**Using Chemical Biology to Modulate Antibody Activity**

By



Jittasak Khowsathit

Submitted to the graduate degree program in Molecular Biosciences and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chair: Mark Richter, Ph.D

_____

Joanna Slusky, Ph.D.

_____

Scott Hefty, Ph.D.

_____

Krzysztof Kuczera, Ph.D.

_____

Chris Fischer, Ph.D.

Date Defended: 1 May 2019

The Dissertation Committee for Jittasak Khowsathit certifies that this is the approved version of the following dissertation:

**Using Chemical Biology to Modulate Antibody Activity**

_____

Chair: Mark Richter, Ph.D.

Date approved: 1 May 2019

## ABSTRACT

Monoclonal antibodies have shown promising results as therapeutic agents, and yet they can also be associated with adverse side effects due to activity outside the disease site. Aiming to reduce these side effects, we have explored the possibility of a tunable antibody, whose activity can be manipulated via the addition of a small molecule. Previously, we incorporated a single cavity-forming mutation (tryptophan to glycine) into an antibody, and observed reduced antigen-binding activity that could be restored by addition of a complementary ligand (indole) — albeit with binding affinity too low for potential therapeutic applications.

Here, I describe a novel computational strategy for enumerating larger cavities in a fluorescein-binding single-chain variable fragment (scFv), leading to a designed variant with three large-to-small mutations (triple mutant) at the domain-domain interface with reduced antigen-binding. Through a complementary virtual screen, we identified a rescuing small molecule (JK43) that enhances binding affinity for antigen. Thorough characterization of this system shows that the loss of activity upon mutation was due to loss of stability and domain dissociation; conversely, addition of JK43 restores stability of the antibody fragment, induces domain re-association, and rescues antigen binding.

Beyond this initial model system, I will also describe the transferability of this design (triple mutant and JK43) from the fluorescein-binding scFv onto an unrelated scFv that shares the same three residues used in this design. We hypothesize that this design will also prove transferable onto the many therapeutic antibodies that also share these three residues, including Ipilimumab (anti-CTLA-4), Atezolimumab (anti-PD-L1), Nivolumab (anti-PD-1) and Adalimumab (anti-TNF-α).

*To the memory of my grandmother, Pattra Khowsathit*

## ACKNOWLEDGEMENTS

I am very grateful to many individuals who helped and supported me though my PhD study and would like to offer my sincerest gratitude to all of them.

First and foremost, I offer my greatest appreciation to my advisor, Dr. John Karanicolas for his advice, inspiration, and most importantly, his patience to bear with my during the course of my PhD study.

I am grateful to all the lab members, past and present, for our friendship, valuable discussions and feedbacks. Special thanks to Dr. Andrea Bazzoli for his help on various technical issues that crossed my path.

I am very thankful to Dr. Mark Richter, Dr. Scott Hefty, Dr. Joanna Slusky, Dr. Krzysztof Kuczera and Dr. Chris Fischer for serving on my committee.

Finally, I would like to convey my deepest gratefulness to my parents and the rest of my family for their unconditional love.

## TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES

# Chapter I.
# Introduction

The emerging use of biologics in medical treatment has been dramatic in recent years, and amongst FDA-approved biologics therapeutic antibodies make up the largest portion (*1*). To date, there have been more than 30 FDA-approved antibody-derived drugs marketed for cancer treatment (*2, 3*). These anti-cancer antibodies can be grouped broadly by the nature of their antigens, such as antibodies that target soluble proteins, surface antigens expressed on cancer cells, or receptors on effector cells (*4*).

A class of cancer targets that have attracted considerable attention is the immune checkpoint family, immune system regulators that plays a role in immunological tolerance. Here, the immune system is capable of recognizing and responding to certain substances but chooses to remain tolerant to these substances, and thus the body makes use of this system to differentiate self from non-self, and preventing autoimmunity (*5, 6*). The discovery of the first immune checkpoint family member, cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), has deepened our understanding of an underlying mechanism by which cancer cells can exploit this system to evade T-cell mediated cell death, posing as a "self" (i.e. non-cancer) cells (*7-10*). This understanding spurred development of the first immune checkpoint inhibiting therapeutic monoclonal antibody, Ipilimumab, which is designed to prevent cytotoxic T-cells from recognizing cancer cells as "self" and is approved for metastatic melanoma treatment in 2011 (*11*).

Since then, more immune checkpoint proteins have been discovered and targeted for cancer treatment. One pair of immune checkpoint proteins that currently occupy the limelight are programmed cell death protein 1 (PD-1) expressed on the surface of

activated T-cells, B cells and macrophages (*12*) and its binding partner programmed death-ligand 1 (PD-L1) that is often upregulated on the cancer cell surface (*13-15*). These two proteins are the targets of multiple new immune checkpoint inhibiting monoclonal antibodies, including FDA-approved drugs Atezolizumab (*16*) and Nivolumab (*17*) that have been shown to prolong overall survival in patients with non-small cell lung cancer (*16, 17*). Although the immune checkpoint inhibitors mentioned above have shown promising early indications, they are also associated with severe, and sometimes fatal, adverse side effects (*18-21*), largely due to the non-specific activation of T-cells (*22*). The ability to selectively blockade immune checkpoints, and thus localize immunological responses to the tumor environment, will be an important step in advancing therapeutic antibodies for safer cancer therapies.

**Allosteric protein switches: a proposed approach for limiting toxicity of therapeutic antibodies**

The use of chemical biology to modulate protein activity is a powerful strategy that can be used to engineer small molecule-dependent protein activity into specific proteins of interest. By tethering protein activity to the presence of a small molecule, one can switch "on and off" the protein using those small molecules. This scheme, if applied to therapeutic antibodies, may help reduce adverse side effects caused by activity of checkpoint inhibitors outside the tumor microenvironment. By re-engineering an allosteric site into therapeutic antibodies, we envision that the antibody can be produced in an "off state" by default, such that administering of this inactive form of the antibody would not trigger an immune response. Upon introduction of the activating small

molecule into the intended (tumor) site, the antibody will be activated and thus elicit toxicity only in the proximity of the injection site **(Figure 1.1)**.

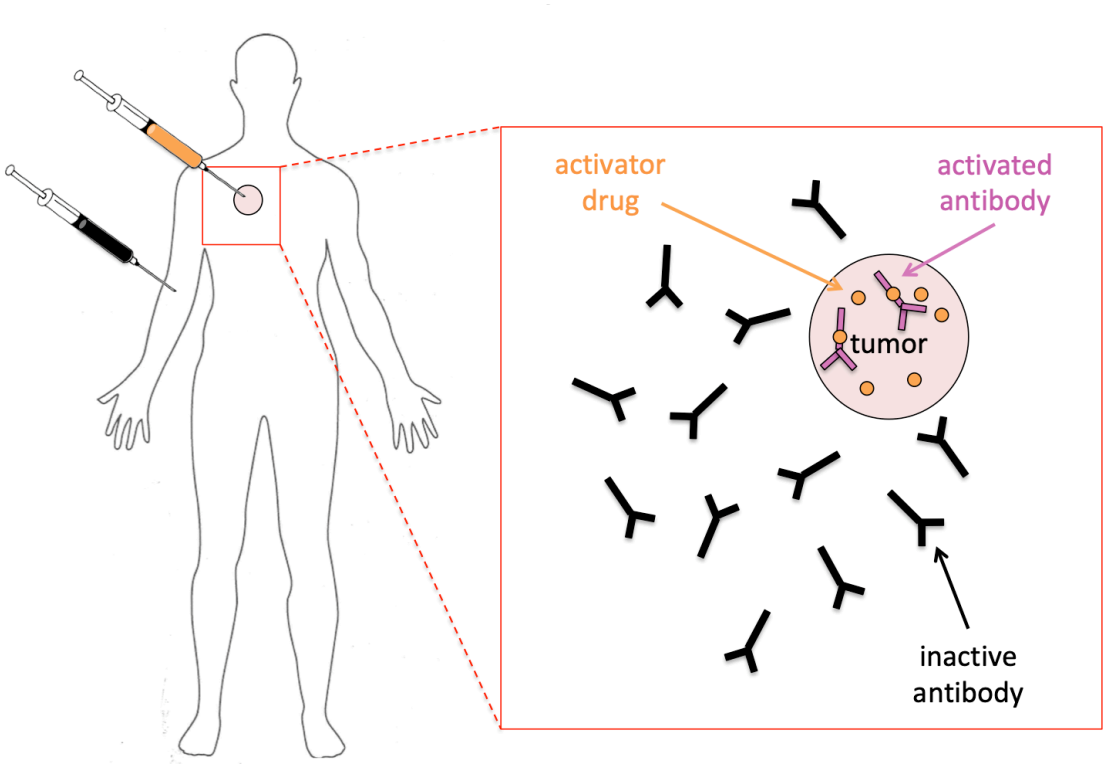


**Figure 1.1: Model of locally-activated therapeutic antibodies.** We envision that a therapeutic antibody could be systemically administered to a patient in an inactive form. An allosteric (small molecule) activator would then be administered locally, which in turn would activate the antibody only in this vicinity. Thus, antibody activity could be restricted to a specific region of the body, for example a tumor microenvironment. We further envision that the ability to connect antibody activity to the presence of a specific ligand will enable additional future refinements, such as using pro-drug forms of the rescuing ligand for enhanced precision with respect to spatial and temporal control.

*Typical Approaches for Constructing Allosteric Switches*

Most conventional methods used to create allosteric switches take advantage of proteins harboring naturally-occurring allosteric domains that change conformation upon binding to their binding partners. Such domains are used as "input domains" and fused to some "output domain" (harboring a readout or effector function), in a process called domain insertion (*23*). In order to achieve productive interdomain coupling upon binding of the allosteric activator, the proper arrangement of the two domains is critical. To pair new combinations of input-output domains, circular permutation is commonly employed to identify the optimal fusing point between the input and output domain, and very large numbers of potential insertion sites are explicitly screened (*23-26*).

Alternatively, a "split strategy" has also been used to create allosteric switches. In this case, the output domain is split into two halves, and each half of the output domain is fused onto a different part of the input domain (*27*). When the input domain is in its apo conformation, the two halves of the output domains are kept apart and are thus inactive. Upon addition of the effector ligand, the input domain undergoes a conformational change that brings the two halves of the output domain into close proximity with one another, thus restoring their activity.

The main challenges for these conventional strategies are in the rational design of allosteric coupling (the detailed mechanism of activation). To date, the accuracy of structure-based models for recapitulating precise conformational changes upon small molecule binding are still limited. For this reason, conventional methods still rely heavily on explicit (wetlab) screening in order to assemble the various components of the system into a working allosteric switch.

*Chemical Rescue of Structure*

As an alternative approach, we have drawn inspiration from the observation that internal cavities formed from mutations can be filled by the binding of hydrophobic small molecules (*28*). We have previously used our approach, "chemical rescue of structure", to introduce allosteric manipulation of protein activity (*29-31*). Rather than relying on naturally occurring allosteric domains, chemical rescue of structure can be used to engineer allosteric sites into an arbitrary protein domain. Briefly, this method identifies a structurally-critical residue located in the interior of the protein of interest (a "buttressing" residue), and substitutes a smaller residue in its place. This mutation destabilizes the protein, and induces residues around the mutation site to adjust and fill in the resulting hydrophobic void; in doing so, the resulting conformational change can affect the geometry of the active site, and thus result in loss of activity. In a subsequent step, this mutation can be complemented with an exogenous small molecule that binds selectively and non-covalently to the pocket created by the mutation. If the ligand is carefully chosen to mimic the original sidechain that was removed upon mutation, it can move the surrounding residues precisely back to their original (WT) conformations, and thus restore protein activity.

This approach has an advantage over conventional approaches, in that the precise mechanism of inactivation and rescue is not required to predict the loss and subsequent recovery of activity. Chemical rescue of structure, instead, relies on the ability to predict loss and recovery of protein stability upon small molecule binding, and protein stability is then used to predict loss or gain of activity (*29*).

To date, chemical rescue of structure has been demonstrated by using indole to rescue designed tryptophan-to-glycine (W→G) substitutions. This strategy has been applied successfully in two different enzymes (*30, 31*), a fluorescent protein (*29*), a transcription factor (*29*), and more recently, in an antibody (*32*). In each case, however, the opportunity to deploy the resulting switches for any real application was prevented by the fact that millimolar concentrations of indole were needed to recover protein activity. For the goal of modulating the activity of therapeutic antibodies described earlier, for example, it is not feasible to introduce millimolar concentrations of indole into a patient. Attempts to explore indole analogs, in search for a rescuing small molecule with tighter binding affinity, did not result in a rescuing ligand with significant improvements in binding affinity but rather showed that none of these rescued better than indole itself (*31, 32*). We surmise that the high concentration of indole needed to activate these designed switches arises due to indole's low molecular weight, which in turn limits the potentially attainable binding affinity (*33*).

With respect to a therapeutic antibody, and an scFv fragment in particular, I envision a scenario in which mutations at the domain-domain interface could destabilize this interface such that the two domains dissociate, and thus this protein no longer binds to antigen. Upon addition of a complementary ligand, the two domains would then re-associate in precisely the orientation needed for recognition of antigen, restoring activity **(Figure 1.2).**

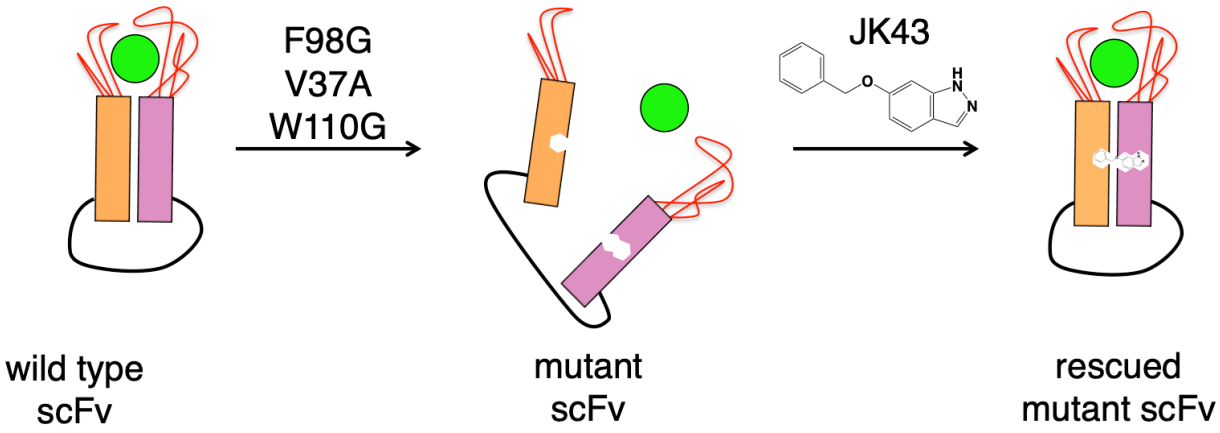


**Figure 1.2: Schematic of design strategy.** Mutations on the light and heavy chain interface lead to dissociations of the domain-domain interface, which leads to loss of antigen binding activity. Subsequent addition of exogenous rescuing ligand (JK43) induces re-association of the domain-domain interface, and thus restores activity.

In the following chapter, I describe a computational strategy for enumerating larger cavities that arise from multiple simultaneous large-to-small mutations at adjacent buried sites, and its application to a model scFv. This mutation is then complemented by a larger rescuing ligand, which binds with higher affinity than any of our earlier designs.

# Chapter II.
# Computational Design of an Allosteric Antibody Switch

## Design Strategy

To improve the binding affinity of the activating ligand, we sought to move towards a small molecule with higher molecular weight. The cavity formed from a single W→G mutation would be too small to accommodate a ligand much larger than indole, thus necessitating extension of our approach to use multiple (simultaneous and adjacent) cavity-forming mutations instead of just one.

Although my end goal is to engineer switchable variants of immune checkpoint inhibiting antibodies, producing monoclonal antibodies (mAbs, a standard antibody format used for therapeutic antibodies) and assays checkpoint inhibition are too expensive and time-consuming for a first proof-of-concept study. For this reason, I have instead selected a fluorescein-binding scFv (4D5Flu) as a first model system. This construct can be conveniently expressed in *E. coli* (as compared to production of monoclonal antibodies made in mammalian hybridomas), and 4D5Flu's antigen, fluorescein, is a dye that allows for convenient monitoring of binding.

Prior to this study, the cavity-forming mutation used in chemical rescue of structure could be manually selected by identifying tryptophan residues in the protein of interest. In the current study, where we are trying to create a larger cavity from multiple residue mutations, an issue arises because there is a much larger number of possible mutations that one could make. We therefore developed a collection of criteria for defining allowed sets of mutations:

1) The mutations must together remove at least 12 heavy atoms from the protein.

2) At least one mutated residue must be aromatic.
3) All mutated residues must be located at the interface between the two domains of the scFv.
4) None of the mutated residues can be located on the CDR loops.
5) We will consider only double- and triple-mutants.
6) The mutated sidechains must be in close proximity to one other, so that the resulting cavity will be contiguous.

Even with these criteria in place, there remain a vast number of possible combinations to test. To address this, we developed a computational approach for collecting and considering them.

We began by writing an application in the Rosetta macromolecular software suite that implements the above criteria. The program loops over pre-defined cavity-forming mutations for all 2- and 3-residue groups, and upon finding a set of mutations that matches the criteria listed above it exports the spatial coordinates of the atoms that would be deleted by the mutation. We call this output a "constellation" of atoms, because this precise *three-dimensional* arrangement of atoms is what we will subsequently seek to mimic using a small molecule (**Figure 2.1A**).

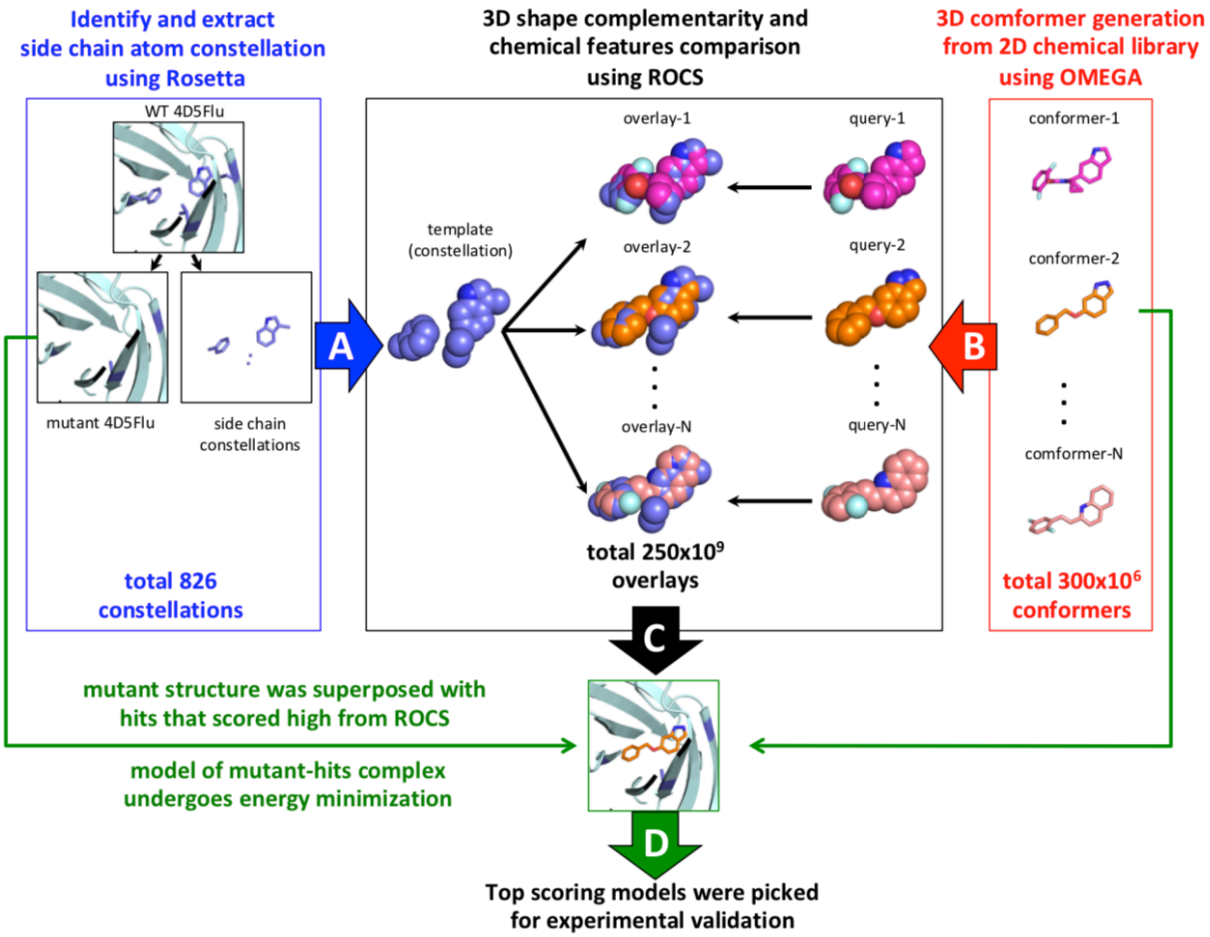


**Figure 2.1: Computational design of antibody switch.** ***(A)*** Groups of two/three cavity-forming mutations are exhaustively considered, in order to identify those located in the domain/domain interface that would yield a suitable cavity for subsequent rescue by a drug-like small molecule. ***(B)*** For each drug-like compound in the ZINC12 database, energetically favorable three-dimensional conformations were generated. ***(C)*** For each constellation of atoms that could be deleted from the protein domain/domain interface, every conformation of every compound was considered as a potential structural match. ***(D)*** The top-scoring structural matches were refined in the context of the (mutant) protein environment, and the top-scoring designs after re-ranking were experimentally characterized.

Applying this strategy to the 4D5Flu scFv yielded a collection of 826 unique constellations – each of these is the outcome of a distinct double/triple mutant, because each mutant removes a specific set of atoms from the scFv's domain-domain interface. Each constellation was used as a template to virtually screen for small molecules that best complement the corresponding cavity formed. In search for small molecules that best recapitulate each of the 826 unique constellations, after virtual screening, it is necessary that we could acquire the hit compounds from these databases for experimental validation.

Even though larger chemical libraries do exist, we specifically picked ZINC12 database (*34*), a library of commercially-available small molecules. We further curated ZINC12 using the following criteria:

1) Must follow Lipinski's rule of 5 (*35*)
2) Must have molecular weight ≥ 180 Da
3) Must not be a pan-assay interference compound (PAINS) (*36-38*)
4) Must not contain chemically-reactive groups (*39, 40*)

Applying these filters reduced the size of ZINC12 from over 20 million entries down to ~3 million entries.

Entries in the ZINC12 database are "flat" chemical structures, rather than 3D structures. In order to use these for screening against the 3D constellations, it was therefore necessary to generate several low-energy 3D models for each of the curated ZINC12 entries. We carried out this step using the OMEGA application (*41*), generating up to 100 energetically-favorable 3D models ("conformers") for each of the 3 million

curated ZINC12 entries. All of these conformers were then compiled together to create a 3D conformer library (300 million conformers total) (**Figure 2.1B**).

Next, we used the ROCS application (*42*) to compare each of the 826 constellations against each of the 300 million conformers in an all-versus-all manner. Each of these 2.5 billion comparisons (overlays) consisted of a fast alignment step, followed by a scoring step. This scoring step includes consideration of both 3D shape complementarity and conservation of chemical features (e.g. position of hydrogen bond donor/acceptors, aromatic rings, etc.) between the template (constellation) and the query molecule (conformers) (**Figure 2.1C**).

For these candidate ligands to rescue activity, they must bind to the designed cavity in the active (WT) conformation of the protein. Due to the nature of constellations, which are combinations of atoms from different sidechains, groups of atoms in the constellation are not covalently connected to one another. While the rescuing ligand is intended to mimic the atoms in the constellation, there are inevitable differences because the rescuing ligand is a single entity; in particular, there are often additional atoms that link together the functional groups that overlap with the constellation. The extra atoms that comprise the linker(s), if not chosen carefully, could cause steric clashes with the walls of the cavity and prevent ligand binding.

To test which of the top-scoring ligand from the ROCS alignments indeed fit within their cognate designed cavity, we used the ligand/constellation alignment to transfer the best scoring ROCS hit ligand from each constellation into the designed cavity and then subjected these models to energy minimization in Rosetta (**Figure 2.1D**).

After minimization, model complexes were filtered based on a combination of physical properties and on the similarity of the minimized complex to the WT structure, using the following criteria:

1) Ligand solvent-accessible surface area (SASA) ≤ 0.13 $Å^2$
2) Number of rotatable bonds in the small molecule ≤ 4
3) Numbers of buried unsatisfied polar atoms ≤ 1
4) Ligand RMSD compared to constellation atoms ≤ 1.38 Å
5) Backbone RMSD compared to WT ≤ 0.60 Å
6) Ligand–cavity interaction score ≤ -10.29 Rosetta energy units (REU)
7) Rosetta score difference compared to WT;
    a. (model of apo mutant) – (WT) ≥ 5 REU
    b. (model of holo mutant) – (WT) ≤ 5 REU

After applying the filters, the remaining designs (cavity forming mutations and their cognate ligands) were ranked by their ligand-cavity interaction score. From the ranking, five designs were selected for experimental validation. One of these in particular **(Figure 2.2)** was characterized in depth through a suite of biochemical and biophysical assays, as described in the following Chapter.

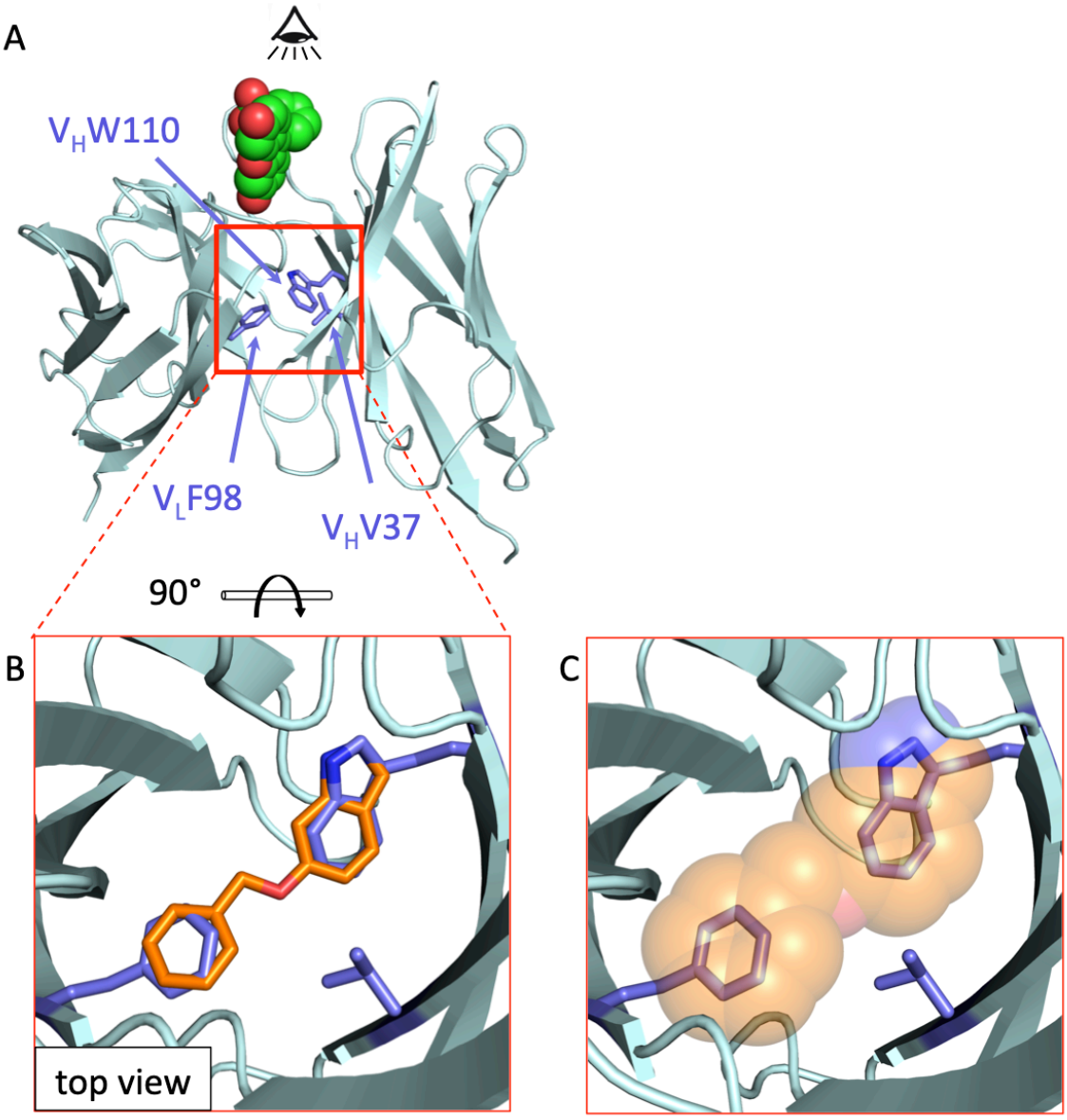


**Figure 2.2: Design model of the rescued complex.** ***(A)*** The crystal structure of 4D5 (which harbors the same framework as 4D5Flu), taken from PDB ID: 1FVC. The location of the fluorescein antigen is modeled into this structure using a related antibody, PDB ID: 4NJ9. The residues that comprise our experimentally characterized triple mutant are indicated in blue sticks. ***(B)*** A superposition of the constellation identified through our computational approach, together with the ligand predicted to rescue this mutation (JK43, *orange*) shown in sticks or shown in spheres ***(C)***.

## Methods

*Compilation of 3D Small Molecule Library*

A library of 3 million commercially-available small molecules, downloaded from ZINC12 (*34*) and subsequently curated for drug-like properties, was used as a starting point to create a 3D library. For each of the 3 million chemical structures, 100 low-energy conformers were created using the OMEGA software (*41*), using the following command:

```
omega2 -in input_file.smi -out output_file.sdf.gz
-prefix ligand_name -warts -maxconfs 100
```

*Rapid Overlay of Chemical Structure*

To match constellations with specific complementary small molecules, we used shape-based chemoinformatic virtual screening at implemented in the ROCS software (OpenEye) (*42*). First, ROCS measures the shape overlap between template structure and each conformation of every compound in our library. The shape of a template molecule and the shape of each query molecule are aligned very rapidly using atom-centered Gaussian functions to maximize the overlap of volume between them. The degree of shape similarity between template and query is calculated using the following equation:

$$Tanimoto_{f,g} = \frac{O_{f,g}}{I_f + I_g - O_{f,g}}$$

The *I* terms are self-volumes of each molecule, while the *O* term is the overlap between two molecules. Second, in addition to shape complementarity (ShapeTanimoto),

ROCS also includes an additional pharmacophoric feature (ColorTanimoto), that takes into account the spatial overlap of features such as hydrogen bond donors, hydrogen bond acceptors, and aromatic rings. It has been shown that the combination of shape and chemical features (TanimotoCombo) yields superior results relative to using one of the component scores alone (*43-45*).

In our study, the template and query molecules were aligned using the following command:

```
rocs -dbase database.sdf.gz -query query.pdb
-prefix structure_name -cutoff -1.0 -oformat sdf -scdbase true
-maxhits 4000 -maxconfs 4000 -outputquery false
```

*Rosetta Software Suite*

Extraction of constellations, refinement protein/ligand complexes, and final re-ranking were carried out using the Rosetta software suite (*46*). The ROCS screen described above provides an alignment of a given ligand onto the template constellation. Since the constellation was extracted from the protein coordinates, this alignment is used to directly build a model of this ligand in complex with the mutant protein. The model was then subjected to energy minimization in Rosetta, using the following command line:

```
minimize.linuxgccrelease -s input_pdb –relax:fast
–in:file:fullatom
```

## Chapter III.
## Experimental validation of switchable antibody design

### Materials and Methods

*Plasmid construction*

Because each of the two immunoglobulin domains of an scFv includes a disulfide bond, and we sought to express soluble scFv with the correct redox state in an *E. coli* system, we elected to add a periplasmic signal sequence at the N-terminus of the scFv. This extra 21-residue signal sequence is auto-cleaved during the secretion of the polypeptide into the periplasm (*47, 48*). In addition, to facilitate purification and to improve expression of soluble fractions, we used a vector that includes a hexa-histidine tag (His-tag) and a maltose binding protein tag (MBP-tag).

The vector DNA selected for cloning was the pET PPL His6 MBP LIC cloning vector (2K-T), purchased from QB3-Berkeley Core Research Facilities (Addgene plasmid # 37183). In addition to the features described above, this vector also contains a tobacco etched virus protease cleavage site (TEV-site) between the three tags and our gene of interest. This vector was originally designed for Ligase Independent Cloning (LIC), but due to the emerging popularity and low failure rates of Gibson Assembly (*49*), it has been adapted for cloning with the latter approach.

Empty vector was transformed into an NEB® Turbo Competent *E. coli* (High Efficiency), plated in a LB agar plate with 100 μg/mL ampicillin, incubated overnight at 37 °C. A single colony from the plate was picked and grown in 50 mL of LB broth with

100 μg/mL carbenicillin, incubated with shaking at 250 RPM, overnight at 37 °C. Plasmid DNA from liquid culture was isolated according to protocol provided from manufacturer (QIAprep Spin Miniprep Kit, QIAGEN). Next, the empty vector DNA was linearized using SspI-HF® (NEB), reaction performed according to protocol provided by the manufacturer and linearized vector was gel purified from a 0.8% agarose gel using a QIAquick Gel Extraction Kit (QIAGEN).

As for insert DNA, WT and mutant genes were synthesized in a double-stranded linear DNA fragment format (GeneArt Strings DNA Fragments, ThermoFisher). The WT amino acid sequence was obtained from a previous published study (*50*). These genes were prepared for Gibson Assembly by adding the following sequence to the 5'-end of the sense strand and antisense strand, respectively:

Sense 5'- TACTTCCAATCCAATGCA-3'

Antisense 5'- TAATAACATTGGAAGTGGATAA-3'

DNA strings were supplied in lyophilized vials and were reconstituted with molecular biology grade water.

Gibson assembly of the linearized 2K-T empty vector and prepared inserts were carried out by mixing 2 μL of vector (65 ng/μL) with 8 μL of insert (100 ng/μL), then 10 μL of Gibson Assembly® Master Mix (NEB) was added to the insert/vector mixture. Gibson assembly reaction was allowed to occur at 50 °C for 15 minutes, then cooled down to 4 °C. Reaction mixture was used to transform into a NEB® Turbo Competent *E. coli* (High Efficiency) and plated using the same protocol mentioned above. Three colonies from each transformation were picked and plasmid DNA prepared with the same protocol mentioned above, an aliquot of the plasmid DNA was sent out for successful

cloning verification, using Sanger sequencing. The primer set used for Sanger sequencing were T7 and T7-term.

| | |
|---|---|
| T7 | 5'-TAATACGACTCACTATAGGG-3' |
| T7-term | 5'-GCTAGTTATTGCTCAGCGG-3' |

*Protein expression*

scFv genes in 2K-T plasmids were transformed into *E. coli* BL21 (DE3) Competent (NEB) using the protocol suggested by the manufacturer and plated accordingly. A single colony was obtained from the plate and inoculated into a starter culture, 200 mL of TB broth, 0.4% glycerol, 100 μg/mL carbenicillin and incubated with shaking at 250 RPM, overnight at 37 °C. Starter culture was inoculated into the main culture (using the same media as starter culture), with the ratio of 1:20 dilution. Main culture was grown with shaking at 200 RPM, at 37 °C, until O.D. reached 0.8 (measured at 600 nM), temperature then lowered to 15 °C and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM, culture was left shaking for 48 hours. Cells were harvested by centrifugations at 3,500xG for 20 minutes. Used media was discarded and cell pellet's surface was rinsed with sterile ddi$H_2O$ (pass this point, all samples were kept at 4 °C at all times). Cells were then resuspended in PBS buffer (6 mM phosphate buffer, 150 mM NaCl, pH 7.4) and applied trough a French Press for 3 rounds. Cell lysate was spun down at 15,000xG for 30 minutes and supernatant (which contains soluble scFv) was filtered through a 0.22-micron PES syringe filter.

*Protein purification*

First, cell lysate was purified with Capto L resin (GE Healthcare) by flowing lysate through the column (flow rate 1 mL/min) using an ÄKTA pure FPLC system (GE Healthcare). Column was washed using PBS until UV signal reached baseline level. Column was eluted using 15 mM NaOH with eluted fractions directly collected into a tube containing 1M sodium citrate, pH 8.0 (5% of fraction volume). Fractions containing scFv with higher than 70% purity was pooled and concentrated down to 50 mL and dialyzed against 50 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole, pH 8.0 for 3 rounds.

Second, 5 mL of TEV protease 0.1 mg/ml was added to the dialyzed sample in the dialysis bag. The dialysis bag was then moved into a TEV reaction buffer 50 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole, 2 mM DTT (freshly prepared) pH 8.0 and left to dialyze for 4 hours before moving the bag into a fresh batch of the same buffer and left to dialyze overnight. Samples were then dialyzed against 50 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole, pH 8.0 for 3 rounds to remove DTT.

Third, samples were then applied though a Dextrin Sepharose® (GE Healthcare) column (flow rate 1 mL/min) using an ÄKTA pure FPLC system (GE Healthcare). Flow-through fractions were collected and was then passed through a Ni Sepharose HisTrap FF (GE Healthcare) to remove TEV protease. Again, the flow-through fractions were collected; at this stage scFv are about 90% pure.

Finally, samples were then concentrated down to 2.5 mL and filtered through a 0.22-micron PES syringe filter. Next, the filtered samples were purified using a Superdex 75 size exclusion column (GE Healthcare), using 50 mM sodium phosphate, 500 mM

NaCl, 5 mM imidazole, pH 8.0 buffer. Purified samples were then dialyzed into PBS, aliquoted into smaller volumes, then flash frozen with liquid nitrogen and stored at -80 °C for further use.

*Fluorescence quenching*

For the protein titration experiment, samples were prepared by titrating varying concentrations of protein (WT or triple mutant) into PBS containing a final concentration of 10 nM fluorescein and 1% DMSO, with or without 100 μM JK43. For the JK43 titration experiment, samples were prepared by titrating varying concentrations of JK43 into PBS with a final concentration of 10 nM fluorescein, 1% DMSO and 500 nM protein (WT or triple mutant). For the SAR experiment, samples were prepared by adding JK43 analogs (to a final concentration of 10 μM) into PBS with a final concentration of 10 nM fluorescein, 1% DMSO and 500 nM triple mutant.

Fluorescence quenching experiments were carried out in a SpectraMax i3x Multi-Mode Microplate Reader (Molecular Devices), using a Costar 96-well black round bottom plates, 200 μL reaction volume, excitation wavelength of 475 nm and emission wavelength of 525 nm, at a temperature of 27 °C. Blank measurements containing all sample components except fluorescein were subtracted to remove background. All samples were normalized to a control of fluorescein alone (0% quenching) and buffer alone (100% quenching).

MTPS quenching was monitored in the same manner as fluorescein quenching except for using MTPS instead of fluorescein, 400 nM triple mutant anti-MTPS antibody, and using excitation wavelength 375 nm, and emission wavelength 435 nm.

*Thermal shift assay*

Differential scanning fluorimetry experiments were carried out in a QuantStudio™ 6 Flex Real-Time PCR System (ThermoFisher), using a 384-well white round bottom plate, 20-30 µL reaction volume, excitation wavelength of 580 nm and emission detected using ROX filter setting (623 nm). Samples were prepared in PBS with 5 µM protein, 5x SYPRO Orange (ThermoFisher), and 1.1% DMSO with or without 100 µM JK43. The scFv melting curve was collected from 25-95 °C (0.5 °C/min) and data were analyzed with PRISM 6 (GraphPad) using the Gibbs-Helmholtz equation (*51*). Samples were run in 8 replicates.

*NMR studies*

NMR spectra were recorded at 20 ºC on a Bruker Avance II 600 MHz NMR instrument equipped with a TCI triple-resonance cryogenic probe. The protein sample was dissolved in PBS buffer (6.1 mM phosphate buffer, 154 mM NaCl, pH 7.4), at a concentration of 100 µM.

NMR data were collected using the Bruker standard pulse sequences hsqct1etf3gpsi3d and hsqct2etf3gpsi3d (*52*), then were analyzed using TopSpin 3.2 (Bruker, USA). The $T_1$ and $T_2$ relaxation times were determined both by fitting peak intensities of the biggest peak at ~8.3 ppm, and by integrating between 7-10 ppm (covering most amide proton peaks) as a function of 11 relaxation delay intervals ($T_1$: 0.02, 0.05, 0.1, 0.2 0.3, 0.4, 0.6, 0.8, 1.0, 1.2, and 1.5s, and $T_2$: 1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 20s) (*53*) based on:

$$I_{(t)} = I_{(t)} + Pe^{(\frac{-t}{T_1})}$$

$$I_{(t)} = Pe^{(\frac{-t}{T_2})}$$

The Igor Pro software (WaveMetrics, USA) was used to calculate the $T_1$ and $T_2$ relaxation times.

**Results**

*JK43 rescues fluorescein binding of the 4D5Flu triple mutant*

4D5Flu is an anti-fluorescein scFv that was generated by grafting the CDR loops from 4-4-20 (anti-fluorescein Fab) (*54, 55*) onto a humanized anti-HER2 scFv scaffold called 4D5 (*50*). The two halves of 4D5Flu are connected with a 30-residue linker $(Gly_4Ser)_6$. Out of the five multi-residue cavity-forming mutations constructs generated from the computational pipeline, one construct ($V_L$F98G/$V_H$V37A/$V_H$W110G) was further characterized in-depth, this construct will be referred to as "triple mutant", from this point onwards **(Figure 2.2)**. We expressed and purified WT and triple mutant scFv and tested the fluorescein-binding activity (quenching) of both constructs.

Binding by 4D5Flu quenches the fluorescence signal from fluorescein; we took advantage of this property to monitor the quenching activity. In the absence of activating ligand, the triple mutant showed diminished quenching relative to WT **(Figure 3.1A)**. Upon addition of 100 μM JK43, activity of the triple mutant is restored nearly to that of WT **(Figure 3.1A)**.

The rescue of fluorescein quenching in the triple mutant exhibited dose-dependency with an $EC_{50}$ of 16 μM **(Figure 3.1B)**. In order to verify that the rescued quenching activity was not due to direct fluorescein quenching by JK43, we next titrated

increasing concentrations of either WT or triple mutant protein into a sample with no JK43 or with a fixed concentration of JK43 (100 μM) **(Figure 3.1C)**. It is evident that the presence or absence of JK43 does not affect WT binding curve, whereas addition of JK43 shifts the binding curve of the triple mutant closer that of the WT curve.

In summary, then, we find that collectively introducing these three mutations leads to partial loss of antigen-binding activity for this scFv, and that this activity can be rescued through addition of JK43. These observations are consistent with the original model underlying our design strategy **(Figure 1.2)**.

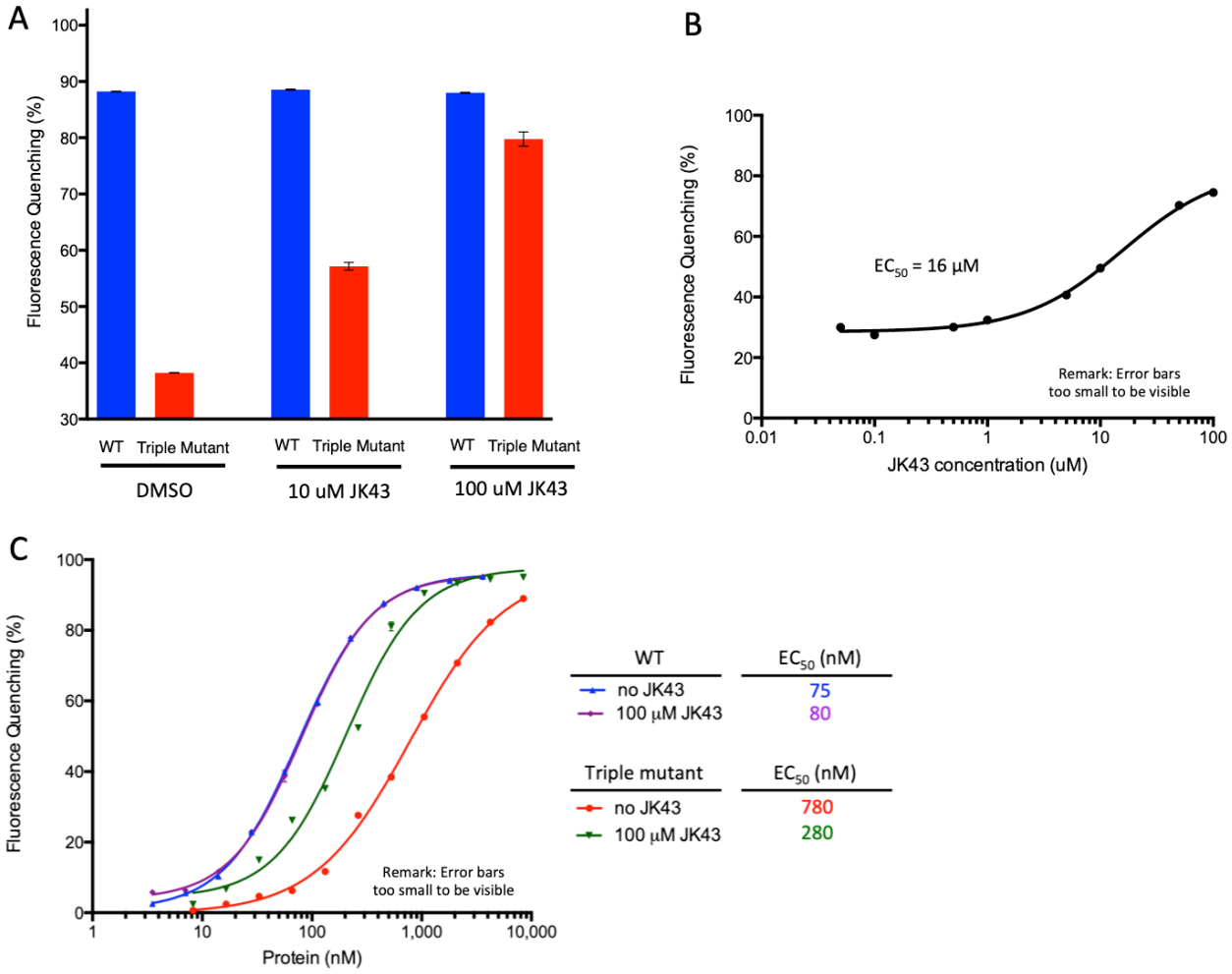


**Figure 3.1: JK43 rescues activity of the 4D5Flu triple mutant.** ***(A)*** Relative to WT 4D5Flu, the designed triple mutant shows diminished quenching of fluorescein fluorescence. Upon addition of 10 or 100 µM JK43, activity of the triple mutant was partially rescued. ***(B)*** Addition of JK43 rescues fluorescence quenching of the triple mutant in a dose-dependent manner, with an $EC_{50}$ of 16 µM. ***(C)*** Protein titrations in the presence or absence of 100 µM JK43. All data are presented as mean ± SEM, n = 8.

*JK43 stabilizes the triple mutant in the presence of antigen*

From previous examples of chemical rescue of structure, we observed that alteration of the protein thermostability could lead to inactivation and rescue (*29, 32*). We therefore used differential scanning fluorimetry (DSF) to examine the effect of this triple mutation on thermostability. SYPRO Orange dye has higher fluorescence intensity when bound to exposed hydrophobic parts of proteins, and thus it provides a probe for proteins' thermal unfolding (*56-58*).

For both the WT and the triple mutant, we used DSF to monitor the thermal unfolding transition that results from increasing temperature. Both unfolding curves were fit using the equation from (*51*), resulting in a melting temperatures ($T_m$) of 56.2 °C and 44.0 °C, respectively **(Figure 3.2A)**.

Upon addition of 5 μM fluorescein (equimolar to protein concentration, and far above the dissociation constant for this interaction), both the WT and the triple mutant show elevated $T_m$, indicating that bound fluorescein stabilizes the folded protein **(Figure 3.2B)**. In contrast, addition of 100 μM JK43 (without fluorescein) destabilizes both the WT and the triple mutant **(Figure 3.2C)**. The change in $T_m$ for the WT scFv (which lacks a designed binding site for JK43) suggests that this compound is destabilizing the protein through a non-specific mechanism, presumably by making the solvent more non-polar and thus stabilizing exposure of the protein's hydrophobic sidechains. The fact that the triple mutant is similarly destabilized suggests that the binding site for JK43 is not present, presumably because the protein does not adopt the necessary conformation in the absence of fluorescein.

To test this, we next monitored unfolding in the presence of both fluorescein (5 μM) and JK43 (100 μM). Under these conditions the WT construct was slightly stabilized, but less than the effect observed with fluorescein alone **(Figure 3.2D)**. In essence, these two ligands appear to act independently on the WT construct, in their own competing directions. By contrast, addition of the two ligands together stabilizes the triple mutant *more* than fluorescein alone. Put another way, JK43 can stabilize the triple mutant, but *only* if it is in a conformation appropriate for antigen binding. In this case, our interpretation is that addition of fluorescein at a concentration far above the dissociation constant reverses the conformational change resulting from introducing the triple mutation, and thus pre-orders the designed binding site for JK43. This explains why addition of JK43 is destabilizing for the fluorescein-bound WT construct, but stabilizing for the triple mutant.

Collectively, these findings point to positive cooperativity between antigen binding and JK43 binding in the triple mutant construct. Our fluorescence quenching experiment demonstrated that JK43 enhanced the triple mutant's binding affinity for antigen, and here we demonstrate the converse: the presence of antigen enhances binding of JK43.

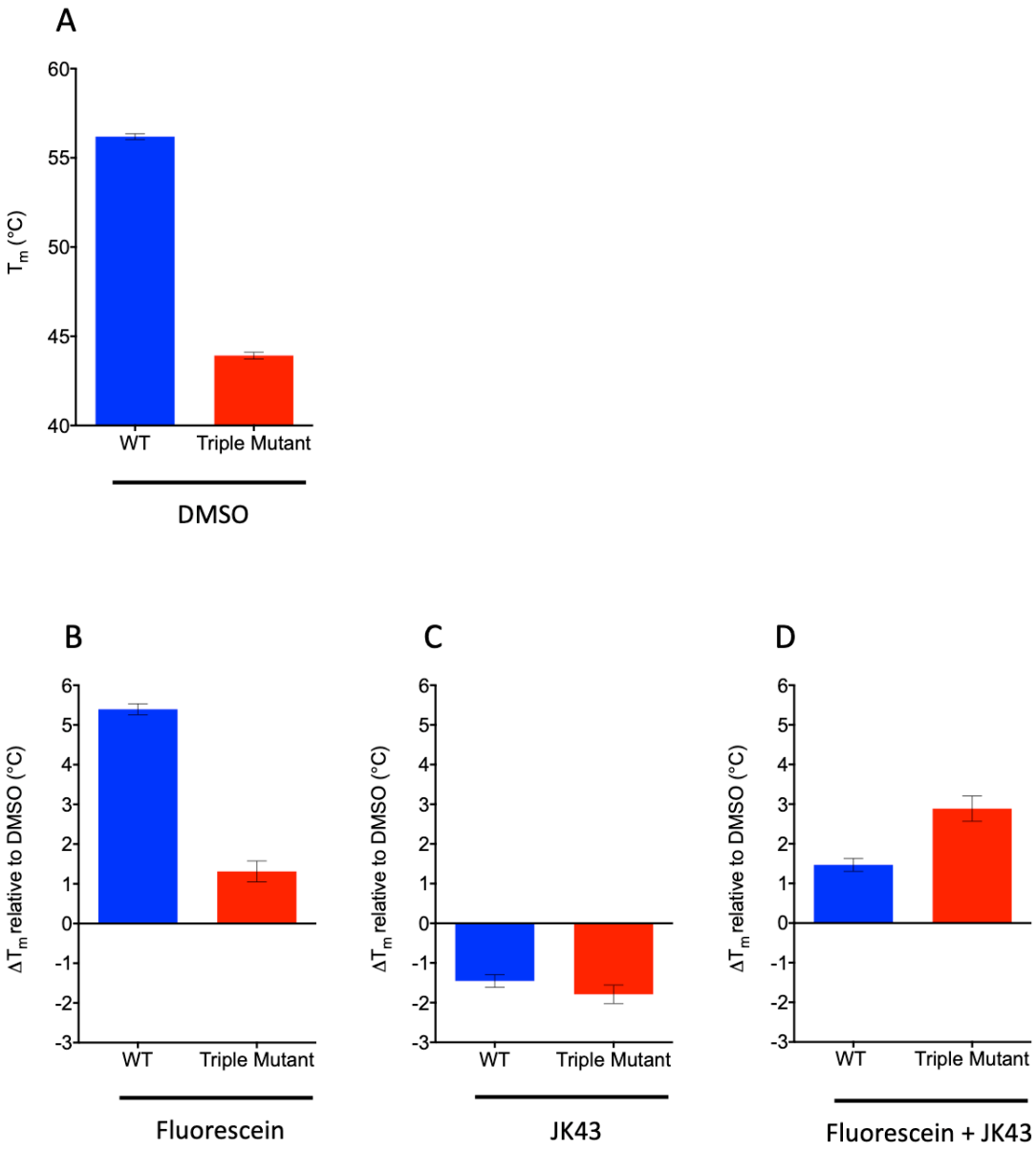
A
60
55
50
45
40
$T_m$ (°C)
WT
Triple Mutant
DMSO
B
C
D
6
5
4
3
2
1
0
-1
-2
-3
$\Delta T_m$ relative to DMSO (°C)
WT
Triple Mutant
Fluorescein
JK43
Fluorescein + JK43

**Figure 3.2: JK43 rescues thermal stability of the 4D5Flu triple mutant only in the presence of fluorescein.** ***(A)*** The apparent thermal unfolding temperature (as measured via differential scanning fluorimetry), for WT 4D5Flu and for the designed triple mutant. ***(B)*** Upon addition of 5 μM fluorescein (at a concentration far above the dissociation constant), both WT and the triple mutant show increased thermal stability. ***(C)*** Upon addition of 100 μM JK43 (without fluorescein), both WT and the triple mutant show decreased thermal stability. ***(D)*** In the presence of both fluorescein and JK43, both WT and the triple mutant show increased thermal stability. However, the increase in stability of the triple mutant extends beyond that observed upon addition of fluorescein alone, while the increase in stability of the WT is less than that observed upon addition of fluorescein alone. All data are presented as mean ± SEM, n = 8.

*JK43 enhances the interaction between the two scFv domains*

Next, we sought to directly test the structural basis for rescue of activity, by characterizing protein conformational changes in the triple mutant associated with addition of JK43. To do so, we produced uniformly $^{15}N$ labeled triple mutant protein and measured longitudinal ($T_1$) and transverse ($T_2$) $^{15}N$ relaxation times at 20 °C. In a stacked plot of the first spectra of $T_1$ relaxation (delay of 0.02 s) of the triple mutant with and without the presence of equimolar JK43 (100 μM), the spectra are dominated by three large peaks around 8.1-8.4 ppm **(Figure 3.3A)**. These peaks mainly arise from glycine and serine residues in the linker, as evidenced by their large intensities and their positions in the $^{15}N$ HSQC spectrum **(Figure 3.3B)**. Addition of JK43 results in a 3.4-fold increase in peak intensity (signal to noise ratio is 31 for the protein alone, and 104 upon addition of JK43). While longitudinal $T_1$ $^{15}N$ relaxation times remain nearly unchanged between the triple mutant with and without JK43, the calculated transverse $T_2$ $^{15}N$ relaxation times are significantly increased after the addition of JK43; this observation is consistent for $T_2$ values obtained both from the calculations based on peak intensity of the glycine residues in the linker (largest glycine peak around 8.3 ppm) and from calculations that integrate values between 7-10 ppm **(Table 3.1)**.

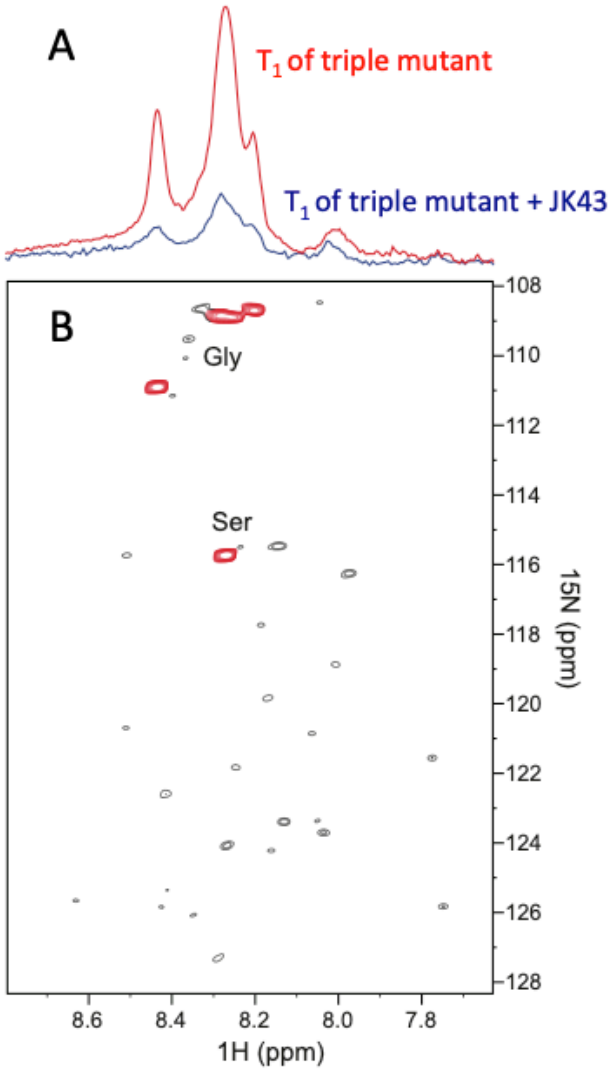


**Figure 3.3: Addition of JK43 induces compaction and ordering of the linker in the 4D5Flu triple mutant.** ***(A)*** 1D NMR spectra of T1 relaxation measurements at 0.02 s relaxation delay, for the 4D5Flu triple mutant alone (in blue) and upon addition of 100 μM JK43 (in red). The spectrum shows three dominant peaks in the 8.1-8.5 ppm range, based on their intensities and positions in the 1H-15N HSQC spectrum below it we infer that these correspond to the glycine and serine residues in the scFv's synthetic linker. ***(B)*** A portion of the 2D 1H-15N HSQC spectrum of the 4D5Flu triple mutant in the presence of 100 μM JK43.

| Sample | Peak Intensity | $T_1$ | $T_2$ | $T_1/T_2$ |
|---|---|---|---|---|
| Apo | ~8.29 ppm | 0.864±0.155 | 18.336±0.477 | 0.0471 |
| 1:1 | ~8.28 ppm | 0.825±0.026 | 22.789±0.421 | 0.0362 |
| 1:1.5 | ~8.27 ppm | 0.875±0.015 | 23.456±0.353 | 0.0373 |
| | | | | |
| | Integrate | $T_1$ | $T_2$ | $T_1/T_2$ |
| Apo | 7~10 ppm | 1.034±0.148 | 15.869±0.477 | 0.0652 |
| 1:1 | 7~10 ppm | 0.909±0.053 | 18.277±0.717 | 0.0497 |
| 1:1.5 | 7~10 ppm | 0.995±0.106 | 20.186±0.257 | 0.0493 |

**Table 3.1: NMR relaxation parameters.**

For rigid protein molecules, in the limit of slow molecular motion ($\tau_c >> 0.5$ ns) in high magnetic field (500 MHz or greater), it has been known (*52, 53, 59*) that a closed form solution for $\tau_c$ as a function of $T_1/T_2$ $^{15}N$ relaxation times exists:

$$\tau_c = \frac{1}{4\pi\nu_N}\sqrt{6\frac{T_1}{T_2} - 7}$$

And according to the Stokes-Einstein equation:

$$\tau_c = \frac{4\pi\eta\alpha^3}{3kT}$$

Where $\nu_N$ is nuclear frequency, $\eta$ is viscosity, k is Boltzmann constant, T is temperature, and $\alpha$ is hydrodynamic radius. Thus, increasing $T_1/T_2$ reflects an increase in $\tau_c$, and an increase in $\tau_c$ reflects an increase in $\alpha$.

With this interpretation, our observation of decreased $T_1/T_2$ upon addition of JK43 suggests that upon binding, this ligand reduces the scFv's hydrodynamic radius and its rotational correlation time. Additionally, we note that the peaks in the glycine and serine region of the HSQC spectrum (**Figure 3.3B**) become narrower upon addition of JK43 than the corresponding peaks in the spectrum of protein alone: this suggests increased uniformity of the linker as JK43 induces the protein to become more compact.

Overall, these results are consistent with a model in which the two domains comprising the scFv are dissociated from one another in the triple mutant, and then associate in response to addition of JK43. Because residues involved in antigen recognition are distributed across the CDR loops of both domains, this explains loss and then recovery of antigen binding upon mutation and subsequent rescue **(Figure 1.2)**.

*Structure-activity relationship of the rescuing ligand*

From our previous studies of (W→G) substitutions, we observed that indole (corresponding exactly to tryptophan's sidechain) provided the most effective rescue of activity; however, we also observed that other ligands yielded partial rescue, with rank order commensurate with their similarity to indole (*31, 32*). These studies provided motivation for selection of JK43 as a complement for this particular triple mutant: we anticipated that an optimal rescuing ligand would closely recapitulate the three-dimensional shape and chemical features of deleted atoms in the original protein structure (i.e. our constellation), and this formed the basis for our computational virtual screening approach.

To test this hypothesis, we used a series of JK43 analogs to explore how varying the structure of the ligand would affect recovery of antigen binding in this particular triple mutant. Given that the two ring systems in JK43 are intended to separately mimic a deleted phenyl ring (from $V_L$F98G) and an indole ring (from $V_H$W110G), we split the JK43 structure in half: the left side consists of the phenyl ring and the ether linker, while the right side is comprised solely of the indole ring. We identified and purchased a series of JK43 analogs that vary either the left side of the compound, the right side, or both **(Figure 3.5A)**.

We then tested the degree to which each of these compounds rescues activity of the triple mutant, using the fluorescence quenching assay described earlier. At a concentration of 10 μM, we find that *none of the analogs* recovers fluorescein binding to the same extent as JK43 itself **(Figure 3.5B)**. This observation is consistent with our model in which the rescuing ligand precisely binds to the cavity left by the cavity-

forming mutations **(Figure 1.2)**. Further, this observation supports the utility of detailed structure-based modeling for identifying suitable ligands for a given cavity: our underlying hypothesis was that an optimal ligand would be one that precisely matches the three-dimensional arrangement of the deleted atoms, and the activity of JK43 relative to these close analogs supports this hypothesis.

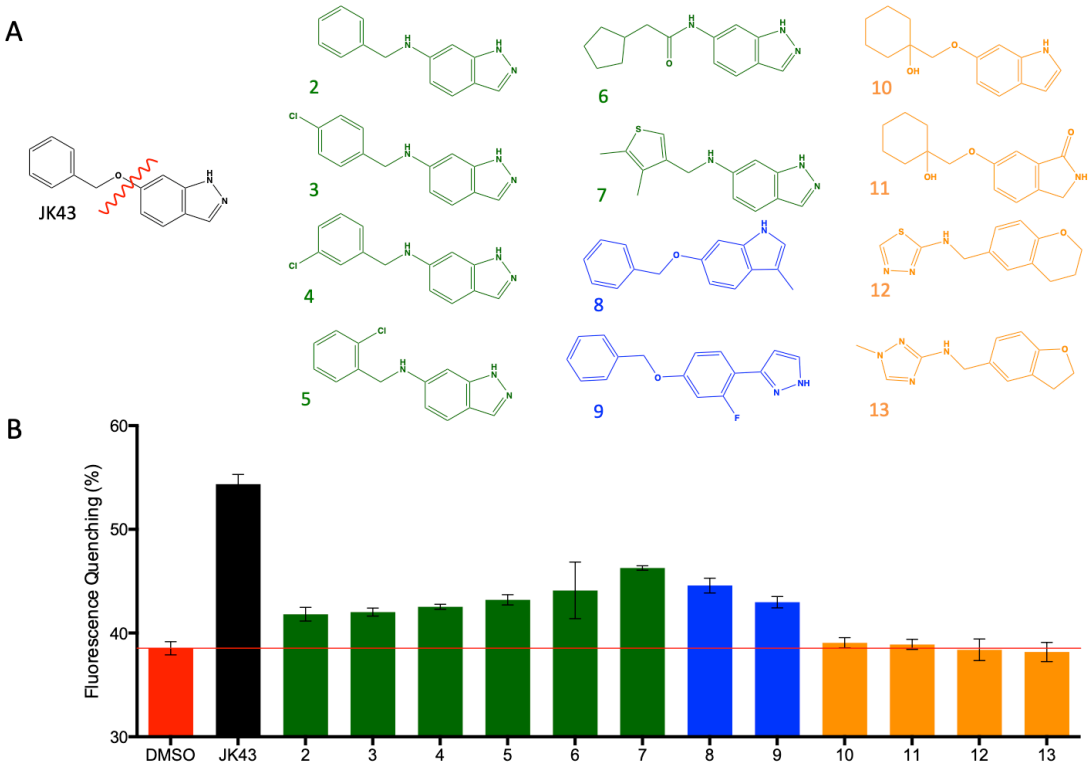


**Figure 3.5: JK43 rescues activity of the 4D5Flu triple mutant more effectively than its chemical analogs.** ***(A)*** The chemical structure of JK43 (*black*). Several analogs were selected for characterization by varying JK43's left side / linker (*green*), by varying JK43's right side (*blue*), or by varying both halves (*orange*). ***(B)*** At a concentration of 10 μM, JK43 rescues fluorescence quenching of the triple mutant more effectively than any of these chemical analogs. All data are presented as mean ± SEM, n = 8.

*Transferability of this designed switch*

A further important consequence of this hypothesis is that JK43 should also serve as a suitable rescuing ligand for the corresponding mutation in other antibodies (or at least, scFv's), provided they share the three deleted sidechains in precisely the same arrangement as in 4D5Flu. We therefore searched for the deleted constellation of atoms across a pair of antibody databases.

Starting with AbYsis (*60*), a database of antibody sequences, we find that all three residues comprising this constellation are present in 63% of the 150,697 total entries **(Figure 3.6A)**. Among antibodies with solved structure, as collected in the SAbDab database (*61*), we find that 78% of the 3,179 entries have this constellation in precisely the same geometry as 4D5Flu **(Figure 3.6B)**. Next, we extracted the constellations from the subset of SAbDab that has all 3 residues conserved; we then aligned these constellations against the original constellation from 4D5Flu using ROCS and found out that most are superposable. The fact that this constellation is so strongly conserved among antibody sequences and structures implies that – if transferrable – this triple mutant coupled with JK43 may serve as a very general means to modulate activity of scFv's, without the need to repeat the computational design calculations for (most) individual antibodies of interest.

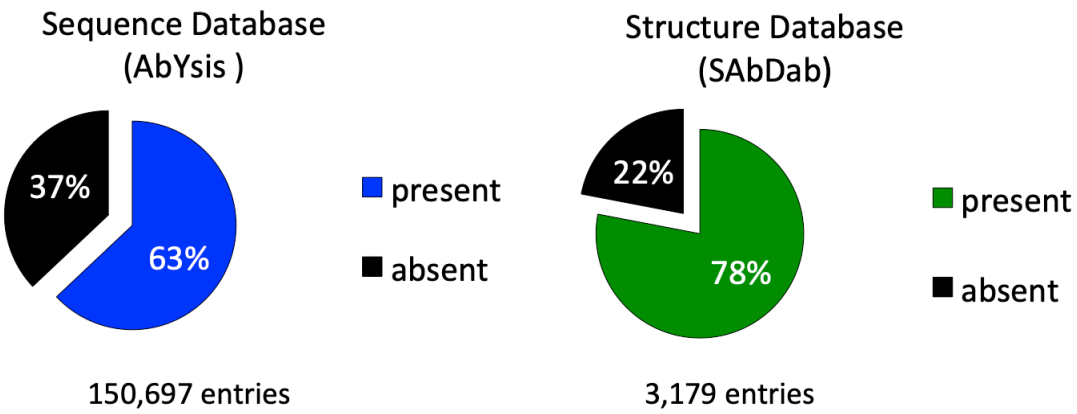


**Figure 3.6: Conservation of the designed constellation across antibody databases.** ***(A)*** In the antibody sequence database AbYsis, 63% of all entries harbor the three residues used in the designed constellation ($V_L$F98, $V_H$V37 and $V_H$W110). ***(B)*** In the antibody structure database SAbDab, 78% of all entries harbor these three residues in precisely the same conformation as in 4D5Flu

To test the transferability of our design, we selected another model scFv that shares this constellation. For ease of testing we chose another dye-binding scFv, this one raised to recognize 8-methoxypyrene-1,3,6-trisulfonic acid (MTPS) as its antigen. Although both fluorescein and MTPS are fluorescent dyes, their chemical structures are quite different **(Figure 3.7A)**. More importantly though, the scFv's that recognize these antigens are very different, sharing only 56% sequence identity outside of the serine-glycine linker **(Figure 3.7B)**. Unsurprisingly, given their different antigens, the CDR loops in both scFv's differ in both their lengths and their conformations **(Figure 3.7C)**. Despite these differences, however, both share the precise arrangement of the key residues that comprise our constellation **(Figure 3.7D)**, suggesting that the anti-MTPS scFv can also be modulated using precisely the same mutation/ligand pairing that we have already characterized using 4D5Flu.

To test this hypothesis, we generated the same triple mutant in the context of this anti-MTPS scFv, and used MTPS quenching as a measure of scFv binding. In the absence of JK43 this mutant scFv quenches MTPS with an $EC_{50}$ value of 1.0 μM; upon addition, however, the $EC_{50}$ value is reduced to 0.1 μM **(Figure 3.7E)**. Thus, our studies of this mutation/ligand pair in the context of 4D5Flu allowed us to rapidly generate an analogous switch in an unrelated antibody, by virtue of this structurally-conserved constellation. These results confirm that indeed this mutation/ligand pairing is indeed transferrable, provided that the scFv of interest harbors these conserved residues in precisely the same geometry as in 4D5Flu.

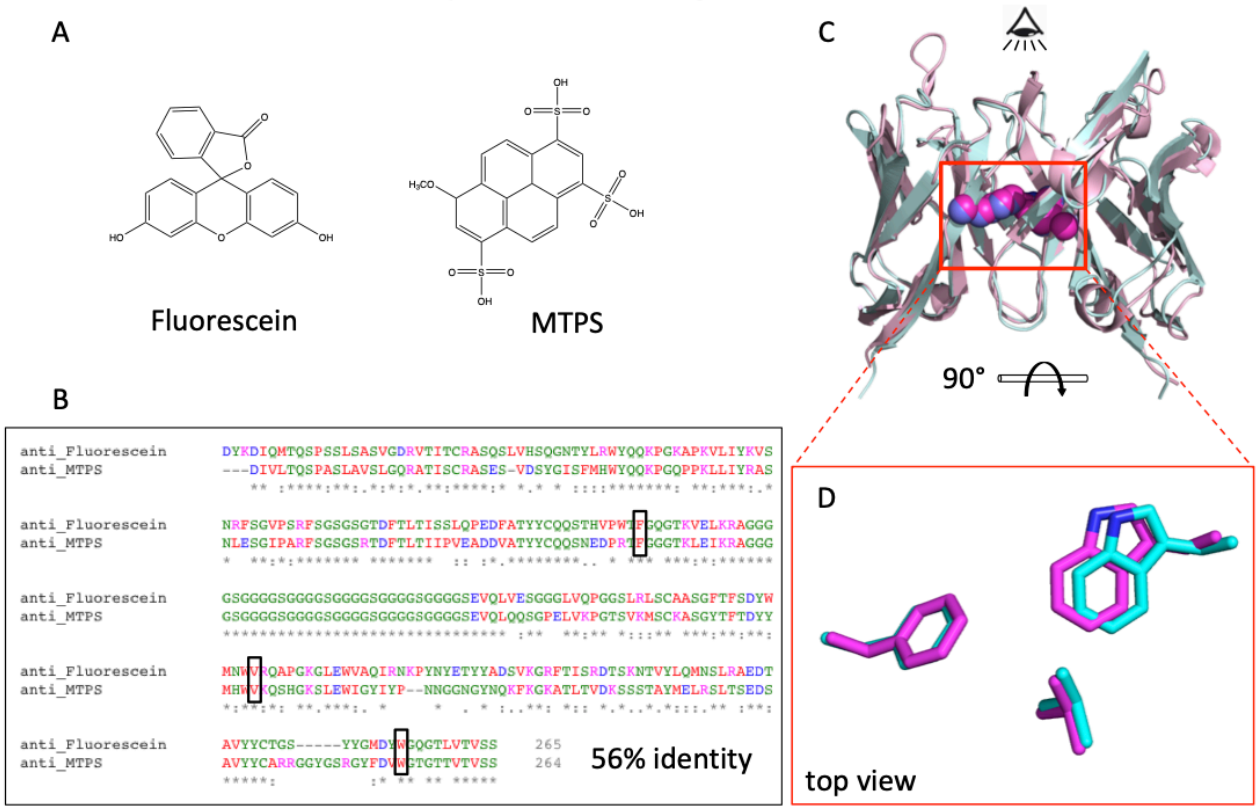
A
Fluorescein
MTPS
B
anti_Fluorescein DYKDIQMTQSPSSLSASVGDRVTITCRASQSLVHSQGNTYLRWYQQKPGKAPKVLIYKVS
anti_MTPS ---DIVLTQSPASLAVSLGQRATISCRASES-VDSYGISFMHWYQQKPGQPPKLLIYRAS
anti_Fluorescein NRFSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSTHVPWTFGQGTKVELKRAGGG
anti_MTPS NLESGIPARFSGSGSRTDFTLTIIPVEADDVATYYCQQSNEDPRTFGGGTKLEIKRAGGG
anti_Fluorescein GSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLSCAASGFTFSDYW
anti_MTPS GSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLQQSGPELVKPGTSVKMSCKASGYTFTDYY
anti_Fluorescein MNWVKQAPGKGLEWVAQIRNKPYNYETYYADSVKGRFTISRDTSKNTVYLQMNSLRAEDT
anti_MTPS MHWVKQSHGKSLEWIGYIYP--NNGGNGYNQKFKGKATLTVDKSSSTAYMELRSLTSEDS
anti_Fluorescein AVYYCTGS-----YYGMDYWGQGTLVTVSS 265
anti_MTPS AVYYCARRGGYGSRGYFDVWGTGTTVTVSS 264
56% identity
C
90°
D
top view
E
no JK43
100 μM JK43
MTPS quenching (%)
30
20
10
0
10
100
1,000
10,000
anti-MTPS scFv concentration (nM)

**Figure 3.7: JK43 rescues activity of the corresponding triple mutant in an unrelated scFv.** ***(A)*** The chemical structures of the two antigens of the characterized scFv's, fluorescein and MTPS. ***(B)*** The amino acid sequence alignment of 4D5Flu and the anti-MTPS scFv. The three residues that comprise the conserved constellation is highlighted with black boxes. ***(C)*** Despite their sequence divergence, these two scFv's have the same overall architecture. The structure of the anti-MTPS crystal structure (*salmon*) (PDB ID: 4NJ9) is aligned to the framework structure of 4D5Flu (*light blue*) is drawn from the crystal structure of the parent antibody, 4D5 (PDB ID: 1FVC). ***(D)*** Despite their sequence divergence, these two scFv's also have precisely the same structural arrangement of the sidechains that comprise the deleted constellation. ***(E)*** Monitoring MTPS quenching allows determination of the scFv's binding affinity in the presence or absence of JK43. In the absence of JK43, the anti-MTPS triple mutant quenches MTPS with $EC_{50}$ 1.0 μM. Upon addition of 100 μM JK43 the $EC_{50}$ is improved ten-fold, to 0.1 μM. All data are presented as mean ± SEM, n = 8.

## Chapter IV.
## Discussion and Future Steps

As a first step towards avoiding toxicity associated with therapeutic antibodies acting outside the desired spatial region, we have designed a method to control antibody activity using chemical biology.

The first study reporting allosteric control of antibody activity directly engineered allosteric sites into the CDR loop (H3) of an anti-fluorescein antibody, and demonstrated regulation of antigen-binding upon addition of $Zn^{2+}$ (*62*). This method has several drawbacks. First, the fact that the effector molecule ($Zn^{2+}$) is not bio-orthogonal precludes this strategy from being used for spatiotemporal control of an antibody in therapeutic settings. Second, the system lacks selectivity, as demonstrated by partial activation of the antibody in the presence of other divalent ions.

By contrast, envisioning future therapeutic applications, we were determined to create an allosteric antibody switch that would be both bio-orthogonal and selective. Looking back at our own previous work, we noted that chemical rescue of structure has been demonstrated as an alternative approach used to engineer allosteric protein switches, by connecting protein activity to the presence of an activating small molecule (*29-31*). As preparation for the current study, we also applied chemical rescue of structure in an antibody setting, by showing that the activity of the mutant antibody (FITC-E2 $V_H$ W118G) was restored upon addition of indole – albeit with an $EC_{50}$ of 220 μM (*32*), too high for potential therapeutic applications.

Here, to overcome the shortcomings of the single W→G design, we have expanded our strategy to include multiple-cavity forming mutations and screen a large

library of commercially-available compounds for re-activating small molecules. We have shown that addition of JK43 rescues the stability, structure, and conformational dynamics of the triple mutant: together, these rescue the scFv's affinity for antigen. Through the findings reported herein, we have shown that the mechanism of inactivation and rescue is modulated through changes in antibody stability, in agreement with many of our previous studies in assorted model systems (*29, 31, 32*). The present study is also consistent with results from previous model systems in that alternate ligands exhibit diminishing extents of rescue as the potential rescuing ligand become more dissimilar to the rescuing ligand that best fits the engineered cavity (*29, 31, 32*).

As hypothesized, we found that using a larger rescuing ligand (and by necessity, engineering a larger cavity into the protein) led to an improvement in binding affinity: JK43 has an $EC_{50}$ of 16 μM, as compared to 220 μM from indole. That said, we also anticipated that designing a larger cavity into the protein would reduce the background antigen binding (i.e. binding prior to addition of the rescuing ligand): we found this not to be the case. Moving forward, we plan to introduce additional destabilizing mutations designed to shift the equilibrium of the apo mutant scFv further towards the inactive state.

Already in these studies we have demonstrated the transferability of this design (triple mutant and JK43) from a fluorescein-binding antibody into an MTPS-binding antibody, even though the two antibodies only share 56% sequence identity. This shows that the overall sequence identity, in general, is less relevant to the transferability of the design, as long as the three residues that comprise this constellation are conserved in the antibody of interest. With respect to therapeutic antibodies already in the clinic, this

particular constellation/ligand pairing could find application in a number of candidates: among these, Ipilimumab (anti-CTLA-4, PDB ID: 5XJ3), Atezolimumab (anti-PD-L1, PDB ID: 5X8L), Nivolumab (anti-PD-1, PDB ID: 5GGR) and Adalimumab (anti-TNF-α, PDB ID: 3WD5). We did a structural alignment, comparing crystal structures of these four therapeutic antibodies against the crystal structure of our design framework, and confirmed that all three residues in the triple mutant design are superposable (**Figure 4.1**). We do note, however, that to date we have only tested this design in an scFv construct; further development may be necessary to use this approach in a Fab or IgG context, in which the dissociated (mutant) variable domains may be more restricted by the presence of the constant regions.

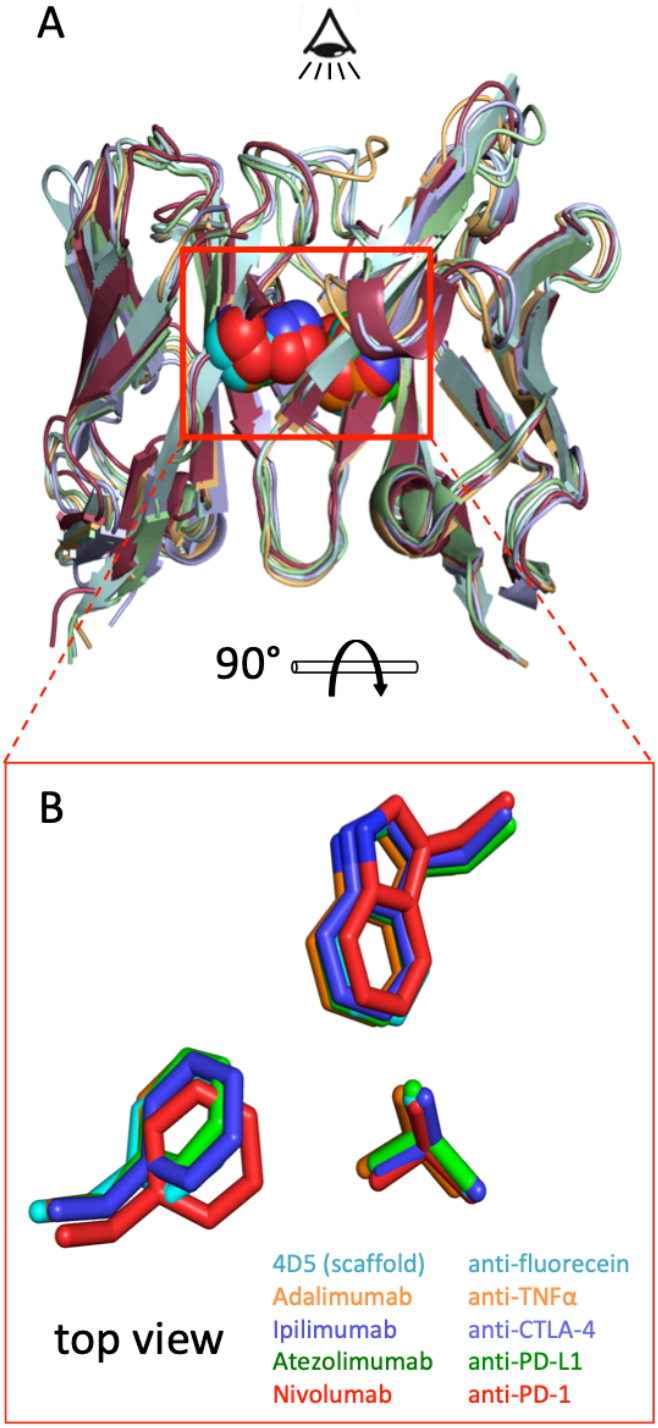
A
90°
B
top view
4D5 (scaffold)
Adalimumab
Ipilimumab
Atezolimumab
Nivolumab
anti-fluorecein
anti-TNFα
anti-CTLA-4
anti-PD-L1
anti-PD-1

**Figure 4.1: Triple mutant constellation is conserved in FDA-approved therapeutic antibodies.** ***(A)*** Despite their sequence divergence, all 4 therapeutic antibodies (Ipilimumab (anti-CTLA-4, PDB ID: 5XJ3), Atezolimumab (anti-PD-L1, PDB ID: 5X8L), Nivolumab (anti-PD-1, PDB ID: 5GGR) and Adalimumab (anti-TNF-α, PDB ID: 3WD5)) have the same overall architecture with the framework structure of 4D5Flu drawn from the crystal structure of the parent antibody, 4D5 (PDB ID: 1FVC). ***(B)*** Despite their sequence divergence, all 5 antibodies also have precisely the same structural arrangement of the sidechains that comprise the deleted constellation.

Thinking ahead, the use of a small molecule to control activity of a therapeutic antibody has unique advantages over traditional modalities. First, as compared to small molecule drugs, this approach maintains the tight and selective binding of therapeutic antibodies to avoid off-target interactions. At the same time, though, the fact that antibody activity relies on the presence of the small molecule activator allows for potential spatiotemporal control in new ways – including the fact that decay of activity may be determined by clearance of the small molecule, which is expected to have more rapid pharmacokinetics than the antibody.

In order to enhance this system beyond the capabilities of the current scheme, I envision drawing further from the chemical biology toolbox in two ways. First, additional spatial control could be achieved by exploiting unique conditions of the tumor microenvironment to further help reduce off-site activation of antibodies. This could be achieved through a pro-drug approach (*63*), in which labile protective groups are added to the activating ligand that prevent it from fitting into the designed cavity. Factors present in the tumor microenvironment, such as specific matrix metalloproteases, would then cleave the protecting group to reveal the original activating ligand. Second, to enhance the effective binding affinity of the rescuing ligand, I envision adapting the rescuing ligand such that it binds covalently to the designed antibody. This could be achieved by using site-specific non-natural amino acid incorporation (*64*) to introduce a bio-orthogonal and bio-compatible "clickable" group into the designed cavity (*65, 66*), and then providing the complementary functional group on the activating ligand. Together, I envision these two future developments will allow additional selectivity and sensitivity surpassing those of the current designed allosteric switch.

## References


1. M. S. Kinch, An overview of FDA-approved biologics medicines. *Drug Discov Today* **20**, 393-398 (2015).
2. S. M. Chiavenna, J. P. Jaworski, A. Vendrell, State of the art in anti-cancer mAbs. *J Biomed Sci* **24**, 15 (2017).
3. K. J. Arlotta, S. C. Owen, Antibody and antibody derivatives as cancer therapeutics. *Wiley Interdiscip Rev Nanomed Nanobiotechnol*, e1556 (2019).
4. L. M. Weiner, R. Surana, S. Wang, Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat Rev Immunol* **10**, 317-327 (2010).
5. J. Sprent, H. Kishimoto, The thymus and central tolerance. *Philos Trans R Soc Lond B Biol Sci* **356**, 609-616 (2001).
6. K. A. Hogquist, T. A. Baldwin, S. C. Jameson, Central tolerance: learning self-control in the thymus. *Nat Rev Immunol* **5**, 772-782 (2005).
7. P. Waterhouse *et al.*, Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. *Science* **270**, 985-988 (1995).
8. E. A. Tivol *et al.*, Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* **3**, 541-547 (1995).
9. J. F. Brunet *et al.*, A new member of the immunoglobulin superfamily--CTLA-4. *Nature* **328**, 267-270 (1987).
10. D. M. Pardoll, The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* **12**, 252-264 (2012).

11. C. Fellner, Ipilimumab (yervoy) prolongs survival in advanced melanoma: serious side effects and a hefty price tag may limit its use. *Pharm Ther* **37**, 503-530 (2012).

12. Y. Agata *et al.*, Expression of the PD-1 antigen on the surface of stimulated mouse T and B lymphocytes. *Int Immunol* **8**, 765-772 (1996).

13. Y. Ishida, Y. Agata, K. Shibahara, T. Honjo, Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J* **11**, 3887-3895 (1992).

14. H. Dong, G. Zhu, K. Tamada, L. Chen, B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion. *Nat Med* **5**, 1365-1369 (1999).

15. K. Bardhan, T. Anagnostou, V. A. Boussiotis, The PD1:PD-L1/2 Pathway from Discovery to Clinical Implementation. *Front Immunol* **7**, 550 (2016).

16. H. A. Blair, Atezolizumab: A Review in Previously Treated Advanced Non-Small Cell Lung Cancer. *Target Oncol* **13**, 399-407 (2018).

17. R. Sundar, B. C. Cho, J. R. Brahmer, R. A. Soo, Nivolumab in NSCLC: latest evidence and clinical potential. *Ther Adv Med Oncol* **7**, 85-96 (2015).

18. A. Winer, J. N. Bodor, H. Borghaei, Identifying and managing the adverse effects of immune checkpoint blockade. *J Thorac Dis* **10**, S480-S489 (2018).

19. G. Myers, Immune-related adverse events of immune checkpoint inhibitors: a brief review. *Curr Oncol* **25**, 342-347 (2018).

20. D. B. Johnson, C. Peng, J. A. Sosman, Nivolumab in melanoma: latest evidence and clinical potential. *Ther Adv Med Oncol* **7**, 97-106 (2015).

21. L. C. Cappelli, A. A. Shah, C. O. Bingham, 3rd, Immune-Related Adverse Effects of Cancer Immunotherapy- Implications for Rheumatology. *Rheum Dis Clin North Am* **43**, 65-78 (2017).

22. S. L. Topalian, C. G. Drake, D. M. Pardoll, Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* **27**, 450-461 (2015).

23. M. Ostermeier, Engineering allosteric protein switches by domain insertion. *Protein Eng Des Sel* **18**, 359-364 (2005).

24. M. Kojima, K. Ayabe, H. Ueda, Importance of terminal residues on circularly permutated Escherichia coli alkaline phosphatase with high specific activity. *J Biosci Bioeng* **100**, 197-202 (2005).

25. A. Galarneau, M. Primeau, L. E. Trudeau, S. W. Michnick, Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat Biotechnol* **20**, 619-622 (2002).

26. G. S. Baird, D. A. Zacharias, R. Y. Tsien, Circular permutation and receptor insertion within green fluorescent proteins. *Proc Natl Acad Sci U S A* **96**, 11241-11246 (1999).

27. S. B. Kim, Y. Otani, Y. Umezawa, H. Tao, Bioluminescent indicator for determining protein-protein interactions using intramolecular complementation of split click beetle luciferase. *Anal Chem* **79**, 4820-4826 (2007).

28. B. W. Matthews, L. Liu, A review about nothing: are apolar cavities in proteins really empty? *Protein Sci* **18**, 494-502 (2009).

29. Y. Xia *et al.*, The designability of protein switches by chemical rescue of structure: mechanisms of inactivation and reactivation. *J Am Chem Soc* **135**, 18840-18849 (2013).

30. K. Deckert, S. J. Budiardjo, L. C. Brunner, S. Lovell, J. Karanicolas, Designing allosteric control into enzymes by chemical rescue of structure. *J Am Chem Soc* **134**, 10055-10060 (2012).

31. S. J. Budiardjo *et al.*, Full and Partial Agonism of a Designed Enzyme Switch. *ACS Synth Biol* **5**, 1475-1484 (2016).

32. C. E. Kaiser *et al.*, Modulating Antibody Structure and Function through Directed Mutations and Chemical Rescue. *ACS Synth Biol* **7**, 1152-1162 (2018).

33. R. D. Smith, A. L. Engdahl, J. B. Dunbar, Jr., H. A. Carlson, Biophysical limits of protein-ligand binding. *J Chem Inf Model* **52**, 2098-2106 (2012).

34. J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, R. G. Coleman, ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* **52**, 1757-1768 (2012).

35. C. A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* **1**, 337-341 (2004).

36. J. L. Dahlin *et al.*, PAINS in the assay: chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J Med Chem* **58**, 2091-2113 (2015).

37. J. B. Baell, Feeling Nature's PAINS: Natural Products, Natural Product Drugs, and Pan Assay Interference Compounds (PAINS). *J Nat Prod* **79**, 616-628 (2016).

38. J. Baell, M. A. Walters, Chemistry: Chemical con artists foil drug discovery. *Nature* **513**, 481-483 (2014).

39. J. L. Dahlin, M. A. Walters, The essential roles of chemistry in high-throughput screening triage. *Future Med Chem* **6**, 1265-1290 (2014).

40. J. L. Dahlin, J. Baell, M. A. Walters, in *Assay Guidance Manual,* G. S. Sittampalam *et al.*, Eds. (Bethesda (MD), 2004).

41. P. C. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, M. T. Stahl, Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model* **50**, 572-584 (2010).

42. P. C. Hawkins, A. G. Skillman, A. Nicholls, Comparison of shape-matching and docking as virtual screening tools. *J Med Chem* **50**, 74-82 (2007).

43. J. Kirchmair *et al.*, How to optimize shape-based virtual screening: choosing the right query and including chemical information. *J Chem Inf Model* **49**, 678-692 (2009).

44. G. B. McGaughey *et al.*, Comparison of topological, shape, and docking methods in virtual screening. *J Chem Inf Model* **47**, 1504-1519 (2007).

45. V. I. Perez-Nueno *et al.*, Comparison of ligand-based and receptor-based virtual screening of HIV entry inhibitors for the CXCR4 and CCR5 receptors using 3D ligand shape matching and ligand-receptor docking. *J Chem Inf Model* **48**, 509-533 (2008).

46. A. Leaver-Fay *et al.*, ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods in enzymology* **487**, 545-574 (2011).

47. T. Baumgarten, A. J. Ytterberg, R. A. Zubarev, J. W. de Gier, Optimizing Recombinant Protein Production in the Escherichia coli Periplasm Alleviates Stress. *Appl Environ Microbiol* **84**, (2018).

48. P. Oelschlaeger *et al.*, Identification of factors impeding the production of a single-chain antibody fragment in Escherichia coliby comparing in vivo and in vitro expression. *Appl Microbiol Biotechnol* **61**, 123-132 (2003).

49. D. G. Gibson *et al.*, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345 (2009).

50. S. Jung, A. Pluckthun, Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng* **10**, 959-966 (1997).

51. D. Matulis, J. K. Kranz, F. R. Salemme, M. J. Todd, Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* **44**, 5258-5266 (2005).

52. L. E. Kay, D. A. Torchia, A. Bax, Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979 (1989).

53. N. A. Farrow *et al.*, Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984-6003 (1994).

54. L. Nieba, A. Honegger, C. Krebber, A. Pluckthun, Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng* **10**, 435-444 (1997).

55. W. D. Mallender, J. Carrero, E. W. Voss, Jr., Comparative properties of the single chain antibody and Fv derivatives of mAb 4-4-20. Relationship between interdomain interactions and the high affinity for fluorescein ligand. *The Journal of biological chemistry* **271**, 5338-5346 (1996).

56. P. Cimmperman *et al.*, A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys J* **95**, 3222-3231 (2008).

57. R. Zhang, F. Monsma, Fluorescence-based thermal shift assays. *Curr Opin Drug Discov Devel* **13**, 389-402 (2010).

58. C. J. Layton, H. W. Hellinga, Thermodynamic analysis of ligand-induced changes in protein thermal unfolding applied to high-throughput determination of ligand affinities with extrinsic fluorescent dyes. *Biochemistry* **49**, 10831-10841 (2010).

59. G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, A. Bax, Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**, 5269-5278 (1992).

60. M. B. Swindells *et al.*, abYsis: Integrated Antibody Sequence and Structure-Management, Analysis, and Prediction. *J Mol Biol* **429**, 356-364 (2017).

61. J. Dunbar *et al.*, SAbDab: the structural antibody database. *Nucleic Acids Res* **42**, D1140-1146 (2014).

62. O. Khersonsky, S. J. Fleishman, Incorporating an allosteric regulatory site in an antibody through backbone design. *Protein Sci* **26**, 807-813 (2017).

63. J. Rautio, N. A. Meanwell, L. Di, M. J. Hageman, The expanding role of prodrugs in contemporary drug design and development. *Nat Rev Drug Discov* **17**, 559-587 (2018).

64. K. Lang, J. W. Chin, Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem Rev* **114**, 4764-4806 (2014).

65. E. M. Sletten, C. R. Bertozzi, From mechanism to mouse: a tale of two bioorthogonal reactions. *Acc Chem Res* **44**, 666-676 (2011).

66. N. J. Agard, J. A. Prescher, C. R. Bertozzi, A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J Am Chem Soc* **126**, 15046-15047 (2004).