Lessons From Cyclosporine A: Structural Determinants of Conformation-Switching and Passive Membrane Penetration

LESSONS FROM CYCLOSPORINE A: STRUCTURAL DETERMINANTS OF CONFORMATION-SWITCHING AND PASSIVE MEMBRANE PENETRATION

A dissertation presented

by

ERIN MALLORY MAY

to

COMMITTEE ON HIGHER DEGREES IN CHEMICAL BIOLOGY

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Chemical Biology

Harvard University
Cambridge, MA

September 2015

Dissertation Advisor: Professor Gregory Verdine Erin Mallory May

LESSONS FROM CYCLOSPORINE A: STRUCTURAL DETERMINANTS OF CONFORMATION-SWITCHING AND PASSIVE MEMBRANE PENETRATION

ABSTRACT

The structural complexity of 'beyond-rule-of-5' compounds, such as peptide macrocycles, may facilitate access to additional biological target space beyond the enzymatic active site. Naturally occurring cyclic peptides, in particular, exhibit a wide variety of unusual and potent biological activities. Cyclosporine A (CsA), the canonical 'rule-breaking' molecule, is heavily N-methylated, and it is widely held that this contributes to its unexpectedly high oral bioavailability. This has been a major motivator of the use of N-methylation to enhance the ability of cyclic peptides to cross lipid bilayers. The evidentiary basis of this, stems largely from structural studies which predict that all the backbone amide bonds of CsA that are not engaged in transannular hydrogen bonds in the suspected 'cell-permeable conformation,' are otherwise N-methylated. However, to our knowledge, no direct, systematic analysis of the effect of N-methylation on cell permeability and target engagement of CsA has been reported. In order to gain a deeper understanding of this structurally-significant molecule, we

have engineered a new tool for the study CypA ligands. A destabilized variant of CsA presenting protein, Cyclophilin A (CypA) that persists in the cell only when bound to a stabilizing ligand, was used to evaluate the target binding and cell permeability of selected CsA anlaogues. Subsequent structure-activity studies show that seemingly small structural changes, such as the loss of a single N-methyl group or the introduction of a less conformationally-constrained side chain, can result in large conformational changes that alter CsA's ability to engage CypA. We also propose that this tool may be used to identify or evaluate novel CypA-binding scaffolds with altered target specificity.

TABLE OF CONTENTS

Abstract……..……..iii

Table of Contents…….……..……v

Acknowledgements……..…….……...vi

Abbreviations……..……..…….....xi

List of Tables and Figures.....……...……xiv

INTRODUCTION.....……...……1

CHAPTER 1

Engineering a selectively destabilized mutant of Cyclophilin A……..…..……..…35

CHAPTER 2

Structure-activity relationship studies of cyclosporine analogues……...……71

CHAPTER 3

Towards identification of naïve immunophilin-binding peptide scaffolds…..……..102

CONCLUSION….....……...……...……125

ACKNOWLEDGEMENTS

I've always valued the saying, "If you're the smartest person in the room, then you're in the wrong room." Well I can say that, since moving to Cambridge five years ago, I've very consistently been in the right room. During my graduate career I have been questioned, questioned myself, and experienced more failures than the collective years prior to my PhD. But I also believe that this has led to immense personal and professional growth. For challenging me, teaching me, collaborating with me, and supporting me, I am truly grateful to the following people.

I am indebted to my academic mentors for their kindness, guidance, and inspiring intellect. I am grateful to my advisor, Greg Verdine, who encourages his students to take on challenging questions that address unmet needs in science and for the benefit of society. In this respect, he leads by example, finding ways to do transformative science in both academic and industrial settings. I am grateful for his confidence in me, even when I questioned my own abilities. I felt very strongly throughout my PhD that he always had my best interest in mind; and through his guidance, I feel encouraged to find the questions that excite me and pursue them with passion.

I am also thankful to my academic advisory committee: David Liu, James Bradner, and Joshua Krtizer. I believe that my committee went above and beyond what was required of them to help me find direction during the challenging middle years of my PhD. They gave me both excellent advice about experiments, as well as cautionary criticism about the

direction of challenging projects, likely saving me time and sanity. They challenged my data and in doing so, improved the quality of my work. Yet the discussions were always very supportive and encouraging. I am also sometimes overwhelmed, in ways good and bad, but mostly good, that I get to have a group of such brilliant people in the same room to discuss my work. I am also truly thankful to Sara Buhrlage for agreeing to join my thesis committee in the eleventh hour, and enabling me to complete my defense.

I have had the good fortune of collaborating with many talented and hardworking scientists, several of whom contributed directly to my projects. Khian Hong Pua and I joined the lab several months apart and it has been a great personal pleasure to share space and conversations with him during this time. Two years ago, he and I began to collaborate on a project that formed the basis for most of the work presented herein. I have learned a great deal from his rigorous and systematic approach to science. Khian and I both were fortunate enough to spend time at Warp Drive Bio, LLC., working in an industrial setting with a passion for basic science. Here, I had access to expertise that added immense value to my projects. Several scientists at Warp Drive have contributed directly to this work including Dylan Stiles, Ping Wang, Mat Sowa, Dan Gray, and Ashutosh Jogalekar. Their specific contributions are outlined in the acknowledgements sections following each chapter. I am also thankful to Chi Zhang for teaching me cell culture techniques and Keith Robison for helpful advice on the design and analysis of genomic sequencing experiments.

I feel so fortunate to have had the most wonderful group of lab-mates at Harvard. The Verdine lab members, through their hard work and scientific insight, have provided a

challenging and enriching academic experience. But moreover, they have been excellent lab citizens as well as kind and supportive friends. Even as I sit here now, glued to my computer screen in the midst of thesis writing, two separate lab-mates have brought me food, and many more have offered to help in any way that they are able. In particular, to So Youn Shim, Kazuhiro Hayashi, Lan Wang, Yvonne Nagel, Khian Hong Pua, Qian Chu, Matt Kolar, Jan Gong, Sam Jiang, John McGee and Jerry Hilinski, I am truly indebted to you, and look forward to seeing the many amazing things that you all accomplish both now and in the future. I am also so thankful for our lab administrator, Lydia Carmosino, without whom the lab would rapidly come to a standstill. After nearly 25 years, Lydia still supports and cares for each individual in the lab, and goes out of her way to help them succeed in their ambitions.

My classmates in the ChemBio program, Sixun Chen, Stephanie Wang, Ahmed Badran, Kevin Davis, and adopted program member, Margie Li, will no doubt become lifelong friends. I'm not sure whether this bond stems from commiserating over the ups and downs of grad school, engaging in meandering discussions about science, or the abundance free food and drink shared at conferences and program events over the years. I am thankful to have shared this journey with them. I am also grateful to program administrators KeyAnna Schmiedl, Samantha Reed, and Jason Millberg. They are largely responsible for establishing the tight knit community that we enjoy in the program.

My friends outside of the lab have provided the much needed means to escape my own head, and gain perspective on my work through time away from the bench. I really can't express

how much they have contributed to my success and happiness throughout my PhD. While my lab-mates and academic mentors may have pushed me over figurative mountains, my friends in the climbing and outdoors communities have literally pushed me up mountains and cliff faces throughout the world. Additionally, when not in the lab, the most likely place that I could be found was nearby at the HLS or Rhino volleyball courts. I am so thankful for the competition, sportsmanship, laughter, and friendship shared over literally thousands of passes, sets, kills, and digs. I also feel very fortunate to have spent the past year with my boyfriend, Alex. He has been a constant and caring source of support as I navigate an exciting and uncertain period of my life. It is a gift to have someone, whose opinion you value so highly, to discuss your ambitions, and know that they are met without judgement, but instead with confidence, interest, and enthusiasm.

Lastly, I want to express my heartfelt gratitude to my family. To my brother, Evan, my sister, Blaize and my Dad, you have shaped me and put my interests above your own. I appreciate your support of my scientific ambitions, but value even more, that you have and will be there for me regardless of the outcome of my experiments. To my mom, I couldn't have achieved this without you. Without any knowledge of the underlying science, my mom consoled me through every frustration, and matched ten-fold my own enthusiasm for every little success. For running around to find dry ice when I demanded it for a middle school science fair, for editing everything I'd ever written up until the point where the scientific jargon consumed the other words, for your selflessness and unconditional support, I share any success I've achieved with you.

To all my family, friends, and mentors, I offer my most sincere appreciation.

I dedicate this work to my friend Jo,

who delighted equally in accomplishments big and small.

ABBREVIATIONS

3T3	NIH3T3 mouse fibroblast cell line
ASA	Solvent-accessible surface area
Bmt	(4R)-4-[(E)-2-butenyl]-4-methyl-L-threonine
bRo5	Beyond rule of five
Caco-2	Human colorectal carcinoma
Cn(A/B)	Calcineurin (subunit A or B)
CsA	Cyclosporine A
CsAlog	Cyclosporine A analogue
CypA	Cyclophilin A
cDD	Cyclophilin A destabilizing domain
DD	Destabilizing domain
DHFR	Dihydrofolate reductase
DSF	Differential scanning fluorimetry
ER	Estrogen receptor
FACS	Fluorescence-activated cell sorting
Hsp90	Heat shock protein 90
LogP	Octanol-water partition coefficient
$LogP_e$	Effective permeability (PAMPA assay)
Ro5	Rule of five
FBLD	Fragment-based lead discovery
FKBP12	FK506 binding protein 12

GPCR	G-protein-coupled receptor
HDM2	Human protein double minute 2
HPV E2	Human papilloma virus transcription factor E2
HTS	High throughput screening
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor
ICW	In cell western blotting
IL-2(R)	Interleukin-2 (receptor)
MG132	Proteasome inhibitor
mTOR	Mammalian target of rapamycin
MW	Molecular weight
NFAT	Nuclear factor of activated T-lymphocytes
NME	New molecular entity
NMR	Nuclear magnetic resonance
NRotB	Number of rotatable bonds
NRPS	Non-ribosomal peptide synthetase
PAMPA	Parallel artificial membrane permeability assay
PPIs	Protein-protein interactions
PPIA	Gene encoding wild-type cyclophilin A
PSA	Polar surface area
R&D	Research and development
Ro5	Rule of 5
SAR	Structure-activity relationship

SPR	Surface plasmon resonance
TNF	Tumour-necrosis factor
WT	Wild-type

LIST OF FIGURES

Figure 0-1	Distribution of mechanistically-defined drugs by target class.	4
Figure 0-2	Structures of selected small molecule and bRo5 inhibitors of PPIs.	11
Figure 0-3	Distribution and physicochemical properties of approved drugs and clinical candidates with MW > 500Da.	13
Figure 0-4	CsA inhibits calcineurin resulting in immunosuppression.	16
Figure 0-5	Xtal structure of CypA-CsA-calcineurin inhibitory trimeric complex.	17
Figure 0-6	Chemical structure of CsA.	18
Figure 0-7	Xtal structures of CsA in bound and free states shows rearrangement of conformation and hydrogen bonds.	20
Figure 1-1	Surface-borrowing affinity enhancement enables formation of inhibitory trimeric complex.	37
Figure 1-2	Microbial natural products bind immunophilin proteins to enable target inhibition.	38
Figure 1-3	Destabilizing domains as a general method to conditionally control protein stability.	40
Figure 1-4	Generalizable workflow for engineering a protein destabilizing domain.	41
Figure 1-5	FACS enrichment of cyclophilin A destabilizing domains (cDDs).	43
Figure 1-6	Ligand dependent stability of mCherry-cDD mutant fusion proteins.	46
Figure 1-7	Highly enriched mutations in cDD population identified by deep sequencing.	48
Figure 1-8	Fluorescence-based characterization of CypA mutants that display CsA dependent stability.	52
Figure 1-9	Fluorescence microscopy of 3T3 cells stably expressing cDD-mCherry fusions.	53

Figure 1-10	Kinetic analysis by surface plasmon resonance implies that cDD mutants I10F/G14C and S40N retain strong interaction with CsA.	55
Figure 1-11	Differential scanning fluorimetry reveals that cDDs are destabilized relative to WT CypA.	57
Figure 1-12	Location of I10F/G14C mutations on cDD, cyclophilin A destabilizing domain.	58
Figure 1-13	Degradation of cDD-mCherry is not mediated by the 26S proteasome.	60
Figure 1-14	Development of cytosensor assay compatible with HTS of CypA ligands.	62
Figure 2-1	Features of primary and secondary protein structure.	74
Figure 2-2	Workflow for CDD cytosensor analysis of CsAlogs.	77
Figure 2-3	Location of modified CsAlog residues MeLeu4, Val5, and MeLeu6 with respect to CypA.	79
Figure 2-4	Effect of structural modifications to CsA on cell penetration and target engagement as determined by cDD cytosensor assay.	81
Figure 2-5	Workflow for structure-activity studies of CsA analogues with reduced bioactivity.	83
Figure 2-6	Structure-activity relationship studies of desmethyl CsAlogs.	84
Figure 2-7	Predicted lowest energy confomers of desmethyl analogues in a high dielectric environment.	86
Figure 2-8	Structure-activity relationship studies of sarcosine CsAlogs.	88
Figure 2-9	Structure-activity relationship studies of D-amino acid CsAlogs.	89
Figure 2-10	Validation of SPR data with CypA-binding CsAlog.	90
Figure 2-11	Correlation of computationally-predicted and experimentally-derived permeabilities.	92
Figure 2-11	Summary for CsAlog SAR studies.	93

Figure 3-1.	Protein-protein interactions may be mediated by chemical inducers of dimerization.	105
Figure 3-2	Distinct protein binding domains of rapamycin and FK506.	107
Figure 3-3	Design of immunophilin-binding cyclic peptide scaffolds.	108
Figure 3-4	Synthetic scheme for solid-phase synthesis of cyclic peptides.	109
Figure 3-5	Solid-phase synthesis of peptides with immunophilin-binding motifs.	110
Figure 3-6	Summary of peptide binding as observed by real-time mass spectrometry.	112
Figure 3-7	Orbitrap mass spectrometry allows for identification of peptide-protein complexes.	113
Figure 3-8	Top-down proteomics shows that FGPDLP-containing peptides selectively bind FKBP12.	114
Figure 3-9	Binding assays confirm low affinity of protein-peptide complexes.	116

INTRODUCTION

Whether an academic chemist, investor in the biotechnology sector, or an informed healthcare consumer, it's likely you've read articles that warn of a long-standing decline in the output from the pharmaceutical industry. The so called, "innovation gap," describes the phenomenon of increased investment with stalling returns, measured in FDA approval of New Molecular Entities (NMEs) brought to market. One 2012 estimate suggests that the number of new drugs approved per billion US dollars spent on research and development (R&D) has roughly halved every 9 years since 1950, falling around 80-fold in inflation adjusted terms.[1] This has led to many to questions featuring the now ubiquitous "fruit" trope, such as, "Have we picked all the *low-hanging fruit*?" – meaning have earlier drug discovery efforts exhausted the more chemically tractable drug targets. Certainly those in the challenging business of drug discovery have good cause to deny the existence of low hanging fruit altogether. Additionally, measuring innovation in terms of NME output is not necessarily the best surrogate, since it does not account for the differential impact of first-in-class drugs as compared to me-too drugs. Still, the important question which follows logically is: What new techniques or applications of existing methods will allow us to reach the *high-hanging fruit?*[2] Both questions have spawned a great deal of debate and remarkable innovation in the field of chemical biology.

To address these questions, one must consider the size and breadth of all proposed disease targets and the corresponding pharmacopeia. In 2006, Hopkins *et al.* estimated 1,357 unique drugs (excluding salt forms, formulations, etc.) were approved for use by the FDA.[3] Of those, protein molecular targets believed to be the sole or major route through which the drug achieves its efficacy were assigned to 1,065 drugs. While these drugs act through 324 distinct

molecular targets, a striking feature emerges when one considers their distribution amongst gene families. Notably, more than 50% of drugs target only four key gene families: Class I G-protein-coupled receptors (GPCRs), nuclear receptors, ligand-gated ion channels, and voltage-gated ion channels **(Fig. 0-1)**. Additionally, 60% of targets exist on the cell surface, as compared to an estimated 22% proteins in the human genome. Structural analysis suggests that approximately 130 '*privileged druggable domains'* cover all current drug targets, which is only a fraction of the estimated 10,000 folds and more than 16,000 protein families.[4,5] Additionally, of the 361 NMEs approved between 1989 and 2001, only 6% targeted a previously undrugged domain, while 76% targeted a previously drugged domain; the remainder have either unknown targets (4%) or are not believed to have a distinct molecular target (17%).[3] This suggests that challenges faced in drug discovery stem not from a lack of disease-relevant proteins, but rather the limited ability to selectively target a wide-range of protein folds and locations.

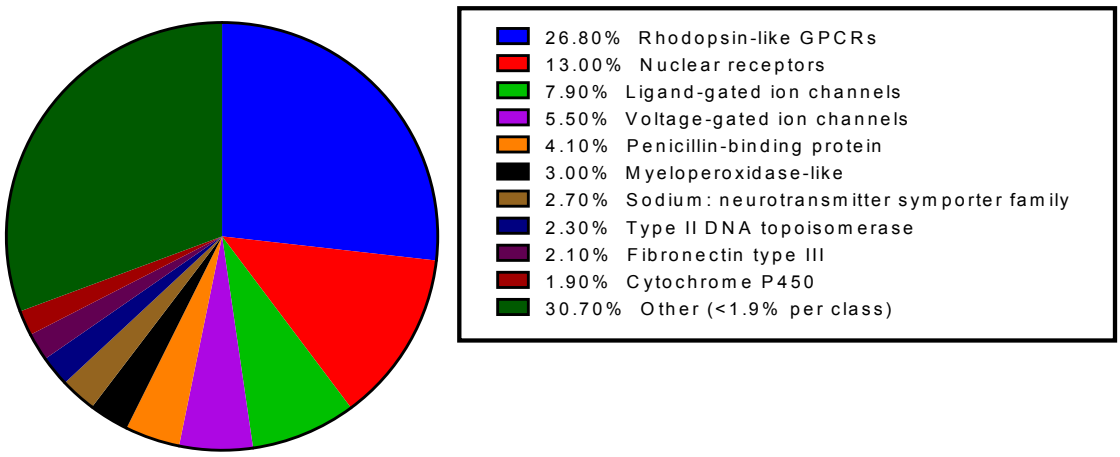


Figure 0-1. Distribution of mechanistically-defined drugs by target class. The protein family share of all FDA-approved drugs (as of 2006) displayed for the top ten families. Data obtained from Overington *et al.,* 2006.[3]

Expanding the scope of "druggable" targets: Targeting protein-protein interactions

While a few gene families, such as those discussed above, account for a disproportionately large number of approved therapeutics, there are structural and mechanistic classes of great import that remain underrepresented. There is no class of macromolecular interactions that rivals the complexity, diversity, and regulatory impact of interactions between proteins.[6,7] From intercellular communication to signal transduction, growth and metabolism, and regulated cell death, protein-protein interactions (PPIs) are central to most cellular processes. Conservative estimates place the total number of PPIs at >100,000, indicating that there is no shortage of disease-relevant interactions to target. And just last year, an international collaboration led by Frederick Roth and Marc Vidal published the *human interactome map*,

the largest-scale map to date of interactions between proteins encoded by the human genome. It describes about 14,000 direct interactions, and has been used to newly predict dozens of genes involved in cancer.[8] While new interactome targets require extensive biochemical and phenotypic validation, development of parallel advances in chemistry will be required to capitalize on this potential.

The ability to selectively target protein interfaces with chemical probes remains a long-sought-after goal. However, the properties of these interfaces make this extremely challenging from a chemical standpoint. Interactions are characterized by large, flat contact surface areas, and lack either a defined hydrophobic pocket or a small molecule ligand, which could otherwise be used as a starting point for drug development.[9] Yet, the previously held notion that these targets are 'undruggable' has shifted significantly with the emergence of more and more successful examples of PPI inhibitors, expanding significantly the scope of potential drug targets.[10]

The anatomy of protein-protein interactions

A thoughtful analysis of the physico-chemical properties of protein interfaces reveals features that complicate traditional drug discovery, but which may also suggests rational approaches to targeting these interactions.

The size of a protein-protein interface is defined as the surface area on both partners that becomes inaccessible to solvent due to protein-protein association. This area is the sum of

the solvent-accessible surface areas (ASAs) of the isolated components less that of the complex.[11] Contact surfaces involved in PPIs are large (~1,500-3,000 Å^2)[11,12] compared to protein-small molecule interactions (~300-1,000 Å ^{2}).[13,14] However, alanine-scanning mutagenesis has been used to define a small subset of residues that contribute the majority of the free energy of binding. These 'hotspots' are defined by sites where alanine-substitution increases the binding energy by greater than 2.0 kcal/mol.[16] Protein hotspots constitute less than half of the interfacial residues, and are frequently found either clustered at the center of the contact surface, or in more recently defined 'hotsegments,' which are extended continuous epitopes that contribute the majority of binding energy.[15,16] The energetic contribution of hotspot residues results from a combination of the following: desolvation of hydrophobic residues, geometric complementarity enabling Van der Walls interactions, favourable electrostatics, and may involve salt bridges between charged residues. Notably, the same protein hotspot can often adapt to mediate binding in different structural contexts. Promiscuous proteins with multiple binding partners may adopt varied structural conformations at the contact surface, but the same hotspot residues are often responsible for the thermodynamic driving force of association.[11,15]

Another frequently discussed hurdle in targeting protein surfaces is the absence of a hydrophobic pocket. This limits structure-based drug design, including *in silico* screening of ligand libraries or rational design[17]. However, molecular dynamics simulations suggest that protein surfaces can form transient hydrophobic pockets that are absent from crystal structures. This is not surprising given the dynamic nature of protein conformations. Small

molecules that either capture, or induce the formation of hydrophobic cavities in protein surfaces may have sufficient affinity to regulate protein-protein interactions.

Progress towards small-molecule inhibitors of protein-protein interactions

A growing number of small molecule success stories indicate that protein interfaces may be more tractable than previously thought.[18] By 2005, about a half-dozen small molecules had been reported to bind with the affinities one would expect for drug leads (<100nM), at binding sites defined by high-resolution structures.[10] In the decade since then, more than 40 PPIs have now been targeted with synthetic molecules, and several inhibitors have reached clinical trials.[19,20]

This success has been enabled by parallel advances in computational and chemical technologies well-suited for PPI inhibitor screening. For instance, fragment-based lead discovery (FBLD) uses biophysical methods including crystallography, surface plasmon resonance (SPR), and nuclear magnetic resonance (NMR), or disulfide tethering, to identify small, low-complexity molecules that bind weakly to subsites on the protein surface.[21,22] A fragment-based approach guided by X-ray structures was used to identify a small molecule that disrupts the interaction between the cytokine interleukin-2 (IL-2) and the α-chain of the IL-2 receptor (IL2-2Rα)[23,24]. Interestingly, this molecule, SP4206, was assembled before the structure of IL-2 in complex with IL-2Rα had been solved. Subsequent studies have shown that SP4206 is not an accurate atomic mimic of the receptor, and induces an alternate conformation of IL-2. Evidently, the dynamic nature of protein surfaces make structure-

activity relationships (SAR) difficult to predict, but also indicates that more than one chemical solution may exist for tight binding.[25]

The effectiveness of high throughput screening (HTS) of small molecule libraries has been questioned, since small molecule libraries used to identify inhibitors bound to hydrophobic pockets have typically yielded low hit rates when screening against PPIs. Yet HTS has also yielded some notable lead-like candidates. HTS was used to identify inhibitors of human protein double minute 2 (HDM2), a cancer target which increases p53 degradation.[26] Subsequent medicinal chemistry efforts at F.Hoffman-La Roche identified a series of tetra-substituted imidazoles, later named Nutlins.[27] Independent HTS at Janssen identified a series of benzodiazepinediones with high affinity for HDM2[28]. Structural data indicates that both small molecule scaffolds bind to the same region of HDM2 as the alpha-helical portion of p53, despite having dissimilar structures.[27,28] HTS has also been successfully applied to identify drug leads that disrupt the interaction of human papilloma virus transcription factor E2 (HPV E2) with the viral E1 helicase, as well as inhibitors of the trimeric tumour-necrosis factor (TNF) by increasing the kinetics of monomer dissociation.[29,30] While these examples, and others, dispel the myth that HTS cannot be used to discover small molecule leads for protein interactions, it remains to be seen whether this strategy will have broad utility. This is generally attributed to the properties of existing libraries, which are dominated by past drug-discovery efforts, and therefore occupy a limited chemical space.[31] With the expanding number of small molecules that bind protein surfaces, we may begin to derive lessons that guide library synthesis for screening resulting in increased hit-rates.

Inhibition of challenging intracellular targets with 'Beyond the Rule of 5' compounds

With advances in small-molecule targeting of PPIs, it is important to consider the distinction between ligandability and our ability to convert PPI inhibitors into drugs, which must consider complex pharmacokinetics, cell permeability for intracellular targets and bioavailability for oral adminsitration. For many compounds, passive diffusion through a cell membrane is central to both transcellular absorption in the gut and subsequent entry into cells at the site of action for intracellular targets. Compounds must be able to partition into the membrane after desolvation from the polar aqueous environment and then diffuse across the nonpolar membrane and resolvate upon exit at the other side.[32] This process is difficult to study *in vivo* but can be assessed using *in vitro* model systems such as human colorectal carcinoma (Caco-2) cells, a model of gastrointestinal absorption, and the parallel artificial membrane permeability assay (PAMPA), for a reductionist assessment of diffusion across a lipid bilayer.[33,34]

The processes of desolvation, diffusion, and resolvation required for passive permeability across a cell membrane are a function of several properties including size, polarity, lipophilicity, and conformational dynamics.[35,36] In order for a compound to be orally bioavailable, these properties need to be balanced; for example, highly polar compounds fail to desolvate and enter the membrane, whereas highly lipophilic compounds may not dissolve or fail to partition out of the membrane.[37] To address the complex process of absorption, in 1997, Lipinski *et al.* reported a set of guidelines designed to enable *in silico* prediction of whether a compound populates the chemical space where solubility and permeability are

likely to allow for oral absorption. According to the Rule of Five (Ro5), approximately 90% of oral compounds pass three of the four following rules: (1) Molecular weight (MW) ≤ 500 Da; (2) Octanol-water partition coefficient (logP), a measure of lipophilicity, is ≤ 5 and ≥ 0; (3) Hydrogen bond acceptors (HBA) ≤ 10; and (4) Hydrogen bond donors (HBD) ≤ 5.[38] Later, additional computationally predicted properties such as polar surface area (PSA) ≤ 140Å^2 and the number of rotatable bonds (NRotB) ≤ 10–20 have been added.[35] The Ro5 and later additions were heavily used in early stages of drug discovery to identify oral, drug-like compounds, and clearly favour smaller, more conformationally rigid molecules **(Fig. 0-2A)**.[39] In fact analyses of average ligand properties by target class reveal that the highly explored classes previously discussed, such as ligand gated ion channels (MW 359 Da, cLogP 3.0) and GPCRs (MW 378 Da, cLogP 3.8) are well-represented within the Ro5 space.[40,41]

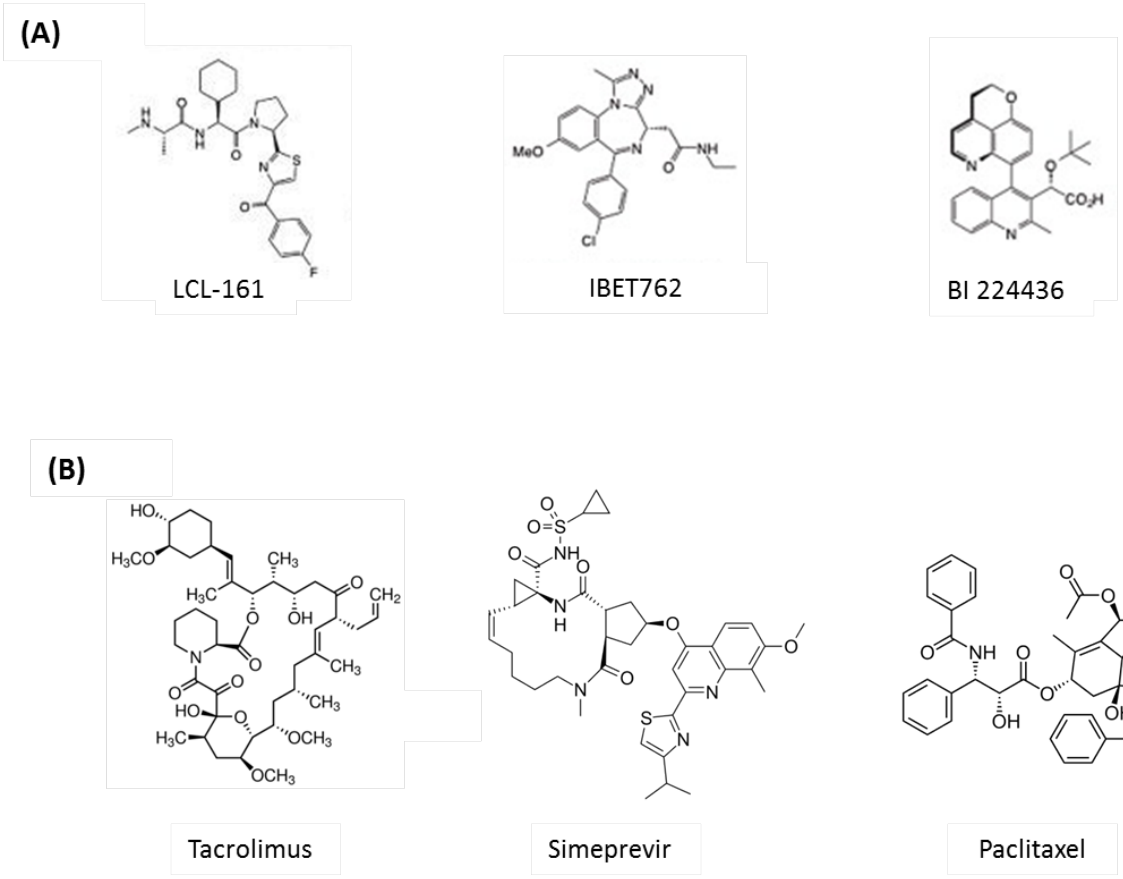
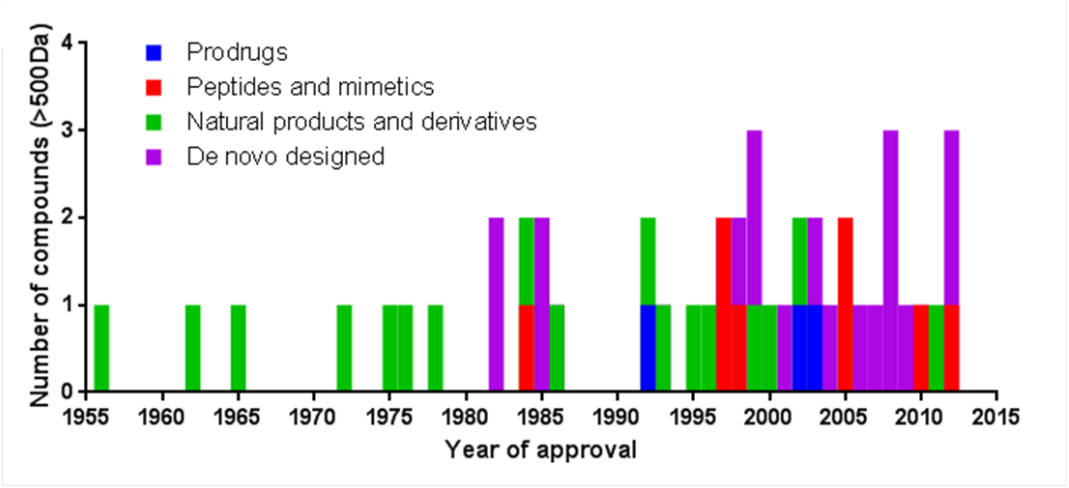


Figure 0-2. Structures of selected small molecule and bRo5 inhibitors of PPIs. (A) Ro5-adherent small molecule inhibitors of BIR domains (LCL161, Novartis); bromodomains (I-BET762, Glaxo Smith-Klein); and HIV integrase (BI 224436, Boehring-Ingelheim). (B) bRo5 inhibitors of calcineurin (tacrolimus or FK506); HCV NS3/4 protease (Simeprevir, Janssen); and a stabilizer of microtubule assembly (Paclitaxel).

However, the fact that a number of orally bioavailable drugs exist far beyond these boundaries, have led to a rejection of strict cutoffs on the basis that adherence is preventing the exploration of larger, therapeutically-relevant molecules.[39,42-43] Kihlberg and colleagues have recently published a comprehensive analysis of drugs and clinical candidates in the so-called *beyond rule-of-five* (bRo5) space **(Fig. 0-2b)**. Their dataset included 85 orally-bioavailable compounds that have at least one calculated physicochemical property within the following ranges: MW > 700 Da, cLogP < 0 or > 7.5, HBD > 5, HBA > 10, PSA > 200 $Å^2$, or NRotB > 20.[32] This set of bRo5 compounds is highly enriched in natural products and derivatives, peptidomimetics, and macrocyclic structures **(Fig. 0-3)**. Hence, there is potential for higher degrees of novelty, diversity, and complexity in bRo5 compounds. This suggests that the likelihood of finding high-affinity cell-permeable ligands for difficult target classes such as PPIs may be enhanced if ligands with physicochemical properties at the borders or beyond Ro5 were explored to a larger extent.

(A)



(B)

Class	Delivery	N	MW (Da)	cLogP	HBD	HBA	PSA (Å^2)	NRotB	Fsp^3
all	Oral	226	642(621-662)	4.1(3.8-4.3)	2.6(2.4-2.9)	7.7(7.3-8.1)	137(130-145)	10.4(9.8-11.1)	0.49(0.46-0.52)
	Parenteral	259	1059(983-1136)	-1.4(-2.3- -0.5)	10.5(9.2-11.8)	15.6(14.2-17.0)	357(320-394)	21.9(19.3-24.6)	0.52(0.50-0.54)
peptides & mimetics	Oral	25	815(700-929)	3.2(2.6-3.9)	4.5(3.5-5.5)	8.7(7.6-9.9)	199(169-229)	15.1(12.5-17.7)	0.56(0.50-0.62)
	Parenteral	100	1473(1336-1610)	-6.9(-8.2- -5.6)	19.3(17.1-21.5)	22.1(19.7-24.5)	582(518-645)	35.6(30.2-40.9)	0.56(0.54-0.58)
natural products	Oral	51	713(670-756)	3.7(3.0-4.3)	3.3(2.8-3.8)	10.2(9.2-11.2)	168(153-182)	9.2(8.0-10.5)	0.68(0.62-0.73)
	Parenteral	73	883(765-1002)	0.7(-0.6-2.0)	6.8(5.1-8.4)	14.3(11.8-16.9)	270(219-322)	13.6(10.1-17.2)	0.59(0.54-0.63)
de novo designed	Oral	136	590(574-606)	4.4(4.1-4.7)	2.1(1.8-2.3)	6.6(6.2-6.9)	113(107-120)	9.3(8.8-9.9)	0.40(0.37-0.43)
	Parenteral	75	733(662-803)	3.0(2.1-3.9)	3.7(2.8-4.6)	9.4(7.8-11.0)	176(146-206)	12.2(10.5-14.0)	0.39(0.34-0.43)
prodrugs	Oral	14	572(534-611)	3.8(2.2-5.4)	2.1(1.6-2.7)	7.6(6.6-8.6)	149(135-164)	17.1(13.6-20.6)	0.62(0.49-0.74)
	Parenteral	11	699(558-840)	4.6(1.6-7.6)	2.2(0.3-4.0)	7.2(5.2-9.2)	126(85-168)	19.6(14.9-24.2)	0.60(0.53-0.66)
CNS	Oral	6	601(474-727)	4.4(2.1-6.7)	0.8(-0.6-2.2)	5.0(3.9-6.2)	81(56-106)	8.8(5.6-12.1)	0.42(0.36-0.49)

Figure 0-3. Distribution and physicochemical properties of approved drugs and clinical candidates with MW > 500Da. (A) Approved oral drugs (44 total) with MW >500Da as a function of year of FDA approval, and compound class. (B) Physicochemical property distribution of all approved and clinical candidates with MW >500Da. Molecular weight (MW); calculated octanol-water partition coefficient (cLogP); hydrogen bond donors (HBD); hydrogen bond acceptors (HBA); polar surface area (PSA); number of rotatable hydrogen bonds (NRotB); and fraction of sp3 carbon atoms (Fsp^3). Data obtained from from Kihlberg *et al.,* 2014.[32]

Conformation-dependent permeability of cyclic peptide natural products

The structural complexity of many natural products sets them apart from common synthetic drugs, allowing them to access a biological target space that lies beyond the enzyme active site and receptors targeted by conventional small molecule drugs. Naturally occurring cyclic peptides, in particular, exhibit a wide variety of unusual and potent biological activities. Many of these compounds penetrate cells by passive diffusion and some, like the clinically important drug cyclosporine A, are orally bioavailable. [44] These natural products tend to have molecular weights and polar group counts that put them firmly in the bRo5 space. Because of their size and complexity, cyclic peptides occupy a chemical "middle space" in drug discovery, between small molecule Ro5 compounds and larger non-permeant biologics, that may provide useful scaffolds for modulating more challenging biological targets such as PPIs and allosteric binding sites.[45] Therefore, there is great value in elucidating the relationship between structure and pharmacokinetic behavior, especially cell permeability, in cyclic peptides.

Cyclosporine A: The archetypal 'rule-breaking' molecule

Cyclosporine A (CsA) is a natural product originally isolated from the fungus *Tolypocladium inflatum* found in a soil sample obtained in 1969 from a mountain plateau in Norway by Sandoz biologist, Hans Peter Frey.[46] This is a typical example of the serendipitous nature of many early natural product discoveries. Originally isolated for its antifungal activities, it was found to exert only marginal effects *in vivo*. It was later characterized by Borel *et al.* as an

immunosuppressant, and has since been widely used in organ transplantation to reduce T-cell-mediated rejection.[47]

CsA lowers T-cell activity by inhibition of intracellular target, calcineurin, a Ca^{2+}/calmodulin-dependent serine threonine protein phosphatase **(Fig 0-4)**.[44] Activated calcineurin dephosphorylates regulatory sites on several transcription factors including nuclear factor of activated T-lymphocytes (NFAT). CsA prevents the dephosphorylation of NFAT and its subsequent translocation from the cytoplasm to the nucleus in an IL-2-mediated process. Inhibition thereby prevents activation of promoters of T-cell activation and overall immune response.

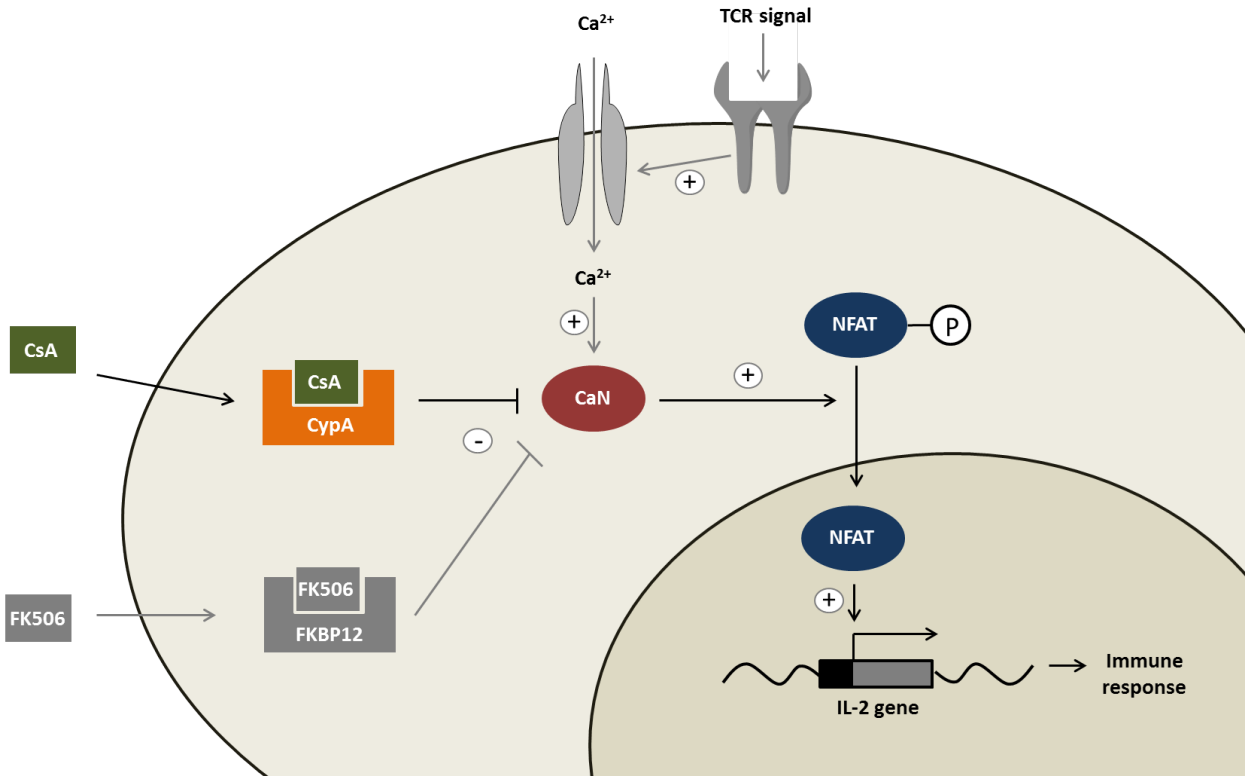


Figure 0-4. CsA inhibits calcineurin resulting in immunosupression. Inhibition of calcineurin is mediated by the CsA/CypA binary complex. FK506 (tacrolimus) complexed with FKBP12 also inhibits calcineurin. Cyclosporine A (CsA, green); cyclophilin A (CypA, orange); FK506-binding protein 12 (FKBP12); calcineurin (CaN, red); TCR (T-cell receptor); nuclear factor of activated T-lymphocytes (NFAT, blue); interleukin-2 (IL-2).

Interestingly, CsA exerts its inhibitory effect by an unusual mechanism of action termed composite surface borrowing. CsA binds first with 30nM affinity to cytoplasmic immunophilin protein, Cyclophilin A (CypA), an 18kD peptidyl-prolyl *cis/trans*-isomerase. Only then does the composite surface have sufficient affinity to form an inhibitory trimeric complex and sequester its target, calcineurin **(Fig. 0-5)**. Thus, CSA has the unusual function of inducing an inhibitory protein-protein interaction.[44] Surface-borrowing has also been

observed with another *cis/trans*-isomerase, FK506 binding protein 12 (FKBP12), which binds immunophilin ligands, rapamycin and FK506, and leads to trimeric inhibition of mammalian target of rapamycin (mTor) and calcineurin, respectively.

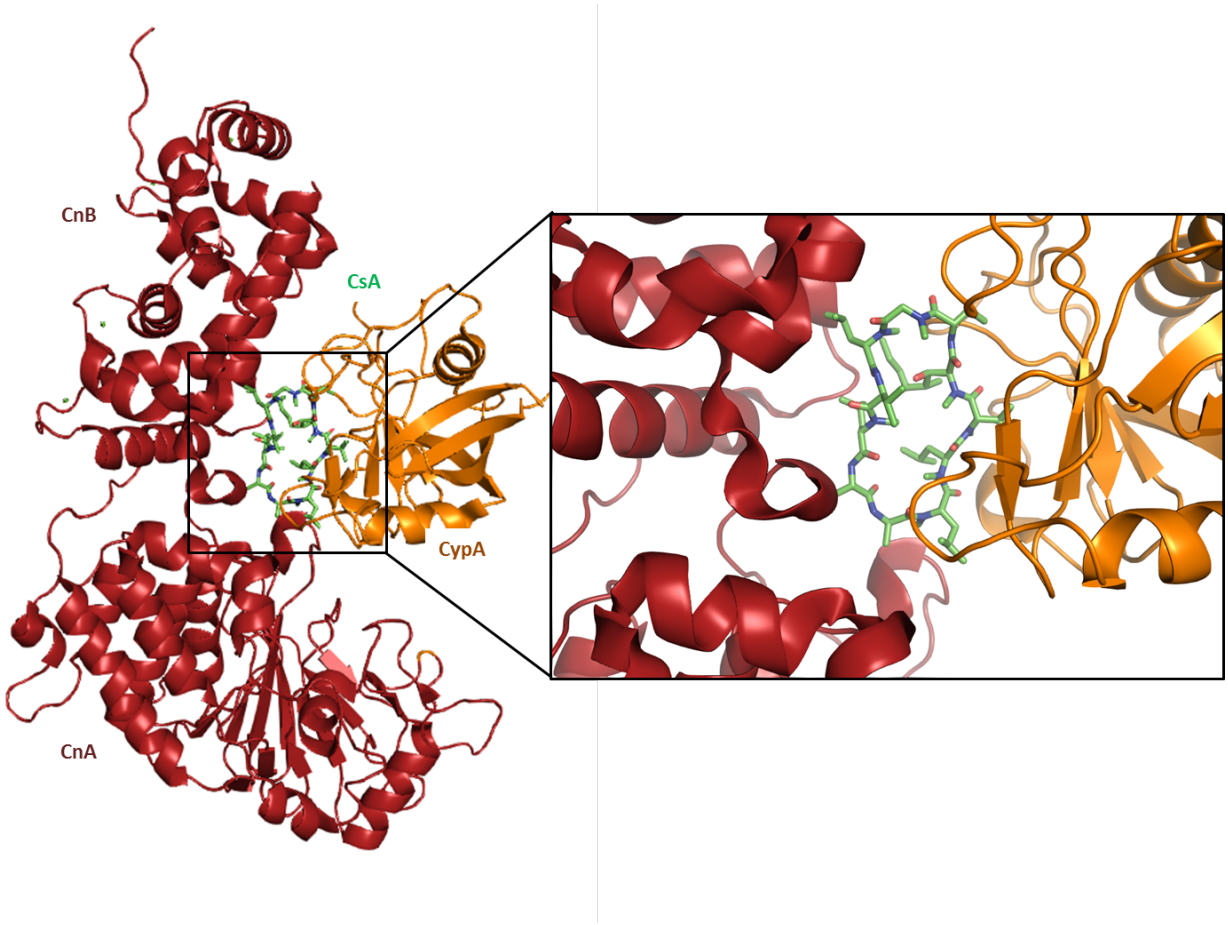


Figure 0-5. Xtal structrure of CypA-CsA-calcineurin inhibitory trimeric complex. Calcineurin subunits A (CnA, red) and B (CnB, red) in complex with cyclosporine A (CsA, green) and cyclophilin A (CypA, orange). [PDB ID: 2z6w]

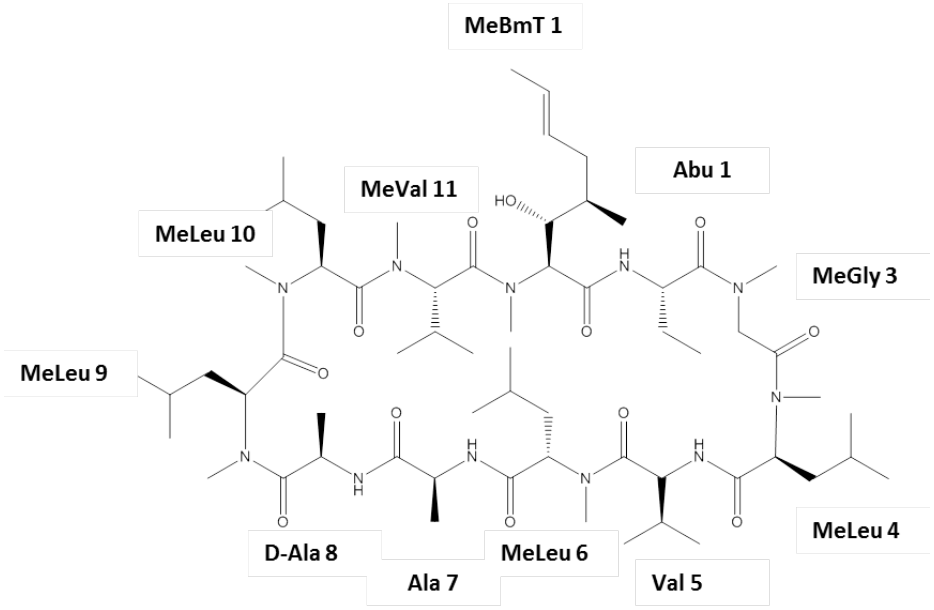


Figure 0-6. Chemical structure of CsA. Shown with standard numbering of amino acid residues. Uncommon amino acids include L-α-aminobutyric acid (Abu) and (4R)-4-[(E)-2-butenyl]-4,N-dimethyl-L-threonine (Bmt), along with 7 N-methylated amino acids and a single D-amino acid.

CsA is also a structurally-significant molecule, especially in the context of bRo5 compounds. Like many cyclic peptide natural products, CsA is the biosynthetic product of non-ribosomal peptide synthetases (NRPS). It is an 11-amino acid, homodetic, non-polar cyclic peptide that contains seven N-methyl groups, a single D-amino acid, and a rare non-proteinogenic prenylated threonine residue, (4R)-4-[(E)-2-butenyl]-4,N-dimethyl-L-threonine (Bmt) **(Fig. 0-6)**. The structure of CsA has been determined both in the free state, by X-ray crystallography and NMR, and in the ternary complex with cyclophilin and calcineurin.[48-50] In the free state, in both the crystal structure and the NMR solution structure in $CDCl_3$, CsA adopts a rectangular conformation characterized by three transannular hydrogen bonds and

an external, γ-turn forming hydrogen bond **(Fig. 0-7a)**. In contrast, in the crystal structure of the ternary complex with cyclophilin and calcineurin, CsA adopts a more open conformation in which backbone polar groups point outward, forming intermolecular hydrogen bonds with both protein partners **(Fig. 0-7b)**.[50] Interestingly, the non-polar conformation of CsA observed in $CDCl_3$ was also significantly populated in Xtal structures in 50% aqueous methanol; however, this may be an artifact of crystallographic methods.[51] This type of conformational flexibility may be common to nonpolar cyclic peptides and has been identified in at least two other natural products.[52] Despite great interest in these molecules, the structure-permeability relationship resulting in the unexpected passive permeability of CsA is still not fully understood.[59] However, in the past 5 years, several systematic analyses of cell-permeable natural products, combined with the evaluation of synthetic cyclic peptide models, have contributed significantly to our understanding of the structure-permeability relationship of large macrocycles.[52-59]

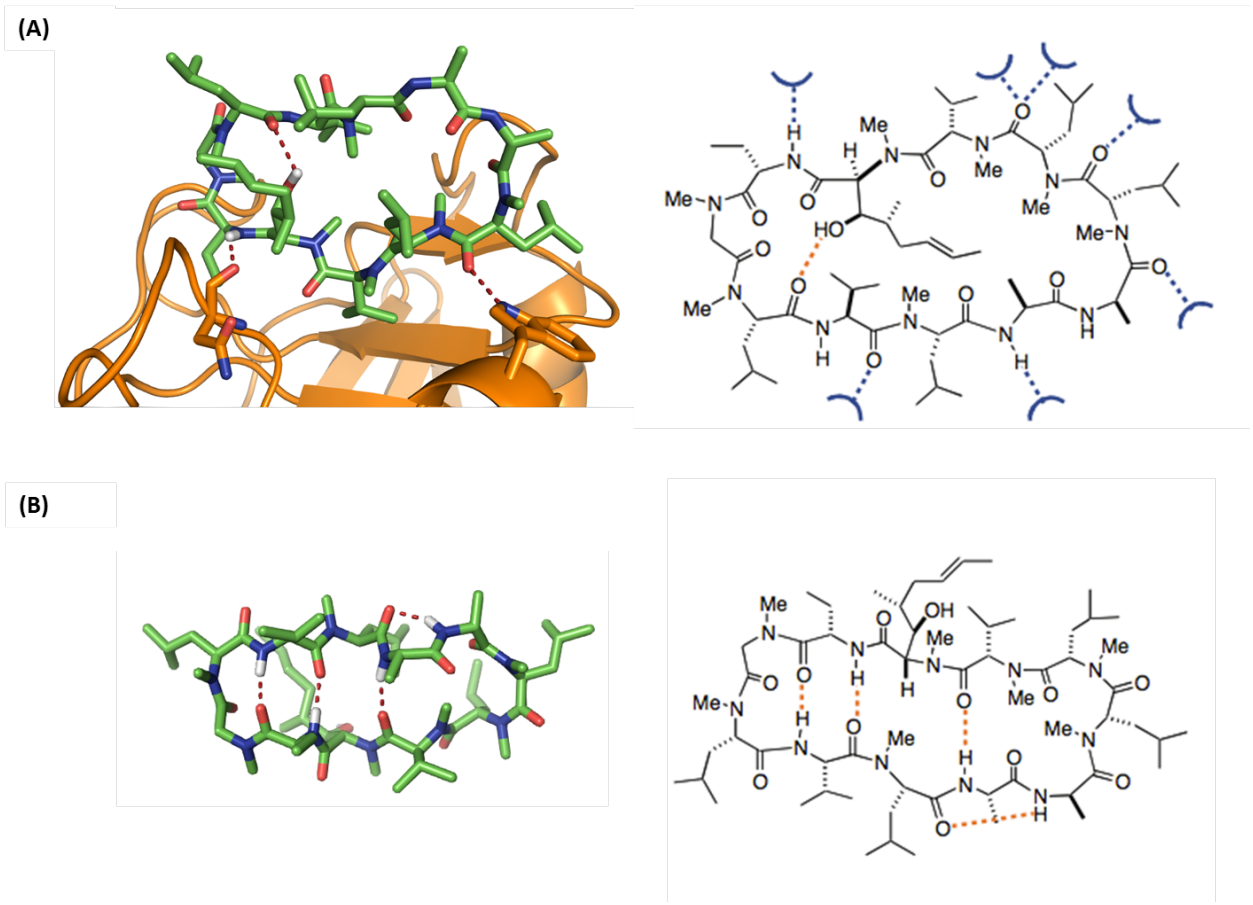


Figure 0-7. Xtal structures of CsA in bound and free states shows rearrangement of conformation and hydrogen bonds. (A) Xtal structure of CsA (green) bound to CypA (orange) and Calcineurin (not shown) in an aqueous environment. Shown with predicted intramolecular and intermolecular hydrogen bonds. [PDB ID: 2z6w] (B) Xtal structure of CsA (green) in organic solvent ($CDCl_3$). [CSD ID: DEK SAN]

Beyond cyclosporine A: Structural features of cell-permeable peptide natural products

CsA has long been considered a prototypical 'rule-breaker': A cell permeable and orally bioavailable drug that violates common rules-of-thumb for drug-likeness. However, a recent study by Matthew Jacobsen, Scott Lokey, and co-workers has demonstrated that many other large cyclic peptides with different scaffolds are also passively permeable. Using PAMPA as a model for diffusion across a lipid-bilayer, they tested the permeability of 32 cyclic peptides for which structural information (X-ray crystallography or NMR) of the compound or a closely-related analogue had been reported.[59] Within the dataset, 23 were classified as passively permeable, with a PAMPA effective permeability ($logP_e$) of greater than -6.4. CsA ($logP_e$ -5.01) and several related analogues were included in their analysis, the results of which will be discussed in greater detail in **Chapter 2**.

Several notable conformational and structural features in this series that may contribute to passive permeability include depsipeptide linkages, N-methylation, thiazole/oxazole rings, and intramolecular hydrogen bonding, all of which either eliminate or sequester backbone amide NH groups. Eight of the permeable compounds include at least one depsipeptide linkage in their peptide backbone, reflecting the prominence of this linkage among cyclic peptide natural products and the largely untapped potential of depsipeptides as cell permeable scaffolds.[58] In addition to masking polar groups, the pattern of N-methylation and transannular hydrogen bonding have been shown to effect cyclic peptide conformations and thereby alter permeability characteristics.[52,54-60] The former appears to alter the later, as shown in controlled experiments with polyalanine hexa- and hepta-peptides.[57,60] We

introduce these two structural features here, and will revisit them in greater detail in **Chapter 2**.

Multiple long-range effects of N-methylation on cyclic peptides

Mono-N-methylation has been used for years to change pharmacological properties of peptides.[61] Recent advances in peptide chemistry have now made multiple N-methylation much more synthetically accessible.[62,63] This has led to the systematic investigation of the effect of N-methylation on the conformational behavior, pharmacokinetic properties, and target selectivity of several model peptide systems.

To demonstrate the utility of N-methylation to highly active, but poorly bioavailable peptides, Horst Kessler and colleagues performed an N-methyl scan of the cyclopeptidic somatostatin analogue cyclo(-PFWKTF-), known as the Veber-Hirschmann peptide.[64,65] A library of 30 N-methylated analogues were synthesized and detailed pharmacokinetic experiments identified 8 peptides with improved GI permeability, as determined by transport across a Caco-2 monolayer. Interestingly, the most bioactive analog, which was mono-N-methylated, was also one of the least permeable of the 8 tested. The authors' conformational analysis showed that, in the most active analogue, externally oriented amide hydrogens of D-Trp8 and Phe11 are involved in a γ-turn with nearby backbone carbonyls, resulting in the solvent shielding of all amide NHs, making the molecule more lipophilic. This "bent" conformation is also observed in cyclo(-PFwKTF-). N-methylation of any other site results in partial to complete loss of this "bent" conformation. Thus, the highest affinity is observed

in the mono-methylated analogue with a gradual decrease in activity with increasing N-methylations, despite increased permeability. In a separate study, Kessler's group found that multiple N-methylation of a cyclic hexapeptide integrin antagonist of GPIIb-IIIa (αIIbβ3 integrin), cyclo(-GRGDfL-), imparts selectivity of the peptide toward different integrin subtypes.[66]

Taken together, these examples suggest a very important point when considering the bioactivity of N-methylated peptides: N-methylation can alter both target engagement and permeability, and incentives are not always aligned.

Much of the interest in N-methylation originated from CsA. Recall that this 11-amino acid peptide has a remarkable 7 N-methylated peptide bonds. It is widely held that N-methylation contributes to its unexpectedly high oral bioavailability (28% ± 18%), resulting in its success as an immunosuppressive drug. However, surprisingly, the mechanistic effect of N-methylation in this prototypical molecule has yet to be fully elaborated. From studies in simpler systems, such as those discussed above, we know that N-methylation can have multiple and at times competing effects, including but not limited to: (1) Masking of polar HBDs, thereby increasing lipohilicity and decreasing desolvation energy; (2) Preventing transannular hydrogen bonds, thereby decreasing lipophilicity due to solvent-exposed backbone carbonyls or other HBAs; (3) Altering the pattern of transannular hydrogen bonds resulting in conformational changes; (4) Conformational changes due to replacement of hydrogen with a more sterically-bulky methyl group; and (5) Increased population of *cis*

peptide bonds resulting in changes to the peptide backbone conformation. Predicting the outcome of N-methylation on bioactivity, therefore, remains challenging.

The interplay of N-methylation and intramolecular hydrogen bonds

Crystal structures of CsA and other N-methylated cyclic peptides (ABA, destruxins) suggest that N-methylation may serve to increase membrane permeability by allowing, or possibly stabilizing, intramolecular hydrogen bonding in the membrane-associated state. [67.68] Based on this conformational hypothesis, Scott Lokey's group investigated whether intramolecular hydrogen bonding could be used to direct the N-methylation of a cyclic peptide scaffold on solid phase, and whether selective N-methylation could be used to improve the passive membrane permeability of the parent scaffolds.[63] They found that the same conformational determinants that influence permeability in non-methylated cyclic peptides could be employed to impart selectivity by directing N-methylation to the most exposed NH groups. Moreover, their results indicate that, for the series of compounds tested, partially N-methylated cyclic peptides are often more permeable than the corresponding permethylated species. Consistent with detailed computational predictions, the permeabilities of cyclic peptides with the same number of N-methyl groups varied widely.

Revisiting the distant past to provide insights for the future

Considering the discussions above, it is clear that peptide macrocycles, whether synthetic or naturally-derived, are an attractive class of molecules that may allow access to

underrepresented drug target classes, such as PPIs. Recent studies have provided structural and mechanistic insight into the conformationally-dependent membrane permeability and target engagement of small cyclic peptides. In the following pages we explore these findings as they relate to the best-known orally bioavailable cyclic peptide, CsA. We hope to gain insight into the generalizability of these principles, by exploring their application to a well-studied, large, macrocyclic natural product.

To this end, we have developed a tool to study the bioactivity of CsA analogues (CsAlogs), which can be used to determine the effect of structural modifications on cell permeability and target-binding. We propose that this bioactivity assay may also be useful for the evaluation of CsAlogs with altered target specificity, and to find novel cyclophilin-binding scaffolds. Based on the results of this assay, we selected several CsAlogs to perform additional structure-activity relationship studies. We focused on modifications known to alter cyclic peptide conformational dynamics, with the goal of elucidating their effect on permeability and target specificity. Finally, we describe preliminary work on the engineering of non-CsA, cyclophilin-binding peptide scaffolds with the potential for altered target specificity.

References

1. Scanell, J.W., Blanckley, A., Boldon, H., & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery.* **11**, 191-200 (2012).

2. Wells, J.A., & McClendon, C.L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature.* **450**, 1001-1009 (2007).

3. Overington, J.P., Al-Lazikani, B., & Hopkins, A.L. How many drug targets are there? *Nature Reviews Drug Discovery.* **5**, 993-996 (2006).

4. Koonin, E.V., Wolf, Y.I. & Karev, G.P. The structure of the protein universe and genome evolution. *Nature.* **420**, 218-223 (2002).

5. Vitkup, D., Melamud, E., Moult, J., & Sander, C. Completeness in structural genomics. *Nature Structural Biology.* **8**, 559-566 (2001).

6. Strong, M. & Eisenberg, D. The protein network as a tool for finding novel drug targets. *Prog. Drug Res*. **64**, 191-215 (2007).

7. Krogan, N.J. *et al*. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637-643 (2006).

8. Rolland, T., Rother, F.R., Vidal, M. *et al.* A proteome-scale map of the human interactome network. *Cell*. **159**, 1212-1226 (2014).

9. Hopkins, A.L. & Groom, C.R. The druggable genome. *Nature Revues Drug Discovery.* **1**, 727-730 (2002).

10. Arkin, M.R., Tang, Y., & Wells, J.A. Small-molecule inhibitors of protein-protein interactions: Progressing toward the reality. *Chemistry & Biology*. **21**, 1102-1114 (2014).

11. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177-2198 (1999).

12. Jones, S. & Thornton, J.M. Principles of protein-protein interactions. *PNAS.* **93**, 13-20 (1996).

13. Cheng, A.C. *et al.* Structure-based maximal affinity model predicts small-molecule druggability. *Nature Biotechnology*. **25**, 71-75 (2007).

14. Smith, R.D. *et al.* Exploring protein-ligand recognition with Binding MOAD. *J. Mol. Graph. Model.* **24**, 414-425 (2006).

15. Moiera, I.S. *et al.* Hot spots – A review of the protein-protein interface determinant amino acid residues. *Proteins.* **68**, 803-812 (2007).

16. London, N., Raveh, B., Schueler-Furman, O., Druggable protein-protein interactions – from hot spots to hot segments. *Current Opinion in Chemical Biology.* **17**, 952-959 (2013).

17. Eyrsich, S. & Helms, V. Transient pockets on protein surfaces involved in protein-protein interactions. *J. Med. Chem.* **50**, 3457-3464 (2007).

18. Arkin, M.R. & Wells, J.A., Small molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Reviews Drug Discovery*. **3**, 301-317 (2004).

19. Basse, M.J., Betzi, S., Bourgeas, R., Bouzidi, S., Chetrit, B., Hamon, V., Morelli, X., & Roche, P. 2P2ldb: A structural database dedicated to orthosteric modulation of protein-protein interactions. *Nucleic Acids Research.* **41**, D824-D827 (2013)

20. Higueruelo, A.P., Schreyer, A., Bickerton, G.R., Pitt, W.R., Groom, C.R. & Blundell, T.L. Atomic interactions and profile of small molecules disrupting protein-protein interfaces: the TIMBAL database. *Chemical & Biological Drug Design.* **74**, 457-467 (2009).

21. Erlanson, D.A., Wells, J.A., & Braistead, A.C. Tethering: fragment-based drug discovery. *Ann. Rev. Biophys. Biomol. Struct.* **33**, 199-223 (2004).

22. Hadjuk, P.J. and Greer, J. A decade of fragment-based drug design: Strategic advances and lessons learned. *Nature Reviews Drug Discovery.* **6**, 211-219 (2007).

23. Braistead, A.C. *et al*. Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *JACS*. **125**, 3714-3715 (2003).

24. Raimundo, B.C. *et al*. Integrating fragment assembly and biophysical methods in the advancement of small-molecule antagonists of IL-2: An approach for inhibiting protein-protein interactions. *J. Med. Chem.* **47**, 3111-3130 (2004).

25. Arkin, M.R & Wells, J.A. Small molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Rev. Drug Discov.* **3**, 301-317 (2004).

26. Levine, A.J., Hu W. & Feng Z. The p53 pathway: What questions to be explored? *Cell Death Differ.* **13**, 1027-1036 (2006).

27. Vassilev, L.T. *et al. In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844-848 (2004).

28. Grasberger, B.L. *et al.* Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. *J. Med. Chem.* **48**, 909-912 (2005).

29. Yoakim, C. *et al.* Discovery of the first series of inhibitors of human pappilomavirus type 11: Inhibition of the assembly of the E1-E2 Origin DNA complex. *Bioorg. Med. Chem. Lett.* **13**, 2539-2541 (2003).

30. He, M.M. *et al*. Small-molecule inhibition of TNF-alpha. *Science* **310**, 1022-1025 (2005).

31. Robin, W.S. High-throughput screening of historic collections: Observations of file size, biological targets, and file diversity. *Biotechnol. Bioeng.* **61**, 61-67 (1998).

32. Doak, C.D., Over, B., Giordanetto, F. & Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chemistry and Biology.* **21**, 1115-1142 (2014).

33. Artursson, P., Palm, K., & Luthman, K. Caco-2 monolayers in experimental and theoretical predictions of drug transport. *Advances Drug Delivery Reviews.* **64**, 280-289 (2012).

34. Kansy, M., Senner, F., & Gubernator, K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of description of passive absorption processes. *Journal of Medicinal Chemistry.* **41**, 1007-1010 (1998).

35. Guimaraes, C.R.W., Mathiowetz, A.M., Shalaeva, M., Goetz, G., & Liras, S. Use of 3D properties to characterize beyond rule-of-5 property space for passive permeation. *J. Chem. Inf. Model.* **52**, 882-890 (2012).

36. Smith, D.A. Metabolism, pharmacokinetics and toxicity of functional groups: Impact of the building blocks of medicinal chemistry in ADMET. Cambridge, UK: Royal Society Publishing. (2010)

37. Yang, Y., Engkvist, O., Llinas, A., & Chen, H. Beyond size, ionization state, and lipophilicity: Influence of molecular topology on absorption, distribution, metabolism, excretion, and toxicity for druglike compounds. *Journal of Medicinal Chemistry*. **55**, 3667-3677 (2012).

38. Lipinski, C.A., Lombardo, F., Dominy, B.W., & Feeney, P.J., Experimental and computation approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Development Reviews*. **23**, 3-25 (1997).

39. Lipinski, C.A., Lombardo, F., Dominy, B.W., & Feeney, P.J., Experimental and computation approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Development Reviews*. **46**, 3-26 (2001).

40. Morphy, R. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. *Journal of Medicinal Chemistry*.**49**, 2969-2978 (2006).

41. Paolini, G.V., Shapland, R.H., van Hoorn, W.P., Mason, J.S., & Hopkins, A.L. Global mapping of pharmacological space. *Nature Biotechnology*. **24**, 805-815 (2006).

42. Lipinski, C.A. Lead and drug-like compounds: The rule of 5 revolution. *Drug Discovery Today Technologies*. **1**, 337-341 (2004).

43. Abad-Zapatero, C. A sorcerer's apprentice and The Rule of Fve: from rule-of-thumb to commandment and beyond. *Drug Discovery Today Technologies*. **12**, 995-997 (2007).

44. Schreiber, S.L., Crabtree, G.R. The mechanism of action of cyclosporine A and FK506. *Immunology Today.* **13**, 136-42 (1992).

45. Hamman, J.H., Enslin, G.M. & Kotze, A.F. Oral delivery of peptide drugs: Barriers and developments. *Biodrugs* **19**, 512-523 (1982).

46. Svarstad, H., Bugge, H.C, & Dhillion, S.S. From Norway to Novartis: Cyclosporin from Tolypocladium inflatum in an open access bioprospecting regime. *Biodiversity and Conservation.* 9, 1521–1541 (2011).

47. Borel, J.F. Comparative study of in vitro and in vivo drug effects on cell mediated cytotoxicity. *Immunological Communications.* **31,** 631–641 (1976).

48. Loosli, H. R., Kessler, H., Oschkinat, H., Weber, H. P., Petcher, T.J., Widmer, A. The conformation of cyclosporin A in the crystal and in solution. *Helv. Chim. Acta,* ***68***, 682-704 (1985).

49. Kessler, H., Kock, M., Wein, T., Gehrke, M. Reinvestigation of the Conformation of Cyclosporin A in Chloroform. *Helv. Chim. Acta,* **73**, 1818-1832 (1990).

50. Jin, L., Harrison, S. C., Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. *Proceedings of the National Academy of Sciences.* **99,** 13522-13526 (2002).

51. Ko, S. Y., Dalvit, C. Conformation of cyclosporin A in polar solvents. *Int. J. Pept. Protein Res.,* **40**, 380-2 (1992).

52. Beck, J.G., Chatterjee, J., Laufer, B., Kiran, M. U., Frank, A. O., Neubauer, S., Ovadia, O., Greenberg, S., Gilon, C., Hoffman, A., Kessler, H. Intestinal permeability of cyclic peptides: common key backbone motifs identified. *JACS.* **134**, 12125−12133 (2012).

53. Thansandote, P., Harris, R. M., Dexter, H. L., Simpson, G. L., Pal, S., Upton, R. J., Valko, K. Improving the passive permeability of macrocyclic peptides: balancing permeability with other physicochemical properties. *Bioorg. Med. Chem*. **23**, 322−327 (2015).

54. Ovadia, O., Greenberg, S., Chatterjee, J., Laufer, B., Opperer, F., Kessler, H., Gilon, C., Hoffman, A. The effect of multiple N-methylation intestinal permeability of cyclic hexapeptides. *Molecular Pharmaceutics.***8**, 479−487 (2011).

55. Bock, J. E., Gavenonis, J., Kritzer, J. A. Getting in shape: controlling peptide bioactivity and bioavailability using conformational constraints. *ACS Chemical Biology*. **8**, 488−499 (2013).

56. Alex, A., Millan, D. S., Perez, M., Wakenhut, F., Whitlock, G. A. Intramolecular hydrogen bonding to improve membrane permeability and absorption in beyond rule of five chemical space. *MedChemComm*. **2**, 669−674 (2011).

57. Chatterjee, J., Mierke, D. F., Kessler, H. Conformational preference and potential templates of N-methylated cyclic pentaalanine peptides. *Chemistry*. **14**, 1508−1517 (2008).

58. Bockus, A.T., McEwen, C.M., Lokey, S.R. Form and function in cyclic peptide natural products: A pharmacokinetic perspective. *Current Topic in Medicinal Chemistry*. **13**, 821-836 (2013).

59. Ahlbach, C.L., Katrina, W.L., Andrew, T.B., Chen, V., Crews, P., Jacobsen, M.P., Lokey, R.S. Beyond cyclosporine A: conformation-dependent passive membrane permeabilities of cyclic peptide natural products. *Future Medicinal Chemistry* (Epub ahead of print).

60. Rezai, T., Bock, J.E., Zhou, M.V., Kalyanaraman, C., Lokey, R.S., Jacobseon, M.P. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: Successful in silico prediction of the relative permeabilities of cyclic peptides. *JACS*. **128**, 14073-14080 (2006).

61. Kessler, H. Detection of intramolecular mobility by NMR spectroscopy and detection of hindered rotation and inversion by NMR spectroscopy. *Angew. Chem. Int. Ed.* **9**, 219–235 (1970).

62. Chatterjee, J., Laufer, B., Kessler, H. Synthesis of N-methylated cyclic peptides. *Nature Protocols*. **7**, 432-444 (2012).

63. White, T.R., Renzelman, C.M., Rand, A.C., Rezai, T., McEwen, C.M., Gelev, V.M., Turner, R.A., Linington, R.G., Leung, S.S., Kalgutkar, A.S., Bauman, J.N., Zhang, Y., Liras, S., Price, D.A., Mathiowetz, A.M., Jacobson, M.P., Lokey, R.S. On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nature Chemical Biology.* **25**, 810-817 (2011).

64. Chatterjee, J., Gilon, C., Hoffman, A. & Kessler, H. N-methylation of peptides: a new perspective in medicinal chemistry. *Acc. Chem. Res.* **41**, 1331–1342 (2008).

65. Ovadia, O., Greenberg, S., Laufer, B., Gilon, C., Hoffman, A., Kessler, H. Improvement of drug-like properties of peptides: the somatostatin paradigm. *Expert Opin. Drug Discov.* **5**, 655–671 (2010).

66. Mas-Moruno, C., Beck, J.G., Doedens, L., Frank, A.O., Marinelli, L., Cosconati, S., Novellino, E., Kessler, H. Increasing αvβ3 selectivity of the anti-angiogenic drug cilengitide by N-methylation. *Angew. Chem. Int. Ed*. **50**, 9496 –9500 (2011).

67. In, Y., Ishida, T., Takesako, K. Unique molecular conformation of aureobasidin A, a highly amide N-methylated cyclic depsipeptide with potent antifungal activity: X-ray

crystal structure and molecular modeling studies. *Journal of Peptide Research*. **53**:492–500 (1999).

68. Steiner, J.R., Barnes, C.L. Crystal and Molecular-Structure of Destruxin-B. *International Journal of Peptide and Protein Research*. **31**, 212–219 (1988).

CHAPTER 1:

ENGINEERING A SELECTIVELY DESTABILIZED MUTANT OF CYCLOPHILIN A

Inspired by nature: Affinity enhancement by surface-borrowing immunophilin ligands

When considering challenging questions in chemical biology, it is this author's belief that insight can be gleaned from evolutionarily-tailored biological mechanisms. The work described herein was inspired by a particularly elegant example of small molecule inhibition of protein surfaces. Selected microorganisms use endogenous proteins to present and enhance the affinity of their toxins. Immunophilin ligands are macrocyclic natural products that bind to cytosolic peptidyl-prolyl *cis/trans*-isomerases, thus enabling the formation of an inhibitory trimeric complex with an endogenous effector protein **(Fig. 1-1)**[1,2]. These molecules have the unusual ability to induce protein dimerization and, in doing so, enhance their affinity for inhibition of the eventual protein target. Display of the ligand by the 'presenting protein' forms a new composite surface and greatly increases the surface area available for interactions with the target. As such, these ligands do not bind to their target in isolation, but rely on the increased affinity that results from surface-borrowing.

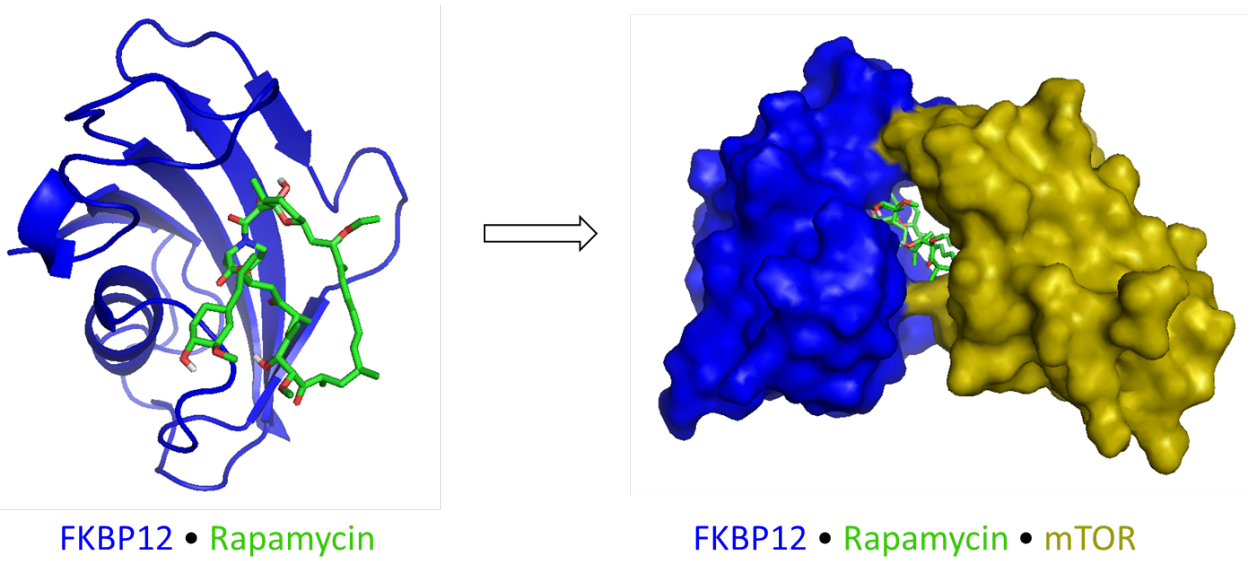


Figure 1-1. Surface-borrowing affinity enhancement enables formation of inhibitory trimeric complex. Rapamycin binds presenting protein FKBP12; the resulting dimeric species binds to and inhibits mTOR. [PDB ID: 1FAP]

Rapamycin, FK506, and cyclosporine A (CsA) are immunosuppressive natural products that have been used to prevent rejection following organ transplantation **(Fig. 1-2)**. FK506 and rapamycin each bind to FK506-binding protein 12 (FKBP12) through a conserved binding domain. A variable effector domain determines the target for inhibition, either calcineurin for the FK506-FKBP12 complex, or mammalian target of rapamycin (mTOR), which is inhibited by the rapamycin-FKBP12 complex[3]. Interestingly, despite a markedly different structural scaffold, CsA also targets calcineurin following complexation with the immunophilin protein, cyclophilin A (CypA). A fourth immunophilin ligand with immunosuppressive activity, sanglifehrin A, has been described[4]. While the target of sanglifehrin A that mediates its biological effect is not yet known, it is thought to act by a similar mechanism due to its high affinity for CypA.

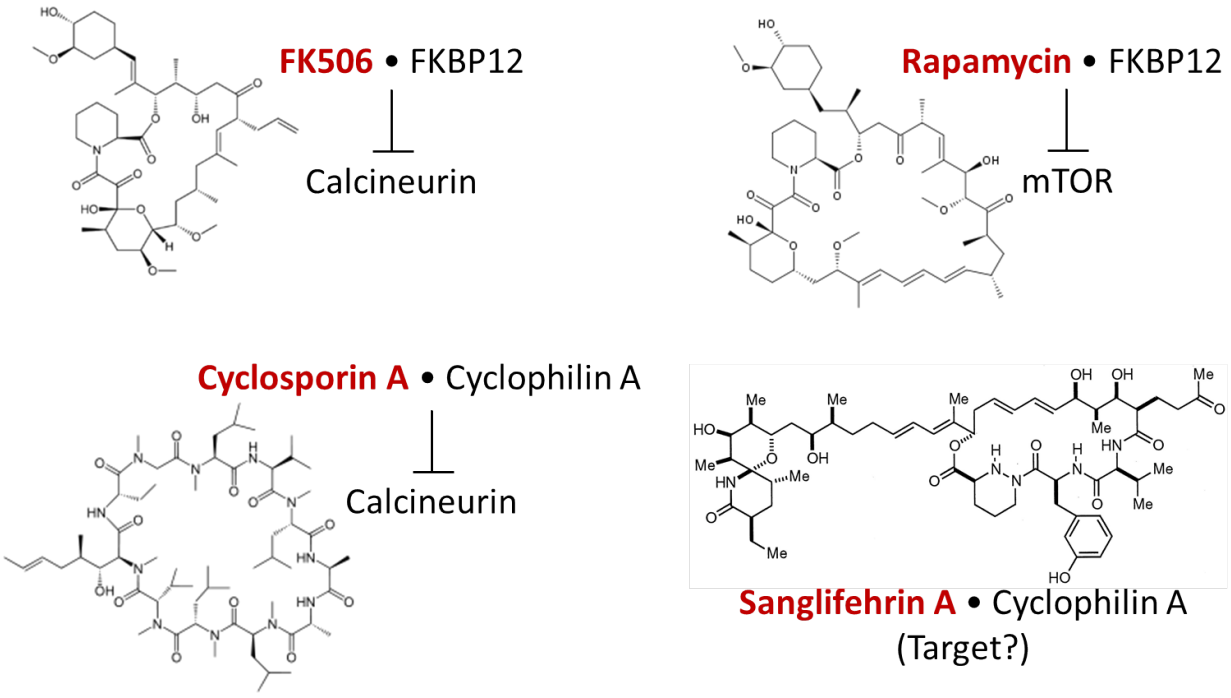


Figure 1-2. Microbial natural products bind immunophilin proteins to enable target inhibition. Four known immunosuppresants (red) act by recruiting an immunophilin protein (CypA or FKBP12) to increase the affinity toward their target. The mechanism of immunosuppressive action of sanghliferin continues to be unraveled.[4]

As described in the introductory chapter, CsA is a structurally-significant molecule thanks to its unexpected permeability characteristics.[5-7] In the following chapter, we present the engineering and validation of a new tool for the study CypA ligands. This tool was used to evaluate the bioactivity of CsA anlaogues (CsAlogs) as a preliminary hypothesis generating read-out. Subsequent structure-activity studies are then presented in **Chapter 2**. We also propose that this tool may be used to identify or evaluate novel CypA-binding scaffolds with altered target specificity. Work in pursuit of this goal is presented in **Chapter 3**.

Destabilizing domains allow conditional small-molecule control of protein-stability

Rapid and reversible methods for perturbing the function of specific proteins are desirable tools for probing complex biological systems. A generalizable technique for conditional perturbation of any protein of interest has been described by Tom Wandless and colleagues.[8] They developed a method to regulate the stability of proteins in mammalian cells using cell-permeable, synthetic molecules. Destabilizing domains (DDs) are proteins or protein domains that have been engineered to contain one or few mutations that confer instability and promote proteosomal degradation. Several DDs have been reported including immunophilin protein, FKBP12; the estrogen receptor (ER), and the *E.coli* enzyme, dihydrofolate reductase (DHFR).[8-10] While these proteins are rapidly degraded in the absence of a stabilizing ligand, when bound by a high affinity small molecule, the protein half-life returns to near wild-type levels. Because these domains have the ability to confer instability onto protein fusions, they have been developed as a general tool to enable temporal control of protein degradation by small molecule ligands **(Figure 1-3A)**. As such, the existing destabilizing domains also contain active site mutations, enabling the use of orthogonal ligands with no biological activity via a bump-hole strategy.

We noted, however, that FKBP12 or CypA destabilizing domains with wild-type, rather than orthogonal, binding characteristics may be useful tools for the study of immunophilin ligands. By fusion to a fluorescent reporter protein, we can assay the permeability and binding characteristics of structurally-related ligand analogs, or screen for new immunophilin-binding scaffolds. This approach is a new use of DD technology, since the persistence of the fusion protein is only a surrogate for ligand stabilization. It is the DD-ligand interaction that is now under scrutiny **(Figure 1-3B)**.

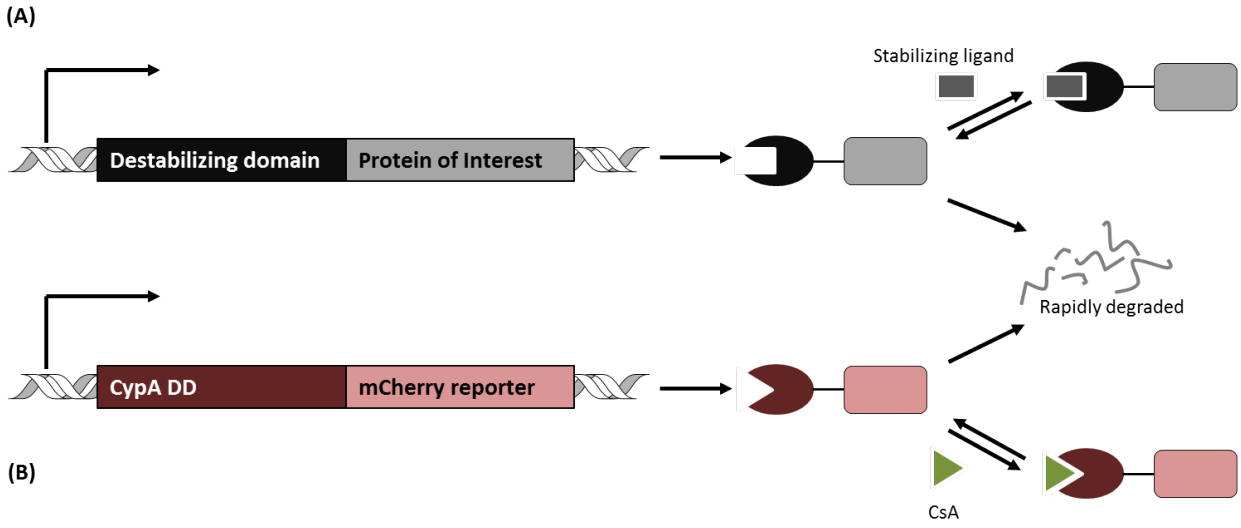


Figure 1-3. Destabilizing domains as a general method to conditionally control protein stability. (A) Genetic fusion of a destabilizing domain (DD) to a protein of interest results in degradation of the entire fusion. Addition of a DD ligand stabilizes the fusion and allows both proteins to persist in the cell. (B) Fusion of a CypA DD to an mCherry reporter gene. Persistence of the resulting fusion protein is controlled by CsA, or other CypA ligands.

We chose to develop a CypA destabilizing domain (cDD) for two primary reasons. Firstly, as noted earlier, CypA ligand, CsA, is a peptidic macrocycle whose structure-permeability relationship has generated intense interest. Secondly, through a collaboration with Warp Drive Bio, LLC., we were given access to a large library of CsAlogs. This facilitated a rapid path from cDD development to gaining additional insights about the structural determinants CsA bioactivity.

Engineering a destabilizing domain of CypA

To identify mutants that display the desired ligand-dependent behavior, we used a cell-based screen, in which a fluorescent reporter protein, mCherry, was used as an indicator of CypA stability. This allowed us to enrich, isolate, and characterize CypA mutants with the desired ligand-dependent stability. An HTS-amenable cell-based assay to identify CypA ligands was developed with the most differentially destabilized mutants. The general workflow for selection of DDs is adapted from Wandless *et al.*, 2006, and is given below (**Fig. 1-4)**.[8]

Figure 1-4. Generalizable workflow for engineering a protein destabilizing domain.

A library based on the wild-type (WT) sequence of CypA, *PPIA gene*, was generated using error-prone PCR. CypA was cloned in-frame in front of the gene-encoding mCherry. The genes were separated by a short, genetically-encoded Glycine-Serine linker, $(GGGGS)_{2X}$. The mutant library was cloned into a puromycin-resistance retroviral vector, under the control of a CMV promoter, which is used to drive high levels of gene expression in mammalian expression vectors.[11] Sanger sequencing was used to estimate the library size and diversity. We proceeded with an estimated library of $3x10^6$ CypA mutants, averaging 1-3 point mutations each.

A murine stem cell retroviral system was used to stably integrate the library of CypA-mCherry fusions into NIH3T3 fibroblasts (3T3) with 10-fold library coverage. Transduced cells were subjected to five rounds of sorting using flow cytometry. Titration of the parental 3T3 cell line with CsA and readout by CellTox Green, a cytotoxicity assay, showed that CsA was well tolerated up to 5uM. In the first round, cells were treated with 2 μM CsA, for 48 hours prior to fluorescence-activated cell sorting (FACS). Fluorescent cells (approximately top 25% of the population) were collected and further cultured in the absence of ligand for 72 hours, with media being changed every 24 hours to remove CsA and allow for degradation of destabilized mutants. Following this ligand-free incubation, cells with a low basal level of fluorescence (approximately bottom 25% of the population) were collected and cultured. Three additional rounds of sorting iteratively for highly- and lowly-fluorescent populations were performed in the presence of absence of a stabilizing ligand, respectively. As enrichment was observed, the stringency of gating was reduced to collect cells exhibiting only the greatest fluorescent dynamic range. The final population was fully enriched for

protein mutants that directed differential fluorescence in the presence of absence of CsA **(Fig. 1-5)**.

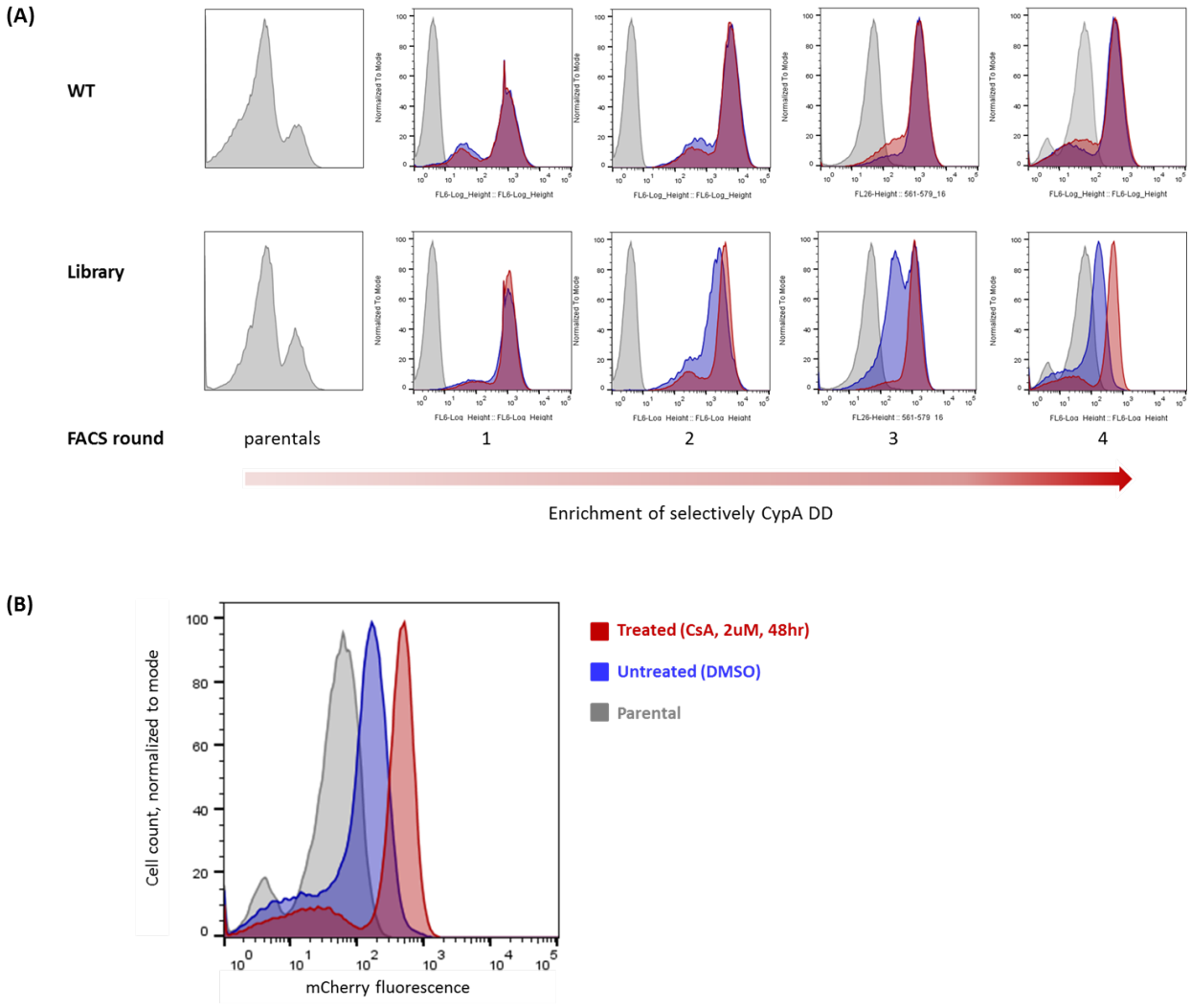


Figure 1-5. FACS enrichment of cyclophilin A destabilizing domains (cDDs). (A) Multiple rounds of sorting, in the presence or absence of CsA, resulted in (B) a population enriched for cDD mutants where CypA-mCherry fusions are selectively stabilized in the presence of CsA.

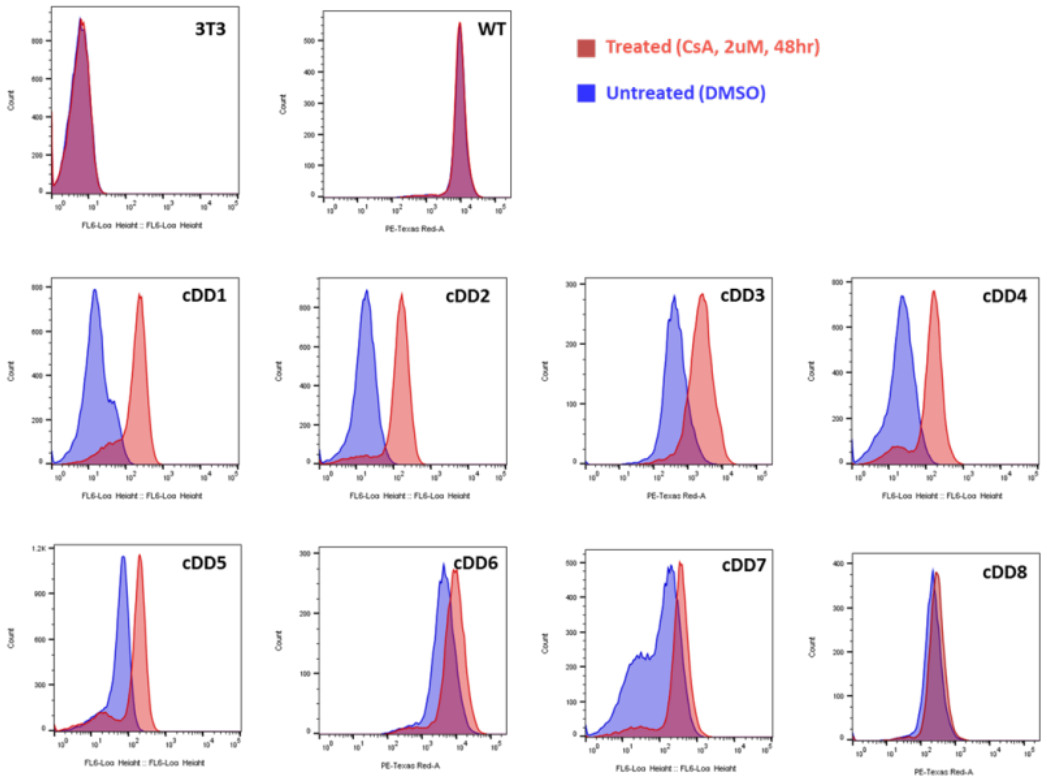
Identification and characterization of CypA destabilizing domains

The resulting population of cells was collected and sequences of the cDDs were identified by one of two methods. Firstly, monoclonal populations were isolated by culturing in limiting dilution conditions. We tested 27 clonal populations, each representing a unique CypA mutant, for destabilizing domain activity. Preliminary characterization was done by reading mCherry fluorescence in a fluorescence plate reader assay. We found this assay to be insufficient for further characterization, as readout of untreated and parental populations produced high levels of background in the emission spectrum of mCherry. The assay was therefore only used as a first-pass to identify 5 mutants which appeared to have high fluorescent signal in the presence of CsA, and a large dynamic range. Subsequent analysis was then performed by FACS, to determine fold-change in fluorescence following addition of the stabilizing ligand. The fold-change and sequences of these cDDs are shown in **Figure 1-6**. In parallel, genomic DNA from the entire resultant population from FACS enrichment was isolated and sent for deep sequencing. The most highly enriched mutations relative to the initial population were V29D and S40N. Therefore, my fellow lab-mate and collaborator, Khian Hong Pua, performed site-directed mutagenesis to generate the single and double mutant versions of CypA. These constructs were transduced into 3T3 cells and added to the panel of mutants to be characterized by subsequent FACS analysis **(Fig. 1-6A)**.

Fluorescence levels were measured in the absence of CsA. All mutants showed decreased fluorescence levels with respect to the WT CypA-mCherry fusion, indicating that the mutants identified from the library screen are destabilizing. The mutants exhibited varying

degrees of destabilization, with the most destabilizing mutant, cDD1 (I10F/G14C), expressing mCherry fluorescence at <5% relative to WT. All mutants showed increased fluorescence upon addition of CsA, with observed efficiencies of rescue varying by over an order of magnitude (**Fig. 1-6B**).

(A)



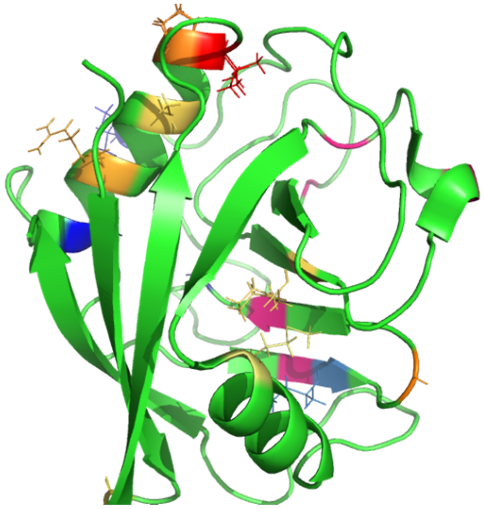
(B)

	cDD mutant	Fold-change over DMSO	Method of identification
3T3 parental cells	absent	1.0 (no change)	
WT	WT	1.0 (no change)	
cDD1	I10F_G14C	15.4x	Clonal isolation and expansion
cDD2	G50C_T116S_M136L	12.1x	Clonal isolation and expansion
cDD3	V29D	8.3x	Population sequencing
cDD4	K118L	8.1x	Clonal isolation and expansion
cDD5	G80N	6.3	Clonal isolation and expansion
cDD6	S40N	2.7x	Population sequencing
cDD7	I78V	2.3x	Clonal isolation and expansion
cDD8	V29D_S40N	1.2x	Engineered based on population sequencing

Figue 1-6. Ligand dependent stability of mCherry-cDD mutant fusion proteins. (A) FACS analysis of clonal cDDs reveals differential mCherry fluorescence after incubation with (red) or without (blue) 3uM CsA for 48 hours. No difference was observed for WT CypA and 3T3 parentals. (B) List of cDD mutant genotypes and method of identification.

The most highly enriched mutations, as determined by Illumina sequencing, were mapped to the protein primary and tertiary sequence **(Fig. 1-7)**. Similar to what was observed by Wandless and colleagues with the FKBP-DD, mutations were distributed fairly evenly over the primary amino acid sequence. They also found that the majority were spatially separated from the ligand binding site, suggesting that ligand binding was an important determinant in selection. By contrast, we found that destabilizing mutations of CypA were more prevalent on solvent-exposed structural elements, located both proximally and distally to the ligand-binding site. Additionally, mutations appeared to cluster on secondary structural elements, such as the series of helix-braking mutations observed on the α-helix formed by residues 29-41. The relative absence of mutations in the protein core suggests that these mutations may be too highly destabilizing to allow for efficient rescue.

(A)



(B)

Mutation	Library	Sort 3	Sort 5	Fold enrichment
V29D	1.351%	1.539%	8.10%	6.00
S40N	1.966%	2.451%	8.06%	4.10
G65C	2.359%	2.575%	7.43%	3.15
I78V	1.969%	2.197%	6.14%	3.12
I56T	0.655%	0.784%	1.68%	2.57
G59E	0.505%	0.633%	1.01%	1.99
P30S	0.911%	1.289%	1.81%	1.99
R37H	0.860%	1.231%	1.70%	1.98
A33T	0.665%	0.827%	1.27%	1.90
G14C	0.501%	0.457%	0.93%	1.87
I114F	0.791%	1.096%	1.47%	1.86
M136T	0.580%	0.613%	0.96%	1.66
G80E	0.701%	0.740%	1.15%	1.65
P30L	0.671%	0.763%	1.04%	1.56

Figue 1-7. Highly enriched mutations in cDD population identified by deep sequencing. Genomic DNA was extracted from the initial library of CypA mutants, after three rounds of FACS sorting, and after the fifth and final round. Illumina sequencing revealed the most highly enriched mutations. (A) Mapping of the 14 most enriched mutations to the CypA tertiary structure. [PDB ID: 2z6w] (B) Fold enrichment indicates the prevalence of a given mutation in the final population (Sort 5) relative to the initial population (Library).

V29D and S40N were the most enriched mutations in the selected population relative to the original library **(Fig 1-7)**. Both sites are located on the solvent-exposed helix (29-41) that proved to be a hotbed for destabilizing mutations. We engineered the V29D/S40N CypA double mutant to assay whether this protein, which we suspected would me more destabilized relative to the single mutants, could be rescued by CsA. We generated a 3T3 line that stably expressed the CypA V29D/S40N-mCherry fusion (cDD8) by retroviral transduction. Following incubation with 3uM CsA, FACS analysis revealed that the basal level of mCherry fluorescence was similar to that of the V29D single mutant (cDD3), and lower than that of the S40N single mutant (cDD6) **(Fig. 1-6)**. However, rescue of mCherry fluorescence was poor (1.2-fold relative to DMSO-treated), as compared to either of the single mutants. We sought to evaluate the protein stability of CypA V29D/S40N to gain greater insight into the relationship between destabilization and the ability to achieve ligand-dependent stability. However, we were unable to express the double mutant. We suspect that it may be too highly destabilized for expression, and therefore, also too unstable for rescue by ligand binding. This suggest that an optimal level of destabilization may be necessary for engineering and regulating DDs. There are, of course, a number of other explanations, such as a high level of proteolytic cleavage that could result in uncoupling mCherry fluorescence from ligand binding. Therefore, evaluation of this hypothesis would require a much more systematic analysis of the relationship between protein stability and the propensity to act as a reversible destabilizing domain.

To further characterize ligand-responsive cDDs, we selected the six mutants with the greatest differential response to CsA stabilization for additional analysis. These included one double mutant (I10F/G14C), one triple mutant (G50C/T116S/M136L), and four single mutants, (V29D, K118L, G80N, and S40N). Each was separately transduced into 3T3 mouse fibroblast cells.

Dose-response curves were generated by single dose ligand titration (**Fig. 1-8A)**. cDD-transduced cells were incubated in the presence of CsA for 72 hours. Cells were then trypsinized and mCherry fluorescence was quantified by FACS. We observed a dose-dependent increase with maximal effect achieved in the 5-10uM range. Concentrations of up to 30uM were tested, however, at >10uM CsA, significant cytotoxicity was observed (data not shown).

In a subsequent kinetic study we found that mCherry fluorescence for all mutants accumulated over a period of four days following a single dose of 3uM CsA (**Fig. 1-8B)**. The persistence of fluorescence following a single dose implies that, upon addition of CsA, cDDs are able to adopt a conformation that approximates the stability of the wild-type protein. Additionally, increases in fluorescence may be a function of the rate of protein synthesis. Time points beyond day 4 could not be reliably measured due to overgrowth of the cultures.

We also assayed the three most responsive destabilizing domains (cDD1, 2 and 3) for reversibility by observing the kinetics of protein degradation. 3T3 cell lines stably expressing cDD-mCherry fusions were first treated for 48 hours with 3uM CsA. We found that, upon

removal of CsA and replacement with CsA-free media, mCherry fluorescence returned to basal levels within 24 hours (**Fig. 1-8C**). The accumulation of fluorescence over time was also observed by fluorescence microscopy. Images were taken over a period of 4 days in parallel to FACS quantification. Representative images of cDD-mCherry or WT-mCherry transduced 3T3 cells, either in the presence or absence of 3uM CsA, are shown in **Figure 1-9.**

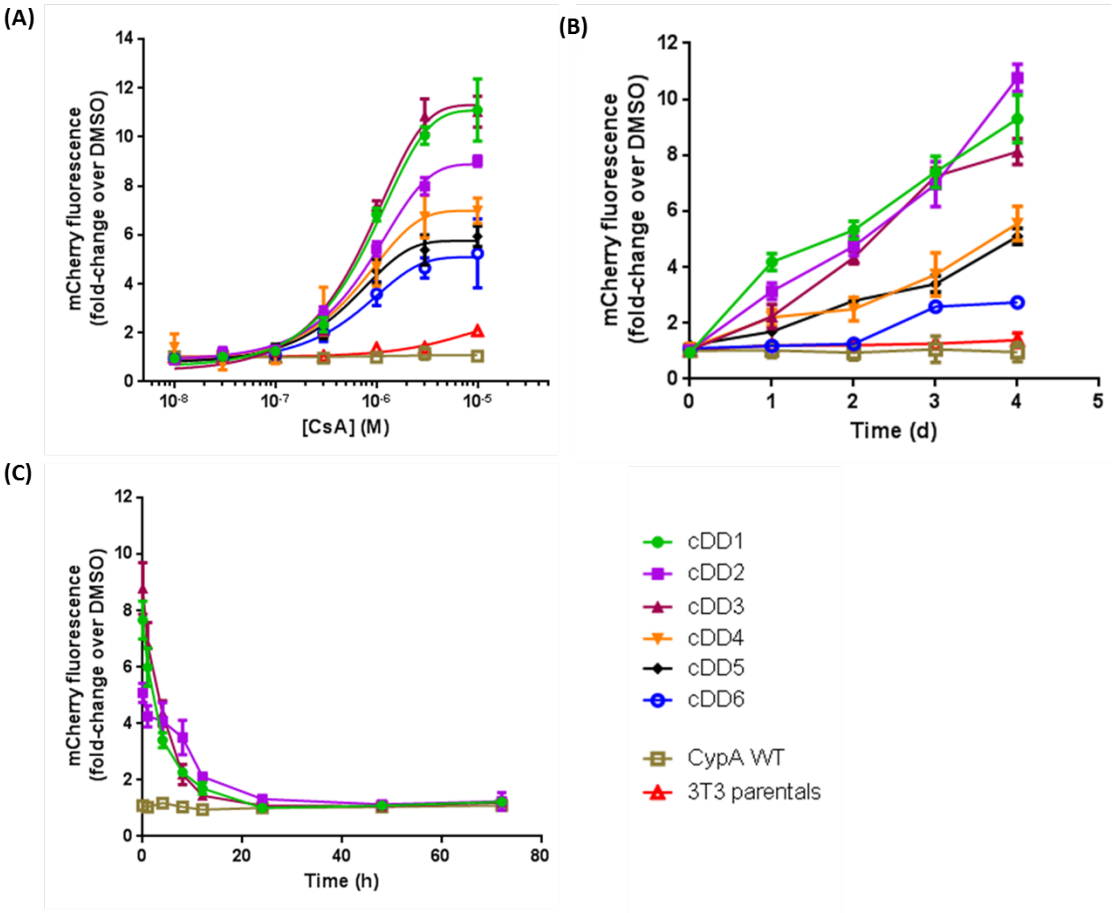


Figure 1-8. Fluorescence-based characterization of CypA mutants that display CsA-dependent stability. All values represent the fold-change increase in fluorescence as compared to mock (DMSO) treated cells of the same mutant fusion. All experiments were performed in triplicate. (A) 3T3 cells stably expressing cDD-mCherry fusions were treated with 3-fold dilutions of CsA (10uM to 10nM) and fluorescence was quantified by flow cytometry. (B) cDD-mCherry stable cell lines were incubated with medium containing 3uM CsA and fluorescence was monitored over time for up to 4 days by both FACS and microscopy. Corresponding images in **Figure 1-9**. (C) cDD-mCherry cell lines were pre-treated with 3uM CsA for 48 hours. At t_0, CsA was removed and replaced with ligand free media. Fluorescence was quantified by FACS over 72 hours.

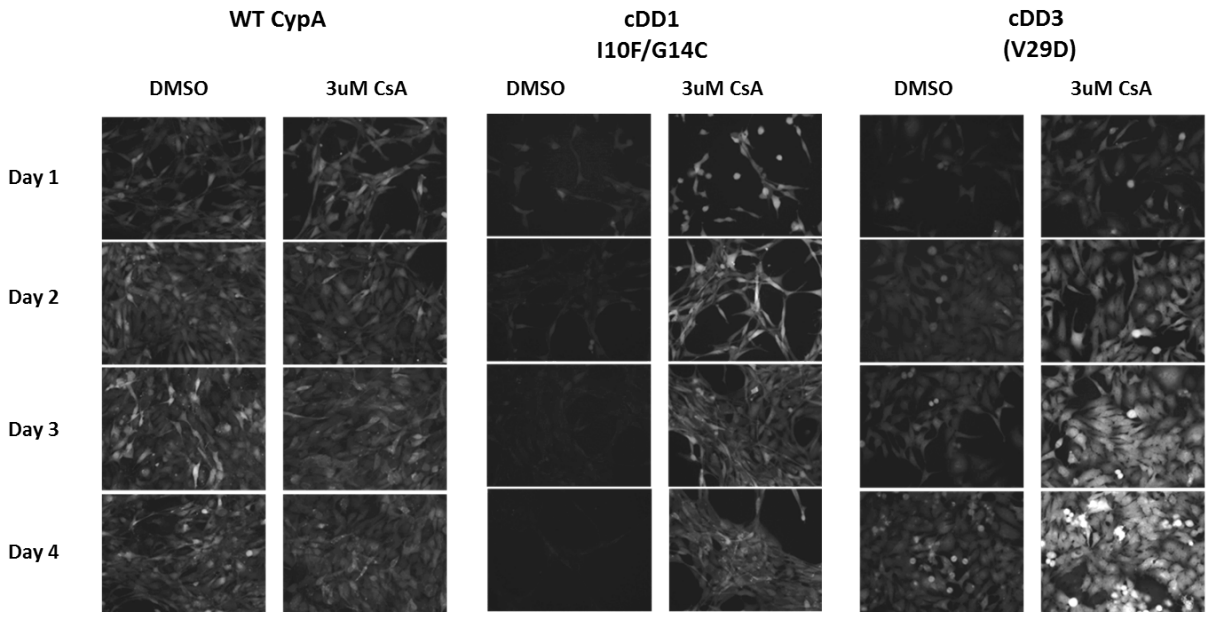


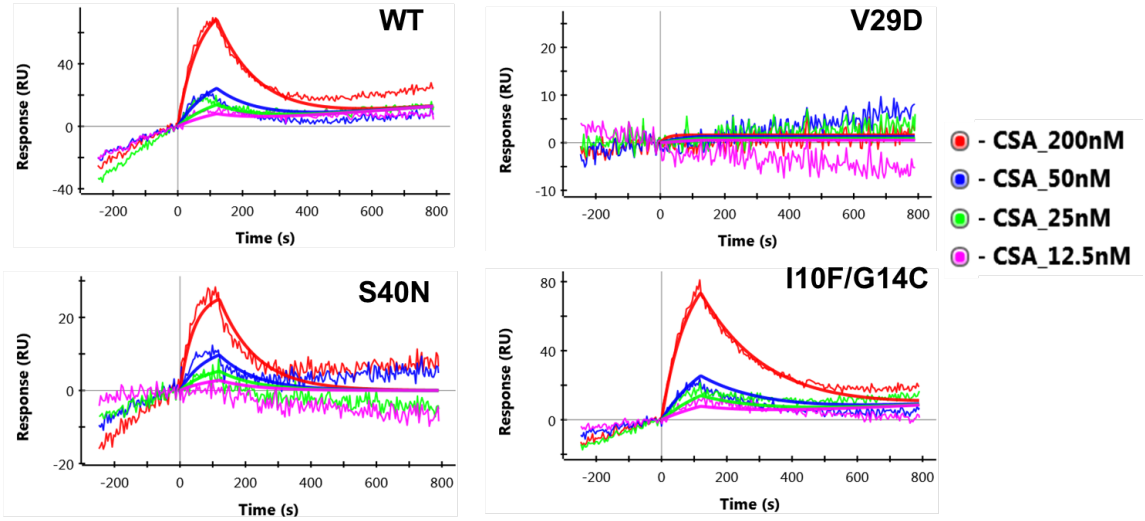
Figure 1-9. Fluorescence microscopy of 3T3 cells stably expressing cDD-mCherry fusions. Cell lines were incubated with medium containing 3uM CsA and fluorescence was monitored over time for up to 4 days by both FACS and microscopy. Corresponding fluorescence quantification is shown in **Figure 1-8C**.

Destabilizing domains were originally devised as a reversible method to conditionally regulate protein stability, thus enabling study of the function of a protein of interest by fusion to the DD. As such, it was desirable to have a protein-ligand pair with orthogonal binding to the corresponding endogenous protein so that the ligand does not elicit undesired responses when administered to cultured cells or animals.

By contrast, our system was designed specifically to study the DD protein-ligand interaction. Therefore, it was necessary for the WT ligand binding characteristics to be maintained. We moved forward only with cDDs, where the destabilizing mutation(s) was/were located far from the CsA-binding face. Mutated *PPIA* genes encoding three CypA destabilizing domains (cDD1, 3, and 6) were subcloned into expression vectors, and the resulting proteins were expressed and purified. We performed surface plasmon resonance (SPR) to determine the affinity of their interactions of with CsA **(Fig, 1-10)**. We found that the I10F/G14C double mutant (cDD1) and the S40N single mutant (cDD6) both retained binding to CypA and the observed K_d for each was similar to that observed with WT CypA.

There are a few necessary notes on the data presented in **Figure 1-10**. The literature reported affinity of the CsA-WT CypA interaction is 30nM. The observed K_d for this interaction in the following experiment, performed on the Bio-Rad ProteON Protein Interaction Array System, at 173nM, was nearly 6-fold weaker. This may be a function of either the system or needed optimization of the protocol. We have, however, since replicated the 30nM interaction using the same protein construct on the more accurate Biacore SPR system (data shown in **Chapter 2**). Therefore, binding data with cDDs is being replicated under those

conditions. Based on the consistency of multiple replicates, we do however feel that while the quantification may off by a significant factor, it is likely that I10F/G14C has retained the ability to bind CsA at or near WT levels. Another discrepancy that remains to be resolved is that CypA V29D (cDD3) appears not to bind CsA, despite observed differential fluorescence of the mCherry fusion during FACS analysis.



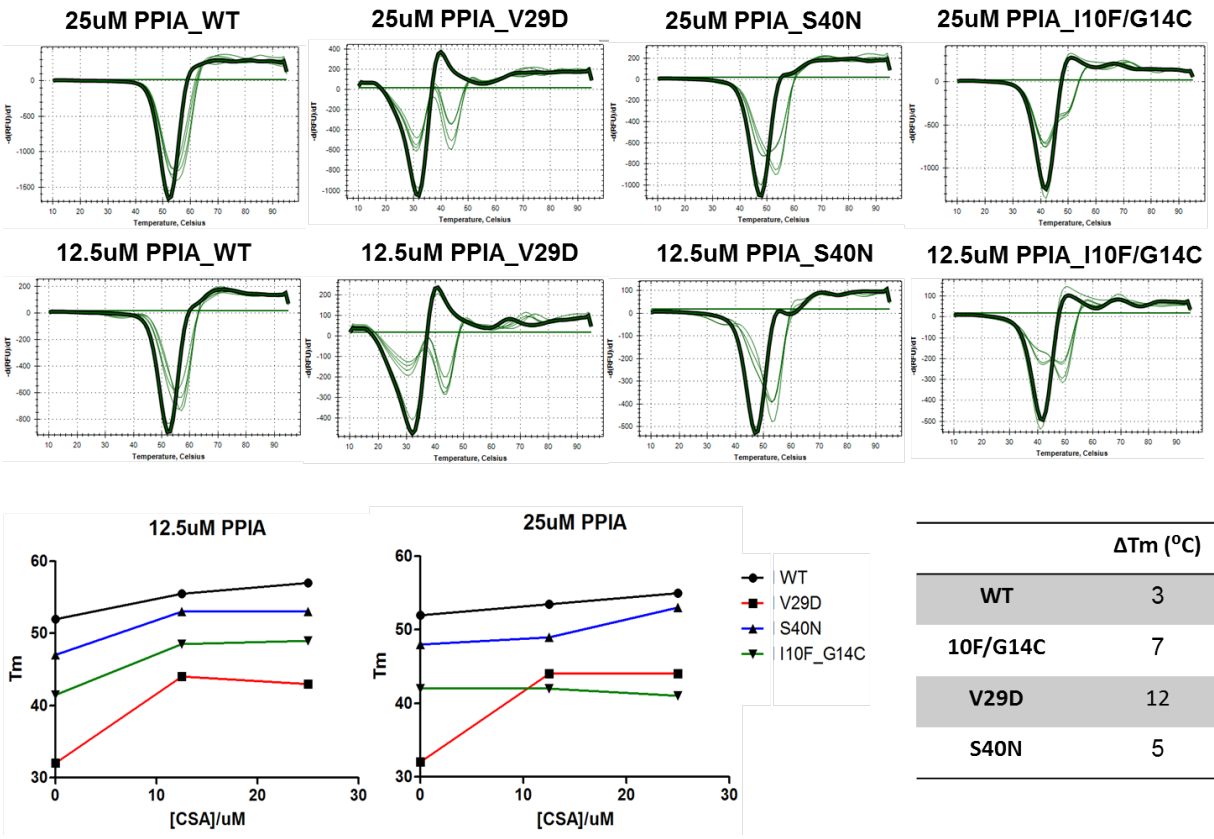
	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_d (nM)
WT	4.40 x 10^4	7.62 x 10^{-3}	173
10F/G14C	7.36 x 10^4	8.87 x 10^{-3}	120
V29D	N.A.	N.A.	N.A.
S40N	4.01 x 10^4	4.79 x 10^{-3}	120

Figure 1-10. Kinetic analysis by surface plasmon resonance implies that cDD mutants I10F/G14C and S40N retain strong interaction with CsA. CypA double mutant, I10F/G14C, and single mutant, S40N, appear to retain affinity for CsA at or near wild-type levels. V29D does appear to bind CsA. (Note: The affinity of WT CypA for CsA reported in the literature is 30nM. Data is being repeated with an optimized protocol on the Biacore system.)

We are also interested, more generally, in understanding the relationship between protein stability and the ability to regulate degradation by a small molecule ligand. Above, we discussed the hypothesis that the V29D/S40N CypA double mutant was destabilized to the point where ligand rescue was no longer possible – either due to rapid degradation, inability to bind to CsA, or inability of CsA binding to stabilize the protein fold. We also performed differential scanning fluorimetry (DSF, thermal shift assay) to determine the melting temperature of three CypA cDD mutants: I10F/G14C (cDD1), V29D (cDD3), and S40N (cDD6) **(Figure 1-11)**.

We observed that the V29D mutation conferred the greatest destabilizing effect, with a melting point nearly 20°C lower than that of WT CypA. A significant rescue of ΔTm 12°C was observed upon the addition of 12.5uM CsA. Perhaps this high degree of destabilization is responsible for the loss of affinity for CsA discussed above. While the observed stabilization upon ligand addition suggests that ligand binding is occurring, the concentration of ligand used is significantly higher than the expected K_d. The ΔTm of the I10F/G14C mutant and the S40N mutant were 7°C and 5°C, respectively, with I10F/G14C having a lower initial Tm, indicating greater destabilization. Unsurprisingly, we observed that ligand binding also conferred a small degree of additional stability to the WT protein (ΔTm 3°C), which, predictably, also has the highest Tm in the absence of ligand. Recall that 3T3 cells stably expressing the I10F/G14C cDD-mCherry fusion had the greatest increase in fluorescence upon addition of ligand as determined by multiple FACS analyses. This suggests that perhaps there is an optimal range of destabilization that allows for constitutive

degradation, while still enabling rescue by ligand binding. If so, this may related to the affinity of the protein-ligand interaction, among other factors.



	ΔTm (°C)
WT	3
10F/G14C	7
V29D	12
S40N	5

Figure 1-11. Differential scanning fluorimetry reveals that cDDs are destabilized relative to WT CypA. Experiment designed and performed by Khian Hong Pua.

We selected the I10F/G14C CypA mutant, hereafter referred to as cDD, to perform all subsequent experiments and for development of a HTS cytosensor assay based on the following factors: (1) Low basal fluorescence of mCherry fusion in the absence of CsA; (2) Maximal response and dynamic range following addition of CsA; (3) The ability to bind to CsA with near-WT affinity; and (4) Moderate level of destabilization and resultant rescue. I10F and G14C are grafted onto the crystal structure of WT CypA to illustrate the location of the mutations on the tertiary protein structure **(Figure 1-12)**. These mutations are located on the turn and a single strand of a 2-stranded N-terminal anti-parallel beta-sheet. Importantly, the mutations are located on the opposite side of the protein relative to the ligand binding face.

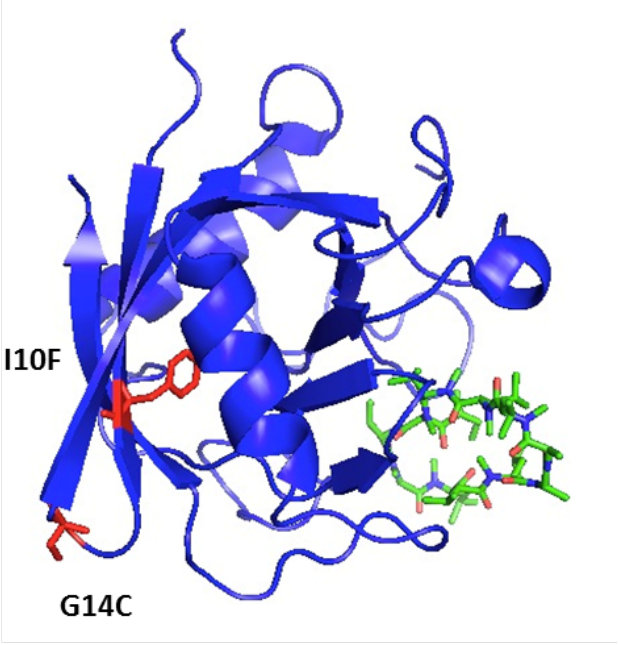


Figure 1-12. Location of I10F/G14C mutations on cDD, cyclophilin A destabilizing domain. Destabilizing mutations I10F and G14C (red) are grafted onto the WT CypA protein crystal structure (blue) bound to CsA (green). [PDB ID: 2z6w]

To gain additional insight into our chosen destabilized domain, we examined the mechanism of degradation for the I10F/G14C cDD mutant. The ubiquitin-proteasome system is a major mediator of intracellular protein degradation.[12] Wandless *et al.* report that degradation of their FKBP12 L106P destabilizing domain is, at least in part, mediated by the 26S proteasome.[8] We treated cells expressing the cDD with MG132, a small peptidomimetic which blocks proteolytic degradation by the 26S proteasome, either in the presence or absence of 3uM CsA. We quantified fusion protein levels by extraction of cell lysate and western blotting for mCherry, since the cells still contain endogenous CypA, and normalized to constitutively expressed chaperone, heat shock protein 90 (Hsp90) **(Fig. 1-13A)**. Expression was also quantified by FACS analysis of mCherry fluorescence **(Fig. 1-13B)**. We expected that if degradation was proteasomal, in the absence of CsA, MG132 would be able to partially or fully restore CypA-mCherry to CsA-treated levels. However, we found no effect of MG132 on cDD-mCherry fusion protein levels, either in the presence or absence of CsA. It is possible that degradation is occurring by a non-proteasomal mechanism. We will continue to investigate the mechanism of degradation by testing additional proteasome inhibitors.

On a separate note, this western blot also serves as direct (ie. not fluorescence-based) validation of the ability of CsA treatment to increase cellular levels of the cDD-mCherry protein fusion.

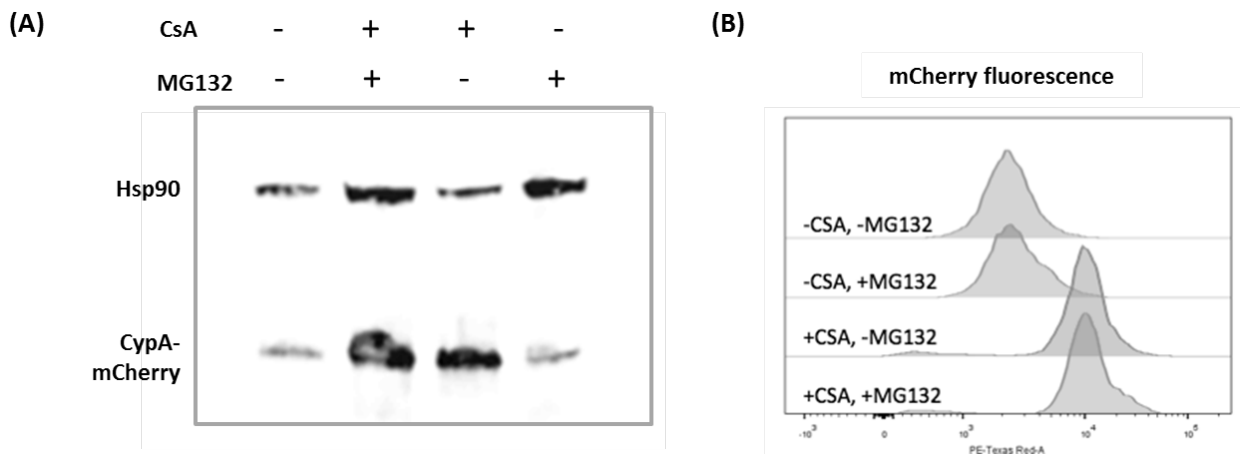


Figure 1-13. Mechanism of degradation of cDD-mCherry remains to be elucidated. Cells stably expressing the cDD-mCherry fusion protein were treated with proteasome inhibitor, MG132, either in the presence or absence of 3uM CsA. (A) Western blot of cellular extracts using anti-mCherry to assess cDD-mCherry fusion protein levels. (B) Fluorescence-based cDD-mCherry quantification by FACS.

Development of a cDD cytosensor assay for structure-function analysis of CsAlogs

Finally, with the fully-characterized cDD in hand, we developed an assay to evaluate the permeability and binding characteristics of structurally-related CsA analogues (CsAlogs). As noted in **Chapter 3**, this tool may also be useful for the identification of new immunophilin-binding scaffolds, either by design or directed evolution.

We seeded either WT- or cDD-mCherry-expressing 3T3 cells into 96-well plates. Cells were treated with 3-fold dilutions of CsA (10uM to 10nM) for 48 hours before being fixed, permeabilized, and stained. CypA-mCherry expression was quantified by in cell western (ICW) blotting using an anti-mCherry antibody conjugated to a fluorescent dye detected in the 800nM channel (green). Protein expression was normalized to cell number using the CellTag700 dye (red). We observed a reproducible dose-dependent effect of CsA on cDD-mCherry fusion protein expression levels, reported as fold-change over DMSO-treated control **(Figure 1-14)**. Note that we also performed similar experiments by blotting for CypA directly, and observed increasing levels of protein with increasing concentration of ligand, thereby confirming that cDD and mCherry levels remain linked (data not shown). However, due to the presence of endogenous CsA, this produced greater background levels, and a smaller dynamic range. A low background was deemed necessary to eventually identify lower affinity interactions; blotting for mCherry was used for all subsequent experiments.

The highest level of cDD stabilization was observed at CsA concentration of ≥3uM. At 10uM, however, cytostatic and cytotoxic effects lead to lower cell counts and a decrease in

well-to-well reproducibility. Therefore, screening experiments of CsAlogs described in **Chapter 2** were performed at a ligand concentration of 3uM.

In parallel, we assayed the effect of CsA on cells expressing the WT-mCherry fusion. Predictably, we found little effect since WT CypA stability and expression are high throughout the range of concentrations. A small increase observed at the highest concentrations of CsA (3uM and 10uM) may reflect the small stabilizing effect (ΔTm) of CsA on WT CypA observed previously by thermal shift assay.

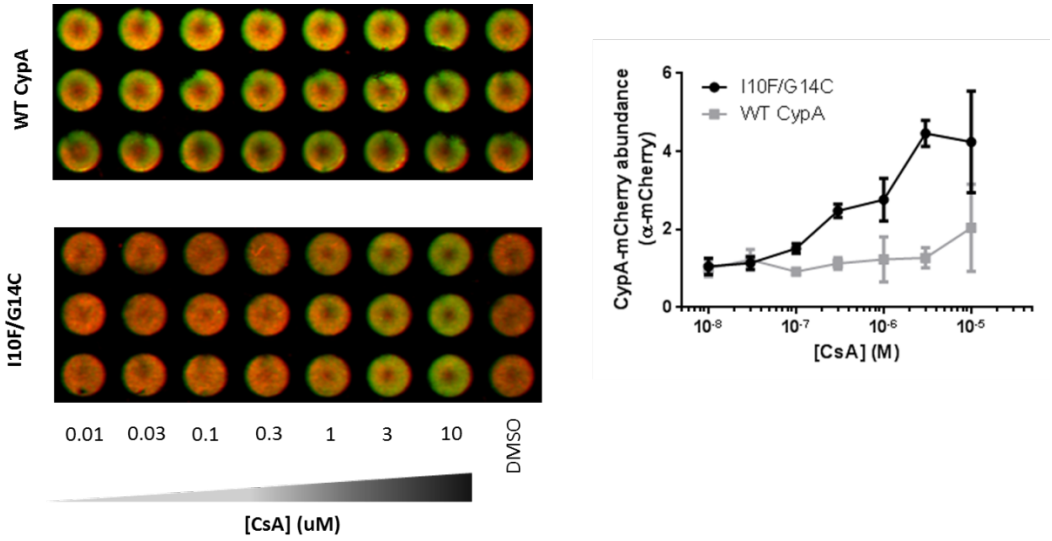


Figure 1-14. Development of cytosensor assay compatible with HTS of CypA ligands. 3T3 cells stably expressing either WT-mCherry or cDD(I10F/G14C)-mCherry fusion proteins were seeded in 96-well plates and treated for 48 hours with increasing concentrations of CsA. Fusion protein levels were determined by ICW against mCherry (green), normalized to cell number (red), and reported as fold-change over DMSO-treated.

In the following chapter, we describe the used of the cDD cytosensor cell line to assay the cell permeability and target binding activity of structurally-related CsA analogues. In the final chapter, we propose how this cytosensor may be used to screen for engineered or genetically-encoded cyclic peptide scaffolds capable of directing CypA/ligand-mediated inhibition to challenging, disease-relevant protein targets.

Acknowledgements

This work is the result an intellectual collaboration between myself, and former Verdine lab member, Khian Hong Pua. Khian also performed several experiments including cloning, expression and purification of the I10F/G14C, V29D and S40N mutants, design of primers for Illumina sequencing analysis, and differential scanning fluorimetry experiments. Initial library construction was performed by Dan Gray at Warp Drive Bio, LLC. We are also grateful to Keith Robison for advice on Illumina sequencing analysis. Finally, we are thankful to Mandy Tam and Patricia Rogers at the Harvard Bauer Core center for help with tuning lasers for FACS experiments, and helpful discussions about experimental design.

Methods

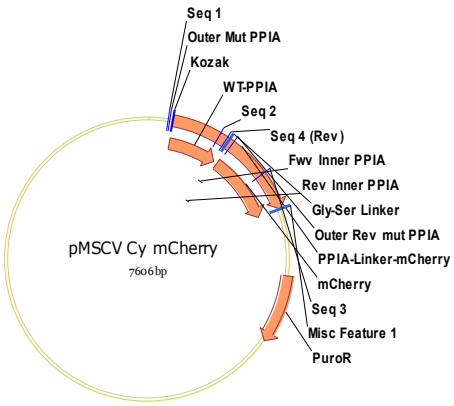
General tissue culture methods

Adherent cell lines, NIH3T3, HEK293T, and HeLa were purchased from ATCC and maintained in T75 or T175 tissue culture flasks in a humidified atmosphere with 5% CO_2 at 27degC. HEK39T and HeLa cells were grown in DMEM supplemented with 10% fetal bovine serum and NIH3T3 cells were cultured supplemented with 10% (FBS) and 1% penicillin/streptomycin. Cell lines were maintained by passaging at confluence, approximately every three days, and P1 aliquots were flash frozen and stored in liquid nitrogen.

All cDD assays (including CsAlog assays from Chapter 2) were performed by seeding 5000K cells per well in a 96 well plate and allowing cells to adhere overnight (min. 12 hours) before addition of compound (DMSO stock diluted into fresh DMEM+FBS, final concentration of DMSO not to exceed 0.1%). Cells were treated with 3uM compound for 48 hours unless otherwise specified before analysis. For FACS analysis, cells were brought to the Bauer Core Facility (Harvard), trypsinized with 0.25% Trypsin-EDTA for 5 min at 37degC, trypsinization was stopped by addition of fresh media. Cells were pelleted at no more than 180xg (RCF) before being resuspended in warm 1X PBS with 1% FBS. Following sorting and collection, cells were further cultured according to standard protocols, above.

Library construction

WT *PPIA* (CypA gene) was cloned in-frame in front of the gene-encoding mCherry. The genes were separated by a short, genetically-encoded Glycine-Serine linker, $(GGGGS)_{2X}$. Three independent condition sets were used to generate diversity. Condition set A utilized 4 ng template, 0.5 mM of each oligonucleotide primer, 5 units Taq polymerase, 5 mM MgCl2, 0.2 mM MnCl2, 0.4 mM dNTPs in equal ratio, and an excess of 0.2 mM dATP and dCTP. Condition set B was identical to A, except that dGTP and dTTP were present in excess. Condition set C utilized the nonnatural nucleotides 8-oxo-dGTP and dPTP to encourage nucleotide misincorporation. The libraries were pooled and ligated into the puromycin-resistant pMSCV Cy mCherry retroviral expression vector (below).



Retroviral transduction of pMSCV plasmid into NIH3T3 cells

*Note: This requires training on viral safety and disposal protocols, and appropriate PPE must be worn at all times. On day 1, HEK293T packaging cells were seeded at 50-70% confluence ($2.4x10^6$ cell in 10cm plates) and incubated overnight. On day 2, packaging cells were transfected using the TransIT-293. 48ul of Mirus TransIT-293 was added to 1ml of serum-free media (Opti-MEM) and allowed to sit for 10 minutes. 15ug of the pMSCV Cy

mCherry viral vector and 3.75ug of each retroviral packaging vector (VSVG, Gag-Pol) were added followed by and additional incubation period of 30 minutes. The transfection mixture was added dropwise to the HEK293T 10cm plate and incubated overnight at 37degC in a viral-designated incubator. On Day 3, target NIH2T2 cells were seeded in 10x15cm dishes with $1x10^7$ cells/dish (to ensure 10-fold library coverage of the $3x10^6$ CypA mutant library). On day 4, 5ml of fresh DMEM +10ul polybrene was added to target 3T3 cells. Media was harvested from HEK293T packaging cells and filtered by 0.45um filter. A 30x dilution of the virus-containing media was prepared and 15ml of 30x viral dilution was used to infect the 3T3s. Mix gently by shaking and place in incubator overnight. Transfected cells were selected for using 2ug/ml puromycin beginning 48 hours post-transduction.

Genomic DNA extraction from enriched libraries/clonal populations and Illumina sequencing design

Genomic DNA was extracted using the epicenter QuikExtract solution. $1x10^4$ 3T3 cells were collected and pelleted in an epidorf tube and resuspended in 250ml of QuickExtract before vortexing for 15 seconds. The tube was then incubated at 65degC for 6 minutes before vortexing again for 15 seconds. The tube was transferred to 98degC for 2 minutes and then stored at -20degC until use. The following primer pairs (next page) were used for Illumina sequencing of the resulting FACS enriched populations.

Primer name	Sequence (5' to 3')
N701-W/F-PPIA	CAAGCAGAAGACGGCATACGAGATTAAGGCGAGTCTCGTGGGCTCGGGCGCCGGA ATTAGATCTCCA
N702-W/F-PPIA	CAAGCAGAAGACGGCATACGAGATCGTACTAGGTCTCGTGGGCTCGGGCGCCGGA ATTAGATCTCCA
N704-W/F-PPIA	CAAGCAGAAGACGGCATACGAGATTCCTGAGCGTCTCGTGGGCTCGGGCGCCGGA ATTAGATCTCCA
N501-W/R-PPIA	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTCCTGCC TCCGCCTCCGAATT
N502-W/R-PPIA	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTCCTGCC TCCGCCTCCGAATT
N507-W/R-PPIA	AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCGGCAGCGTCCTGCC TCCGCCTCCGAATT

Expression and purification of CypA and cDD mutants

PPIA variants encodeing WT, I10F/G14C, and V29D mutants were cloned into PQE-60 plasmid, which was subsequently transformed into the *E. coli* M15 bacterial strain. The bacterial colonies that were resistant to both kanamycin and ampicillin. The cells were harvested by centrifugation at 5000 rpm and lysed by ultrasonication at 4 °C in a buffer containing 50 mM Tris–Cl, 100 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF) pH 8.0 and 10 mM β-mercaptoethanol. After the cell debris was discarded by centrifugation (12 000 rpm, 20 min, 4 °C), 40 ml of the supernatant was loaded onto the Sephacryl S-100 column (2.6 × 100 cm). The material was then eluted with a buffer containing 20 mM Tris–Cl pH 7.8, 1 mM EDTA, 1 mM PMSF and 10 mM β-mercaptoethanol. The fractions that contained CypA were pooled and loaded onto a DEAE-Sepharose column. The elution buffer was the same as that used for the Sephacryl S-100 column, with the exception that it did not contain EDTA and PMSF. The flow-through fractions were analysed by SDS–PAGE, and the homogenous protein fractions were combined and concentrated in a dialysis bag by polyethylene glycol (PEG).

MG132 proteasome inhibitor assay

NIH3T3 cells stably expressing $CypA^{I10F/G14C}$ were treated with 3 uM CsA for 24 hr. Cells were then washed with media and treated with 10 mM MG132 in the presence or absence of 3uM CsA for 4 hr. Immunoblotting was performed with an anti-mCherry antibody (AbCam ab167453).

References

1. Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807-815 (1991).

2. Griffith, J.P. *et al.* X-ray structure of calcineurin inhibited by the immunophilin-immunosuppresant FKBP12-FK506 complex. *Cell* **82**, 507-522 (1995).

3. Choi, J. *et al.* Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239-242 (1996).

4. Zhang, L.H. & Liu, J.O. Sanglifehrin A, a novel cyclophilin-biding immunosuppressant, inhibits IL-2-dependent T cell proliferation at the G1 phase of the cell cycle. *J. Immunol.* **166**, 5611-5618 (2001).

5. Kessler, H., Kock, M., Wein, T., Gehrke, M. Reinvestigation of the Conformation of Cyclosporin A in Chloroform. *Helv. Chim. Acta,* **73**, 1818-1832 (1990).

6. Jin, L., Harrison, S. C., Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. *Proceedings of the National Academy of Sciences.* **99,** 13522-13526 (2002).

7. Ko, S. Y., Dalvit, C. Conformation of cyclosporin A in polar solvents. *Int. J. Pept. Protein Res.,* **40**, 380-2 (1992).

8. Banaszynski, L.A., Chen, L-C., Maynard-Smith, L.A., Ooi, A.G.L, Wandless, T.J. A Rapid, Reversible, and Tunable Method to Regulate Protein Function in Living Cells Using Synthetic Small Molecules. *Cell,***126**, 995-1004 (2006).

9. Maynard-Smith, L.A., Chen, L-C., Banaszynski, L.A., Ooi, A.G.L, Wandless, T.J. A Directed Approach for Engineering Conditional Protein Stability using Biologically Silent Small Molecules. *J. Biol. Chem.* **282**, 24866-24872 (2007).

10. Miyazaki, Y., Imoto, H., Chem, L.C., Wandless, T.J. Destabilizing domains derived from the human estrogen receptor. *JACS*. **134**, 3942-3945 (2012).

11. Alexopoulou A.N., Couchman J.R., Whiteford J.R. The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors. *BMC Cell Biology*. **9**: 2, (2008).

12. Pickart, C.M. Back to the Future with Ubiquitin. *Cell*. **116**, 181-190 (2004).

CHAPTER 2:

STRUCTURE-ACTIVITY RELATIONSHIP STUDIES OF CYCLOSPORINE ANALOGUES

Heavy N-methylation of natural product macrocycle, CsA, is suspected to contribute to its unexpectedly high cell permeability and bioavailability.[1] This has been a major motivator of the use of N-methylation to enhance the ability of cyclic peptides to cross lipid bilayers.[2-5] The evidentiary basis of this work, stems largely from structural studies which predict that all the backbone amide bonds of CsA that are not engaged in transannular hydrogen bonds in the 'cell-permeable conformation,' are otherwise N-methylated.[6] This results in a masking of polar NH bonds, which are disfavoured for cellular permeability, since hydrogen bonding with water molecules in the aqueous environment of the extracellular matrix increases the enthalpic cost of desolvation, a major determinant of permeability. However, to our knowledge, no direct, systematic analysis of the effect of N-methylation on cell permeability and target engagement of CsA has been reported.

In this chapter, we describe the use of the cDD-mCherry cytosensor assay as a proxy for the permeability and target-binding of cyclosporine A analogues (CsAlogs). Furthermore, compounds in which small structural modifications were suspected to affect the global conformational dynamics of CsA were selected for further characterization to better understand the interplay of cell permeability and target engagement in this canonical peptide macrocycle. Firstly, I offer here a short review of selected primary and secondary structural elements of peptides in the context of their interaction with protein binding partners, which will help guide the subsequent discussion of structure-activity relationships in CsA.

Features of peptide structure that dictate function

The architecture of a protein can be described in terms of its primary sequence and secondary structure, and tertiary structure. The linear sequence of amino acids that makes up a peptide is referred to as the primary sequence. **(Fig. 2-1A).** Frequently, the function of proteins is determined by small, structurally defined regions known as secondary structure elements, which are stabilized by hydrogen bonding between NH and CO groups within the peptide backbone. Common secondary structural motifs include α-helices, β-strands, β-sheets, and α-, β-, and γ-turns (**Fig. 2-1, B-D)**. These features are often endowed with an entropic advantage through preorganization to a specific, though often flexible, shape recognized by a binding partner.[7] Finally, to form motifs or domains, elements of secondary structure combine in a variety of manners to produce what is known as tertiary structure.

Short synthetic peptide sequences corresponding to small folded protein epitopes do not tend to form thermodynamically stable structures in an aqueous environment, due to competing hydrogen bonding from water for peptide backbone polar atoms. [8] However, cyclization can restrict peptide conformational flexibility, resulting in increased molecular constraints, fewer rotatable bonds, and more stable peptide conformations than corresponding acyclic analogues.[9] Thus, cyclic peptides may exhibit or mimic the structural elements described above. The nature of these constraints can also promote biological activity.

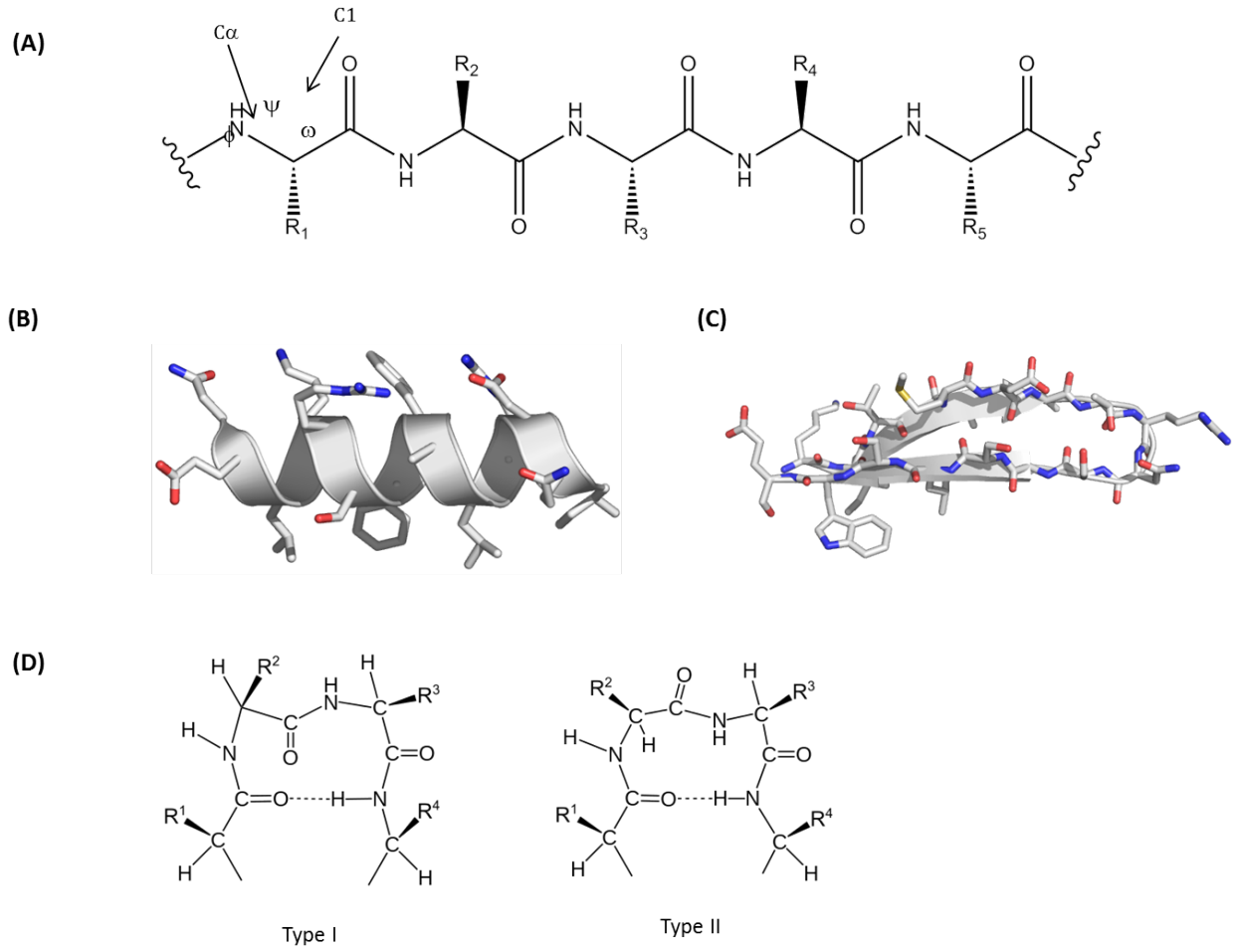


Figure 2-1. Features of primary and secondary protein structure. (A) Primary structure of a protein indicated by the linear sequence of constituent amino acids. R_n denotes a generic amino acid side chain. ϕ is the dihedral angle between the backbone N and Cα. ψ is the dihedral angle between C_α and C1. ω is the dihedral angle between C1 and the backbone N. (B) A right-handed α-helix with sidechains projecting perpendicularly to the helical axis. [PDB: 1NKP] (C) β-strands form a pseudo-flat sheet with sidechains projecting above and below the plane of the sheet. [PDB: 3LPW] (D) Type I (left) and Type II (right) β-turns reverse the direction of strands and helices and are defined by torsion angles.

Turns are motifs that reverse the direction of peptide strands and helices. Turns of increasing size are defined as γ, β, or α subtypes by adjacent phi (ϕ) and psi (ψ) angles of 3, 4 or 5 consecutive amino acids in a peptide sequence; by 7-, 10- and 13-membered hydrogen-bonded rings; and by distances between $C\alpha_i \ldots C\alpha_{i+2}$, $C\alpha_i \ldots C\alpha_{i+3}$, and $C\alpha_i \ldots C\alpha_{i+4}$ residues, respectively. Many small molecule turn mimics are known with potent biological activities, and some have been developed into drugs.[10] However, cyclization alone is often not enough to stabilize a turn structure.[11] Other turn-inducing features are often needed to help constrain a cyclic peptide into a turn. One method is to introduce a *cis*-amide bond into a peptide chain, forming a motif analogous to a β-turn.[12] This is often achieved by incorporating a proline, glycine, D-amino acid, or N-methyl into peptide structures. Proline has the highest tendency of all amino acids to occur in reverse turns in nature, while glycine has the smallest side chain and has the most conformational freedom, enabling it to mimic either, D- or L-amino acids. Incorporating D-amino acids into peptides has a turn-inducing effect, which can be enhanced or locked in place by subsequent cyclization.[14] N-methylation can have a similar effect as proline in favouring some proportion of *cis*-amide conformation that enables reverse turn formation in cyclic peptides.[13]

Structure-activity relationship studies of cyclosporine A analogues (CsAlogs)

To investigate the effect of structural modification on the structure of function of peptidic macrocycle, CsA, we used the selectively destabilized CypA mutant cytosensor (cDD-mCherry) described in **Chapter 1**. Through a collaboration with Dylan Stiles and Ping Wang at Warp Drive Bio, LLC, we were fortunate to have access to a library of CsAlogs at our

disposal. Thus, the cytosensor cell line was seeded into 96 well plates, and cells were treated with 3uM CsA or analogue in triplicate for 48 hours, at which time cells were fixed and western blot quantification of mCherry levels (as a surrogate for CypA stabilization) was performed. Readouts were normalized to cell number, and are reported as a total of CsA-induced fluorescence. The workflow for which is given in **Figure 2-2.** It is important to note that this readout represents the cumulative effect on cell permeability and changes in affinity for CypA, which may be differentially affected by structural changes. To isolate these features, subsequent *in vitro* experiments were performed on selected analogues.

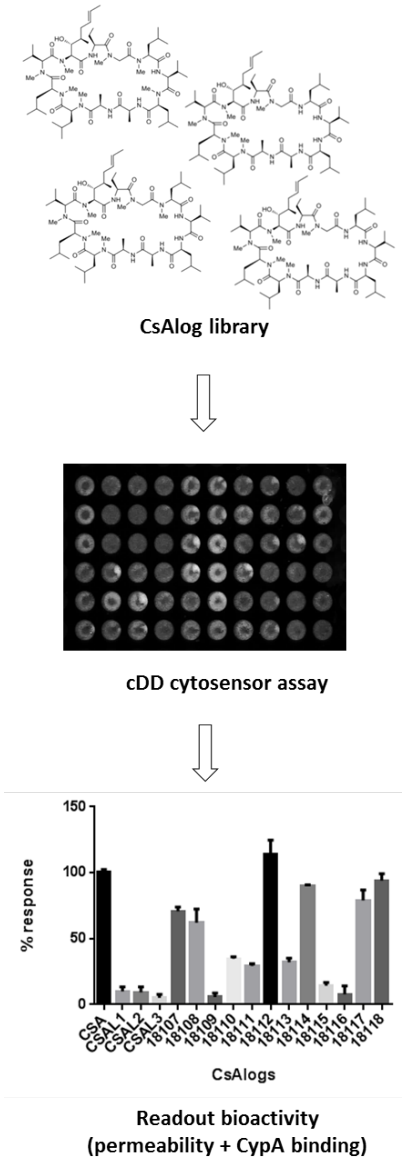


Figure 2-2. Workflow for cDD cytosensor analysis of CsAlogs. Selected compounds from a library of CsA analogs were tested for their ability to penetrate 3T3 cells and bind to CypA (incubation with 3uM CsA, CsAlog, or DMSO for 48 hours). In cell western blotting for CypA DD fusion, cDD-mCherry, enabled quantification of relative ligand binding, resulting in increased protein stability. Experiments were performed in triplicate and percent response was first normalized to DMSO control, and then expressed as a percent of CsA-induced stabilization.

We selected 33 analogues to test, all of which had either backbone or side chain modifications at MeLeu4, Val5, or MeLeu6. These residues are located on the face of CsA that contacts calcineurin in the crystal structure of the trimeric inhibitory complex. Neither the backbone, nor side chains, of these residues is expected to make direct contacts with CypA **(Fig. 2-3A)**. Therefore, to affect the affinity of complexation with CypA, the modification would need to affect the energy landscape of the preferred conformations in a high dielectric environment, such as water. Changes to overall conformational dynamics, particularly those that interfere with or otherwise alter the formation of transannular hydrogen bonds, could also serve to increase or decrease membrane permeability. The location of residues 4-6 in the crystal structure of CsA in an apolar solvent, $CDCl_3$, is shown in **Figure 2-3B**. For reference, we have also once again provided the annotated structure of CsA **(Fig. 2-3C)**.

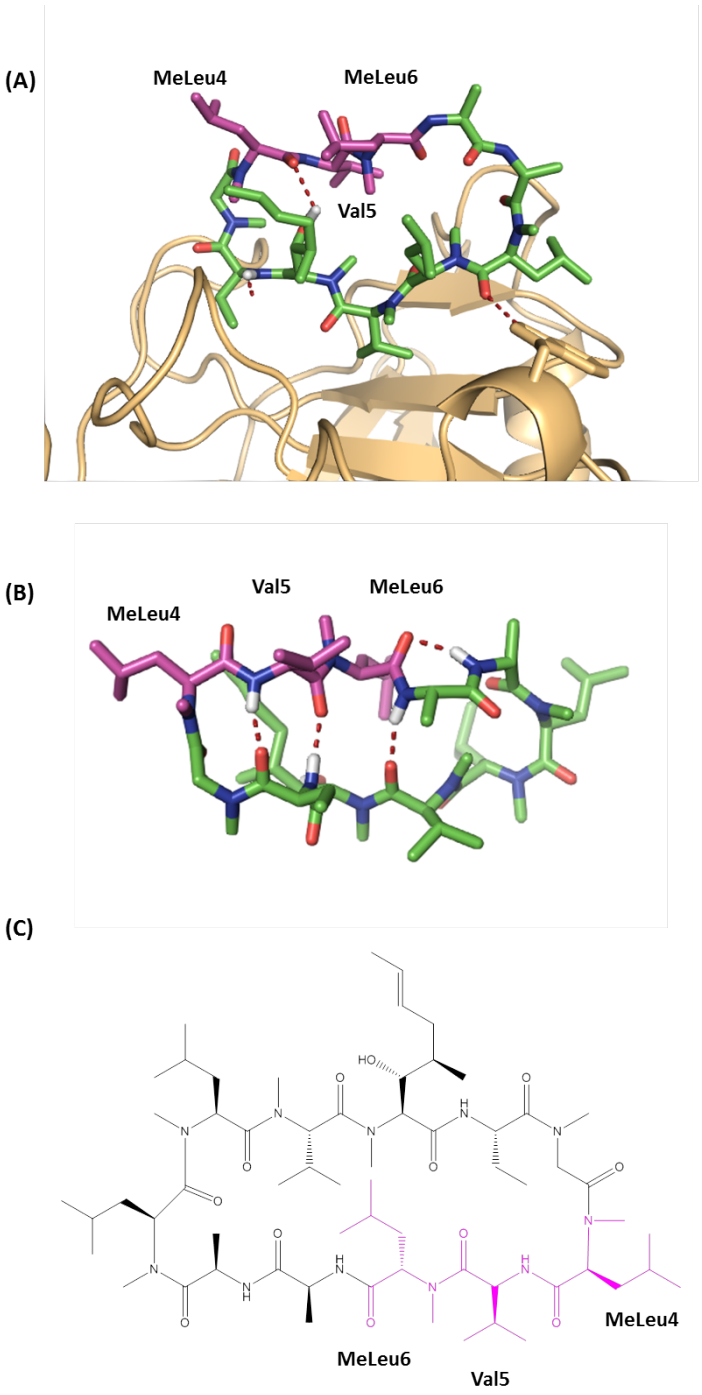


Figure 2-3. Location of modified CsAlog residues MeLeu4, Val5, and MeLeu6 with respect to CypA. (A) Complex of CsA bound to CypA. Residues 4-6 (magenta) do not contact CypA, but do form the contact surface with target, calcineurin (not shown). [PDB ID: 2Z6W] (B) Crystal structure of CsA in $CDCl_3$ [CSD ID: DEK SAN] (C) Structure of CsA.

We tested three desmethyl CsAlogs (CSAL1-3) to determine the effect of N-methylation at MeLeu4 and MeLeu6 on target binding and cell permeability. We also evaluated substitutions at the side chains of MeLeu6 (CSAL4-15), Val5 (CSAL16-28), and MeLeu4 (CSALs 29-33). In so doing, we sought to understand how both the nature and position of side chain modifications affects binding of CsA to CypA. The results of the cytosensor assay are normalized to DMSO-treated controls and reported as a percent of CsA treated (**Fig. 2-4)**.

Considering first the side chain modifications, conservative mutations of MeLeu or Val to other hydrophobic and/or beta-branched amino acids (valine, norvaline, leucine, norleucine, isoleucine, threonine, cyclopropyl-gly, etc.), predictably, gave high responses (>70%), indicating little penalty on bioactivity from these substitutions (bioactivity for the purpose of this discussion is defined as the cumulative effect of cell penetration and CypA binding, not immunosuppressive action). Substitution of bulkier groups, such as phenylalanine and tyrosine, resulted in a significant reduction in bioactivity (<30%) for all positions tested. It is possible that steric constraints imposed by large aromatic rings could increase the energy barrier for interconversion between different states (cell-permeable and target-binding) or make one or the other state less accessible. However, additional studies would be required to verify this hypothesis. We were, however, most interested in the modifications that resulted in a large loss in bioactivity (<30%) and that, based on the discussion of peptide structure above, have the potential to induce large conformational changes that alter either target-binding or cell permeability. An analysis of the structure-activity relationships in the context of what is known about CsA structure is given for the following substitutions: desmethyl (CSAL1-3), sarcosine (CSAL12,32), and D-amino acids (CSAL14-15, 27-28, 33).

CsAlog	AA4	AA5	AA6	cDD cytosensor % Response
CSA	MeLeu	Val	MeLeu	100.0 ± 2.3
CSAL1	**Leu**	Val	MeLeu	10.1 ± 3.3
CSAL2	MeLeu	Val	**Leu**	9.3 ± 4.0
CSAL3	**Leu**	Val	**Leu**	5.0 ± 2.6
CSAL4	MeLeu	Val	**MeThr**	113.8 ± 11.1
CSAL5	MeLeu	Val	**MeNva**	94.0 ± 5.3
CSAL6	MeLeu	Val	**MeVal**	89.9 ± 0.9
CSAL7	MeLeu	Val	**MeNle**	79.0 ± 7.9
CSAL8	MeLeu	Val	**MeIle**	70.9 ± 3.0
CSAL9	MeLeu	Val	**MeAla**	62.0 ± 10.4
CSAL10	MeLeu	Val	**MePhe**	34.7 ± 1.5
CSAL11	MeLeu	Val	**MeTyr**	32.1 ± 3.1
CSAL12	MeLeu	Val	**Sar**	29.4 ± 1.6
CSAL13	MeLeu	Val	**MeTyr(PO3H2)**	14.3 ± 2.5
CSAL14	MeLeu	Val	**D-MeVal**	7.6 ± 6.5
CSAL15	MeLeu	Val	**D-MeAla**	5.5 ± 3.1
CSAL16	MeLeu	**Cyclopropyl-Gly**	MeLeu	101.1 ± 9.0
CSAL17	MeLeu	**Ile**	MeLeu	85.6 ± 6.2
CSAL18	MeLeu	**Nva**	MeLeu	76.5 ± 6.3
CSAL19	MeLeu	**4-pyridylAla**	MeLeu	69.6 ± 1.1
CSAL20	MeLeu	**Nle**	MeLeu	68.9 ± 2.3
CSAL21	MeLeu	**Leu**	MeLeu	65.5 ± 4.1
CSAL22	MeLeu	**CyclobutylAla**	MeLeu	63.5 ± 4.4
CSAL23	MeLeu	**4-thiazolylAla**	MeLeu	58.4 ± 4.5
CSAL24	MeLeu	**Tyr(PO3H2)**	MeLeu	40.7 ± 5.3
CSAL25	MeLeu	**Ala**	MeLeu	26.4 ± 8.2
CSAL26	MeLeu	**Phe**	MeLeu	8.8 ± 4.0
CSAL27	MeLeu	**D-Val**	MeLeu	5.2 ± 0.9
CSAL28	MeLeu	**D-Ala**	MeLeu	3.1 ± 1.4
CSAL29	**MeVal**	Val	MeLeu	125.9 ± 4.2
CSAL30	**MeNle**	Val	MeLeu	89.9 ± 10.5
CSAL31	**MePhe**	Val	MeLeu	25.1 ± 6.8
CSAL32	**Sar**	Val	MeLeu	9.1 ± 3.9
CSAL33	**D-MeLeu**	Val	MeLeu	2.2 ± 4.2

Figure 2-4. Effect of structural modifications to CsA on cell penetration and target engagement as determined by cDD cytosensor assay. cDD cytosensor % response is normalized to DMSO and reported relative to CsA response (100%). Structural changes to relative to CsA are shaded in grey.

Recall that the conformation of CsA is widely dependent on its environment. In apolar solvents, such as $CHCl_3$, THF, or CH_3CN, one main stable conformer dominates. This suspected 'cell-permeable' conformation is stabilized by four stable intramolecular H-bonds involving all the backbone amides that have not been masked by N-methylation. These H-bonds cause residues 11–7 to form an antiparallel β-sheet and residues 3–4 a type-II′ β-turn. Moreover one *cis*-peptide bond between MeLeu9 and MeLeu10 is present and all hydrophobic carbon chains are exposed to solvent.[14,15] By contrast, the crystal structure of CypA-CsA-calcineurin shows an all-*trans* peptide backbone conformation when in engaged in the ternary complex. Computational modeling suggests that when the polarity of the solvent increases, such as in acetone, dimethylsulfoxide, or MeOH, hydrogen bonds disappear and the structure loses its rigidity, leading to the coexistence of various stable conformers.[16,17]

To tease out the mechanistic details that alter the bioactivity of CsAlogs described above, it was necessary to separate cell permeability from protein binding. We determined the ligand affinity for the interaction with CypA by surface plasmon resonance (SPR), the ability to diffuse through a lipid bilayer by parallel membrane permeability assay (PAMPA), and are engaged in on-going structural investigation by computational modeling and nuclear magnetic resonance (NMR) (**Fig. 2-5)**.

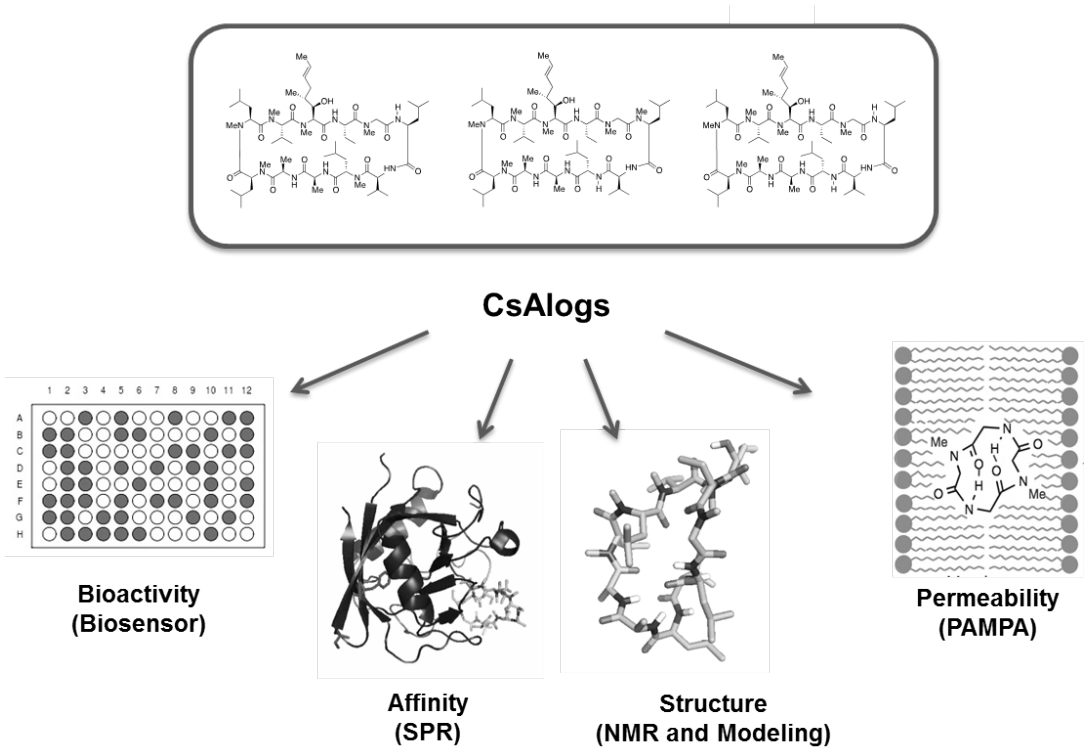


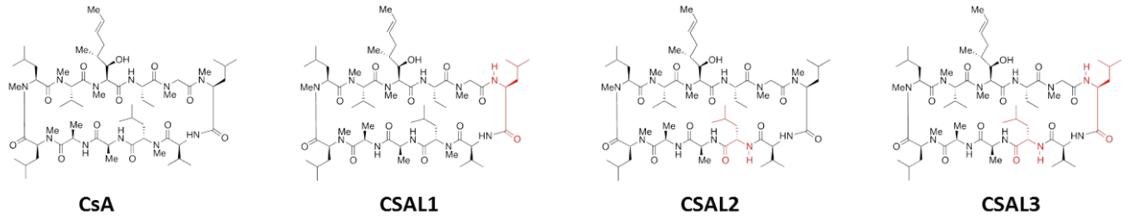
Figure 2-5. Workflow for structure-activity studies of CsA analogues with reduced bioactivity.

Conformational control by N-Methylation is critical for target engagement

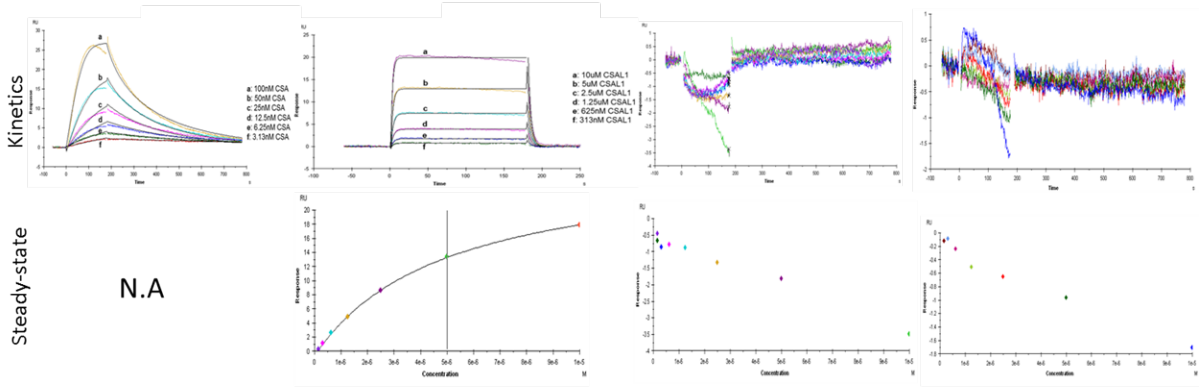
Desmethyl CsAlogs were evaluated for their ability to bind to CypA by SPR, and to permeate cell membranes by PAMPA **(Fig. 2-6)**. It is hypothesized that N-methylation of backbone amides in the CsA cell permeable state mask polar H-bonds, resulting in a decreased energetic penalty of desolvation in an aqueous environment. We were somewhat surprised, however, to find that loss of one or two N-methyl groups at MeLeu4 or MeLeu6 resulted in little reduction in cell permeability, but did completely abrogate binding to the

protein target, CypA. Considering that N-methylation increases the accessibility of *cis* peptide bonds, it is possible that this could result in alterations in the backbone conformation that disfavor the target binding conformation.

(A)



(B)



(C)

Compound	AA4	AA5	AA6	Biosensor activity (% relative to CsA)	Kd by SPR (uM)	PAMPA $logP_e$ (cm/s)	ΔG (CHCl3-water) (kcal/mol)
CSA	MeLeu	Val	MeLeu	100.0 ± 2.3	0.025	-5.48	14.1
CSAL1	Leu	Val	MeLeu	10.1 ± 3.3	5	-5.61	16.8
CSAL2	MeLeu	Val	Leu	9.3 ± 4.0	>10	-5.33	9.4
CSAL3	Leu	Val	Leu	5.0 ± 2.6	>10	-5.39	6.8

Figure 2-6. Structure-activity relationship studies of desmethyl CsAlogs. (A) Structure of CsA and desmethyl analogues (CSAL1-3). (B) Binding affinities determined by SPR show that dimethyl analogues CSAL2 and CSAL3 no longer bind CypA. CSAL1 appears to bind weakly, with at K_d of approximately 5uM. (C) Summary table for desmethyl analogues including permeability coefficients (less negative indicated better permeability), and predicted energy for transition from $CHCl_3$ to water (lower predicts more permeable).

We have begun structural studies to determine the effect of N-methylation at MeLeu4 and MeLeu6 on the conformation of CsA in an aqueous environment. We used MacroModel to predict the lowest energy conformers of the desmethyl analogues in a high dielectric environment **(Fig. 2-7)**. We observed significant structural changes in the preferred state, particularly in the non-binding CSAL2 and CSAL3, which adopt a more compact fold, owing to a predicted bifurcated hydrogen bond between the backbone carbonyl of Val5 and the backbone amides of Ala7 and D-Ala8. Notably, however, these new transannular interactions do not occur directly with the newly unmasked amide hydrogens.

This serves only as a first pass analysis, since we expect that many interconverting states are likely to exist in water and that a static picture does not provide a complete depiction of CsA's dynamic nature. Therefore, we plan to do additional modeling experiments including an interrogation of the number of predicted conformations, molecular dynamics simulations, and identifying if particular peptide bonds are more likely to exist as *cis*. Recall that, Xtal structures of CsA in complex with CypA and calcineurin show an all *trans* backbone. Therefore, if all *trans* is less energetically accessible it could result in a reduction in binding. Alternately, if a greater number of states have become accessible by relieving restraints imposed by N-methylation such that the target binding state is populated less frequently, this too could affect the affinity for CypA.

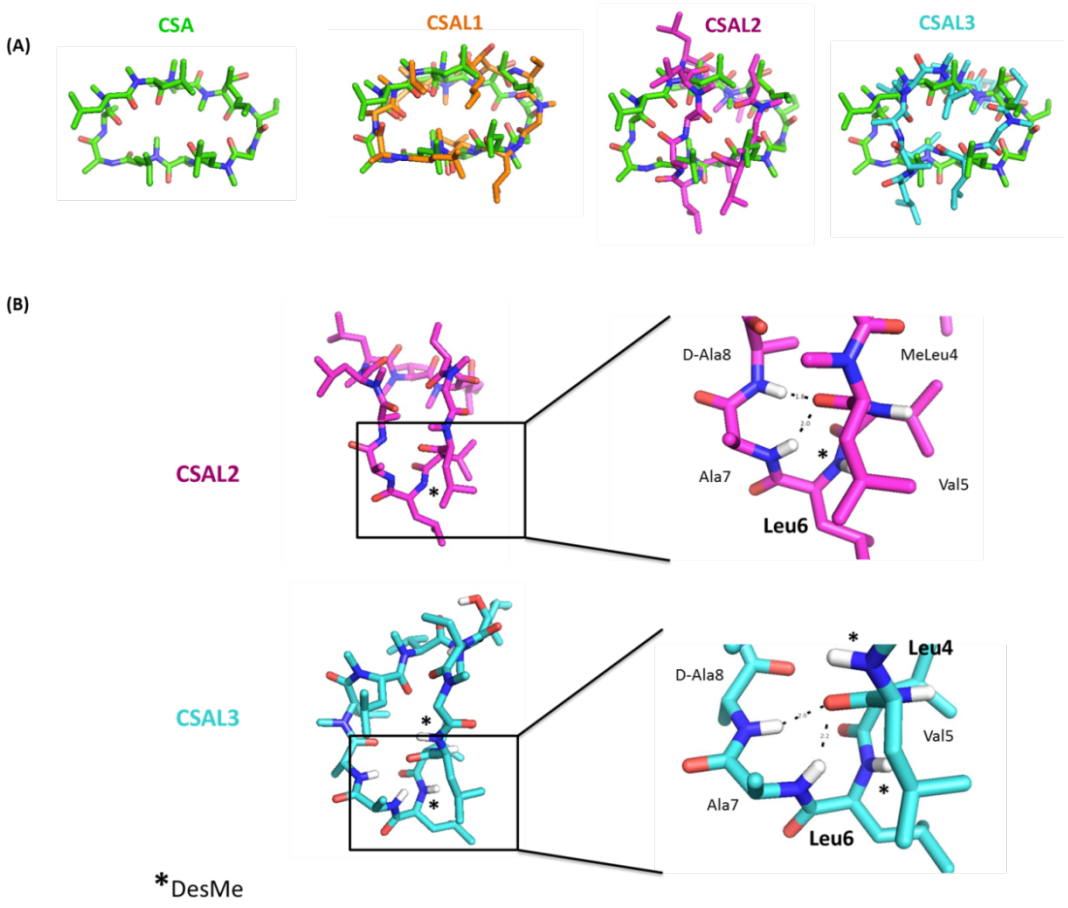
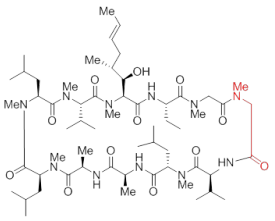


Figure 2-7. Predicted lowest energy confomers of desmethyl analogues in a high dielectric environment. (A) Macromodel predictions of the lowest energy conformers of CSAL2 and CSAL3 show significant global conformational changed. (B) Modeling predicts the presence of a new bifurcated hydrogen bond between the backbone carbonyl of Val5 and the backbone amides of Ala7 and D-Ala8 resulting in a more conformationally-restrained, 'buckled' structure.

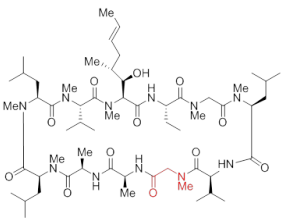
Conformational constraints may result in structural preorganization that favours bioactive conformations

Next, sarcosine-substituted analogs were evaluated for their ability to engage CypA and for their rate of diffusion across a lipid membrane **(Fig, 2-8)**. Since the side chain of sarcosine (N-methyl glycine), is only a proton, there is loss of stereochemistry at Cα and a greater range of psi and phi angles are accessible, resulting in increased conformational flexibility. We observed a significant, although not complete, reduction in both target binding and rate of membrane penetration as a result of this modification. Like the desmethyl analogues, this may suggest that structural preorganization is important for target binding of CsA to CypA. Computational simulations to assess the predicted conformational flexibility of CsAlogs relative to CsA will be valuable to test this hypothesis.

(A)

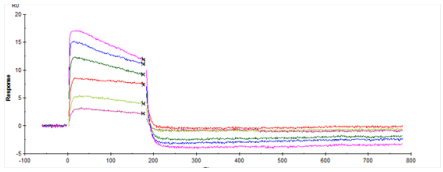


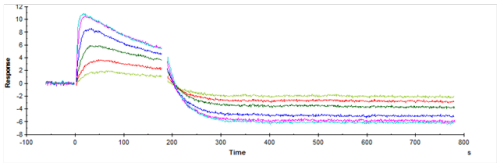
CSAL12



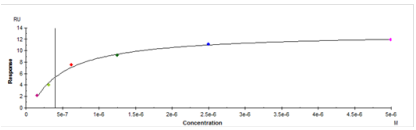
CSAL32

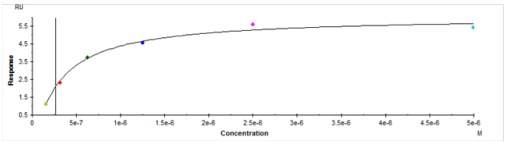
(B)





Steady-state



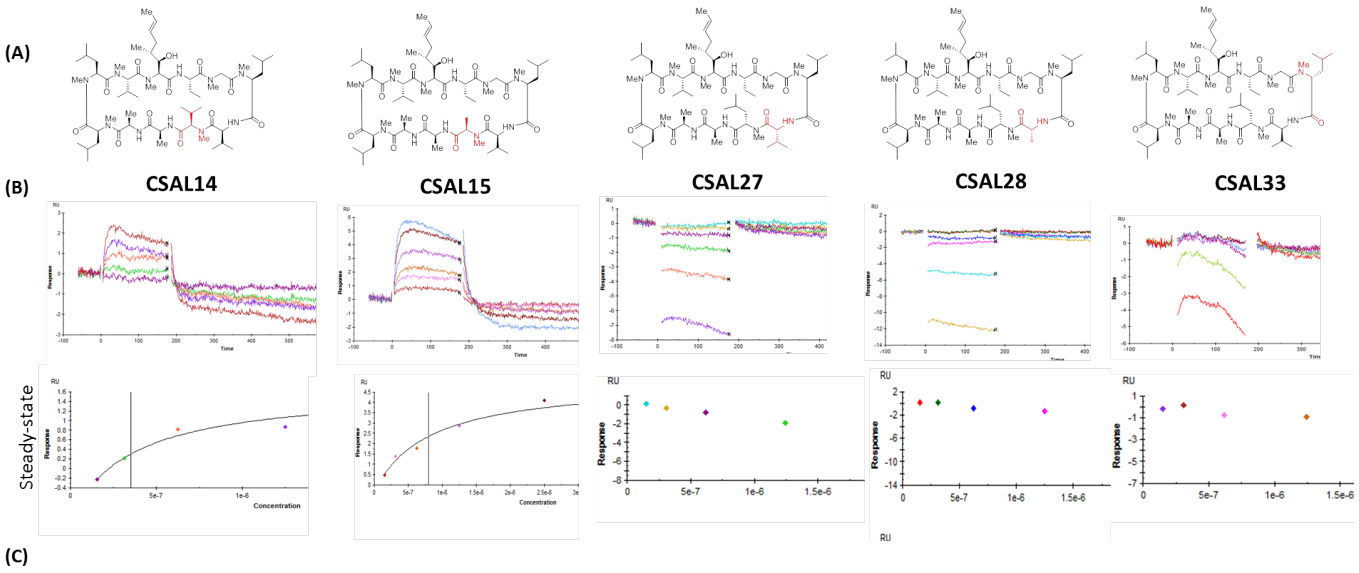


(C)

Compound	AA4	AA5	AA6	**Biosensor activity (% relative to CsA)**	**Kd by SPR (uM)**	**PAMPA logP$_e$ (cm/s)**	**ΔG $_{(CHCl3-water)}$ (kcal/mol)**
CSA	**MeLeu**	**Val**	**MeLeu**	100.0 ± 2.3	0.025	-5.48	14.1
CSAL12	**MeLeu**	**Val**	**Sar**	29.4 ± 1.6	0.396	-5.61	25.2
CSAL32	**Sar**	**Val**	**MeLeu**	9.1 ± 3.9	0.268	-5.93	29.2

Figure 2-8. Structure-activity relationship studies of sarcosine CsAlogs. (A) Structure of CsA sarcosine analogues (CSAL12,32). (B) Binding affinities, determined by SPR show a 10-fold reduction in affinity relative to CsA. (C) Summary table for sarcosine analogues including permeability coefficients (less negative indicated better permeability), and predicted energy for transition from $CHCl_3$ to water (lower predict more permeable).

Finally, we assayed the effect of incorporation of an N-methylated D-amino acid at positions 4 or 6 **(Fig. 2-9)**. D-amino acids have a turn-inducing effect in constrained cyclic peptides as well as other protein secondary structures. While we observed a reduction in permeability of nearly a log unit for most compounds tested, notably, two analogues (CSAL14, 15) reproducibly retained sub-micromolar affinity for CypA. We are continuing to reconcile this relatively recent data.



Compound	AA4	AA5	AA6	**Biosensor activity (% relative to CsA)**	**Kd by SPR (uM)**	**PAMPA logP$_e$ (cm/s)**	**ΔG (CHCl3-water) (kcal/mol)**
CSA	**MeLeu**	**Val**	**MeLeu**	100.0 ± 2.3	0.025	-5.48	14.1
CSAL14	**MeLeu**	**Val**	**D-MeVal**	7.6 ± 6.5	0.596	-6.33	-0.2
CSAL15	**MeLeu**	**Val**	**D-MeAla**	5.5 ± 3.1	0.805	-5.99	7.2
CSAL27	**MeLeu**	**D-Val**	**MeLeu**	5.2 ± 0.9	> 5uM	-6.68	49.6
CSAL28	**MeLeu**	**D-Ala**	**MeLeu**	3.1 ± 1.4	> 5uM	-6.69	18.4
CSAL33	**D-MeLeu**	**Val**	**MeLeu**	2.2 ± 4.2	> 5uM	-6.86	20.7

Figure 2-9. Structure-activity relationship studies of D-amino acid CsAlogs. (A) Structure of CsAlogs D-amino acid substituted (CSAL14-14, 27-28, 33). (B) Binding affinities determined by SPR. (C) Summary table for D-amino acid CsAlogs.

As an additional validation of the reliability of out binding data, we selected an analogue from the original panel of 33 (CSAL29) that, based on high activity in the cytosensor assay and conservative nature of the amino acid change (MeLeu6 to MeVal6), should retain binding to CypA. We found that, as expected, this analogue did retain binding. However, we note also that, with a K_d several fold less than the reported literature value for CsA (30nM), this suggests, perhaps unexpectedly, that even small side chain modifications can affect the protein-ligand interaction. We are currently quantifying the binding of additional CsAlogs with conservative substitutions to identify if this is a trend or an outlier.

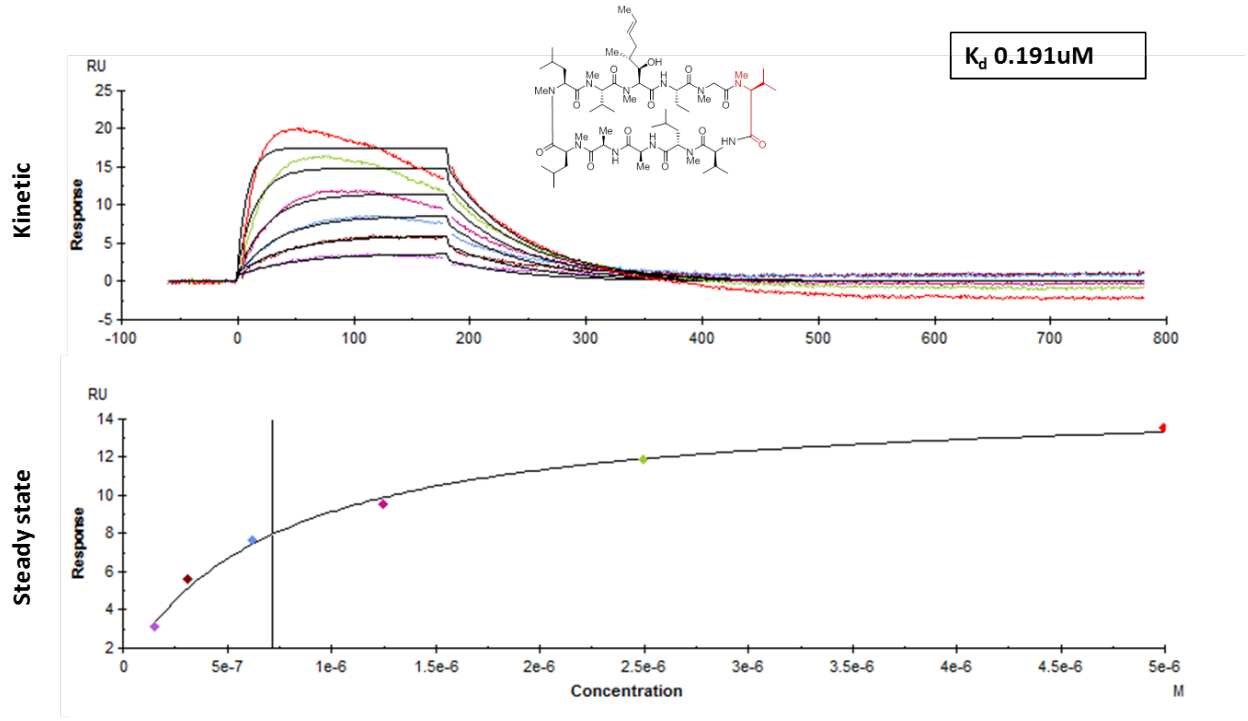


Figure 2-10. Validation of SPR data with CypA-binding CsAlog. CSAL29 contains a MeVal in place of a MeLeu in position amino acid position six, resulting in a relatively modest, but significant, loss of binding to protein target, CypA.

Validation of a computational method for the prediction of peptide permeability

In parallel to the PAMPA assays, we also employed a computational method developed by the Jacobson lab at UCSF to predict the permeabilities of CsAlogs.[18] This computation was performed by Ashutosh Jogalekar at Warp Drive Bio, LLC., and Ashutosh and I were not aware of each other's results until both assays were completed, in order to have two independent assessments. This atomistic physical model has been used to accurately predict the relative permeabilities of small N-methylated cyclic peptides, therefore we felt it would be useful to know whether it could be extended to larger, more complex and conformationally-labile peptide macrocycles. Like other similar computations, this model relies primarily on calculating the compound's desolvation energy; additionally, the Jacobsen lab has added extensive conformational sampling, making it particularly applicable to our interest in CsA structure-permeability relationships. The predicted energy differences between the $CHCl_3$ and aqueous states ($\Delta G_{CHCl3\text{-}water}$) for the CsAlogs are given in their respective structure-activity summaries above **(Fig. 2-6, 2-8, 2-9)** and the correlation with experimental effective permeability, as determined by PAMPA, is given below **(Fig. 2-10)**. We observed a very weak correlation between $\log P_e$ and $\Delta G_{CHCl3\text{-}water}$ ($R^2 = 0.14$). It is possible that, for larger macrocycles such as CsA, the large number of possible states in a high dielectric environment makes it more challenging to predictlowest energy conformers, and therefore permeability. Looking at only the D-amino acid CsAlogs, we see that PAMPA results indicate relatively poor permeability for all compounds tested, whereas the Jacobson permeability prediction places some analogues well above and some well below the expected permeability of CsA. However, if you disregard the D-amino acid analogues, the correlation

coefficient for the DesMethyl analogues and sarcosine analogues alone is much closer to the values reported by Jacobson ($R^2 = 0.81$). We plan to share our results with the Jacobson lab to aid in development of a more broadly applicable method to predict the permeabilities of large peptidic macrocycles.

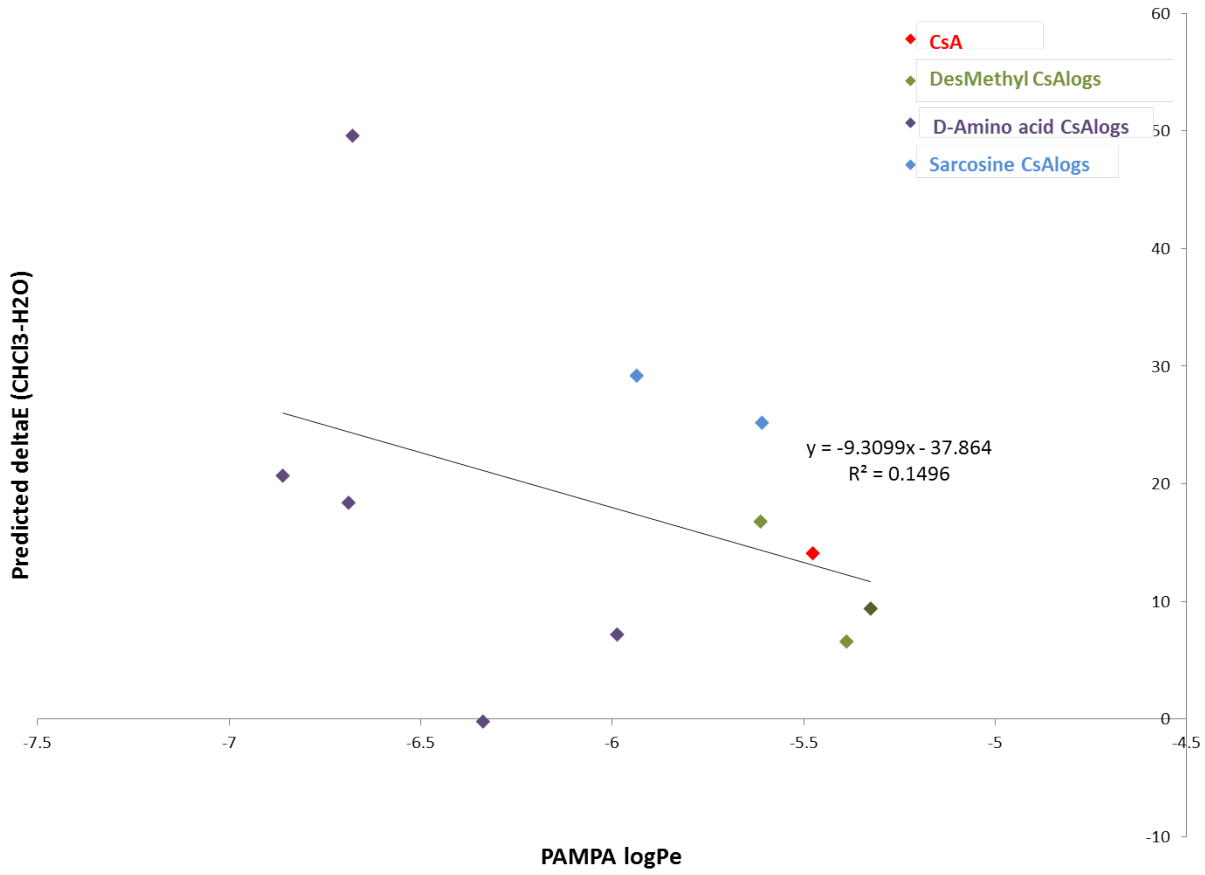


Figure 2-10. Correlation of computationally-predicted and experimentally derived permeabilities. A weak correlation ($R^2 = 0.14$) between the experimentally-determined effective permeability ($logP_{e)}$ and the *in silico* predicted permeabilty ($\Delta G_{CHCl3\text{-}water}$) was observed.

CsAlog				Bioactivity	Binding to CypA	Permeability	
Compound	AA4	AA5	AA6	Biosensor activity (% relative to CsA)	Kd by SPR (uM)	PAMPA $logP_e$ (cm/s)	ΔG (CHCl3-water) (kcal/mol)
CsA	MeLeu	Val	MeLeu	100 ±2.3	0.025	-5.48	14.1
CSAL1	Leu	Val	MeLeu	10.1 ±3.3	~ 5	-5.61	16.8
CSAL2	MeLeu	Val	Leu	9.3 ±4.0	> 10	-5.33	9.4
CSAL3	Leu	Val	Leu	5.0 ±2.6	> 10	-5.39	6.6
CSAL14	MeLeu	Val	D-MeVal	7.6 ±6.5	0.596	-6.33	-0.2
CSAL15	MeLeu	Val	D-MeAla	5.5 ±3.1	0.805	-5.99	7.2
CSAL27	MeLeu	D-Val	MeLeu	5.2 ±0.9	> 5	-6.68	49.6
CSAL28	MeLeu	D-Ala	MeLeu	3.1 ±1.4	> 5	-6.69	18.4
CSAL33	D-MeLeu	Val	MeLeu	2.2 ±4.3	>5	-6.86	20.7
CSAL12	MeLeu	Val	Sar	29.4 ±1.6	0.396	-5.61	25.2
CSAL32	Sar	Val	MeLeu	9.1 ±3.9	0.268	-5.93	29.2

DesMethyl; D-Amino acid; Sarcosine

Figure 2-11. Summary for CsAlog SAR studies thus far.

Figure 2-11 summarizes the bioactivity, target binding, and experimental and predicted permeabilities for the CsAlogs evaluated thus far. Taken together, our analysis of CsAlog structure and function suggests strongly that precise evolutionary tailoring is required to achieve such a molecule exquisitely tuned molecule. We hope that this work, in concert with many excellent studies on cyclic peptide structure and function, will provide greater insight into the complex energy landscape that dictates the activity of these macrocycles, and that resulting guidelines will be useful in the *de novo* design of cell-penetrating peptides with high molecular complexity.

We plan to continue this line of inquiry and gain greater mechanistic insight with NMR studies in an apolar solvent, and computational modeling to understand the relationship between the multiple the conformers expected in an aqueous environment. For the desmethyl project in particular, we will probe the prevalence of *cis* and *trans* bonds, as compared to the native CsA ligand. We are eager to know how this data will fit with our working hypothesis that structural preorganization is important for both cell membrane permeability and target selectivity.

Acknowledgements

I am very grateful for my collaborators at Warp Drive Bio, LLC., without whom this work would not have been possible. Khian Hong Pua collected the initial SPR of CSAL1-3 that demonstrated a loss of affinity and altered the direction of the project. Dylan Stiles and Ping Wang synthesized the CsAlog compounds and provided helpful discussions throughout the project. Ashutosh Jokalegar ran the computational structural analysis and has provided many stimulating conversations on computational chemistry.

Methods

In cell western

General cell culture protocols were performed as described in the methods section in Chapter 1. NIH3T3 cells stably transfected with the cDD-mCherry retroviral expression construct were seeded in 96-well costar clear bottom plates at 5000K cells/well in DMEM (Gibco) medium with 10% fetal bovine serum. Cells were incubated at 37degC overnight (min. 12 hours) to allow for attachment prior to treatment. Media was removed and replaced with 100ul/well fresh medium containing 3uM CsA, or CsAlog, or mock treatment (final concentration of 0.1% DMSO). Plates were incubated at 37degC for 48 hours. Following the incubation period, media was removed and cells were fixed with 150ul/well fresh 4% paraformaldehyde in 1X PBS for 20 min at room temperature with no shaking. Paraformaldehyde was removed and cells were washed 5 times with permeabilization buffer, 1X PBS + 0.1%Triton X-100 for 5 min per wash. Blocking was done with 50ul/well of LI-COR Odyssey blocking buffer for 90 min at RT with moderate shaking. 50ul/well primary antibody at 1:500 dilution in blocking buffer (mouse monoclonal anti-mCherry, Abcam 1C51) was incubated 4 hr-to-overnight at RT (4degC if overnight). Plate was washed with 200ul/well 5 times with 1X PBS +0.1%Tween for 5 min each. Seconday antibody (IRDye 800CW G-anti-M, LifetTechnologies, 1:800 dilution) and cell stain (CellTag700, LifeTechnologies, 1:500 dilution) were diluted into blocking buffer, and cells were incubated with 50ul/well for 60min with gentle shaking at RT. Cell were washed a final 5x with 1XPBS +0.1Tween and visualized on the LI-COR Odyssey CLx infrared imaging system in 800 and 700nm channels. Data processing was accomplished by background subtraction of

DMSO control, and normalization of the protein levels in the 800 channel (anti-mCherry-IRDye) to the cell number as determined 700 channel (CellTag700).

Surface plasmon resonance

Cyclophilin construct, PPIA-H12, was expressed and purified as described in the methods in Chapter 1. All runs were performed with HBSP+ buffer (GE Healthcare) with 5mM EDTA. Stable capture was achieved on a Biacore Ni-NTA chip following nickel activation. EDTA regeneration was performed in between cycles. Biacore run was performed with multi-cycle kinetics. CsA was evaluated from 0.1nM to 200nM range. CsAlogs were evaluated first in the 50nM to 10uM range, and if bound, subsequently narrowed in to the range of K_d. Kinetic analysis was performed where binding kinetics were sufficiently slow, and K_d reported is a function of koff/kon. For ligand with low rapid kinetics, steady state affinity analysis was performed.

Parallel artificial membrane permeability assay (PAMPA)

***** PAMPA was perfomed using Corning GenTest Pre-Coated PAMPA plates, as standard EMD Millipore plates were found to retain hydrophobic small molecules in dodecane solvent layer.

96-well Corning GenTest Pre-Coated PAMPA were warmed to RT for 30min prior to use in the PAMPA permeability test. The acceptor plate was prepared by adding 300 uL of 5% DMSO/PBS to each well. 10-50 uM solutions of the cyclic peptides were prepared in 5% DMSO/PBS buffer. 150 uL of the peptide solutions were added to the donor wells. The

donor plate was then placed on top of the acceptor plate so that the artificial membrane was in contact with the buffer solution below. A lid was placed on the donor well, and the system was covered with a glass evaporating dish and left 16hr at room temperature.

Acceptor and donor well concentrations were measured by LCMS (Thermo LTQ) using selected ion monitoring (SIM) mode. An internal standard of H2N-Tyr(O*t*Bu)- CO2H was run with each sample so that compound-to-standard peak area ratios from the TIC detector could be used to determine relative concentrations.

Computational methods.

In brief, the conformational space of the cyclic peptides was searched in a semi-exhaustive manner using technology originally developed for protein loop prediction. The conformational search occurs in torsion angle space, with the backbone and side chain conformations selected from libraries: backbone-independent rotamer libraries for the side chains, and discretized versions of the 'Ramachandran' maps for the backbone degrees of freedom, with separate libraries for Gly, Pro, and all other amino acids (and D-chirality versions of these). Conformations that result in closed backbone structures and are free of steric clashes were subjected to additional energy-based side chain rotamer optimization and energy minimization. In the work reported here, a minimum of 200 energy minima (with distinct backbone conformations) were generated, except for cyclic peptides with 4 or more N-methyl groups. These cyclic peptides were found to be highly constrained, such that it was frequently impossible to generate 200 distinct backbone conformations. For these peptides, the minimum number of conformations was set to 25 or smaller until a distinct conformation was identified. The computational procedures described previously were modified to

accommodate N-methyl amino acids. Jacobsen *et al.* developed new backbone libraries to to represent the more restricted space of N-methyl amino acids (both L and D chiralities). Although the N-methyl groups can restrict side chain conformations, we did not attempt to modify the side chain rotamer libraries; side chain orientations with steric clashes are automatically eliminated during rotamer optimization. Both *cis* and *trans* peptide conformations were sampled for the peptide bonds preceding N-methyl amino acids, as with Pro. A much more detailed explanation of the methodology is provided in reference 18.

References

1. Kessler, H., Kock, M., Wein, T., Gehrke, M. Reinvestigation of the Conformation of Cyclosporin A in Chloroform. *Helv. Chim. Acta,* **73**, 1818-1832 (1990).

2. Thansandote, P., Harris, R. M., Dexter, H. L., Simpson, G. L., Pal, S., Upton, R. J., Valko, K. Improving the passive permeability of macrocyclic peptides: balancing permeability with other physicochemical properties. *Bioorg. Med. Chem*. **23**, 322−327 (2015).

3. Ovadia, O., Greenberg, S., Chatterjee, J., Laufer, B., Opperer, F., Kessler, H., Gilon, C., Hoffman, A. The effect of multiple N-methylation intestinal permeability of cyclic hexapeptides. *Molecular Pharmaceutics.***8**, 479−487 (2011).

4. Alex, A., Millan, D. S., Perez, M., Wakenhut, F., Whitlock, G. A. Intramolecular hydrogen bonding to improve membrane permeability and absorption in beyond rule of five chemical space. *MedChemComm*. **2**, 669−674 (2011).

5. Chatterjee, J., Mierke, D. F., Kessler, H. Conformational preference and potential templates of N-methylated cyclic pentaalanine peptides. *Chemistry*. **14**, 1508−1517 (2008).

6. Bockus, A.T., McEwen, C.M., Lokey, S.R. Form and function in cyclic peptide natural products: A pharmacokinetic perspective. *Current Topic in Medicinal Chemistry.* **13**, 821-836 (2013).

7. Arkin, M.R. & Wells, J.A., Small molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Reviews Drug Discovery.* **3**, 301-317 (2004).

8. Scholtz, J.M., Baldwin, R.L. The mechanism of alpha-helix formation by peptides. *Annu. Rev. Biophys. Biolmol. Stuct.***21**, 95-1118 (1992).

9. Fairlie, D.P., Abbenante, G., March, D.R. Macrocyclic peptidomimetics: Forcing peptides into bioactive conformations. *Curr. Med. Chem.* **2**, 654-686 (1995).

10. Jones, R.M., Boatman, P.D., Semple, G., Shin, Y.J., Tamura, S.Y. Clinically validated peptides as templates for de novo peptidomimetic drug design at G-protein-coupled receptors. *Curr. Opin. Pharmacol.* **3**, 530-43 (2003).

11. Clark, R.J., Craik, D.J. Native chemical ligation applied to the synthesis and bioengineering of circular peptides and proteins. *Biopolymers.* **94**, 414-422 (2010).

12. Tamaki, M., Akabori, S., Muramatsu, I. Biomimetic synthesis of Gramcidin S. Direct formation of antibiotic pentapeptide active esters having no protecting group on the side chain of the Orn residue. *JACS.* **115**, 10492-10496 (1993).

13. Chatterjee, J., Mierke, D., Kessler, H. N-methylated cyclic pentaalanine peptides as template structures. *JACS*. **47**, 15164-1516472 (2006).

14. Loosli, H. R., Kessler, H., Oschkinat, H., Weber, H. P., Petcher, T.J., Widmer, A. The conformation of cyclosporin A in the crystal and in solution. *Helv. Chim. Acta,* ***68***, 682-704 (1985).

15. Kessler, H., Kock, M., Wein, T., Gehrke, M. Reinvestigation of the Conformation of Cyclosporin A in Chloroform. *Helv. Chim. Acta,* **73**, 1818-1832 (1990).

16. Jin, L., Harrison, S. C., Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. *Proceedings of the National Academy of Sciences.* **99,** 13522-13526 (2002).

17. Ko, S. Y., Dalvit, C. Conformation of cyclosporin A in polar solvents. *Int. J. Pept. Protein Res.,* **40**, 380-2 (1992).

18. Rezai, T., Bock, J.E., Zhou, M.V., Kalyanaraman, C., Lokey, R.S., Jacobseon, M.P. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: Successful in silico prediction of the relative permeabilities of cyclic peptides. *JACS*. **128**, 14073-14080 (2006).

CHAPTER 3:

TOWARDS IDENTIFICATION OF NAÏVE IMMUNOPHILIN-BINDING PEPTIDE SCAFFOLDS

Immunophilin ligands recruit endogenous protein surfaces for affinity-enhancement

In the complex milieu of the cytoplasm, proximity matters. It matters when dimerization or multimerization of cell surface receptors occurs in response to extracellular stimuli; when enzymes are recruited to protein scaffolds setting off a chain of chemical modifications which alter protein function; when dissociation or association with a chaperone enables protein translocation to the nucleus to alter transcription.[1,2] Protein-protein interactions direct the cell's central lines of communication. So, it's no surprise that when these signaling cascades become deregulated, cells may lose control over central processes such as growth and metabolism, leading to progression of disease.

The ability to selectively bind to protein surfaces within the cell remains a long-sought-after goal with enormous therapeutic potential[3]. Unfortunately, the properties of these interfaces make this goal extremely challenging from a chemical standpoint. These interactions are characterized by large, flat contact surface areas, and often lack either an obvious hydrophobic pocket or a small molecule ligand, which could otherwise be used as a starting point for drug development **(Fig. 3-1, A-B)**[4]. Nonetheless, successful examples of small-molecule modulators of PPIs have been growing in recent years.[5,6] Alternately, the evolution of monoclonal antibodies and minimized protein domains has been remarkably successful at targeting proteins on the cell surface.[7] However, their inability to permeate cellular membranes limits most biologics to extracellular targets.

To address this challenge, the following chapter describes a preliminary work on a project that draws inspiration from the immunophilin proteins, FK506-binding protein 12 (FKBP12) and Cyclophilin A (CypA). Given that the Greek root "philin" means friend, it is no surprise that these chaperone proteins are highly promiscuous and help to guide protein folding by catalysis of proline *cis/trans* peptide bond isomerization of structurally distinct proteins. As previously noted, these proteins also bind to microbial toxins, Rapamycin, FK506, and Cyclosporine A (CsA).[8-10] Binding of these macrocyclic natural products to cognate immunophilin proteins generates a new composite surface. This enables the formation of an inhibitory trimeric complex with downstream effector proteins, resulting in host immunosuppression (**(Fig. 3-1, C-F)**. The ability to reset the specificity of an endogenous protein surface is an attractive approach that could circumvent the limiting size of cell-permeable therapeutics. This affinity-enhancing protein surface recruitment strategy has been previously validated by the inhibition of SH2 domains, amyloid beta aggregation, and RNA targeting[11-13].

The preceding chapters describe the engineering of a new CypA destabilizing domain (cDD) tool for the study of CypA ligands. We also probed the structure activity relationship of CsA analogues to gain greater insight into the mechanism of this structurally-significant molecule. In this final chapter, we describe work towards the identification of naïve peptide macrocycles that can bind to, and ultimately, reset the specificity of FKBP12 or CypA to target a disease modifying protein of interest. We expect that, our increased understanding of CsA, and use of this new cDD tool, may be useful towards the identification of evolvable, target-agnostic immunophilin-binding scaffolds.

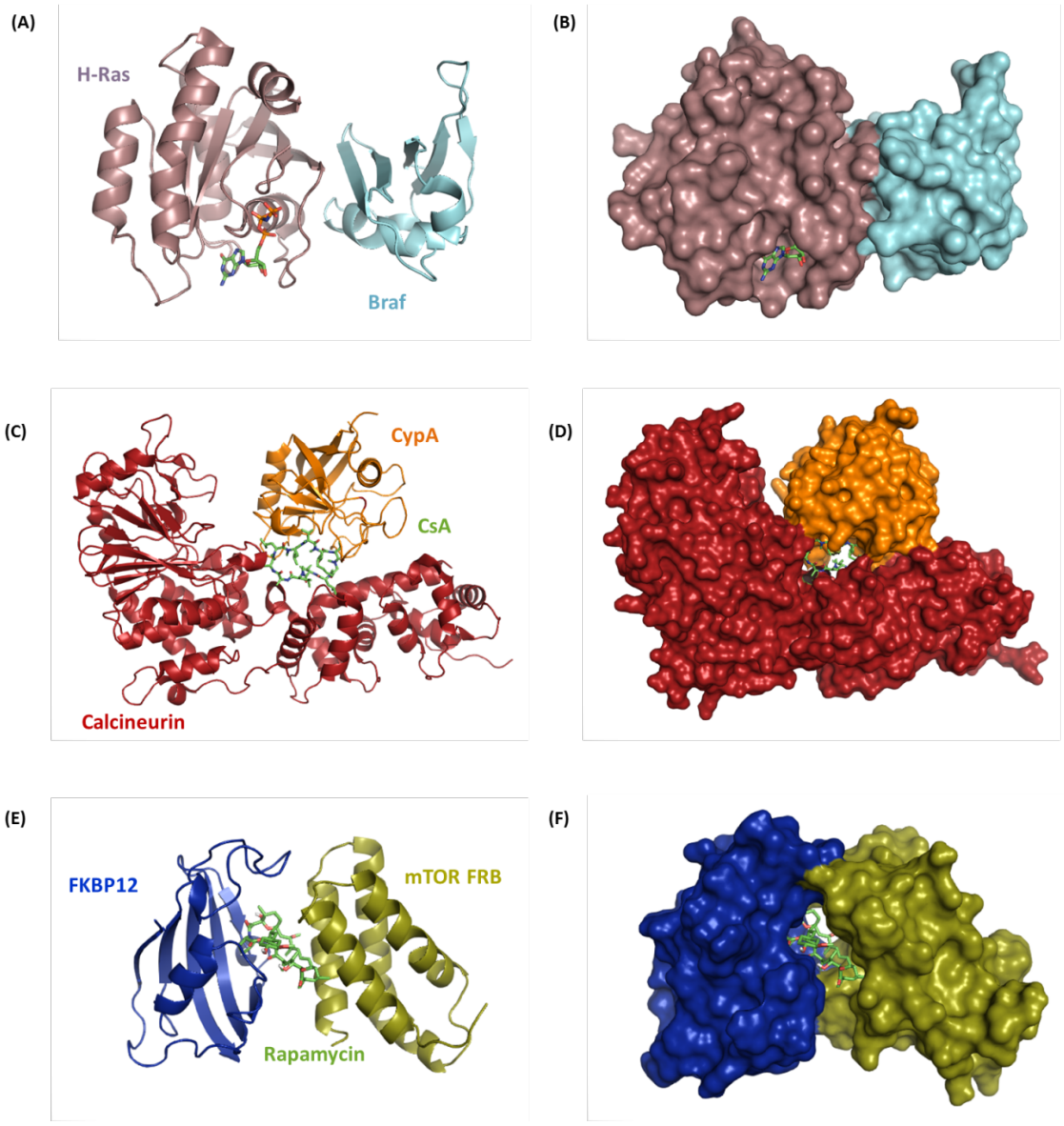


Figure 3-1. Protein-protein interactions may be mediated by chemical inducers of dimerization. (A,B) H-Ras (violet), bound to a non-hydrolyzable GTP analogue and locked into the active conformation, is bound to effector protein, Braf (cyan) [PDB ID: 4G0N]. The protein interaction is mediated by beta-sheet complementation and provides no obvious small molecule strategy for inhibition. (C, D) Interaction of CypA (orange) with calcineurin (red) is mediated by dimmerizer, CsA [PDB ID: 1MF8]. (E, F) FKBP12 (blue) and the mTOR FRB (olive) domain form a trimeric complex with dimerizer, rapamycin [PDB ID: 1FAP].

Having spent a great deal of time on the analysis of CsA structure and function, let's now turn our attention to the structure of two other well-characterized immunophilin ligands, FK506 and rapamycin **(Figure 3-2)**. Both of these molecules bind to presenting protein, FKBP12, through a conserved binding domain. However the target for inhibition is set by a variable effector domain. When this work was initially conducted, we considered these macrocyclic inducers of protein dimerization as having two distinct domains resulting in bifunctional binding capacity. For FK506 and rapamycin, this framework is valid. The synthesis of a 200-member library of rapamycin analogs (rapalogs), which include the conserved FKBP12 binding domain, demonstrated that immunophilin-binding is maintained despite widespread changes to the effector domain.[14]

As discussed in **Chapter 2**, CsA target engagement may be sensitive to small structural changes on which global conformational dynamics are dependent, even if they do not make direct contact with the protein. Therefore, as a cautionary note, *any future work towards novel immunophilin binding scaffolds should also consider conformational changes as a strong determinant in library design.*

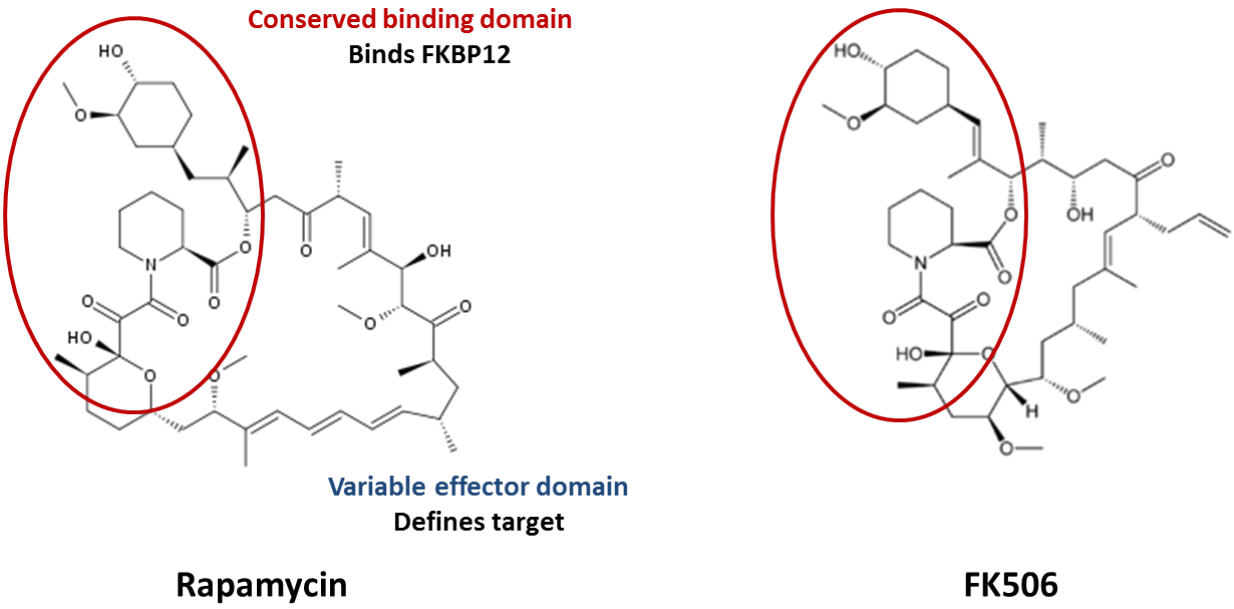


Figure 3-2. Distinct protein binding domains of rapamycin and FK506. The conserved domain (red) is shared between rapamycin, FK506, and rapalog libraries. The effector domain (blue) is variable and sets the target specificity.

A method to generate libraries of genetically-encoded cyclic peptides that employs the *Synechocystis sp PCC6803 dnaE* intein has been developed in *Escherichia coli.*[15,16] This method, termed split-intein mediated circular ligation of peptides and proteins (SICLOPPS), has been used to select inhibitors of protein-protein interactions. SICLOPPS can be combined with an *in vivo* selection to identify biologically active peptides. This method has a distinct advantage over solid-phase synthesis of peptide libraries, since genetically-encoded peptides with low target binding activity can by evolved to reach much higher affinities. It was this that motivated us to find a homodetic cyclic peptide scaffold that is able to bind immunophilin presenting proteins, either FKBP12 or CypA. We envisioned that such a

scaffold could be incorporated into a SICLOPPS selection to find dimerizers that alter the target specificity of immunophilin presenting proteins.

Design of modular immunophilin-binding cyclic peptide scaffolds

There are two elements of cyclic peptide scaffolds, size and sequence, that can be altered to increase the chances of finding a peptide with the desired function. We combed the literature to identify linear or 'loop-like' motifs known to bind either FKBP12 or CypA or both, the results of which are described in **Figure 3-3**. We identified two hexapeptide motifs, a loop from the HIV gp120 coat protein that interacts with immunophilins in blood and a phage display peptide against CypA.[17,18] We also included the concensus GP dipeptide motif, which is found in the interaction region of many protein substrates of FKBP12 and CypA.

(A)

Library	Sequence	Bias	Reason for bias
GP	**GP**N_n	CypA and FKBP12	Concensus GP motif exists in many immunophilin binding partners
FGP	**FGPDLP**N_n	CypA	Peptide motif selected by phage display against CypA
IGP	**IGPGRA**N_n	CypA and FKBP12	Contiguous peptide motif from the HIV gp120 coat protein known to interact with immunophlins in the blood

(B)

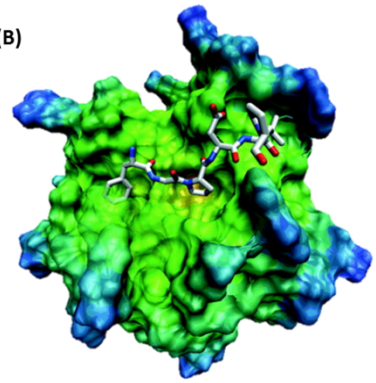


Figure 3-3. Design of potential immunophilin-binding cyclic peptide scaffolds. (A) Synthetic cyclic peptides included a linear peptide motif known to bind either FKBP12 or CypA, and a flexible poly-alanine linker. Cyclic peptide of 11-15 residues were synthesized. (B) Modeled complex of CypA with the FGPNLP phage-display peptide.[19,20]

We incorporated these motifs into peptide macrocycles by solid-phase synthesis. Cyclic peptides were synthesized by linking the side chain of a glutamate residue to solid resin and elongating the peptide chain by standard Fmoc chemistry, combined with orthogonal Alloc protection of the C-terminus, as described in **Figure 3-4**. The Alloc group is selectively removed by palladium catalysis prior to cyclization and cleavage from the resin. For residues beyond the glutamate required for cyclization, and the incorporated hexapeptide motif, we substituted a flexible poly-alanine linker, which we propose could be replaced by randomized residues for library construction.

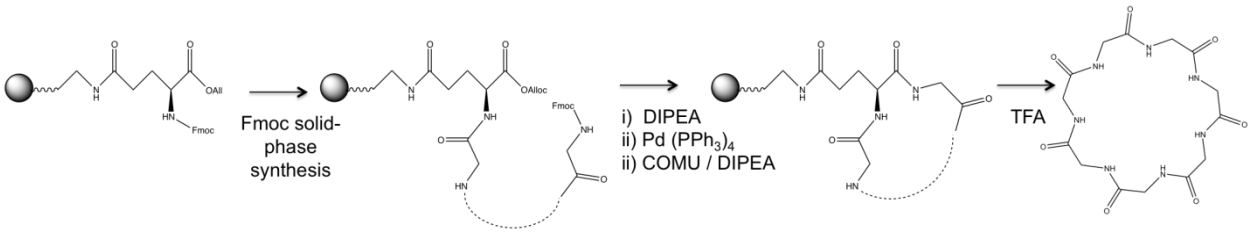


Figure 3-4. Synthetic scheme for solid-phase synthesis of cyclic peptides.

We synthesized cyclic peptides ranging from 11 to 15 residues each, incorporating one of the immunophilin-binding hexapeptide motifs. We recognized that a smaller macrocycle, which is therefore more conformationally-constrained, has a greater chance of being cell permeable, and may increase binding affinity and selectivity. However, we did not want to interfere with the binding modality of the linear hexapeptide motif to the immunophilin protein. Therefore, we identified ring cycle size as an important criterion for optimization. We also included the following control peptides: a negative control in which the required proline residue is

mutated to alanine; a linear version of the hexapeptide to evaluate the effect of conformational constraint by cyclization; and the full length linear peptide reported in the literature **(Fig. 3-5)**.

Figure 3-5. Solid-phase synthesis of peptides with immunophilin-binding motifs. The above peptides were synthesized on solid-phase where, CP indicates a cyclic peptide; Neg indicates a negative control which lacks required proline; numbers denote number of total amino acid residues; and reference refers to the full-length linear peptide sequence reported in the literature.

Evaluation of immunophilin-binding motifs by top-down proteomics, alphaLISA, and fluorescence polarization

Both peptide motifs were evaluated for their ability to bind to immunophilin proteins, FKBP12 and CypA. With help from Mat Sowa at Warp Drive Bio, LLC., a top-down proteomic approach by orbitrap mass spectrometery was used to visualize the peptide-protein complexes in real time. A summary of binding, as observed by mass spectrometry, is given in **Figure 3-6**. Notably, both peptide motifs were shown to selectively bind to FKBP12, but neither bound CypA when incorporated into a macrocycle. The strongest association under these conditions was observed between FKBP12 and the FGPDLP motif; this was surprising since this motif was generated by phage display against CypA. This is not implausible, however, since both proteins recognize similar GP-containing peptides. A legend to aid in the interpretation of mass spectrometry binding data is provided as **Figure 3-7**, and a representative data set is shown in **Figure 3-8**.

Peptide Name	Sequence	Description	Reported Kd (CypA)	Reported Kd (FKBP12)	Binds to CypA?	Binds to FKBP12?
FGP13_neg_CP	WAAA**FGADLP**AAQ	Negative control, cyclic peptide, P→A mutation	-	-	n/d	
FGP_ref	WEGE**FGPDLP**AGDS	Positive control, linear phage display peptide against CypA	1uM	-		
FGP12_lin	WAA**FGPDLP**AAQ	Linear	-	-		
FGP11_CP	WA**FGPDLP**AAQ	Cyclic	-	-	n/d	
FGP13_CP	WAAA**FGPDLP**AAQ	Cyclic	-	-		
IGP12_neg_CP	WAA**IGAGRA**AAQ	Negative control, cyclic peptide, P→A mutation	-	-	n/d	
IGP_ref	WNTRKSIH**IGPGRA**FYTTGE	Positive control, V3 loop from HIV gp120 coat protein	0.330uM	0.036uM		
IGP12_lin	WAA**IGPGRA**AAQ	Linear	-	-	n/d	
IGP11_CP	WA**IGPGRA**AAQ	Cyclic	-	-		weak
IGP13_CP	WAAA**IGPGRA**AAQ	Cyclic	-	-	n/d	weak

Binding observed; No binding observed

Figure 3-6. Summary of peptide binding as observed by real-time mass spectrometry.

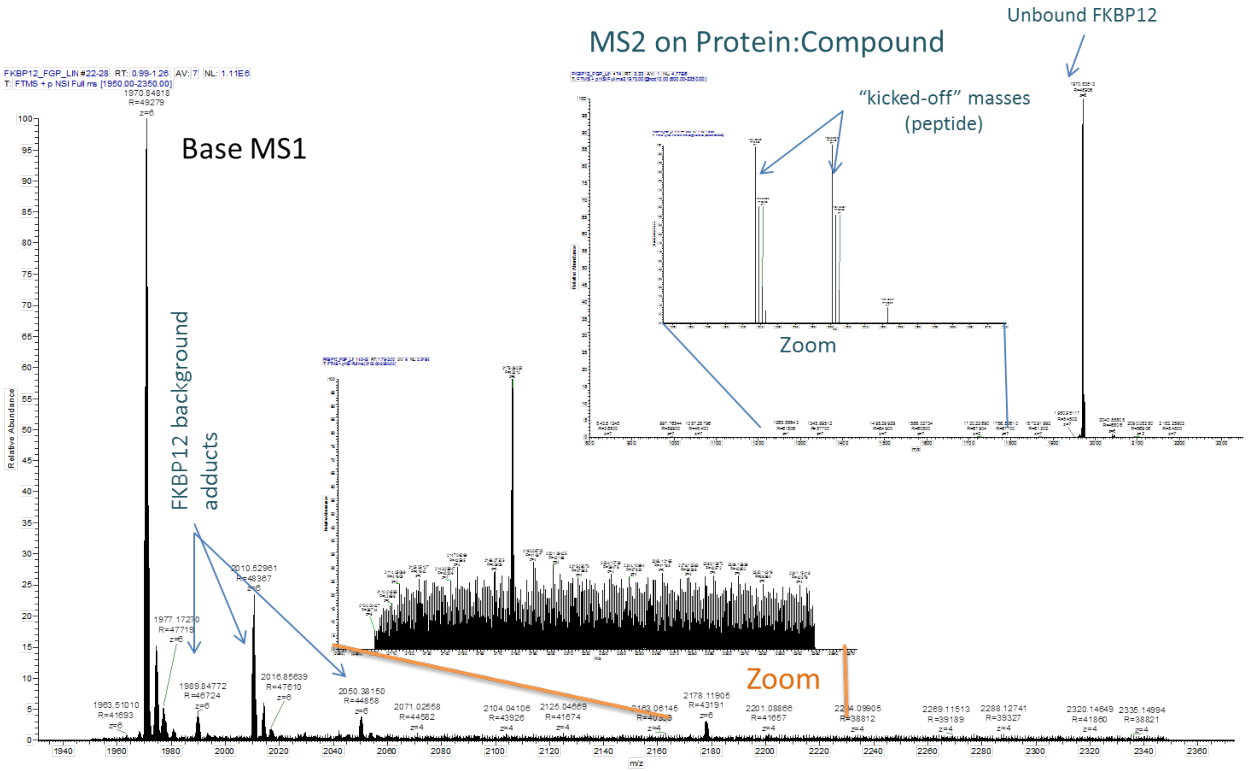


Figure 3-7. Orbitrap mass spectrometry allows for identification of peptide-protein complexes.

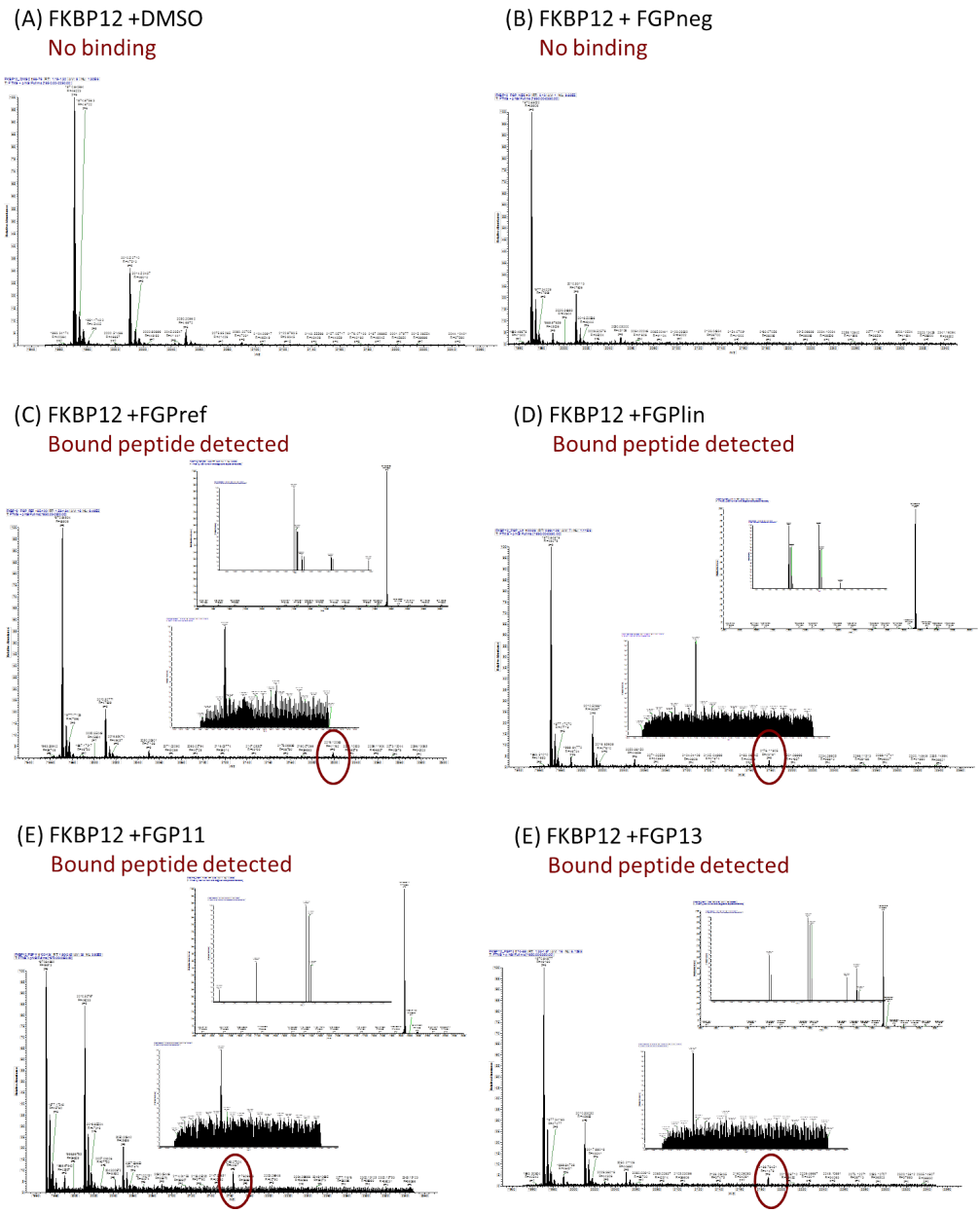


Figure 3-8. Top-down proteomics shows that FGPDLP-containing peptides selectively bind FKBP12. A representative data set from the top-down proteomics determination of complex formation. Peaks in the DMSO and FGPneg controls are non-specific FKBP12 adducts. MS2 insets in FGPref and FGP11 correctly identify the peptide-FKBP12 complexes. However, the amount of bound complex observed is only a small fraction of the free protein and free peptide.

Since the mass spectrometry experiments were performed in the gas phase, it is difficult to get an accurate assessment of the strength of the interaction. Therefore, we chose to quantify the association by two binding assays: 1) AlphaLISA, a competition-based assay, where the readout is the dissociation of either the FKBP12-rapamycin-FRB complex, and 2) a direct binding fluorescence polarization assay with fluorescently labeled peptide (**Figure 3**). Unfortunately, these assays confirmed that the interaction is very weak (>10uM), and as a result, we were unable to get an accurate K_d. Thus far, there is no evidence to indicate that these biasing motifs would be a good starting point for an immunophilin-binding peptide library.

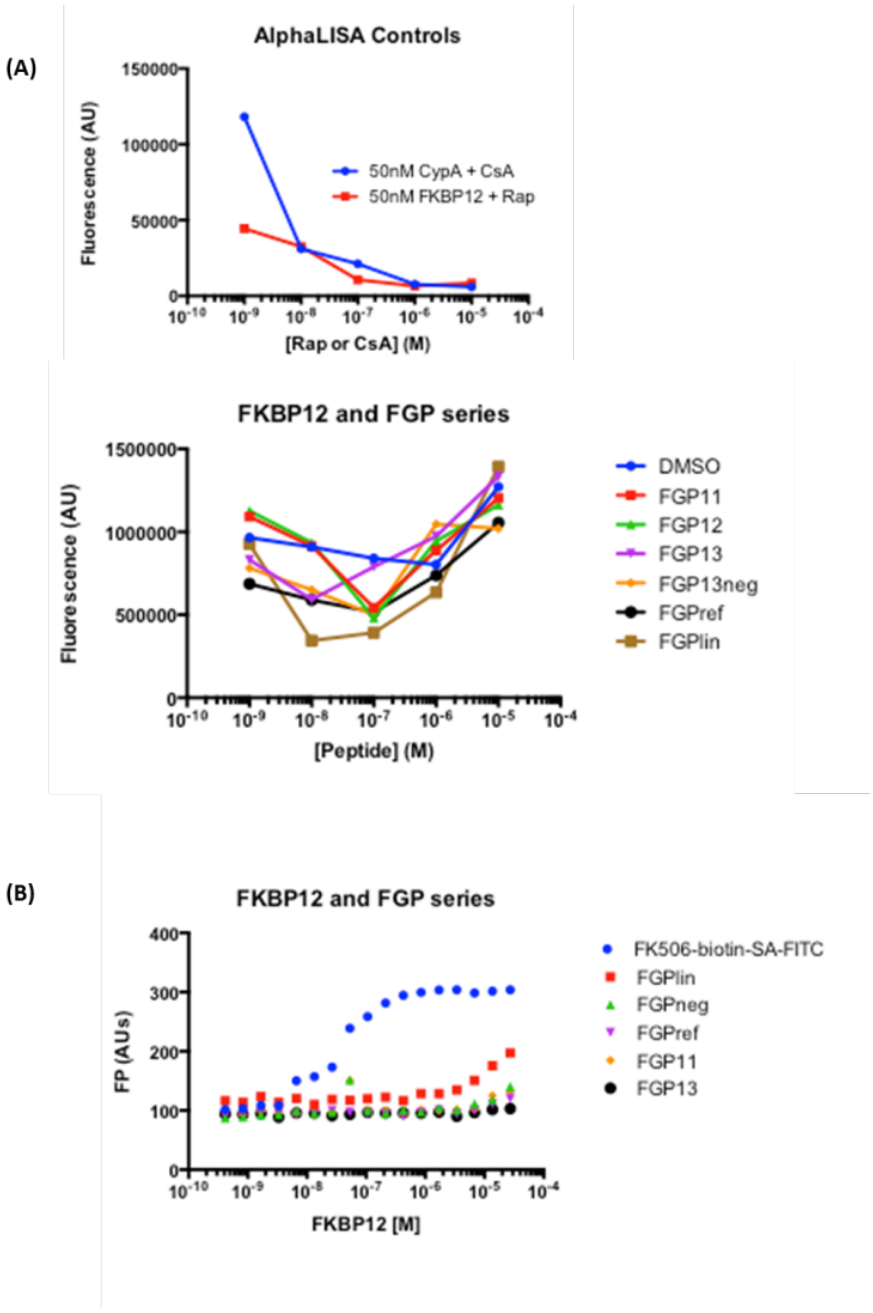


Figure 3-9. Figure 3. Binding assays confirm low affinity of protein-peptide complexes. Expected K_d of the FGP peptide-FKBP12 complexes observed by mass spectrometry is >10uM as determined by (A) AlphaLISA and (B) fluorescence polarimetry of FITC labeled peptides. IGP-FKBP12; IGP-CypA; and FGP-FKBP12 complexes showed a similarly low level of binding (results not shown).

As I final thought for this chapter, I would like to put these efforts in context with respect to the discussions from Chapters 2 and 3.

In Chapter 2, we described the engineering of a CypA destabilized domain to evaluate the binding and permeability characteristics of cyclosporine analogues. The cDD was also engineered with another purpose in mind. Using SICLOPPs expression of cyclic peptides, one could envision screening a genetically-encoded peptide library using the cDD-mCherry reporter to identify cyclic peptides that bind CypA. Structural information of such a peptide bound to CypA *might* provide information on sites tolerant to mutagenesis. A subsequent library could then be screened for the formation of new CypA-peptide trimeric complexes in either a target-directed or phenotypic (target-agnostic) manner.

In Chapter 3, however, we found that peptide macrocycles may not be very tolerant to structural modifications, even if they don't directly mediate interaction with the target protein; although, such molecules are likely more tolerant to side chain mutations than backbone modification. Therefore taking a more holistic approach to screening may increase the likelihood of success. For example, considering the cooperativity observed in the presenting protein-ligand-target interactions, one could screen for the formation of a trimeric complex without first establishing a scaffold known to bind the presenting protein. This is a more challenging screen because it seeks to find a molecule with bifunctional binding specificity, that also orients bound proteins such that the interaction is cooperative. Whether this is possible would be highly dependent on the size and quality of the peptide library. Thankfully, SICLOPPs cyclic peptide expression is well positioned to fill the role of library

generation since it is possible to create very large ($>1x10^8$) libraries. Moreover, if any hits were found, they could be evolved to select for the highest affinity interaction.

Acknowledgements

Mass spectrometry was performed by Mat Sowa at Warp Drive Bio, LLC. I am also grateful to Brian Bowman for technical advice and reagents for the AlphaLISA assay. FK506 biotin for fluorescence polarimetry was obtained from Dylan Stiles at Warp Drive Bio, LLC.

Methods

General solid-phase peptide synthesis

Unless otherwise noted, all reagents were obtained from Sigma-Aldrich or EMD Chemicals and used as received. Typical procedures for manual syntheses using polypropylene chromatography columns (Bio-Rad), solvent-resistant three-way stopcocks (Bio-Rad), and a Vac-Man vacuum manifold (Promega) connected to house vacuum and house N_2 stream were used in concert with N-a-Fluorenylmethoxycarbonyl (Fmoc) peptide solid phase synthesis (described). Peptide amides were synthesized using Rink Amide MBHA resin (100-200 mesh, loading 0.4-0.8 mmol/g) and Fmoc-protected amino acids with acid-labile side chain protecting groups when necessary. Manual coupling reactions for standard amino acids were performed with 6 equivalents of amino acid, 5.7 equivalents of 2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) and 12 equivalents of N,N-diisopropylethylamine (DIPEA). Amino acids were prepared fresh daily as 0.4M solutions in *N*-methyl-2-pyrrolidone (NMP) and stored at 4°C as 0.4M solutions and HCTU was prepared fresh daily as a 0.38M solution in NMP and stored protected from light at 4°C. DIPEA was stored at room temperature. All steps are performed with N_2 bubbling for mixing. Prior to use, Rink Amide MBHA resin was swelled in NMP for approximately 30 minutes. Amino acids were then sequentially coupled from C-terminal to N-terminal residues using the following cycle:

1) Deprotect Fmoc by 3 x 12 minute treatments with 25% (v/v) piperidine in NMP.
2) Wash deprotected resin with 5 x 1 minute treatments with NMP

3) Combine and vortex amino acid, HCTU, and DIPEA then add to washed, deprotected resin. Couple standard amino acids for approximately 45 minutes and couple β-branched/sterically hindered amino acids (His, Ile, Pro, Thr, Trp, Val) or non-natural amino acids (Fmoc-S_5-OH, Fmoc-R_8-OH) for 60 minutes or longer.

4) Wash resin with 5 x 1 minute treatments with NMP

5) Repeat to elongate peptide.

Solid-phase synthesis of cyclic peptides

Solid phase synthesis of cyclic peptides followed the standard Fmoc chemistry described above for chain elongation. However, the first residue was typically an O-alloc-carboxy-protected glutamate (Fmoc-Glu(OAll)-OH, Sigma), which was coupled to the resin through the unprotected side chain carboxyl. Although other reactive side chains could be used with different resins. The chain was elongated before eventual alloc deprotection of the C-terminus, Fmoc deprotection of the N-terminus, cyclization (optimized conditions) and cleavage according to the following steps.

O-all deprotection with Pd/PhSiH3 (Joshi lab protocol)

1) Wash resin 3x with DCM under nitrogen.
2) Add phenylsilane (25eq in degassed DCM) to resin and shake 5min under nitrogen.
3) Leaving phenylsilane in, add Tetrakis(triphenylphosphine)palladium (0.4eq in degassed DCM) under nitrogen to resin and deprotect 30min.

4) Wash 3x DCM, 3x DMF, 3x sodium diethyldithiocarbamate/DMF (2min), 3x MeOH.
5) Repeat deprotection for 2 more cycles.
6) Do a test cleavage to make sure OAll group is completely off.

Fmoc deprotection and cyclization (efficiency dependents on length and sequence)

7) Dissolve in DMF and shake for 12-24 hours with 4eq PyBOP, 4eq HOBT, and 20eq ml DIPEA

Expression and purification of FKBP12

Received FKBP constructs FKBP-SBP (streptaviding binding tag) in pET41 vector from Khian Hong Pua (see Verdine lab vector database). Transformed FKBP-SBP into BL21(DH3) competent cells and plated on LB-Kan. 4 x 1L LB-Kan were inoculated with 2ml each of overnight culture. Cells were grown with shaking to an OD_{600} of between 0.6-0.7 at 37degC and 220rpm. Protein expression was induced at a final concentration of 0.25mM IPTG and grown for an additional 4.5 hours at 30degC. Cells were spun down at 5000xg for 15min at 4degC. Pellet swere re-suspended in 10ml FKBP lysis buffer (100mM phosphate pH8, 300mM NaCl), combined, and then split into 4 x 50ml Falcon tubes. Pellets were flash-frozen in liquid nitrogen, and stored at -80degC until ready for purification.

FKBP-SBP Pellet was thawed in cold room, resuspended in 30ml total of FKBP lysis buffer with protease inhibitor cocktail. Cells were lysed by sonication (9x10sec, Power 7.0) and lysate was clarified by centrifugation at 14000rpm for 30min at 4degC. Lysate was applied to

2ml equilibrated Cobalt resin, washed 2x15ml with FKBP lysis buffer + 10mM Imidazole, and eluted in 3x5ml lysis buffer + 250mM Imidazole. The protein was either centrifuged at 16,000g or passed through a 0.2mm syringe filter to remove any particulates then gel filtration was performed on a preparative scale Superdex 75 Gel Filtration column (buffer: 50mM HEPES pH7.2, 150mM Li_2SO_4, 10% Glycerol, 0.2mM TCEP).

Fluorescence polarimetry (FP)

FP assay buffer was the same as the buffer used for gel filtration of FKBP12. Prior to setting up the assay, the protein was concentrated using a Centriprep YM-10 centrifugal concentrator if necessary and then centrifuged or filtered through a 0.2um syringe filter to remove particulates prior to quantification by A_{280}. FITC-labeled peptides were diluted to 1mM in DMSO then further diluted to 20nM in FP Assay Buffer. Serial dilutions of FKBP12 were then also prepared in FP assay buffer.

Corning black opaque 384-well plates were used for the assay. 40uL of 20nM peptide in FP Assay Buffer was added to the wells. 40uL of FKBP12 was then added to the wells containing peptide and the sample in each well was mixed by pipetting. The plate was sealed with adhesive film, covered with foil, and incubated at room temperature for 45 minutes. The fluorescence polarization was read using the SpectraMax M5 plate reader (Molecular Devices) with the following settings: excitation: 485nm, emission: 525nm, cutoff filter: 515nm, high sensitivity/100 reads per well, slow carriage speed, 500ms settling time. Plates were read twice to protect against instrument error during a read.

References

1. Strong, M. & Eisenberg, D. The protein network as a tool for finding novel drug targets. *Prog. Drug Res*. **64**, 191-215 (2007).

2. Krogan, N.J. *et al*. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637-643 (2006).

3. Wells, J.A. & McClendon, C.L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001-1009 (2007).

4. Hopkins, A.L. & Groom, C.R. The druggable genome. *Nature Rev. Drug Discov.* **1**, 727-730 (2002).

5. Arkin, M.R. & Wells, J.A., Small molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Reviews Drug Discovery.* **3**, 301-317 (2004).

6. Arkin, M.R & Wells, J.A. Small molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Rev. Drug Discov.* **3**, 301-317 (2004).

7. Breedveld, F.C. Therapeutic monoclonal antibodies. *Lancet* **355**, 735-740 (2000).

8. Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807-815 (1991).

9. Griffith, J.P. *et al.* X-ray structure of calcineurin inhibited by the immunophilin-immunosuppresant FKBP12-FK506 complex. *Cell* **82**, 507-522 (1995).

10. Choi, J. *et al.* Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239-242 (1996).

11. Briesewitz, R. *et al.* Affinity modulation of small-molecule ligands by borrowing endogenous protein surfaces. *PNAS* **96**, 1953-1958 (1999).

12. Gestwicki, J.E., Crabtree, G.R. & Graef, I.A. Harnessing chaperones to generate small-molecule inhibitors of Amyloid Beta aggregation. *Science* **306**, 865-869 (2004).

13. Plummer, K.A. *et al. In vitro* selection of RNA aptamers against a composite small molecule-protein surface. *Nucleic Acids Res.* **33**, 5602-5610 (2005).

14. Wu, X. *et al.*Creating diverse target-binding surfaces on FKBP12: Synthesis and evaluation of a rapamycin analogue library. *ACS Comb. Sci.* **13**, 486-495 (2011).

15. Benkovic, S.J. & Tavassoli, A. Split-intein mediated circular ligation used in the synthesis of cyclic peptde libraries in *E. coli*. *Nature Protocols* **2**, 1126-1133 (2007).

16. Horswill A.R. & Benkovic, S.J. Identifying small-molecule modulators of protein-protein interactions. *Curr. Prot. Protein Sci.* **19**, 1-19 (2006).

17. Piotukh, K. et al. Cyclophilin A binds to linear peptide motifs containing a consensus that is present in many human proteins. *Journ. Biol. Chem*. 280, 23668-23674 (2005).

18. Endrich, M.M. & Gehring, H. The V3 loop of human immunodefiency virus type-1 envelope protein is a high-affinity ligand for immunophilins present in human blood. *Europ. Journ. Biochem*. **252**, 441-446 (2001).

19. Horswill A.R. & Benkovic, S.J. Identifying small-molecule modulators of protein-protein interactions. *Curr. Prot. Protein Sci.* **19**, 1-19 (2006).

20. Piotukh, K. *et al.* Cyclophilin A binds to linear peptide motifs containing a consensus that is present in many human proteins. *Journ. Biol. Chem.* **280**, 23668-23674 (2005).

CONCLUSION

In recent years there has been a noticeable increase in the molecular complexity of therapeutic targets, including challenging protein-protein interactions. This has coincided with a revived interest in the development of both *de novo* and naturally-derived macrocyclic inhibitors. Preorganization of binding elements in the course of a biologically relevant protein-ligand interaction has been the topic of particularly intense efforts over the past decade.

Among different classes of macrocycles, cyclic peptides and peptidomimetics have received a large share of attention in drug discovery. This is likely explained by the existence of synthetic and biological methods to rapidly assemble the constituent amino acid building blocks. There are also other reasons for which peptide-based molecules have received so much attention. Cyclic peptides offer a predictable increase in affinity over their linear counterparts owing to inherent structural preorganization, which decreases the entropic penalty of target binding. Additionally, cyclization addresses a central liability of linear peptides – their propensity to undergo rapid proteolytic degradation in cells – since cyclic peptides are more resistant proteolysis by peptidases.

Based on these properties, some have suggested that cyclic peptides and other macrocycles represent a privileged class of molecules for therapeutic intervention. Unfortunately, the large polar surface areas that accompany high amide content typically come at the expense of cellular permeability. This often limits the bioavailability of peptide macrocycles. As a result, one big challenge facing the field is how to devise effective synthetic guidelines that consider both target engagement and retention of drug-like properties.

To gain insight into the principles that guide such drug discovery efforts, we have revisited the canonical 'rule-breaking' molecule, CsA. This molecule is known for its relatively good passive membrane permeability and oral bioavailability, resulting from a network of intramolecular hydrogen bonds that are presumed to form while CsA passes through the lipid bilayer. A molecular-level analysis of the interaction between CsA and its cellular target, CypA, paints a different conformational picture, in which the amide groups are involved in target recognition and thus relinquish the intramolecular interactions that contribute to its membrane permeability. Many researchers have been taking notes from this example to guide their discovery of large therapeutic agents, the goal being to design molecules with favorable drug-like properties by minimizing their effective polar surface areas.

One such strategy has been the use of N-methylation, which can enhance cellular permeability by masking polar amide groups. Our work, when considered with the work of others before us, suggests that N-methylation has consequences, which may be favourable or not depending on the context, for the global conformational dynamics of flexible peptide macrocycles.

In efforts to come up with new synthetic tools aimed at peptide macrocycles, we need to be aware of the propensity of macrocycles to form transannular interactions. We found that N-methylation of CsA is not just important for cellular permeability, but also, surprisingly, CypA target engagement. We suspect that this is due to rearrangement of the preferred transannular hyrdogen bonding pattern. Therefore, medicinal chemistry efforts that rely on N-methylation must also consider the interplay of cell permeability and target binding. These

interactions can deliver dividends in areas that require conformational constraint, but attention must be paid to the potentially detrimental consequences of particularly strong transannular interactions that can lead to unanticipated intramolecular reactivity and molecular "collapse". The availability of broadly applicable methods that address these long-standing goals will further facilitate synthesis-driven improvement of macrocyclic lead molecules in drug discovery.