaime-Figueroa, S.; McQuaid, D. C.; Bondeson, D. P.; Toure, M.; Dong, H.; Qian, Y.; Wang, J.; Crew, A. P.; Hines, J.; Crews, C. M. The Advantages of Targeted Protein Degradation Over Inhibition: An RTK Case Study. *Cell Chemical Biology* **2018**, *25* (1), 67-77.e3. https://doi.org/10.1016/j.chembiol.2017.09.009.

(10) Uhlen, M.; Fagerberg, L.; Hallstrom, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, A.; Kampf, C.; Sjostedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.-K.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P.-H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.; von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Ponten, F. Tissue-Based Map of the Human Proteome. *Science* **2015**, *347* (6220), 1260419–1260419. https://doi.org/10.1126/science.1260419.

(11) Lundstrom, K. An Overview on GPCRs and Drug Discovery: Structure-Based Drug Design and Structural Biology on GPCRs. *G Protein-Coupled Receptors in Drug Discovery* **2009**, *552*, 51–66. https://doi.org/10.1007/978-1-60327-317-6_4.

(12) Lobo, E. D.; Hansen, R. J.; Balthasar, J. P. Antibody Pharmacokinetics and Pharmacodynamics. *Journal of Pharmaceutical Sciences* **2004**, *93* (11), 2645–2668. https://doi.org/10.1002/jps.20178.

(13) Lowman, H. B.; Bass, S. H.; Simpson, N.; Wells, J. A. Selecting High-Affinity Binding Proteins by Monovalent Phage Display. *Biochemistry* **1991**, *30* (45), 10832–10838. https://doi.org/10.1021/bi00109a004.

(14) Hornsby, M.; Paduch, M.; Miersch, S.; Sääf, A.; Matsuguchi, T.; Lee, B.; Wypisniak, K.; Doak, A.; King, D.; Usatyuk, S.; Perry, K.; Lu, V.; Thomas, W.; Luke, J.; Goodman, J.; Hoey, R. J.; Lai, D.; Griffin, C.; Li, Z.; Vizeacoumar, F. J.; Dong, D.; Campbell, E.; Anderson, S.; Zhong, N.; Gräslund, S.; Koide, S.; Moffat, J.; Sidhu, S.; Kossiakoff, A.; Wells, J. A High Through-Put Platform for Recombinant Antibodies to Folded Proteins. *Mol Cell Proteomics* **2015**, *14* (10), 2833–2847. https://doi.org/10.1074/mcp.O115.052209.

(15) Kontermann, R. Dual Targeting Strategies with Bispecific Antibodies. *mAbs* **2012**, *4* (2), 182–197. https://doi.org/10.4161/mabs.4.2.19000.

(16) Rodriguez-Gonzalez, A.; Cyrus, K.; Salcius, M.; Kim, K.; Crews, C. M.; Deshaies, R. J.; Sakamoto, K. M. Targeting Steroid Hormone Receptors for Ubiquitination and Degradation in Breast and Prostate Cancer. *Oncogene* **2008**, *27* (57), 7201–7211. https://doi.org/10.1038/onc.2008.320.

(17) Lu, J.; Qian, Y.; Altieri, M.; Dong, H.; Wang, J.; Raina, K.; Hines, J.; Winkler, J. D.; Crew, A. P.; Coleman, K.; Crews, C. M. Hijacking the E3 Ubiquitin Ligase Cereblon to

Efficiently Target BRD4. *Chemistry & Biology* **2015**, *22* (6), 755–763. https://doi.org/10.1016/j.chembiol.2015.05.009.

(18) Ward, C. C.; Kleinman, J. I.; Brittain, S. M.; Lee, P. S.; Chung, C. Y. S.; Kim, K.; Petri, Y.; Thomas, J. R.; Tallarico, J. A.; McKenna, J. M.; Schirle, M.; Nomura, D. K. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem. Biol.* **2019**, *14* (11), 2430–2440. https://doi.org/10.1021/acschembio.8b01083.

(19) Spradlin, J. N.; Hu, X.; Ward, C. C.; Brittain, S. M.; Jones, M. D.; Ou, L.; To, M.; Proudfoot, A.; Ornelas, E.; Woldegiorgis, M.; Olzmann, J. A.; Bussiere, D. E.; Thomas, J. R.; Tallarico, J. A.; McKenna, J. M.; Schirle, M.; Maimone, T. J.; Nomura, D. K. Harnessing the Anti-Cancer Natural Product Nimbolide for Targeted Protein Degradation. *Nature Chemical Biology* **2019**, *15* (7), 747–755. https://doi.org/10.1038/s41589-019-0304-8.

(20) Zhang, X.; Crowley, V. M.; Wucherpfennig, T. G.; Dix, M. M.; Cravatt, B. F. Electrophilic PROTACs That Degrade Nuclear Proteins by Engaging DCAF16. *Nat Chem Biol* **2019**, *15* (7), 737–746. https://doi.org/10.1038/s41589-019-0279-5.

(21) Lu, M.; Liu, T.; Jiao, Q.; Ji, J.; Tao, M.; Liu, Y.; You, Q.; Jiang, Z. Discovery of a Keap1-Dependent Peptide PROTAC to Knockdown Tau by Ubiquitination-Proteasome Degradation Pathway. *Eur J Med Chem* **2018**, *146*, 251–259. https://doi.org/10.1016/j.ejmech.2018.01.063.

(22) Zebisch, M.; Xu, Y.; Krastev, C.; MacDonald, B. T.; Chen, M.; Gilbert, R. J. C.; He, X.; Jones, E. Y. Structural and Molecular Basis of ZNRF3/RNF43 Transmembrane Ubiquitin Ligase Inhibition by the Wnt Agonist R-Spondin. *Nat Commun* **2013**, *4* (1), 2787. https://doi.org/10.1038/ncomms3787.

(23) Koo, B.-K.; Spit, M.; Jordens, I.; Low, T. Y.; Stange, D. E.; van de Wetering, M.; van Es, J. H.; Mohammed, S.; Heck, A. J. R.; Maurice, M. M.; Clevers, H. Tumour Suppressor RNF43 Is a Stem-Cell E3 Ligase That Induces Endocytosis of Wnt Receptors. *Nature* **2012**, *488* (7413), 665–669. https://doi.org/10.1038/nature11308.

(24) Tsukiyama, T.; Fukui, A.; Terai, S.; Fujioka, Y.; Shinada, K.; Takahashi, H.; Yamaguchi, T. P.; Ohba, Y.; Hatakeyama, S. Molecular Role of RNF43 in Canonical and Noncanonical Wnt Signaling. *Mol Cell Biol* **2015**, *35* (11), 2007–2023. https://doi.org/10.1128/MCB.00159-15.

(25) Giannakis, M.; Hodis, E.; Jasmine Mu, X.; Yamauchi, M.; Rosenbluh, J.; Cibulskis, K.; Saksena, G.; Lawrence, M. S.; Qian, Z. R.; Nishihara, R.; Van Allen, E. M.; Hahn, W. C.; Gabriel, S. B.; Lander, E. S.; Getz, G.; Ogino, S.; Fuchs, C. S.; Garraway, L. A. RNF43 Is Frequently Mutated in Colorectal and Endometrial Cancers. *Nat Genet* **2014**, *46* (12), 1264–1266. https://doi.org/10.1038/ng.3127.

(26) Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7* (11), 1848–1857. https://doi.org/10.1021/cb3002478.

(27) Koerber, J. T.; Hornsby, M. J.; Wells, J. A. An Improved Single-Chain Fab Platform for Efficient Display and Recombinant Expression. *Journal of Molecular Biology* **2015**, *427* (2), 576–586. https://doi.org/10.1016/j.jmb.2014.11.017.

(28) Riching, K. M.; Mahan, S.; Corona, C. R.; McDougall, M.; Vasta, J. D.; Robers, M. B.; Urh, M.; Daniels, D. L. Quantitative Live-Cell Kinetic Degradation and Mechanistic

Profiling of PROTAC Mode of Action. *ACS Chemical Biology* **2018**, *13* (9), 2758–2770. https://doi.org/10.1021/acschembio.8b00692.

(29) Freeman, G. J.; Long, A. J.; Iwai, Y.; Bourque, K.; Chernova, T.; Nishimura, H.; Fitz, L. J.; Malenkovich, N.; Okazaki, T.; Byrne, M. C.; Horton, H. F.; Fouser, L.; Carter, L.; Ling, V.; Bowman, M. R.; Carreno, B. M.; Collins, M.; Wood, C. R.; Honjo, T. Engagement of the PD-1 Immunoinhibitory Receptor by a Novel B7 Family Member Leads to Negative Regulation of Lymphocyte Activation. *J. Exp. Med.* **2000**, *192* (7), 1027–1034. https://doi.org/10.1084/jem.192.7.1027.

(30) Escors, D.; Gato-Cañas, M.; Zuazo, M.; Arasanz, H.; García-Granda, M. J.; Vera, R.; Kochan, G. The Intracellular Signalosome of PD-L1 in Cancer Cells. *Sig Transduct Target Ther* **2018**, *3* (1), 26. https://doi.org/10.1038/s41392-018-0022-9.

(31) Li, C.-W.; Lim, S.-O.; Chung, E. M.; Kim, Y.-S.; Park, A. H.; Yao, J.; Cha, J.-H.; Xia, W.; Chan, L.-C.; Kim, T.; Chang, S.-S.; Lee, H.-H.; Chou, C.-K.; Liu, Y.-L.; Yeh, H.-C.; Perillo, E. P.; Dunn, A. K.; Kuo, C.-W.; Khoo, K.-H.; Hsu, J. L.; Wu, Y.; Hsu, J.-M.; Yamaguchi, H.; Huang, T.-H.; Sahin, A. A.; Hortobagyi, G. N.; Yoo, S. S.; Hung, M.-C. Eradication of Triple-Negative Breast Cancer Cells by Targeting Glycosylated PD-L1. *Cancer Cell* **2018**, *33* (2), 187-201.e10. https://doi.org/10.1016/j.ccell.2018.01.009.

(32) Zhang, F.; Qi, X.; Wang, X.; Wei, D.; Wu, J.; Feng, L.; Cai, H.; Wang, Y.; Zeng, N.; Xu, T.; Zhou, A.; Zheng, Y. Structural Basis of the Therapeutic Anti-PD-L1 Antibody Atezolizumab. *Oncotarget* **2017**, *8* (52), 90215–90224. https://doi.org/10.18632/oncotarget.21652.

(33) Ridgway, J. B. B.; Presta, L. G.; Carter, P. 'Knobs-into-Holes' Engineering of Antibody CH3 Domains for Heavy Chain Heterodimerization. *Protein Eng Des Sel* **1996**, *9* (7), 617–621. https://doi.org/10.1093/protein/9.7.617.

(34) Ovacik, A. M.; Li, J.; Lemper, M.; Danilenko, D.; Stagg, N.; Mathieu, M.; Ellerman, D.; Gupta, V.; Kalia, N.; Nguy, T.; Plaks, V.; Johnson, C. D.; Wang, W.; Brumm, J.; Fine, B.; Junttila, T.; Lin, K.; Carter, P. J.; Prabhu, S.; Spiess, C.; Kamath, A. V. Single Cell-Produced and in Vitro-Assembled Anti-FcRH5/CD3 T-Cell Dependent Bispecific Antibodies Have Similar in Vitro and in Vivo Properties. *mAbs* **2019**, *11* (2), 422–433. https://doi.org/10.1080/19420862.2018.1551676.

(35) Li, C.-W.; Lim, S.-O.; Xia, W.; Lee, H.-H.; Chan, L.-C.; Kuo, C.-W.; Khoo, K.-H.; Chang, S.-S.; Cha, J.-H.; Kim, T.; Hsu, J. L.; Wu, Y.; Hsu, J.-M.; Yamaguchi, H.; Ding, Q.; Wang, Y.; Yao, J.; Lee, C.-C.; Wu, H.-J.; Sahin, A. A.; Allison, J. P.; Yu, D.; Hortobagyi, G. N.; Hung, M.-C. Glycosylation and Stabilization of Programmed Death Ligand-1 Suppresses T-Cell Activity. *Nat Commun* **2016**, *7* (1), 12632. https://doi.org/10.1038/ncomms12632.

(36) Luca, V. C.; Miao, Y.; Li, X.; Hollander, M. J.; Kuo, C. J.; Garcia, K. C. Surrogate R-Spondins for Tissue-Specific Potentiation of Wnt Signaling. 14.

2. Engineering Antibody-based PROTACs (AbTACs) to elucidate characteristics necessary for optimal degradation of cell surface proteins

2.1 Abstract

Targeted protein degradation is a promising therapeutic strategy capable of overcoming limitations of traditional occupancy-based inhibitors. By ablating all of protein's associated functions at once, the event-driven pharmacology of degrader technologies has enabled targeting of proteins that have been historically deemed "undruggable". However, most degradation strategies utilize the ubiquitin-proteasome system to mediate target degradation and are thus limited to targeting proteins with ligandable cytoplasmic domains. While some of these strategies such as PROTACs have shown great promise, there is a need for additional modalities that can be applied to specifically target cell surface proteins. We previously published the development of Antibody-based PROTACs (AbTACs), which utilize an IgG bispecific antibody scaffold to bring the cell surface E3 ligase RNF43 into proximity of a membrane protein of interest (POI) to mediate its degradation. Here, we employ rational protein engineering strategies to interrogate the properties necessary for efficient degradation and apply these to degrade PD-L1 and EGFR. We show that the specific antibody binding epitopes on RNF43 and the POI are more important than the individual affinities of the AbTAC antibodies. Furthermore, we show that AbTACs are highly modular and amenable to a variety of different scaffolds that can range in flexibility, valency, and orientation of the binding arms. Lastly, we increase the available repertoire of E3 ligases by

showing that that we can co-opt the E3 ligase ZNRF3 to degrade both PD-L1 and EGFR. Taken together, we provide a roadmap for optimizing the development of future AbTACs thereby expanding their utility for targeted cell surface protein degradation.

2.2 Introduction

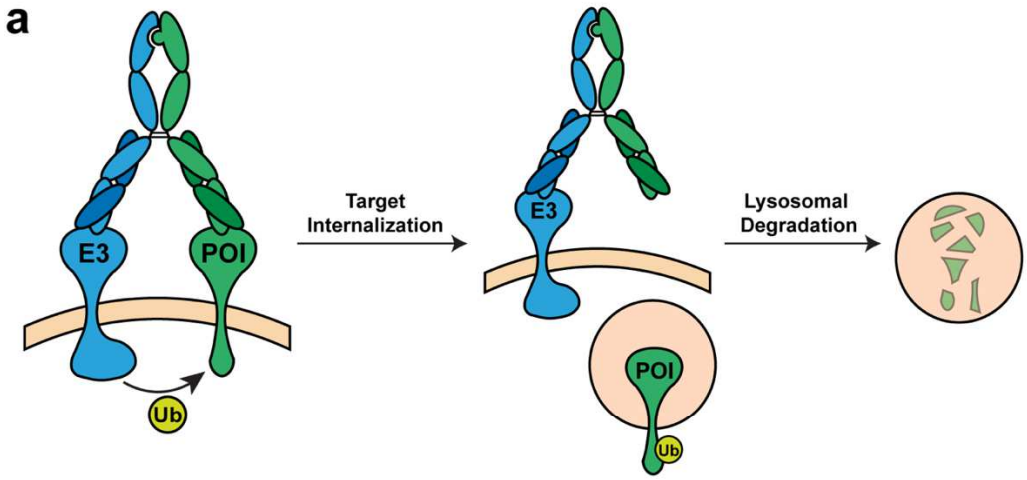
Targeted protein degradation has garnered increasing attention over the last decade as a promising therapeutic strategy due to its various advantages over conventional inhibition-based therapeutics[1]. Inhibitors rely on sustained occupancy-driven pharmacology necessitating high affinity binders capable of abrogating catalytic function[2]. As such, proteins that mediate disease states through alternate or even multiple independent functions have been traditionally deemed "undruggable". In contrast, degraders are catalytic and utilize event-driven pharmacology, alleviating the need for high affinity binders and durably abrogating all protein functions at once[2]. As such, degrader technologies such as proteolysis targeting chimeras (PROTACs) have had great success in targeting traditionally challenging proteins and several are currently being tried in the clinic[3]. However, most degrader technologies, including PROTACs, utilize an intracellular mechanism of action and are thus limited to targeting proteins with ligandable cytoplasmic domains[1]. To address this limitation and expand the scope of targeted protein degradation, two recent degradation strategies have been described, lysosomal targeting chimeras (LYTACs), and our previous work, Antibody-based PROTACs (AbTACs)[4]. LYTACS mediate degradation of cell surface proteins by co-opting lysosome shuttling receptors using an IgG-glycan bioconjugate[5,6]. AbTACs utilize an IgG bispecific antibody to bring a cell surface E3 ligase into proximity of a membrane protein of interest (POI) to mediate its degradation[4]. While LYTACs show great promise, we believe AbTACs to have several advantages. LYTACs require complex chemical

synthesis and *in vitro* bioconjugation of glycans, which are preferentially cleared in the liver. In contrast, AbTACs are fully recombinant, allowing for rapid genetic optimization of both binding properties and scaffold. Furthermore, the IgG scaffold on which the AbTAC is built, possesses favorable pharmacokinetic properties relative to LYTACS and other small molecule based degraders [7]. As such, we believe AbTACs can both complement and overcome limitations of current targeted protein degradation strategies and thus hold great therapeutic potential.

2.3 Results

AbTACs utilize a bispecific IgG scaffold to bring a cell surface E3 ligase in proximity of a protein of interest (POI) to mediate its degradation through the lysosomal degradation pathway (**Figure 2.1a**). In our recent AbTAC proof of concept publication, we developed an AbTAC to target and degrade the immune-checkpoint protein programmed death-ligand 1 (PD-L1) by co-opting the cell surface E3 ligase RNF43. Although we were able to successfully degrade PD-L1 on a variety of cell lines, the maximal degradation we were able to achieve was ~60% (**Figure 2.1b**). We hypothesized that low cell surface levels of RNF43 relative to PD-L1 might be preventing efficient POI degradation. To this end, we found that the relative abundance of RNF43 to PD-L1 RNA transcripts was low in the cell lines we had previously tested[8] **(Figure 2.1b)**. To determine whether we could improve degradation by increasing the relative surface levels of RNF43, we generated T24 cell lines overexpressing either WT RNF43 (T24 R-WT) or an RNF43 mutant in which the intracellular domain has been replaced with eGFP (T24 R-MUT) **(Figure 2.S1)**. Next, we used the Knob-into-Hole engineering strategy[9] to generate a bispecific IgG AbTAC R0/Atz, as previously described (**Figure 2.S2**)[4]. Briefly, the RNF43 binding arm is an antigen binding fragment (Fab) previously isolated from our in-house Fab phage library (clone

R0) and the PD-L1 binding arm is the clinically approved Atezolizumab (Atz). To measure levels of PD-L1 degradation, we chose to utilize a flow-based system using a commercially available fluorescent antibody that binds a separate epitope from Atezolizumab **(Figure 2.S3)**. Gratifyingly, when we performed a degradation assay with R0/Atz and measured surface levels of PD-L1 after 24hrs we found that PD-L1 degradation increased to 80% for the WT but not the mutant RNF43 overexpression cells **(Figure 2.1c)**. Taken together, these data indicate that degradation can be improved by either increasing the surface level of functional RNF43, or via careful cell-line selection.



b

Target	Cell Line	RNF43 TPM	PD-L1 TPM	% degradation
PD-L1	MDA-MB-231	5	35	63%
PD-L1	T24	1	35	45%
PD-L1	HCC827	3	35	40%

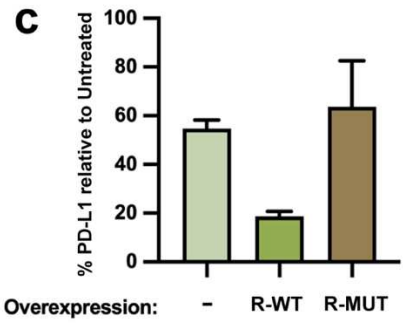


Figure 2.1: Relative levels of RNF43 and POI affect AbTAC mediated degradation. (a) Cartoon depiction of the AbTAC mediated degradation mechanism. (b) Table depicting RNA transcript levels for RNF43 and PD-L1 in different cell lines and the maximal degradation of PD-L1 achieved in reach respective cell line using an R0/Atz AbTAC. (c) Graph of degradation assays depicting surface levels of PD-L1 as measured by flow cytometry following 24hr incubatin of 10nM R0/Atz on either WT T24 cells (-), T24 cells overexrpressing WT RNF43 (T24 R-WT), or T24 cells overexpressing an RNF43 mutant in

which the intracellular domain has been replaced with eGFP (T24 R-MUT). PD-L1 levels are relative to un-treated cells. Data indicate that PD-L1 degradation is enhanced when surface levels of WT RNF43 are increased and are representative of at least three independent biological replicates.

Using the T24 R-WT cell line, we next sought to interrogate how changing the epitope to which the AbTAC binds on either the E3 ligase or POI might affect degradation efficiency. We utilized our in-house Fab phage display library to identify additional Fabs that bind non-overlapping RNF43 epitopes compared to our original RNF43 binder, clone R0. We isolated two additional RNF43 binders, clones R3 and R6, that bound by bio-layer interferometry (BLI) to purified RNF43 ectodomain with *in vitro* KDs of 60nM and 22nM respectively (**Figure 2.S4**). Additionally, as assessed by flow-cytometry, these clones bind to RNF43 on the surface of cells (**Figure 2.S5**). Importantly, we determined by epitope binning that each of the RNF43 binders, R0, R3, and R6, bound distinct epitopes (**Figure 2.S6**). Next, we generated R3/Atz and R6/Atz AbTACs as well as an Atz Dummy lacking an E3 ligase binding arm to use as a control for our degradation assays (**Figure 2.S2**). Interestingly, we observed robust PD-L1 degradation for all three AbTACs (R0/Atz, R3/Atz, R6/Atz), but the level of degradation did not correlate with the K_D of the E3 binder (**Figure 2.2a**). These data indicate that epitope could be playing a more substantial role in determining efficient POI degradation than affinity. To further test this hypothesis, we wondered how detuning the affinity for a given epitope binder might affect degradation. We performed an alanine scan of the heavy chain complementary determining region 3 (CDR3) of each clone. We chose mutants with incrementally lower affinities from their parental Fab (**Figure 2.S4**) and generated AbTACs for each of these mutants (**Figure 2.S2**). When we tested these AbTACs for their ability to degrade PD-L1, we found that decreased affinity correlated with decreased degradation (**Figure 2.2b**). However, affinity for RNF43 could be detuned

substantially without major loss of degradation, indicating that the specific epitope on the E3 ligase to which the AbTAC binds appears to be the driving property in promoting effective POI degradation.

We next wondered how changing the affinity and epitope on the POI arm of the AbTAC might affect degradation. Based on the known structure of atezolizumab in complex with PD-L1, we made alanine mutations in key interacting residues of Atezolizumab's CDRs[10]. We chose mutants with a range of binding affinities (K_D, 0.33-458 nM) (**Figure 2.S4**) and generated the corresponding R0/Atz AbTACs (**Figure 2.S2**). When we tested these constructs for their ability to degrade PD-L1, we found that degradation correlated strongly to the K_D ($R^2 = 0.93$) (**Figure 2.2c**). Interestingly, whereas detuning the affinity on the E3 binding arm to the high nM range only slightly abrogated degradation, detuning affinity to 450nM on the POI arm almost completely abolished degradation suggesting a greater importance of affinity on the POI side. We next sought to test how changing the epitope on the POI might affect degradation. However, due to the lack of unique PD-L1 epitope binders, we chose to interrogate a new target EGFR, for which an abundance of different EGFR antibodies and their corresponding epitope binding and CDR sequence information is readily available. We performed an initial proof of degradation assay by generating an AbTAC utilizing the clinically approved Cetuximab for the EGFR binding arm, R0/Ctx, as well as a Dummy control (**Figure 2.S2**). Gratifyingly, when we tested this construct for its ability to degrade EGFR on T24 R(WT) cells we observed roughly 50% degradation after 24hrs as determined by western blot analysis. (**Figure 2.2d**). Therefore, we generated five additional R0/EGFR AbTACs and the corresponding Dummy controls (**Figure 2.S2**) using EGFR binders annotated in the literature to bind separate epitopes from each other[11]. Interestingly, when we tested these constructs for their ability to degrade EGFR on T24 R(WT) cells, two AbTACs showed no

EGFR degradation while the three other degraded EGFR between 40% and 50%, equitable to R0/Ctx (**Figure 2.2e, 2.S7**). Importantly, Dummy versions of each EGFR binder did not show any degradation of eGFR (**Figure 2.S7**). Interestingly, there was little correlation between affinity of the EGFR binding arm and the level of EGFR degradation (**Figure 2.2f**) indicating that the epitope on the POI to which the AbTAC binds also plays a substantial role in degradation. Next, we wondered whether degradation of PD-L1 and EGFR could be recapitulated on a non-engineered cell line to ensure the general applicability of our previous observations. We chose to use HCC2935 cells, an adenocarcinoma tumor cell line, as they exhibit higher levels of RNF43 expression based on RNA transcript data[8]. Gratifyingly, we found that both R0/Atz and R3/Atz could robustly degrade PD-L1 (**Figure 2.2g**) and EGFR (**Figure 2.2h**); interestingly we observed higher levels of degradation than for the over-expression line, T24 R(WT).

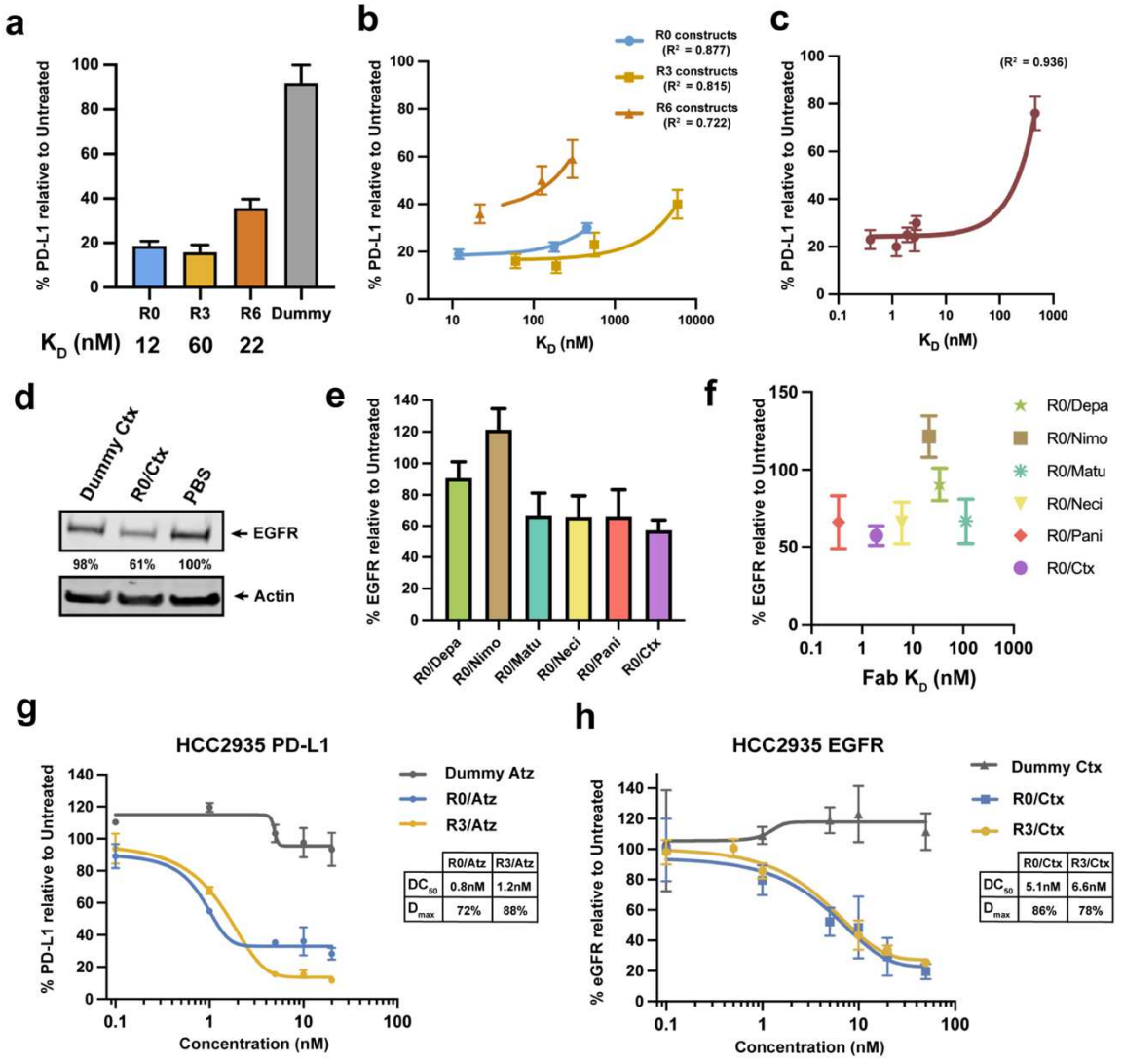


Figure 2.2 Epitope and affinity of E3 ligase and POI binding arms affect AbTAC mediated degradation. (a-c) Graphs of degradation assays depicting surface levels of PD-L1 as measured by flow cytometry following 24hr incubation of 10nM of the indicated AbTAC on T24 R-WT cells. PD-L1 levels are relative to un-treated cells. (a) Graph showing that ABTACs with binding arms that bind different RNF43 epitopes degrade PD-L1 to different levels. (b) Graph showing that affinity of RNF43 binding arm is correlated to degradation efficiency. (c) Graph showing that affinity of the POI binding arm correlates to degradation efficiency. (b-c) Linear regression analysis was utilized to determine correlation. (d-f) Degradation assays depicting levels of EGFR as measured by western blot following 24hr incubation of 10nM of the indicated AbTAC on T24 R-WT cells. EGFR levels are relative to un-treated cells. (d) Western blot showing that R0/Ctx can degrade EGFR. (e-f) Depa = Depatuxizumab[12], Nimo = Nimotuzumab[13], Matu = Matuzumab[14], Neci = Necitumumab[15], Pani = Panitumumab[16], Ctx = Cetuximab[13] (e) Graph showing that ABTACs with binding arms that bind different EGFR epitopes degrade EGFR to different levels. (f) Graph depicting that affinity of a specific EGFR binder arm does not correlate with degradation

efficiency. (g-h) Graphs of degradation assays depicting levels of PD-L1 (g) or EGFR (h) as measured by western blot following 24hr incubation of the indicated AbTAC at the indicated concentration on HCC2935 cells. PD-L1 and EGFR levels are relative to un-treated cells. (a-c, e-h) Data are representative of at least three independent biological replicates.

Given our new understanding of the importance of epitope and affinity we were interested in how altering the valency, flexibility, and orientation of the AbTAC binding arms might affect degradation efficiency. To this end, we generated 5 different constructs by fusing a Atezolizumab scFv domain to either the N- or C-terminus of a Fab or IgG scaffold (**Figure 2.3**). Because AbTACs are fully recombinant we were able to rapidly generate each of these constructs for both the R0 and R3 RNF43 binders as well as a Dummy control binder (**Figure 2.S2**). Gratifyingly, when T24 R(WT) cells were treated with each of the R0 and R3 bispecific constructs we observed PD-L1 degradation (**Figure 2.3**), whereas no change in protein levels were seen for the control constructs (**Figure 2.S8**). Interestingly, while the IgG scFv fusions have a higher D_{max}, the original Knob-into-Hole constructs have a lower DC_{50}. The light chain scFv fusions appear to be the exception to this observation as they exhibit both the highest maximal degradation (97% and 99% for R0 and R3 constructs respectively) as well as the most favorable DC_{50}. Taken together, these data indicate that AbTACs are amenable to a variety of different scaffolds, and thus highly modular.

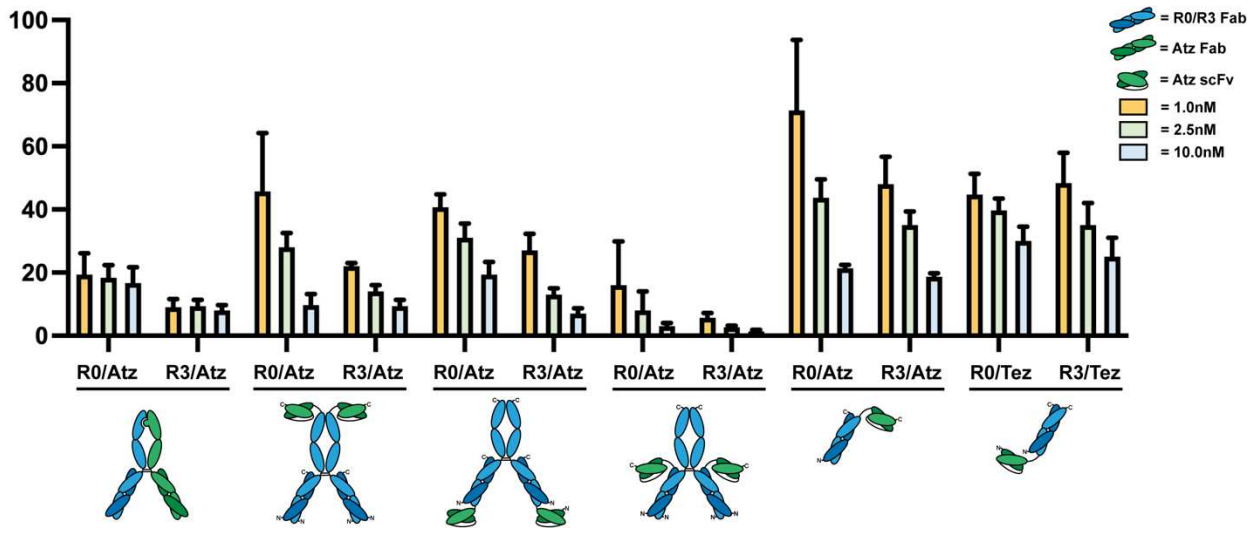


Figure 2.3: R0 and R3 AbTAC mediated degradation of PD-L1 is amenable to a variety of different scaffolds. Graphs of degradation assays depicting surface levels of PD-L1 as measured by flow cytometry following 24hr incubation of the indicated AbTAC at different concentrations on T24 R-WT cells. PD-L1 levels are relative to un-treated cells. Data are representative of at least three independent biological replicates.

Next, we hypothesized that ZNRF3, another member of the PA-TM-RING family of E3 ligases, and close homolog of RNF43 could be co-opted by our AbTAC strategy to degrade PD-L1 and EGFR. Similar to RNF43, ZNRF3 also negatively regulates Wnt signaling by inducing degradation of the membrane receptor Frizzled. Utilizing our in-house Fab phage display library, we were able to isolate a single Fab, Z18, that both bound ZNRF3 with an *in vitro* KD of 21nM (**Figure 2.S4**) and to ZNRF3 on the surface of cells (**Figure 2.S9**). Next, we generated a Z18/Atz AbTAC (**Figure 2.S2**) and found that it could degrade roughly 55% of PD-L1 on WT T24 cells (**Figure 2.4a**). To try and improve degradation, we generated a T24 cell line overexpressing WT ZNRF3 (**Figure 2.S10**). Gratifyingly, when we incubated the Z18/Atz AbTAC with this overexpression cell line, degradation of PD-L1 increased to 95% (**Figure 2.4a**). Next, we decided to apply the same engineering strategies on the ZNRF3 AbTAC as we had for the RNF43 AbTACs. To this end, we utilized alanine scanning to generate two additional Z18 mutants with detuned

affinities for ZNRF3 **(Figure 2.S2, 2.S4**). Similar to what we observed for the RNF43 AbTACs, the affinity of the Z18 AbTACs could be substantially detuned without much loss of degradation, again indicating that the specific epitope on the E3 ligase to which the AbTAC binds might be more important than having a high affinity binder (**Figure 2.4c**). Next, we generated the same 5 alternate AbTAC scaffolds we had previously used (**Figure 2.S2**) and tested these new constructs for their ability to degrade PD-L1. All constructs were able to robustly degrade PD-L1 (**Figure 2.4c**) showing that Z18 AbTACs are also highly modular and amenable to a variety of different scaffolds. Gratifyingly, we were also able to apply the Z18/Atz and a Z18/Ctx AbTAC to effectively degrade PD-L1 (**Figure 2.4d**) and EGFR (**Figure 2.4e**) respectively on non-engineered WT HCC2935 cells. Lastly, we performed a luciferase reporter assay to test the effects AbTACs have on potentiating WNT signaling. We treated HEK 293 Super Top Flash (STF) WNT reporter cells (HEK 293 STF) that had been transduced to over-express PD-L1, with R0/Atz and/or Z18/Atz AbTACs. Importantly, in the AbTAC treated conditions for both cell-lines, we observed no significant change in signal indicating that the AbTACs do not interfere with canonical WNT signaling **(Figure 2.S11)**.

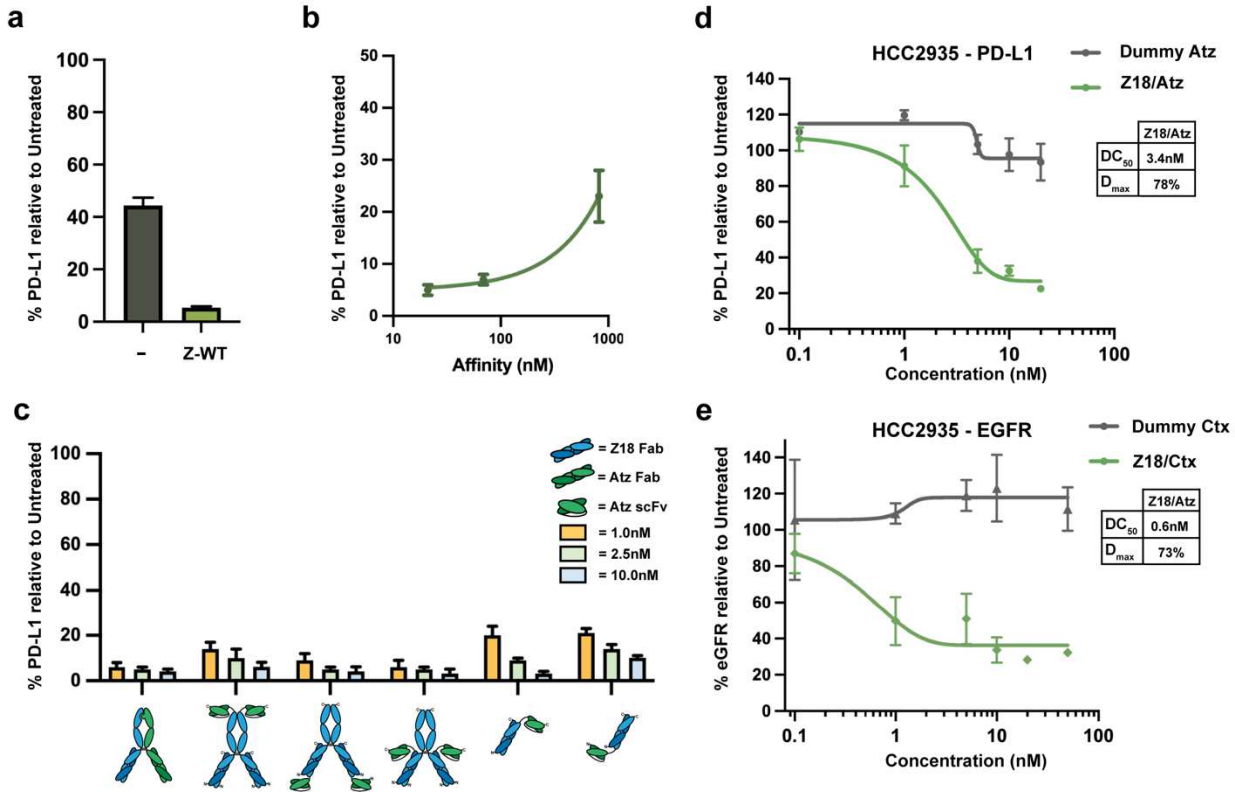


Figure 2.4: ZNRF3 AbTACs can be utilized to effectively degrade PD-L1 and eGFR. (a-c) Graphs of degradation assays depicting surface levels of PD-L1 as measured by flow cytometry following 24hr incubation of the indicated AbTAC. 10nM of AbTAC and T24 Z-WT cells were utilized unless otherwise specified. PD-L1 levels are relative to un-treated cells. (a) Graph indicating Z18/Atz mediated PD-L1 degradation is enhanced on T24 cells overexpressing WT ZNRF3 (Z-WT) compared to just WT T24 cells (-) (b) Graph showing that the affinity of the ZNRF3 binding arm correlates to degradation efficiency. Linear regression analysis was utilized to determine correlation. (c) Graph of degradation assays indicating that ZNRF3 mediated degradation of PD-L1 is amenable to a variety of different scaffolds. (g-h) Graphs of degradation assays depicting levels of PD-L1 (g) or EGFR (h) as measured by western blot following 24hr incubation of the indicated AbTAC at the indicated concentration on HCC2935 cells. PD-L1 and EGFR levels are relative to un-treated cells. (a-e) Data are representative of at least three independent biological replicates. Indicated PD-L1 or EGFR levels are relative to un-treated cells.

2.4 Discussion

In Summary, we employed a variety of antibody engineering strategies to interrogate the properties necessary for efficient AbTAC mediated degradation. Because the epitopes to which

the AbTAC binds plays a more substantial role than affinity for efficient degradation, future AbTAC discovery campaigns would benefit by focusing on generating unique binders. Indeed, the affinity of an AbTAC can be detuned substantially without losing much degradation efficiency, eliminating the need for high affinity binders. We've shown that the fully recombinant nature of the AbTAC scaffold can be taken advantage of by rapidly generating AbTACs with different E3 ligase and POI binding arms. Additionally, we've improved the modularity of AbTACs by showing that they are amenable to a variety of different scaffolds, further simplifying their production by eliminating the need for Knob-into-Hole assemblies. Lasty, we've shown that other cell surface E3 ligases can be co-opted by creating AbTACs that utilize ZNRF3 to degrade PD-L1 and EGFR. Taken together, we provide a roadmap for optimizing the development of future AbTACs thereby expanding their utility for targeted cell surface protein degradation.

2.5 Materials and Methods

Cell culture

HEK 293T, T24, and HCC2935 cell lines were grown and maintained at 37°C and 5% $C0_2$. HEK293T cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (P/S). T24 cells were grown in McCoy's 5a supplemented with 10% FBS and 1% P/S. HCC2935 cells were grown in RPMI supplemented with 10% FBS and 1% P/S. T24 and HEK 293T WT cells were obtained from the UCSF Cell Culture Facility. HEK 293 Super Top Flash (STF) WNT reporter cells (HEK 293 Stf) and HCC2935 cells were obtained from the American Type Culture Collection (ATCC).

Bio-Layer Interferometry (BLI)

BLI data was measured using an Octet RED384 (ForteBio) instrument. RNF43(Extracellular domain)-Fc or ZNRF3(Extracellular domain)-Fc fusions were immobilized on a streptavidin biosensor and loaded until 1.0 nm signal was achieved. After blocking with 10μM Biotin, purified Fabs at the indicated concentrations in Phosphate buffered saline containing 0.1% Tween and 0.2% Bovine serum albumin (PBS-T 0.2% BSA) were used as the analytes. Data were analyzed using the ForteBio Octet analysis software and kinetic parameters for each Fab were determined using either a 1:1 monovalent (For RNF43 R0 clones and ZNRF3 Z18 clones) or 1:2 heterobifunctional binding model (For RNF43 R3 and R6 clones). Epitope binning experiments for RNF43 clones R0, R3, and R6 were performed by first incubating RNF43-Fc bound tips with 50nM of Fab1 and then with both 50nM Fab1 and 25nM of Fab2.

Generation of Overexpression cell lines

T24 R-WT, T24 R-MUT, T24 Z-WT, HEK 293T R-MUT, and HEK 293T Z-MUT overexpression cells were made using a lentiviral transduction system: R-WT = N-terminally Myc

Tagged RNF43, R-MUT = N-terminally Myc Tagged RNF43 in which the C-terminal intracellular domain has been replaced with eGFP, Z-WT = N-terminally Myc Tagged ZNRF3, Z-MUT = N-terminally Myc Tagged ZNRF3 in which the C-terminal intracellular domain has been replaced with eGFP. The indicated protein sequences were cloned into a pCDH-EF1-FHC (Addgene plasmid #64874) vector and then each was transfected along with standard packaging vectors into HEK 293T cells to generate lentivirus. Media containing virus was collected 72hrs after transfection and filtered using a 0.45μM filter. The filtered lentivirus containing media was then used to transduce either T24 or HEK293T cells and stably transduced cells were selected for with Puromycin (0.8μg/ml). Successful transduction and expression of each protein was confirmed using flow cytometry.

Fab-Phage display selections

Phage display was performed as previously described[1]. In brief, selections with Fab-phage Library E and Library UCSF were performed using biotinylated RNF43-Fc or ZNRF3-Fc fusions as the positive antigens and Biotinylated Fc for the negative selections. A 'Catch and release' strategy was utilized with streptavidin-coated magnetic beads (Promega) and TEV protease. Four rounds of selections were used with each successive round using a decreased concentration of Antigen (1000nM,100nM, 50nM, 10nM) to selective for higher affinity binders. 96 clones were then chosen to be analyzed using a Fab-phage ELISA as previously described[1]. Clones that looked promising via ELISA were then sequenced to determine the identity of the CDRs. Unique clones were cloned into a Fab Expression vector and expressed and purified as Fabs and analyzed by BLI.

Expression and Purification of Fabs

Fabs were expressed and purified using an optimized autoinduction protocol that has been previously described[2]. In brief, C43 (DE3) Pro + *E. coli* containing expression plasmids for the

indicated Fabs were grown in TB autodinduction media at 37°C for 6hrs and then switched to 30°C for roughly 18hrs. Fabs were purified by Protein A affinity chromatography and buffer exchanged into PBS. Purity was assessed by SDS/PAGE.

Expression and Purification of AbTACs

RNF43-Fc, ZNRF3-Fc, Fc, Half-IgGs, Dummy IgGs, IgG/scFv fusions, and scFv/Fab fusions were expressed in Expi293 BirA cells (ThermoFisher) according to an established protocol from the manufacturer. Briefly, a ratio of 30ug vector/75 millions cells was used for transient transfections using the Expifectamine kit (Thermofisher). After roughly 20hrs, enhancer was added to the transfection and cells were then grown for an additional 72hrs at 37°C and 8% CO_2. Cells were then pelleted by centrifugation (4000rpm, 4°C, 20 minutes) and the supernatant was collected and filtered using a 45μM filter. Proteins were then purified using standard Protein A affinity chromatography followed by concentration and buffer exchange into PBS containing 20% glycerol. Knob and Hole half IgGs were mixed together in a 1:1.5 ratio of Knob:Hole in buffer containing 10mM Tris, 100mM NaCl and 160mM L-Arginine, pH 8.5 and 200 fold excess (relative to concentration of Knob+Hole) of reduced glutathione. Excess Hole was added to try to react with all available Knob. Unreacted Hole as well as Hole-Hole dimers, which are more likely to form during the assembly process, are removed during a later Nickel purification step using a Histidine tag located at the C-terminus of the Knob construct. Assembly mixtures were incubated overnight at 35°C with 150rpm shaking for 12 hrs. Mixtures were then diluted 5 fold into PBS containing 10mM Imidazole and the pH was adjusted to pH 8. Each reaction mixture was then incubated with Ni-NTA resin for 1hr at 4°C and gentle shaking. Assembly reactions were then added to a gravity column and the mixture was allowed to flow through. Ni-NTA beads were then washed 3X with PBS containing 10mM Imidazole pH 8. Knob-into-Hole constructs were then eluted by adding

1ml of PBS containing 300mM Imidazole. Elutions were then concentrated and buffer exchanged into PBS containing 20% glycerol. For all AbTACs, following concentration and buffer exchange into PBS containing 20% glycerol, samples were aliquoted, flash frozen and stored at -80°C.

Flow cytometry

Cells were washed with room temperature PBS, and then lifted using Versene. Cells were then added to wells of a 96 well dish and pelleted by centrifugation (500g, 5minm, 4°C). Cell pellets were washed 1X with 200ul cold PBS containing 3% BSA and then pelleted again by centrifugation. Cells were then resuspended and incubated with 200ul of cold PBS containing 3% BSA and the indicated primary antibody at a concentration of 10mg/ml for 20 minutes on ice. For all PD-L1 degradation assays, cells were incubated with 200ul of cold PBS containing 3% BSA and a final concentration of 50nM Atezolizumab for 20 minutes. Following incubation with primary antibody, cells were washed three times with 200ul of cold PBS containing 3% BSA and then resuspended and incubated with 200ul of cold PBS containing 3% BSA and the indicated secondary antibody at a dilution of 1:1000. For all PD-L1 degradation assays, cells were incubated with 200ul of cold PBS containing 3% BSA and a 1:500 dilution of (D8T4X)-647 (Cell Signaling) or a Rabbit IgG-647 (Cell signaling) isotype control. Following 30 minutes of incubation on ice, cells were washed 3X with 200ul of cold PBS containing 3% BSA. Flow cytometry was performed on a CytoFLEX cytometer (Beckman Coulter) and gating was performed on single cells and live cells before acquisition of cells. Analysis was performed using the FLowJo software package.

Western blotting

Cells were washed with room temperature PBS, lifted using Versene and pelleted by centrifugation (500g, 5minm, 4°C). Cells were then lysed with RIPA lysis buffer containing cOmplete mini protease inhibitor (Sigma) on ice for 20 minutes. Lysates were pelleted (17000g,

10min, 4°C) and the supernatant was removed and mixed with 4X NuPAGE LDS sample buffer (Invitrogen) and 1% BME. Equal amounts of lysates were loaded onto a 4-12% Bis-Tris gel and ran at 200V for 40 minutes. The gel was then transferred to a polyvinylidene difluoride (PVDF) membrane using an iBLot2 and standard manufacturer protocol (Thermo). The membrane was then blocked in Odyssey Blocking Buffer (TBS) (LICOR) for 45 minutes at room temperature with gentle shaking. The membranes were then incubated overnight with primary antibodies for anti-β-Actin (8H10D10: Cell Signaling), and either anti-PD-L1 (E1L3N: Cell Signaling) or anti-EGFR (D38B1: Cell Signaling). Following overnight incubation, membranes were washed three times with TBS + 0.1% Tween (TBS-T) and then incubated with secondary antibodies for 1hr at room temperature: 680RD Goat anti-mouse IgG and 800RD Goat anti-Rabbit IgG (LI-COR Biosciences). Membranes were washed again 3X with TBS-T and then analyzed using an OdysseyCLxImager (LI-COR Biosciences). Band intensities were quantified using Image Studio Software (LI-COR Biosciences).

Degradation experiments

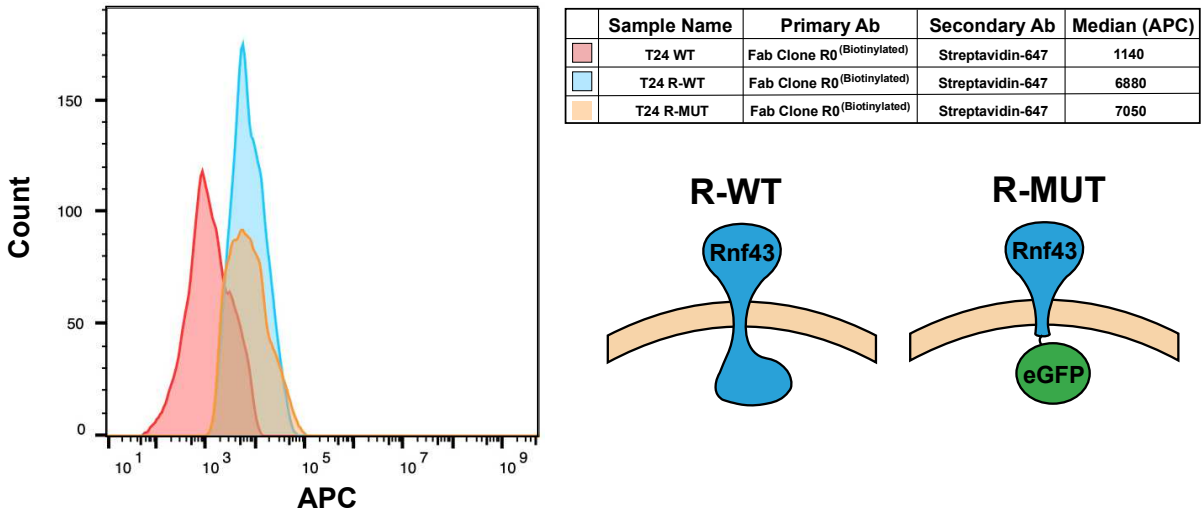
For all degradation assays using T24 cells, the indicated T24 cells were plated into wells of a 12 well dish at a density of between 10,000 and 15,000 cells per well and allowed to grow for roughly 72 hrs. Cells were then treated with 10nM (unless otherwise indicated) of AbTAC for 24hrs. For all PD-L1 degradation assays on T24 cells, cells were analyzed using the previously described flow cytometry workflow. Background APC-A signal from cells incubated with isotype control was subtracted for each of the samples and the relative surface levels of PD-L1 were then determined by dividing the signal of each treated sample by the signal of the untreated sample. For all EGFR degradation assays using T24 cells, the previously described western blotting protocol was used. We determined the relative total levels of EGFR, by calculating the ratio of EGFR/β-

Actin for each individual sample and dividing this by the ratio of EGFR/β-Actin for the untreated sample. For all degradation assays using HCC2935 cells, cells were plated into wells of 6 well dish. When cells were at roughly at 70% confluency, they were treated with AbTAC for 24hrs. Cells were then harvested and analyzed using the previously described western blot workflow to determine relative levels of PD-L1 and EGFR.

Wnt Activation assay

HEK 293T cells that have been stably tranduced with Firefly Luciferase Reporter under the control of seven LEF/TCF binding sites (HEK 293 STF) were purchased from ATCC. These cells were further transiently transfected to express PD-L1 under a CMV promoter. 36 hours after transfection, cells were stimulated with 20% WNT3a conditioned media (gift from the Mattis Lab at UCSF) supplemented with 25 nM RSPO2 (R&D Systems), R0/Atz (10 nM or 100 nM), Z18/Atz (10 nM or 100 nM), or a 1:1 mixture of R0/Atz and Z18/Atz (10 nM or 100 nM). Cells were cultured in the presence of reagents for another 24 hours before the addition of ONE-Glo Luciferase reagent (Promega). Cells were incubated in the dark for 15 minutes before being transferred to a white 96-well plate and imaged using a Tecan Infinite M200 Pro plate reader and analyzed by GraphPad Prism 7.

2.6 Supplemental Figures



	Sample Name	Primary Ab	Secondary Ab	Median (APC)
	T24 WT	Fab Clone R0 (Biotinylated)	Streptavidin-647	1140
	T24 R-WT	Fab Clone R0 (Biotinylated)	Streptavidin-647	6880
	T24 R-MUT	Fab Clone R0 (Biotinylated)	Streptavidin-647	7050

Figure 2.S1: Cell surface levels of RNF43. Flow cytometry measurement of the cell surface levels of RNF43 on either T24 WT cells, T24 cells overexrpressing WT RNF43 (T24 R-WT), or T24 cells overexpressing an RNF43 mutant in which the intracellular domain has been replaced with eGFP (T24 R-MUT).

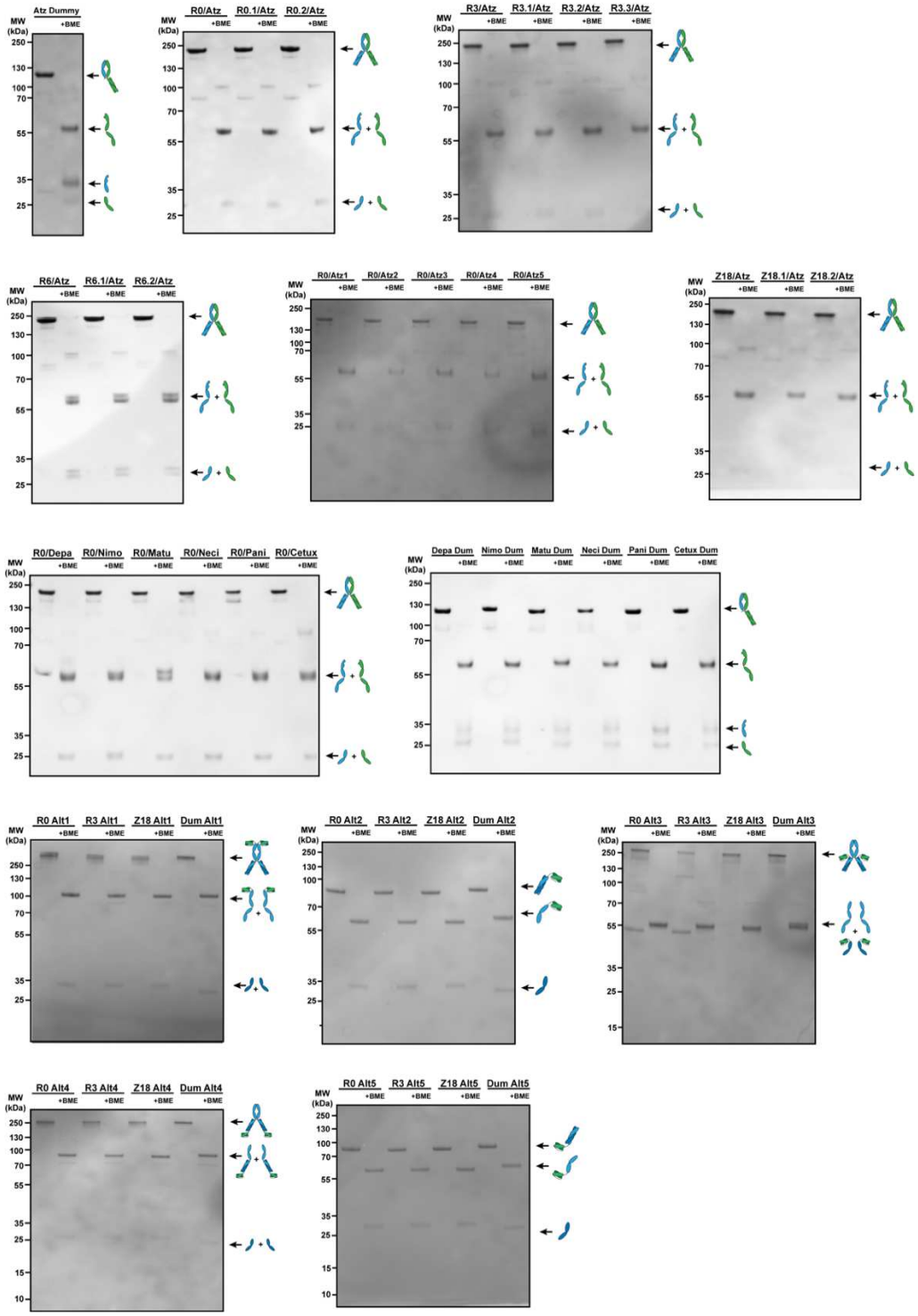
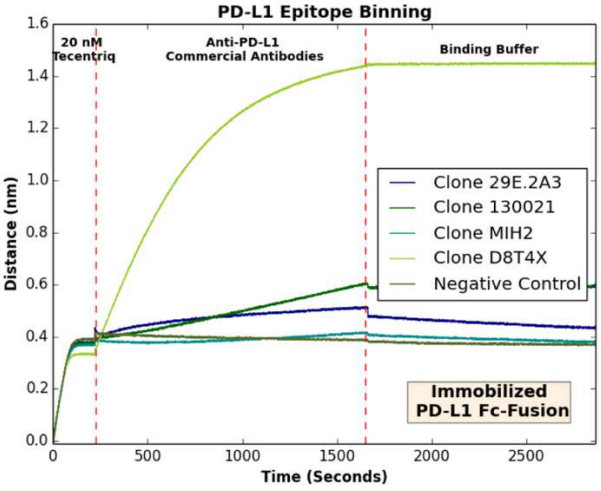
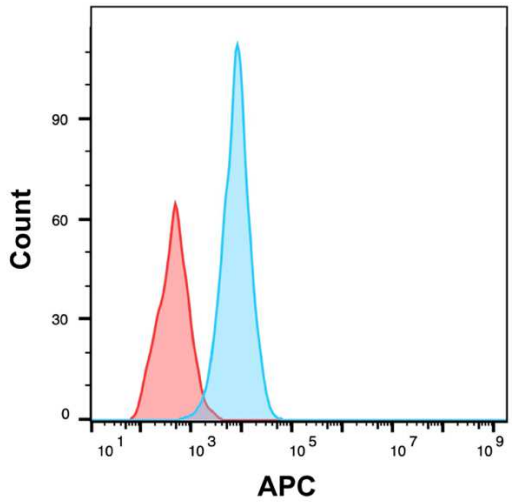


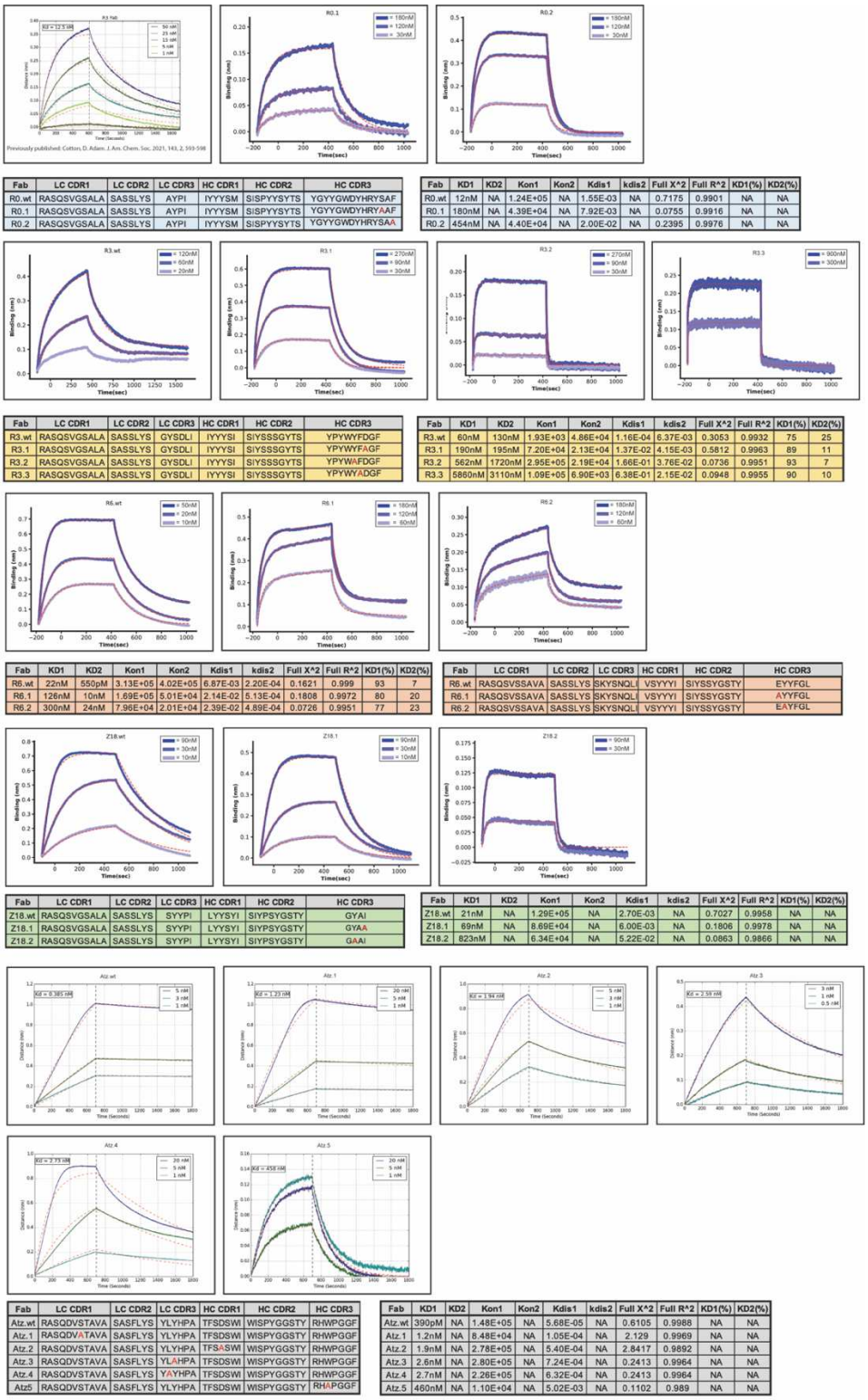
Figure 2.S2: SDS-PAGE of AbTACs used in this study. SDS-Polyacrylamide gel electrophoresis (PAGE) analysis of each purified AbTAC utilized in this study.





	Cells	Primary Ab	Secondary Ab	Median (APC)
	T24 R-WT	Atezolizumab Fab (50nM)	Isotype Control-647	480
	T24 R-WT	Atezolizumab Fab (50nM)	PD-L1 (D8T4X)-647	8120

Figure 2.S3: PD-L1 antibody used in flow-based degradation readout binds separate epitope than Atezolizumab. BLI (Top) and Flow cytometry (Bottom) based analysis indicating that the PD-L1 (D8T4X)-647 antibody used for flow cytometry-based measurement of PD-L1 levels in degradation assays binds a different epitope than Atezolizumab.



Fab	LC CDR1	LC CDR2	LC CDR3	HC CDR1	HC CDR2	HC CDR3
R0.wt	RASQSVGSALA	SASSLYS	AYPI	IYYYSM	SISPYYSYTS	YGYYGWDYHRYSAF
R0.1	RASQSVGSALA	SASSLYS	AYPI	IYYYSM	SISPYYSYTS	YGYYGWDYHRYAAF
R0.2	RASQSVGSALA	SASSLYS	AYPI	IYYYSM	SISPYYSYTS	YGYYGWDYHRYSAA

Fab	KD1	KD2	Kon1	Kon2	Kdis1	kdis2	Full X^2	Full R^2	KD1(%)	KD2(%)
R0.wt	12nM	NA	1.24E+05	NA	1.55E-03	NA	0.7175	0.9901	NA	NA
R0.1	180nM	NA	4.39E+04	NA	7.92E-03	NA	0.0755	0.9916	NA	NA
R0.2	454nM	NA	4.40E+04	NA	2.00E-02	NA	0.2395	0.9976	NA	NA

Fab	LC CDR1	LC CDR2	LC CDR3	HC CDR1	HC CDR2	HC CDR3
R3.wt	RASQSVGSALA	SASSLYS	GYSDLI	IYYYSI	SIYSSSGYTS	YPYWYFDGF
R3.1	RASQSVGSALA	SASSLYS	GYSDLI	IYYYSI	SIYSSSGYTS	YPYWYFAGF
R3.2	RASQSVGSALA	SASSLYS	GYSDLI	IYYYSI	SIYSSSGYTS	YPYWAFDGF
R3.3	RASQSVGSALA	SASSLYS	GYSDLI	IYYYSI	SIYSSSGYTS	YPYWYADGF

Fab	KD1	KD2	Kon1	Kon2	Kdis1	kdis2	Full X^2	Full R^2	KD1(%)	KD2(%)
R3.wt	60nM	130nM	1.93E+03	4.86E+04	1.16E-04	6.37E-03	0.3053	0.9932	75	25
R3.1	190nM	195nM	7.20E+04	2.13E+04	1.37E-02	4.15E-03	0.5812	0.9963	89	11
R3.2	562nM	1720nM	2.95E+05	2.19E+04	1.66E-01	3.76E-02	0.0736	0.9951	93	7
R3.3	5860nM	3110nM	1.09E+05	6.90E+03	6.38E-01	2.15E-02	0.0948	0.9955	90	10

Fab	KD1	KD2	Kon1	Kon2	Kdis1	kdis2	Full X^2	Full R^2	KD1(%)	KD2(%)
R6.wt	22nM	550pM	3.13E+05	4.02E+05	6.87E-03	2.20E-04	0.1621	0.999	93	7
R6.1	126nM	10nM	1.69E+05	5.01E+04	2.14E-02	5.13E-04	0.1808	0.9972	80	20
R6.2	300nM	24nM	7.96E+04	2.01E+04	2.39E-02	4.89E-04	0.0726	0.9951	77	23

Fab	LC CDR1	LC CDR2	LC CDR3	HC CDR1	HC CDR2	HC CDR3
R6.wt	RASQSVSSAVA	SASSLYS	SKYSNQLI	VSYYYI	SIYSSYGSTY	EYYFGL
R6.1	RASQSVSSAVA	SASSLYS	SKYSNQLI	VSYYYI	SIYSSYGSTY	AYYFGL
R6.2	RASQSVSSAVA	SASSLYS	SKYSNQLI	VSYYYI	SIYSSYGSTY	EAYFGL

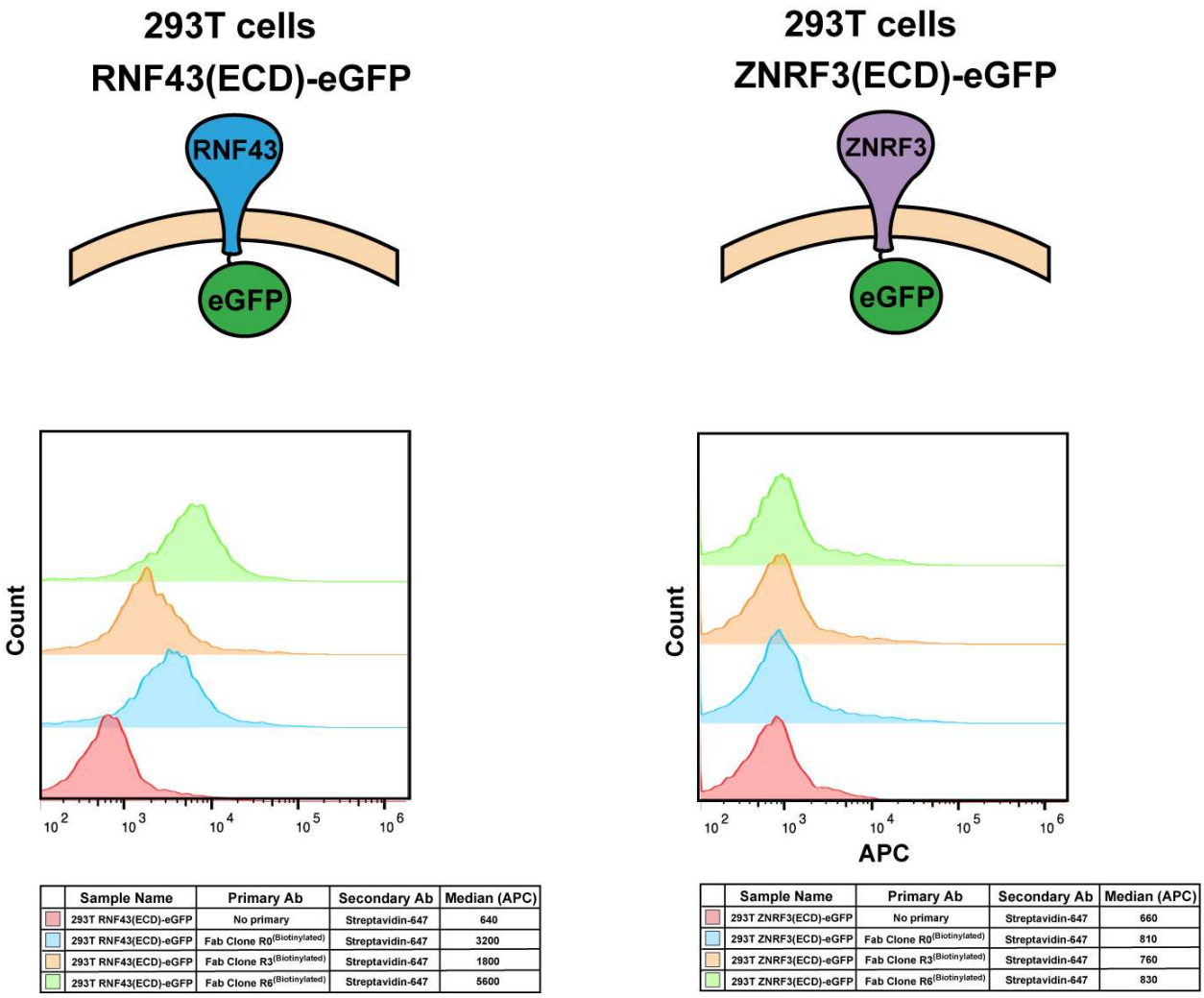
Fab	LC CDR1	LC CDR2	LC CDR3	HC CDR1	HC CDR2	HC CDR3
Z18.wt	RASQSVGSALA	SASSLYS	SYYPI	LYYSYI	SIYPSYGSTY	GYAI
Z18.1	RASQSVGSALA	SASSLYS	SYYPI	LYYSYI	SIYPSYGSTY	GYAA
Z18.2	RASQSVGSALA	SASSLYS	SYYPI	LYYSYI	SIYPSYGSTY	GAAI

Fab	KD1	KD2	Kon1	Kon2	Kdis1	kdis2	Full X^2	Full R^2	KD1(%)	KD2(%)
Z18.wt	21nM	NA	1.29E+05	NA	2.70E-03	NA	0.7027	0.9958	NA	NA
Z18.1	69nM	NA	8.69E+04	NA	6.00E-03	NA	0.1806	0.9978	NA	NA
Z18.2	823nM	NA	6.34E+04	NA	5.22E-02	NA	0.0863	0.9866	NA	NA

Fab	LC CDR1	LC CDR2	LC CDR3	HC CDR1	HC CDR2	HC CDR3
Atz.wt	RASQDVSTAVA	SASFLYS	YLYHPA	TFSDSWI	WISPYGGSTY	RHWPGGF
Atz.1	RASQDVATAVA	SASFLYS	YLYHPA	TFSDSWI	WISPYGGSTY	RHWPGGF
Atz.2	RASQDVSTAVA	SASFLYS	YLYHPA	TFSASWI	WISPYGGSTY	RHWPGGF
Atz.3	RASQDVSTAVA	SASFLYS	YLAHPA	TFSDSWI	WISPYGGSTY	RHWPGGF
Atz.4	RASQDVSTAVA	SASFLYS	YAYHPA	TFSDSWI	WISPYGGSTY	RHWPGGF
Atz5	RASQDVSTAVA	SASFLYS	YLYHPA	TFSDSWI	WISPYGGSTY	RHAPGGF

Fab	KD1	KD2	Kon1	Kon2	Kdis1	kdis2	Full X^2	Full R^2	KD1(%)	KD2(%)
Atz.wt	390pM	NA	1.48E+05	NA	5.68E-05	NA	0.6105	0.9988	NA	NA
Atz.1	1.2nM	NA	8.48E+04	NA	1.05E-04	NA	2.129	0.9969	NA	NA
Atz.2	1.9nM	NA	2.78E+05	NA	5.40E-04	NA	2.8417	0.9892	NA	NA
Atz.3	2.6nM	NA	2.80E+05	NA	7.24E-04	NA	0.2413	0.9964	NA	NA
Atz.4	2.7nM	NA	2.26E+05	NA	6.32E-04	NA	0.2413	0.9964	NA	NA
Atz.5	460nM	NA	1.10E+04	NA	5.02E-03	NA	0.1102	0.989	NA	NA

Figure 2.S4: Binding kinetics and CDR sequences of Fabs utilized in this study. Bio-layer interferometry kinetic measurements for each of the Fabs utilized in this study along with the respective CDR sequences.



	Sample Name	Primary Ab	Secondary Ab	Median (APC)
	293T RNF43(ECD)-eGFP	No primary	Streptavidin-647	640
	293T RNF43(ECD)-eGFP	Fab Clone R0 (Biotinylated)	Streptavidin-647	3200
	293T RNF43(ECD)-eGFP	Fab Clone R3 (Biotinylated)	Streptavidin-647	1800
	293T RNF43(ECD)-eGFP	Fab Clone R6 (Biotinylated)	Streptavidin-647	5600

	Sample Name	Primary Ab	Secondary Ab	Median (APC)
	293T ZNRF3(ECD)-eGFP	No primary	Streptavidin-647	660
	293T ZNRF3(ECD)-eGFP	Fab Clone R0 (Biotinylated)	Streptavidin-647	810
	293T ZNRF3(ECD)-eGFP	Fab Clone R3 (Biotinylated)	Streptavidin-647	760
	293T ZNRF3(ECD)-eGFP	Fab Clone R6 (Biotinylated)	Streptavidin-647	830

Figure 2.S5: RNF43 Fabs specifically bind RNF43 on cells. RNF43 Clones 0, 3, and 6 specifically bind RNF43 on cells. Fabs were incubated with HEK 293T cells overexpressing either RNF43(ECD)-eGFP (RNF43 in which the intracellular domain is replaced with eGFP) or ZNRF3(ECD)-eGFP (ZNRF3 in which the intracellular domain is replaced with eGFP).

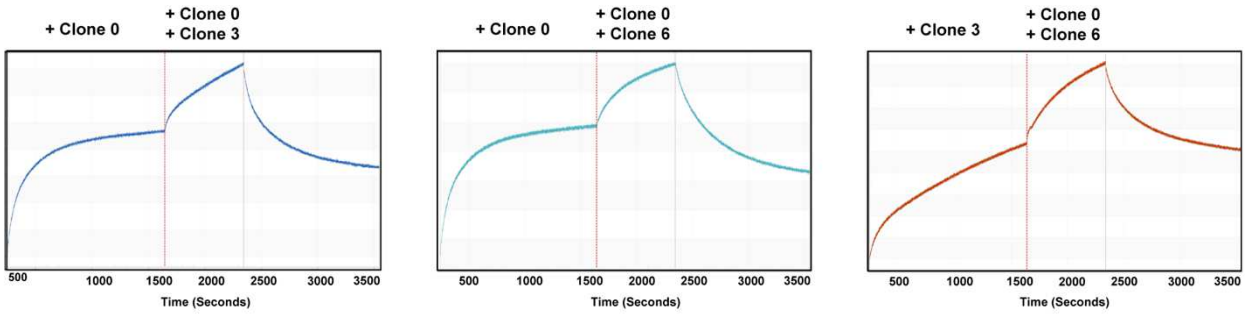


Figure 2.S6: RNF43 Fab clones 0, 3, and 6 bind unique epitopes on RNF43. Fabs bind unique epitopes based on BLI. For each graph, 100nM of the first construct was added followed by 100nM of the first construct and 100nM of the second construct.

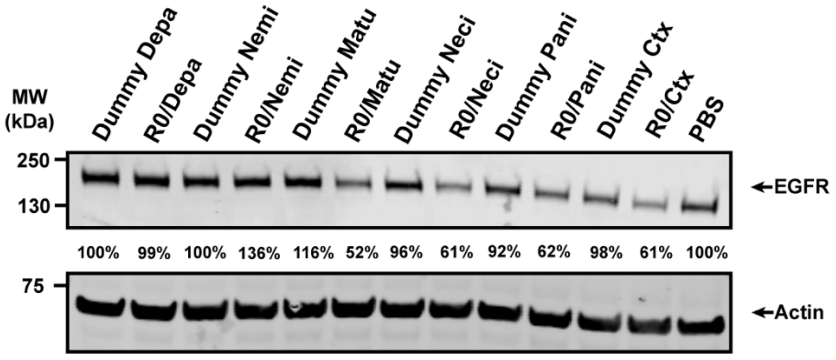


Figure 2.S7: Representative Western blot of EGFR degradation on T24 R(WT) cells. Western blot of degradation assays depicting levels of EGFR following 24hr incubation of 10nM of the indicated AbTAC on T24 R-WT cells. EGFR levels are relative to un-treated cells.

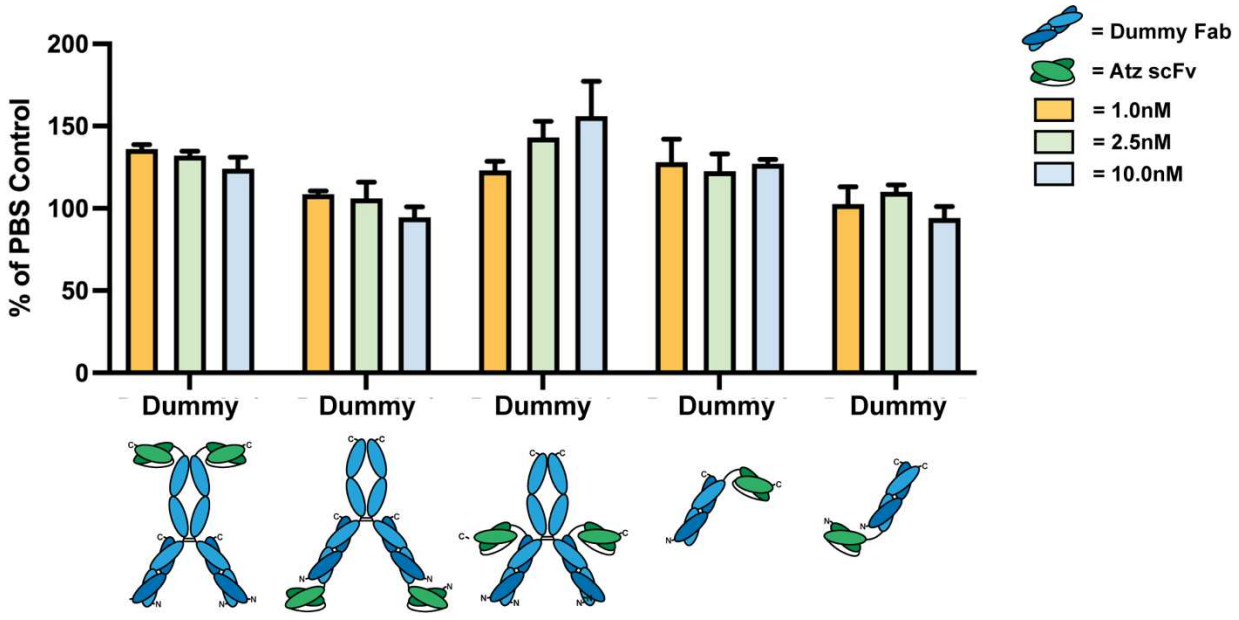
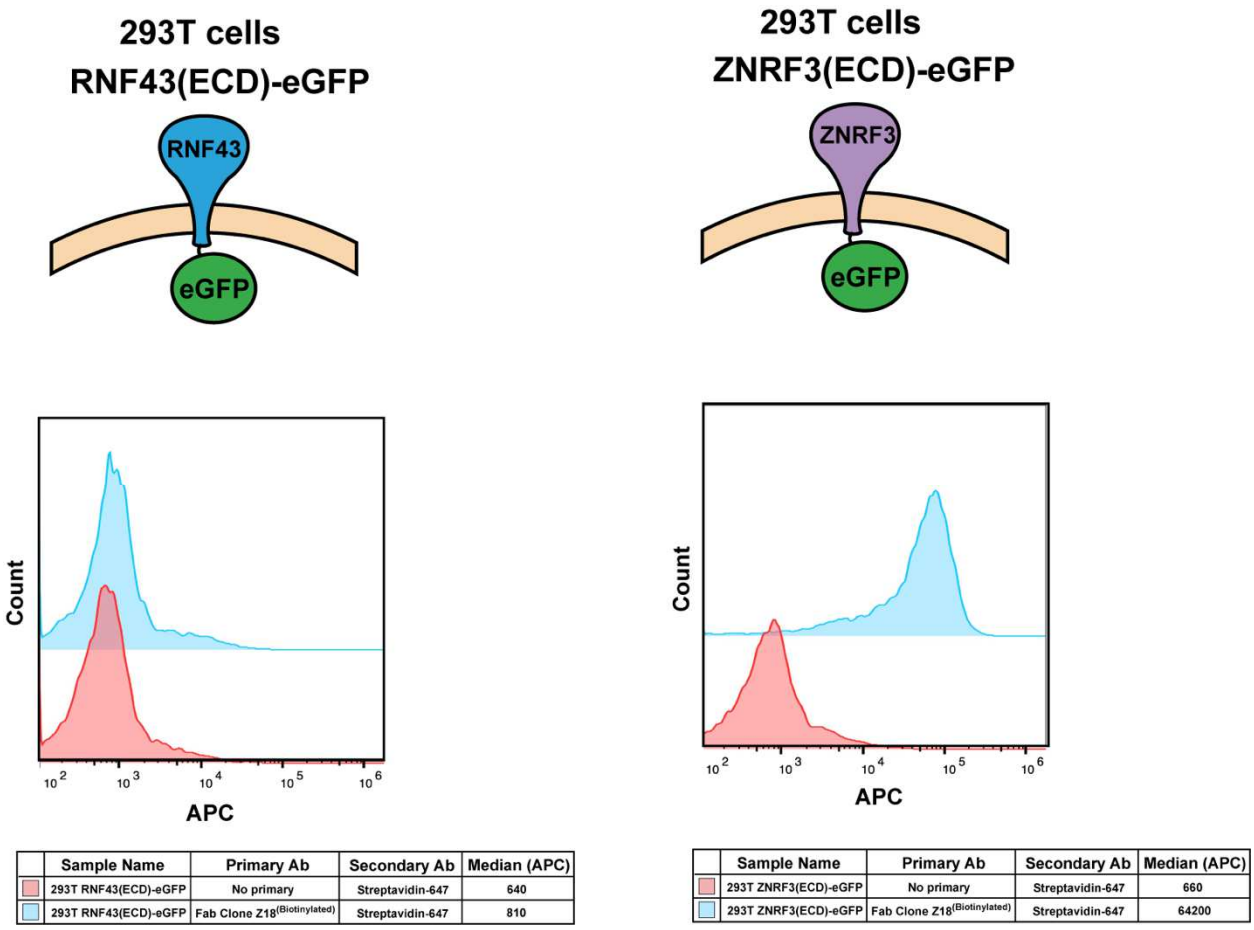


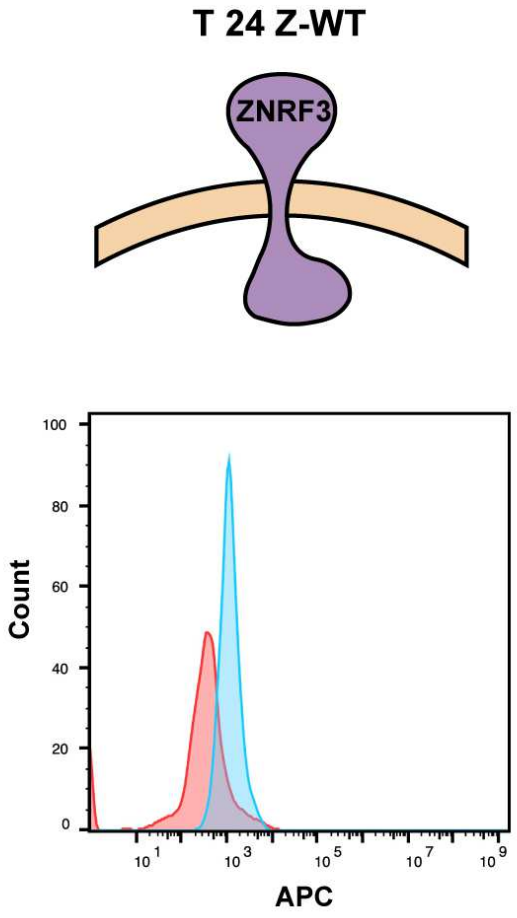
Figure 2.S8: Dummy AbTACs do not degrade PD-L1. Graphs of degradation assays depicting surface levels of PD-L1 as measured by flow cytometry following 24hr incubation of the indicated Dummy AbTACs at different concentrations on T24 R-WT cells. PD-L1 levels are relative to un-treated cells. Data are representative of at least three independent biological replicates. Dummy AbTACs contain an E3 ligase binding arm that has been replaced with a Covid-19 RBD binding arm[1].



	Sample Name	Primary Ab	Secondary Ab	Median (APC)
	293T RNF43(ECD)-eGFP	No primary	Streptavidin-647	640
	293T RNF43(ECD)-eGFP	Fab Clone Z18$^{(Biotinylated)}$	Streptavidin-647	810

	Sample Name	Primary Ab	Secondary Ab	Median (APC)
	293T ZNRF3(ECD)-eGFP	No primary	Streptavidin-647	660
	293T ZNRF3(ECD)-eGFP	Fab Clone Z18$^{(Biotinylated)}$	Streptavidin-647	64200

Figure 2.S9: Fab Z18 specifically binds ZNRF3 on cells. ZNRF3 Clone 18 specifically binds ZNRF3 on cells. Fab was incubated with HEK 293T cells overexpressing either RNF43(ECD)-eGFP (RNF43 in which the intracellular domain is replaced with eGFP) or ZNRF3(ECD)-eGFP (ZNRF3 in which the intracellular domain is replaced with eGFP).



	Sample Name	Primary Ab	Secondary Ab	Median (APC)
	T24 WT cells	Fab Clone Z18$^{(Biotinylated)}$	Streptavidin-647	370
	T24 Z-WT cells	Fab Clone Z18$^{(Biotinylated)}$	Streptavidin-647	1174

Figure 2.S10: T24 ZNRF3 overexpression cells. Flow cytometry measurement of the cell surface levels of ZNRF3 on either T24 WT cells or T24 cells overexpressing WT ZNRF3 (T24 Z-WT).

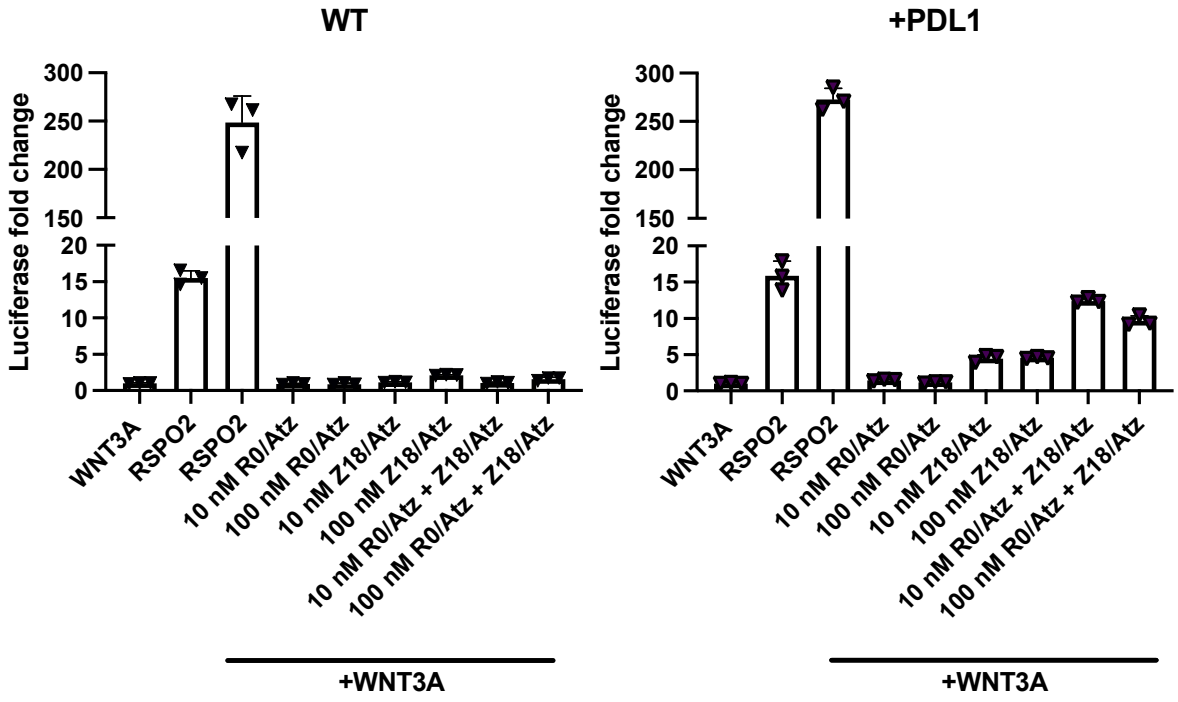


Figure 2.S11: R0/Atz and Z18/Atz AbTACs do not affect canonical WNT signaling on HEK293 STF cells with or without overexpression of PDL1

2.7 References

(1) Bond, M. J.; Crews, C. M. Proteolysis Targeting Chimeras (PROTACs) Come of Age: Entering the Third Decade of Targeted Protein Degradation. *RSC Chem. Biol.* **2021**, *2* (3), 725–742. https://doi.org/10.1039/d1cb00011j.

(2) Pettersson, M.; Crews, C. M. PROteolysis TArgeting Chimeras (PROTACs) - Past, Present and Future. *Drug Discov. Today Technol.* **2019**, *31*, 15–27. https://doi.org/10.1016/j.ddtec.2019.01.002.

(3) Mullard, A. First Targeted Protein Degrader Hits the Clinic. *Nat. Rev. Drug Discov.* **2019**. https://doi.org/10.1038/d41573-019-00043-6.

(4) Cotton, A. D.; Nguyen, D. P.; Gramespacher, J. A.; Seiple, I. B.; Wells, J. A. Development of Antibody-Based PROTACs for the Degradation of the Cell-Surface Immune Checkpoint Protein PD-L1. *J. Am. Chem. Soc.* **2021**, *143* (2), 593–598. https://doi.org/10.1021/jacs.0c10008.

(5) Banik, S. M.; Pedram, K.; Wisnovsky, S.; Ahn, G.; Riley, N. M.; Bertozzi, C. R. Lysosome-Targeting Chimaeras for Degradation of Extracellular Proteins. *Nature* **2020**, *584* (7820), 291–297. https://doi.org/10.1038/s41586-020-2545-9.

(6) Ahn, G.; Banik, S. M.; Miller, C. L.; Riley, N. M.; Cochran, J. R.; Bertozzi, C. R. LYTACs That Engage the Asialoglycoprotein Receptor for Targeted Protein Degradation. *Nat. Chem. Biol.* **2021**, *17* (9), 937–946. https://doi.org/10.1038/s41589-021-00770-1.

(7) Chames, P.; Van Regenmortel, M.; Weiss, E.; Baty, D. Therapeutic Antibodies: Successes, Limitations and Hopes for the Future. *Br. J. Pharmacol.* **2009**, *157* (2), 220–233. https://doi.org/10.1111/j.1476-5381.2009.00190.x.

(8) Papatheodorou, I.; Moreno, P.; Manning, J.; Fuentes, A. M.-P.; George, N.; Fexova, S.; Fonseca, N. A.; Füllgrabe, A.; Green, M.; Huang, N.; Huerta, L.; Iqbal, H.; Jianu, M.; Mohammed,

S.; Zhao, L.; Jarnuczak, A. F.; Jupp, S.; Marioni, J.; Meyer, K.; Petryszak, R.; Prada Medina, C. A.; Talavera-López, C.; Teichmann, S.; Vizcaino, J. A.; Brazma, A. Expression Atlas Update: From Tissues to Single Cells. *Nucleic Acids Res.* **2020**, *48* (D1), D77–D83. https://doi.org/10.1093/nar/gkz947.

(9) Ridgway, J. B.; Presta, L. G.; Carter, P. "Knobs-into-Holes" Engineering of Antibody CH3 Domains for Heavy Chain Heterodimerization. *Protein Eng.* **1996**, *9* (7), 617–621. https://doi.org/10.1093/protein/9.7.617.

(10) Zhang, F.; Qi, X.; Wang, X.; Wei, D.; Wu, J.; Feng, L.; Cai, H.; Wang, Y.; Zeng, N.; Xu, T.; Zhou, A.; Zheng, Y. Structural Basis of the Therapeutic Anti-PD-L1 Antibody Atezolizumab. *Oncotarget* **2017**, *8* (52), 90215–90224. https://doi.org/10.18632/oncotarget.21652.

(11) Cai, W.-Q.; Zeng, L.-S.; Wang, L.-F.; Wang, Y.-Y.; Cheng, J.-T.; Zhang, Y.; Han, Z.-W.; Zhou, Y.; Huang, S.-L.; Wang, X.-W.; Peng, X.-C.; Xiang, Y.; Ma, Z.; Cui, S.-Z.; Xin, H.-W. The Latest Battles Between EGFR Monoclonal Antibodies and Resistant Tumor Cells. *Front. Oncol.* **2020**, *10*, 1249. https://doi.org/10.3389/fonc.2020.01249.

(12) Garrett, T. P. J.; Burgess, A. W.; Gan, H. K.; Luwor, R. B.; Cartwright, G.; Walker, F.; Orchard, S. G.; Clayton, A. H. A.; Nice, E. C.; Rothacker, J.; Catimel, B.; Cavenee, W. K.; Old, L. J.; Stockert, E.; Ritter, G.; Adams, T. E.; Hoyne, P. A.; Wittrup, D.; Chao, G.; Cochran, J. R.; Luo, C.; Lou, M.; Huyton, T.; Xu, Y.; Fairlie, W. D.; Yao, S.; Scott, A. M.; Johns, T. G. Antibodies Specifically Targeting a Locally Misfolded Region of Tumor Associated EGFR. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (13), 5082–5087. https://doi.org/10.1073/pnas.0811559106.

(13) Talavera, A.; Friemann, R.; Gómez-Puerta, S.; Martinez-Fleites, C.; Garrido, G.; Rabasa, A.; López-Requena, A.; Pupo, A.; Johansen, R. F.; Sánchez, O.; Krengel, U.; Moreno, E.

Nimotuzumab, an Antitumor Antibody That Targets the Epidermal Growth Factor Receptor, Blocks Ligand Binding While Permitting the Active Receptor Conformation. *Cancer Res.* **2009**, *69* (14), 5851–5859. https://doi.org/10.1158/0008-5472.CAN-08-4518.

(14) Schmiedel, J.; Blaukat, A.; Li, S.; Knöchel, T.; Ferguson, K. M. Matuzumab Binding to EGFR Prevents the Conformational Rearrangement Required for Dimerization. *Cancer Cell* **2008**, *13* (4), 365–373. https://doi.org/10.1016/j.ccr.2008.02.019.

(15) Bagchi, A.; Haidar, J. N.; Eastman, S. W.; Vieth, M.; Topper, M.; Iacolina, M. D.; Walker, J. M.; Forest, A.; Shen, Y.; Novosiadly, R. D.; Ferguson, K. M. Molecular Basis for Necitumumab Inhibition of EGFR Variants Associated with Acquired Cetuximab Resistance. *Mol. Cancer Ther.* **2018**, *17* (2), 521–531. https://doi.org/10.1158/1535-7163.MCT-17-0575.

(16) Sickmier, E. A.; Kurzeja, R. J. M.; Michelsen, K.; Vazir, M.; Yang, E.; Tasker, A. S. The Panitumumab EGFR Complex Reveals a Binding Mechanism That Overcomes Cetuximab Induced Resistance. *PloS One* **2016**, *11* (9), e0163366. https://doi.org/10.1371/journal.pone.0163366.

3. Biotin as a reactive handle to selectively label proteins and DNA with small molecules

3.1 Abstract

Biotin is a common functional handle for bioconjugation to proteins and DNA, but its uses are limited to protein-containing conjugation partners such as streptavidin and derivatives thereof. Recently, oxaziridine reagents were developed that selectively conjugate the thioether of methionines on the surface of proteins, a method termed redox-activated chemical tagging (ReACT). These reagents generate sulfimide linkages that range in stability depending on solvent accessibility and substitutions on the oxaziridine. Here we show that oxaziridine reagents react rapidly with the thioether in biotin to produce sulfimide products that are stable for more than 10 days at 37 ºC. This method, which we call biotin redox-activated chemical tagging (BioReACT) expands the utility of biotin labeling and enables predictable and stable chemical conjugation to biomolecules without the need to screen for a suitable methionine conjugation site. We demonstrate the versatility of this approach by producing a fluorescently labeled antibody, an antibody–drug conjugate, and a small molecule-conjugated oligonucleotide. We anticipate that BioReACT will be useful to researchers to rapidly introduce biorthogonal handles into biomolecules using biotin, a functional group that is widespread and straightforward to install.

3.2 Introduction

Chemical modification of biomolecules has widespread applications ranging from antibody-drug conjugates to in vivo localization assays.[1,2,3] Researchers have devised numerous methods to conjugate small molecules to biomolecules such as amino acid modification or introduction of a biorthogonal reactive handle.[4] Due to their prevalence and reactivity, lysine and cysteine are the most widely used amino acids. While methionine has been the focus of other labelling strategies[5], its relatively unreactive side chain makes bioconjugation challenging. Recently, the Chang and Toste groups described oxaziridine chemistry for selective methionine modification on proteins called redox-activated chemical tagging (ReACT).[6] Methionine is rarely on the surface of proteins, enabling accurate and site-specific labelling[7,8]. Although the ReACT method is selective, the conjugation product is prone to hydrolysis, especially when the label is installed at a C–terminal methionine, limiting its potential in applications where stability is critical.[9] To address this limitation, more stable oxaziridine reagents were synthesized and our group identified sites on the general antibody scaffold that both label well and are stable for several days.[9,10] The most stable sites were not entirely predictable, and were often in partially buried regions of the antibody. A method to generate oxaziridine–biomolecule linkages at accessible sites with predictable stability would be advantageous.

Biotin is ubiquitous throughout biological research, and proteins and DNA can usually be biotinylated with no effect on function. It is regularly used in conjunction with avidin-based reagents, enabling applications such as immobilization or fluorescent tagging.[11,12,13,14,15] Avidin is a tetrameric protein that can introduce unwanted avidity effects and additional protein complications to numerous biophysical assays.[15] Furthermore, unless it is pre-modified, avidin cannot be used to introduce small molecule conjugates. Oxaziridine reagents such as those used in

ReACT function by nucleophilic attack of methionine into the N-O σ^* followed by elimination of an aldehyde resulting in a sulfimidated methionine (Figure 1a). We hypothesized that oxaziridine reagents might also react with the tetrahydrothiophene in biotin,[16] enabling rapid small molecule bioconjugation to proteins and other biomolecules. In addition, we hypothesized that the increased steric hinderance of the biotin bicycle would alter the reactivity and potentially decrease the hydrolysis rate of the resulting sulfimide conjugate.

Here we report the reactivity of biotin and azide-labelled oxaziridine reagents. This method, which we call biotin redox-activated chemical tagging (BioReACT), achieves efficient labeling in aqueous conditions at ambient temperature and provides stable conjugation products. We functionally validate the method by generating an antibody-drug conjugate and numerous flow-cytometry reagents. Finally, we conjugate a functional click handle to a biotinylated oligonucleotide. These studies show that the biotin–oxaziridine reaction is a powerful approach for the efficient synthesis of stable protein and DNA bioconjugates.

3.3 Results

We first evaluated the reactivity of unconjugated biotin with oxaziridine reagents by exposing biotin methyl ester **1** to oxaziridine **2** (prepared from known protocol)[9] in 1:1 CD_3OD/D_2O (**Fig. 3.1b**). The two protons alpha to the sulfur in biotin appear as distinct peaks between 2.7-3.0 ppm in ^1H NMR, allowing us to assess the conversion and reaction kinetics. Remarkably, at 20 mM biotin methyl ester with 1.5 equivalents of oxaziridine **2**, the reaction proceeds to >80% conversion within 3 minutes, and 100% conversion by 20 minutes (**Fig. 3.1c, 3.S1**). We observed a 17:3 ratio of N-transfer product (NTP, sulfimide) to O-transfer product (OTP, sulfoxide), which proved challenging to separate, resulting in an 85% NMR yield and 52% isolated yield of desired sulfimide

3. It should be noted that while we observed OTP in this reaction, it was never seen for any of the biological labeling experiments described below (perhaps due to solvent or concentration effects).

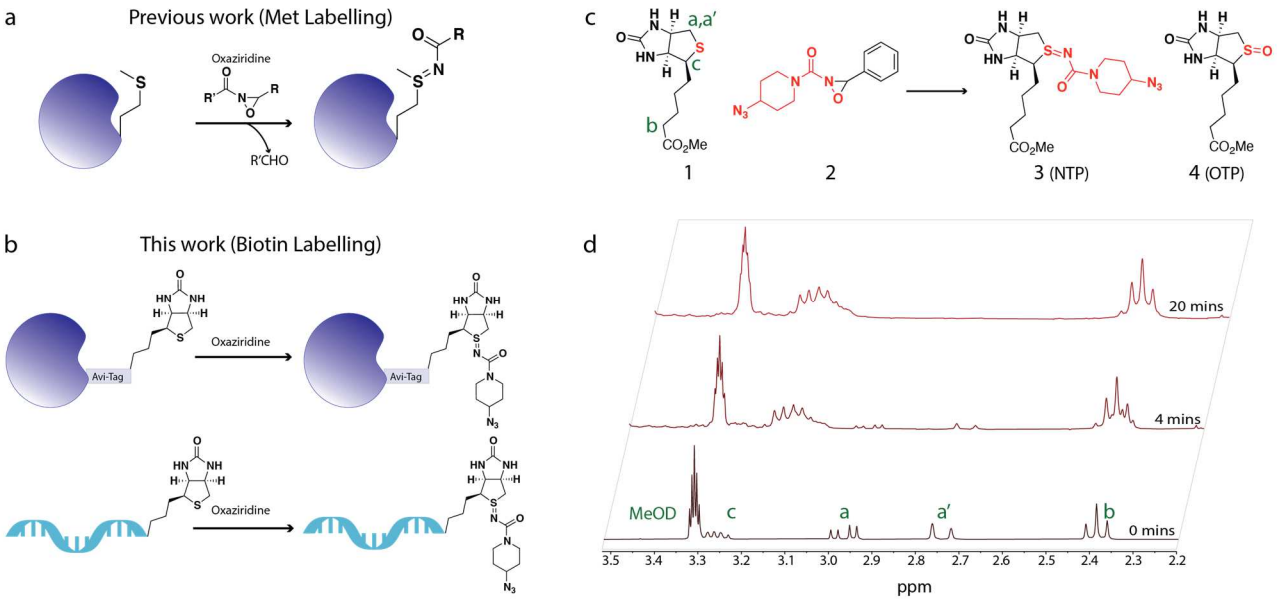


Figure 3.1: Repurposing oxaziridines to label biotin. a) ReACT strategy for methionine bioconjugation. Oxaziridine compounds react with the thioether in methionine to form a sulfimide conjugation product. b) Hypothesized reactivity of oxaziridines with biotin to form a sulfimidated bicycle. BioReACT enables labeling of biomolecules such as DNA that contain no methionines. c) The reaction of biotin methyl ester with 1.5 equivalents of oxaziridine 2 in 1:1 CD_3OD/D_2O to form the resulting NTP sulfimide (85% NMR yield, 52% isolated yield). d) 1H-NMR series of the reaction in (c). Conversion was monitored by detecting the chemical shift of the protons alpha to the sulfur at ~2.75 and ~2.97 ppm. 20 mM biotin methyl ester with 1.5 equivalents of oxaziridine 2 achieved complete conversion within 20 minutes.

We next evaluated the reactivity of biotin attached to an antigen-binding fragment (Fab) derived from the anti-Her2 antibody trastuzumab as a model protein. The Fab scaffold has been thoroughly characterized with the methionine strategy (ReACT) to identify reactive and stable sites, making it a privileged candidate for comparison of the two labelling strategies.[8] We chose trastuzumab as it expresses well and has high thermal stability. To generate a biotinylated protein, we introduced an Avi-tag to the C-terminus of the Fab sequence. The Avi-Tag allows for enzymatic, site-specific biotinylation using a BirA enzyme[17]. This biotinylation event can be performed during or post-expression. We incubated 50 µM biotinylated trastuzumab Fab in PBS at 25 ºC with several

concentrations of oxaziridine **2** and used whole protein mass spectrometry to measure labelling (**Fig. 3.2a**). Gratifyingly, we saw labelling at equimolar concentrations of Fab and oxaziridine, and 10-fold excess oxaziridine **2** leads to the complete conversion of Fab to labeled product containing either 1 or 2 sulfimides in a 3.2:1 ratio. (**Fig. 3.S2**). 20% of WT unbiotinylated trastuzumab Fab reacted with 5 equivalents of oxaziridine reagent (**Fig. 3.2b red curve**), although at high molar excess **2**, we saw increased labelling of the native protein (**Fig. 3.S2)**. We hypothesized that the three native methionines were causing the background labelling of the native Trastuzumab Fab. [18] To test this, we mutated the native methionines to leucines and repeated the labelling experiment. Surprisingly, we observed the same level of labelling as the methionine-containing unbiotinylated protein, indicating a non-methionine residue is being labelled. To determine the site of labeling, we performed peptide mapping using a Trypsin/Lys-C digest. This analysis identified the heavy-chain N-terminus and a histidine in the Avi-Tag as the sites of modification by the Oxaziridine reagent (**Fig. 3.S5**). This result is in agreement with a recent report from Zanon and coworkers that identified similar background reactivity with N-termini and histidines.[19] This minor side product was not noted in previous ReACT studies, and may aid in the understanding of additional labeling in similar systems. This peptide mapping experiment also confirmed that the majority of labelling events occurred at the predicted biotinylated lysine in the Avi-Tag. To minimize double labeling, we use 5 equivalents of oxaziridine throughout the rest of this study.

To gauge the working protein concentration range of labelling we incubated trastuzumab Fab at several concentrations with 5-fold molar excess of oxaziridine probe (**Fig. 3.S3**). We chose a range of 10-130 μM as a representative working range for proteins. Interestingly we saw comparable labelling at 10 μM to that of 130 μM, indicating this method can be used to label proteins at a range of concentrations. To assess labelling efficiency, we incubated 50 μM biotinylated

trastuzumab Fab with 250 μM oxaziridine reagent. The biotinylated Fab reacted rapidly, achieving a maximum labelling efficiency of 80% (65% single, 15 % double labelled) after 1 hour (**Fig. 3.2c**). We were interested in the relative reaction rates of oxaziridine with biotin vs. methionine. Due to the positioning of the C-terminal Avi-tag, a C-terminal methionine mutant was chosen as a direct comparison. Each Fab at 50 μM was incubated with 5 molar excess oxaziridine reagent and relative rate constants calculated (**Fig. 3.S6)**. Interestingly oxaziridine reacts with the C-terminal methionine 3.8x faster than with biotin (see Supporting Information for details).

Modifications to biotin (e.g., oxidation) can drastically reduce the affinity of the biotin-streptavidin interaction.[20] To evaluate whether the Oxaziridine-biotin sulfimide product can still interact with streptavidin, we performed a flow-cytometry experiment with biotinylated anti-Her2 Fab and oxaziridine-biotin-anti-Her2 Fab and cells containing cell-surface Her2. After incubating SKBR3 cells with either Fab, we used a Streptavidin-647 conjugate to report on the interaction. There was a marked decrease (81%) in fluorescent signal for cells treated with the oxaziridine-labelled Fab, indicating the oxaziridine modification can block the biotin-streptavidin interaction (**Fig. 3.S8**). To test if neutravidin will prevent the biotin-oxaziridine reaction from occurring we pre-incubated biotinylated Fab with equimolar amounts of neutravidin followed by treatment with oxaziridine and used whole protein mass spectrometry to assess labeling. Only a small amount of background labeling occurred, equivalent to the amount of labeling that occurs in the absence of a biotin handle, indicating that binding to avidin precludes biotin modification with oxaziridine **2** (**Fig. 3.S9**).

To test the stability of the resulting biotin sulfimide to hydrolysis, we incubated labelled Fab in PBS at 37 ºC for several days (**Fig. 3.2d**). Similar to above, a C-terminal methionine was used as a direct comparison. Using the ReACT strategy, C-terminal methionine conjugates were 50% hydrolyzed within 3 days, whereas conjugates using the BioReACT approach were 100% stable 10-days at 37 ºC. The stability of the biotin conjugate is comparable to the most stable mutants seen from an extensive methionine scan of the same Fab scaffold.[9] These results demonstrated that BioReACT is an efficient method to rapidly and stably label proteins that contain the common Avi-Tag.

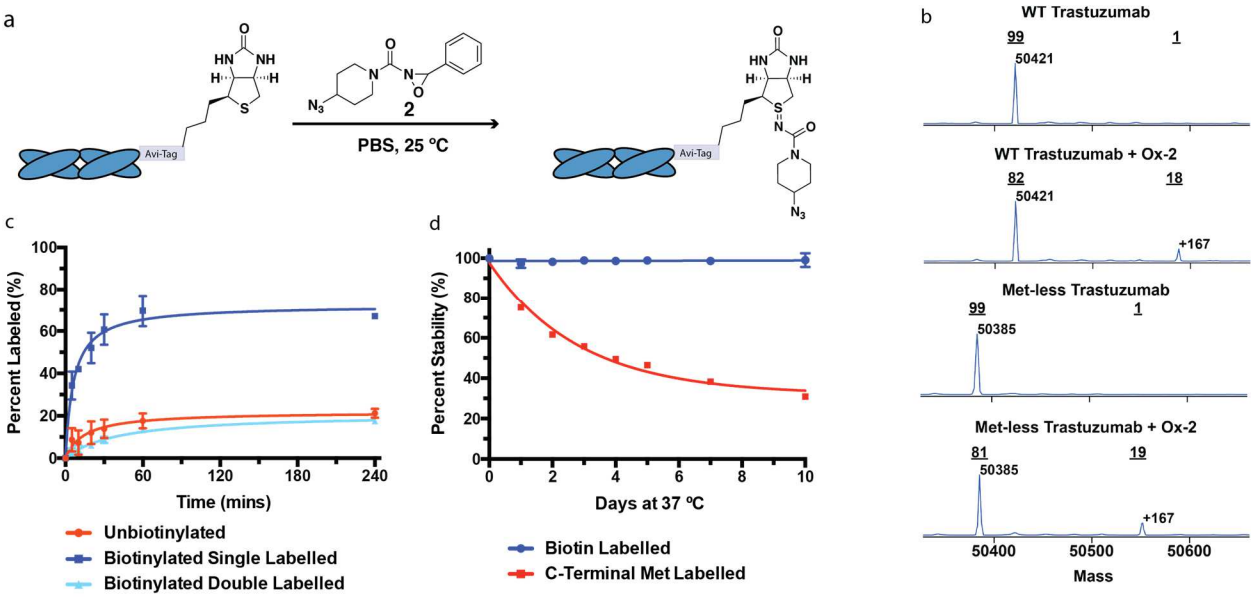


Figure 3.2: Biotin is a reactive handle for stable bioconjugation to proteins. a) Trastuzumab Fab was allowed to react with oxaziridine 2 in PBS for 1 hour at 25 ºC. b) The second labelling event is not caused by a native buried methionine. WT trastuzumab or methionine-less trastuzumab (in which all native methionines have been removed) was allowed to react with 5 equiv 2; +167 indicates an oxaziridine modification. c) Whole protein mass spectrometry was used to quantify labelling efficiency over time with 5 equivalents oxaziridine 2. Single labelling represents a singly biotinylated Fab with one oxaziridine adduct, while double labelling represents two oxaziridine adducts on a singly biotinylated Fab d) The singly labeled protein is 100% stable after incubation in PBS at 37ºC for 10 days, whereas a C-terminal methionine oxaziridine conjugate is rapidly hydrolyzed.

To evaluate practical applications of BioReACT we generated flow cytometry and antibody-drug conjugate reagents. We first confirmed that Fabs labelled with oxaziridine reagents maintained the ability to bind to the corresponding antigens. The affinity of labelled or unlabelled anti-PD-L1 fab for a PD-L1 Fc fusion was measured using biolayer interferometry (BLI); for both Fabs, the K_D was unperturbed (**Fig. 3.S7)**. Flow cytometry is a method often used to evaluate protein levels on the surface of cells[21]. Most protocols involve multiple incubation steps with either a primary antibody or a fluorophore-conjugated secondary antibody. We saw an opportunity to reduce the number of required incubations and wash steps by labelling biotinylated Fabs directly with a fluorescent dye in a controlled and site-specific manner. We first labelled biotinylated Fabs with oxaziridine **2**, and the resulting azide-containing conjugate was allowed to react with a DBCO-Cy5.5 dye in a strain-promoted azide–alkyne cycloaddition (**Fig. 3.S12)**.[22,23] We labelled Fabs that target three different cell surface proteins, Her2, PD-L1, and CDCP1.[18,24,25] An anti-GFP Fab was fluorescently labelled as an isotype control. Each labelled Fab was able to detect its corresponding antigen on known cell lines (SKBR3 cell line for Her2 and MDA-MB-231 cells for PD-L1 and CDCP1, **Fig. 3.3a, 3.S10, 3.S11**).

We next applied BioReACT to the synthesis of an antibody drug conjugate (ADC). Common methods to generate ADCs can result in a wide range of drug-to-antibody ratios (DARs).[26] We anticipated BioReACT would give a well-defined DAR (assuming minimal reactivity with the native scaffold) due to the predefined incorporation of a single biotin handle. We allowed azide-labelled trastuzumab Fab to react with DBCO-conjugated monomethyl auristatin F (MMAF), a cell-impermeable analog of MMAE (**Fig. 3.S12)**.[27] The resulting ADC selectively killed SKBR3 cells over 72 h more effectively than the warhead (MMAF) or the Fab alone (**Fig. 3.3b**). These

results demonstrate that BioReACT is a useful tool for common applications that require protein–small molecule conjugation.

Halo, SNAP, and CLIP tags are common methods used to chemically modify proteins.[28, 29] While these methods are utilized, Avi-tagged proteins are considerably more prevalent. Addgene, a commonly used plasmid repository contains over 5,000 plasmids containing an Avi-tag, with the number of Halo/SNAP/CLIP-tagged proteins totaling 264/247/108 respectively. Secondly, the Halo-tag is 33 kDa, and SNAP/CLIP-tags are ~20 kDa, we hypothesized that the short 15 amino acid Avi-tag would have comparatively favorable properties as compared to the other approaches. To test this, we took two clinically validated antibodies trastuzumab and atezolizumab and generated C-terminal heavy chain fusions of each tag. Fabs are commonly expressed in bacteria followed by a 60ºC lysis step.[25] To not bias the results away from a less thermally stable construct we used either a heat or sonication step to lyse the cells. Following protein-A purification no desired Halo/SNAP/CLIP tagged protein was observed, only the Avi-tagged protein expressed (**Fig. 3.S13**). We performed whole protein mass spectrometry to confirm the lack of expected products (**Fig. 3.S14**). This result in union with the 20-fold higher prevalence of Avi-tagged proteins demonstrates the potential utility and use for a method to chemically modify proteins containing an Avi-tag.

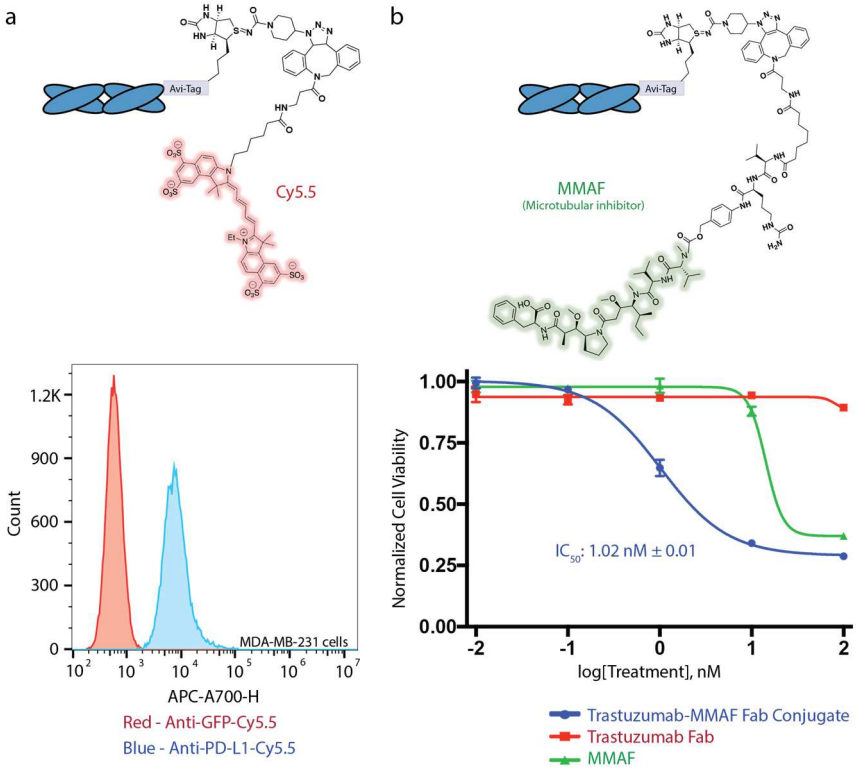


Figure 3.3. Biotinylated Fabs conjugated using BioReACT can be used to detect protein levels on cells and to form antibody drug conjugates. a) (Top) DBCO-Cy5.5 was allowed to react with azide-labelled Anti-PD-L1 Fab to produce a fluorescently labelled Fab. (Bottom) Cy5.5 conjugated Fab detected PD-L1 by flow cytometry on the surface of MDA-MB-231 cells. b) (Top) DBCO-Val-Cit-PABA-MMAF was reacted with azide labelled trastuzumab Fab to form an antibody drug conjugate (ADC). (Bottom) The resulting ADC was incubated with SKBR3 cells for 3 days after which a CellTiter-Glo® assay was used to confirm dose-dependent killing.

Due to the methionine dependence, to date the ReACT method has only been applied to labelling proteins. Biotinylation is very common in biological research, and we sought to exploit this to label other biomolecules. Oligonucleotide based therapeutics have gained traction as a way to modulate protein expression levels, and strategies for delivering these biomolecules often involve attachment to delivery vehicles such as carbohydrate polymers or peptides.[3] We purchased a single stranded DNA oligonucleotide modified with a biotin tag at the 5' end and allowed it to react with a 5-fold molar excess of oxaziridine 2 in water at 25 ºC for one hour (**Fig. 3.4**). Oxaziridine 2 was

added as a methanol solution to prevent subsequent signal suppression caused by either DMSO or DMF. MALDI-TOF Mass spectrometry revealed highly efficient introduction of the azide handle onto DNA. In a parallel experiment, we showed that non-biotinylated DNA does not label. Since biotin-conjugated oligonucleotides are cheap, stable, and readily available for purchase through commercial suppliers, BioReACT provides a rapid and practical method to generate stable small molecule–oligonucleotide conjugates.

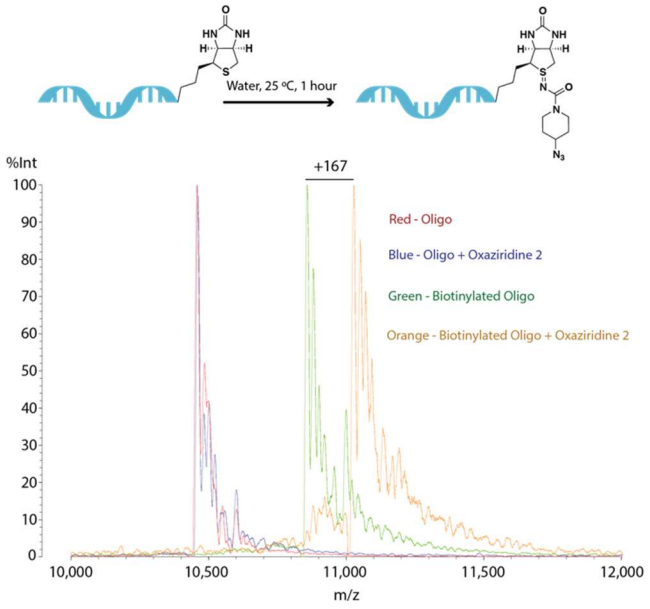


Figure 3.4: Expanding the applications of ReACT to label biotinylated DNA. (Top) Biotinylated DNA was exposed to 5 molar equivalents of oxaziridine 2 for 1 hour at 25 ºC in water to form the resulting sulfimide conjugate. (Bottom) MALDI-TOF Mass Spectrometry indicates efficient conjugation to biotinylated DNA (orange trace) and no labeling in the absence of biotin (blue trace).

3.4 Discussion

The ReACT method provided a way to selectively label methionine in proteins. Since methionine is rarely found on the surface of proteins, ReACT relies on the introduction of the amino acid by site-directed mutagenesis. Exhaustive screening is often required to find sites that readily react

with oxaziridine reagents and provide hydrolytically stable conjugates, two properties that are critical for many bioconjugation applications. We have expanded the utility of these reagents by demonstrating their reactivity with biotin, which is commonly incorporated on surfaces, antibodies, proteins, oligonucleotides, and polymers in modern biological research. Additionally, biotin can be readily incorporated into proteins by a genetically encoded Avi-tag or by chemical conjugation. BioReACT works rapidly and efficiently with biotinylated proteins and DNA without the need to screen for a suitable site, and the resulting sulfimide is extremely stable to hydrolysis. While biotinylation is typically considered a synthetic endpoint, BioReACT enables rapid modification and customization through chemical conjugation.

3.5 Materials and Methods

Fab Expression

Fabs were expressed and purified by an optimized autoinduction protocol previously described.[25] In brief, C43 (DE3) + BirA Pro + *E. coli*-containing expression plasmids were grown in TB autoinduction media containing biotin at 37ºC for 6 h, then cooled to 30ºC for 18 h. Fabs were purified by Protein A affinity chromatography. Purity was assessed by SDS/PAGE and intact protein mass spectrometry.

Fab Biotinylation

Where 100% biotinylation was not required Fabs were expressed as above with biotin at 20 μM final concentration present in the TB autoinduction media. For quantification experiments where 100% biotinylation was required in-vitro biotinylation was used. The Avidity reaction conditions were used with 1:1 BiomixA (0.5M bicine buffer solution, pH 8.3), BiomixB (100 mM ATP, 100 mM $Mg(OAc)_2$, 500 μM d-biotin), 5 μg BirA enzyme, and 10% reaction volume of Avi-tagged Fab in TBS at 50 μM. The reaction was incubated at RT for 1 hour, and whole protein mass spectrometry was used to assess biotinylation levels.

Synthetic materials

Oxaziridine 2 was synthesized using previously described protocol[9]. Biotin methyl ester was purchased from AK Scientific and was used as received.

Biotinylated Fab reactivity

Fabs at either 50 μM or described concentration in PBS were incubated with oxaziridine reagents at room temperature for either 1 hour or for described length of time. Samples were diluted into water to 250 nM final concentration before injection onto mass spectrometry. Conjugation levels

were assessed using intact protein mass spectrometry using a Xevo G2-XS Mass Spectrometer (Waters). Spectra were integrated to calculate labeling efficiency.

Oxaziridine Stability assay

Biotinylated Fab or a C-terminal Met mutant Fab was labeled with oxaziridine reagent and buffer exchanged into PBS to a final concentration of 10 μM. The resulting labeled Fab was incubated at 37 ºC and remaining conjugation levels were assessed using intact protein mass spectrometry using a Xevo G2-XS Mass Spectrometer (Waters).

Conjugation of biotinylated Fabs with oxaziridine and DBCO reagents

Fabs were incubated at 50 μm with 10 molar equivalents of oxaziridine 2 for 1 hour at room temperature in PBS. The reaction was buffer exchanged into PBS using a 0.5 mL 30 kDa spin concentrator (Thermo Fisher Scientific), to remove excess oxaziridine reagent. Then 5 molar equivalents of either DBCO-PEG4-valcit-MMAF (Levana Biosciences) or DBCO-Cy5.5 (BroadPharm) was added. The click reaction was incubated overnight at room temperature. The conjugate was buffer exchanged into PBS using a 0.5 mL 30 kDa spin concentrator (Thermo Fisher Scientific), to remove excess unconjugated DBCO reagent. Conjugation was monitored by intact protein mass spectrometry using a Xevo G2-XS Mass Spectrometer (Waters).

Cell lines

Cells were grown and maintained in T75 flasks (Thermo Fisher) at 37ºC and 5% CO_2. MDA-MB-231 cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. SKBR3 cells were grown in in McCoy's 5a supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. HCC-2935 cells were grown in RPMI supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. MDA-MB-231

and SKBR3 cells were obtained from the UCSF Cell Culture Facility. HCC-2935 cells were purchased from American Type Culture Collection (ATCC).

Antibody drug conjugate cell viability assay

10,000 SKBR-3 cells were plated in each well of a 96-well plate. After 24 hours, Fab, Conjugate, or drug alone was added in a dilution series. Cells were incubated for 3 days at 37ºC under 5% CO_2. 40 µL of 2.5 mg/mL of Thiazolyl Blue Tetrazolium Bromide (Sigma Aldrich) was added to each well and incubated at 37ºC for 4 h. Next, 100 µL of 10% SDS 0.01M HCl was added. After 4 hours, absorbance at 600 nM was quantified using an Infinite M200 PRO-plate reader (Tecan). Data points were plotted using GraphPad Prism (version 9.1), and curves were generated by using nonlinear regression with Sigmoidal 4PL parameters.

Flow cytometry

500,000 cells were washed with cold phosphate-buffered saline (PBS), lifted with Versene solution and pelleted by centrifugation (500g, 5 min, 4ºC). Cell pellets were washed with cold PBS, before pelleting (500g, 5 min, 4ºC). Cells were blocked with cold PBS + 3% BSA at 4ºC for 10 min. Cells were washed three times with cold PBS + 3% BSA and 20 µg/mL conjugate was added and incubated at 4ºC for 30 min. Cells were washed three times with cold PBS + 3% BSA and finally resuspended in cold PBS. Flow cytometry was performed on a CytoFLEX cytometer (Beckman Coulter), and gating was performed on single, and live cells, before acquisition of 10,000 cells. Analysis was performed using the FlowJo software package.

Peptide mapping

A Preomics iST sample preparation kit was used. In short, 50 µg of protein was incubated in 50 µL of 'LYSE' solution for 10 mins, at 60ºC and 1000 rpm. Sample transferred to a Preomics column and 50 µL of resuspended 'Digest' was added to the sample and incubated at 37ºC for 90

mins at 500 rpm. 100 µL of 'Stop' was added. Sample was spun at 3800xg for 3 minutes and washed twice with 200 µL 'Wash 1' and 'Wash 2'. Sample was eluted using 100 µL of 'Elute'. Sample was concentrated using a GeneVac. Sample was resuspended in 25 µL 2% ACN + 0.1% FA. A Pierce Peptide Quant kit was used to quantify peptide levels and diluted to 100 ng/µL. 200 ng of peptide was run on a TimsTOF mass spectrometer. PEAKS Online was used to analyze data. Input parameters: 20 ppm precursor mass error tolerance, 0.03 fragment mass error tolerance, 3 missed cleavages per sample, peptide length 6-45. Peptide results were triaged using, > 20 spectra per peptide, and the same peptide must be seen in duplicate samples.

Mass spectrometry data acquisition

LC/MS/MS was performed using a Bruker NanoElute chromatography system coupled to a Bruker TimsTOF Pro mass spectrometer. Peptides were separated using a pre-packed IonOpticks Aurora (25 cm x 75 µM) C18 reversed phased column (1.6 µM pore size) fitted with a CaptiveSpray emitter for the TimsTOF Pro CaptiveSpray source. For all samples, 200 ng of resuspended peptides were injected and separated using a linear gradient of 2-15% solvent B (solvent A: 0.1% formic acid + 2% acetonitrile, solvent B: acetonitrile with 0.1% formic acid) over 20 minutes at 400 µL/minute, with a second ramp to 23% over five minutes and a final ramp to 33% B over 5 minutes. Column temperature was held at 50ºC for all separations. Data-dependent acquisition was performed using a TimsTOF PASEF MS/MS method (TIMS mobility scan range per 1.17 seconds; active exclusion 24 seconds; charge range 0-5; minimum MS1 intensity 500). The normalized collision energy was set at 20

DNA Oligos

Single stranded DNA oligos were purchased from Integrated DNA Technologies with or without a 5' biotin modification. The DNA sequence used in this study: 5'-CAT CGT TAC TCT GCT GCG GAC TAC TGG GGT CAA G-3'

Heavy Chain Tag fusions

Tags with a short linker were fused to the C-terminus of the Fab heavy chain, directly after -EPSKC.

HaloTag:

GSGGSGAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPH
VAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGF
HWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIEGTLP
MGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQS
PVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTL
EISG*

SNAP-Tag:

GSGGSGDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAVEVPAPAAVL
GGPEPLMQATAWLNAYFHQPEA
IEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTAL
SGNPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLG*

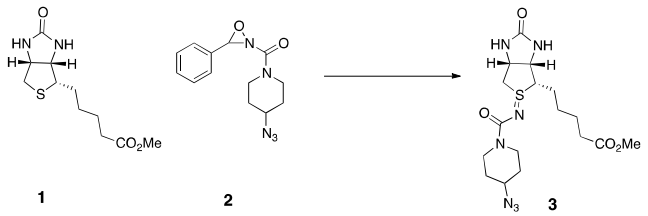
CLIP-Tag:

GSGGSGDKDCEMKRTTLDSPLGKLELSGCEQGLHRIIFLGKGTSAADAVEVPAPAAVLG

GPEPLIQATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISE
SHLAALVGNPAATAAVNTALDGNPVPILIPCHRVVQGDSDVGPYLGGLAVKEWLLAHE
GHRLGKPGLG*

Avi-Tag: DKTHTGGSGSAGGLNDIFEAQKIEWHE*

Reaction of oxaziridine 2 with biotin methyl ester



A 10-mL round-bottom flask was charged with biotin methyl ester 1 (59.0 mg, 228 μmol, 1.0 equiv). Oxaziridine 2 (93.6 mg, 343 μmol, 1.5 equivs) dissolved in 0.5 mL CD_3OD was added. 1 mL D_2O and 0.5 mL CD_3OD were added, resulting in a colorless solution. The reaction mixture was stirred for 45 min and was then concentrated under reduced pressure at 30 ºC on a rotary evaporator. The resulting crude residue was analyzed by ^{1}H-NMR with maleic acid as an internal standard, indicating an 85% yield pre-purification. The residue was then purified by flash column chromatography (silica gel, eluent: MeOH:DCM = 1:9), followed by preparatory TLC in 10 mg batches (silica gel, eluent: MeOH:DCM = 1:9), to afford sulfimide product **3**. (50.8 mg, 52% isolated yield) as a light-yellow oil.

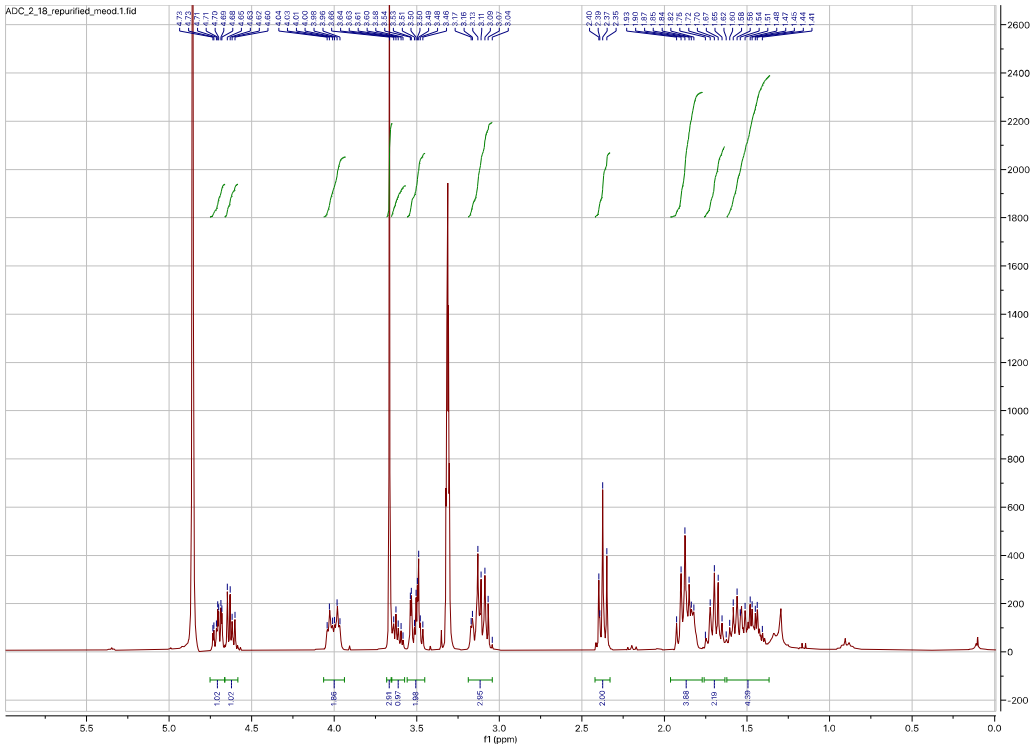
TLC (MeOH:DCM = 1:9): R_f = 0.35 (UV)

^{1}H NMR (300 MHz, CD_3OD) δ 4.71 (ddd, *J* = 8.1, 6.2, 1.9 Hz, 1H), 4.62 (dd, *J* = 8.6, 4.9 Hz, 1H), 4.06 – 3.94 (m, 2H), 3.66 (s, 3H), 3.65 – 3.57 (m, 1H), 3.56 – 3.45 (m, 2H), 3.19 – 3.04 (m, 3H), 2.37 (t, *J* = 7.3 Hz, 2H), 1.96 – 1.77 (m, 4H), 1.69 (q, *J* = 7.2 Hz, 2H), 1.62 – 1.36 (m, 4H).

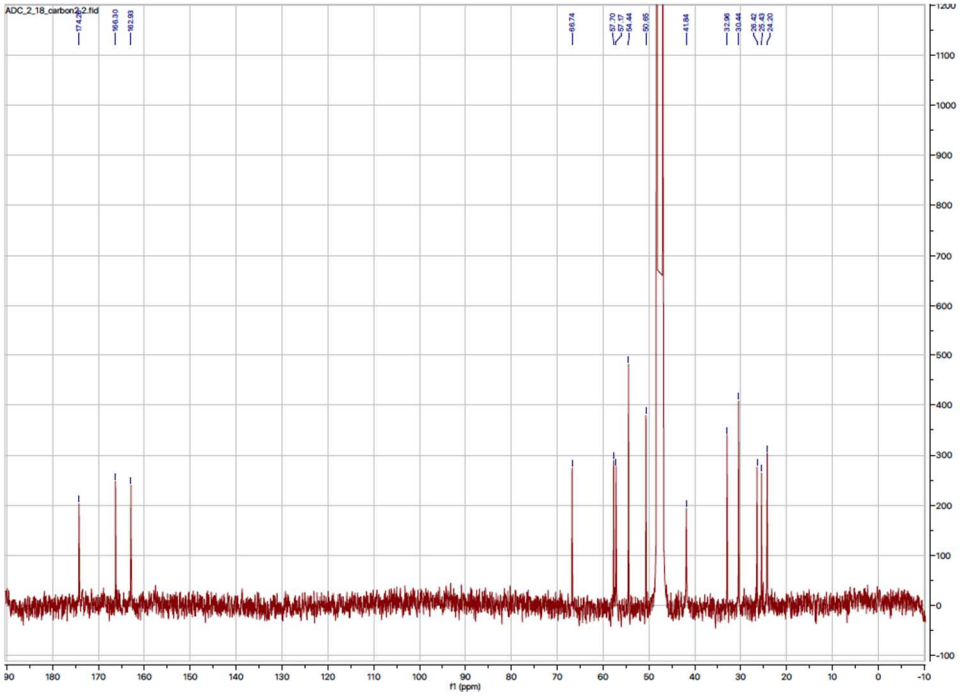
^{13}C NMR (150 MHz, CD_3OD) δ 174.3, 166.3, 162.9, 66.7, 57.7, 57.2, 54.4 (2C), 50.7, 41.8, 33.0, 30.4, 26.4, 25.4, 24.2.

HRMS-ESI m/z calcd for $C_{17}H_{27}N_7O_4S^+$ $[M+H]^+$ 426.1918, found 426.1910

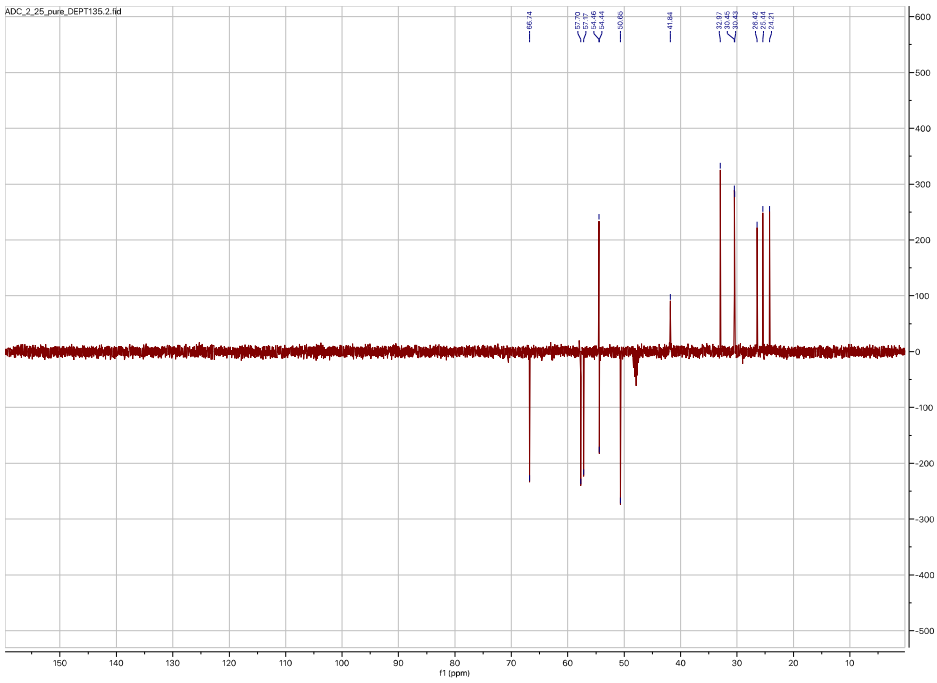
^{1}H-NMR spectrum of biotin conjugate 3 (300 Mhz, CD_3OD)



^{13}C-NMR spectrum of biotin conjugate 3 (150 Mhz, CD_3OD)



^{13}C-APT spectrum of biotin conjugate 3 (150 Mhz, CD_3OD)



3.6 Supplemental Figures

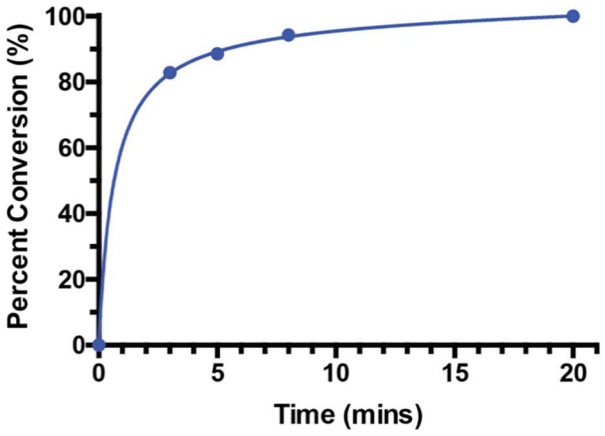


Figure 3.S1: Quantitative NMR was used to track reaction conversion between methyl ester biotin 1, and oxaziridine 2. The reaction is >80% within 3 minutes, and complete conversion was observed within 20 minutes.

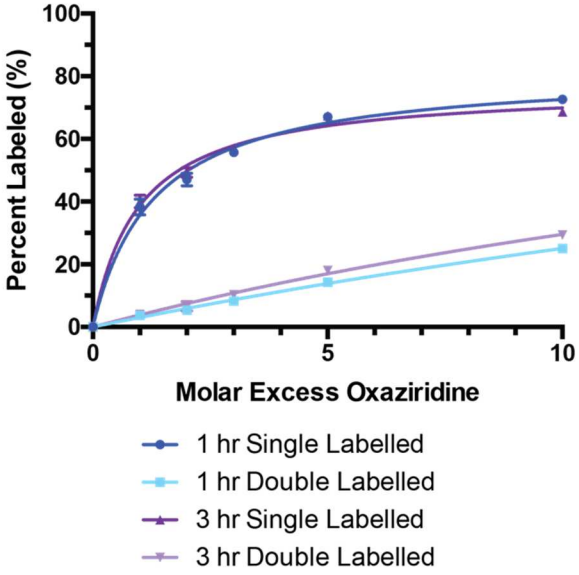


Figure 3.S2: Whole protein mass spectrometry showed that biotinylated Fabs were labeled with oxaziridine 2 in an oxaziridine dependent manner. The labeling reaction observed complete conversion at 10 molar equivalence of oxaziridine reagent, within 1 hour.

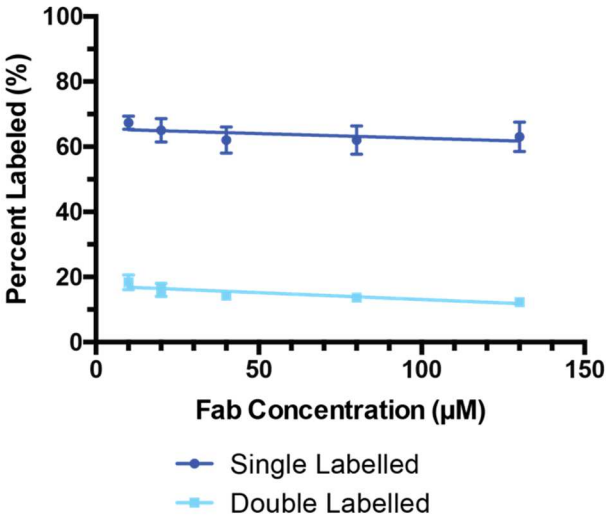


Figure 3.S3: Whole protein mass spectrometry showed that biotinylated Fabs was reacted with 5 molar equivalence oxaziridine 2 at a range of protein concentrations.

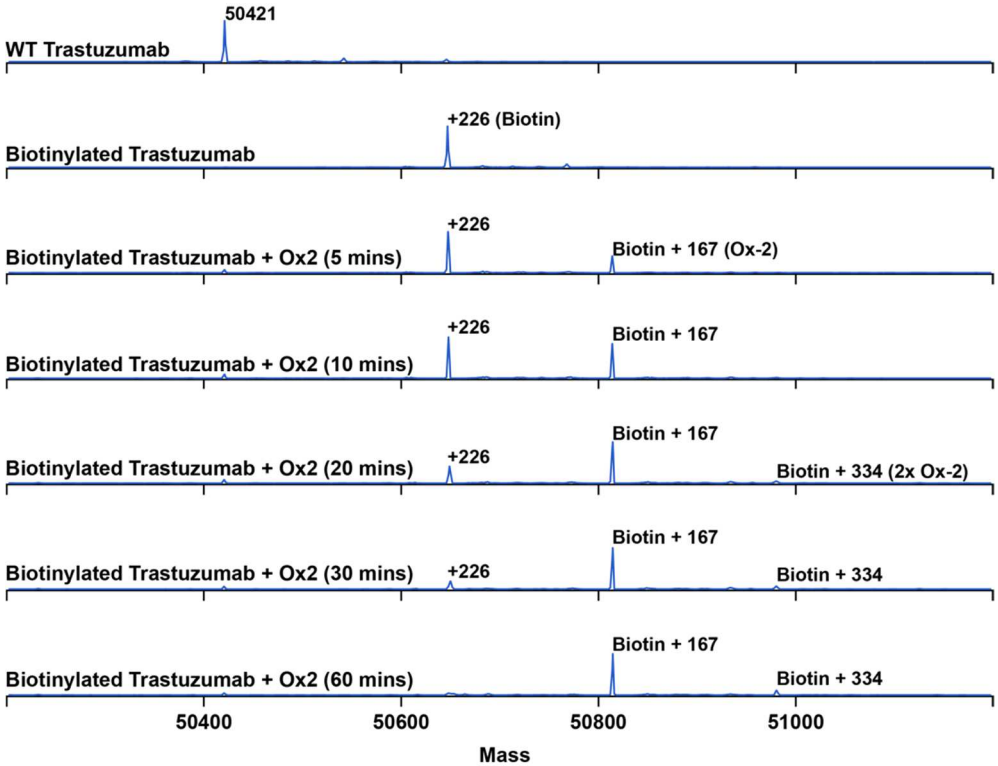


Figure 3.S4: Whole protein mass spectrometry raw data shows a time dependence of labeling. Biotinylated Trastuzumab was incubated with 5x Ox-2. Integration of peaks was used to quantify conversion.

Unbiotinylated Fab

Trastuzumab Heavy Chain:

E(+Oxaziridine)ISEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLE

WVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDG

FYALDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN

SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS

CDKTHTGGSGSAGGLNDIFEAQKIEW**H(+Oxaziridine)**E

Trastuzumab Light Chain:

SDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGV

PSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP

SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST

LTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Biotinylated Fab

Trastuzumab Heavy Chain:

E(+Oxaziridine)ISEVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLE

WVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDG

FYALDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN

SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS

CDKTHTGGSGSAGGLNDIFEAQ**K(+Biotin+Oxaziridine)**IEW**H(+Oxaziridine)**E

Trastuzumab Light Chain:

SDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGV

PSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP

SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST

LTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE

Peptide Count containing each PTM (Average of 2 samples):

Unbiotinylated Fab:

-EISE**-** 77 (269 unmodified)

-IEW**H**- 40 (33 unmodified)

Biotinylated Fab:

-EISE**-** 47.5 (243 unmodified)

-EAQ**K-** 266 (56 unmodified)

-IEW**H**- 30 (205.5 unmodified)

Figure 3.S5: Peptide mapping of labeled Trastuzumab Fab. In addition to reacting with the biotinylated lysine in the Avi-Tag, oxazirdine-2 also reacts with the heavy chain N-terminus and a histidine. Underlining indicates peptide coverage. For peptide counts, values reported are an average of two samples, with bracketed values indicating the number of unmodified peptides containing the modified amino acid.

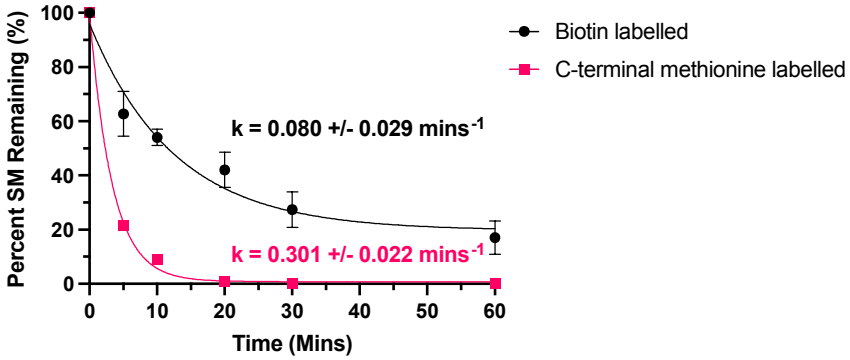


Figure 3.S6: Relative reaction rates for oxaziridine's with either biotinylated proteins or a C-terminal methionine mutant. 50 µM either biotinylated Fab or a C-terminal methionine mutant was incubated with 250 µM Ox-2. Disappearance of starting material was quantified using whole protein mass spectrometry and plotted against time. A non-linear exponential model was used to calculate relative reaction rates.

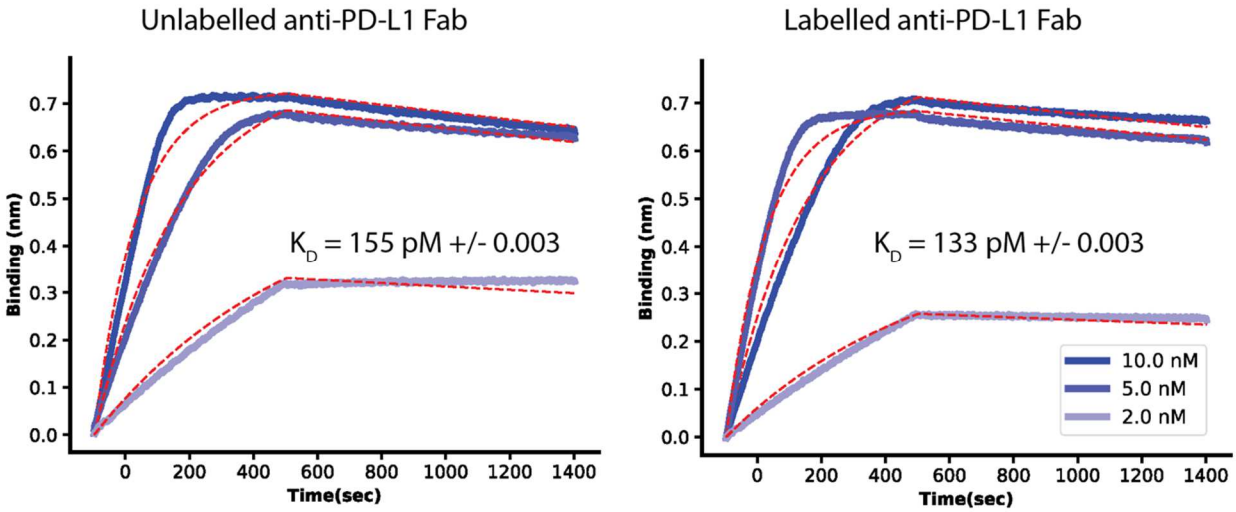


Figure 3.S7: BioReACT labeling of Fabs do not disrupt the Fab antigen interaction. Bio-layer interferometry was used to measure the binding affinity of a labeled or unlabeled anti-PD-L1 Fab to bind to an immobilized PD-L1 ecto domain Fc fusion.

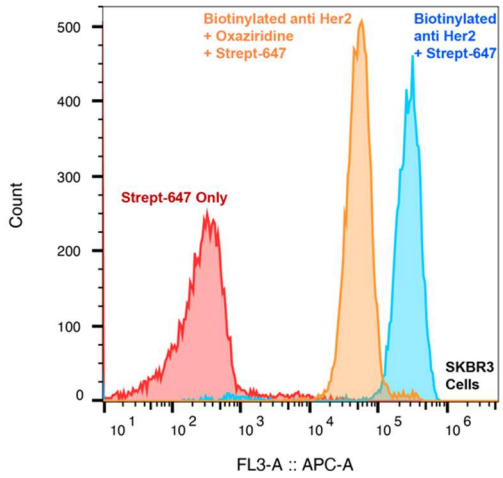


Figure 3.S8: BioReACT modification of biotin disrupts the biotin-streptavidin interaction. SKBR3 cells were incubated with biotinylated anti-Her2 Fab +/- Oxaziridine modification, with a Strept-647 conjugate. The leftward shift after labeling represents an 81% decrease in fluorescent signal.

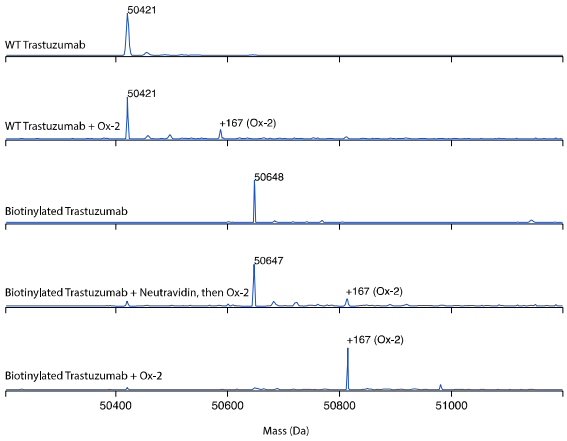


Figure 3.S9: Whole protein mass spectrometry was used to show that 5-minute preincubation with neutravidin blocks biotinylated Fabs from reacting with 10-fold molar excess oxaziridine reagent after 1 hour. The small amount of labeling seen is identical to the background labeling seen in the absence of biotin.

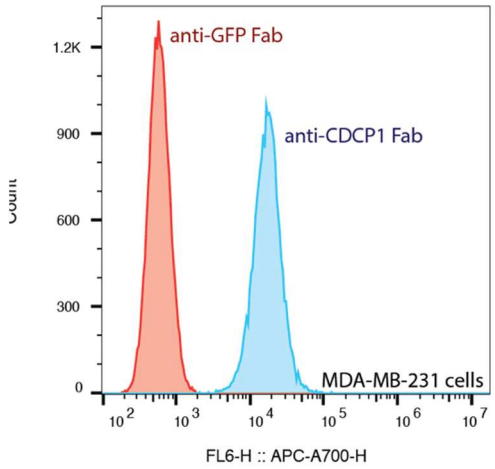


Figure 3.S10: BioReACT was used to generate flow cytometry reagents. DBCO-Cy5.5 was reacted with azide labeled Anti-CDCP1 Fab to produce a fluorescently labeled Fab. Cy5.5 conjugated Fab detected CDCP1 by flow cytometry on the surface of MDA-MB-231 cells

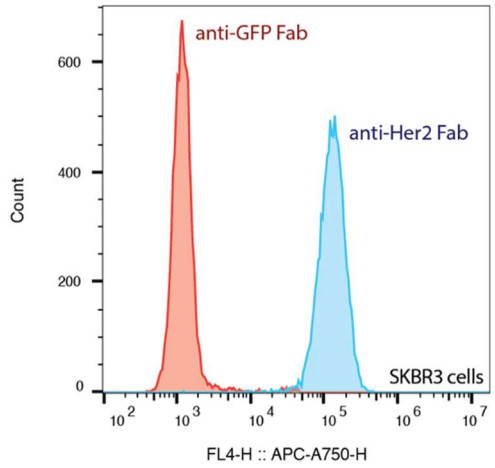


Figure 3.S11: BioReACT was used to generate flow cytometry reagents. DBCO-Cy5.5 was reacted with azide labeled Anti-Her2 Fab to produce a fluorescently labeled Fab. Cy5.5 conjugated Fab detected Her2 by flow cytometry on the surface of SKBR3 cells.

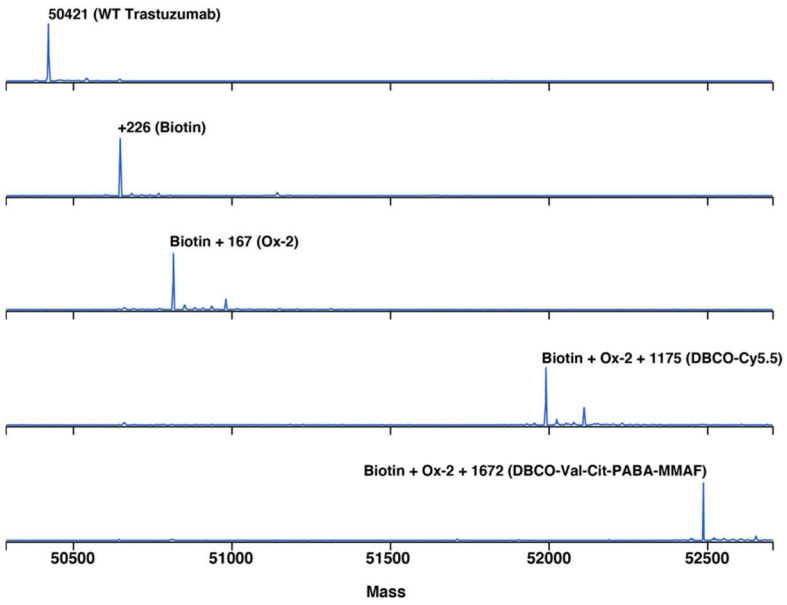


Figure 3.S12: Whole protein mass spectrometry raw data showing the characterization of each construct used in this study. WT Fab, Biotinylated Fab, Biotinylated Fab + Ox-2, Biotinylated Fab + Ox-2 + DBCO-Cy5.5, and Biotinylated Fab + Ox-2 + DBCO-Val-Cit-PABA-MMAF.

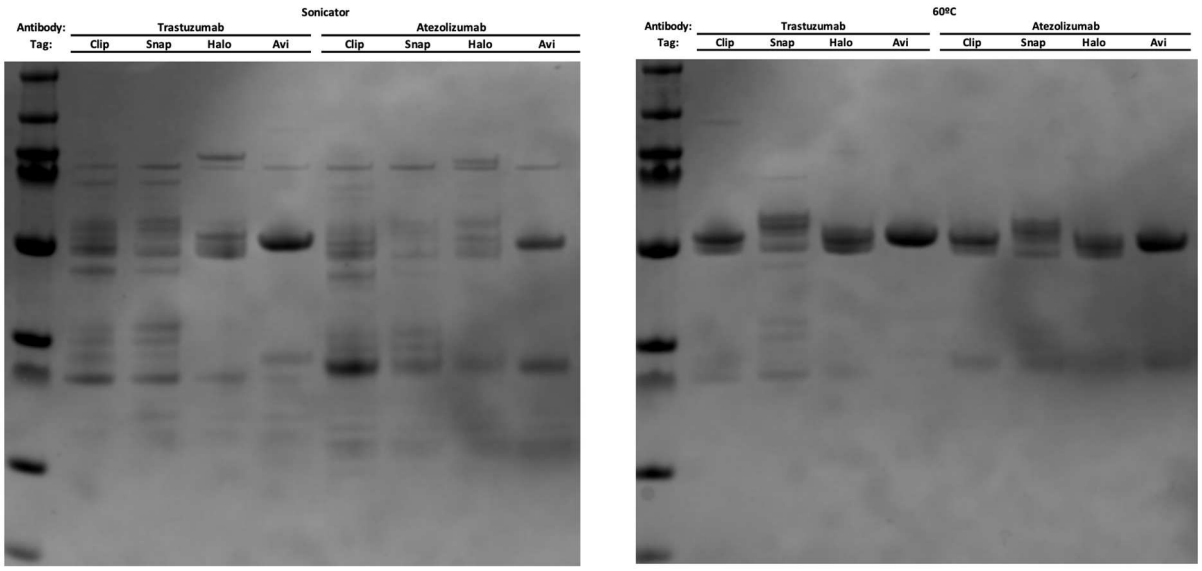


Figure 3.S13: SDS-PAGE gel showing the Avi-tag Fab was the only fusion protein to express purely and with the expected mass. C-terminal heavy chain Fab fusions with either Halo, SNAP, CLIP, or Avi tags were expressed in bacterial cells. Either a 60ºC incubation (right) or sonication in the presence of protease inhibitors (left) was used to lyse the cells. Following a protein-A purification, 2 µg of protein were loaded onto an SDS-PAGE gel. Expected masses in kDa in order from left to right: 67.3, 67.1, 81.4, 50.4, 67.6, 67.5, 81.8, 50.8.

Figure S14: Whole protein mass spectrometry was used to assess the mass products from C-terminal Fab (Trastuzumab – left, Atezolizumab – right) fusions with either Halo, SNAP, CLIP, or Avi tags after a 60ºC lysis step.

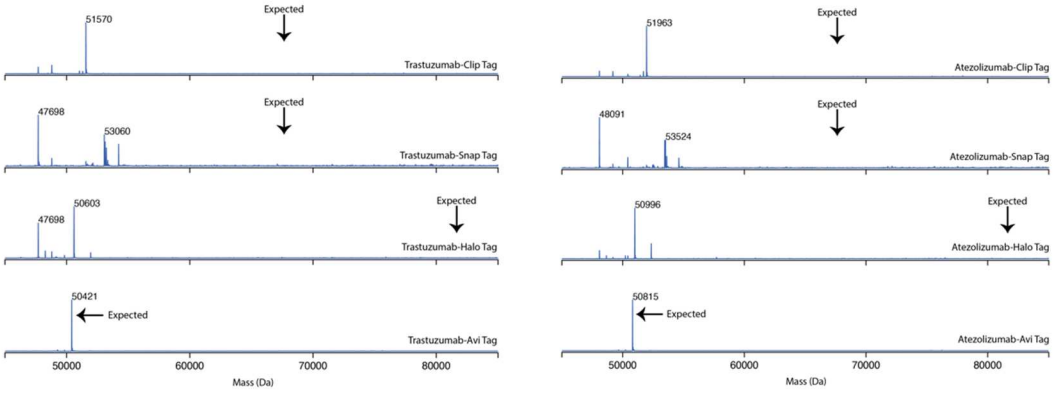


Figure 3.S14: Whole protein mass spectrometry was used to assess the mass products from C-terminal Fab (Trastuzumab – left, Atezolizumab – right) fusions with either Halo, SNAP, CLIP, or Avi tags after a 60ºC lysis step.

3.7 References

(1) Beck, A.; Goetsch, L.; Dumontet, C.; Corvaïa, N. Strategies and Challenges for the next Generation of Antibody–Drug Conjugates. *Nature Reviews Drug Discovery* **2017**, *16* (5), 315–337. https://doi.org/10.1038/nrd.2016.268.

(2) Ohata, J.; Krishnamoorthy, L.; Gonzalez, M. A.; Xiao, T.; Iovan, D. A.; Toste, F. D.; Miller, E. W.; Chang, C. J. An Activity-Based Methionine Bioconjugation Approach To Developing Proximity-Activated Imaging Reporters. *ACS Cent. Sci.* **2020**, *6* (1), 32–40. https://doi.org/10.1021/acscentsci.9b01038.

(3) Benizri, S.; Gissot, A.; Martin, A.; Vialet, B.; Grinstaff, M. W.; Barthélémy, P. Bioconjugated Oligonucleotides: Recent Developments and Therapeutic Applications. *Bioconjugate Chem.* **2019**, *30* (2), 366–383. https://doi.org/10.1021/acs.bioconjchem.8b00761.

(4) Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat Commun* **2014**, *5* (1), 4740. https://doi.org/10.1038/ncomms5740.

(5) Taylor, M. T.; Nelson, J. E.; Suero, M. G.; Gaunt, M. J. A Protein Functionalization Platform Based on Selective Reactions at Methionine Residues. *Nature* **2018**, *562* (7728), 563–568. https://doi.org/10.1038/s41586-018-0608-y.

(6) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; Chang, C. J. Redox-Based Reagents for Chemoselective Methionine Bioconjugation. *Science* **2017**, *355* (6325), 597–602. https://doi.org/10.1126/science.aal3316.

(7) Dyer, K. F. The Quiet Revolution: A New Synthesis of Biological Knowledge. *Journal of Biological Education* **1971**, *5* (1), 15–24. https://doi.org/10.1080/00219266.1971.9653663.

(8) Shaytan, A. K.; Shaitan, K. V.; Khokhlov, A. R. Solvent Accessible Surface Area of Amino Acid Residues in Globular Proteins: Correlation of Apparent Transfer Free Energies with Experimental Hydrophobicity Scales. *Biomacromolecules* **2009**, *10* (5), 1224–1237. https://doi.org/10.1021/bm8015169.

(9) Elledge, S. K.; Tran, H. L.; Christian, A. H.; Steri, V.; Hann, B.; Toste, F. D.; Chang, C. J.; Wells, J. A. Systematic Identification of Engineered Methionines and Oxaziridines for Efficient, Stable, and Site-Specific Antibody Bioconjugation. *Proc Natl Acad Sci USA* **2020**, *117* (11), 5733–5740. https://doi.org/10.1073/pnas.1920561117.

(10) Christian, A. H.; Jia, S.; Cao, W.; Zhang, P.; Meza, A. T.; Sigman, M. S.; Chang, C. J.; Toste, F. D. A Physical Organic Approach to Tuning Reagents for Selective and Stable Methionine Bioconjugation. *J. Am. Chem. Soc.* **2019**, *141* (32), 12657–12662. https://doi.org/10.1021/jacs.9b04744.

(11) Sung, K.; Maloney, M. T.; Yang, J.; Wu, C. A Novel Method for Producing Mono-Biotinylated, Biologically Active Neurotrophic Factors: An Essential Reagent for Single Molecule Study of Axonal Transport. *J Neurosci Methods* **2011**, *200* (2), 121–128. https://doi.org/10.1016/j.jneumeth.2011.06.020.

(12) Williams, J. G. K.; Steffens, D. L.; Anderson, J. P.; Urlacher, T. M.; Lamb, D. T.; Grone, D. L.; Egelhoff, J. C. An Artificial Processivity Clamp Made with Streptavidin Facilitates Oriented Attachment of Polymerase-DNA Complexes to Surfaces. *Nucleic Acids Res* **2008**, *36* (18), e121. https://doi.org/10.1093/nar/gkn531.

(13) Valadon, P.; Darsow, B.; Buss, T. N.; Czarny, M.; Griffin, N. M.; Nguyen, H. N.; Oh, P.; Borgstrom, P.; Chrastina, A.; Schnitzer, J. E. Designed Auto-Assembly of

Nanostreptabodies for Rapid Tissue-Specific Targeting in Vivo. *J Biol Chem* **2010**, *285* (1), 713–722. https://doi.org/10.1074/jbc.M109.061838.

(14) Sims, S.; Willberg, C.; Klenerman, P. MHC-Peptide Tetramers for the Analysis of Antigen-Specific T Cells. *Expert Rev Vaccines* **2010**, *9* (7), 765–774. https://doi.org/10.1586/erv.10.66.

(15) Livnah, O.; BAYERt, E. A.; WILCHEKt, M.; Sussman, J. L. Three-Dimensional Structures of Avidin and the Avidin- Biotin Complex. *Proc. Natl. Acad. Sci. USA* **1993**, 5.

(16) Glasel, J. A. A Nuclear Magnetic Resonance Investigation of Biotin. The Biotin Sulfonium Ion *. *Biochemistry* **1966**, *5* (6), 1851–1855. https://doi.org/10.1021/bi00870a010.

(17) Beckett, D.; Kovaleva, E.; Schatz, P. J. A Minimal Peptide Substrate in Biotin Holoenzyme Synthetase-Catalyzed Biotinylation. *Protein Sci* **1999**, *8* (4), 921–929.

(18) Cho, H.-S.; Mason, K.; Ramyar, K. X.; Stanley, A. M.; Gabelli, S. B.; Jr, D. W. D.; Leahy, D. J. Structure of the Extracellular Region of HER2 Alone and in Complex with the Herceptin Fab. **2003**, *421*, 5.

(19) Zanon, P. R. A.; Yu, F.; Musacchio, P.; Lewald, L.; Zollo, M.; Krauskopf, K.; Mrdović, D.; Raunft, P.; Maher, T. E.; Cigler, M.; Chang, C.; Lang, K.; Toste, F. D.; Nesvizhskii, A. I.; Hacker, S. M. Profiling the Proteome-Wide Selectivity of Diverse Electrophiles. **2021**. https://doi.org/10.26434/chemrxiv.14186561.v1.

(20) Rebhan, M. A. E.; Brunschweiger, A.; Hall, J. Measurement by SPR of Very Low Dissociation Rates: Oxidation-Mediated Loss of Biotin–Streptavidin Affinity. *ChemBioChem* **2013**, *14* (16), 2091–2094. https://doi.org/10.1002/cbic.201300468.

(21) McKinnon, K. M. Flow Cytometry: An Overview. *Curr Protoc Immunol* **2018**, *120*, 5.1.1-5.1.11. https://doi.org/10.1002/cpim.40.

(22) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126* (46), 15046–15047. https://doi.org/10.1021/ja044996f.

(23) Dommerholt, J.; Rutjes, F. P. J. T.; van Delft, F. L. Strain-Promoted 1,3-Dipolar Cycloaddition of Cycloalkynes and Organic Azides. *Top Curr Chem (Z)* **2016**, *374* (2), 16. https://doi.org/10.1007/s41061-016-0016-4.

(24) Zhang, F.; Qi, X.; Wang, X.; Wei, D.; Wu, J.; Feng, L.; Cai, H.; Wang, Y.; Zeng, N.; Xu, T.; Zhou, A.; Zheng, Y. Structural Basis of the Therapeutic Anti-PD-L1 Antibody Atezolizumab. *Oncotarget* **2017**, *8* (52), 90215–90224. https://doi.org/10.18632/oncotarget.21652.

(25) Martinko, A. J.; Truillet, C.; Julien, O.; Diaz, J. E.; Horlbeck, M. A.; Whiteley, G.; Blonder, J.; Weissman, J. S.; Bandyopadhyay, S.; Evans, M. J.; Wells, J. A. Targeting RAS-Driven Human Cancer Cells with Antibodies to Upregulated and Essential Cell-Surface Proteins. *eLife* **2018**, *7*, e31098. https://doi.org/10.7554/eLife.31098.

(26) Drago, J. Z.; Modi, S.; Chandarlapaty, S. Unlocking the Potential of Antibody-Drug Conjugates for Cancer Therapy. *Nat Rev Clin Oncol* **2021**. https://doi.org/10.1038/s41571-021-00470-8.

(27) Doronina, S. O.; Mendelsohn, B. A.; Bovee, T. D.; Cerveny, C. G.; Alley, S. C.; Meyer, D. L.; Oflazoglu, E.; Toki, B. E.; Sanderson, R. J.; Zabinski, R. F.; Wahl, A. F.; Senter, P. D. Enhanced Activity of Monomethylauristatin F through Monoclonal Antibody Delivery: Effects of Linker Technology on Efficacy and Toxicity. *Bioconjugate Chem.* **2006**, *17* (1), 114–124. https://doi.org/10.1021/bc0502917.

(28) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3* (6), 373–382. https://doi.org/10.1021/cb800025k.

(29) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chemistry & Biology* **2008**, *15* (2), 128–136. https://doi.org/10.1016/j.chembiol.2008.01.007.

4. When are SARS-CoV-2 neutralizing antibodies good enough as a therapeutic, and can we make enough to have an impact?

4.1 Abstract

A significant challenge in preclinical drug development is knowing when a lead candidate is good enough. This challenge is especially apparent for groups developing antibody-based therapeutics, as there are numerous ways to systematically improve the potency of an antibody using well-trodden protein engineering efforts. While well established, these efforts are both time and capital consuming. However, during the global SARS-CoV-2 pandemic there is a critical need to rapidly develop effective therapeutics to both treat and prevent the spread of this deadly virus. With each passing week, there are mounting reports of new antibodies that neutralize the virus *in vitro*, with neutralization IC_{50}'s ranging from low picomolar to tens of nanomolar. Although there is the potential to further enhance the potencies of these constructs, we wanted to study at what point these efforts should transition to clinical candidacy.

4.2 Introduction

Therapeutic antibody efforts towards treating SARS-CoV-2 viral infections have so far targeted the viral entry pathway[1,2]. The virus utilizes its spike protein receptor-binding-domain (RBD) to recognize and interact with the angiotensin converting enzyme-2 (ACE2) protein on the surface of host cells. To date, the most efficacious antibodies have targeted the spike RBD domain to disrupt

ACE2 engagement[3]. In considering the advancement of these recombinant antibodies for clinical use as passive immune anti-virals, we envisioned three key questions that could be addressed with pharmacokinetic modeling: 1. Will there be molar excess of antibody to RBD that reaches the lungs at steady-state? 2. Given a set dosage, will the antibody concentration in the lungs at steady-state reach a therapeutically relevant level as estimated from *in vitro* viral neutralization assays? 3. Given the current large number of COVID-19 patients each month, would a single pharmaceutical company have the production capacity to meet US demand?

4.2 Results

To begin our analysis, we developed an abbreviated target product profile (TPP) for an antibody-based therapeutic for acute COVID-19 infection. We imagined a single dose of an IgG1, with standard half-life and Fc effector functions. The values selected are in-line with other non-viral therapeutic IgG's. We thought it apt to choose virion variables on the conservative side to reflect that these are estimates. To model the pharmacokinetic parameters, we used a two-compartment antibody biodistribution model[4]. The antibody biodistribution coefficient for the lung tissue is estimated at ~15% as compared to steady-state plasma concentrations[5]. However, based on studies of the FDA approved antibody, palivizumab, directed to the respiratory syncytial virus (RSV) for upper and lower respiratory tract infections in infants, the concentration of antibody in the lung sputum is estimated to be 500-fold lower than the plasma concentration[6]. We considered therapeutically relevant doses ranging from 1-15 mg/kg. As a reference, palivizumab antibody is dosed at 15 mg/kg (subcutaneously) monthly[7].

Biologic Variables	
Number of Doses	1
Biologic Mass	150 kDa
Desired IC_{value}	IC_{90} or IC_{95}
Duration of effect	14 days
Biologic Half-Life	18 days
Production yield	2 g/L

Virion Variables	
Viral Titer	10^{10} per/mL
Spike trimers per virion	100
Number of RBDs accessible per trimer	1

PK Variables	
Central Compartment	46 mL/kg
Peripheral Tissue	31 mL/kg
Antibody biodistribution coefficient from plasma to sputum	0.2%
Clearance Model	Linear

Figure 4.1: Definition of variables used in this study.

1. Will there be molar excess of antibody to RBD that reaches the lungs at a steady-state?

At a traditional antibody dose of 8 mg/kg, our model predicts at day 14 post-administration there will be a 510-fold molar excess of antibody to viral RBD in the lung sputum (10^{10} virions per mL) (**Fig: 4.2**). In fact, a 1:1 ratio of antibody to spike trimer is not achieved until dosing at 16 µg/kg. This result suggests we will always be in antibody molar excess per virion RBD in the lung sputum of patients.

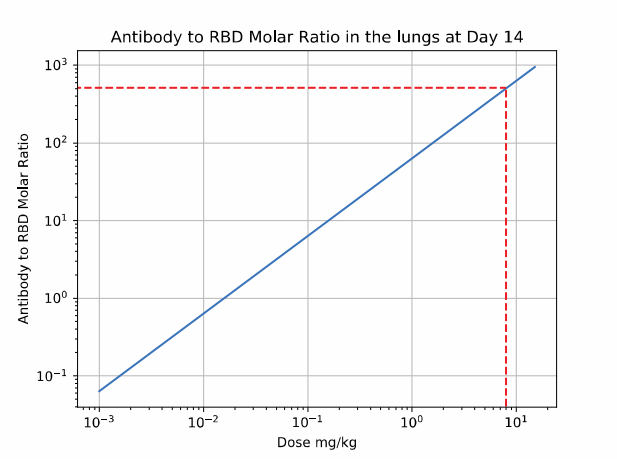


Figure 4.2: Antibody to RBD Molar Ratio in the lungs at Day 14.

2. Given a set dosage, will the antibody concentration in the lungs at steady-state reach a therapeutically relevant level as estimated from *in vitro* viral neutralization assays?

To begin this analysis, we must first state the biggest assumption when predicting requisite *in vivo* IC_{50}'s is that we assume that viral neutralization assays in cell culture are an accurate depiction of

events in the patient lung. At a dose of 8 mg/kg the day 14 sputum concentration is estimated to be 847 pM (**Fig. 4.3a**). Next, given this information and a goal to achieve either IC_{90} or IC_{95} (assuming a hill slope of 1) at day 14, we asked what these *in vivo* concentrations would relate to in terms of *in vitro* viral neutralization IC_{50}. For a dose of 8 mg/kg, the IC_{90} and IC_{95} targeted values are 94 and 45 pM respectively (**Fig. 4.3b**). The estimates provide useful target *in vitro* viral neutralization IC_{50}'s the community should be aiming for. Overall, the modeling suggests that one should be targeting a low pM IC_{50} for a systemically delivered antibody.

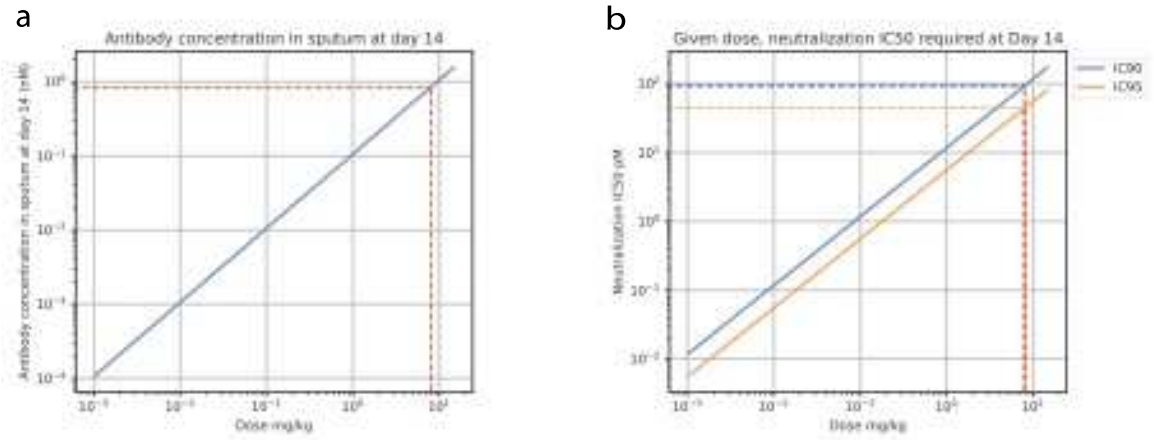


Figure 4.3: Given a dose, will the antibody concentration in the lungs reach a therapeutically relevant level. (a) Antibody concentration in the sputum at day 14 as a function of dose. (b) Neutralization IC50 required at Day 14 as a function of dose.

3. Given the current large number of COVID-19 patients each month, would a single pharmaceutical company have the production capacity to meet US demand?

The final question examines if the first company to take a therapy to market will be able to produce enough to achieve wide market penetration, thus rendering other efforts futile. This analysis led us to develop a second model based on three different cohorts of patients receiving an antibody-based SARS-CoV-2 therapeutic. 1. Only the sickest patients admitted to intensive care units in the US.

2. All patients who test positive for the virus in an effort to reduce the viral load in the community. 3. Prophylactic administration to all front-line workers, those over 60, and all immuno-comprised people. We chose to use the per monthly numbers for the above populations, with front-line workers including healthcare workers, law enforcement, and the fire service. The last population parameter was the average mass of the US population[8]. We chose a model pharmaceutical company named COVID Inc with a total fermenter capacity of 150,000 liters per two weeks. This represents 100% capacity for a typical pharmaceutical company which would be required to halt all other production operations.

Population Variables	
Average male mass	89.7 kg
Average female mass	77.3 kg
Total cases per day	25,000
Percent critical cases	1.48%

Production Variables	
Total fermenter capacity	150,000
Production turn-over	2 weeks
Antibody yield	2 g/L

Figure 4.4: Assumed population and production variables.

The model predicts that treating cohort 1 is viable given all doses calculated, with a dose of 8 mg/kg reaching a mere 1.2% capacity (**Fig. 4.5a**). For cohort 2, at a dose of 8 mg/kg, the total percent capacity is 84%. 100% is reached at a dose of 9.6 mg/kg (**Fig. 4.5b**). Finally, the drastic increase in size to cohort 3 shows that COVID Inc can't produce enough antibody to meet the prophylactic demands, with a dose of 8 mg/kg corresponding to 10,354% capacity (**Fig. 4.5c**). 100% capacity correlates to a dose of 77 μg/kg; at this dose our IC_{50} model predicts required *in vitro* viral neutralization IC_{50}'s of 905 or 429 fM to attain an IC_{90} or IC_{95} respectively[3]. These values are an order of magnitude below the current best neutralizing antibody reported. This dose corresponds to a 5-fold molar excess of antibody to RBD.

a

Cohort	Population	Mass of antibody needed (kg)	Percent Capacity (%)
1	11,100	7.4	1.2
2	750,000	501	84
3	93,000,000	62,124	10,354

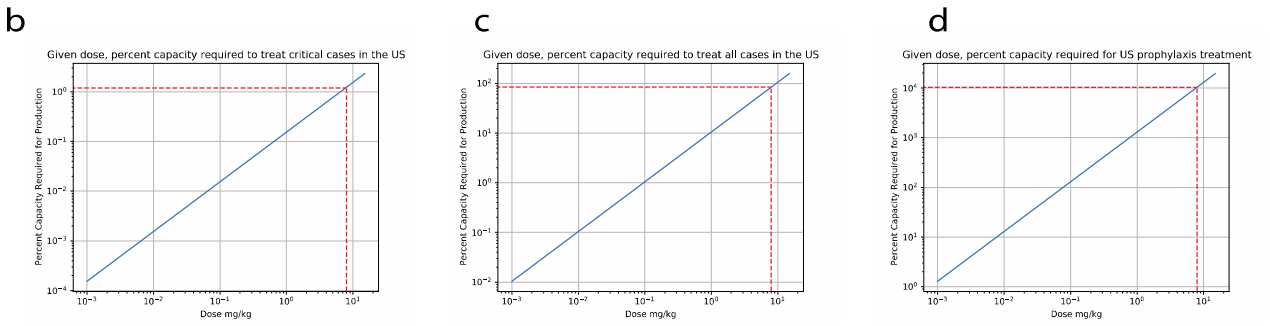


Figure 4.5: Percent pharmaceutical production capacity required to produce enough Antibody to treat a range of populations to as a function of dose. (a) Population sizes and percent capacity of COVID Inc at 8 mg/kg. (b) Percent capacity as a function of dose for treating US critical cases. (c) Percent capacity as a function of dose for treating all US cases. (d) Percent capacity as a function of dose for US prophylaxis treatment.

4.3 Discussion

Our models have these important takeaways for the development of antibodies towards the treatment of the COVID-19 pandemic:

1. At all therapeutically relevant doses, there is a large molar excess of antibodies to RBD domains in the sputum on day 14.
2. Protein engineering efforts should proceed to optimize lead compounds until they achieve neutralization IC_{50}'s in the tens of picomolar range.
3. A single pharmaceutical company has the potential capacity to meet the demands of acute treatment, indicating that the conventional advantages of being first to market will prevail in this case.

4. No single company has the capacity to administer their antibodies as a prophylactic, indicating joint efforts and collaborations will need to be forged, or the more likely reliance on a vaccine.

4.4 References

1. Yuan, A. Q. et al. Isolation of and Characterization of Neutralizing Antibodies to Covid-19 from a Large Human Naïve scFv Phage Display Library. http://biorxiv.org/lookup/doi/10.1101/2020.05.19.104281 (2020) doi:10.1101/2020.05.19.104281.

2. Zeng, X. et al. Blocking antibodies against SARS-CoV-2 RBD isolated from a phage display antibody library using a competitive biopanning strategy. http://biorxiv.org/lookup/doi/10.1101/2020.04.19.049643 (2020) doi:10.1101/2020.04.19.049643.

3. Robbiani, D. F. et al. Convergent Antibody Responses to SARS-CoV-2 Infection in Convalescent Individuals. http://biorxiv.org/lookup/doi/10.1101/2020.05.13.092619 (2020) doi:10.1101/2020.05.13.092619.

4. Betts, A. et al. Linear pharmacokinetic parameters for monoclonal antibodies are similar within a species and across different pharmacological targets: A comparison between human, cynomolgus monkey and hFcRn Tg32 transgenic mouse using a population-modeling approach. mAbs 10, 751–764 (2018).

5. Shah, D. K. & Betts, A. M. Antibody biodistribution coefficients: Inferring tissue concentrations of monoclonal antibodies based on the plasma concentrations in several preclinical species and human. mAbs 5, 297–305 (2013).

6. Hart, T. K. et al. Preclinical efficacy and safety of mepolizumab (SB-240563), a humanized monoclonal antibody to IL-5, in cynomolgus monkeys. Journal of Allergy and Clinical Immunology 108, 250–257 (2001).

7. MedImmune. SYNAGIS® (PALIVIZUMAB) for Intramuscular Administration. (1999).

8. National Health Statistics Reports, Number 122, December 20, 2018. 16 (2018).

5. Examining Gender Imbalance in Chemistry Authorship

5.1 Abstract

Despite decades of progress towards a more equitable society, gender representation in the sciences continues to be heavily skewed towards men. We were interested in gender representation in chemistry through the lens of scientific publishing. Publications are a central academic currency and are critical for funding, recruiting, and promotion in academia. Here we report the results of an analysis that compared the percentage of female first and last authors across 10 chemistry, 3 chemical biology, and 3 general journals over the past 15 years. We show that women are substantially underrepresented in chemistry authorship even when compared to their relative populations in academia and are not predicted to achieve parity within the next 50 years at the current rate in any journal. Our findings highlight the need for changes to the publishing process to achieve a more equitable publishing environment.

5.2 Introduction

There have been numerous validated studies demonstrating that diversity in teams improves creativity and productivity.[1,2] Beyond team structures, increasing diversity is critical for the progression of robust scientific systems[3]. Recently, many publications contributed by leaders in both industry and academia on this subject have provided thoughtful steps we can take to improve the diversity in our field[4,5,6]. Here we investigate gender representation in chemistry authorship,

chemistry departments, and editorial boards. Our goal with this study is to provide data on trends in the representation of women as authors in chemistry.

PubMed contains an abundance of publicly available authorship data. Several recent studies have emerged that use large data sets of authors names in combination with name-based gender predictive software to analyze trends in authorship, we were interested in performing a similar analysis for chemistry journals.[7,8,9,10,11,12] We wrote a python-based web scraper to extract the first (given) name of the first and last authors from papers published on PubMed each year since 2005. Since some chemistry journals are not archived in PubMed until the mid- to late 2000's, we chose 2005 as a year to start the analysis[19]. PubMed does not provide metadata for cases in which there are multiple first or multiple corresponding authors; as a result additional authors beyond first and last listed are not included in our analysis. There are papers where the last author is not a corresponding author, but this instance is rare; from a random sampling of 100 papers in our dataset, this event occurred 3 times. Therefore, we refer to last author as corresponding author throughout this work. We focused our efforts on organic chemistry journals and the closely related chemical biology field. We also included Nature, Cell, and Science (NCS) as general comparators. Self-identification is the only accurate way to assign gender, but such data is not currently widely available for authors in chemistry journals. In the absence of sufficient self-identification data and given that many first names strongly correlate with gender, we sought to use first names to predict gender probabilistically. We utilized Gender API[16] to determine a gender probability and confidence score for each name, using country to improve accuracy for common names. Due to the binary nature of the API, we could only account for two genders: male and female. There were occasions where the API couldn't interpret the name or where the name was not in the database, and these instances were removed from the analysis. Gender API outputs an accuracy score and

sampling size, we set strict data quality requirements to calculate population parameters. Any names with <95% accuracy score and <100 sample size were excluded to minimize uncertainty. With our high accuracy criteria, our process removes a larger proportion of non-western names, meaning that any conclusions drawn are based upon mainly western names (**Figure 5.S8-11)**. Additionally, gender-neutral names result in low accuracy scores, so many of these were excluded from the analysis (**Figure 5.S4**). Any years for a journal with fewer than 25 names with >95% confidence were removed from any time-series analyses. Of 404,676 authors, we removed 125,151 that did not meet the inclusion criteria (**Figure 5.S3-4**). To ensure we are not biasing female names in the exclusion groups, we performed detailed analysis of the included and excluded datasets, showing that the gender ratios in the excluded group are similar to that of the unfiltered dataset (**figure 5.S5-7**). Details, graphical figures, individual journal graphs, and further discussion about the limitations of our approach can be found in the supporting information. All Python scripts used are available on github[17]. Raw data files for each journal, by year and author type can be found on Zenodo[18].

5.3 Results

We first assessed cumulative gender distribution for first authors and for last (corresponding) authors over the past 16 years (**Figure 5.1**). In *Journal of the American Chemical Society* (JACS*)*, a leading chemistry journal, women have represented 9.5% of corresponding authors and 21.4% of first authors (**Figure 5.1**). These values were consistent across the different chemistry-specific journals, ranging from 6.8% and 19.7% in Tetrahedron to 11% and 24.9% in Chemical Science. Interestingly medicinal chemistry journals such as *Journal of Medicinal Chemistry* (J Med Chem) had a higher representation of women, who made up 15.7% of corresponding authors and 29.9%

of first authors. Chemical biology is a closely related field to Chemistry, but gender ratios in the three chemical biology journals included in this study grouped separately from chemistry journals. ACS Chemical Biology had the highest representation of women, who comprised 18.4% of corresponding authors and 34.6% of first authors. It is important to note that these values are still far from representative of the gender demographics seen in the general population. Interestingly the NCS journals separated in two groups throughout our analyses, with Nature and Science closely correlating with each other while Cell had a higher representation of female authors. Finally, there is a large disparity between first authors and corresponding authors (Figure 1a vs 1b). This trend tracks with faculty and student gender percentages in academia: 39% of chemistry graduate students and only 12% of faculty are women.[13,14] However, Female authors are still under-represented compared with these population values (vide infra). This aligns with the documented decrease in the percentage of women in science between PhD and faculty jobs[15].

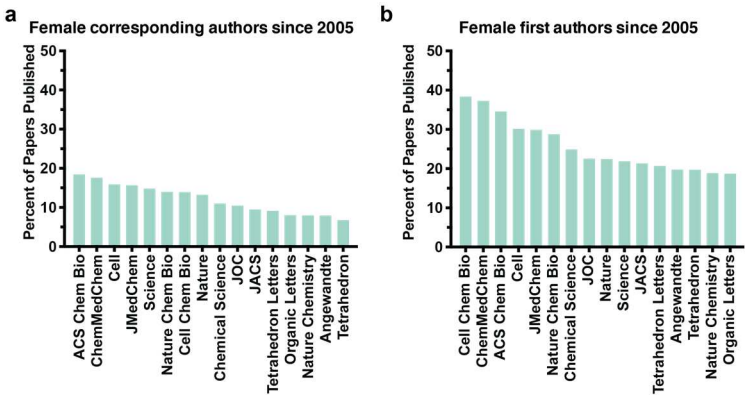


Figure 5.1: Rank-ordered breakdown of female representation as corresponding and first authors for papers published since 2005. A) Percentage of female corresponding authors for papers published in chemistry, chemical biology and NCS journals since 2005. B) Percentage of female first authors for papers published in chemistry, chemical biology and Nature, Cell, Science (NCS) journals since 2005. *Name-predicted gender–see methods section for details.

Next, we plotted the percentage of female and male corresponding authors for each journal since 2005. Given the results from the collated data we separated the journals into chemistry (**Figure 5.2a**), medicinal chemistry (**Figure 5.2b**), chemical biology (**Figure 5.2c**) and NCS (**Figure 5.2d**). Due to the high proportion of papers originating the USA (28% of the included dataset; Germany is second with 9% of the share), we have included the US female academic faculty percentages in synthetic chemistry in 2020 as a comparison (**Figure 5.S8-10**).[11] We next performed a linear regression of each time series and assessed the statistical significance of a non-zero slope of the resulting fit (**Table 5.S1**). Across all 16 journals, only 6 had a statistically significant positive trend that would indicate an increase in the percentage of female authors. *Journal of Medicinal Chemistry* had the highest representation of women of all journals, with 23.7% female corresponding authors for Jan-June 2020. This journal also had the greatest positive trend at 0.5% per year. At this rate, *J. Med. Chem.* would reach equal representation of female authors in 2080. On the other end of the spectrum, Angewandte had 7.9% female corresponding authors and if current trends hold, it will take until the year 2283 before parity is reached.

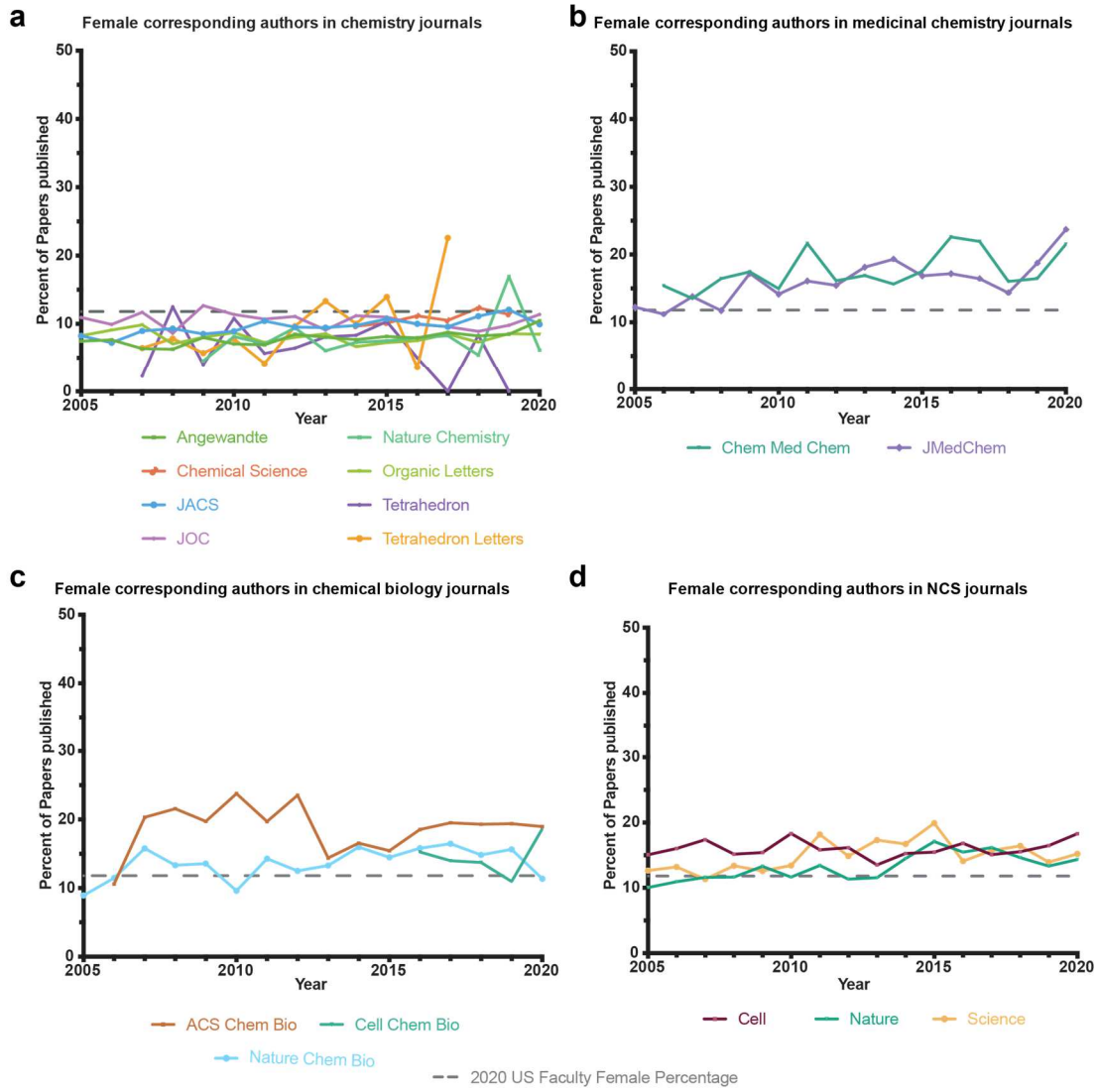


Figure 5.2: Percentage of female corresponding authors for different journal types per year since 2005. Dashed lines represent current (2020) faculty percentages in US chemistry departments. A) Chemistry journals. B) Medicinal chemistry journals. C) Chemical biology journals. D) Nature, Cell, Science. *Name-predicted gender–see methods section for details.

We completed the same time-series analysis for first authors. We included the NSF self-reported gender percentages for US graduate student and postdoctoral fellows as a comparator (dashed line). Only 4 of the 16 journals have showed a statistically significant positive trend (**Figure 5.3, table 5.S2**). Similar to corresponding authors, underrepresentation of women is most pronounced in chemistry-specific journals (**Figure 5.3a**). Both medicinal chemistry and chemical biology first

author percentages have tracked closely with the percentages seen in graduate schools, suggesting women have more equitable opportunities to be first authors in these fields.

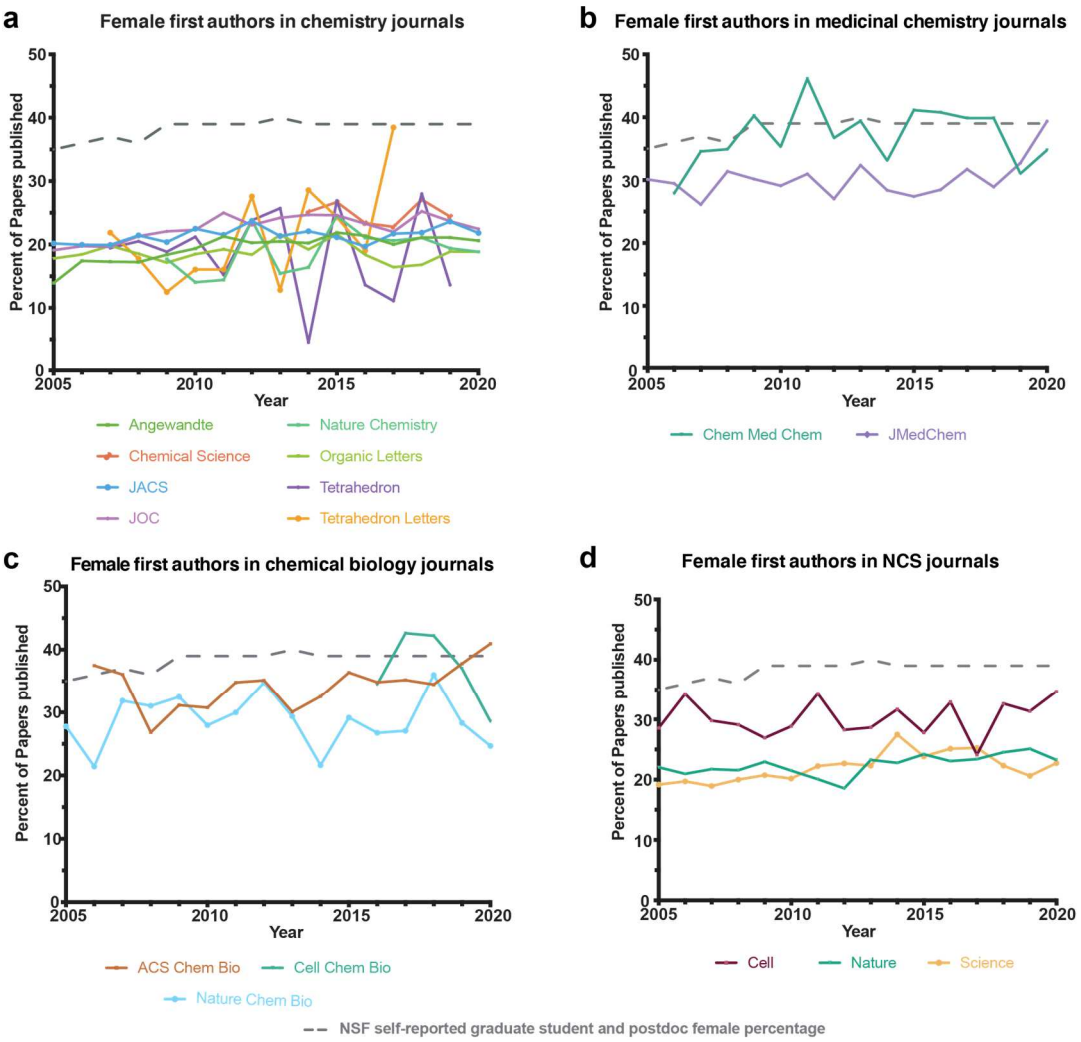


Figure 5.3: Percentage of female first authors for different journal types per year since 2005. Dashed lines represent the NSF self-reported gender percentages of US graduate students and postdoctoral fellows. A) Chemistry journals. B) Medicinal chemistry journals. C) Chemical biology journals. D) Nature, Cell, Science. *Name-predicted gender–see methods section for details.

We next evaluated whether female corresponding authors were more likely to have female first authors. We show four representative journals in figure 4, and the remaining data can be accessed

in the Supporting Information. We performed a binomial test to assess if a first authors were more likely to be women if the corresponding author was a woman. Except for Nature Chemistry, female corresponding authors had a significantly higher chance of having a female first author than for male corresponding authors (**Table 5.S3**). The greatest disparity was for ACS Chemical Biology where 48.1% of first authors from female corresponding authors was female, whereas only 32.1% were female for male corresponding authors. In two cases, Cell Chemical Biology and ChemMedChem, female corresponding authors had greater than 50% of their first authors being female (50.8 and 51.4% respectively). Three out of the four journals that showed statistically significant positive trends in female first author percentages also showed a positive trend in female corresponding percentages (**Table 5.S1, 5.S2**). This suggests that progress towards parity in first author gender ratios (**Figure 5.3**) may be due in part to the increasing number of female corresponding authors.

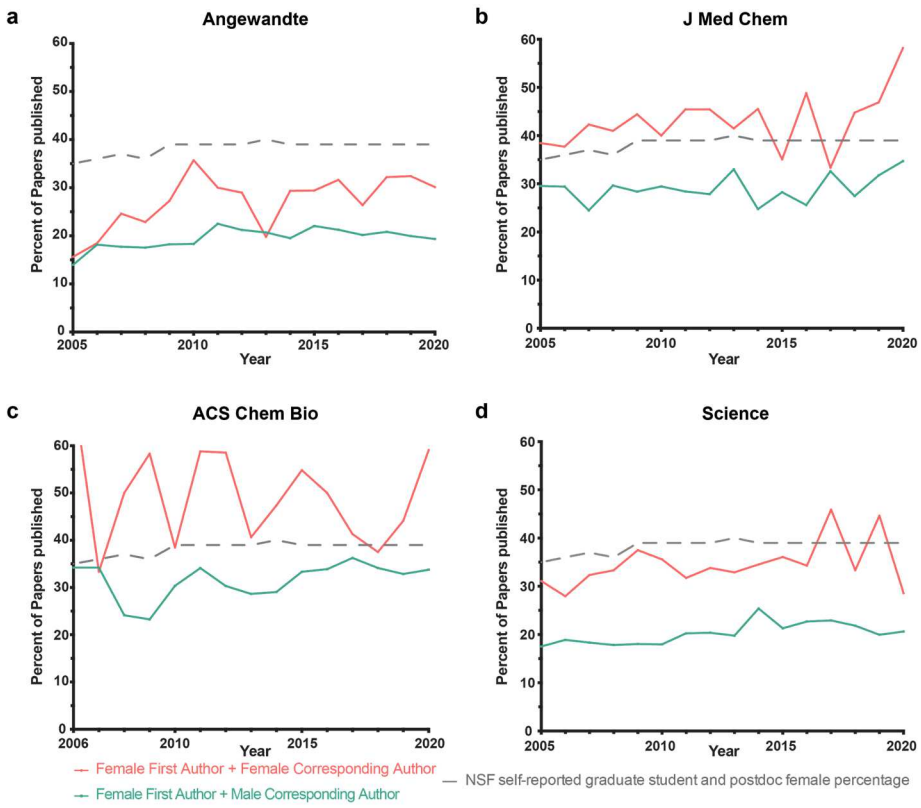


Figure 5.4: Are female corresponding authors more likely to have female first authors? Dashed lines represent the NSF self-reported gender representation of US graduate students and postdoctoral fellows. a) Female first author percentage in Angewandte with either a male (green) or female (red) corresponding author. B) Female first author percentage in J Med Chem with either a male (green) or female (red) corresponding author. C) Female first author percentage in ACS Chemical Biology with either a male (green) or female (red) corresponding author. D) Female first author percentage in Science with either a male (green) or female (red) corresponding author. *Name-predicted gender–see methods section for details.

In addition to authorship data, we were interested in gender representation on journal editorial boards. We manually curated names and countries of current editorial board members from journal websites and used the same protocol with Gender API to estimate gender percentages. Due to our focus on the Chemistry field and the wide breadth of topics covered in Nature, Cell, and Science, the editorial boards and editorial teams for these journals were not analyzed. Editorial board gender ratios ranged from ~1:9 to 1:1 female: male. (**Figure 5.5**). Tetrahedron and Tetrahedron Letters have journal advisory boards of over sixty members in which less than 10% of the members are

women. The editorial teams of Nature Chemistry and Nature Chemical Biology are 50% and 60% female, but it should be noted that these journals differ from the others in that they have a small team of full-time editors. In addition, the roles of editorial boards (broadly defined) in the screening process of manuscripts for any given journal may vary.

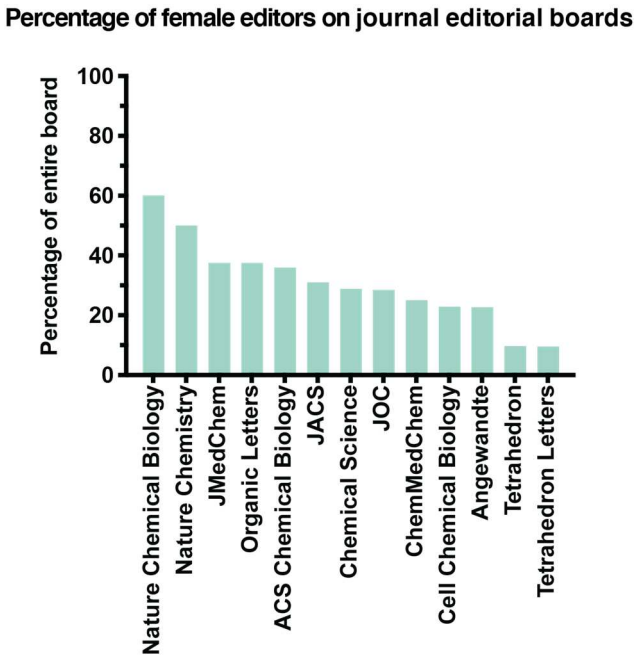


Figure 5.5: Rank-ordered female percentages of journal editorial boards and editorial teams in 2020. Data collected from respective journal websites and analyzed with Gender API.

5.4 Discussion

We recognized some salient trends in our dataset**.** Chemistry journals have a substantial disparity in gender representation in authorship, especially for corresponding authors. First author representation is more likely to be equitable when the corresponding author is a woman, and with few exceptions, the gender ratios of editorial boards heavily favor men. Furthermore, over the past 16 years, the representation of women as authors has not changed substantially. At the current rate,

the representation of women as first or corresponding authors will not equal that of men in any journal for over 50 years.

The goal of our analyses is to provide data on authorship that we hope will foster further discussion about how to improve gender representation in chemistry. Our analysis was focused on gender, but we recognize that science is replete with other primary and secondary diversity imbalances (e.g., racial, socio-economic). Open discussion on the topic of diversity is a necessary step towards a more equitable publishing environment and we hope our data allows informed conversations about how we as a community can improve representation of women in chemistry.

5.5 Materials and Methods

On June 10th 2020 a web-scraper was used to search PubMed for all papers published in a journal each year since 2005. For journals that were created post-2005, their date of creation was used as the first year. The scraper pulled the first name of both the first and corresponding authors. The country of the corresponding author was extracted, and this was assumed to be the same for the first author. The country information improves the accuracy of the gender predicting algorithm. Any papers that have a solo author were discarded from the data. The data sets containing a first name and country were uploaded to Gender API to predict the gender of the author based on their first name. Gender API outputs a predicted gender, an accuracy score out of 100 and data size sampled. The accuracy score is a probability of the gender being correctly assigned, thus any data points below 95% confidence or small sample size (below 100, with confidence below 100) were removed from the data set. This resulted with an increasing set of high-quality data points for 279,525 authors (**Figure 5.S1**). A histogram demonstrating the quality of the data set is found in supporting information (**Figure 5.S2**). Any years for a journal with less than 25 quality data points was removed from any time-series analysis. Each author was then assigned a binary gender and population totals calculated.

NSF self-reported data were used to calculate the US graduate student and postdoc gender percentages.[14] These data sets are available for 2005-2014; post-2014 values were extrapolated from the publicly available data.

ACS Division of Organic Chemistry has a curated list of 518 current US-based faculty who work on synthetic organic chemistry.[13] Similar to above, a web-scraper was used to extract first-names, and Gender API was used to assign a name-predicted gender.

Journal websites were used to obtain their editorial boards. These lists were manually extracted, and Gender API was used to assign a name-predicted gender

Our approach has some inherent limitations. While the use of a gender API is an efficient high throughput method to attempt to address gender prediction, gender is not conclusively discernible by name, so this approach is inherently reductionist[16]. Furthermore, due to the limitations of first name-based gender APIs, we could only account for two genders: male and female. We desired to include other genders in the analysis; however, we could not find a method to include them that was high throughput. We are aware that only considering male and female as genders can be considered an act of erasure of other genders, and as more data becomes available, we will include nonbinary genders to remedy this. Our approach does not account for gender transition, differing gender identity and gender expression, or nonbinary identities, which can give rise to several additional dynamics that are not captured with our data or discussion.

Additionally, we used the affiliated country of the last (corresponding) author to inform gender assignment by the API, which uses country to improve gender prediction for common names. There are likely cases in which that affiliated country will not reflect the origins of the names of the first or last authors. Additionally, there are likely cases where authors intentionally use a gender-neutral or male first name or publish only under their initials to avoid publishing bias. Finally, there are many cases where there are multiple first or multiple corresponding authors, and, more rarely, cases where the last author is not a corresponding author. Our approach does not account for these cases. Gender neutral names were also eliminated from the data sets in order to minimize uncertainty.

5.6 Supplemental Figures

5.6.1 Analysis of data set

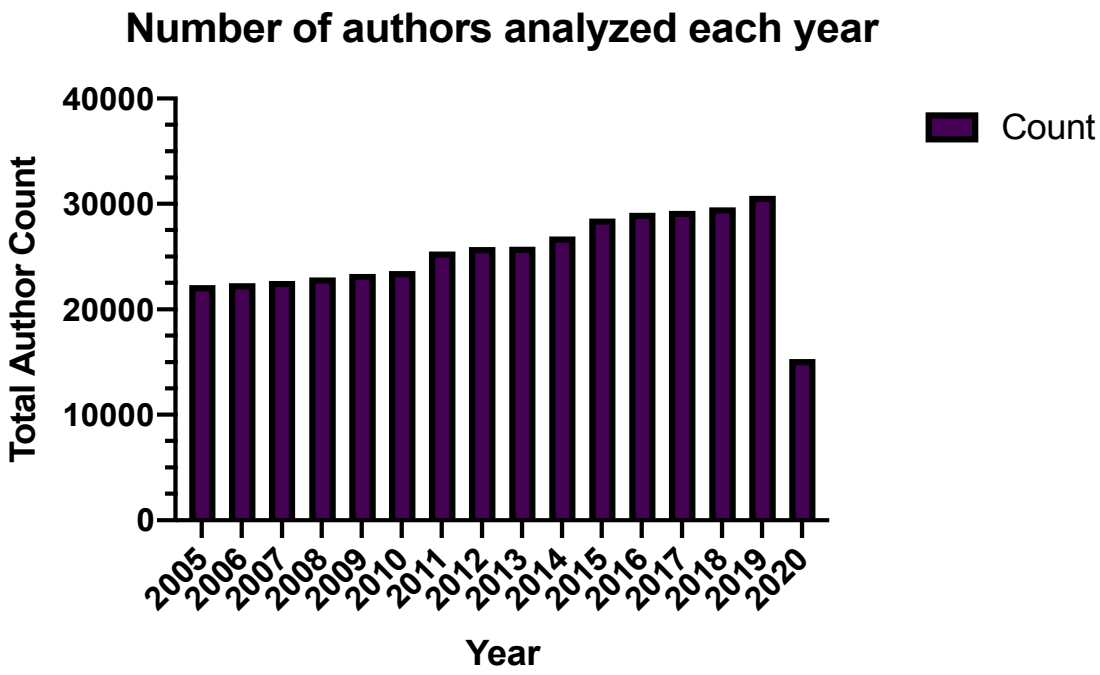


Figure 5.S1: Total number of authors' names analyzed each year.

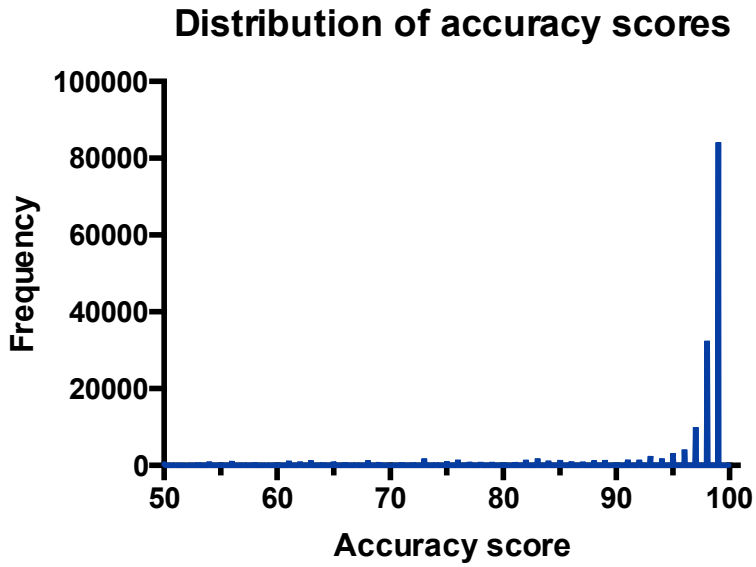


Figure 5.S2: Distribution of all accuracy scores for all types of authors

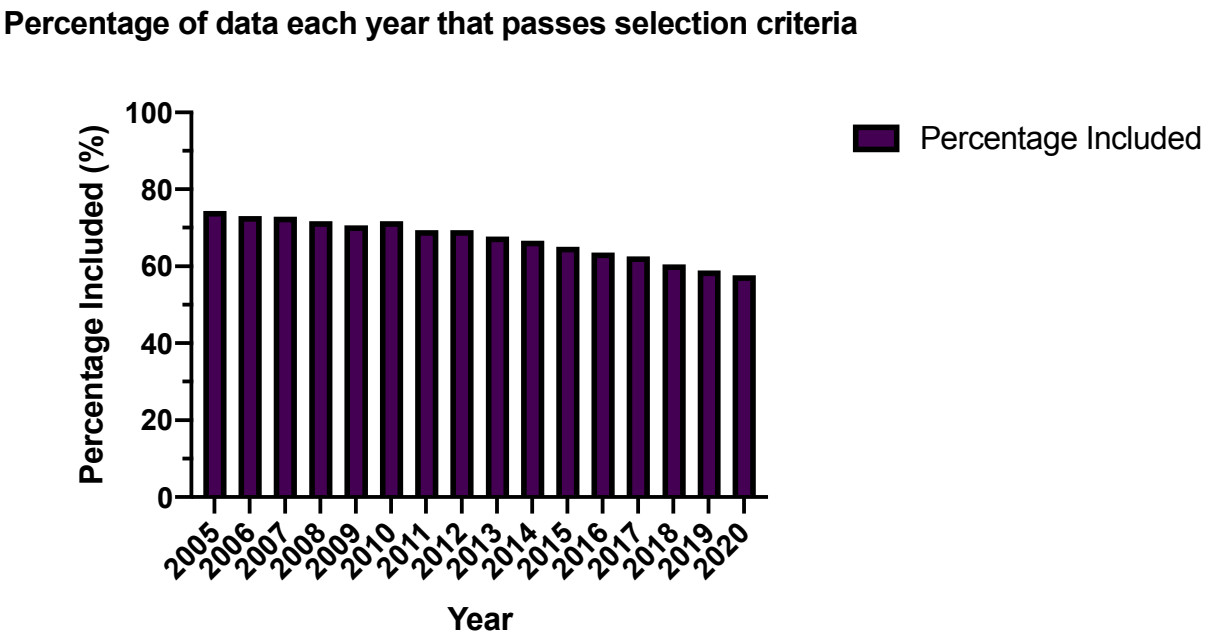


Figure 5.S3: Percentage of author names that passes the inclusion criteria each year.

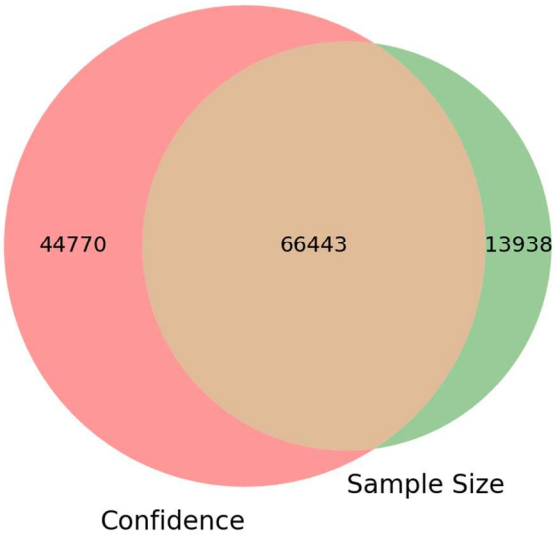


Figure 5.S4: Venn diagram showing the reason for exclusion with total author counts for each category. Confidence-only portion indicates likely gender-neutral names.

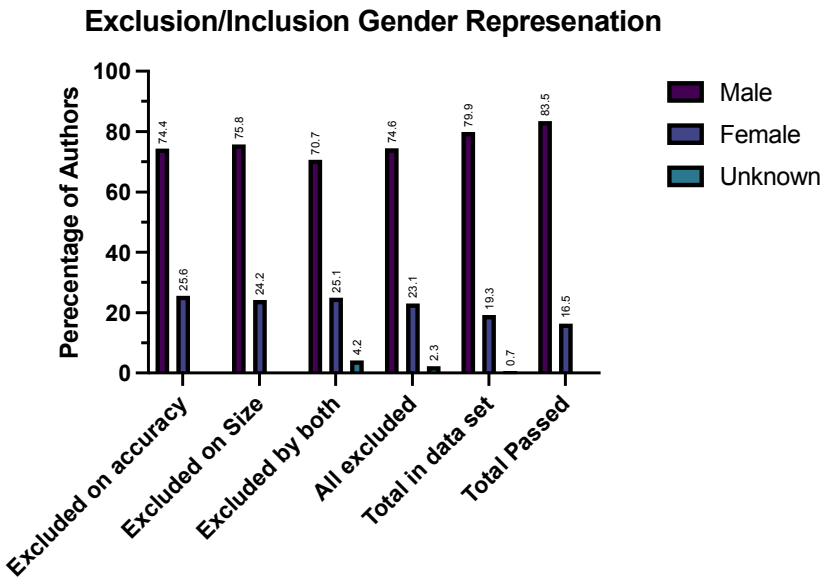


Figure 5.S5: Gender breakdown for each exclusion group, total data set, and passing author names. Gender for this figure is assigned by >50% accuracy score. 'Unknown' indicates an accuracy score of 50%.

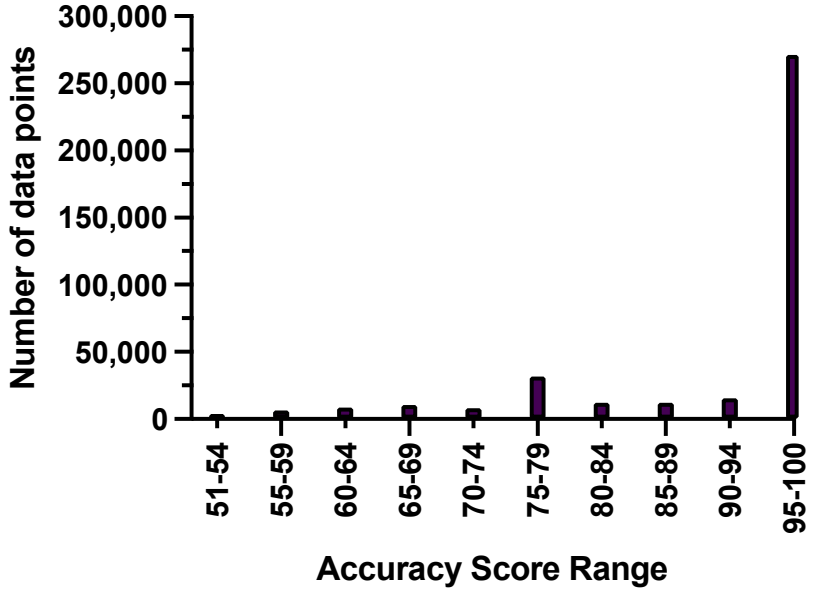


Figure 5.S6: Number of data points for each 5% increment of accuracy scores

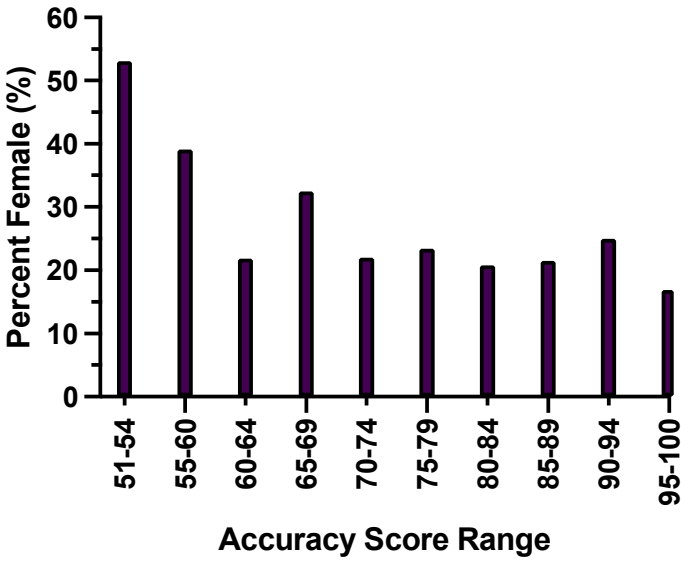


Figure 5.S7: Percentage of female names for each 5% increment of accuracy scores. Gender for this figure is assigned by >50% accuracy score. Lower accuracy scores represent lower confidence in assigned gender. Note that, with the way that Gender API assigns accuracy scores, the lowest score is 50% (meaning that it is a 50/50 chance the author is male or female); after that point, it inverts to being a probability of being the other gender. At lower accuracy scores (more uncertainty), the expected trend to ~50% of authors being female and 50% being male, because there is no confidence that either gender is correct.

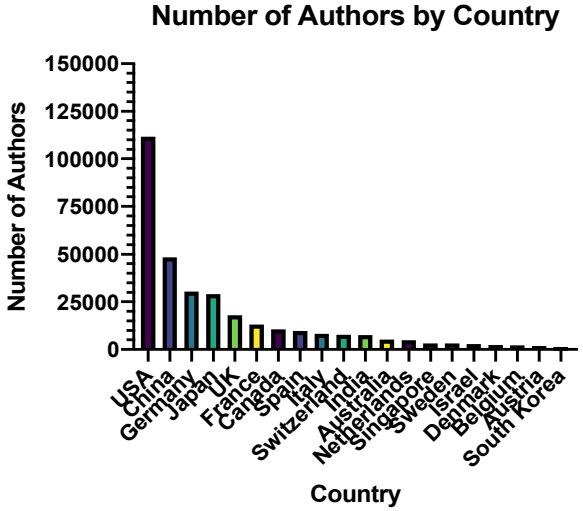


Figure 5.S8: Number of authors from the top 20 countries represented in the total dataset.

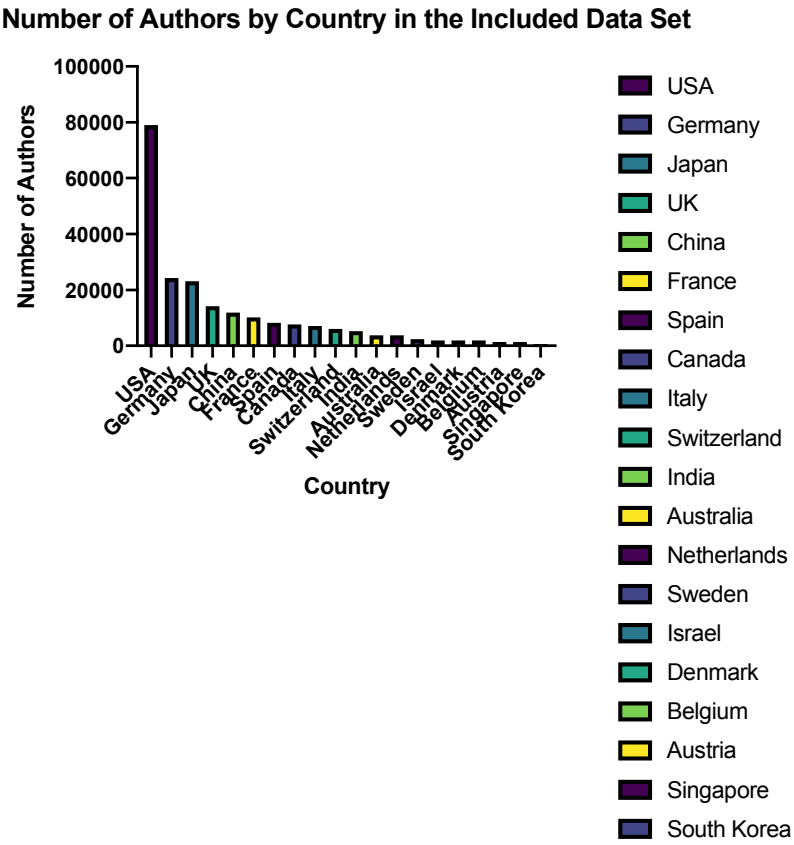


Figure 5.S9: Number of authors from the top 20 countries represented in the included dataset.

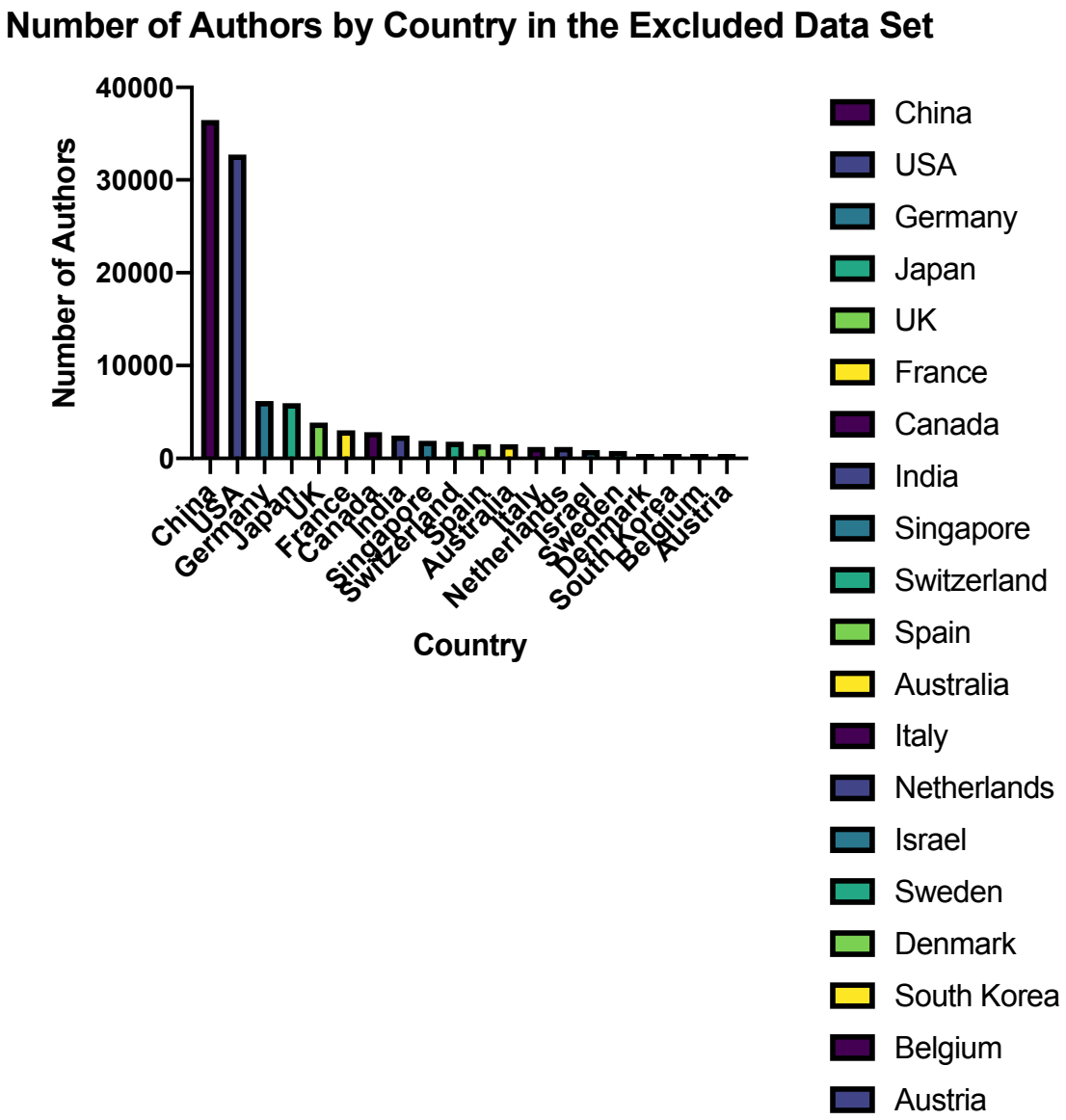


Figure 5.S10: Number of authors excluded from the top 20 most represented countries in the total dataset.

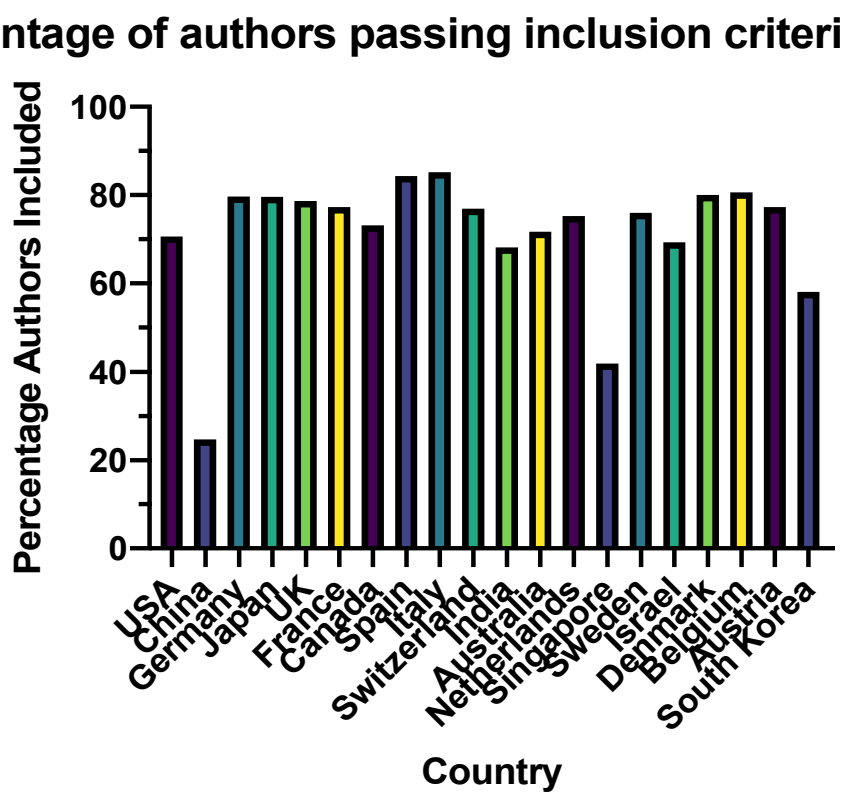


Figure 5.S11: Percentage of authors passing the inclusion criteria for each of the top 20 represented countries in the dataset. Note the low value for China, which results from low accuracy scores due to a higher frequency of gender-neutral names when translated

5.6.2 Statistical Analysis

Table 5.S1: Results of a linear regression analysis performed on time-series data for each journal for corresponding authors. P-value is a representation of the significance of each slope being non-zero.

Corresponding Authors					
Journal	**equation**	**slope**	**p value?**	**Significant ?**	**Date of 50%**
ACS Chemical Biology	Y = 0.03413*X - 49.94	0.03413	0.8741	No	-
Angewandte	Y = 0.1555*X - 305.1	0.1555	0.0011	Yes	2283

Corresponding Authors					
Journal	**equation**	**slope**	**p value?**	**Significant ?**	**Date of 50%**
Cell	Y = 0.03588*X - 56.23	0.03588	0.6138	No	-
Cell Chemical Biology	Y = 0.3727*X - 737.5	0.3727	0.7332	No	-
Chemical Science	Y = 0.4200*X - 836.0	0.42	0.0495	Yes	2110
ChemMedChem	Y = 0.3239*X - 634.4	0.3239	0.0555	No	-
JACS	Y = 0.1949*X - 382.7	0.1949	0.0002	Yes	2220
JMedChem	Y = 0.5051*X - 1000	0.5051	0.0007	Yes	2080
JOC	Y = -0.05688*X + 125.0	-0.05688	0.3692	No	-
Nature	Y = 0.3161*X - 623.0	0.3151	0.0011	Yes	2129
Nature Chemical Biology	Y = 0.2276*X - 444.4	0.2276	0.0655	No	-
Nature Chemistry	Y = 0.2976*X - 591.5	0.2976	0.2784	No	-
Organic Letters	Y = -0.04531*X + 99.26	-0.04531	0.3381	No	-
Science	Y = 0.2436*X - 475.3	0.2436	0.0473	Yes	2156

Corresponding Authors					
Journal	**equation**	**slope**	**p value?**	**Significant ?**	**Date of 50%**
Tetrahedron	Y = -0.3113*X + 632.9	-0.3113	0.3111	No	-
Tetrahedron Letters	Y = 0.9256*X - 1853	0.9256	0.0721	No	-

Table 5.S2: Results of a linear regression performed on time-series data for each journal for first authors. P-value is a representation of the significance of each slope being non-zero.

First Authors					
Journal	**equation**	**slope**	**p value?**	**Significant?**	**Date of 50%**
ACS Chemical Biology	Y = 0.3628*X - 696.0	0.3628	0.0855	No	-
Angewandte	Y = 0.3617*X - 708.4	0.3617	0.0001	Yes	2097
Cell	Y = 0.1019*X - 174.8	0.1019	0.5512	No	-
Cell Chemical Biology	Y = -1.750*X + 3568	-1.75	0.4196	No	-
Chemical Science	Y = -0.09096*X + 208.3	-0.09096	0.8535	No	-
ChemMedChem	Y = 0.1904*X - 346.3	0.1904	0.511	No	-
JACS	Y = 0.1228*X - 225.6	0.1228	0.0603	No	-
JMedChem	Y = 0.2816*X - 536.5	0.2816	0.0938	No	-
JOC	Y = 0.2778*X - 536.5	0.2778	0.0038	Yes	2111

Nature	Y = 0.2162*X - 412.7	0.2162	0.0124	Yes	2140
Nature Chemical Biology	Y = -0.01678*X + 62.53	-0.01678	0.9426	No	-
Nature Chemistry	Y = 0.4066*X - 800.2	0.4066	0.1689	No	-
Organic Letters	Y = -0.003601*X + 25.97	-0.003601	0.965	No	-
Science	Y = 0.3148*X - 611.4	0.3148	0.011	Yes	2101
Tetrahedron	Y = -0.2584*X + 538.7	-0.2584	0.6314	No	-
Tetrahedron Letters	Y = 1.320*X - 2635	1.32	0.075	No	-

Table 5.S3: Results of a binomial test for statistical significance of if a female corresponding author is more likely than their male counterparts to have a female first author on their papers. Expected value used is the percentage of female first authors for each respective journal. P-value represents the probability that a data point is below or equal to the expected value.

Journal	Corresponding author gender	Percentage of first authors female	p value	Significant?
ACS Chemical Biology	Female	48.10%	$1.3x10^{-8}$	Yes
ACS Chemical Biology	Male	32.10%	0.983	No
Angewandte	Female	27.90%	$9.9x10^{-14}$	Yes
Angewandte	Male	19.70%	0.629	No
Cell	Female	39.10%	$1.0x10^{-7}$	Yes
Cell	Male	29.10%	0.914	No

Cell Chemical Biology	Female	50.80%	0.0167	Yes
Cell Chemical Biology	Male	38.10%	0.531	No
Chemical Science	Female	34.60%	0.000179	Yes
Chemical Science	Male	25.50%	0.27	No
ChemMedChem	Female	51.40%	$2.8x10^{-8}$	Yes
ChemMedChem	Male	36.00%	0.843	No
JACS	Female	30.50%	$<1.0x10^{-14}$	Yes
JACS	Male	21.20%	0.737	No
JMedChem	Female	43.10%	$<1.0x10^{-14}$	Yes
JMedChem	Male	28.90%	0.962	No
JOC	Female	30.40%	$3.7x10^{-11}$	Yes
JOC	Male	22.50%	0.526	No
Nature	Female	28.30%	$1.5x10^{-6}$	Yes
Nature	Male	22.20%	0.678	No
Nature Chemical Biology	Female	38.20%	0.00121	Yes
Nature Chemical Biology	Male	27.80%	0.779	No
Nature Chemistry	Female	21.60%	0.212	No
Nature Chemistry	Male	19.80%	0.21	No
Organic Letters	Female	27.60%	$8.9x10^{-12}$	Yes
Organic Letters	Male	18.90%	0.302	No
Science	Female	34.50%	$<1.0x10^{-14}$	Yes

Science	Male	20.00%	0.999	No

5.6.3 Corresponding author gender percentages since 2005 for each journal

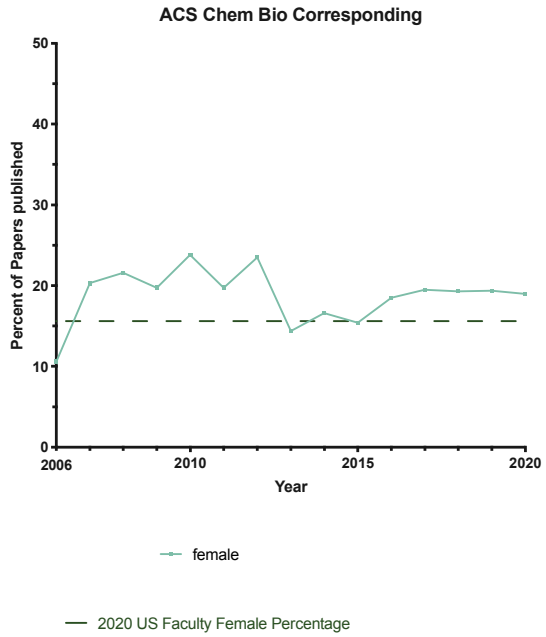


Figure 5.S12: Percentage of female corresponding authors in ACS Chemical Biology since 2006

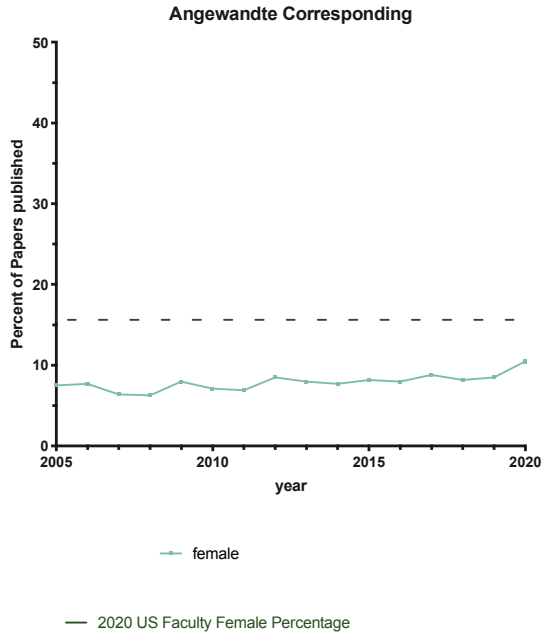


Figure 5.S13: Percentage of female corresponding authors in Angewandte since 2005

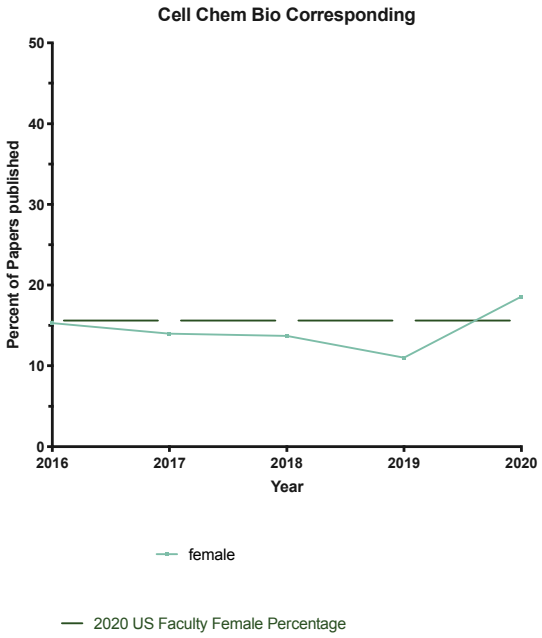


Figure 5.S14: Percentage of female corresponding authors in Cell Chemical Biology since 2016

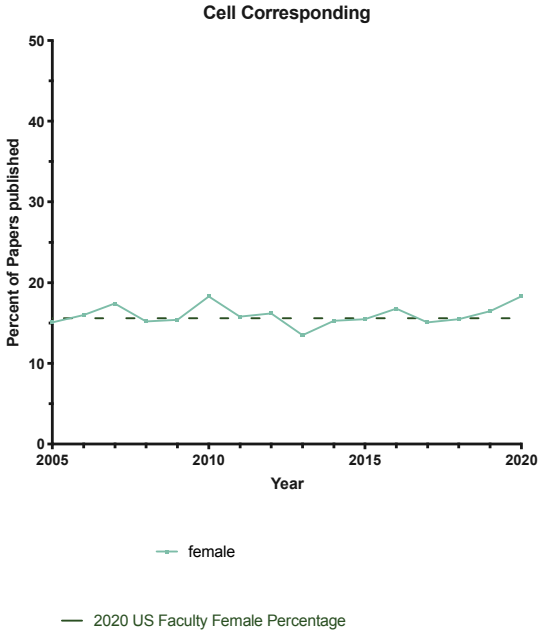


Figure 5.S15: Percentage of female corresponding authors in Cell since 2005

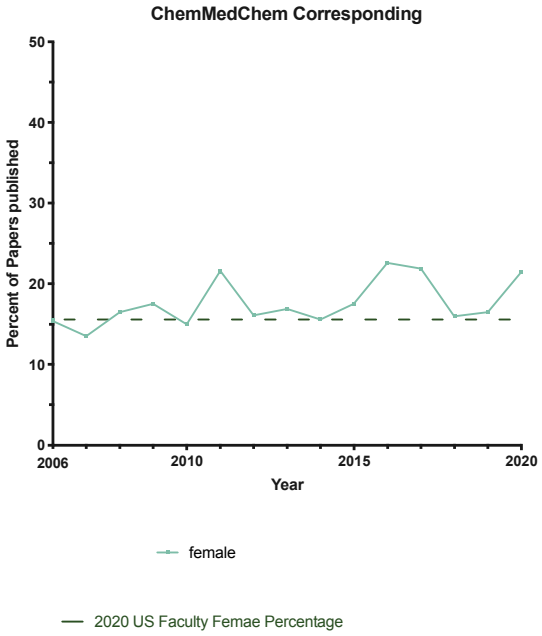


Figure 5.S16: Percentage of female corresponding authors in ChemMedChem since 2006

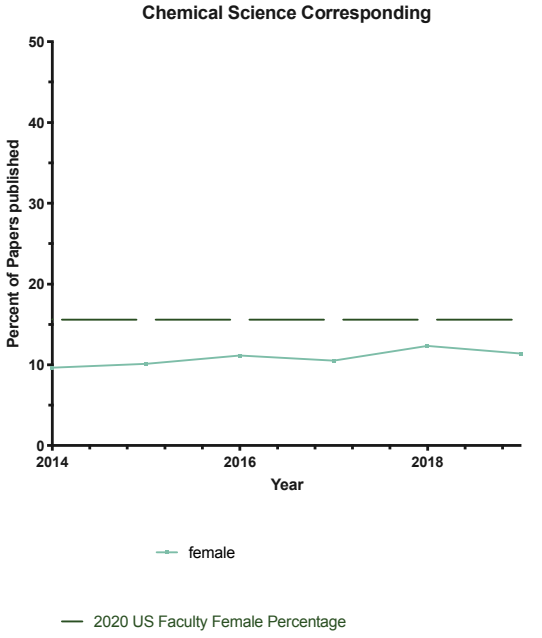


Figure 5.S17: Percentage of female corresponding authors in Chemical Science since 2014

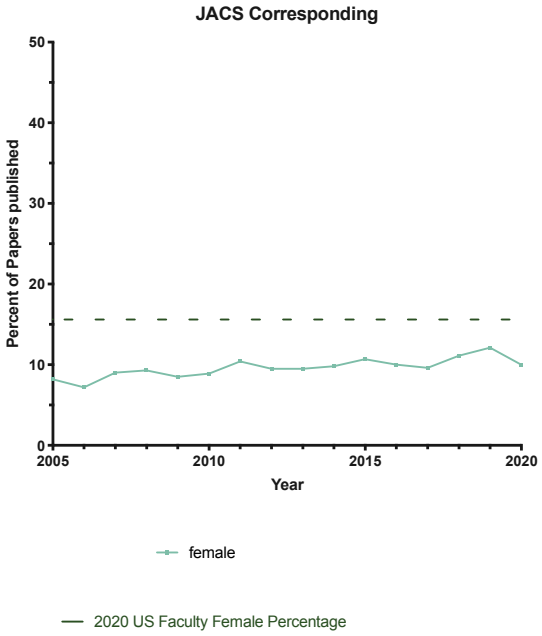


Figure 5.S18: Percentage of female corresponding authors in JACS since 2005

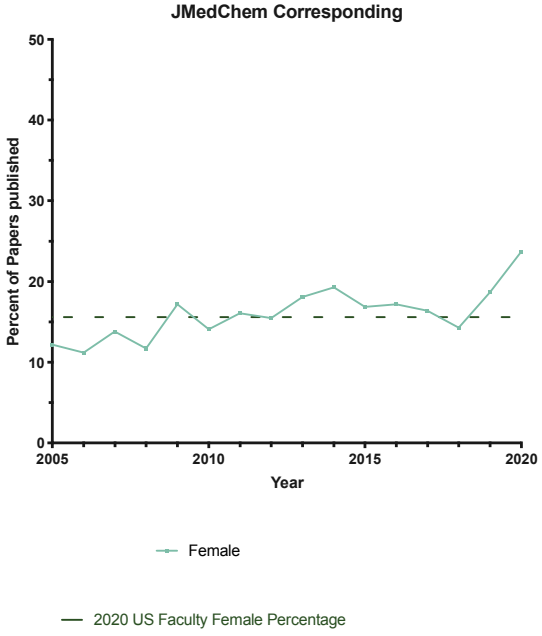


Figure 5.S19: Percentage of female corresponding authors in JMedChem since 2005

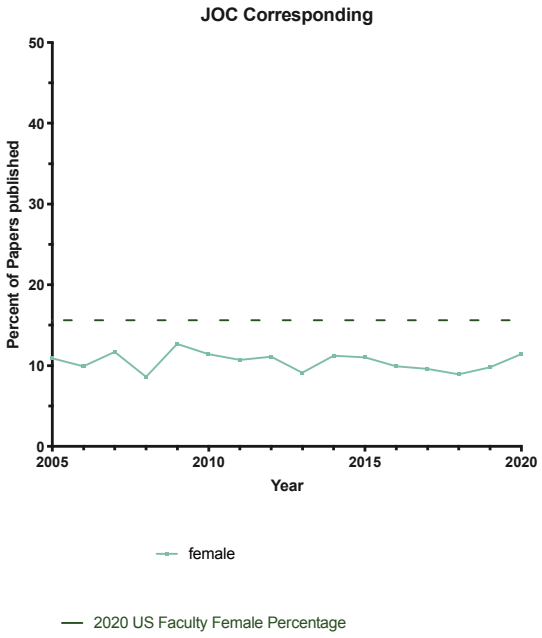


Figure 5.S20: Percentage of female corresponding authors in JOC since 2005

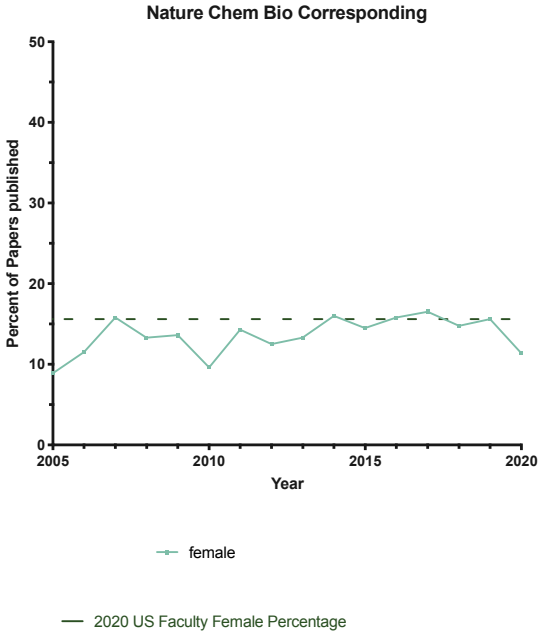


Figure 5.S21: Percentage of female corresponding authors in Nature Chemical Biology since 2005

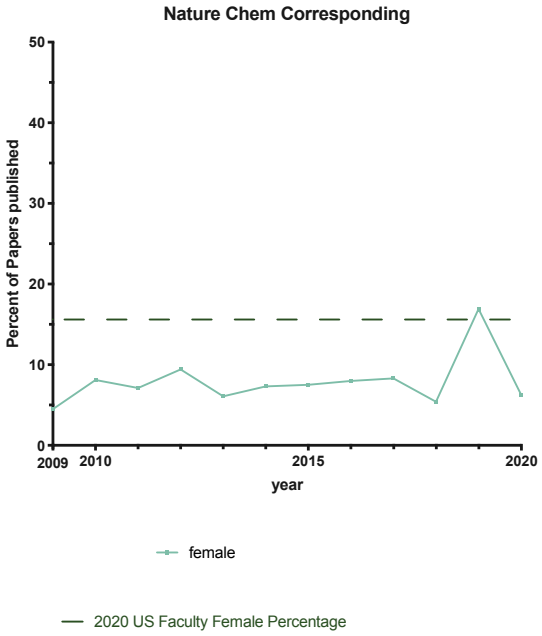


Figure 5.S22: Percentage of female corresponding authors in Nature Chemistry since 2005

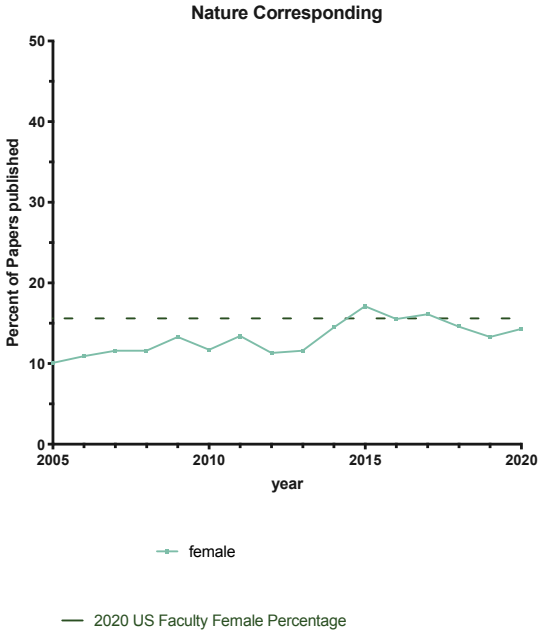


Figure 5.S23: Percentage of female corresponding authors in Nature since 2005

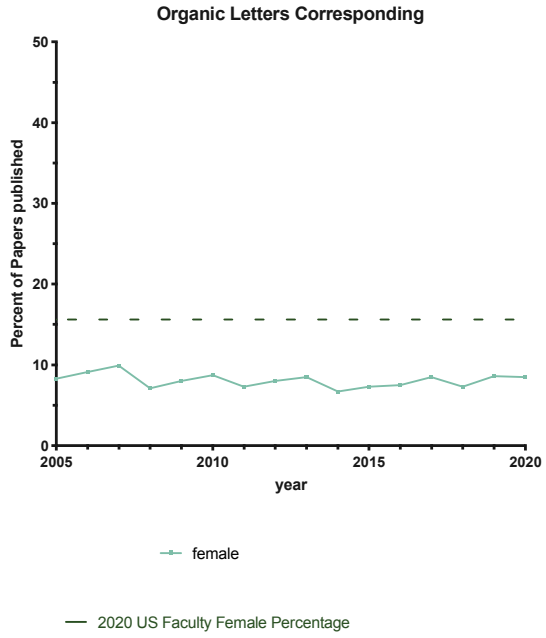


Figure 5.S24: Percentage of female corresponding authors in Organic Letters since 2005

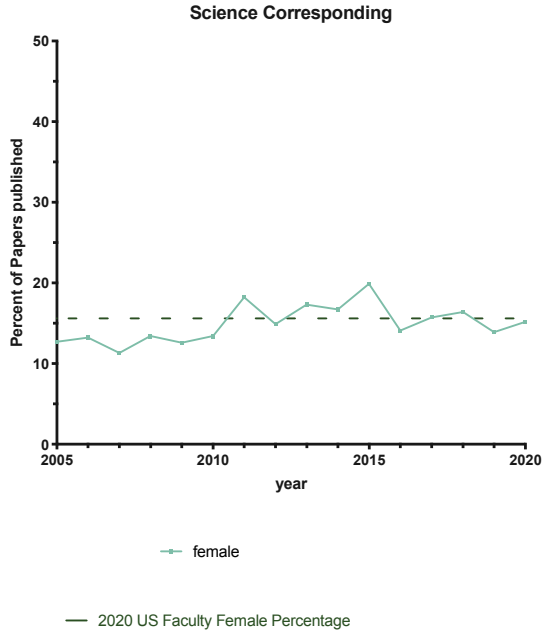


Figure 5.S25: Percentage of female corresponding authors in Science since 2005

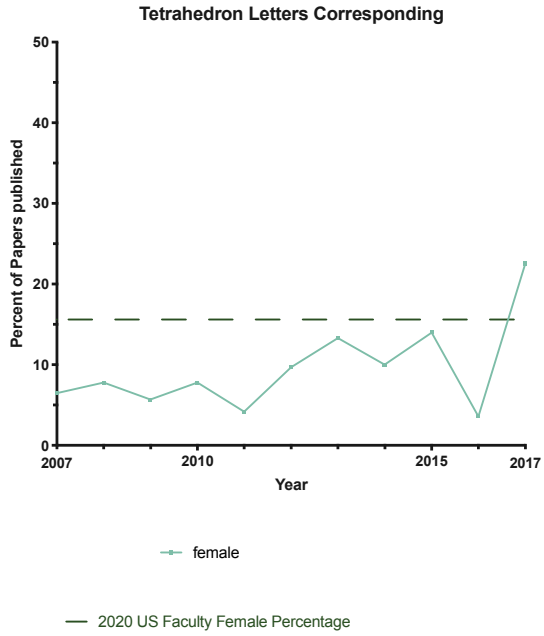


Figure 5.S26: Percentage of female corresponding authors in Tetrahedron Letters since 2007

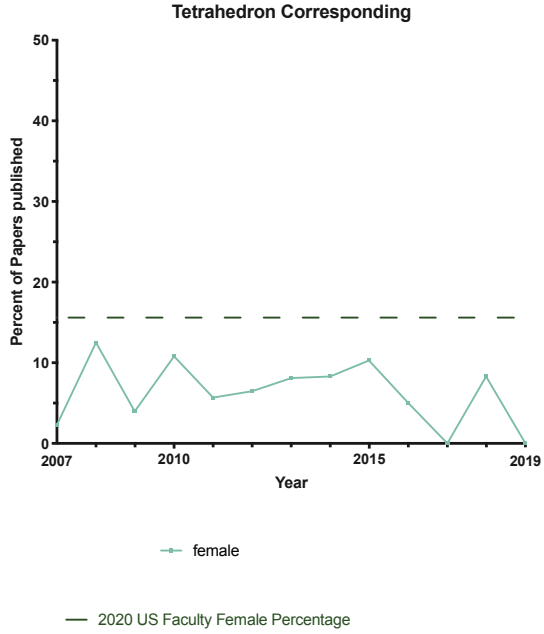


Figure 5.S27: Percentage of female corresponding authors in Tetrahedron since 2007

5.6.4 First author gender percentages since 2005 for each journal

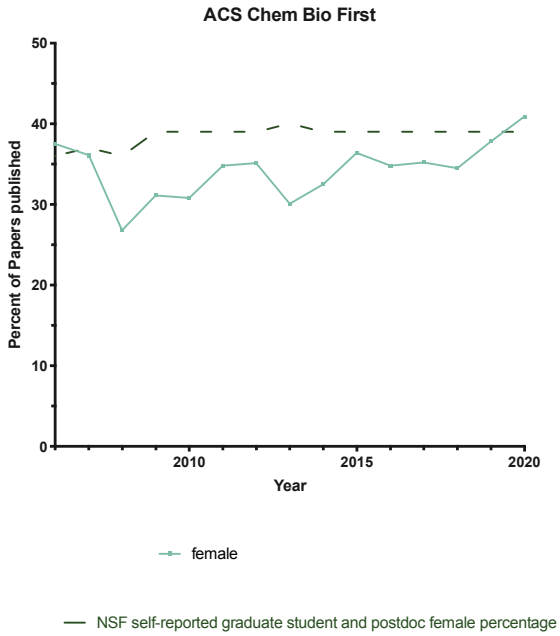


Figure 5.S28: Percentage of female first authors in ACS Chemical Biology since 2006

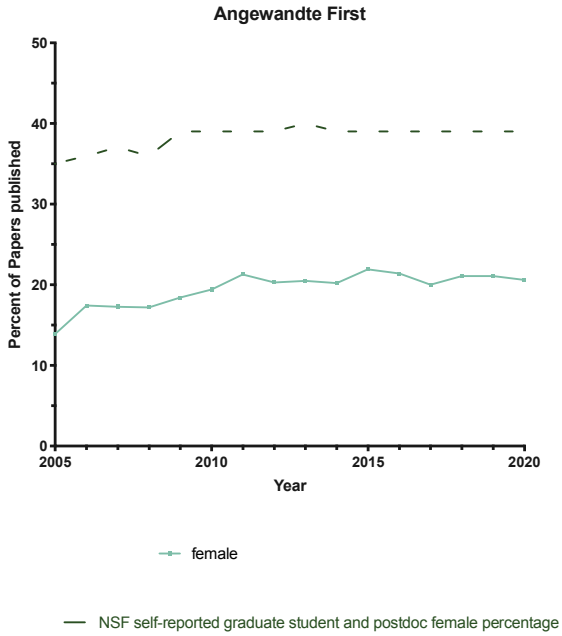


Figure 5.S29: Percentage of female first authors in Angewandte since 2005

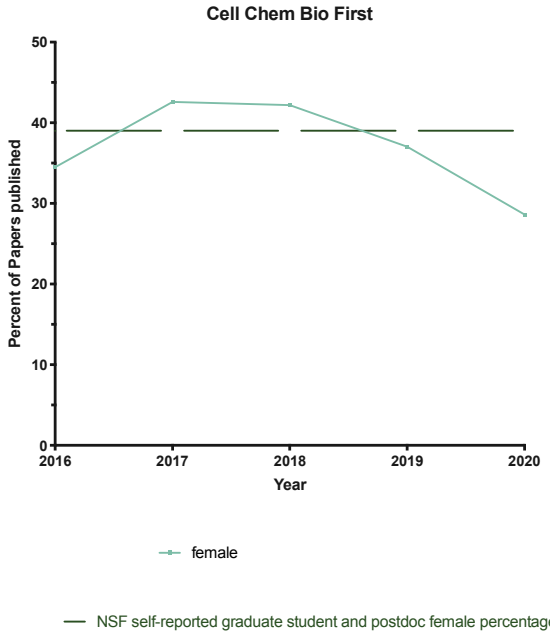


Figure 5.S30: Percentage of female first authors in Cell Chemical Biology since 2016

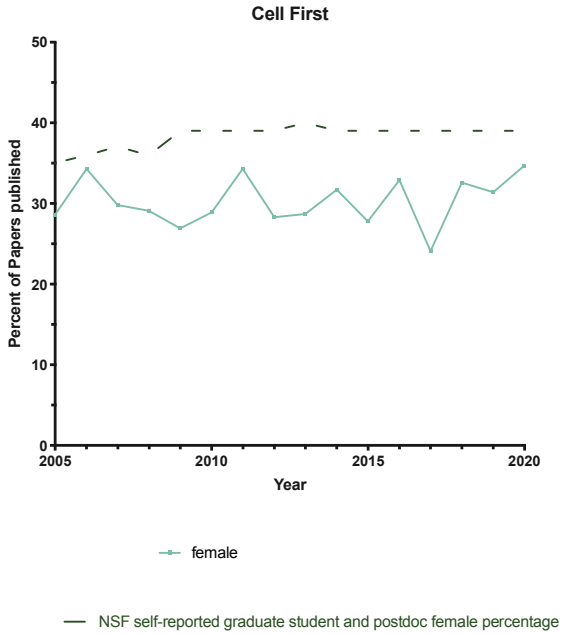


Figure 5.S31: Percentage of female first authors in Cell since 2005

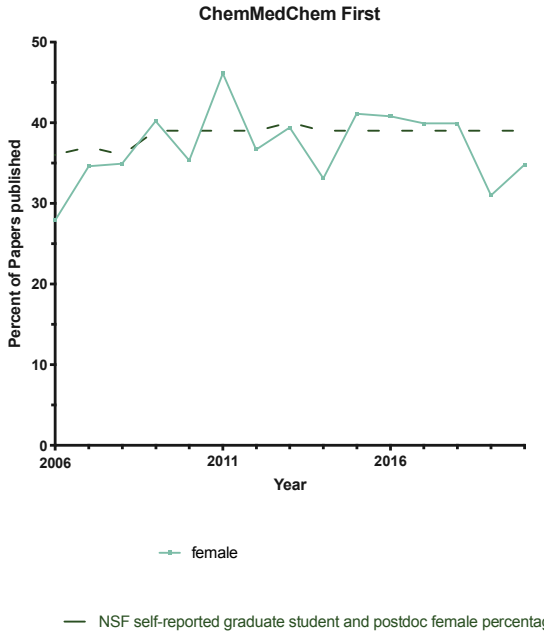


Figure 5.S32: Percentage of female first authors in ChemMedChem since 2006

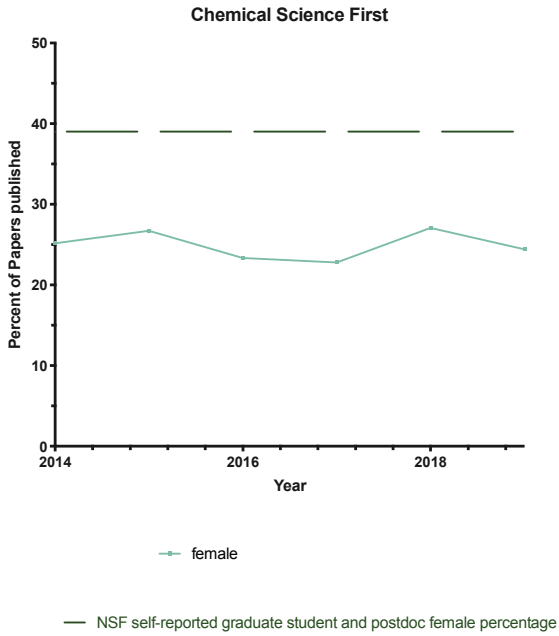


Figure 5.S33: Percentage of female first authors in Chemical Science since 2014

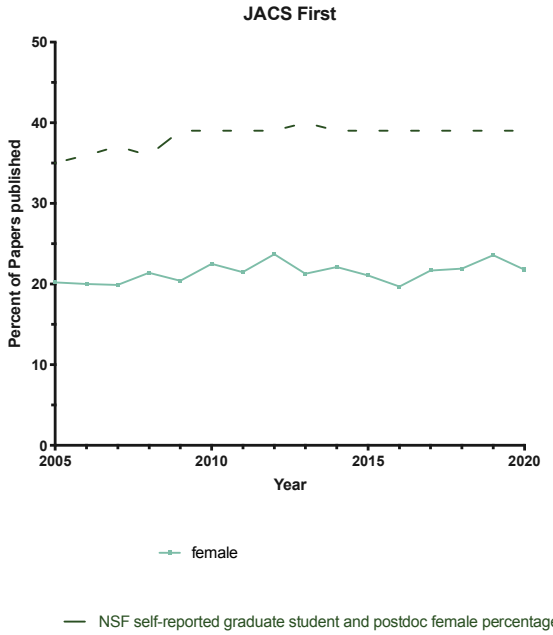


Figure 5.S34: Percentage of female first authors in JACS since 2005

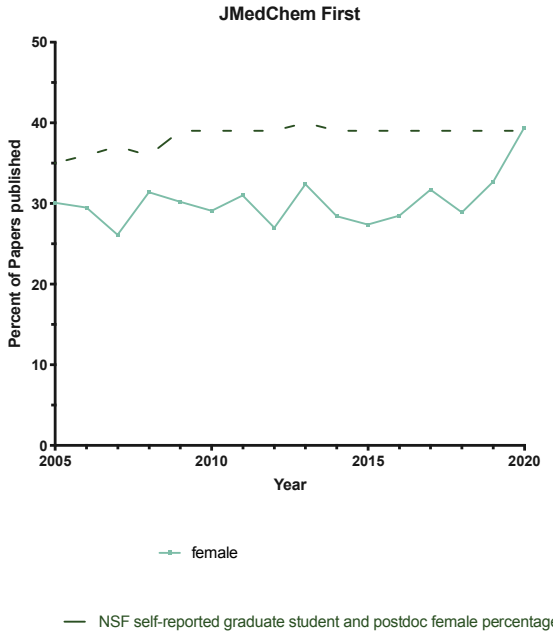


Figure 5.S35: Percentage of female first authors in JMedChem since 2005

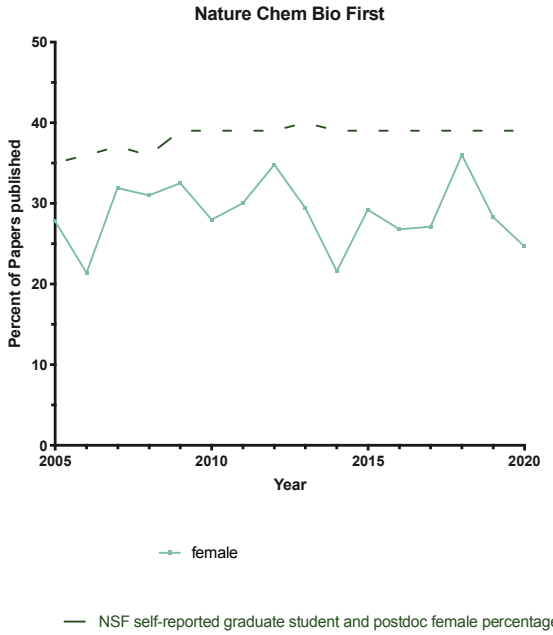


Figure 5.S36: Percentage of female first authors in Nature Chemical Biology since 2005

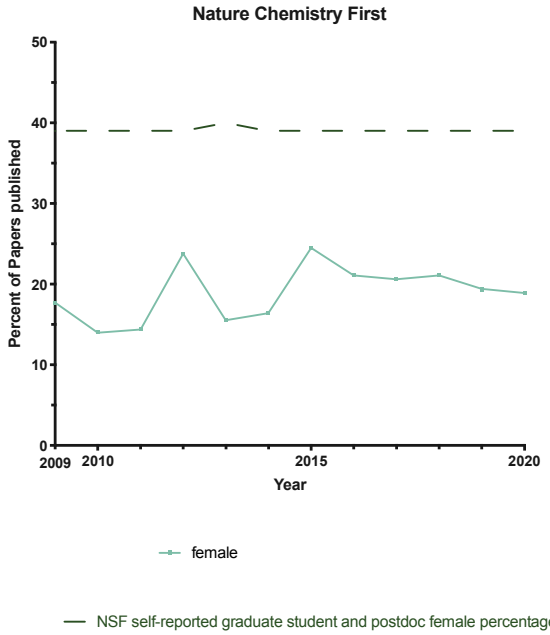


Figure 5.S37: Percentage of female first authors in Nature Chemistry since 2009

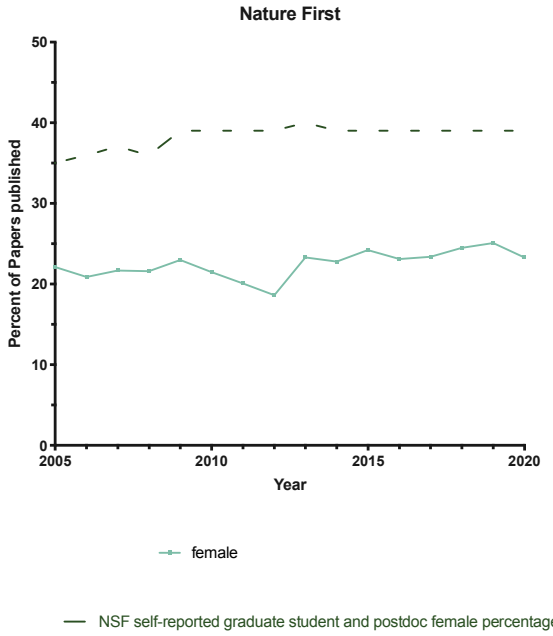


Figure 5.S38: Percentage of female first authors in Nature since 2005

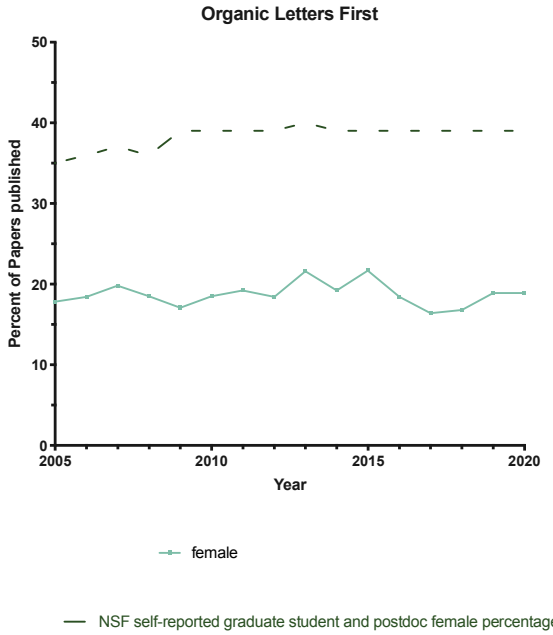


Figure 5.S39: Percentage of female first authors in Organic Letters since 2005

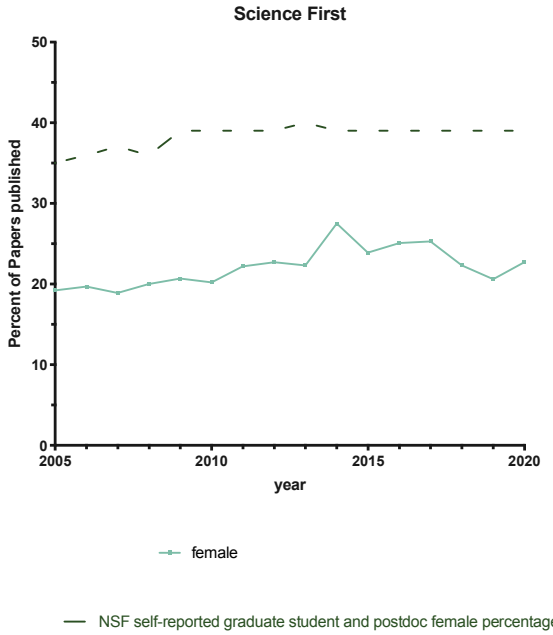


Figure 5.S40: Percentage of female first authors in Science since 2005

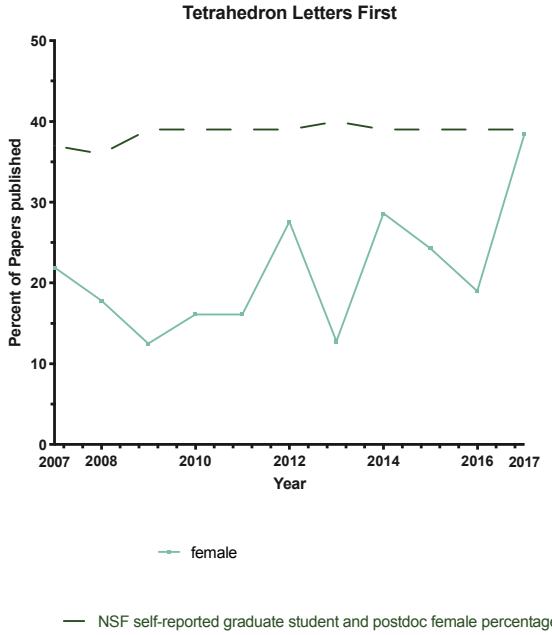


Figure 5.S41: Percentage of female first authors in Tetrahedron Letters since 2007

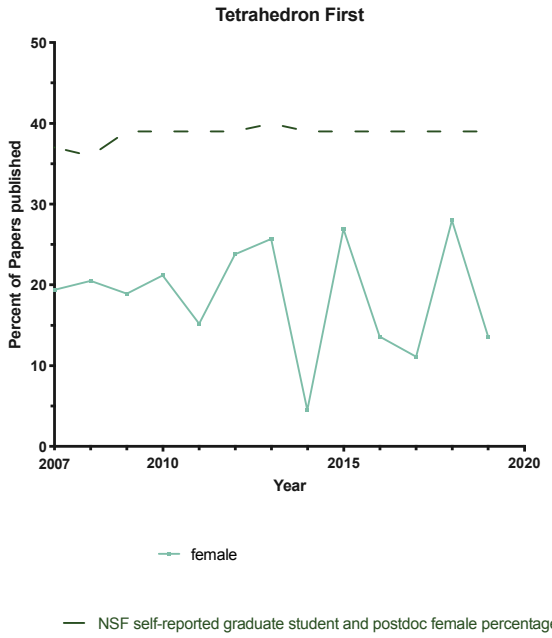


Figure 5.S42: Percentage of female first authors in Tetradedron since 2007

5.6.5 Percentage of female first authors with either female or male corresponding authors for each journal since 2005

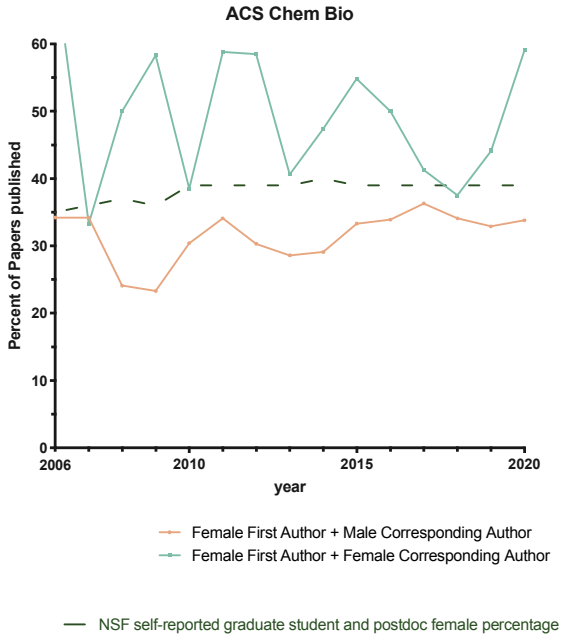


Figure 5.S43: Percentage of female first authors with either a female (green) or male (orange) corresponding author in ACS Chemical Biology since 2006

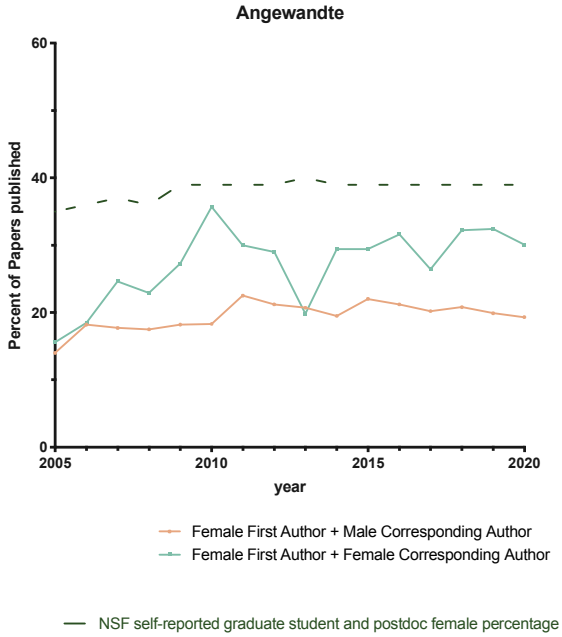


Figure 5.S44: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Angewandte since 2005

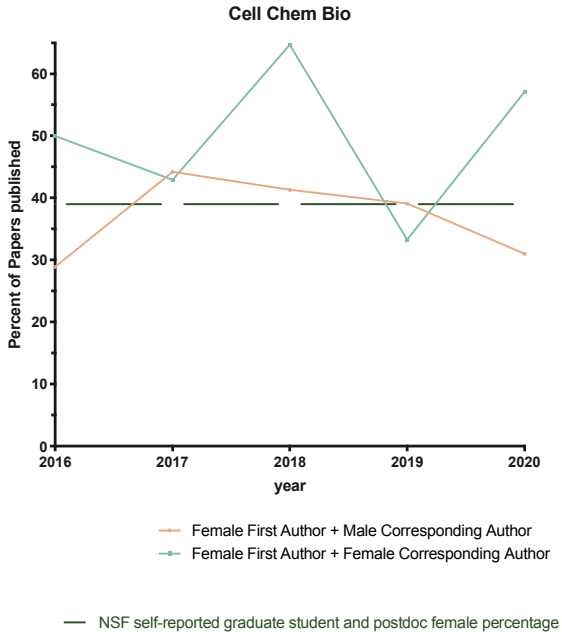


Figure 5.S45: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Cell Chemical Biology since 2016

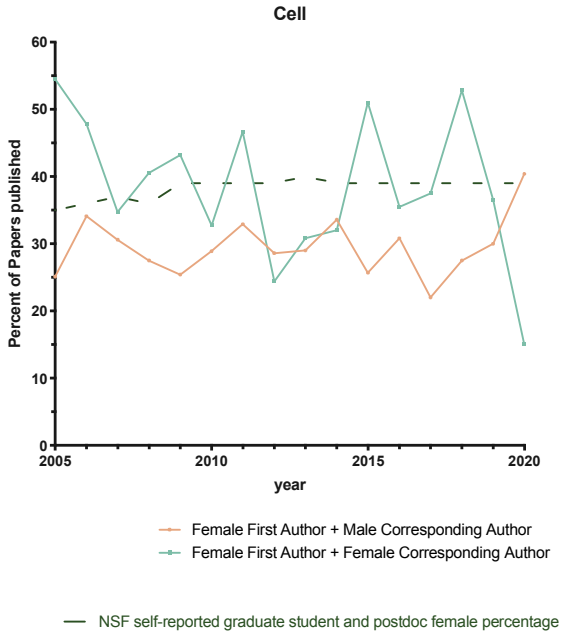


Figure 5.S46: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Cell since 2005

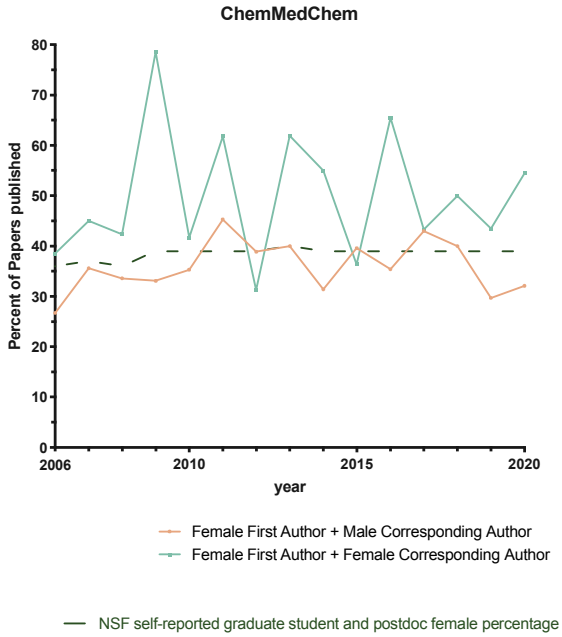


Figure 5.S47: Percentage of female first authors with either a female (green) or male (orange) corresponding author in ChemMedChem since 2006

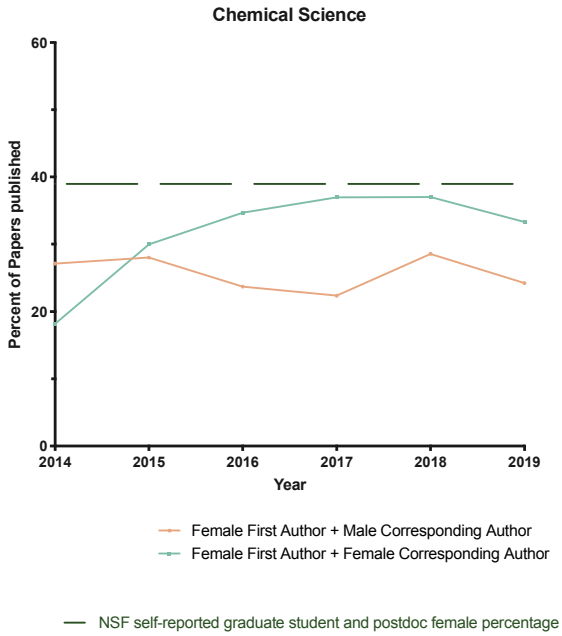


Figure 5.S48: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Chemical Science since 2014

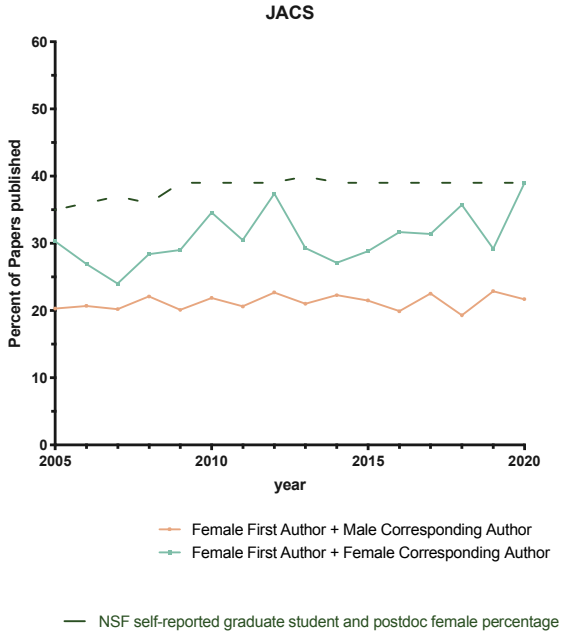


Figure 5.S49: Percentage of female first authors with either a female (green) or male (orange) corresponding author in JACS since 2005

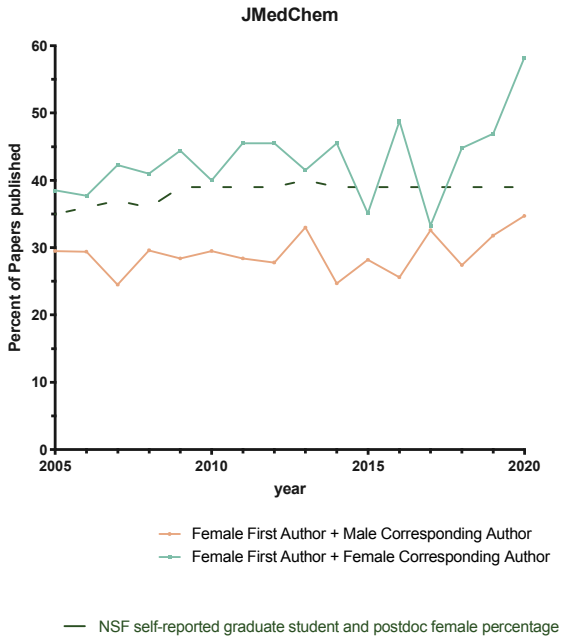


Figure 5.S50: Percentage of female first authors with either a female (green) or male (orange) corresponding author in JMedChem since 2005

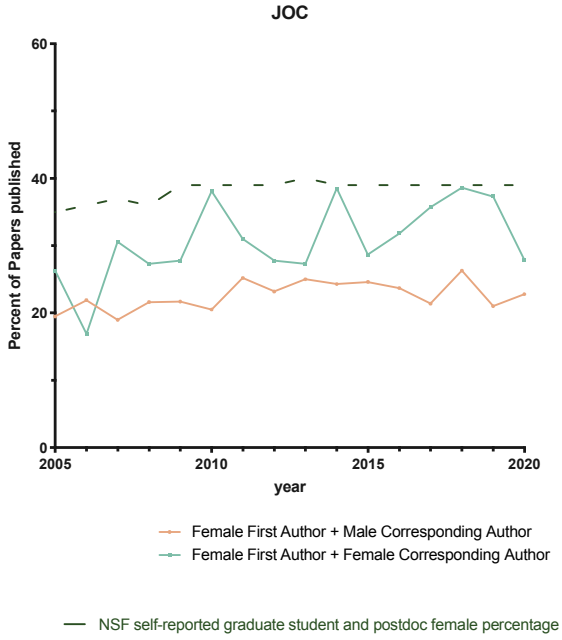


Figure 5.S51: Percentage of female first authors with either a female (green) or male (orange) corresponding author in JOC since 2005

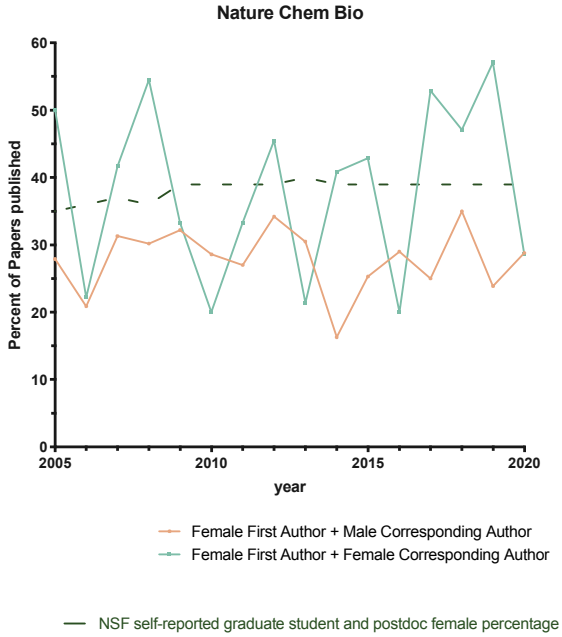


Figure 5.S52: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Nature Chemical Biology since 2005

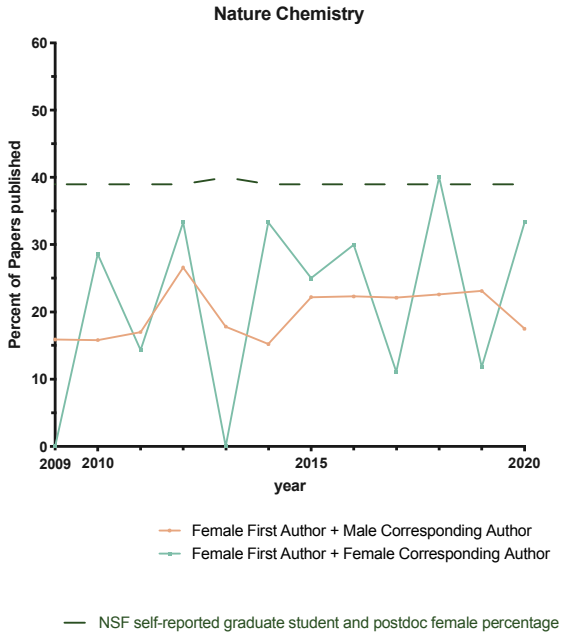


Figure 5.S53: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Nature Chemistry since 2009

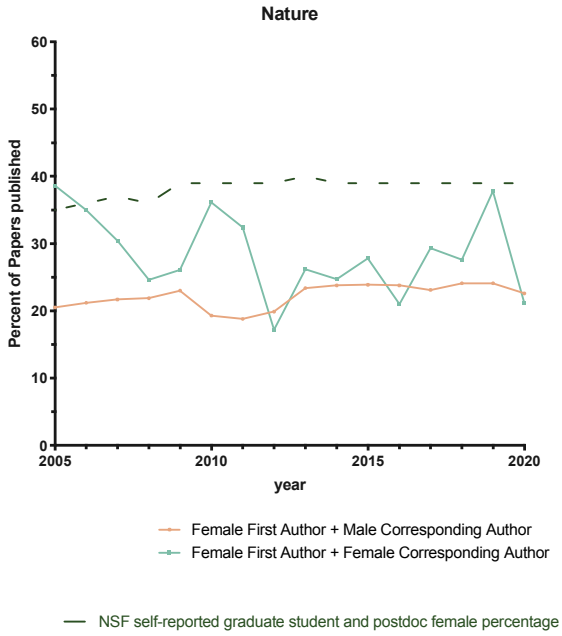


Figure 5.S54: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Nature since 2005

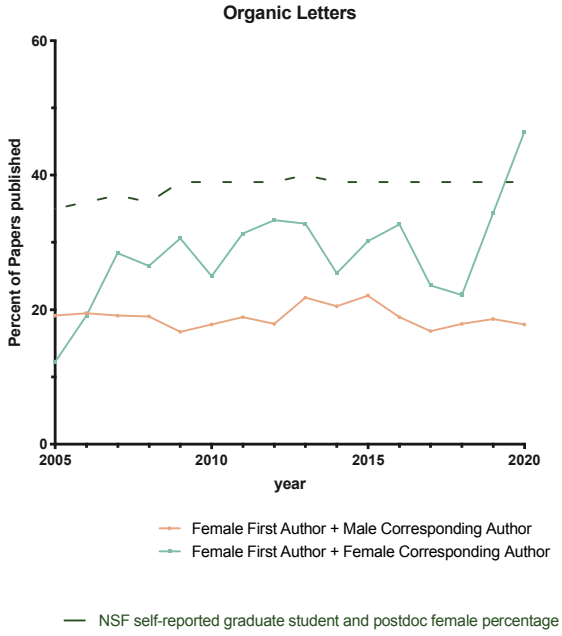


Figure 5.S55: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Organic Letters since 2005

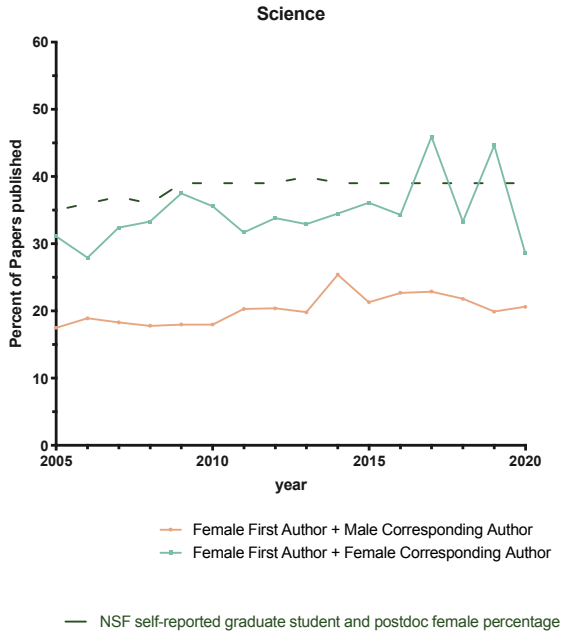


Figure 5.S56: Percentage of female first authors with either a female (green) or male (orange) corresponding author in Science since 2005

5.6.6 Number of data points (total and passing) each year for each journal

Table 5.S4: Number of authors analyzed and passing inclusion criteria each year for different journals.

Year	Journal	Count	Count Passing Inclusion Criteria
2005	ACS Chemical Biology	0	0
2006	ACS Chemical Biology	158	123
2007	ACS Chemical Biology	162	131
2008	ACS Chemical Biology	124	99
2009	ACS Chemical Biology	170	138
2010	ACS Chemical Biology	234	185
2011	ACS Chemical Biology	376	289
2012	ACS Chemical Biology	574	439

Year	Journal	Count	Count Passing Inclusion Criteria
2013	ACS Chemical Biology	756	578
2014	ACS Chemical Biology	800	589
2015	ACS Chemical Biology	742	540
2016	ACS Chemical Biology	854	618
2017	ACS Chemical Biology	804	576
2018	ACS Chemical Biology	774	580
2019	ACS Chemical Biology	696	498
2020	ACS Chemical Biology	404	289
2005	Angewandte	2744	2124
2006	Angewandte	2868	2201
2007	Angewandte	3194	2437
2008	Angewandte	3432	2553
2009	Angewandte	3188	2169
2010	Angewandte	3584	2589
2011	Angewandte	4580	3030
2012	Angewandte	4914	3217
2013	Angewandte	5034	3257
2014	Angewandte	5576	3631
2015	Angewandte	5968	3814
2016	Angewandte	5900	3549
2017	Angewandte	5908	3593
2018	Angewandte	6098	3547

Year	Journal	Count	Count Passing Inclusion Criteria
2019	Angewandte	6672	3607
2020	Angewandte	4492	2307
2005	Cell	739	586
2006	Cell	894	703
2007	Cell	924	726
2008	Cell	874	699
2009	Cell	940	736
2010	Cell	818	642
2011	Cell	852	665
2012	Cell	1006	764
2013	Cell	1044	765
2014	Cell	1086	849
2015	Cell	1140	870
2016	Cell	1194	926
2017	Cell	998	761
2018	Cell	1190	876
2019	Cell	1070	817
2020	Cell	466	330
2005	Cell Chemical Biology	0	0
2006	Cell Chemical Biology	0	0
2007	Cell Chemical Biology	0	0
2008	Cell Chemical Biology	0	0

Year	Journal	Count	Count Passing Inclusion Criteria
2009	Cell Chemical Biology	0	0
2010	Cell Chemical Biology	0	0
2011	Cell Chemical Biology	0	0
2012	Cell Chemical Biology	0	0
2013	Cell Chemical Biology	0	0
2014	Cell Chemical Biology	0	0
2015	Cell Chemical Biology	0	0
2016	Cell Chemical Biology	324	250
2017	Cell Chemical Biology	368	278
2018	Cell Chemical Biology	382	297
2019	Cell Chemical Biology	388	285
2020	Cell Chemical Biology	156	107
2005	ChemicalScience	0	0
2006	ChemicalScience	0	0
2007	ChemicalScience	0	0
2008	ChemicalScience	0	0
2009	ChemicalScience	0	0
2010	ChemicalScience	0	0
2011	ChemicalScience	0	0
2012	ChemicalScience	0	0
2013	ChemicalScience	0	0
2014	ChemicalScience	546	352

Year	Journal	Count	Count Passing Inclusion Criteria
2015	ChemicalScience	2352	1355
2016	ChemicalScience	2336	1284
2017	ChemicalScience	2336	1306
2018	ChemicalScience	2252	1247
2019	ChemicalScience	2100	1257
2020	ChemicalScience	0	0
2005	ChemMedChem	0	0
2006	ChemMedChem	260	226
2007	ChemMedChem	336	292
2008	ChemMedChem	388	319
2009	ChemMedChem	394	341
2010	ChemMedChem	392	329
2011	ChemMedChem	390	331
2012	ChemMedChem	396	308
2013	ChemMedChem	388	293
2014	ChemMedChem	384	293
2015	ChemMedChem	390	315
2016	ChemMedChem	384	291
2017	ChemMedChem	392	312
2018	ChemMedChem	394	312
2019	ChemMedChem	394	321
2020	ChemMedChem	314	258

Year	Journal	Count	Count Passing Inclusion Criteria
2005	EJOC	0	0
2006	EJOC	0	0
2007	EJOC	0	0
2008	EJOC	20	16
2009	EJOC	30	23
2010	EJOC	22	13
2011	EJOC	22	15
2012	EJOC	30	21
2013	EJOC	42	30
2014	EJOC	38	25
2015	EJOC	48	37
2016	EJOC	40	31
2017	EJOC	38	33
2018	EJOC	50	34
2019	EJOC	54	42
2020	EJOC	10	9
2005	JACS	6746	4904
2006	JACS	6470	4717
2007	JACS	6178	4525
2008	JACS	6454	4532
2009	JACS	6646	4523
2010	JACS	6584	4700

Year	Journal	Count	Count Passing Inclusion Criteria
2011	JACS	6734	4783
2012	JACS	6424	4490
2013	JACS	5880	3973
2014	JACS	5454	3543
2015	JACS	4980	3098
2016	JACS	5030	3139
2017	JACS	5524	3422
2018	JACS	5178	3090
2019	JACS	5310	3106
2020	JACS	2732	1576
2005	JMedChem	1722	1326
2006	JMedChem	1704	1368
2007	JMedChem	1484	1110
2008	JMedChem	1706	1300
2009	JMedChem	1640	1261
2010	JMedChem	1650	1228
2011	JMedChem	1510	1132
2012	JMedChem	1850	1368
2013	JMedChem	1694	1255
2014	JMedChem	1674	1288
2015	JMedChem	1554	1150
2016	JMedChem	1638	1224

Year	Journal	Count	Count Passing Inclusion Criteria
2017	JMedChem	1518	1052
2018	JMedChem	1620	1127
2019	JMedChem	1624	1066
2020	JMedChem	908	578
2005	JOC	3176	2392
2006	JOC	2894	2045
2007	JOC	3264	2209
2008	JOC	2924	2056
2009	JOC	2876	1975
2010	JOC	2610	1822
2011	JOC	2682	1918
2012	JOC	2774	1920
2013	JOC	2902	1967
2014	JOC	2826	1749
2015	JOC	2946	1885
2016	JOC	2944	1791
2017	JOC	3098	1855
2018	JOC	3324	1921
2019	JOC	3472	2007
2020	JOC	1704	976
2005	Nature	2214	1559
2006	Nature	1962	1358

Year	Journal	Count	Count Passing Inclusion Criteria
2007	Nature	1728	1221
2008	Nature	1882	1364
2009	Nature	1770	1280
2010	Nature	1804	1308
2011	Nature	1844	1307
2012	Nature	1872	1355
2013	Nature	2042	1453
2014	Nature	2428	1702
2015	Nature	2420	1722
2016	Nature	2240	1592
2017	Nature	2130	1534
2018	Nature	1906	1259
2019	Nature	2108	1411
2020	Nature	1028	640
2005	Nature Chemical Biology	130	113
2006	Nature Chemical Biology	212	182
2007	Nature Chemical Biology	236	190
2008	Nature Chemical Biology	228	181
2009	Nature Chemical Biology	326	264
2010	Nature Chemical Biology	298	251
2011	Nature Chemical Biology	330	261
2012	Nature Chemical Biology	264	211

Year	Journal	Count	Count Passing Inclusion Criteria
2013	Nature Chemical Biology	312	247
2014	Nature Chemical Biology	368	290
2015	Nature Chemical Biology	348	280
2016	Nature Chemical Biology	434	330
2017	Nature Chemical Biology	450	333
2018	Nature Chemical Biology	366	288
2019	Nature Chemical Biology	390	274
2020	Nature Chemical Biology	248	169
2005	Nature Chemistry	0	0
2006	Nature Chemistry	0	0
2007	Nature Chemistry	0	0
2008	Nature Chemistry	0	0
2009	Nature Chemistry	234	174
2010	Nature Chemistry	342	266
2011	Nature Chemistry	318	244
2012	Nature Chemistry	286	195
2013	Nature Chemistry	332	252
2014	Nature Chemistry	364	274
2015	Nature Chemistry	340	247
2016	Nature Chemistry	468	355
2017	Nature Chemistry	428	300
2018	Nature Chemistry	406	267

Year	Journal	Count	Count Passing Inclusion Criteria
2019	Nature Chemistry	374	238
2020	Nature Chemistry	174	122
2005	Organic Letters	2930	2199
2006	Organic Letters	3038	2043
2007	Organic Letters	2990	2118
2008	Organic Letters	2798	1796
2009	Organic Letters	2856	1942
2010	Organic Letters	3210	2171
2011	Organic Letters	3624	2166
2012	Organic Letters	3372	2180
2013	Organic Letters	3456	2061
2014	Organic Letters	3436	2045
2015	Organic Letters	3314	1891
2016	Organic Letters	3500	1903
2017	Organic Letters	3606	1879
2018	Organic Letters	3932	1946
2019	Organic Letters	4390	2081
2020	Organic Letters	1904	963
2005	Science	1914	1393
2006	Science	1962	1407
2007	Science	1952	1409
2008	Science	1930	1390

Year	Journal	Count	Count Passing Inclusion Criteria
2009	Science	1978	1445
2010	Science	1838	1269
2011	Science	1900	1319
2012	Science	1904	1321
2013	Science	1874	1275
2014	Science	1780	1196
2015	Science	1782	1209
2016	Science	1754	1185
2017	Science	1622	1042
2018	Science	1678	1068
2019	Science	1666	1087
2020	Science	748	497
2005	Tetrahedron Letters	12	9
2006	Tetrahedron Letters	40	29
2007	Tetrahedron Letters	112	63
2008	Tetrahedron Letters	138	99
2009	Tetrahedron Letters	186	131
2010	Tetrahedron Letters	170	124
2011	Tetrahedron Letters	180	130
2012	Tetrahedron Letters	156	124
2013	Tetrahedron Letters	120	89
2014	Tetrahedron Letters	86	60

Year	Journal	Count	Count Passing Inclusion Criteria
2015	Tetrahedron Letters	222	163
2016	Tetrahedron Letters	68	51
2017	Tetrahedron Letters	80	61
2018	Tetrahedron Letters	58	41
2019	Tetrahedron Letters	30	24
2020	Tetrahedron Letters	12	11
2005	Tetrahedron	6	4
2006	Tetrahedron	20	14
2007	Tetrahedron	108	80
2008	Tetrahedron	116	90
2009	Tetrahedron	130	104
2010	Tetrahedron	96	71
2011	Tetrahedron	134	106
2012	Tetrahedron	78	54
2013	Tetrahedron	88	75
2014	Tetrahedron	68	48
2015	Tetrahedron	78	56
2016	Tetrahedron	56	44
2017	Tetrahedron	46	23
2018	Tetrahedron	68	52
2019	Tetrahedron	60	45
2020	Tetrahedron	0	0

5.7 References

1. Joshi, A. By Whom and When Is Women's Expertise Recognized? The Interactive Effects of Gender and Education in Science and Engineering Teams. *Administrative Science Quarterly* **59**, 202–239 (2014).

2. Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. & Malone, T. W. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* **330**, 686–688 (2010).

3. Sugimoto, C. R., Ahn, Y.-Y., Smith, E., Macaluso, B. & Larivière, V. Factors affecting sex-related reporting in medical research: a cross-disciplinary bibliometric analysis. *The Lancet* **393**, 550–559 (2019).

4. Reisman, S. E., Sarpong, R., Sigman, M. S. & Yoon, T. P. Organic Chemistry: A Call to Action for Diversity and Inclusion. *J. Org. Chem.* **85**, 10287–10292 (2020).

5. Wilson-Kennedy, Z. S., Payton-Stewart, F. & Winfield, L. L. Toward Intentional Diversity, Equity, and Respect in Chemistry Research and Practice. *J. Chem. Educ.* **97**, 2041–2044 (2020).

6. Fadeyi, O. O., Heffern, M. C., Johnson, S. S. & Townsend, S. D. What Comes Next? Simple Practices to Improve Diversity in Science. *ACS Cent. Sci.* **6**, 1231–1240 (2020).

7. Andersen, J. P., Nielsen, M. W., Simone, N. L., Lewiss, R. E. & Jagsi, R. COVID-19 medical papers have fewer women first authors than expected. *eLife* **9**, e58807 (2020).

8. Day, A. E., Corbett, P. & Boyle, J. Is there a gender gap in chemical sciences scholarly communication? *Chem. Sci.* **11**, 2277–2301 (2020).

9. Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C. R. Bibliometrics: Global gender disparities in science. *Nature News* **504**, 211 (2013).

10. Macaluso, B., Larivière, V., Sugimoto, T. & Sugimoto, C. R. Is Science Built on the Shoulders of Women? A Study of Gender Differences in Contributorship. *Academic Medicine* **91**, 1136–1142 (2016).

11. Dung, S. K. *et al.* Illuminating Women's Hidden Contribution to Historical Theoretical Population Genetics. *Genetics* **211**, 363–366 (2019).

12. Royal Society of Chemistry. Breaking the barriers: Women's retention and progression in the chemical sciences. (2018).

13. Synthetic Organic Faculty in the U.S. | ACS Division of Organic Chemistry. https://www.organicdivision.org/organicsyntheticfaculty/.

14. Data Tables - nsf.gov - Women, Minorities, and Persons with Disabilities in Science and Engineering - NCSES - US National Science Foundation (NSF). https://www.nsf.gov/statistics/2017/nsf17310/data.cfm.

15. Finkelstein, M. J., Conley, V. M. & Schuster, J. H. Taking the measure of faculty diversity; TIAA Institute: New York, 2016; pp 1 - 18

16. https://gender-api.com/

17. https://github.com/adamcotton/gender-scraper

18. Adam D Cotton. Raw genderized data for 'examining gender imbalance in chemistry authorship' [Data set]. http://doi.org/10.5281/zenodo.3939710 (2020).

19. Selected journals and years they were added to PubMed: JACS: 2000, Tetrahedron and Tetrahedron Letters: 2005, ACS Chemical Biology and ChemMedChem: 2006, Nature Chemistry: 2009.

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:
Adam Cotton
6CD3871C990E411...
Author Signature

12/9/2021
Date