Exploring molecular recognition: de novo design of doxorubicin binding proteins

by
Stephanie Ouchida

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:

Signed by:
582727E949C7441...

William DeGrado
Chair

DocuSigned by:
Michelle Arkin

Michelle Arkin

DocuSigned by:
0E22FCF3CD53444...

Alan Ashworth

_____

_____
Committee Members

DEDICATION

To every version of myself who questioned whether she could do it.

You did it.

ACKNOWLEDGEMENTS

This work would not have been possible without the incredible support network I have been fortunate enough to have throughout my scientific career.

First and foremost, I want to thank my advisor, Bill DeGrado. I feel incredibly lucky to have trained under a true trailblazer in protein design. I will miss clicking through PyMOL with you and learning to see proteins through your eyes. You encouraged me through the many ups and downs of research, always reminding me to keep going. Beyond your scientific wisdom, you showed me what it means to be both an exceptional scientist and a kind, thoughtful person.

To the DeGrado Lab, thank you for creating such an amazing community. Hyunil Jo, our second Bill, thank you for always being there to help and for keeping the lab running. A special thank you to Lee Schnaider, Ethel Tackie-Yarboi, Katie Hatstat, Sam Mann, Hyunjun Yang, Huong Kratochvil, Ian Bakanas, and Maggie Horst. Having such incredible postdocs to learn from made all the difference. Thank you for being such incredible scientists, mentors, and role models. Thank you as well to the rest of the past and present DeGrado Lab members: Kehan Chen, Yuda Chen, Henry Pan, Sophia Tan, Sagar Bhattacharya, Nam Hyeong Kim, Rose Yang, Henry Scott, Angelika Arada, Nicholas Freitas, Giovanni Aviles, Haoyu Fan, Sungwon Jung, Dylan Boselli, Lorena Šimunić, Yibing Wu, Sam Schneider, Lei Lu, Matthew Chalkley, Hailey Wallace, Rian Kormos, and Sarah Yorke. Thank you all for your mentorship, scientific discussions, yap sessions, and friendship, and for making the lab such a wonderful place to grow as a scientist. Finally, William Fish DeGrado, our lab betta fish, thank you for giving me a reason to show up during one of the hardest periods of graduate school.

To the Ashworth Lab, especially Xuxu Gou and Morgan Diolaiti, thank you for being wonderful collaborators and for helping bring the translational aspect of this project to life. To my committee members, Michelle Arkin and Alan Ashworth, thank you for your guidance, thoughtful feedback, and support throughout my PhD.

I'm incredibly grateful for the friends I made at UCSF, who made graduate school brighter and far less lonely. Thank you to Amy Flis, Jonathan Borowsky, Jonathan Zhang, Matthew Smith, Haley Ogasawara, and Colton Sanders for your friendship throughout graduate school. Andrew Ecker, thank you for being my cohort bestie through it all. Your love for chemistry and our countless scientific conversations reminded me why I love science. I'll never forget that the doxorubicin sandwich idea came from one of our spontaneous conversations in CVRI. Thank you for your encouragement and friendship through every high and low of graduate school. Katie Holland, thank you for being one of the greatest gifts graduate school gave me. From our weekly Friday hangouts to countless conversations about life and science, thank you for always showing up for me. You helped me through some of the hardest moments of my PhD.

I've been incredibly fortunate to have many mentors outside of UCSF. Thank you, Yi Lu, for giving me my first opportunity to do research as an undergraduate. In Dr. Lu's lab, I was fortunate to have an incredible graduate student mentor, Sudharsan Dwaraknath. Thank you for pushing me, introducing me to the world of proteins, and sparking my love for research. Donnie Mackenzie, thank you for your mentorship during my time at BioMarin. Thank you to James Spudich, who went from being one of my scientific heroes to someone I had the privilege of working with during both my

undergraduate years and gap year. Makenna Morck was another amazing mentor I met during undergrad who later became a wonderful friend. You showed me there are many ways to build a meaningful career in science. Thank you also to Darshan Trivedi and Sia Karabina for our time together at Kainomyx. A special thank you to Alice Racca, who began as a mentor and became a deeply cherished friend. Mei Hong, thank you for always taking the time to check in on me, offer guidance, and make sure I was progressing both scientifically and personally.

I also want to thank my therapists, Alexandra and Jacqueline, for helping me navigate one of the most challenging periods of my life and for teaching me how to manage burnout, care for my mental health, and persevere through graduate school.

Finally, I want to thank the people who have shaped me beyond the scientist I've become.

Ayan and Melody, I'm so grateful our paths crossed when they did. Thank you for sharing your love of EDM and for creating a space where I could forget about graduate school and simply enjoy life. Jeffrey and Stella, despite the distance to Chicago, thank you for always cheering me on and for staying close throughout graduate school.

Shreya, my best friend since seventh grade, it's still surreal that we get to go through this crazy PhD journey together. Middle school us would be so proud of how far we've come. Thank you for being one of my biggest supporters through every stage of life. We're really doing it, girlypop.

Tyler and Momo, being your older sister has always motivated me to become the best version of myself because I've always wanted to set a good example for you. Watching you both grow into such kind and incredible people has made me so proud,

and I couldn't ask for better siblings. Tofu, you are a constant source of joy and light. Bupbup, I wish you had made it to the end of graduate school, but thank you for keeping me company in the tidepod.

Mom, thank you for always picking up the phone. No matter how stressed, overwhelmed, or discouraged I was, you always listened to me cry, vent, and doubt myself, and somehow, by the end of every conversation, you reminded me that I could keep going. Watching you work so hard and thrive in a male-dominated field taught me that there is no limit to what a woman can accomplish with determination and grit. Dad, thank you for nurturing my scientific curiosity from such a young age. Watching you go through nursing school sparked my own love for biomedical science and ultimately set me on the path that led me here. Together, you both taught me to dream big and work hard enough to make those dreams a reality. Everything I've accomplished is built on the foundation of your love, sacrifices, and unwavering support.

Finally, Steven. There are no words that can fully express how grateful I am for you. Your family welcomed me with open arms and made Chicago feel like a second home, and I'm very thankful for their love and support. Despite being long distance for nearly all of graduate school, I always felt your love and support. Thank you for standing beside me through every high and low of graduate school, celebrating every success, helping me through every setback, and listening to every failed experiment, exciting result, and existential graduate school crisis. Your love, patience, unwavering support, and ability to make me laugh gave me the strength to keep going. I truly could not have done this without you.

Contributions

This dissertation was supervised by Dr. William F. DeGrado. Additional guidance was provided by Dr. Michelle Arkin and Dr. Alan Ashworth

Chapter 2 is reproduced in full from the following manuscript currently in preparation:

Ouchida S, Horst M, Gou X, Bakanas I, Diolaiti M, Ashworth A, DeGrado WF**.** *De novo* design of a novel protein fold for small-molecule binding through aromatic π stacking. Manuscript in preparation.

Exploring molecular recognition: *de novo* design of doxorubicin binding proteins

Stephanie Ouchida

ABSTRACT

The ability of proteins to recognize and bind small molecules underlies numerous biological processes and has broad applications in biotechnology and medicine. In this dissertation, I investigate the *de novo* design of proteins that recognize chemically complex small molecules. Chapter 1 introduces the principles of molecular recognition and *de novo* protein design that motivate this work. Chapter 2 describes the computational design, experimental optimization, and structural validation of *de novo* proteins that bind the anthracycline chemotherapeutic doxorubicin, resulting in a compact 85-residue protein with nanomolar affinity and a novel protein fold. Chapter 3 explores the translational potential of these proteins through engineering of an albumin-binding fusion that preserves doxorubicin binding while introducing recognition of human serum albumin. Together, these studies demonstrate that *de novo* protein design can create structurally novel proteins with programmable molecular recognition and establish a foundation for future therapeutic and biotechnological applications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABD – albumin-binding domain

ABDCon – consensus albumin-binding domain

AFDB – AlphaFold Database

ALS – Advanced Light Source

AUC – area under the curve

BSA – bovine serum albumin

CD – circular dichroism

CI – confidence interval

COMBS – Convergent Motifs for Binding Sites

DFT – density functional theory

DMSO – dimethyl sulfoxide

DNA – deoxyribonucleic acid

DSSP – Define Secondary Structure of Proteins

ESMFold – Evolutionary Scale Modeling Fold

HRP – horseradish peroxidase

HSA – human serum albumin

IC50 – half-maximal inhibitory concentration

IPTG – isopropyl β-D-1-thiogalactopyranoside

LB – Luria-Bertani medium

MASTER – Method of Accelerated Search for Tertiary Ensemble Representatives

MRE – mean residue ellipticity

Ni-NTA – nickel–nitrilotriacetic acid

NRSA – National Research Service Award

OD600 – optical density at 600 nm

ORCA – Orca quantum chemistry software

PBS – phosphate-buffered saline

PDB – Protein Data Bank

pLDDT – predicted local distance difference test

ProteinMPNN – protein message-passing neural network

PyMOL – Python Molecular Graphics System

RFdiffusionAA – RFdiffusion All-Atom

RFU – relative fluorescence units

RMS – root mean square

RMSD – root-mean-square deviation

SASA – solvent-accessible surface area

SDS-PAGE – sodium dodecyl sulfate–polyacrylamide gel electrophoresis

SEM – standard error of the mean

SMILES – Simplified Molecular Input Line Entry System

TEV – tobacco etch virus protease

UV – ultraviolet

XDS – X-ray Detector Software

LIST OF SYMBOLS

α – alpha

β – beta

π – pi

π–π – pi-pi

Å – angstrom

Å² – square angstrom

°C – degrees Celsius

μg – microgram

μL – microliter

μM – micromolar

mL – milliliter

mM – millimolar

nM – nanomolar

Δ – change in

ΔRFU – change in relative fluorescence units

ΔSASA – change in solvent-accessible surface area

K_D – equilibrium dissociation constant

IC_{50} – half-maximal inhibitory concentration

OD_{600} – optical density at 600 nm

R_{free} – crystallographic free R factor

R_{merge} – crystallographic merging R factor

σ – standard deviation; electron density contour level (context dependent)

CHAPTER 1 Introduction

Proteins are highly versatile molecular machines that perform a vast range of biological functions, from catalyzing chemical reactions to mediating cellular signaling and molecular recognition.[1–4] Their functional diversity arises from their ability to adopt precise three-dimensional structures and dynamic conformational states encoded by their amino acid sequences.[4–7] This intimate relationship between sequence, structure, and function has long made proteins a central target for molecular engineering, with goals ranging from understanding fundamental biological mechanisms to developing new therapeutics, catalysts, and biomolecular tools.[8,9] In recent years, *de novo* protein design has emerged as an approach that seeks to construct entirely new protein sequences that fold into predetermined structures and perform desired functions without relying on naturally occurring templates.[10,11] By constructing proteins from first principles instead of modifying existing scaffolds, *de novo* protein design expands the accessible landscape of protein structures and functions and enables proteins to be engineered with properties tailored to specific applications.[10,12]

De novo protein design represents a stringent test of our understanding of protein biophysics. Unlike traditional protein engineering, which modifies existing proteins, *de novo* design seeks to construct proteins from first principles by designing amino acid sequences that fold into predetermined structures.[11,12] This remains a formidable challenge because of the vastness of sequence space: even a protein of 100 amino acids has approximately 20^{100} possible sequences.[11,12] Early efforts relied on simplified physical models and manual design strategies that established fundamental design principles, particularly the importance of hydrophobic patterning in directing

protein folding.[13–17] The incorporation of computational methods into *de novo* protein design enabled more systematic exploration of sequence–structure relationships and culminated in the development of α3D, the first *de novo* globular protein whose experimentally determined structure closely matched its design model.[18,19] Subsequent advances in backbone generation, energy-based optimization, and fragment-based design expanded the diversity and complexity of accessible protein architectures.[20] A landmark example was TOP7, the first *de novo* protein with a novel fold whose experimentally determined structure closely matched its design model.[21] These advances established *de novo* protein design as a robust approach for constructing stable protein structures, but extending these successes to functional proteins capable of precise molecular recognition has remained a considerably greater challenge.

The *de novo* design of proteins that bind small molecules remains a major challenge because successful molecular recognition requires precise specification of both the overall protein scaffold and the atomic interactions that define ligand binding.[10,22] Unlike protein–protein interactions, which are typically stabilized by extensive binding interfaces, small molecules are recognized through a limited number of highly specific interactions within compact binding pockets.[23] Successful design therefore requires the precise organization of complementary shape, hydrogen-bonding interactions, electrostatics, and solvation while maintaining the overall stability of the protein scaffold.[23,24] These challenges are further compounded for structurally complex ligands that present multiple functional groups requiring coordinated molecular recognition. A particularly challenging design target is doxorubicin, a widely used anthracycline chemotherapeutic.[25] Doxorubicin has a rigid planar anthracycline scaffold

with multiple hydroxyl, carbonyl, and amino substituents, presenting competing opportunities for hydrogen bonding and aromatic interactions.[26] Successful design of proteins that recognize molecules of this complexity provides a foundation for applications in biosensing, drug delivery, and therapeutic regulation.[10,22]

This work was carried out during a period of exceptionally rapid development in computational protein design, particularly in machine learning–based methods for backbone generation, sequence design, and structure prediction.[27–33] Consequently, rather than relying on a single static computational workflow, the design strategy evolved continuously as new methodologies became available. The initial computational pipeline constructed ligand-binding sites by placing COMBS-derived hydrogen-bonding motifs within parameterized helical bundle scaffolds looped together using MASTER.[24,34,35] Protein sequences were then generated with Rosetta flexible-backbone design, and sequences were predicted using ESMFold.[24,33,36,37] Although this framework provided precise control over ligand geometry, scaffold diversity was inherently constrained by the limited range of parameterized backbone architectures.

RFdiffusion, a diffusion-based generative model for protein backbone generation, dramatically expanded the accessible structural space by generating diverse protein backbones around predefined binding motifs.[30] ProteinMPNN, a graph neural network trained for inverse protein folding, was integrated into sequence design, substantially accelerating the process.[29] Structure prediction advanced through protein language model–based methods including ESMFold, OmegaFold, and RaptorX-Single.[31–33] Despite these advances, each method considered only the protein. As a result, structural diversity and expression improved, but none of the resulting designs exhibited

measurable ligand binding. Ligand-conditioned models, including RFdiffusionAA and LigandMPNN, subsequently incorporated ligand atoms into backbone generation and sequence design.[27,28] Structure prediction advanced further with multimolecular models such as Chai-1, which predict protein–ligand complexes directly.[38] These advances enabled a computational workflow in which backbone generation, sequence design, and structure prediction all explicitly considered the target ligand, ultimately leading to the successful design of doxorubicin-binding proteins described in Chapter 2.

Chapter 2 describes the computational design, experimental characterization, and structural validation of *de novo* proteins that bind doxorubicin with high affinity. Chapter 3 explores the translational potential of these proteins through preliminary studies toward therapeutic applications, highlighting ongoing work and remaining challenges. Together, these studies illustrate the potential of *de novo* protein design to address fundamental questions in molecular recognition while establishing a foundation for future therapeutic applications.

CHAPTER 2 *De novo* design of a novel protein fold for small-molecule binding through aromatic π stacking

2.1 ABSTRACT

The *de novo* design of proteins that bind chemically complex small molecules has broad chemical and biological implications, but strategies typically rely on a small set of protein scaffolds and require extensive experimental screening. Here, we computationally designed proteins around a minimal aromatic π-stacking motif to bind the anthracycline anticancer drug doxorubicin. Experimental characterization of twelve proteins revealed a μM doxorubicin binder; two additional design cycles improved scaffold stability and binding affinity to yield an 85-residue protein that binds doxorubicin with a dissociation constant of 85 nM. An X-ray crystal structure of the protein-drug complex confirmed the accuracy of the designed π-π stacking interactions. The designed protein could act to protect cultured cells from doxorubicin-induced cytotoxicity. Unlike previous ligand-binding protein designs based on repeat proteins or naturally occurring folds, the designed protein adopts a previously unobserved 5-helix globular fold, indicating that a broader space of folded, functional proteins exists even for compact tertiary structures smaller than 100 residues. These results demonstrate that motif-guided generative protein design can discover compact *de novo* protein folds capable of high-affinity recognition of chemically complex small molecules.

2.2 INTRODUCTION

The *de novo* design of proteins that bind small molecules remains a central challenge in chemistry because successful molecular recognition requires simultaneous specification of both the overall protein scaffold and the atomic-level interactions that define ligand binding.[39–41] A growing collection of design strategies has enabled the *de novo* generation of proteins that bind increasingly diverse small molecules, with a broad range of structures and functional properties.[28,37,42–52] Many successful binders have been identified through extensive experimental screening,[27,28,44,47,50,53] whereas only a few studies have reported high computational hit rates.[37,42,43,48] Despite this progress, experimentally determined structures of *de novo* proteins bound to small molecule ligands remains largely limited to predetermined scaffolds including parameterized helical bundles and repeat proteins[37,42,48,52] and NTF2-like folds.[43,47] Extending *de novo* design to discover new protein architectures capable of high-affinity small molecule recognition therefore remains an important opportunity for expanding both the scope of molecular recognition and the diversity of functional protein folds.[39,40]

Doxorubicin (1, Fig. 2.1A) is a therapeutically important small molecule whose clinical utility is limited by cumulative, dose-dependent cardiotoxicity.[54–56] Although liposomal formulations reduce the systemic toxicity of many chemotherapeutics, their relatively large size and heterogeneous biodistribution can limit penetration into target tissues, particularly solid tumors, necessitating the development of alternative drug sequestration and delivery strategies.[57–59] Compact *de novo* proteins engineered to bind small molecules could provide a complementary platform for these applications owing to their small size, precisely defined molecular composition, and programmable molecular recognition properties.[40,60–62]

From a design perspective, doxorubicin presents a complex small molecule target. The ligand combines a rigid planar anthracycline scaffold with multiple hydroxyl, carbonyl, and amino substituents, presenting numerous competing opportunities for hydrogen bonding and aromatic stacking without an obvious dominant recognition motif.[63,64] To facilitate design we prioritized recognition of the anthracycline ring, since analysis of crystal structures of anthracycline-bound proteins revealed a recurring interaction in which aromatic residues pack against the planar anthracycline core while surrounding residues accommodate the remaining substituents.[65,66] We therefore expected that pi-pi interactions could establish the overall binding geometry while allowing secondary interactions responsible for affinity and specificity to be optimized independently. Recent advances in generative protein design make this strategy feasible by enabling diverse protein scaffolds to be generated around predefined protein–ligand interaction motifs.[28,49,67]

Here we combine motif-guided binding-site design with RFdiffusion All-Atom,[28] LigandMPNN,[27] Rosetta, and structure-prediction-guided filtering to generate *de novo* proteins that bind doxorubicin (Fig 1B). Starting from an interaction motif derived from a crystal structure of an anthracycline-bound protein, we identified a low-micromolar binders, which could be rapidly optimized to yield an 85-residue protein that binds doxorubicin with nanomolar affinity. These results are noteworthy because there are few examples of the design of small molecule binders which have been validated through the determination of high-resolution structures.[37,42,43,47,48] Moreover, our results address the question of the extent to which the space of folded, functional structures have been exhaustively sampled during evolution.[68–70] There are only a limited number of ways

that secondary structures can be arranged to create a folded structure, which becomes successively smaller as the size of the protein decreases.[71,72] While novel structures can be designed in small proteins[73,74], and small proteins can be designed to bind to protein surfaces,[75–77] it has been unclear whether Nature has sampled the universe of small proteins (< 90 residues) that are stable, cooperatively folded and capable of binding small molecules within a deeply buried pocket. The design of doxorubicin binders, termed DoxSands, shows that the universe of small, entirely novel miniaturized small molecule proteins has not been exhaustively discovered. This is particularly important because proteins of this size can be readily accessed by chemical synthesis, enabling incorporation of noncanonical amino acids and backbone modifications that are difficult to achieve biosynthetically.[78–83] These capabilities provide opportunities to engineer properties including proteolytic stability, pharmacokinetics, and immune recognition while introducing molecular functions beyond those available to the genetically encoded amino acids.[62,84–89] Thus, it should be also be possible to achieve entirely new functions that extend well beyond biology.

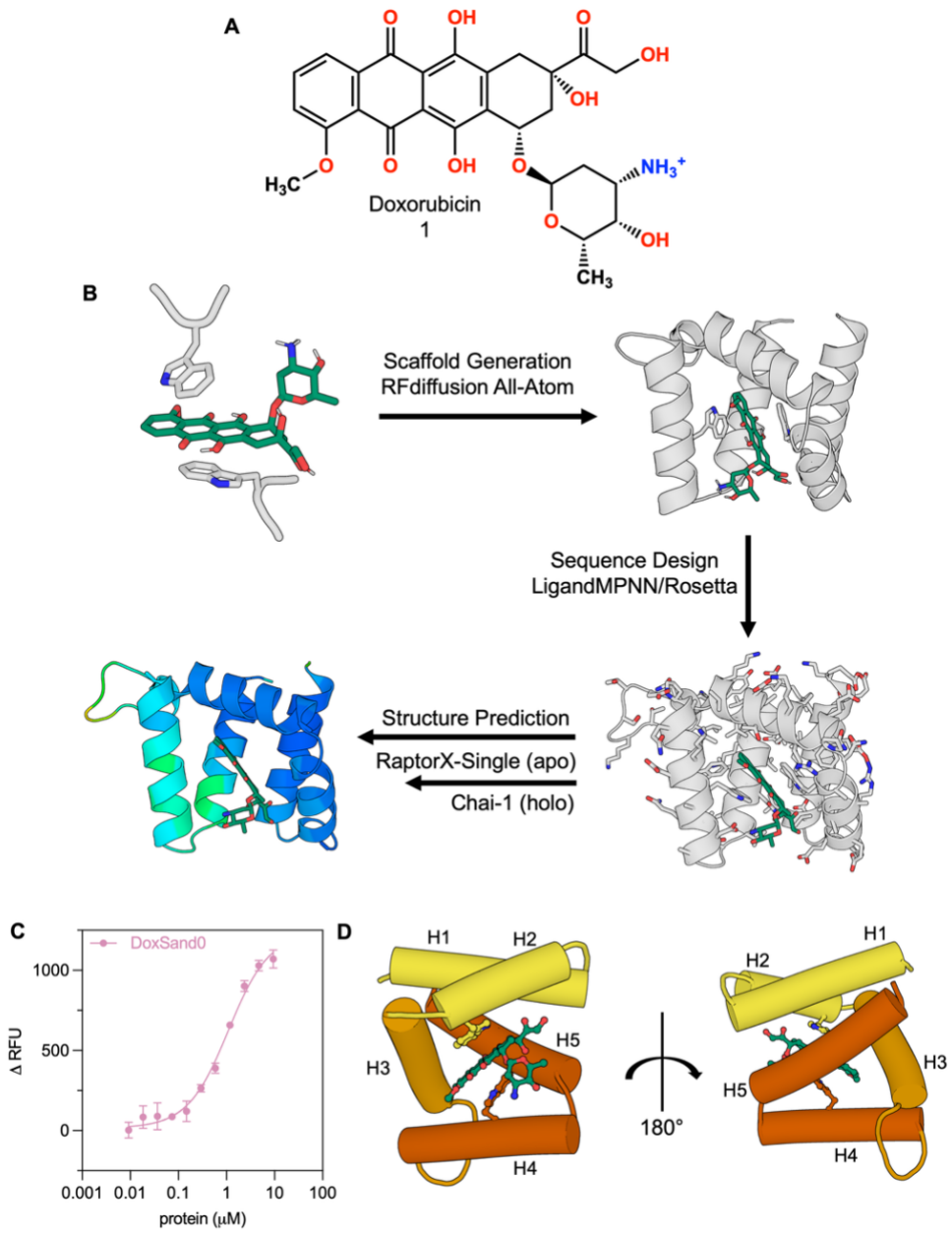

Figure 2.1 Computational workflow for the design of *de novo* doxorubicin-binding proteins.
(A) Chemical structure of doxorubicin. (B**)** An initial protein–ligand complex containing doxorubicin was used as input for backbone generation with RFdiffusion All-Atom. Protein sequences were designed with LigandMPNN and Rosetta FastRelax, and the resulting designs were evaluated by structure prediction with RaptorX-Single and Chai-1 prior to experimental characterization. (C) Fluorescence titrations of DoxSand0 with doxorubicin. Data were fit to a quadratic binding model, yielding dissociation constants of K_D=1.05±0.11 μM for DoxSand0 (best-fit ± parameter SEM from nonlinear regression, n = 3) (D) Chai-1 prediction of DoxSand0. Two views related by a 180° rotation highlight the five-helix scaffold (H1–H5) and the designed doxorubicin-binding pocket.

2.3 RESULTS

2.3.1 Motif-guided computational design identifies a novel fold for binding doxorubicin.

To test whether aromatic π-stacking interactions are sufficient to mediate binding of doxorubicin, we generated *de novo* protein scaffolds around a six-residue motif derived from the crystal structure of a protein with moderate affinity for daunomycin, an anthracycline analogue of doxorubicin [65] (K_D = 250 nM, Fig. 2.S1). The motif consists of two three-residue fragments centered on tryptophan residues that position the anthracycline core between opposing aromatic side chains (Fig. 2.1B). Using this Trp sandwich motif, RFdiffusion All-Atom[28] was used to generate 500 candidate protein backbones. Scaffolds were initially filtered based on ligand burial and Define Secondary Structure of Proteins (DSSP)[90]-derived assignments, reducing the scaffold pool to 66 candidates. Twelve scaffolds were then selected based on overall structural quality, preservation of the aromatic interaction motif, and accessibility of the designed binding pocket prior to sequence design with LigandMPNN and Rosetta FastRelax (Fig. 2.S2). Candidate sequences were evaluated by sequential structure prediction. Apo structures were first screened using RaptorX-Single[31] to eliminate poorly folded designs before ligand-bound complexes were evaluated with Chai-1[38]. Designs were prioritized based on predicted structural confidence, self-consistency with the design model predictions, and preservation of the intended protein–ligand geometry, resulting in a final set of twelve designs for experimental characterization ranging in length from 93 to 101 residues (Fig. 2.S3).

Nine of the twelve designs showed soluble expression in *E. coli* and were purified for biophysical characterization. One construct, DoxSand0, exhibited measurable binding to doxorubicin with a dissociation constant of K_D= 1.05 ± 0.11 μM (Fig. 2.1C).

The fold of DoxSand0 appears to be well optimized to serve as a miniaturized protein binder with a jaw-like architecture (Fig. 2.1D). Starting from the N-terminus, the upper piece of the jaw (colored in yellow) is composed of a helical hairpin (H1 and H2), an intervening helix (H3) forms a hinge (colored in orange), and a C-terminal helical hairpin (H4 and H5) forms the lower jaw (colored in dark orange). The “back” two helices (H1 and H5) are tightly interacting, forming the bottom of the cavity, and likely also imparting thermodynamic and structural stability (Fig. 2.1D, right). The hinge helix pulls the two front helices apart, ,creating the open cavity that houses the Trp sandwich.

DoxSand0 demonstrated that a motif-guided design strategy could generate a functional doxorubicin binder directly from a small computational design set without experimental affinity maturation. However, biophysical characterization revealed that the scaffold remained incompletely optimized (Fig. 2.2A-C), prompting structure-guided improvement of the protein’s stability and binding affinity while preserving the designed recognition mechanism.

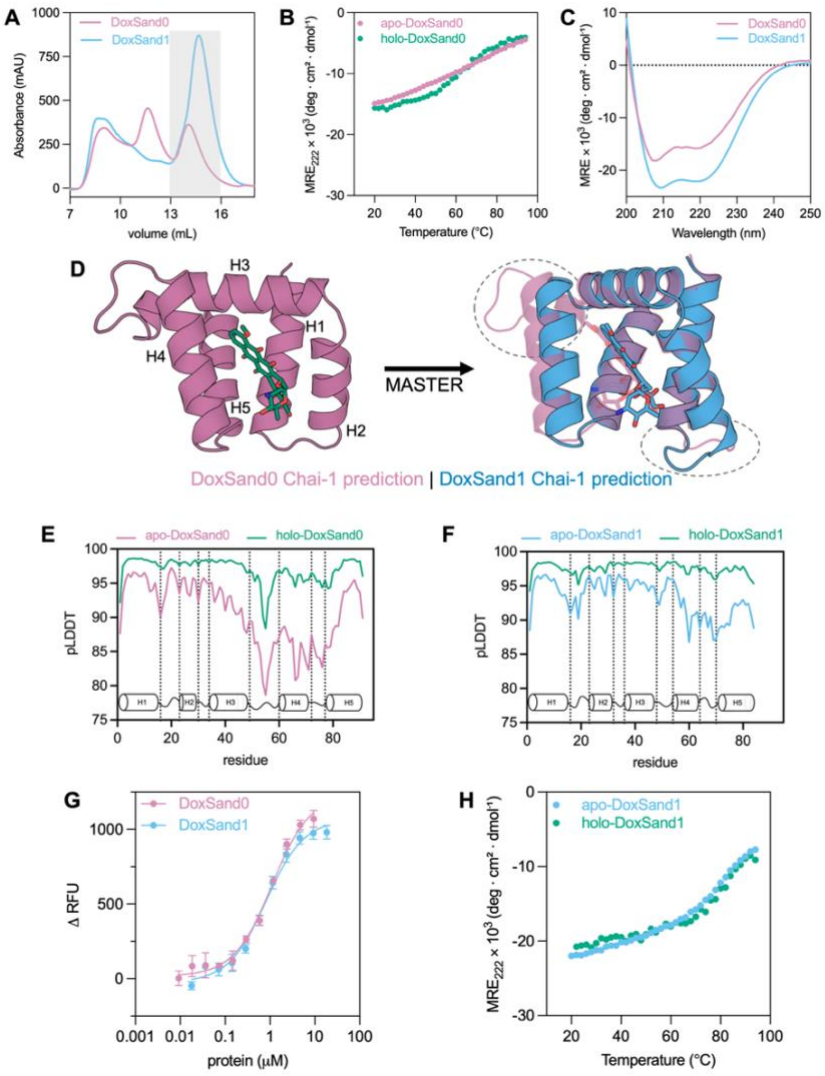

Figure 2.2 In silico scaffold optimization and biophysical characterization of doxorubicin-binding proteins.
(A) Size-exclusion chromatography traces of DoxSand0 and DoxSand1. The monomeric fraction used for quantification is indicated by the gray shaded region. (B) Thermal denaturation of apo- and doxorubicin-bound DoxSand0 monitored by circular dichroism at 222 nm. (C) Far-UV circular dichroism spectra of DoxSand0 and DoxSand1. (D) Loop remodeling of DoxSand0 using MASTER generated the DoxSand1 scaffold. Structures shown are Chai-1 predictions; remodeled loop regions are indicated by dashed circles. (E,F) Chai-1-predicted local distance difference test (pLDDT) scores for the apo and holo forms of DoxSand0 (E) and DoxSand1 (F). Secondary structure elements are shown below each plot. (G) Fluorescence titrations of DoxSand0 and DoxSand1 with doxorubicin. Data were fit to a quadratic binding model, yielding dissociation constants of K_D=1.05±0.11 μM for DoxSand0 and K_D=0.776±0.080 μM for DoxSand1 (best-fit ± parameter SEM, n = 3). (H) Thermal denaturation of apo- and doxorubicin-bound DoxSand1 monitored by circular dichroism at 222 nm.

2.3.2. Structure-guided scaffold optimization improves stability while preserving ligand recognition. DoxSand0 demonstrated that the motif-guided design strategy successfully generated a functional doxorubicin binder, but biophysical characterization indicated that the scaffold had several limitations. Size-exclusion chromatography revealed that monomeric species constituted only 29% of eluted protein, with the remainder forming higher-order oligomers (Fig. 2.2A). Consistent with this observation, circular dichroism spectroscopy showed limited cooperative thermal, suggesting that the designed fold lacked the stability required for a well-defined monomeric scaffold (Fig. 2.2B,C).

Because the binding site was already functional, we sought to improve the scaffold while preserving the designed ligand-recognition geometry. We hypothesized that the structural heterogeneity arose primarily from backbone geometry. To test this possibility, the loops connecting helices H1–H2 and H3–H4 were remodeled using Method of Accelerated Search for Tertiary Ensemble Representatives (MASTER)[35] to identify structurally compatible fragments from experimentally determined protein structures while maintaining the surrounding ligand-binding architecture (Fig. 2.2D). The redesigned scaffold contained 85 amino acids, reducing the overall protein size while preserving the intended binding pocket.

The redesigned scaffold was subsequently optimized through three cycles of LigandMPNN sequence design and Rosetta FastRelax refinement. During sequence design, 31% of the residues were held fixed to preserve the optimized scaffold geometry and ligand-binding motif (Fig. 2.S4). The resulting sequences were evaluated using the same RaptorX-Single and Chai-1 filtering procedure employed in the initial

design campaign. Structure prediction indicated increased confidence throughout the remodeled loop regions (Fig. 2.2E,F). Characterization of DoxSand1 by size exclusion chromatography showed that the monomeric fraction increased from 29% to 51% (Fig 2.2A). DoxSand1 also showed improved thermal stability while maintaining a similar binding affinity to doxorubicin, with K_D=0.78±0.08 μM (comparable to the starting DoxSand0, which had a K_D=1.05±0.11; Fig. 2.2C,G,H). These results indicate that scaffold stability could be substantially improved without disrupting the designed ligand-recognition mechanism, suggesting that scaffold optimization and binding-site optimization could be pursued independently.

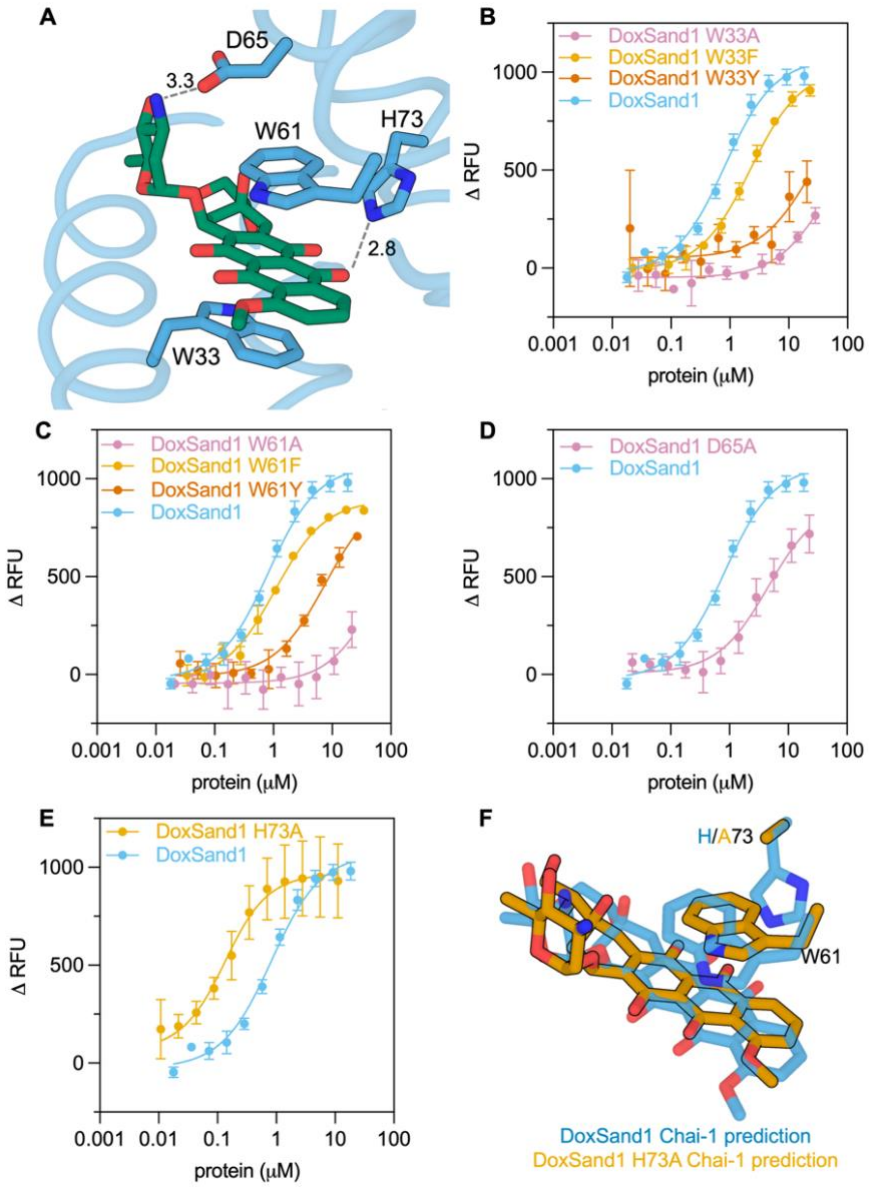

Figure 2.3 Mutational analysis of the DoxSand1 binding pocket.
(A) Predicted binding-site interactions between DoxSand1 and doxorubicin. Residues selected for mutational analysis are shown as sticks. Hydrogen-bonding interactions are shown as dashed lines, with heavy atom–heavy atom (donor–acceptor) distances indicated in angstroms. (B–E) Fluorescence titrations of DoxSand1 variants with doxorubicin. Data were fit to a quadratic binding model. Dissociation constants (K_D) were: W33F, 2.18 ± 0.17 μM (B) W61F, 1.03 ± 0.10 μM; W61Y, 7.57 ± 1.45 μM (C); D65A, 4.51 ± 1.06 μM (D); H73A, 95.0 ± 31.0 nM (E). (best-fit ± parameter SEM from nonlinear regression, n=3). Variants W33A, W33Y, and W61A exhibited substantially reduced binding but did not reach saturation over the concentration range tested, preventing reliable determination of K_D. (F) Overlay of Chai-1 predictions for DoxSand1 (blue) and the H73A mutant (orange), showing the predicted binding-site architecture

2.3.3 Binding-site optimization yields a nanomolar-affinity doxorubicin binder.

Having established a stable scaffold that binds doxorubicin, we next tested the designed binding mode by individually mutating residues predicted to contact the ligand and evaluating the resulting variants by fluorescence binding assays (Fig. 2.3A). Mutation of Trp33 revealed a strong preference for a bulky aromatic residue at this position. Substitution with phenylalanine largely preserved binding (K_D= 2.18 ± 0.17 μM), whereas tyrosine and alanine substantially reduced affinity (Fig. 2.3B). Trp61 was more tolerant to substitution, with phenylalanine (K_D= 1.03 ± 0.10 μM) and tyrosine (K_D= 7.57 ± 1.45 μM) retaining measurable binding, while alanine markedly weakened binding (Fig. 2.3C). These results suggest that favorable packing interactions at both positions are important for doxorubicin binding, with stricter steric requirements at Trp33 than at Trp61 (Fig. 2.3D). Consistent with the design model, mutational analysis of Trp33, Trp61, and Asp65 supports a binding mode in which the two tryptophan residues form the central ligand-binding interface, while surrounding polar interaction further stabilizes the protein–ligand complex. In contrast, mutation of His73 to alanine unexpectedly increased affinity, yielding a dissociation constant of 95± 31 nM (Fig. 2.3E). The Chai-1 predictions for the H73A variant suggest a subtle but meaningful rearrangement of the binding site, including a slight shift in the ligand pose and a corresponding change in the Trp61 rotamer, which together improve ligand burial and the geometry of the aromatic interaction network (Fig. 2.3F).

Despite its improved binding affinity, the H73A variant exhibited a reduced monomeric fraction (36%) compared to DoxSand1 (51%) by size-exclusion chromatography (Fig. 2.4A), suggesting that removal of the His73 side chain

compromised scaffold stability even as it enhanced ligand recognition. To restore stability while maintaining the improved binding properties of the H73A variant, we introduced a structure-guided V46I substitution adjacent to the binding pocket; the increased bulk of the Ile relative to the initial Val sidechain was selected to improve van der Waals packing by filling a small cavity at the helix-helix interface as well as increasing the hydrophobic driving force for folding. Because H73A emerged as the lead variant, we subsequently recharacterized its affinity using fluorescence titrations spanning three doxorubicin concentrations and globally fit the resulting datasets with a shared dissociation constant, yielding a more robust affinity estimate of K_D = 58 ± 17 nM (Fig. 2.4C). The resulting construct, DoxSand2, exhibited an increased monomeric fraction (61%) and retained nanomolar affinity for doxorubicin (K_D = 85 ± 16 nM), while displaying improved biophysical properties relative to DoxSand1 H73A (Fig. 2.4D,E-G). Overall, iterative structure-guided optimization transformed the initial computational hit into a higher-affinity binder with improved biophysical properties while preserving the designed binding mode.

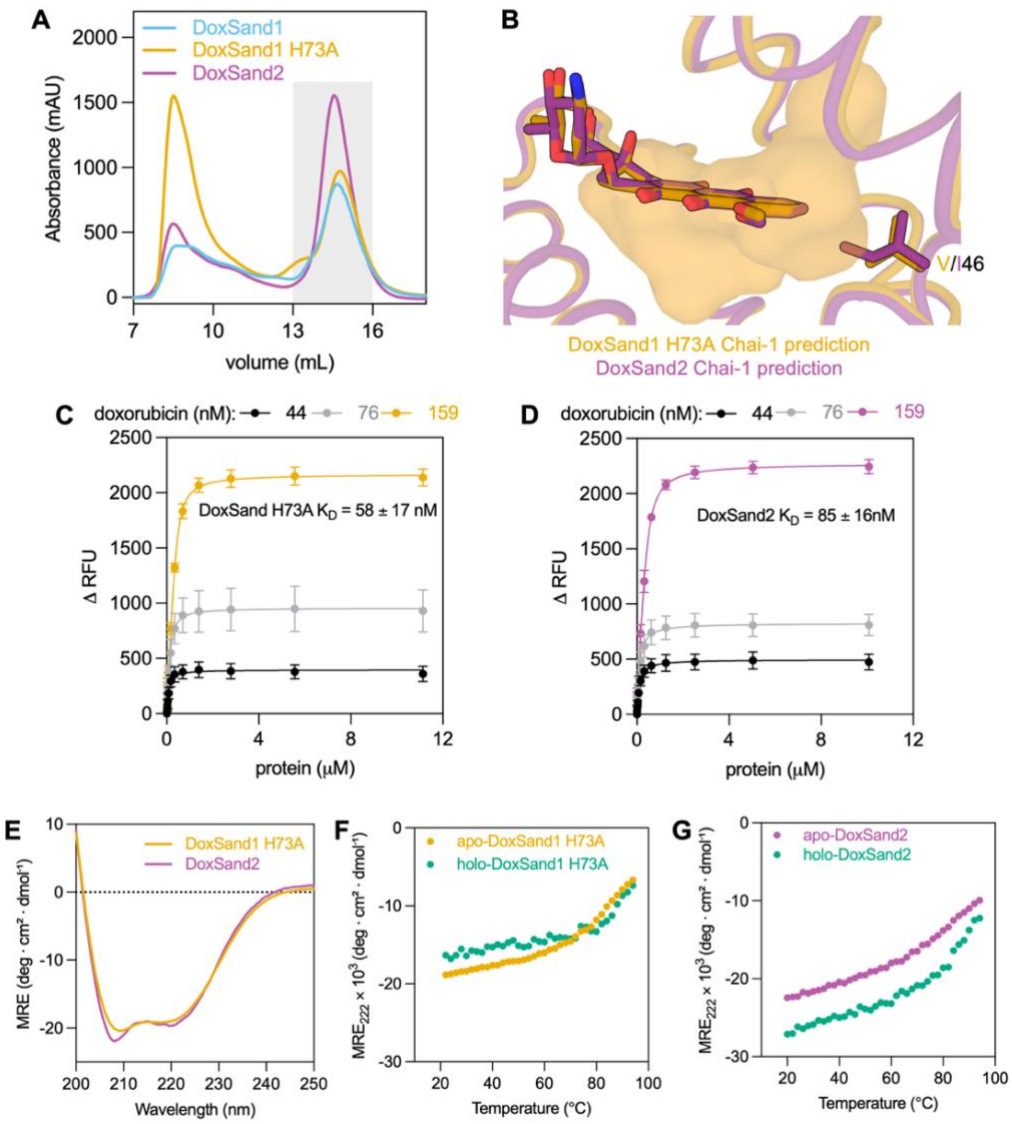

Figure 2.4 Structure-guided optimization of the DoxSand scaffold.
(A) Size-exclusion chromatography traces of DoxSand1, DoxSand1 H73A, and DoxSand2. Gray shaded regions denote the monomeric fractions used for subsequent assays, representing 51% (DoxSand1), 36% (DoxSand1 H73A), and 61% (DoxSand2) of the total eluted protein. (B) Overlay of Chai-1 predictions for DoxSand1 H73A (orange) and DoxSand2 (magenta), showing that the V46I substitution fills previously unoccupied space within the doxorubicin-binding pocket, resulting in improved predicted packing around the ligand. (C,D) Fluorescence titrations of DoxSand1 H73A (C) and DoxSand2 (D) with 44, 76, and 159 nM doxorubicin. Data were globally fit by nonlinear regression using a quadratic binding model with a shared dissociation constant, yielding K_D = 58 ± 17 nM for DoxSand1 H73A and K_D= 85 ± 16 nM for DoxSand2 (best-fit ± parameter SEM from nonlinear regression, n = 3). (E) Far-UV circular dichroism spectra of DoxSand1 H73A and DoxSand2. (F,G) Thermal denaturation monitored by circular dichroism for apo- and doxorubicin-bound DoxSand1 H73A (F) and DoxSand2 (G).

2.3.4. Crystal structure validates the designed recognition mechanism. To evaluate the accuracy of the computational design, we determined the crystal structure of the doxorubicin-bound DoxSand2 complex at 1.56 Å resolution (Fig. 2.5). The experimentally determined structure agrees closely with the design model, with an overall Cα RMSD of 0.57 Å relative to the Chai-1 prediction (Fig. 2.5A). The crystal structure confirms that DoxSand2 adopts the desired fold: a compact five-helix topology composed of an N-terminal helix-hairpin (H1-H2), a central helix (H3), and a second helix-hairpin (H4-H5) that packs across the central helix to form a tightly integrated tertiary structure (Fig. 2.5A). The helices position the designed ligand-binding pocket at the interface between the two helix-hairpin motifs.

Inspection of the binding pocket further confirmed that the principal design features were preserved. The anthracycline core of doxorubicin is positioned between the two designed tryptophan residues (W33 and W61), reproducing the aromatic sandwich interaction that served as the foundation for the computational design (Fig. 2.5B). Likewise, Asp65, identified by mutational analysis as an important determinant of binding, occupies the predicted position adjacent to the amino sugar of doxorubicin and participates in the expected interaction. The agreement between the mutational analysis and the crystal structure demonstrates that the dominant interactions responsible for ligand recognition were accurately encoded during computational design.

Comparison of the crystal structure with the Chai-1 also shows that the ligand is predicted with good accuracy (ligand RMSD = 1.34 Å), with only a small, Ångstrom-level shift of the ligand in the binding pocket (Fig. 2.5A,B). As in the model, the anthracycline core is sandwiched between Trp33 and Trp61 in the designed aromatic recognition

geometry. A small shift in the ligand position establishes several interactions absent from the design model, including a direct hydrogen bond between the doxorubicin hydroxyl group and the backbone carbonyl of Leu12, as well as a direct interaction with Arg39. In addition, an ordered water molecule bridges the doxorubicin carbonyl group and Trp33, creating a water-mediated interaction that was not explicitly modeled during design which did not consider solvent(Fig. 2.5C,D).

To assess the structural novelty of DoxSand2, the experimentally determined structure was searched against both the PDB100 and AlphaFold/UniProt50 databases using Foldseek (3Di/AA mode).[91] No statistically significant structural homologs were identified in either database. Searches against PDB100 yielded only low-confidence matches (9–26% sequence identity; Foldseek probability <0.3; E-value >1) (Fig. 2.S5A). Similarly, searches against AlphaFold/UniProt50 identified no statistically significant homologs; although 24 hits exhibited Foldseek probabilities ≥0.5, all retained E-values >1 (Fig. 2.S5B). Moreover, the highest-scoring matches were distributed across diverse α-helical proteins, with no enrichment for a single structural family. These results indicate that DoxSand2 adopts an α-helical architecture that is distinct from previously characterized protein scaffolds.

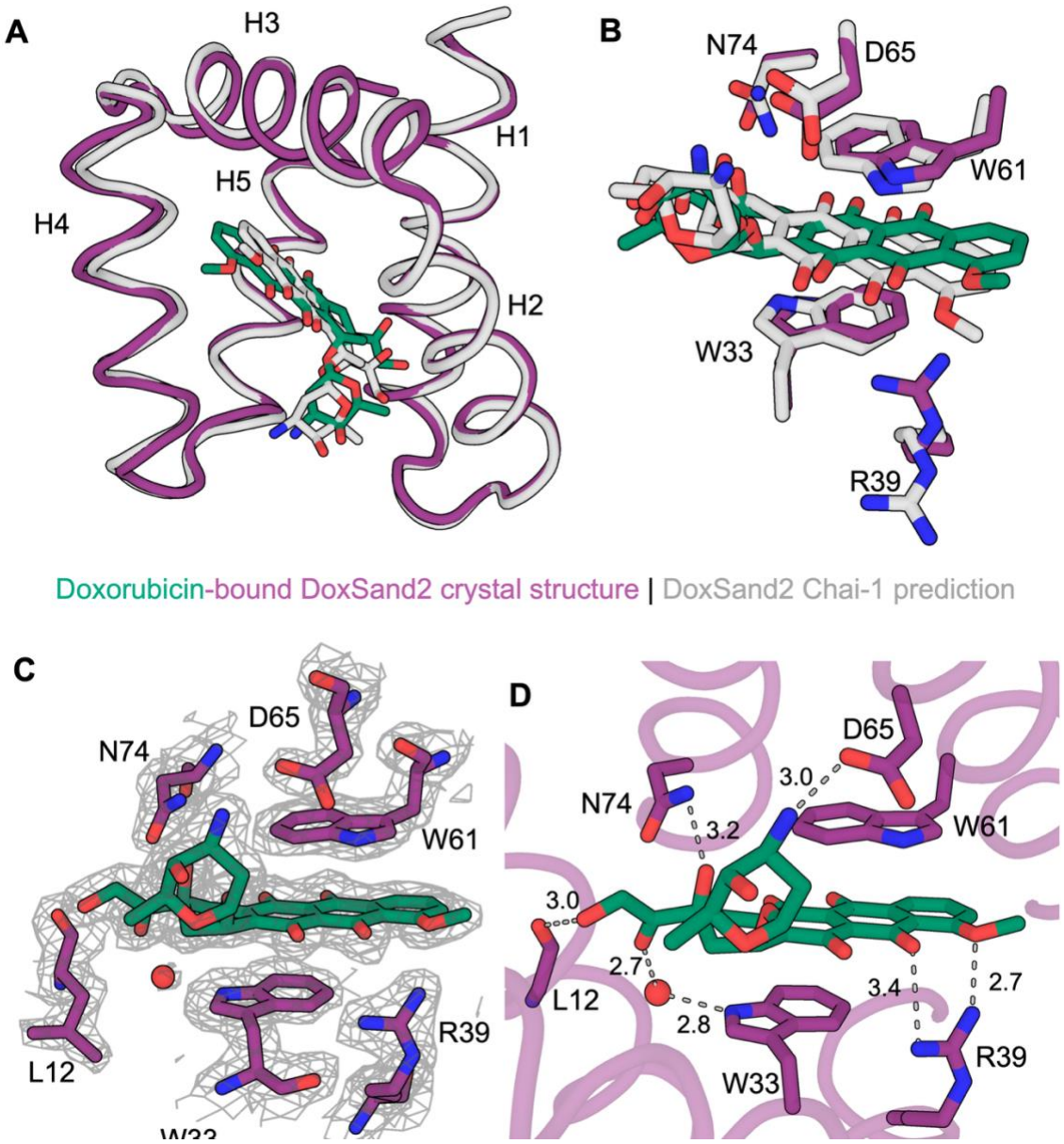

Figure 2.5 Crystal structure of the doxorubicin-bound DoxSand2 complex.
(A) Superposition of doxorubicin-bound DoxSand2 crystal structure (doxorubicin colored in dark green, DoxSand2 colored in magenta) and the DoxSand2 Chai-1 prediction (doxorubicin and DoxSand2 colored gray). Doxorubicin is shown as sticks**.** (B) Overlay of the binding-site residues from the crystal structure and Chai-1 prediction after backbone superposition. (C) Binding pocket of the doxorubicin-bound DoxSand2 crystal structure. The 2Fo−Fc composite omit electron density map for doxorubicin and surrounding binding-site residues is contoured at 1σ**.** (D) Binding-site interactions observed in the crystal structure. Hydrogen-bonding interactions are shown as dashed lines, with heavy atom–heavy atom (donor–acceptor) distances indicated in angstroms. Ordered water molecule is shown as red sphere**.**

2.3.5. Designed binders attenuate doxorubicin cytotoxicity in COV362 cells.

To determine whether the designed proteins modulate the biological activity of doxorubicin, we evaluated their effects on the proliferation of COV362 ovarian cancer cells (Fig. 2.6). Dose–response curves were first established using free doxorubicin and then repeated using doxorubicin preincubated with either DoxSand1 or DoxSand2, each at a fixed protein concentration of 20 μM. Preincubation with DoxSand1 increased the half-maximal inhibitory concentration (IC_{50}) of doxorubicin approximately 17-fold, from 6.9 to 116 nM. DoxSand2 produced an even larger effect, increasing the apparent IC_{50} to 951 nM, corresponding to a ~140-fold shift relative to free doxorubicin. The greater protective effect of DoxSand2 is consistent with its higher affinity for doxorubicin relative to DoxSand1. These results demonstrate that the designed proteins effectively sequester doxorubicin in a cellular environment and substantially attenuate its cytotoxic activity.

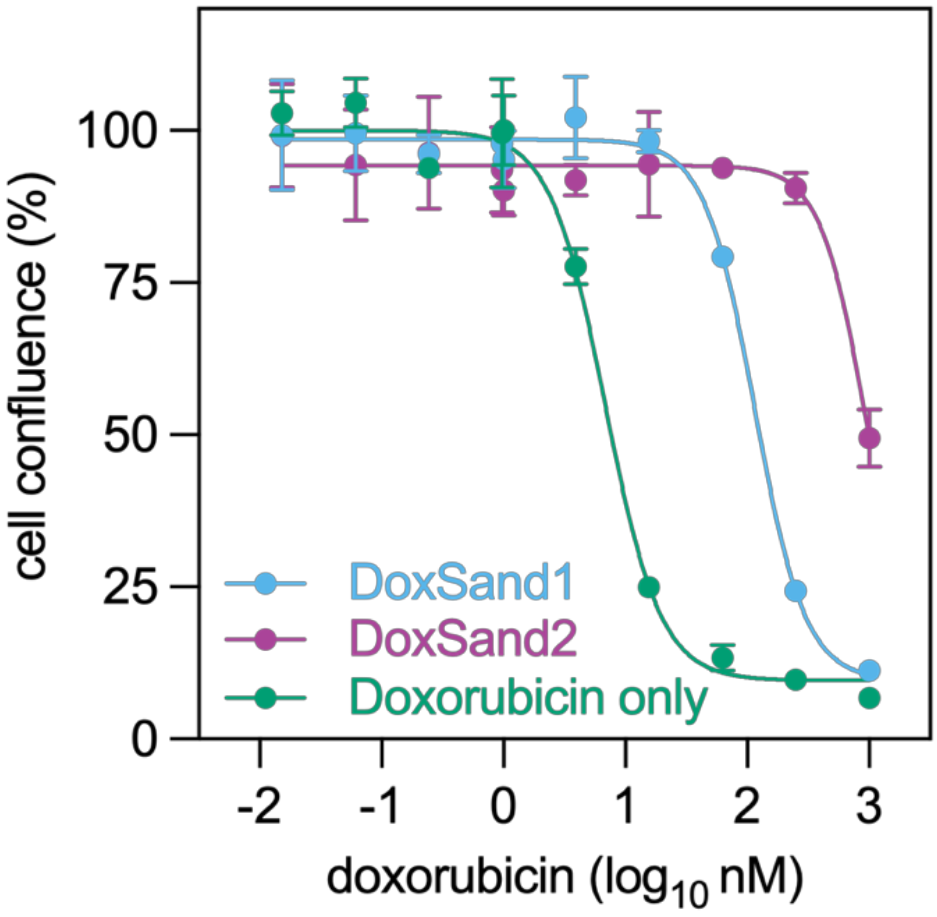

Figure 2.6 DoxSand proteins reduce doxorubicin cytotoxicity in COV362 cells. COV362 cell confluence following treatment with increasing concentrations of free doxorubicin or doxorubicin preincubated with 20 μM DoxSand1 or DoxSand2. Curves represent fits to a four-parameter dose–response model. Data are shown as mean ± SD (n = 2). The apparent IC_{50} values were 6.85 nM (95% CI, 5.80–8.38 nM) for free doxorubicin, 115.7 nM (95% CI, 95.6–139.6 nM) for DoxSand1, and 950.7 nM (95% CI, 805.8–1122 nM) for DoxSand2.

2.4. DISCUSSION

We designed a compact protein architecture *de novo* that binds a chemically complex therapeutic molecule with nanomolar affinity. Though only 85 residues long, the protein adopts a well-defined five-helix topology that accurately positions doxorubicin within a buried ligand-binding pocket. Most structurally characterized *de novo* proteins bound to small molecule ligands have adopted a relatively limited set of scaffold architectures, primarily parameterized helical bundles and proteins based on the NTF2 fold. (Table 2.S1).[37,42,43,47] In contrast, DoxSand2 adopts a distinct five-helix topology organized around two helical hairpins connected by a central hinge helix, producing a jaw-like binding cavity that encloses the ligand while maintaining an exceptionally compact scaffold. This organization demonstrates that compact ligand-binding cavities can be supported by substantially different tertiary architectures than those represented among current *de novo* binders. More broadly, these findings suggest that the current collection of structurally characterized *de novo* small-molecule binders likely underrepresents the diversity of protein architectures capable of supporting high-affinity molecular recognition, even among proteins smaller than 90 residues.

A key aspect of this work is the use of an aromatic π-stacking interaction to guide binder design. Unlike many previous *de novo* binder design efforts, which optimize extensive interaction networks within predefined scaffold architectures,[37,42,43,47,48] we ask whether an aromatic π-stacking sandwich interaction could seed the discovery of an entirely new ligand-binding fold. The resulting nanomolar binder demonstrates that establishing a dominant geometric interaction can substantially reduce the complexity of designing binders for chemically complex small molecules while leaving secondary interactions to be optimized during subsequent refinement.

The crystal structure provides an opportunity to evaluate the current capabilities of ligand-aware computational design.[27,28,38] The experimentally determined protein closely matches the design model, whereas the bound ligand adopts a shifted pose that preserves the designed π-stacking interaction while introducing additional hydrogen-bonding interactions and ordered water molecules not explicitly modeled during design. These observations indicate that current computational methods accurately specify overall protein architecture and the dominant geometric features governing molecular recognition while highlighting opportunities to improve prediction of ligand placement and explicit treatment of solvent. More broadly, the progression from an initial computational hit to a nanomolar binder demonstrates that computational design can provide experimentally useful starting points for rational optimization without requiring extensive empirical screening.[37,42,43,48] The ability of DoxSand2 to reduce doxorubicin cytotoxicity demonstrates that compact *de novo* binders can meaningfully modulate the activity of clinically relevant small molecules, illustrating the functional potential of these designed proteins beyond structural validation.[37,62]

2.5. CONCLUSIONS

A minimal aromatic π-stacking motif was sufficient to seed the *de novo* design of a previously unobserved protein fold that binds a chemically complex small molecule with nanomolar affinity. These findings demonstrate that compact ligand-binding proteins need not be limited to naturally occurring folds or previously established design architectures. DoxSand demonstrates that the structural diversity of functional ligand-binding proteins remains far from fully explored.

2.7. DISCLOSURES

A.A. is a co-founder of Azkarra Therapeutics, Ovibio Corporation, Tango Therapeutics and Tiller Tx; a member of the board of Cambridge Science Corporation, Cytomx and Ovibio; a member of the scientific advisory board of Ambagon, Bluestar/Clearnote Health, Circle, Deciphera, Genvivo, GLAdiator, Interdict Bio Inc., Earli, Leapfrog Bio, ORIC, Phoenix Molecular Designs, Trial Library, Yingli/280Bio and Z prime; a consultant for Catalio, Longitude, Novartis, Overwater, ProLynx and YK bioventures; and holds patents on the use of PARP inhibitors held jointly with AstraZeneca from which he has benefited financially (and may do so in the future).

2.8. MATERIALS AND METHODS

2.8.1. Overview of Computational Design Process. An aromatic sandwich motif was extracted from the crystal structure of a daunorubicin-binding protein complex (PDB ID: 3F8F). The motif consisted of two tryptophan-containing three-residue fragments that form π-stacking interactions with the anthracycline core of daunorubicin. Doxorubicin was positioned by pair-fitting its anthracycline core onto the daunorubicin coordinates from the crystal structure, and the resulting protein–ligand motif was used as a fixed input for scaffold generation.

Protein backbones were generated using RFdiffusion All-Atom[28] with scaffold lengths ranging from 100 to 120 residues. Scaffolds were filtered based on ligand burial and secondary structure content. The remaining scaffolds were subjected to sequence design using LigandMPNN followed by Rosetta FastRelax (RosettaCommons, Linux binary release 315) refinement.[27] Structures were evaluated using RaptorX-Single[31] and subsequently analyzed using Chai-1 (version 0.5.2)[38] holo structure prediction. Designs

were manually selected for experimental characterization based on preservation of the aromatic sandwich geometry and consistency of ligand placement across multiple Chai-1 predictions.

Following identification of the initial binder DoxSand0, a second round of design was performed to improve scaffold stability and monomeric stability through loop replacement using MASTER[35]. Redesigned scaffolds were subsequently subjected to the same sequence design and structure prediction workflow.

Custom Python scripts were developed to implement the computational protein design workflow and perform downstream structural analyses. ChatGPT (OpenAI) was used as an interactive programming assistant to assist with code generation, debugging, and refactoring during pipeline implementation and data analysis. All computational workflows, parameter selection, analyses, and interpretation of results were designed, validated, and verified by the authors.

Representative inputs from design pipeline are shown below:

RFdiffusion all-atom inference command:

```
run_inference.py \
     inference.deterministic=True \
     diffuser.T=100 \
     inference.output_prefix=<output_prefix> \
     inference.input_pdb= <input_pdb> \
     contigmap.contigs=['100-150'] \
     inference.ligand=DM2 \
     inference.num_designs=1 \
     inference.design_startnum=0
```

LigandMPNN run command:

```
LigandMPNN/run.py \
     --model_type "ligand_mpnn" \
     --seed 111 \
```

```
     --pdb_path "input_pdb_path" \
     --out_folder "output_path" \
     --batch_size 2 \
     --number_of_batches 5 \
     --fixed_residues "lmpnn_fixed_resis" \
     --omit_AA "CM" \
```

Rosetta FastRelax run command:

```
rosetta.binary.linux.release-
315/main/source/bin/relax.static.linuxgccrelease \
    -database {rosetta_path}/main/database/ \
     -in:file:s $input_pdb \
     -extra_res_fa {ligand_params_file} \
     -relax:quick \
     -constraints:cst_fa_file {constraints_file} \
        -constraints:cst_fa_weight 5.0 \
     -out:no_nstruct_label \
```

RaptorX-Single run command:

```
RaptorX-Single/pred.py \
     fasta_path \
     params/RaptorX-Single-ESM1b-ESM1v-ProtTrans.pt \
     --out_dir={output_directory_path} \
     --device_id=0
```

Chai1 run command:

```
chai1_lab.chai1.run_inference(
     fasta_file=fasta_path,
     output_dir=output_dir,
     num_trunk_recycles=3,
     num_diffn_timesteps=200,
     seed=42,
     device=torch.device("cuda:0"),
     use_esm_embeddings=True)
```

2.8.2. Scaffold Generation. Protein backbones were generated using RFdiffusion All-Atom with scaffold lengths ranging from 100 to 120 residues. The input motif was derived from the daunorubicin-binding protein structure (PDB ID: 3F8F) and consisted of two fixed three-residue fragments containing the tryptophan residues

responsible for aromatic stacking interactions with the anthracycline core. Doxorubicin coordinates were generated by pair-fitting the anthracycline core of doxorubicin onto the daunorubicin molecule present in the crystal structure.

A total of 500 scaffolds were generated. Scaffolds were filtered based on ligand burial and secondary structure content. Ligand burial was quantified using solvent-accessible surface area (SASA) calculations performed with the Shrake–Rupley algorithm implemented in Biopython using a probe radius of 1.0 Å. Scaffolds were required to exhibit a ligand ΔSASA greater than 220 Å² upon binding. Secondary structure assignments were determined using DSSP through Biopython, and scaffolds containing more than six consecutive residues assigned as loop were removed. Application of these filters reduced the design pool from 500 to 66 scaffolds. The 66 filtered scaffolds were manually evaluated in PyMOL for overall scaffold geometry, preservation of the aromatic interaction motif, and accessibility of the ligand-binding pocket. Based on this assessment, 12 scaffolds were selected for subsequent iterative sequence design.

2.8.3. Sequence Design and Structural Filtering. The 12 selected scaffolds were subjected to three iterative rounds of fixed-backbone sequence design using LigandMPNN, each followed by backbone relaxation with Rosetta FastRelax. During sequence design, the aromatic sandwich residues (W33 and W61) were held fixed. A total of 512 sequences were generated for each scaffold, affording 6144 sequences. Designed sequences were refined using Rosetta FastRelax with ligand movement enabled during relaxation. Apo structures were subsequently predicted using RaptorX-Single and filtered based on average predicted confidence (pLDDT > 85), backbone

RMSD relative to the design model (<1.5 Å), and net charge between −3 and −7. Application of these criteria reduced the design pool to 170 sequences.

2.8.4. Chai-1 Structure Prediction and Design Selection. The remaining sequences were evaluated using Chai-1 holo structure prediction. Five independent Chai-1 predictions were generated for each sequence using doxorubicin represented by the SMILES string:

```
C[C@H]1[C@H]([C@H](C[C@@H](O1)O[C@H]2C[C@@](Cc3c2c(c4c(c3O)C(=O)
c5cccc(c5C4=O)OC)O)(C(=O)CO)O)N)O
```

Predicted structures were analyzed using custom scripts and manually inspected. Designs were prioritized based on preservation of the aromatic sandwich geometry, consistency of ligand placement, and agreement among the five Chai-1 predictions. Emphasis was placed on maintaining the orientation of the tryptophan sandwich residues and the self-consistency of the predicted ligand pose across all five predictions. Chai-1 predictions were used for ranking and design selection; no additional computational filtering was applied at this stage. Twelve sequences derived from six scaffolds were selected for experimental characterization.

For the final DoxSand2 structure prediction used for comparison with the experimentally determined crystal structure, Chai-1 was run using an alternative doxorubicin SMILES representation,

```
COc1cccc2C(=O)c3c(O)c4C[C@](O)(C[C@H](O[C@H]5C[C@H](N)[C@H](O)[C
@H](C)O5)c4c(O)c3C(=O)c12)C(=O)CO
```

because this representation consistently yielded Chai-1 predictions with the correct stereochemistry of doxorubicin.

2.8.5. Loop Remodeling. Following experimental characterization of the initial binder DoxSand0, a second round of design was performed to improve scaffold stability and monomeric stability. Loop replacement was performed using MASTER. Two six-residue helical fragments flanking each loop were used as search queries against structures in the Protein Data Bank. Loop lengths ranging from 3 to 10 residues were sampled using RMSD cutoffs of 1.0–1.5 Å. The wgap option in MASTER was used to identify loops of the specified length connecting the two helices within the defined RMSD threshold. The resulting loop candidates were clustered based on RMSD and ranked according to cluster size. Two loop regions, connecting helices H1–H2 and H3–H4, were replaced with structurally compatible fragments identified through this MASTER search procedure. The redesigned scaffold was reduced from 92 residues to 85 residues while preserving the binding-site architecture. Redesigned structures were subjected to three iterative rounds of fixed-backbone sequence design with LigandMPNN, during which 31% of scaffold residues were held fixed, each followed by Rosetta FastRelax refinement, with Rosetta constraints applied to residues W33 and W61 relative to the bound doxorubicin to preserve the binding-site geometry during sequence optimization. Refined designs were subsequently filtered using RaptorX-Single and evaluated with Chai-1 as described above.

2.8.6. Structural similarity search. Structural similarity searches were performed using the Foldseek[91] web server in 3Di-AA search mode with default settings. The holo monomer structure of DoxSand2 (PDB: 36IQ) was used as the query model. Searches were performed against the PDB100 and AFDB50 databases using default

parameters. Hits were evaluated based on Foldseek probability, E-value, and sequence identity.

2.8.7. Protein expression and purification. Gene fragments encoding the designed proteins followed by TEV protease cut site (sequence: ENLYFQS) and C-terminal 6×His tag were obtained from Twist Bioscience and cloned into pET28a(+) by Gibson assembly. Plasmid constructs were verified by Sanger sequencing (Azenta Life Sciences) using the T7 primer. Plasmids were transformed into *E. coli* BL21(DE3) cells and grown in Luria Broth (Miller formulation) medium supplemented with kanamycin to final concentration of 50 μg/mL at 220 rpm, 37 °C. Cultures were induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM at an OD_{600} of 0.6–0.8 and grown for an additional 4 hours at 220 rpm, 37 °C before harvesting by centrifugation.

Cell pellets were resuspended in 1× phosphate buffered saline (PBS), pH 7.4 supplemented with 20 mM imidazole and lysed by sonication (Sonic Dismembrator Model 500, Fisher Scientific) before clarification by centrifugation. Clarified lysates were loaded onto Ni-NTA resin (HisPur™ Ni-NTA Resin, ThermoFisher) equilibrated 1x PBS, 20mM imidazole and washed with 1× PBS, 20 mM imidazole. Bound proteins were eluted with 1× PBS containing 250 mM imidazole. Eluted fractions were concentrated and further purified by size exclusion chromatography on a Superdex 75 Increase 10/300 GL column equilibrated in 1× PBS (pH 7.4). Protein purity was assessed by SDS-PAGE (Invitrogen NuPAGE 4-12% Bis-Tris Mini Protein Gel, ThermoFisher) and concentrations were determined by UV–visible spectroscopy. Untagged (e.g.

hexahistidine tag removed by TEV cleavage) and His_6-tagged proteins showed no significant difference in binding; therefore, His-tagged protein was used for experiments.

2.8.8. Determination of binding dissociation constants. Fluorescence quenching experiments were performed in black, non-binding surface (NBS) 384-well microplates (Corning®, Cat. No. CLS3575BC) with 1X PBS, pH 7.4 as the assay buffer. Protein solutions were prepared at an initial target concentration of ~10 μM and serially diluted in a 96-well plate (100 μL per well). Exact protein concentrations were determined by UV–visible spectroscopy and used for all subsequent calculations. Aliquots (85 μL) of each protein dilution were transferred to the 384-well plate. Doxorubicin hydrochloride (Cayman Chemical, Cat. No. 15007) was dissolved in dimethyl sulfoxide (DMSO) to prepare a 20 mM stock solution and stored at –20 °C. For fluorescence quenching assays, a working stock solution was prepared by diluting the DMSO stock into Milli-Q water. The concentration of the working stock was verified by UV–visible spectroscopy using an extinction coefficient of 13,500 M^{-1} cm^{-1} at 480 nm.[92] Unless otherwise indicated, the working stock was prepared to yield a final doxorubicin concentration of 100 nM upon addition of 5 μL to each well (final assay volume, 90 μL). For global fitting experiments, the working stock concentration and corresponding final assay concentration were adjusted as required. Final in-well concentrations were calculated based on measured stock concentrations and dilution factors (protein: 85/90; doxorubicin: 5/90), corresponding to an approximate final doxorubicin concentration of ~0.10 μM. Following doxorubicin addition, plates were shaken for 30 s, incubated for 2 min at room temperature, and fluorescence was measured on a BioTek Synergy Neo2

multimode plate reader using top-read optics (excitation 480/20 nm; emission 590/20 nm; gain 100; read height 8.0 mm).

Fluorescence quenching was quantified as ΔRFU, calculated as the fluorescence intensity of the corresponding doxorubicin only (no protein) control well minus the fluorescence intensity of the sample well, such that increased quenching yielded larger positive ΔRFU values. Mean ΔRFU values ± SEM were used for graphical presentation. Nonlinear regression analyses were performed using individual replicate measurements. Binding data were analyzed in GraphPad Prism 10 (GraphPad Software) with a quadratic one-site binding model accounting for ligand depletion using the following equation:

$$Y = Y_0 + (Y_{max} - Y_0) \times \left(\frac{-b - \sqrt{b^2 - (4 \times a \times c)}}{2 \times a \times L}\right)$$

Where Y is ΔRFU, Y_0 is ΔRFU_{min}, Y_{max} is ΔRFU_{max} , X is protein concentration, L is ligand concentration, n is binding stoichiometry parameter, $a = 1$, $b = -K_D - \frac{L}{n} - X$, $c = \frac{L \times X}{n}$. For individual binding curves, the stoichiometry parameter was fixed at n = 1. For global analyses, datasets collected at different ligand concentrations were fit simultaneously with a shared K_D, while all other parameters were allowed to vary independently between datasets.

2.8.9. Circular dichroism. Circular dichroism spectra were collected on a Jasco J-810 spectropolarimeter using a 0.1 cm path length quartz cuvette. Protein samples were prepared in 1× PBS (pH 7.4) at concentrations of 10–15 μM, and concentrations were verified by UV–visible spectroscopy prior to measurement. Holo samples were

prepared by addition of a two-fold molar excess of doxorubicin and left to incubate at room temperature for 30 min.

Spectra were collected from 190–250 nm in continuous scanning mode at 50 nm/min with a bandwidth of 1 nm and five accumulations per sample. Thermal denaturation experiments were performed by monitoring ellipticity at 222 nm from 20–95 °C at 2 °C temperature intervals with a heating rate of 2 °C/min.

2.8.10. Monomeric fraction analysis. Monomeric fractions were quantified directly from the preparative size-exclusion chromatography chromatograms obtained during protein purification. Protein samples from 500 mL expression cultures were purified by preparative size-exclusion chromatography as described above, and chromatograms were exported and analyzed using GraphPad Prism 10. Baseline correction was performed by subtracting the average absorbance between 6 and 7 mL from each trace. To ensure consistent analysis across constructs, a fixed integration range of 7–18 mL was used to define the total eluted protein, and a predefined integration range of 13–16 mL was used to define the monomeric fraction. Areas under the curve (AUCs) for each region were calculated using the AUC analysis in GraphPad Prism. The percentage of monomeric protein was calculated as:

$$Monomeric\ fraction\ (\%) = 100 \times \frac{AUC_{13-16mL}}{AUC_{7-18mL}}$$

2.8.11. Cell viability assay. COV362 cells were cultured in DMEM media (Gibco, Cat# 10569010), supplemented with 10% fetal bovine serum (Corning, 35-010-CV) and 1% Pen Strep (Gibco, 15140-122). Cells were seeded in a 96-well plate at a density of 2,000 cells per well. After overnight, doxorubicin of serial 4-fold dilutions, ranging from 0

to 1,000 nmol/L, was preincubated with 20 μM doxorubicin binder proteins in media at room temperature for 5 min, and added to cells. After 7 days, cell confluence was measured using an IncuCyte Live Cell Imager (Sartorius; mean ± SD, n = 2). The values were normalized to those of cells treated by DMSO vehicle. Half maximal inhibitory concentration $(IC)_{50}$ concentrations were calculated by Nonlinear regression (curve fit) (log(inhibitor) vs. response-variable slope (four parameters)) using GraphPad Prism 10 (RRID:SCR_002798).

2.8.12. X-ray crystallography. A 3-fold molar excess of doxorubicin hydrochloride (stock solution in DMSO) was added to DoxSand2 in 1× PBS (pH 7.4) and incubated for 30 min at room temperature. Because residual free doxorubicin interfered with accurate spectrophotometric determination of the holo-protein concentration, the protein concentration was estimated from the initial apo-DoxSand2 concentration, which was determined by absorbance at 280 nm using an extinction coefficient of 13,980 M^{-1} cm^{-1}, assuming complete protein recovery throughout the loading and concentration steps.

Following incubation, the sample was centrifuged at 21,100 × g for 10 min at room temperature to pellet any precipitated material. The clarified supernatant was then transferred to an Amicon Ultra 0.5 mL centrifugal filter (10 kDa molecular weight cutoff). The sample was concentrated by centrifugation and washed three times with 300–400 μL of 1× PBS by repeatedly concentrating the retentate and replenishing the filter with fresh buffer. This reduced the concentration of unbound doxorubicin, although residual free ligand was expected to remain. The sample was subsequently concentrated to a final estimated protein concentration of 12.5 mg/mL Crystallization conditions were

screened against the Hampton JCSG+ Screen, Peg/Ion Screen, and Classics 1&2 Screens using a Mosquito Crystal instrument.

Based on these screening conditions, additional hanging drop optimization conditions were explored. The reported structure was obtained by diffraction of a ruby, hexagonally prismatic crystal grown in a hanging drop containing 500 mM HEPES, 489 mM Sodium Citrate, 68.5 mM Sodium Chloride, 5 mM Sodium phosphate dibasic, 0.9 mM Potassium phosphate monobasic, and 1.35 mM Potassium chloride at pH 7.25. Crystals were harvested and flash-frozen in liquid nitrogen without cryo protection.

2.8.13. X-ray crystallography data collection and processing. Diffraction data was collected at ALS Beamline 8.3.1 on a PILATUS3 S 6M detector with a 1.11582 Å beam wavelength at 100 K. Data were processed with XDS,[93] and statistics for data processing and structural refinement can be found in Table 2.3. Using the Phenix[94] suite, the structure was solved by molecular replacement with Phaser[95] using a DoxSand2 Chai-1 design model. Phenix and Servalcat[96] were used for refinement, and Elbow[97] was used for ligand restraint construction based on optimized coordinates. Cofactor placement was based on residual electron density in the binding site. When ligand occupancy was allowed to vary, full occupancy was found to best explain the data. Protein model building and adjustments were performed in Coot1[98], and MolProbity[99] was used in validation. The protein structure was deposited to the protein database as entry 36iq. All raw data were deposited to the Integrated Resource for Reproducibility in Macromolecular Crystallography[100] (proteindiffraction.org), from which they can be freely downloaded (doi:10.18430/M336IQ).

2.9. SUPPLEMENTARY FIGURES

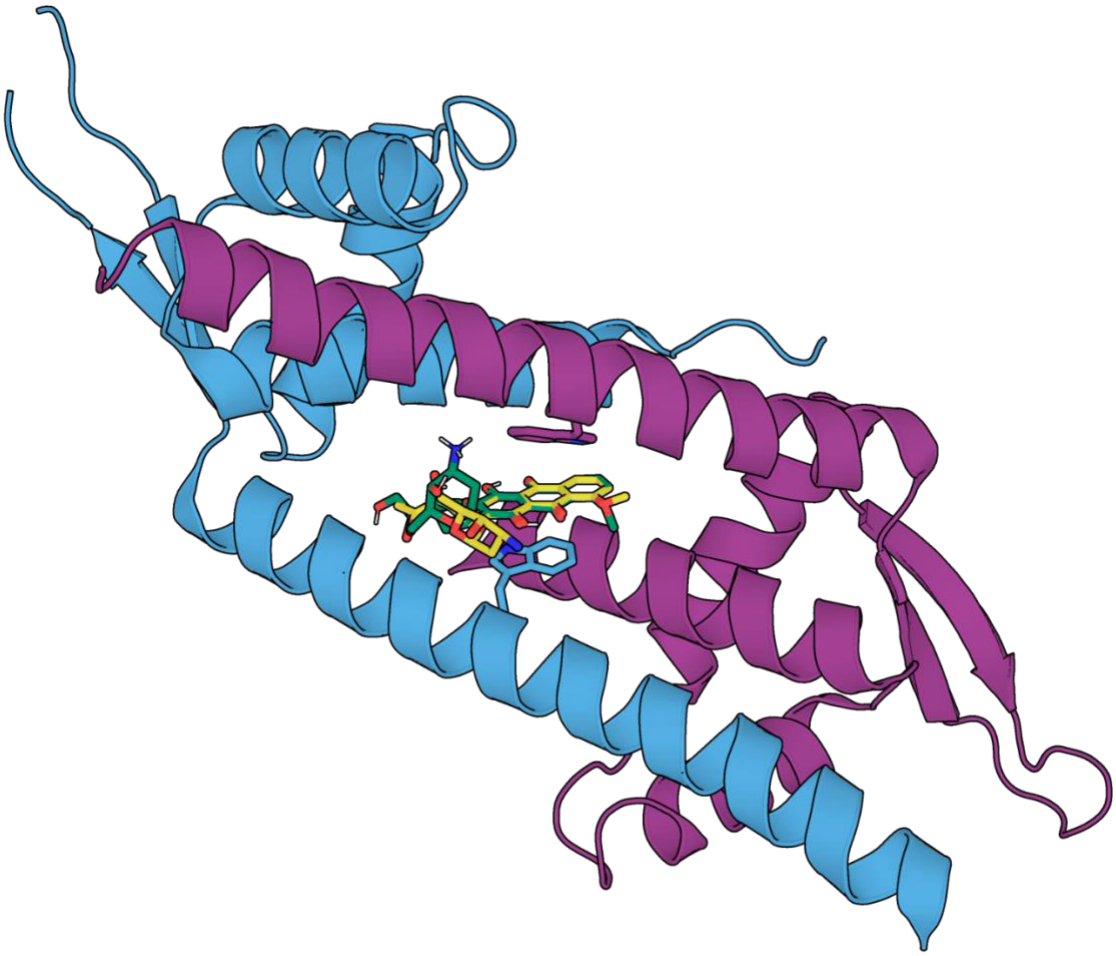

Figure 2.7 *Crystal structure of the daunomycin-bound LmrR dimer*
(PDB: 3F8F) The protein is colored by chain (blue and magenta). Daunomycin from the crystal structure is shown as green sticks. The relative ligand pose of doxorubicin (yellow) for motif-guided protein design was defined by overlaying its anthracycline core onto the crystallographic daunomycin core.

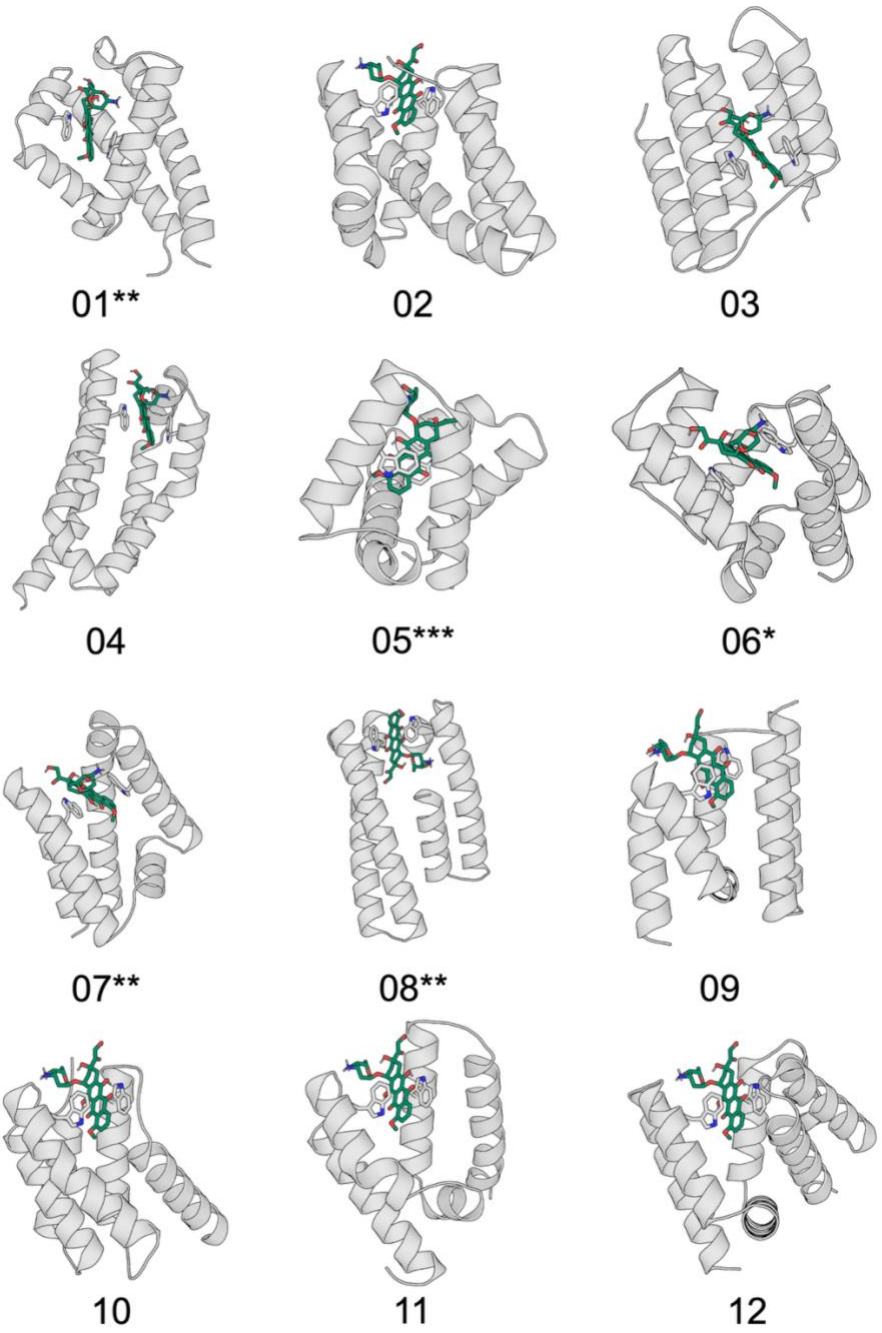

Figure 2.8 RFdiffusion All-Atom scaffold architectures selected for sequence design and experimental evaluation.
Twelve backbone scaffolds generated by RFdiffusion All-Atom and advanced through the sequence design and structure-prediction pipeline are shown. Asterisks denote experimental outcomes: (*) scaffold selected for gene synthesis; (**) scaffold expressed as soluble protein; (***) scaffold expressed as soluble protein and exhibited measurable doxorubicin binding.

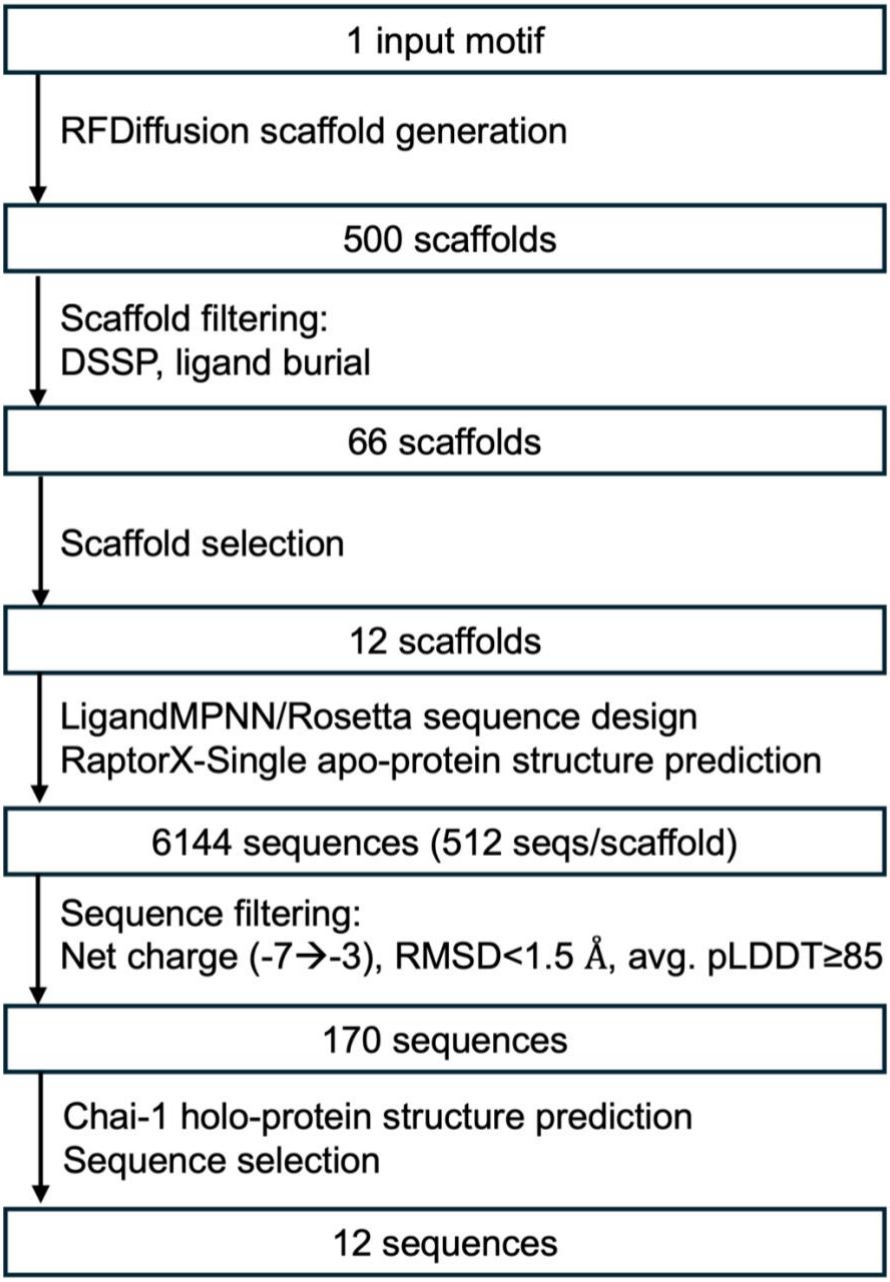

Figure 2.9 Progressive filtering of the computational design search space.
A six-residue aromatic π-stacking motif derived from the daunomycin-bound LmrR crystal structure (PDB: 3F8F) was provided as the fixed input motif to RFdiffusion All-Atom. Sequential filtering based on scaffold quality, ligand burial, secondary structure content, sequence design, and structure prediction progressively reduced the design space to 12 sequences representing six scaffold architectures that were selected for experimental characterization. Each box indicates the design entity and the number carried forward to the next stage.

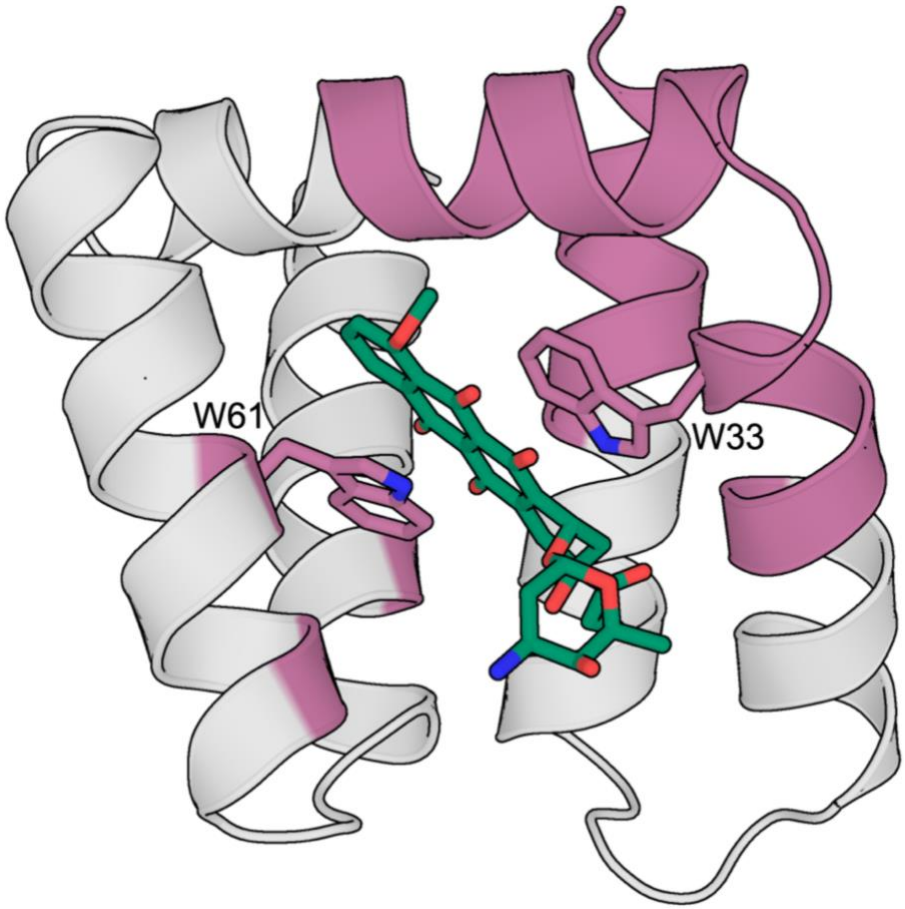

Figure 2.10 DoxSand0 scaffold after loop optimization using MASTER. Residues held fixed during sequence design (31% of the scaffold) are highlighted in pink, while the remaining residues were allowed to vary. Doxorubicin is shown as green sticks. Residues W33 and W61, shown as sticks, were constrained relative to doxorubicin using Rosetta AtomPair constraints during sequence design to preserve the binding-site geometry.

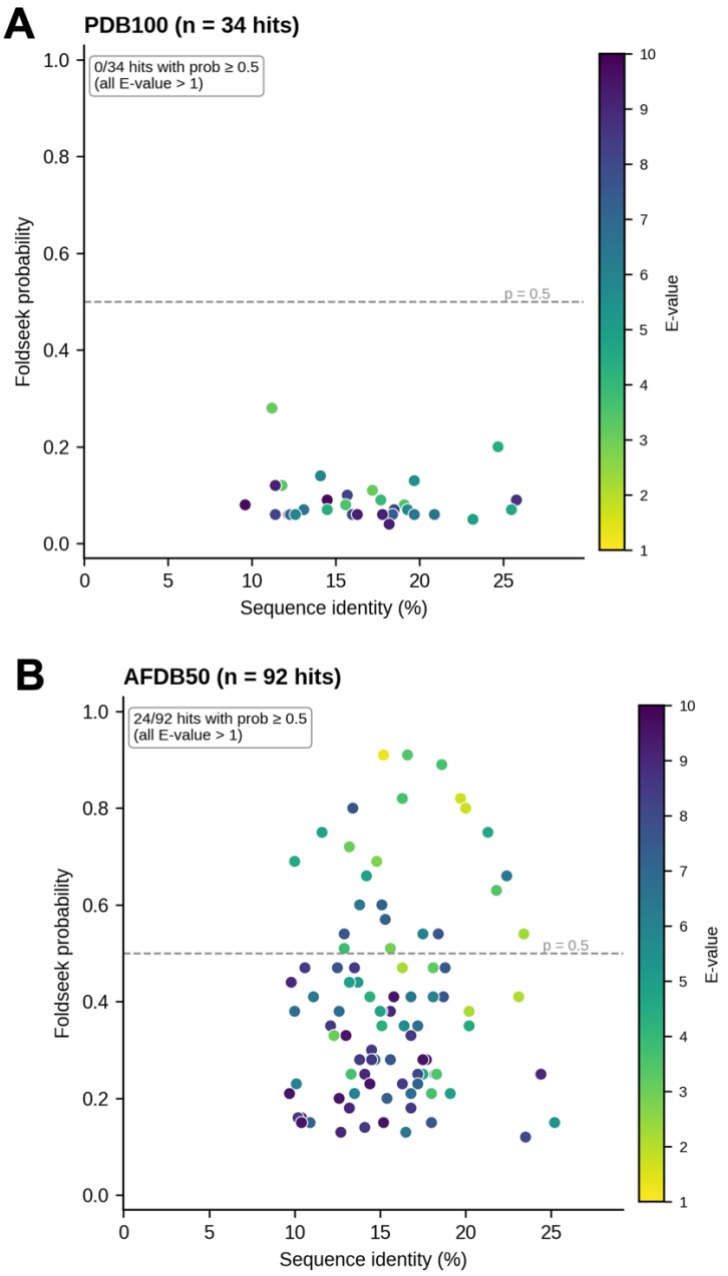

Figure 2.11 Foldseek (3Di/AA mode) search.
Queries against the PDB100 (A) and AlphaFold/UniProt50 (B) databases identified only weak structural matches. Sequence identity is plotted against Foldseek probability, with point color indicating E-value.[101,102] No statistically significant structural homologs were identified in either database (all E-values > 1).

2.10. SUPPLEMENTARY TABLES

Table 2.1 *De novo* Small-Molecule Binding Proteins with Experimentally Determined Holo Structures

protein	ligand	sequence length	scaffold	K_D (nM)
DoxSand2	doxorubicin	85	mini helical jaw	85
mFAP1[18]	DFHBI	112	β-barrel	560
A1E[19]	progesterone	113	NTF2 fold	5000-10000
apx1049[20]	apixaban	119	NTF2 fold	650
ABLE[21]	apixaban	126	4-helix bundle	5000
hcy129_mpnn5[20]	cortisol	135	NTF2 fold	2300
PiB[22]	rucaparib	147	4-helix bundle	<5

Table 2.2 Protein and DNA Sequences of DoxSand Variants

Protein	Protein sequence	DNA sequence
DoxSand0	MEKEIVEEALAILK AVLDGEPGAEER AAAFWANPENRK VVTDVVVKYANKL GLTKGSTDPAQAV KDWWAQADAAG DVAAGNAVGVKA LELELENLYFQSH HHHHH*	CTTTAAGAAGGAGATATACCATGGATGGAGAAGG AAATAGTGGAAGAAGCTCTGGCGATCCTGAAGGC GGTGCTTGATGGCGAACCTGGAGCAGAGGAGCG TGCCGCAGCCTTCTGGGCAAACCCCGAAAACCG AAAGGTCGTTACCGACGTAGTGGTTAAGTACGCG AACAAACTTGGCTTAACTAAGGGCTCTACGGACC CAGCCCAAGCCGTTAAGGACTGGTGGGCCCAGG CAGATGCCGCCGGTGACGTGGCCGCGGGTAATG CGGTAGGTGTAAAAGCTTTGGAGTTAGAGCTTGA GAACCTGTACTTTCAAAGTCATCACCACCATCAC CACTGACTCGAGCACCACCACCACCACCAC
DoxSand1	MEKEIVEEALKLV QGFLDDPNDKAV LEAAAAFWANPE NRKVVTDTVAKEL GISSEELEARWRE YDAAGRLAEHNEI VAKGLRKALENLY FQSHHHHHH*	CTTTAAGAAGGAGATATACCATGGATGGAAAAGG AGATTGTGGAAGAGGCTCTGAAGCTGGTGCAGG GCTTTCTCGACGACCCGAACGACAAGGCAGTGC TCGAAGCCGCGGCCGCGTTCTGGGCCAATCCAG AGAACCGTAAAGTGGTTACCGATACAGTAGCGAA GGAGTTAGGTATTTCGTCCGAGGAGCTTGAGGC CCGCTGGCGCGAGTACGACGCAGCAGGCAGATT AGCAGAACATAATGAGATCGTCGCGAAGGGTTTA CGTAAGGCCCTCGAAAATCTGTACTTCCAATCTC ATCATCACCACCATCATTAGCTCGAGCACCACCA CCACCACCAC
DoxSand2	MEKEIVEEALKLV QGFLDDPNDKAV LEAAAAFWANPE NRKVVTDTIAKEL GISSEELEARWRE YDAAGRLAEANEI VAKGLRKALENLY FQSHHHHHH*	CTTTAAGAAGGAGATATACCATGGATGGAGAAAG AAATCGTTGAAGAGGCCCTTAAGTTAGTCCAAGG CTTCCTGGACGATCCAAACGATAAGGCAGTTCTG GAAGCTGCAGCTGCATTCTGGGCAAACCCTGAA AATCGGAAGGTAGTAACGGACACGATCGCCAAAG AACTGGGTATCTCAAGCGAGGAATTAGAGGCGC GCTGGAGAGAGTATGATGCCGCTGGCCGACTCG CCGAGGCAAACGAGATCGTCGCGAAGGGCCTTC GCAAGGCACTGGAAAATTTGTATTTCCAGTCACAT CACCATCACCATCACTGACTCGAGCACCACCACC ACCACCAC

Table 2.3 Crystallographic data processing and refinement statistics for holo DoxSand2

	Value
PDB ID	36IQ
Space group	P 32 2 1
Cell constants a, b, c, α, β, γ	42.70 Å, 42.70 Å, 174.71 Å 90.00°, 90.00°, 120.00°
Resolution (Å)	58.24 – 1.56
% Data completeness	99.7 (58.24-1.56)
Rmerge	0.18
$< I/\sigma(I) >$	1.56
R, R_{free}	0.218, 0.255
R_{free} test set	1999 reflections (6.77%)
Wilson B-factor (Å^2)	25.1
Anisotropy	0.256
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37, 28.6
L-test for twinning	< |L| > = 0.47, < L2 > = 0.30
Estimated twinning fraction	0.046 for -h,-k,l
Fo,Fc correlation	0.95
Total number of atoms	3071
Average B, all atoms (Å^2)	38.0
Protein residues	180
RMS (bonds)	0.007
RMS (angles)	0.79
Ramachandran favored (%)	97.73
Ramachandran allowed (%)	2.27
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.69
Clashscore	2.33

2.11 SUPPLEMENTARY TEXTS

Quantum chemical calculations

Density functional theory optimization of the cationic doxorubicin geometry was carried out in ORCA 6.1.1[103,104] using the B3LYP functional with a D3BJ[105,106] dispersion correction and def2-SVP set. The reported structure was concluded to be a ground-state minimum based on the analysis of the harmonic vibrational analytical frequencies, which show no imaginary frequencies.

Doxorubicin coordinates:

C	-21.09693843452550	-2.30349715515109	10.72262167574020
C	-20.79468975060255	-3.47656711563832	9.81163732826365
O	-21.93740181405846	-3.67670252351532	8.96019234677085
C	-20.52628542018282	-4.78202881835712	10.56580332752132
O	-21.60057538733238	-5.14578753810384	11.41999901088420
C	-20.28503823548590	-5.91126567344913	9.54859530443355
N	-20.15898426480298	-7.18846789587201	10.32245821844999
C	-21.44339779655918	-6.01880372649053	8.56157005024540
C	-21.72602347213343	-4.64914158043469	7.94500059353509
C	-17.13629240435094	-6.71522001278077	-0.11058125414148
C	-17.41810963911604	-7.89307292240745	-0.80041166549987
C	-18.49779491646132	-8.68900301891866	-0.42914869070723
C	-19.32787751343310	-8.32296837041791	0.64988362955414
O	-20.36937901965412	-9.07049656740681	1.02937907414792
C	-19.05772593494920	-7.11927040524125	1.36758241153209
C	-19.89627058438804	-6.67569154398671	2.50374477736919
O	-20.88415538240861	-7.32586383517695	2.89678060116937
C	-19.54350432689929	-5.42170851453473	3.18385140155990
C	-20.33161887466574	-4.98129221758535	4.26396619832619
O	-21.39559090172265	-5.66468640513066	4.68274230877933
C	-20.00974279013777	-3.76158753785334	4.93182634296934
C	-20.88045718860018	-3.36128770012380	6.09933419591702
O	-20.64718718051054	-4.32996081991562	7.12598242060150
C	-20.65042918723356	-1.92109265769042	6.55586272295201
C	-19.16477730860662	-1.53297675611269	6.59394261030909
O	-18.52010506560866	-2.36868795370375	7.54631242507089
C	-19.09218527590996	-0.06596615609807	7.03283908419579
O	-19.39543509536898	0.84766248717994	6.28973571582890
C	-18.67878163364835	0.23691192791372	8.45729157198461
C	-18.55708748445511	-1.69812967415249	5.19856936578319
C	-18.92166146893204	-2.99847968292543	4.53395214438609

C	-18.12149667323065	-3.43549841253278	3.43706579724489
C	-18.42968617580821	-4.63709756210853	2.77310215605001
C	-17.60043948984298	-5.06497105977099	1.65630573888087
O	-16.62144344509488	-4.39051364490054	1.28414729078038
C	-17.94653717286498	-6.33133904456175	0.96106942040707
C	-20.67372187513569	-10.27435532789593	0.34049131345024
H	-19.28356150608560	-7.22508471806262	10.86373559502561
H	-20.94198089165320	-7.23834899491197	10.99580722250825
H	-17.56235260578633	-2.30970594895921	7.40774604471369
H	-22.42467038692296	-4.86905827631136	10.98509377408131
H	-21.44333331461864	-6.46626988899691	4.07252465369209
H	-16.29526678625109	-6.08007061869547	-0.38844157712256
H	-16.79075337651188	-8.20026610795967	-1.64063322894662
H	-18.69881827187165	-9.60468395199494	-0.98296597788819
H	-22.66634457667861	-4.66990278039658	7.36944952760049
H	-21.93539372107437	-3.45173687441778	5.78651877257764
O	-17.08937757923668	-2.67547956430412	3.07398945985512
H	-19.33668358451593	-5.74905817441032	9.02010600957075
H	-19.63314977973288	-4.66260837459468	11.20056701617475
H	-19.91454407679745	-3.25077815583157	9.18930411087640
H	-21.11889253035856	-1.77416226814696	7.53786236148043
H	-21.21754412497226	-6.73498436259120	7.75834300944037
H	-21.15237170813612	-1.24413582418294	5.84829657597909
H	-22.35255159504149	-6.36138794503422	9.08061326066983
H	-19.30975036115986	-0.37141992389533	9.13430189110308
H	-17.64416369701893	-0.14013995576244	8.59570197216914
O	-18.77975248220344	1.59780582206059	8.73254764448641
H	-18.89280378177246	-0.85694430015677	4.56931338282568
H	-17.46025756198830	-1.60353240812577	5.25340850524468
H	-20.91488060876335	-10.08637422353297	-0.71944178347139
H	-19.84512803566693	-10.99991456095369	0.40425359911444
H	-21.55700840312724	-10.69358690685340	0.84010025561144
H	-21.30298024662072	-1.40439225382596	10.12333578537467
H	-20.23344192827501	-2.09626420140046	11.37316656101464
H	-21.97117186679913	-2.50331327673373	11.36065798440931
H	-16.65881632464369	-3.14977755969153	2.30262384451004
H	-20.19149697322086	-8.02287999087298	9.72048227713987
H	-19.04538316747202	2.01367622569897	7.89130600673360

Doxorubicin ligand Rosetta parameters files (DM2_params_file):

```
NAME DM2
IO_STRING DM2 Z
TYPE LIGAND
AA UNK
ATOM  C15 aroC  X   -0.02
ATOM  C14 aroC  X   -0.02
ATOM  O5  OH    X   -0.56
ATOM  H12 Hpol  X   0.53
ATOM  C13 aroC  X   -0.02
ATOM  C12 aroC  X   -0.02
ATOM  O4  OOC   X   -0.66
ATOM  C11 aroC  X   -0.02
ATOM  C10 aroC  X   -0.02
ATOM  O3  OH    X   -0.56
ATOM  C27 CH3   X   -0.17
ATOM  H27 Hapo  X   0.19
ATOM  H28 Hapo  X   0.19
ATOM  H29 Hapo  X   0.19
ATOM  C9  aroC  X   -0.02
ATOM  C8  aroC  X   -0.02
ATOM  C7  aroC  X   -0.02
ATOM  C26 aroC  X   -0.02
ATOM  C25 aroC  X   -0.02
ATOM  O11 OOC   X   -0.66
ATOM  C24 aroC  X   -0.02
ATOM  C23 aroC  X   -0.02
ATOM  O10 OH    X   -0.56
ATOM  H19 Hpol  X   0.53
ATOM  C22 aroC  X   -0.02
ATOM  C21 CH2   X   -0.08
ATOM  C18 CH1   X   0.01
ATOM  O7  OH    X   -0.56
ATOM  H15 Hpol  X   0.53
ATOM  C17 CH2   X   -0.08
ATOM  C16 CH1   X   0.01
ATOM  O6  OH    X   -0.56
ATOM  C6  CH1   X   0.01
ATOM  O1  OH    X   -0.56
ATOM  C2  CH1   X   0.01
ATOM  C1  CH3   X   -0.17
ATOM  H1  Hapo  X   0.19
ATOM  H2  Hapo  X   0.19
ATOM  H3  Hapo  X   0.19
```

Rosetta constraints (constraints_file) used for re-design of DoxSand0

```
AtomPair NE1 60A C12 1X HARMONIC 3.4 0.2
AtomPair CH2 60A C23 1X HARMONIC 3.1 0.2
AtomPair CG 60A C10 1X HARMONIC 4.1 0.2
AtomPair CD2 60A C26 1X HARMONIC 4.0 0.2
AtomPair CZ2 60A C22 1X HARMONIC 3.5 0.2
AtomPair CD2 32A C14 1X HARMONIC 3.8 0.2
AtomPair CZ2 32A C24 1X HARMONIC 3.2 0.2
AtomPair CD1 32A C15 1X HARMONIC 3.8 0.2
AtomPair CH2 32A C25 1X HARMONIC 3.5 0.2
```

CHAPTER 3 Engineering DoxSand2 for Therapeutic Applications

3.1 INTRODUCTION

Chapter 2 established that DoxSand2 is a compact *de novo* protein capable of binding the chemotherapeutic doxorubicin with nanomolar affinity. While these studies established successful molecular recognition, therapeutic application requires additional properties beyond ligand binding alone. Proteins intended for in vivo use must exhibit favorable pharmacokinetic behavior, maintain structural integrity under physiological conditions, and remain stable in complex biological environments. One widely used strategy for extending serum half-life is fusion to an albumin-binding domain (ABD), which promotes reversible association with human serum albumin (HSA), the most abundant circulating protein.[107–109] To extend the therapeutic potential of DoxSand2, an albumin-binding domain was genetically fused to the protein to extend its circulating half-life.

This chapter describes the initial engineering and characterization of the DoxSand-ABD fusion protein, including evaluation of albumin and doxorubicin binding, structural characterization, and preliminary assessment of serum stability. In addition to these experimental studies, this chapter discusses ongoing pharmacokinetic studies and future engineering strategies aimed at further evaluating the therapeutic potential of DoxSand2.

3.2 RESULTS

3.2.1 Design of a DoxSand-ABD fusion protein. To extend the therapeutic potential of DoxSand2, an albumin-binding domain was genetically fused to the protein. The ABD was derived from the ABDCon sequence, a consensus albumin-binding domain engineered for high-affinity binding to serum albumin.[110] Four fusion constructs

were evaluated, differing in the orientation of the fusion (ABD–DoxSand2 or DoxSand2–ABD) and the length of the connecting linker (7 or 15 amino acids). Each construct contained a C-terminal TEV protease cleavage site followed by a polyhistidine purification tag. Following preliminary characterization, the construct consisting of an N-terminal DoxSand2 domain connected through a 15-residue linker to a C-terminal ABDCon domain was selected for all subsequent studies and is referred to hereafter as DoxSand-ABD (Fig. 3.1).

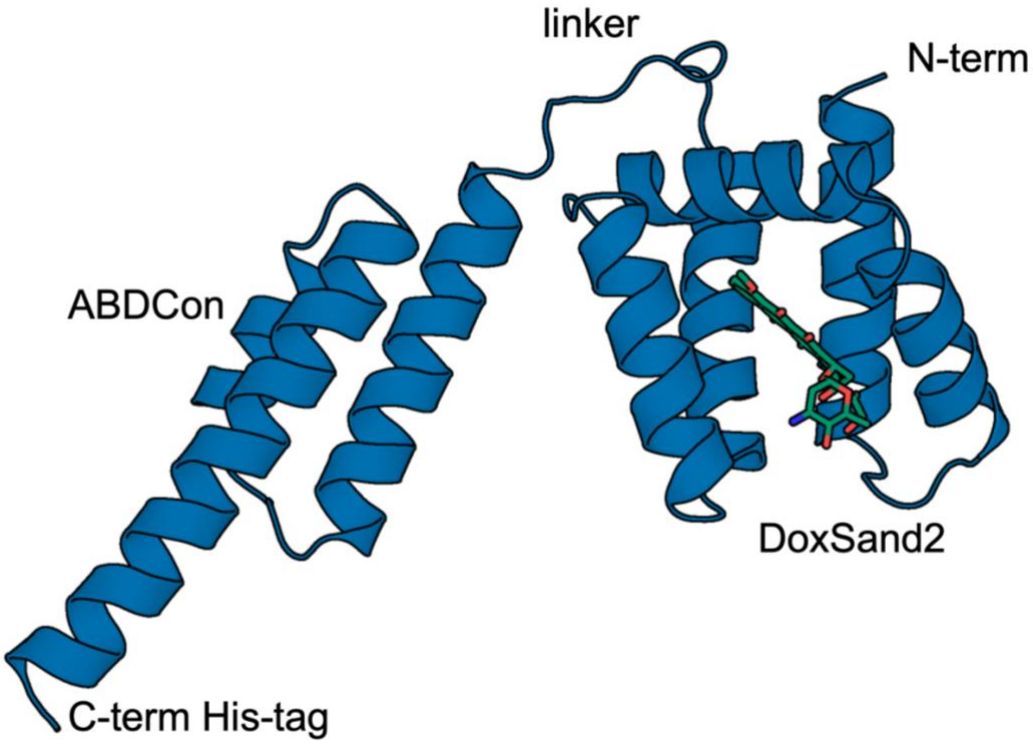

Figure 3.1 Chai-1 predicted structure of DoxSand-ABD.
The selected fusion construct consists of DoxSand2 connected through a 15-residue linker to ABDCon. The model predicts that the two domains remain structurally independent, with the doxorubicin-binding pocket of DoxSand remaining accessible. A C-terminal TEV protease cleavage site and his-tag were included for purification.

3.2.2. Characterization of DoxSand-ABD. Analytical size-exclusion chromatography demonstrated formation of a stable complex between DoxSand2-ABD and HSA, with the observed elution profile consistent with complex formation (Figure 3.2A). To determine whether addition of the albumin-binding domain disrupted the original ligand-binding function, fluorescence binding assays were performed using doxorubicin. DoxSand2-ABD retained doxorubicin binding with an affinity comparable to the parent protein (Figure 3.2B), indicating that fusion of the albumin-binding domain did not substantially perturb the designed binding pocket. Together, these results demonstrate that DoxSand2-ABD simultaneously recognizes both human serum albumin and doxorubicin, representing an initial step toward extending the utility of the designed protein beyond in vitro characterization.

Because fusion proteins frequently alter protein stability and folding, the structural properties of DoxSand2-ABD were compared with those of the parent DoxSand2 scaffold using circular dichroism spectroscopy. Far-UV CD spectra indicated that DoxSand2-ABD remained predominantly α-helical, with only modest differences relative to DoxSand2 (Figure 3.2C). Thermal denaturation monitored by circular dichroism similarly demonstrated that the fusion protein retained cooperative unfolding behavior over the measured temperature range (Figure 3.2D). Although minor differences in thermal behavior were observed, the overall secondary structure and folding characteristics remained similar to those of the parent scaffold. These results indicate that fusion of the albumin-binding domain does not substantially compromise

the structural integrity of DoxSand2 and supports its continued evaluation for translational applications.

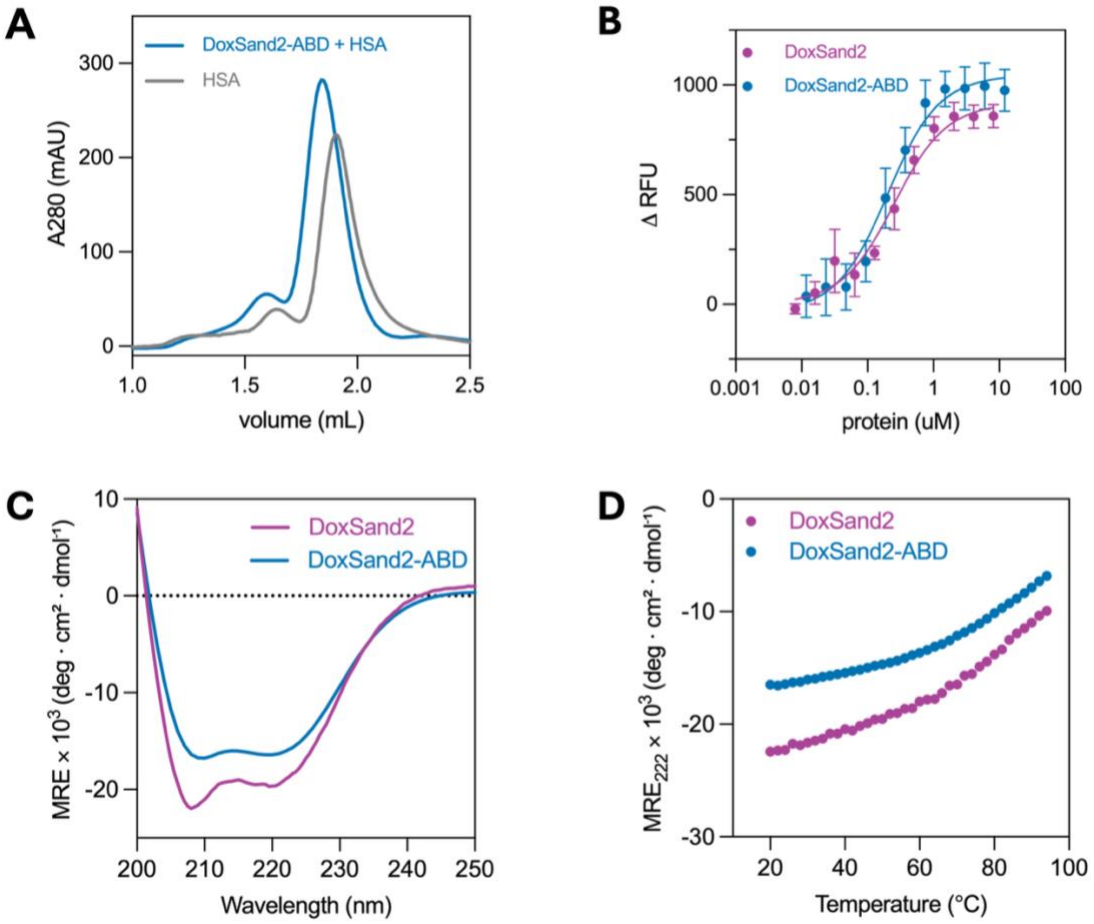

Figure 3.2 Biophysical characterization of DoxSand-ABD.
(A) Analytical size-exclusion chromatography of human serum albumin (HSA) alone (gray) and a preincubated mixture of DoxSand-ABD and HSA (blue). (B) Fluorescence titrations of DoxSand2 (magenta) and DoxSand-ABD (blue) with doxorubicin. Data were fit to a quadratic binding model. (C) Far-UV circular dichroism spectra of DoxSand2 and DoxSand-ABD, showing that the fusion protein retains the predominantly α-helical secondary structure of the parent scaffold. (D) Thermal denaturation monitored by circular dichroism at 222 nm.

A critical requirement for therapeutic development is maintaining structural integrity under physiological conditions. To obtain an initial assessment of biological stability, DoxSand2 and DoxSand2-ABD were incubated in mouse serum and analyzed over time by SDS-PAGE (Figure 3.3). Both proteins remained detectable following

incubation in serum (Figure 3.4). Although qualitative, these experiments suggest that the engineered proteins remain intact under serum conditions.

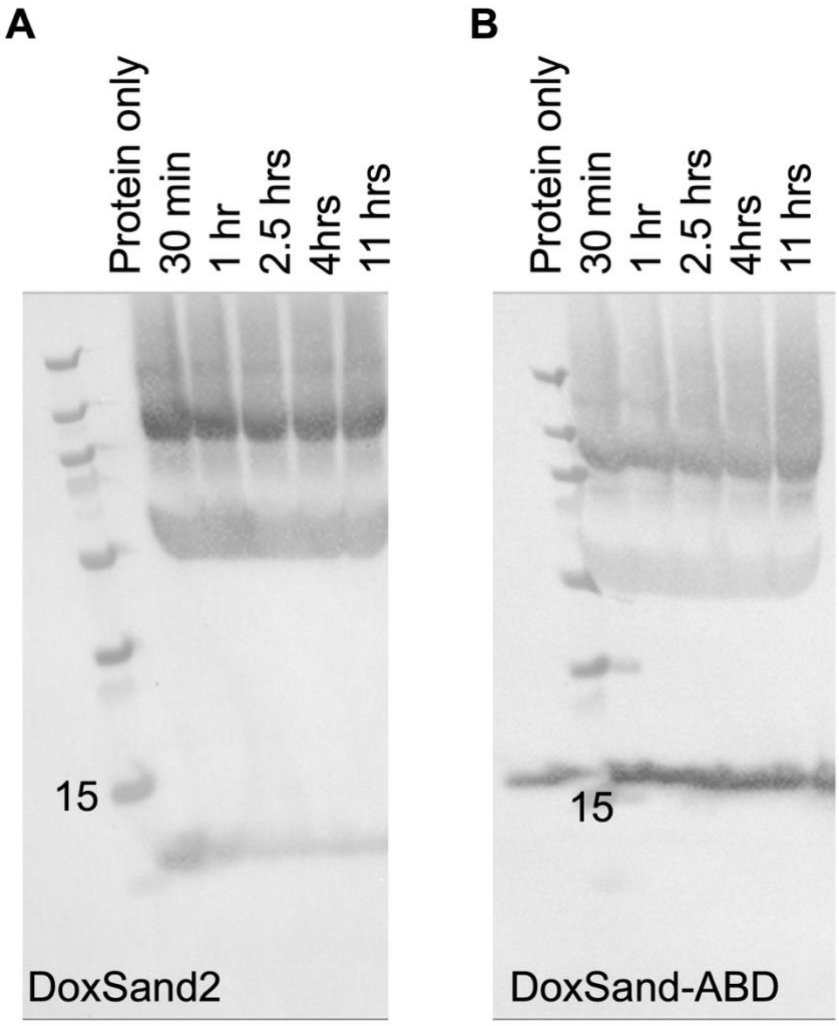

Figure 3.3 In vitro serum stability of DoxSand2 and DoxSand-ABD.
(A) Western blot analysis of DoxSand2 following incubation in mouse serum at 37 °C for 30 min, 1 h, 2.5 h, 4 h, and 11 h. (B) Western blot analysis of DoxSand-ABD under the same conditions. Proteins were detected using an HRP-conjugated anti-6×His tag antibody. Both proteins remained detectable throughout the incubation period, consistent with stability under the conditions examined.

3.3 CONCLUSIONS

The experiments described in this chapter represent initial efforts to translate DoxSand2 from a computationally designed ligand-binding protein toward a potential therapeutic scaffold. Fusion to an albumin-binding domain successfully introduced recognition of human serum albumin while preserving high-affinity binding to doxorubicin and maintaining the overall structural properties of the parent protein. Preliminary serum stability experiments further support the feasibility of this approach.

These results also highlight the substantial work that remains before therapeutic application can be realized. Ongoing studies are comparing the in vivo pharmacokinetics of DoxSand2, DoxSand-ABD, and free doxorubicin to determine how albumin binding influences the circulation of the protein and its bound cargo. Evaluation of tissue distribution and in vivo drug sequestration will also be important for assessing the therapeutic potential of DoxSand-ABD. Additional half-life extension strategies, including Fc fusion proteins and other pharmacokinetic engineering approaches, may further improve the therapeutic potential of DoxSand2.[111,112]

Together, these studies demonstrate that *de novo* designed small-molecule-binding proteins can be engineered beyond proof-of-concept molecular recognition toward properties required for therapeutic applications.

3.4 MATERIALS AND METHODS

3.4.1 Construction and Purification of DoxSand-ABD. Construction, expression, and purification of DoxSand-ABD fusion proteins were performed using the same procedures described for DoxSand2 in Chapter 2. Four fusion constructs were

generated by varying the orientation of the albumin-binding domain (ABDCon)[110] relative to DoxSand2 and the length of the connecting linker (7 or 15 amino acids). Following preliminary characterization, the construct consisting of DoxSand2 fused through a 15-residue linker to a C-terminal ABDCon domain (DoxSand-ABD) was selected for all subsequent studies.

3.4.2 Analytical Size-Exclusion Chromatography. Binding of DoxSand-ABD to human serum albumin (HSA) was evaluated by analytical size-exclusion chromatography using a Superdex 200 Increase 5/150 GL column (Cytiva) equilibrated in 1× phosphate-buffered saline (PBS, pH 7.4) at a flow rate of 0.2 mL min^{-1}. Human serum albumin (Sigma-Aldrich, Cat. No. A1653) was reconstituted in 1× PBS, and its concentration was determined using an extinction coefficient of 35,700 M^{-1} cm^{-1}. For HSA-only controls, 150 µg of HSA was injected onto the column. For complex formation experiments, 50 µg of DoxSand-ABD was mixed with 150 µg of HSA (1.3:1 molar ratio of DoxSand-ABD:HSA), incubated at room temperature for 30 min, and injected onto the column. Complex formation was assessed by comparing the elution profiles of HSA, DoxSand-ABD, and the preincubated protein mixture.

3.4.3. Doxorubicin Binding Assays. Binding of DoxSand-ABD to doxorubicin was measured by fluorescence spectroscopy using the same experimental procedures and data analysis described in Chapter 2.

3.4.4. Circular Dichroism Spectroscopy. Secondary structure and thermal stability of DoxSand-ABD were characterized by circular dichroism spectroscopy using the same experimental procedures described for DoxSand2 in Chapter 2.

3.4.5. Serum Stability Assays. Serum stability was evaluated by incubating DoxSand2 or DoxSand-ABD in mouse serum (Thermo Fisher Scientific, Cat. No. 10410) as previously described.[37,62] Samples were prepared by combining 12 µL of 2.08 mg mL^{-1} protein with 88 µL of mouse serum to give a final volume of 100 µL. For samples containing doxorubicin, 0.8 molar equivalents of doxorubicin were added before incubation, yielding a final protein concentration of 0.25 mg mL^{-1}. Samples were incubated at 37 °C and collected after 30 min, 1 h, 2.5 h, 4 h, and 11 h.

Protein stability was evaluated by SDS-PAGE followed by western blotting. Proteins were transferred using the iBlot 2 Dry Blotting System (Thermo Fisher Scientific, Cat. No. IB24002). Membranes were blocked with 1–3% bovine serum albumin (BSA) prior to incubation with an HRP-conjugated anti-6×His tag antibody (Abcam, Cat. No. ab1269). Immunoreactive bands were detected by chemiluminescence to qualitatively assess protein stability following serum incubation.

REFERENCES

(1) Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis*. *Proceedings of the National Academy of Sciences* **1958**, *44* (2), 98–104. https://doi.org/10.1073/pnas.44.2.98.

(2) Alberts, B. The Cell as a Collection of Protein Machines: Preparing the next Generation of Molecular Biologists. *Cell* **1998**, *92* (3), 291–294. https://doi.org/10.1016/s0092-8674(00)80922-8.

(3) Pawson, T.; Nash, P. Assembly of Cell Regulatory Systems Through Protein Interaction Domains. *Science* **2003**, *300* (5618), 445–452. https://doi.org/10.1126/science.1083653.

(4) Anfinsen, C. B. Principles That Govern the Folding of Protein Chains. *Science* **1973**, *181* (4096), 223–230. https://doi.org/10.1126/science.181.4096.223.

(5) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The Energy Landscapes and Motions of Proteins. *Science* **1991**, *254* (5038), 1598–1603. https://doi.org/10.1126/science.1749933.

(6) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338* (6110), 1042–1046. https://doi.org/10.1126/science.1219021.

(7) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450* (7172), 964–972. https://doi.org/10.1038/nature06522.

(8) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nat Rev Mol Cell Biol* **2019**, *20* (11), 681–697. https://doi.org/10.1038/s41580-019-0163-x.

(9) Listov, D.; Goverde, C. A.; Correia, B. E.; Fleishman, S. J. Opportunities and Challenges in Design and Optimization of Protein Function. *Nat Rev Mol Cell Biol* **2024**, *25* (8), 639–653. https://doi.org/10.1038/s41580-024-00718-y.

(10) Kortemme, T. *De novo* Protein Design—From New Structures to Programmable Functions. *Cell* **2024**, *187* (3), 526–544. https://doi.org/10.1016/j.cell.2023.12.028.

(11) Korendovych, I. V.; DeGrado, W. F. *De novo* Protein Design, a Retrospective. *Q Rev Biophys* **2020**, *53*, e3. https://doi.org/10.1017/S0033583519000131.

(12) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of *de novo* Protein Design. *Nature* **2016**, *537* (7620), 320–327. https://doi.org/10.1038/nature19946.

(13) Desjarlais, J. R.; Handel, T. M. *De novo* Design of the Hydrophobic Cores of Proteins. *Protein Sci* **1995**, *4* (10), 2006–2018. https://doi.org/10.1002/pro.5560041006.

(14) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. Protein Design by Binary Patterning of Polar and Nonpolar Amino Acids. *Science* **1993**, *262* (5140), 1680–1685. https://doi.org/10.1126/science.8259512.

(15) Regan, L.; DeGrado, W. F. Characterization of a Helical Protein Designed from First Principles. *Science* **1988**, *241* (4868), 976–978. https://doi.org/10.1126/science.3043666.

(16) Roy, S.; Ratnaswamy, G.; Boice, J. A.; Fairman, R.; McLendon, G.; Hecht, M. H. A Protein Designed by Binary Patterning of Polar and Nonpolar Amino Acids Displays Native-like Properties. *J. Am. Chem. Soc.* **1997**, *119* (23), 5302–5306. https://doi.org/10.1021/ja9700717.

(17) Ho, S. P.; DeGrado, W. F. Design of a 4-Helix Bundle Protein: Synthesis of Peptides Which Self-Associate into a Helical Protein. *J. Am. Chem. Soc.* **1987**, *109* (22), 6751–6758. https://doi.org/10.1021/ja00256a032.

(18) Walsh, S. T.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. Solution Structure and Dynamics of a *de novo* Designed Three-Helix Bundle Protein. *Proc Natl Acad Sci U S A* **1999**, *96* (10), 5486–5491. https://doi.org/10.1073/pnas.96.10.5486.

(19) Dahiyat, B. I.; Mayo, S. L. *De novo* Protein Design: Fully Automated Sequence Selection. *Science* **1997**, *278* (5335), 82–87. https://doi.org/10.1126/science.278.5335.82.

(20) Simons, K. T.; Kooperberg, C.; Huang, E.; Baker, D. Assembly of Protein Tertiary Structures from Fragments with Similar Local Sequences Using Simulated Annealing and Bayesian Scoring Functions1. *Journal of Molecular Biology* **1997**, *268* (1), 209–225. https://doi.org/10.1006/jmbi.1997.0959.

(21) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science* **2003**, *302* (5649), 1364–1368. https://doi.org/10.1126/science.1089427.

(22) Chu, A. E.; Lu, T.; Huang, P.-S. Sparks of Function by *de novo* Protein Design. *Nat Biotechnol* **2024**, *42* (2), 203–215. https://doi.org/10.1038/s41587-024-02133-2.

(23) Tinberg, C. E.; Khare, S. D.; Dou, J.; Doyle, L.; Nelson, J. W.; Schena, A.; Jankowski, W.; Kalodimos, C. G.; Johnsson, K.; Stoddard, B. L.; Baker, D. Computational Design of Ligand-Binding Proteins with High Affinity and Selectivity. *Nature* **2013**, *501* (7466), 212–216. https://doi.org/10.1038/nature12443.

(24) Polizzi, N. F.; DeGrado, W. F. A Defined Structural Unit Enables *de novo* Design of Small-Molecule–Binding Proteins. *Science* **2020**, *369* (6508), 1227–1233. https://doi.org/10.1126/science.abb8330.

(25) Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L. Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity. *Pharmacol Rev* **2004**, *56* (2), 185–229. https://doi.org/10.1124/pr.56.2.6.

(26) Mathivathanan, L.; Yang, G.; Leng, F.; Raptis, R. G. Crystal Structure and Conformational Analysis of Doxorubicin Nitrate. *Acta Cryst E* **2018**, *74* (3), 400–405. https://doi.org/10.1107/S2056989018002955.

(27) Dauparas, J.; Lee, G. R.; Pecoraro, R.; An, L.; Anishchenko, I.; Glasscock, C.; Baker, D. Atomic Context-Conditioned Protein Sequence Design Using LigandMPNN. *Nat Methods* **2025**, *22* (4), 717–723. https://doi.org/10.1038/s41592-025-02626-1.

(28) Krishna, R.; Wang, J.; Ahern, W.; Sturmfels, P.; Venkatesh, P.; Kalvet, I.; Lee, G. R.; Morey-Burrows, F. S.; Anishchenko, I.; Humphreys, I. R.; McHugh, R.; Vafeados, D.; Li, X.; Sutherland, G. A.; Hitchcock, A.; Hunter, C. N.; Kang, A.; Brackenbrough, E.; Bera, A. K.; Baek, M.; DiMaio, F.; Baker, D. Generalized Biomolecular Modeling and Design with RoseTTAFold All-Atom. *Science* **2024**, *384* (6693), eadl2528. https://doi.org/10.1126/science.adl2528.

(29) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.;

Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. Robust Deep Learning–Based Protein Sequence Design Using ProteinMPNN. *Science* **2022**, *378* (6615), 49–56. https://doi.org/10.1126/science.add2187.

(30) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.; De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. *De novo* Design of Protein Structure and Function with RFdiffusion. *Nature* **2023**, *620* (7976), 1089–1100. https://doi.org/10.1038/s41586-023-06415-8.

(31) Jing, X.; Wu, F.; Luo, X.; Xu, J. Single-Sequence Protein Structure Prediction by Integrating Protein Language Models. *Proceedings of the National Academy of Sciences* **2024**, *121* (13), e2308788121. https://doi.org/10.1073/pnas.2308788121.

(32) Wu, R.; Ding, F.; Wang, R.; Shen, R.; Zhang, X.; Luo, S.; Su, C.; Wu, Z.; Xie, Q.; Berger, B.; Ma, J.; Peng, J. High-Resolution *de novo* Structure Prediction from Primary Sequence. bioRxiv July 22, 2022, p 2022.07.21.500999. https://doi.org/10.1101/2022.07.21.500999.

(33) Deng, Y.; Sun, J. PymolFold: A PyMOL Plugin for API-Driven Structure Prediction and Quality Assessment. bioRxiv October 7, 2025, p 2025.10.03.680230. https://doi.org/10.1101/2025.10.03.680230.

(34) Grigoryan, G.; DeGrado, W. F. Probing Designability via a Generalized Model of Helical Bundle Geometry. *J Mol Biol* **2011**, *405* (4), 1079–1100. https://doi.org/10.1016/j.jmb.2010.08.058.

(35) Zhou, J.; Grigoryan, G. Rapid Search for Tertiary Fragments Reveals Protein Sequence–Structure Relationships. *Protein Science* **2015**, *24* (4), 508–524. https://doi.org/10.1002/pro.2610.

(36) Mann, S. I.; Nayak, A.; Gassner, G. T.; Therien, M. J.; DeGrado, W. F. *De novo* Design, Solution Characterization, and Crystallographic Structure of an Abiological Mn–Porphyrin-Binding Protein Capable of Stabilizing a Mn(V) Species. *J. Am. Chem. Soc.* **2021**, *143* (1), 252–259. https://doi.org/10.1021/jacs.0c10136.

(37) Lu, L.; Gou, X.; Tan, S. K.; Mann, S. I.; Yang, H.; Zhong, X.; Gazgalis, D.; Valdiviezo, J.; Jo, H.; Wu, Y.; Diolaiti, M. E.; Ashworth, A.; Polizzi, N. F.; DeGrado, W. F. *De novo* Design of Drug-Binding Proteins with Predictable Binding Energy and Specificity. *Science* **2024**, *384* (6691), 106–112. https://doi.org/10.1126/science.adl5364.

(38) Discovery, C.; Boitreaud, J.; Dent, J.; McPartlon, M.; Meier, J.; Reis, V.; Rogozhnikov, A.; Wu, K. Chai-1: Decoding the Molecular Interactions of Life. bioRxiv October 15, 2024, p 2024.10.10.615955. https://doi.org/10.1101/2024.10.10.615955.

(39) Kortemme, T. *De novo* Protein Design—From New Structures to Programmable Functions. *Cell* **2024**, *187* (3), 526–544. https://doi.org/10.1016/j.cell.2023.12.028.

(40) Chu, A. E.; Lu, T.; Huang, P.-S. Sparks of Function by *de novo* Protein Design. *Nat Biotechnol* **2024**, *42* (2), 203–215. https://doi.org/10.1038/s41587-024-02133-2.

(41) Pan, X.; Kortemme, T. Recent Advances in *de novo* Protein Design: Principles, Methods, and Applications. *J Biol Chem* **2021**, *296*, 100558. https://doi.org/10.1016/j.jbc.2021.100558.

(42) Polizzi, N. F.; DeGrado, W. F. A Defined Structural Unit Enables *de novo* Design of Small-Molecule–Binding Proteins. *Science* **2020**, *369* (6508), 1227–1233. https://doi.org/10.1126/science.abb8330.

(43) Galvin, C. V.; Guo, A. B.; Chen, M. N.; Alfonso, I. L.; Grisingher, D.; Kretschmer, S.; Kranthi, D.; Kelly, M. J. S.; Kortemme, T. A Hybrid Physics-Deep Learning Framework for Combinatorial *De novo* Design of Small-Molecule Binding Proteins. bioRxiv April 15, 2026, p 2026.04.12.717919. https://doi.org/10.64898/2026.04.12.717919.

(44) An, L.; Said, M.; Tran, L.; Majumder, S.; Goreshnik, I.; Lee, G. R.; Juergens, D.; Dauparas, J.; Anishchenko, I.; Coventry, B.; Bera, A. K.; Kang, A.; Levine, P. M.; Alvarez, V.; Pillai, A.; Norn, C.; Feldman, D.; Zorine, D.; Hicks, D. R.; Li, X.; Sanchez, M. G.; Vafeados, D. K.; Salveson, P. J.; Vorobieva, A. A.; Baker, D. Binding and Sensing Diverse Small Molecules Using Shape-Complementary Pseudocycles. *Science* **2024**, *385* (6706), 276–282. https://doi.org/10.1126/science.adn3780.

(45) Chen, Y.; Bhattacharya, S.; Bergmann, L.; Correy, G. J.; Tan, S. K.; Hou, K.; Biel, J. T.; Lu, L.; Bakanas, I.; Gestwicki, J. E.; Volkov, A. N.; Korendovych, I. V.; Polizzi, N. F.; Fraser, J. S.; DeGrado, W. F. Emergence of Specific Binding and Catalysis from a Designed Generalist Binding Protein. *Nat. Chem.* **2026**, *18* (7), 1334–1344. https://doi.org/10.1038/s41557-026-02125-6.

(46) Petrenas, R.; Ożga, K.; Chubb, J. J.; Romanyuk, A. V.; Alibhai, D.; McManus, J. J.; Leggett, G. J.; Scrutton, N. S.; Oliver, T. A. A.; Woolfson, D. N. Rapid Assessment of Chemical Complementarity of Ligands for Protein Design. bioRxiv May 23, 2026, p 2025.06.30.662286. https://doi.org/10.1101/2025.06.30.662286.

(47) Lee, G. R.; Pellock, S. J.; Norn, C.; Tischer, D.; Dauparas, J.; Anishchenko, I.; Mercer, J. A. M.; Kang, A.; Bera, A. K.; Nguyen, H.; Brackenbrough, E.; Sankaran, B.; Goreshnik, I.; Vafeados, D.; Roullier, N.; Han, H. L.; Coventry, B.; Haddox, H. K.; Liu, D. R.; Yeh, A. H.-W.; Baker, D. Small-Molecule Binding and Sensing with a Designed Protein Family. *Nat Commun* **2026**, *17* (1), 4533. https://doi.org/10.1038/s41467-026-70953-8.

(48) Fry, B.; Slaw, K.; Polizzi, N. F. Zero-Shot Design of Drug-Binding Proteins via Neural Iterative Selection−expansion. *Nature* **2026**, 1–13. https://doi.org/10.1038/s41586-026-10670-w.

(49) Stark, H.; Faltings, F.; Choi, M.; Xie, Y.; Hur, E.; O'Donnell, T.; Bushuiev, A.; Uçar, T.; Passaro, S.; Mao, W.; Reveiz, M.; Bushuiev, R.; Portnoi, T.; Pluskal, T.; Sivic, J.; Kreis, K.; Vahdat, A.; Ray, S.; Goldstein, J. T.; Savinov, A.; Hambalek, J. A.; Gupta, A.; Taquiri-Diaz, D. A.; Zhang, Y.; Snyder, S. J.; Hatstat, A. K.; Arada, A.; Kim, N. H.; Fan, H.; Tackie-Yarboi, E.; Boselli, D.; Schnaider, L.; Liu, C. C.; Li, G.-W.; Hnisz, D.; Sabatini, D. M.; DeGrado, W. F.; Wohlwend, J.; Corso, G.; Barzilay, R.; Jaakkola, T. BoltzGen: Toward Universal Binder Design. bioRxiv June 16, 2026, p 2025.11.20.689494. https://doi.org/10.1101/2025.11.20.689494.

(50) Dou, J.; Vorobieva, A. A.; Sheffler, W.; Doyle, L. A.; Park, H.; Bick, M. J.; Mao, B.; Foight, G. W.; Lee, M. Y.; Gagnon, L. A.; Carter, L.; Sankaran, B.; Ovchinnikov, S.;

Marcos, E.; Huang, P.-S.; Vaughan, J. C.; Stoddard, B. L.; Baker, D. *De novo* Design of a Fluorescence-Activating β-Barrel. *Nature* **2018**, *561* (7724), 485–491. https://doi.org/10.1038/s41586-018-0509-0.

(51) Glasgow, A. A.; Huang, Y.-M.; Mandell, D. J.; Thompson, M.; Ritterson, R.; Loshbaugh, A. L.; Pellegrino, J.; Krivacic, C.; Pache, R. A.; Barlow, K. A.; Ollikainen, N.; Jeon, D.; Kelly, M. J. S.; Fraser, J. S.; Kortemme, T. Computational Design of a Modular Protein Sense-Response System. *Science* **2019**, *366* (6468), 1024–1028. https://doi.org/10.1126/science.aax8780.

(52) Zhu, J.; Liang, M.; Sun, K.; Wei, Y.; Guo, R.; Zhang, L.; Shi, J.; Ma, D.; Hu, Q.; Huang, G.; Lu, P. *De novo* Design of Transmembrane Fluorescence-Activating Proteins. *Nature* **2025**, *640* (8057), 249–257. https://doi.org/10.1038/s41586-025-08598-8.

(53) Tinberg, C. E.; Khare, S. D.; Dou, J.; Doyle, L.; Nelson, J. W.; Schena, A.; Jankowski, W.; Kalodimos, C. G.; Johnsson, K.; Stoddard, B. L.; Baker, D. Computational Design of Ligand-Binding Proteins with High Affinity and Selectivity. *Nature* **2013**, *501* (7466), 212–216. https://doi.org/10.1038/nature12443.

(54) Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L. Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity. *Pharmacol Rev* **2004**, *56* (2), 185–229. https://doi.org/10.1124/pr.56.2.6.

(55) Cardinale, D.; Colombo, A.; Lamantia, G.; Colombo, N.; Civelli, M.; De Giacomi, G.; Rubino, M.; Veglia, F.; Fiorentini, C.; Cipolla, C. M. Anthracycline-Induced

Cardiomyopathy. *JACC* **2010**, *55* (3), 213–220. https://doi.org/10.1016/j.jacc.2009.03.095.

(56) Nadas, J.; Sun, D. Anthracyclines as Effective Anticancer Drugs. *Expert Opinion on Drug Discovery* **2006**, *1* (6), 549–568. https://doi.org/10.1517/17460441.1.6.549.

(57) Goins, B.; Phillips, W. T.; Bao, A. Strategies for Improving the Intratumoral Distribution of Liposomal Drugs in Cancer Therapy. *Expert Opin Drug Deliv* **2016**, *13* (6), 873–889. https://doi.org/10.1517/17425247.2016.1167035.

(58) Kohli, A. G.; Kivimäe, S.; Tiffany, M. R.; Szoka, F. C. Improving the Distribution of Doxil® in the Tumor Matrix by Depletion of Tumor Hyaluronan. *Journal of Controlled Release* **2014**, *191*, 105–114. https://doi.org/10.1016/j.jconrel.2014.05.019.

(59) Jain, R. K.; Stylianopoulos, T. Delivering Nanomedicine to Solid Tumors. *Nat Rev Clin Oncol* **2010**, *7* (11), 653–664. https://doi.org/10.1038/nrclinonc.2010.139.

(60) Chalkley, M. J.; Mann, S. I.; DeGrado, W. F. *De novo* Metalloprotein Design. *Nat Rev Chem* **2022**, *6* (1), 31–50. https://doi.org/10.1038/s41570-021-00339-5.

(61) Woolfson, D. N. A Brief History of *De novo* Protein Design: Minimal, Rational, and Computational. *Journal of Molecular Biology* **2021**, *433* (20), 167160. https://doi.org/10.1016/j.jmb.2021.167160.

(62) Chen, G.; Zhao, K.; Li, M.; Zhang, Y.; Li, T.; Yuan, J.; Lu, L.; Valdiviezo, J.; Liu, D.; Polizzi, N. F.; DeGrado, W. F.; Dang, B. Rapid Clearance of Achiral Small-Molecule Drugs Using *de novo*-Designed Proteins and Their Cyclic and Mirror-

Image Variants. *Nat. Biomed. Eng* **2025**, *9* (10), 1775–1787. https://doi.org/10.1038/s41551-025-01404-w.

(63) Mathivathanan, L.; Yang, G.; Leng, F.; Raptis, R. G. Crystal Structure and Conformational Analysis of Doxorubicin Nitrate. *Acta Crystallogr E Crystallogr Commun* **2018**, *74* (Pt 3), 400–405. https://doi.org/10.1107/S2056989018002955.

(64) Agudelo, D.; Bourassa, P.; Bérubé, G.; Tajmir-Riahi, H. A. Review on the Binding of Anticancer Drug Doxorubicin with DNA and tRNA: Structural Models and Antitumor Activity. *Journal of Photochemistry and Photobiology B: Biology* **2016**, *158*, 274–279. https://doi.org/10.1016/j.jphotobiol.2016.02.032.

(65) Madoori, P. K.; Agustiandari, H.; Driessen, A. J. M.; Thunnissen, A. W. H. Structure of the Transcriptional Regulator LmrR and Its Mechanism of Multidrug Recognition. *EMBO J* **2009**, *28* (2), 156–166. https://doi.org/10.1038/emboj.2008.263.

(66) Schmidt, K.; Gardill, B. R.; Kern, A.; Kirchweger, P.; Börsch, M.; Muller, Y. A. Design of an Allosterically Modulated Doxycycline and Doxorubicin Drug-Binding Protein. *Proceedings of the National Academy of Sciences* **2018**, *115* (22), 5744–5749. https://doi.org/10.1073/pnas.1716666115.

(67) Butcher, J.; Krishna, R.; Mitra, R.; Brent, R. I.; Li, Y.; Corley, N.; Kim, P. T.; Funk, J.; Mathis, S.; Salike, S.; Muraishi, A.; Eisenach, H.; Thompson, T. R.; Chen, J.; Politanska, Y.; Sehgal, E.; Coventry, B.; Zhang, O.; Qiang, B.; Didi, K.; Kazman, M.; DiMaio, F.; Baker, D. *De novo* Design of All-Atom Biomolecular Interactions with RFdiffusion3. bioRxiv: The Preprint Server for Biology November 19, 2025. https://doi.org/10.1101/2025.09.18.676967.

(68) Soskine, M.; Tawfik, D. S. Mutational Effects and the Evolution of New Protein Functions. *Nat Rev Genet* **2010**, *11* (8), 572–582. https://doi.org/10.1038/nrg2808.

(69) Maynard Smith, J. Natural Selection and the Concept of a Protein Space. *Nature* **1970**, *225* (5232), 563–564. https://doi.org/10.1038/225563a0.

(70) Rives, A.; Meier, J.; Sercu, T.; Goyal, S.; Lin, Z.; Liu, J.; Guo, D.; Ott, M.; Zitnick, C. L.; Ma, J.; Fergus, R. Biological Structure and Function Emerge from Scaling Unsupervised Learning to 250 Million Protein Sequences. *Proceedings of the National Academy of Sciences* **2021**, *118* (15), e2016239118. https://doi.org/10.1073/pnas.2016239118.

(71) Söding, J.; Lupas, A. N. More than the Sum of Their Parts: On the Evolution of Proteins from Peptides. *BioEssays* **2003**, *25* (9), 837–846. https://doi.org/10.1002/bies.10321.

(72) Alva, V.; Söding, J.; Lupas, A. N. A Vocabulary of Ancient Peptides at the Origin of Folded Proteins. *eLife* **2015**, *4*, e09410. https://doi.org/10.7554/eLife.09410.

(73) Yang, R.; Yang, H.; Davydenko, A.; Mawaldi, Z.; Kormos, R.; Myerscough, D.; Wu, Y.; DeGrado, W. F. The Emergence of Novel versus Known Three-Dimensional Structures from Random Sequences. *Proceedings of the National Academy of Sciences* **2026**, *123* (27), e2535076123. https://doi.org/10.1073/pnas.2535076123.

(74) Noji, M.; Fujiwara, T.; Sugita, Y.; Suzuki, Y. Thermally Tunable Heterodimeric Linkers Control Protein Tube Morphology. bioRxiv July 1, 2026, p 2025.11.24.689909. https://doi.org/10.1101/2025.11.24.689909.

(75) Zambaldi, V.; La, D.; Chu, A. E.; Patani, H.; Danson, A. E.; Kwan, T. O. C.; Frerix, T.; Schneider, R. G.; Saxton, D.; Thillaisundaram, A.; Wu, Z.; Moraes, I.; Lange, O.; Papa, E.; Stanton, G.; Martin, V.; Singh, S.; Wong, L. H.; Bates, R.; Kohl, S. A.; Abramson, J.; Senior, A. W.; Alguel, Y.; Wu, M. Y.; Aspalter, I. M.; Bentley, K.; Bauer, D. L. V.; Cherepanov, P.; Hassabis, D.; Kohli, P.; Fergus, R.; Wang, J. *De novo* Design of High-Affinity Protein Binders with AlphaProteo. arXiv September 12, 2024. https://doi.org/10.48550/arXiv.2409.08022.

(76) Pacesa, M.; Nickel, L.; Schellhaas, C.; Schmidt, J.; Pyatova, E.; Kissling, L.; Barendse, P.; Choudhury, J.; Kapoor, S.; Alcaraz-Serna, A.; Cho, Y.; Ghamary, K. H.; Vinué, L.; Yachnin, B. J.; Wollacott, A. M.; Buckley, S.; Westphal, A. H.; Lindhoud, S.; Georgeon, S.; Goverde, C. A.; Hatzopoulos, G. N.; Gönczy, P.; Muller, Y. D.; Schwank, G.; Swarts, D. C.; Vecchio, A. J.; Schneider, B. L.; Ovchinnikov, S.; Correia, B. E. One-Shot Design of Functional Protein Binders with BindCraft. *Nature* **2025**, *646* (8084), 483–492. https://doi.org/10.1038/s41586-025-09429-6.

(77) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.; De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. *De novo* Design of Protein Structure and Function with RFdiffusion. *Nature* **2023**, *620* (7976), 1089–1100. https://doi.org/10.1038/s41586-023-06415-8.

(78) Dey, S.; Roy, S.; Puneeth Kumar, Drgk. R.; Nalawade, S. A.; Singh, M.; Toraskar, S. U.; Mahapatra, S. P.; Gopi, H. N. Metal-Directed Hierarchical Superhelices from Hybrid Peptide Foldamers. *Chem. Commun.* **2025**, *61* (13), 2770–2773. https://doi.org/10.1039/d4cc05770h.

(79) George, K. L.; Horne, W. S. Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. *Acc. Chem. Res.* **2018**, *51* (5), 1220–1228. https://doi.org/10.1021/acs.accounts.8b00048.

(80) Chatterjee, S.; Roy, R. S.; Balaram, P. Expanding the Polypeptide Backbone: Hydrogen-Bonded Conformations in Hybrid Polypeptides Containing the Higher Homologues of Alpha-Amino Acids. *J R Soc Interface* **2007**, *4* (15), 587–606. https://doi.org/10.1098/rsif.2006.0203.

(81) Guichard, G.; Huc, I. Synthetic Foldamers. *Chem. Commun.* **2011**, *47* (21), 5933–5941. https://doi.org/10.1039/c1cc11137j.

(82) Wang, P. S. P.; Schepartz, A. β-Peptide Bundles: Design. Build. Analyze. Biosynthesize. *Chem. Commun.* **2016**, *52* (47), 7420–7432. https://doi.org/10.1039/c6cc01546h.

(83) Balaram, P. Non-Standard Amino Acids in Peptide Design and Protein Engineering. *Current Opinion in Structural Biology* **1992**, *2* (6), 845–851. https://doi.org/10.1016/0959-440X(92)90110-S.

(84) Azam, A.; Mallart, S.; Illiano, S.; Duclos, O.; Prades, C.; Maillère, B. Introduction of Non-Natural Amino Acids Into T-Cell Epitopes to Mitigate Peptide-Specific T-Cell Responses. *Front Immunol* **2021**, *12*, 637963. https://doi.org/10.3389/fimmu.2021.637963.

(85) Bautista, A. D.; Craig, C. J.; Harker, E. A.; Schepartz, A. Sophistication of Foldamer Form and Function in Vitro and in Vivo. *Curr Opin Chem Biol* **2007**, *11* (6), 685–692. https://doi.org/10.1016/j.cbpa.2007.09.009.

(86) Lin, Q.; Lan, H.; Ma, C.; Stendall, R. T.; Shankland, K.; Musgrave, R. A.; Horton, P. N.; Baldauf, C.; Hofmann, H.-J.; Butts, C. P.; Müller, M. M.; Cobb, A. J. A. Crystal Structure and NMR of an α,δ-Peptide Foldamer Helix Shows Side-Chains Are Well Placed for Bifunctional Catalysis: Application as a Minimalist Aldolase Mimic. *Angewandte Chemie International Edition* **2023**, *62* (36), e202305326. https://doi.org/10.1002/anie.202305326.

(87) Maayan, G.; Ward, M. D.; Kirshenbaum, K. Folded Biomimetic Oligomers for Enantioselective Catalysis. *Proceedings of the National Academy of Sciences* **2009**, *106* (33), 13679–13684. https://doi.org/10.1073/pnas.0903187106.

(88) Ganesh Kumar, M.; Gopi, H. N. γ- and β-Peptide Foldamers from Common Multifaceted Building Blocks: Synthesis and Structural Characterization. *Org. Lett.* **2015**, *17* (19), 4738–4741. https://doi.org/10.1021/acs.orglett.5b02263.

(89) Ganesh Kumar, M.; Thombare, V. J.; Katariya, M. M.; Veeresh, K.; Raja, K. M. P.; Gopi, H. N. Non-Classical Helices with Cis Carbon–Carbon Double Bonds in the Backbone: Structural Features of α,γ-Hybrid Peptide Foldamers. *Angewandte Chemie International Edition* **2016**, *55* (27), 7847–7851. https://doi.org/10.1002/anie.201602861.

(90) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637. https://doi.org/10.1002/bip.360221211.

(91) van Kempen, M.; Kim, S. S.; Tumescheit, C.; Mirdita, M.; Lee, J.; Gilchrist, C. L. M.; Söding, J.; Steinegger, M. Fast and Accurate Protein Structure Search with Foldseek. *Nat Biotechnol* **2024**, *42* (2), 243–246. https://doi.org/10.1038/s41587-023-01773-0.

(92) Nguyen, T. N.; Nguyen, T. T.; Nghiem, T. H. L.; Nguyen, D. T.; Tran, T. T. H.; Vu, D.; Nguyen, T. B. N.; Nguyen, T. M. H.; Nguyen, V. T.; Nguyen, M. H. Optical Properties of Doxorubicin Hydrochloride Load and Release on Silica Nanoparticle Platform. *Molecules* **2021**, *26* (13), 3968. https://doi.org/10.3390/molecules26133968.

(93) Kabsch, W. XDS. *Acta Cryst D* **2010**, *66* (2), 125–132. https://doi.org/10.1107/S0907444909047337.

(94) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Cryst D* **2019**, *75* (10), 861–877. https://doi.org/10.1107/S2059798319011471.

(95) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J Appl Cryst* **2007**, *40* (4), 658–674. https://doi.org/10.1107/S0021889807021206.

(96) Yamashita, K.; Wojdyr, M.; Long, F.; Nicholls, R. A.; Murshudov, G. N. GEMMI and Servalcat Restrain REFMAC5. *Acta Cryst D* **2023**, *79* (5), 368–373. https://doi.org/10.1107/S2059798323002413.

(97) Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. Electronic Ligand Builder and Optimization Workbench (eLBOW): A Tool for Ligand Coordinate and Restraint Generation. *Acta Cryst D* **2009**, *65* (10), 1074–1080. https://doi.org/10.1107/S0907444909029436.

(98) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Cryst D* **2010**, *66* (4), 486–501. https://doi.org/10.1107/S0907444910007493.

(99) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall III, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Science* **2018**, *27* (1), 293–315. https://doi.org/10.1002/pro.3330.

(100) Grabowski, M.; Langner, K. M.; Cymborowski, M.; Porebski, P. J.; Sroka, P.; Zheng, H.; Cooper, D. R.; Zimmerman, M. D.; Elsliger, M.-A.; Burley, S. K.; Minor, W. A Public Database of Macromolecular Diffraction Experiments. *Acta Cryst D* **2016**, *72* (11), 1181–1193. https://doi.org/10.1107/S2059798316014716.

(101) *BLAST Statistics: The Expect Value*. https://www.nlm.nih.gov/ncbi/workshops/2023-08_BLAST_evol/e_value.html?utm_source=chatgpt.com (accessed 2026-07-14).

(102) Karlin, S.; Altschul, S. F. Methods for Assessing the Statistical Significance of Molecular Sequence Features by Using General Scoring Schemes. *Proceedings of the National Academy of Sciences* **1990**, *87* (6), 2264–2268. https://doi.org/10.1073/pnas.87.6.2264.

(103) Neese, F. Software Update: The ORCA Program System—Version 6.0. *WIREs Computational Molecular Science* **2025**, *15* (2), e70019. https://doi.org/10.1002/wcms.70019.

(104) Neese, F. The SHARK Integral Generation and Digestion System. *Journal of Computational Chemistry* **2023**, *44* (3), 381–396. https://doi.org/10.1002/jcc.26942.

(105) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *Journal of Computational Chemistry* **2011**, *32* (7), 1456–1465. https://doi.org/10.1002/jcc.21759.

(106) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104. https://doi.org/10.1063/1.3382344.

(107) Sleep, D.; Cameron, J.; Evans, L. R. Albumin as a Versatile Platform for Drug Half-Life Extension. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830* (12), 5526–5534. https://doi.org/10.1016/j.bbagen.2013.04.023.

(108) Dennis, M. S.; Zhang, M.; Meng, Y. G.; Kadkhodayan, M.; Kirchhofer, D.; Combs, D.; Damico, L. A. Albumin Binding as a General Strategy for Improving the

Pharmacokinetics of Proteins. *J Biol Chem* **2002**, *277* (38), 35035–35043. https://doi.org/10.1074/jbc.M205854200.

(109) Gapizov, S. Sh.; Petrovskaya, L. E.; Shingarova, L. N.; Kryukova, E. A.; Boldyreva, E. F.; Lukashev, E. P.; Yakimov, S. A.; Svirshchevskaya, E. V.; Dolgikh, D. A.; Kirpichnikov, M. P. Fusion with an Albumin-Binding Domain Improves Pharmacokinetics of an Avβ3-Integrin Binding Fibronectin Scaffold Protein. *Biotechnology and Applied Biochemistry* **2019**, *66* (4), 617–625. https://doi.org/10.1002/bab.1762.

(110) Jacobs, S. A.; Gibbs, A. C.; Conk, M.; Yi, F.; Maguire, D.; Kane, C.; O'Neil, K. T. Fusion to a Highly Stable Consensus Albumin Binding Domain Allows for Tunable Pharmacokinetics. *Protein Eng Des Sel* **2015**, *28* (10), 385–393. https://doi.org/10.1093/protein/gzv040.

(111) Nilsen, J.; Aaen, K. H.; Benjakul, S.; Ruso-Julve, F.; Greiner, T. U.; Bejan, D.; Stensland, M.; Singh, S.; Schlothauer, T.; Sandlie, I.; Andersen, J. T. Enhanced Plasma Half-Life and Efficacy of Engineered Human Albumin-Fused GLP-1 despite Enzymatic Cleavage of Its C-Terminal End. *Commun Biol* **2025**, *8* (1), 810. https://doi.org/10.1038/s42003-025-08249-8.

(112) Kuo, T. T.; Baker, K.; Yoshida, M.; Qiao, S.-W.; Aveson, V. G.; Lencer, W. I.; Blumberg, R. S. Neonatal Fc Receptor: From Immunity to Therapeutics. *J Clin Immunol* **2010**, *30* (6), 777–789. https://doi.org/10.1007/s10875-010-9468-4.

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.

Signed by:
7EF186E1FA1A4B0...

Author Signature

07/28/2026

Date