# Advancing the Development of Macrocyclic Peptide Therapeutics Using High-Throughput Mass Spectrometry Approaches

Jonathan Palmer

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Miklos Guttman, Co-Chair
Gaurav Bhardwaj, Co-Chair
Abhinav Nath

Program Authorized to Offer Degree:

Medicinal Chemistry

Jonathan Palmer

University of Washington

## Abstract

Advancing the Development of Macrocyclic Peptide Therapeutics Using High-Throughput Mass Spectrometry Approaches

Jonathan Palmer

Chair of the Supervisory Committee:
Miklos Guttman
Gaurav Bhardwaj
Department of Medicinal Chemistry

Cyclic peptides are poised to revolutionize the therapeutic market as designed protein modulators. These peptides are characterized by a circular sequence of amino acids, which creates unique structural properties. Their potential as drugs is supported by existing natural products and their synthetic analogues, like vancomycin and cyclosporine[2–4]. The proven success of these therapeutics has stimulated interest in the discovery of new cyclic peptide drugs using advanced methods. Historically, the discovery of cyclic peptide therapeutics has relied on exploring natural product analogues and random high throughput screening[2,5,6]. However, these methods can be

improved to test a more diverse set of cyclic peptide structures and protein targets. Recent advancements in computational protein prediction and design have made *in silico* screening a powerful tool for drug discovery, boosted by improvements in both physics-based and deep learning approaches. Integrating computational tools with enhanced drug discovery techniques allows for the identification of novel cyclic peptide therapeutics[7,8]. This strategy represents a shift from random screening to a focused approach that can uncover new structures and interactions.

High-throughput screening techniques offer an efficient way to evaluate thousands of compounds for target interaction simultaneously. To match the speed of computationally identified peptide candidates, high-throughput screening is essential for cyclic peptides. One such technique, affinity selection mass spectrometry (ASMS), leverages the affinity between protein and peptide to isolate the peptide from a pool of binding candidates[9]. The isolated peptides are identified by characteristic fragmentation inside the mass spectrometer. However, for cyclic peptides, this characteristic fragmentation can be complex and difficult to interpret. To improve interpretation, we expanded mass spectrometry data collection with multistage $MS^n$ fragmentation and developed computational sequencing to read cyclic peptide fragmentation. This approach enables the computational assignment of unknown cyclic peptide peptide sequences, unlocking the ability to test cyclic peptides for target interaction in high throughput. Furthermore, we adapted ASMS to screen against protein targets in complex environments, including membrane-bound proteins like G-protein coupled receptors (GPCRs). Using virus-like particles to anchor GPCRs in their membranes helped preserve their native environment and allowed affinity selection against properly

folded proteins. Adapting ASMS in this way opens new possibilities for discovering cyclic peptide drugs against membrane-bound proteins—a longstanding barrier in the field. Beyond discovery, high-throughput mass spectrometry approaches like HDX-MS (Hydrogen-Deuterium Exchange Mass Spectrometry) allow for the study of structural dynamics in cyclic peptides. HDX-MS is a benchtop labeling experiment that, when paired with mass spectrometry, provides insights on conformational dynamics and solvent accessibility. Applying HDX-MS to cyclic peptide pools enabled us to gain structural insights into dozens of peptides simultaneously. The data generated revealed relationships between conformational flexibility and permeability, helping identify peptide scaffolds for designing membrane-traversing peptides. Overall, these high throughput approaches significantly enhance our capability to identify and assess cyclic peptides as therapeutic candidates.

# TABLE OF CONTENTS

# ACKNOWLEDGEMENTS

Working towards the completion of a thesis dissertation is, at all times, a dichotomy. There is an ever-present tension between being a master of your craft and an unlearned student. An all-consumed researcher and a dedicated family member. Of a person climbing a mountain and the mountain growing before you. Of fraternizing community, and singular isolation. The culmination of years of work is nothing to be scoffed at, and nothing to be enjoyed.

Amongst this tension there has been a constant: the never-ceasing support of those around me. I could not have survived without the gifts of laughter, love, and luxury that you have provided. I like to think that I carry with me pieces of my role models that I have stolen for myself, and there has been no better source of role models than my family.

First and foremost, to my parents. To my mom, the best role model I could ever ask for. Your hard work and long nights enabled me to achieve this accomplishment and I am forever in awe of you. Thank you for making me sit at the dining room table until I could write my Rs correctly. Thank you for teaching me that if something is within my control, I can control it. Thanks for wiping my butt, too. Your lessons have been so so important. To my stepdad Jeff, who taught me to bite the onion first thing in the morning. The ways that you have supported me are longer than the list of errands on a Sunday morning, even though it's supposed to be just a quick run to Home Depot.

To my dad Shawn, thank you for being a constant fountain of support. Your openness to express belief in me has carried me through times of uncertainty and self-doubt, and I'm not sure you've ever known the importance of your words. I have never taken for granted the encouragement you have given me to pursue my dreams, even though they took me to Wisconsin, and then to Washington. I'm never too far away. To my stepmom Susan, thank you for always bringing me joy and positivity. You have been a force of can-do attitude that I could not operate without, and I can't wait to pick up some arts & crafts from you.

To my siblings Nora and Aidan, the world is yours. From coming home so tired that all I could play was teddy bears, to helping you with homework or soccer, to celebrating your high school graduation-- being a part of your story has been an incredible honor. I have wanted nothing more than to be a good role model, support system, and first-child-trailblazer for you two. Have fun in college, make mistakes, and learn from them.

To my grandparents: I am blessed to have so much wisdom, care, and laughter in my life. To Kris and Craig: you have been such strong influences in my life. You have taught me the importance of treating everyone with patience, love, and stewardship. You have shown me what it is to find joy in life and in work, and to never stop singing your own song. To Tom and Carol: thank you for giving me your best. For filling my life with piano and checkers and soccer in the backyard. For giving me a love of cooking, and for nurturing it with recipes and fancy imported oils. You have given so much to my life, and I am better off for it. To Norm and Earlene: thank you for supporting me through every stage. You welcomed me with open arms into your family and made sure I fit right into whatever team sport was going on in your backyard. Thank you for giving me confidence in myself and imparting wisdom and trivia.

To the rest of my family, thank you for the pieces of yourself that you've given me: The joy that my cousins bring me and the work ethic that my aunts and uncles possess. Shout out to Paul, Laurie, and Brian, who have done more than can reasonably be asked to fill my years with role models and joy.

To my mentors: Mike and Gaurav. Mike, you have had such a hands-on impact on my scientific development that I would not know half the mass spec tricks without you, but I wouldn't have broken so many mass specs, either. Your ability as a scientific mentor is only matched by your care for your students, and if you're not careful half the department will be co-mentored by you. Gaurav, thank you for instilling in me a raw excitement for the possibilities of science. The eagerness with which you attack new problems is infectious and I will carry it with me always. Thank you for always pushing me

to be the best scientist possible. Your two mentor styles have blended to give me an incredible pairing of scientific rigor and personal support.

To my labmates: I have had such a wide network of support. Moreso than any other source, this network is responsible for my training as a scientist and coworker (be it for better or worse). To Mike Watson, Clint, and Taylor, thank you guys for trailblazing. For teaching me how to hold a pipette, and for listening to me vent when the robot left the sample tray open overnight. To Gizem and Stephen, we are a unique cohort for setting up a lab during a pandemic. Your joy in lab and inquisitive nature has been a constant reinforcement of why I pursued graduate school. To Ellie, Charlie, and Alesi, who have taken up the mantle of keeping me sane over the recent months (years). Thank you for putting up with my messy lab bench and messier crumbs from honey dijon Kettle chips. To the newer wave of students: Christi, Isabelle, Sarah, Joel. If I have one regret about graduating it is that I will not be able to help in your training and watch you grow as scientists. Alas, you are in good hands, To Chimeh, thanks for always coming ready to problem-solve, as it was unclear if everything would work perfectly or if we would have to reinvent the wheel on any given day.

To Alex Wiley: how cool is it that we were able to turn our dream of collaborating into a reality?!? Even cooler, we didn't kill each other. Long days of working with magnetic beads turned from dreadful to kinda-less-dreadful. I thought (and still think) that medicinal chemistry is sadistic for locking you in a closet/office with your fellow-first years through the firey crucible of first year academic work, but man does that trauma bond go hard. In our first year I was terrified of the start of the summer semester when we would go our separate ways in academic labs. Imagine my joy when we became friend-friends, not just work-friends. Since then, we have survived global pandemics, hikes that are too long, and forged a weird cult-y friend group in Seattle. You have been the lone person who can truly appreciate what I'm feeling in both my work and personal life, and your presence has made Seattle feel like home. and also thanks for teaching me that bruschetta is not just a tomatoes and balsamic thing?

To my friends: man, what a wonderful source of joy you are. I feel incredibly lucky to have a large group of friends from all phases of my life. To the Hopkins guys: when we all left for college, I was terrified things would change. Now here we are, 20 years into knowing each other and spread across the country and I still laugh daily because of you. To Clary R: I would like to acknowledge that the first funding source I had for grad school was venmo transfers from you guys. I'm pretty sure Nick venomed the most, but it's OK because Jere and Dakotah, you guys fly in to visit tomorrow. To Sarah and Courtney and Alex, I could not have made it through either the biochemistry major or the collegiate soccer season without you. Courtney, I'm still waiting on my award for getting a better grade than you in Evolution. To the Groundhogs: do we call it peer pressure, or do we call it work-life balance? Either way, my life is now full of incredible hikes and raucous happy hours and food events that we spend waayyy too much time on thanks to you.

To my best friend Nick: it turns out our collegiate adventures as roommates were just a prelude to our time in Seattle. Moreso than almost anyone else, you see me without any filter. Tired and hungry after long work nights, sleepy and lazy late at night, and (recently) irritated with Microsoft Word formatting. I could not have handcrafted a better friend who will play co-op video games with me because I'm a poor loser, host ridiculous parties with me, and eat whatever I feel like cooking that night (even if it has dairy).

To my partner Riley: the world spins and the sky tilts and yet I am planted in my feet, for having you next to me keeps life sane. You are the reason I work hard to get home at night, and I will never take for granted the sacrifices you made alongside me in the pursuit of this degree. The support and understanding you have given me over the course of this graduate program has only intensified in the last few months, and I feel it never so acutely as now, writing this at 2 AM, with you asleep on my chest, our feet intertwined. Wanna get breakfast tomorrow?

# DEDICATION

This work is dedicated to my aunt, Chrissy Palmer
Let the human impact of better medicine and health not go unnoticed,
as people are behind percentages.

# CHAPTER 1.

# Introduction to Cyclic Peptides and Their Utility and Challenges as a Therapeutic Modality

Portions of this work are reproduced with permission from:

Ramelot, T. A.; **Palmer, J**.; Montelione, G. T.; Bhardwaj, G. Cell-Permeable Chameleonic Peptides: Exploiting Conformational Dynamics in *de Novo* Cyclic Peptide Design. *Curr. Opin. Struct. Biol.* **2023**, *80*, 102603. https://doi.org/10.1016/j.sbi.2023.102603.[1]

## 1.1 Development of Cyclic Peptide Structure for Function

Amino acids, the building blocks of proteins and peptides, dictate structural and functional properties. Through both the individual chemical propensities of an amino acid and the concerted structural composition of strings of amino acids, amino acid sequence drives structure and function. Although amino acids share a common repeating backbone structure, they possess unique side chains that influence their bond angles and energetics. These structural preferences manifest as three broad secondary structure classes: helices, sheets, and loops. Helices and sheets represent stable, energetically favored structures, while loops are more flexible and less structured,

allowing for more conformational movement[10]. Peptides, which are shorter chains of 2 to 50 amino acids, have various levels of structure: primary (sequence), secondary (sub-structures like helices), tertiary (overall 3D shape), and quaternary (complexes of multiple peptide chains). Understanding each level of structure is crucial for predicting how a peptide might fold and interact with proteins, but all structural levels derive from the primary sequence of amino acids.

Cyclic peptides are a subclass of peptides characterized by a circular sequence of amino acids, typically formed either through N-to-C terminal cyclization or sidechain-sidechain bridging. This cyclization imposes structural constraints, favoring loop-like structures that are inherently flexible and challenging to analyze. Recent advancements in computational modeling have significantly improved our ability to predict loop-like cyclic peptide structures[7]. This capability led to the exploration of cyclic peptide structural motifs that promoted accuracy and stability during design. In 2016, computational cyclic peptide design successfully identified peptides with designed states that closely matched experimentally solved crystal structures[11]. By incorporating specific structural constraints, such as internal hydrogen bonds or non-canonical amino acids, researchers can design cyclic peptides with enhanced stability and functionality[11–13].

Computational design can access a structural space using noncanonical amino acids that is inaccessible to biological peptides, allowing for unique structures and interactions[12]. These studies identified design rules that limit the conformational flexibility of cyclic peptides, enabling the creation of more stable structures[13]. Design of accurate structures unlocks cyclic peptides for directed application.

## 1.2 Design of Cyclic Peptides for Protein Interaction

Cyclic peptides offer several advantages for interacting with protein targets. Their larger size allows them to interact across diffuse binding regions, while their variability enables high target specificity. These properties make them particularly effective for targeting "undruggable" proteins and disrupting protein-protein interactions, a capability that small molecule drugs often lack[14,15]. Compared to linear peptides, cyclic peptides have an energetic advantage, as they are pre-arranged in a binding-competent structure, reducing the entropic penalty associated with binding[4,16].

The ability to design cyclic peptides with specific interactions with protein surfaces advances their therapeutic potential. Early successes in this field, such as peptides designed to disrupt the MDM2-p53 interaction, demonstrated the power of computational design in developing peptide therapeutics. Modelling the p53 peptide allowed the development of a series of peptides that disrupted the p53-MDM2 interaction, using either p53-like helical peptides or alternative beta-sheet structures[17–19]. This approach has since evolved, incorporating noncanonical amino acids and stabilizing modifications to enhance peptide binding and stability[20].

Modern peptide design continues to build on these foundations. Two papers from 2021 demonstrated the importance of building structurally stable cyclic peptides around a known binding interaction to create cyclic peptides with strong target affinity[21,22]. One study developed a cyclic peptide binder against New Delhi metallo-β-lactamase 1 and observed that structural stability was important for target interaction[21]. A second study

developed cyclic peptides that inhibit the function of histone deacetylases 2 and 6 using structure design around a known noncanonical amino acid with target interaction, demonstrating the importance of both replicating existing known interactions and utilizing noncanonical amino acids[22]. Both of these studies employed a cyclic peptide design technique known as generalized kinematic closure, which uses Monte Carlo-based sampling of peptide conformation perturbations to cyclize and predict structure in the computational modeling program Rosetta. This technique was popularized as the main design strategy for cyclic peptides in Rosetta and has only recently been integrated and replaced by deep learning programs.

## 1.3 Experimentally Testing Protein Interaction of Designed Cyclic Peptides

Experimentally testing designed cyclic peptides for protein interaction can be challenging. Gold standard methods like surface plasmon resonance (SPR), biolayer interferometry (BLI), and isothermal calorimetry (ITC) provide accurate binding affinity measurements but are time-consuming and require significant amounts of protein[20]. The desire to test many peptides quickly for drug discovery demands a higher throughput technique. Library display techniques like yeast and phage display can be paired with sorting techniques like fluorescence-activated cell sorting to rapidly evaluate thousands of compounds. Display techniques are characterized by their appropriation of cellular machinery to produce peptides from coded DNA or RNA[23–26]. Display techniques excel at large diversity libraries but have limited chemical control of the constituent peptides. Nonetheless, customized display techniques like PeptiDream have reached high levels

of success identifying peptides with target interaction due to the large number of peptides that can be screened and the identification ability of a DNA/RNA tag[27–29]. The key limitation involved in display techniques is the peptide production machinery, which is restricted to biological cyclization strategies like disulfide cyclization. This inability to produce N-to-C cyclized peptides restricts what kind of peptides can be designed computationally and tested with display techniques.

To overcome these limitations, designed cyclic peptides are synthesized chemically using solid-phase peptide synthesis (SPPS). This robust synthesis method is well established, standard, and allows for the generation of large peptide libraries with complete chemical control and composition that can be screened for target interaction[30,31]. SPPS facilitates integrating high-throughput screening with computationally designed peptides (**Figure 1A**). This targeted approach to high throughput screening not only improves the efficiency of drug discovery but also refines the computational models, creating a synergistic feedback loop that enhances both design and testing.

In addition to drug discovery, SPPS and high throughput screening can be employed to optimize lead candidates. For example, deep learning methods have shown remarkable success in peptide design but cannot incorporate nonstandard (“noncanonical”) amino acids which can enhance the pharmacokinetic profiles of peptides[32]. By generating libraries of noncanonicalized analogues, SPPS and high-throughput screening together can optimize drug-like properties of deep learning identified lead candidates. These strategies pair well with high throughput screening techniques that utilize direct identification of peptide binders.

## 1.4 Introduction to Affinity Selection Mass Spectrometry (ASMS)

Affinity selection mass spectrometry (ASMS) is a powerful high-throughput screening technique that identifies peptide-protein interactions by retaining binding peptides for mass spectrometry analysis[9] (**Figure 1B**). Unlike display methods, ASMS utilizes chemically synthesized SPPS libraries and has no cyclization or amino acid limitations. ASMS involves tagging protein targets to an insoluble support bead, allowing selective retention of the protein throughout the screening process. The retained peptides are then eluted and identified via mass spectrometry-based sequencing.

Affinity selection is a proven and versatile technique which has been used to identify linear peptides from libraries as large as $10^7$ compounds[33]. Standard affinity selection approaches use magnetic bead tagged proteins and rigorous washing, but protocol customization can be optimized for a variety of applications. Applied ASMS studies have used robotic prepping and handling to standardize procedures and navigated complex buffers[34,35]. Alternative strategies for affinity selection and binder retention like size exclusion chromatography or pulsed filtration utilize different assays to retain protein and select protein-interacting peptides but still lead to direct MS identification[36]. This versatility makes ASMS a customizable tool for peptide discovery, particularly when combined with follow-up complementary techniques like BLI to validate peptide-protein interactions.

Mass spectrometry is well-established for identifying linear peptides[37,38], but cyclic peptides present unique challenges due to their cyclic structure. Tandem fragmentation

of linear peptides for sequence identification applies energy to break peptide bonds and generate characteristic ions in the resulting mass spectra[37]. Cyclic peptides have a more stable structure and lack charged termini, requiring greater energy to fragment. In combination with greater energy, cyclic peptides require two peptide bond breaking events to generate fragment ions. The first peptide bond breakage linearizes the cyclic peptide, and this event can happen throughout the peptide and generate many different linearized fragments. The second peptide bond breakage will generate characteristic fragments for sequencing (**Figure 2**). The presence of multiple linearized ions fragmenting within a single spectrum complicates interpretation of tandem MS data for cyclic peptide sequencing. This additional complexity both illustrates the need for computational assistance in cyclic peptide sequencing and reduces the accuracy of existing linear peptide sequencing programs.

## 1.5 Computational Peptide Sequencing

Computational tools for sequencing linear peptides are effective and fast. In addition to ASMS, linear peptide sequencing is widely applicable to the field of proteomics to identify peptides and proteins. Contributions from this field have driven accurate database matching, de novo, and deep learning sequencing programs[39,40]. Comet, part of the Trans-Proteomic Pipeline, is a particularly customizable and robust database matching program. Comet scores similarity between theoretical spectra and experimental spectra to identify peptides[37,41,42]. Its success in rapid identification of tryptic

and non-tryptic peptides in proteomics opens the door for further application, and its function as a database matching program is ideal for ASMS design[43].

Attempts to develop cyclic peptide sequencing have been sporadic but successful. Naturally occurring antibiotic cyclic peptides like vancomycin and daptomycin sparked development into the field of natural product peptidogenomics to identify more success stories[44,45]. The programs CycloBranch, CYCLONE, and NRPQuest are all examples of nonribosomal cyclic peptide sequencing programs that integrate nonribosomal building blocks and genomic data to discover new peptides; in the case of NRPQuest this information is used to filter mass spectra directly[46–48]. These programs were intended for discovery of natural product peptides and have poor translation to high throughput screening.

Previous attempts geared towards cyclic peptide sequencing for high throughput screening resulted in the program CycLS, which attempts to match several linearized cyclic peptide fragment ions to a single spectrum[49]. This tool showed accuracy in whole library sequencing but is limited to tandem MS/MS interpretation and no follow up studies have demonstrated the efficacy of this program in an ASMS-like application. Despite the success of these programs, a cyclic peptide sequencing tool for ASMS is needed in order to truly unlock high throughput screening for designed synthetic cyclic peptides.

## 1.6 Introduction to G-Protein Coupled Receptors (GPCRs)

ASMS offers an alternative drug discovery assay for difficult protein targets, which makes it a very valuable and adaptable tool. There is no better example than G-protein coupled receptors (GPCRs), which are notoriously difficult to study because of their dynamic nature and membrane association[50]. Traditional drug discovery screening methods against GPCRs mutate the GPCR to make it soluble and stable, employ model membranes like nanodiscs and micelles, or utilize complex signaling assays[51]. Any route requires complex assay development and significant optimization before screening. Despite this, interest in GPCR drug discovery has remained due to the varied and important biological roles GPCRs play. ASMS offers a customizable alternative that can be tuned to GPCRs.

To stabilize the structure of GPCRs in their native state, virus-like particles (VLPs) can be expressed which display GPCR proteins on its surface. VLPs hijack the functional Gag protein of viruses to create self-assembling structures that mimic the organization of native cell membranes[52]. These particles have found recent use in vaccine development and have benefitted from the resulting interest and standardization[53–55]. By embedding GPCRs in VLPs, we can more accurately study their interactions with cyclic peptides using ASMS, opening new avenues for discovering GPCR-targeting therapeutics. Cyclic peptides offer a promising solution due to their size and structural flexibility, allowing for specific interactions with GPCRs without the permeability concerns that affect small molecules.

The GPCR apelin receptor (APJ) is a high value biological target due to its involvement in signaling pathways related to blood pressure regulation[56]. APJ's unique dual peptide ligands, apelin and elabela, bind to the receptor and trigger distinct biological responses[57,58]. These factors make APJ an ideal candidate to develop GPCR ASMS for peptide binding. Understanding these interactions provides valuable insights for designing cyclic peptides that can modulate GPCR activity in a therapeutic context.

## 1.7 Therapeutic Potential of Flexible Peptides

While GPCRs are accessible targets for cyclic peptides, many other proteins of therapeutic interest are located inside cells, necessitating peptides that can cross cell membranes. Cyclic peptides often face challenges related to permeability due to their size and structural complexity, which place them in the "beyond rule of 5" (bRo5) space. Lipinski's original Rule of 5[59] provides guidelines for drug-like compounds, but cyclic peptides frequently exceed these limits while still exhibiting membrane permeability through structural adaptations[60].

Cyclic peptide therapeutics have seen success in two particular therapeutic applications: hormone drugs and infectious disease treatment. These applications are highlighted by existing natural products and the success of new therapeutic treatments is largely driven by analogues and synthetic modifications. Oxytocin and vasopressin in particular serve as two cyclic peptide hormones that were co-opted as therapeutic treatments after synthesis and administration became viable[3,61]. Vasopressin is an example of a natural cyclic peptide that was optimized as a pharmaceutical under the

name desmopressin. Desmopressin stands out as one of the few orally bioavailable cyclic peptides[62]. Infectious disease cyclic peptides are largely derived from natural products found in microorganisms themselves. Vancomycin and its derivatives are some of the most popular cyclic peptides available for antibiotics, and antifungal cyclic peptides like caspofungin are common yeast infection treatments[2,61]. Both of these applications benefit from targeting extracellular proteins and do not need to be permeable.

Endogenous and natural product therapeutics serve as templates for permeable cyclic peptide development. The endogenous flexible peptide, cyclosporine A (CSA) exhibits a remarkable ability to switch between membrane traversing and binding competent conformational states[63] (**Figure 3A**). CSA's dual structural properties make it an ideal example of a cyclic peptide that can possess protein interaction as well as membrane permeability and bioavailability[63–65]. This conformational flexibility, driven by a single proline isomerization, causes a binding competent conformation with exposed polar sites to become much more hydrophobic by creating a network of internal hydrogen bonds. Designing peptides with similar state-switching capabilities could revolutionize therapeutic strategies, though challenges remain in understanding the mechanisms underlying these conformational changes.

Aside from cyclosporine A, there are very few known cyclic peptides which exhibit conformational flexibility between two stable structures. Structural studies to induce conformational flexibility have met limited success. Cyclosporine mutational studies have increased structural stability but have had limited translation to creating new flexible peptides[66]. A 2021 study identified three new conformationally flexible

peptides through extensive computational sampling, including initial design of cyclic peptides and prediction of energetically favored conformations[67] (**Figure 3B & 3C**). These peptides were found to retain membrane permeability using *in vitro* assays despite the presence of a conformation with exposed polar surface area (**Figure 3D**). These peptides were confirmed to have conformational flexibility using NMR and showed a dependence on the dielectric coefficient of the solvent (**Figure 3E**).

Computational prediction of conformational landscapes can suggest flexiblility , but verification requires experimental evaluation[1]. Recent studies have attempted to measure and identify cyclic peptides with flexibility using techniques like liquid chromatography and suggested that flexibility should be reported as a standard metric for predicting permeability[68,69]. Continued research into these properties will be crucial for the successful design of new flexible peptide therapeutics.

## 1.8 Introduction to Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS)

Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) is a technique used to study the solvent accessibility of backbone amide protons. In a typical HDX-MS experiment, deuterium replaces hydrogen atoms in solvent exposed positions, and the number of incorporated deuterium atoms can be measured using mass spectrometry. It is traditionally used to measure the dynamic movement of proteins but has widespread applications. Development of HDX-MS has led to the automation and standardization of experimental protocols, data collection, and data processing[70,71]. Hydrogen deuterium

exchange is performed at room temperature under physiologically relevant conditions before being quenched and flash frozen in supercooled ethanol. To prevent back exchange from deuterated positions with environment hydrogen, samples are analyzed immediately or kept very cold from exchange up until data collection. Automation of this process has led to a robotic handling, thawing, and injection system that is programmable and can be paired with various mass spectrometers[70].

Data collection and analysis is performed by measuring the shift in centroided mass profiles for deuterated peptides. These shifts can be plotted over time to obtain a rate of deuterium uptake. Deuterium exchange has a known theoretical exchange rate for each amino acid dependent on its solution conditions and neighboring amino acids. This theoretical rate can be compared to the experimentally observed rate and the difference is reported as a protection factor[72–74]. Taken together, this method allows for the rapid evaluation conformational dynamics.

Applying HDX-MS to cyclic peptides measures deuterium incorporation over time and can give insight into the strength of internal hydrogen bonds and structural stability (**Figure 4**). This method is particularly valuable for evaluating the structural dynamics of peptides, providing insights into their conformational flexibility. Previous studies have used NMR to measure the importance of optimized internal H-bonds for therapeutic compounds, especially in the model flexible peptide cyclosporine[75,76]. HDX-MS has had some exploratory applications to cyclosporine as well, though these studies are limited to cyclosporine and its analogues[63,77]. With a larger library of synthetic cyclic peptides, HDX-MS provides a medium-throughput screening approach to evaluate peptides for internal H-bonding and conformational flexibility. Putative conformationally flexible

peptides can be later confirmed with more intense NMR work. This rapid assessment of peptide conformational dynamics can help build a library of highly valuable conformationally flexible peptides and can prioritize candidates for further drug development.

## 1.9 CONCLUSION

In conclusion, cyclic peptides represent a versatile class of molecules with significant potential for therapeutic development. Pairing design of synthetic cyclic peptide libraries with high throughput screening can aid targeted drug discovery and lead maturation. In chapter two, we propose a cyclic peptide sequencing program to enable ASMS for cyclic peptides. This program is capable of identifying cyclic peptide sequences from complex $MS^n$ fragmentation data. We detail a mass spectrometry method that provides high quality and simplified spectra where possible. Improvements to the Comet program specifically designed for cyclic peptides show the utility of including $MS^3$ to pair nested sequencing spectra and increase sequence assignment confidence. Chapter three focuses on the application of chapter two's sequencing program and the creation of an affinity selection mass spectrometry drug discovery pipeline. We collaborated with the Raj (Emory University, Discovery and Developmental Therapeutics) and Totah (University of Washington, Medicinal Chemistry) groups to screen cyclic peptides in our lab. The work within this chapter will focus on the screening and mass spectrometry work we are responsible for. We optimized ASMS against the model protein 12ca5 with the Raj group and identified novel cyclic peptide binders. In collaboration with the Totah group, we applied our assay to difficult GPCR

protein targets using magnetic bead-tagged virus-like particles. ASMS was used to enrich endogenous binding ligands from the complex VLP environment, including a demonstration of feasibility from a mock-library screening. Chapter four focuses on conformationally flexible peptides, which are known to be valuable therapeutic compounds but are difficult to design and identify. We show that HDX-MS is a unique technique that is capable of illustrating the differences in structural dynamics across designed cyclic peptides and provides information in medium throughput. This tool is utilized to identify peptides with conformational flexibility with the hope that these peptides can aid the development of future membrane traversing flexible peptides. The integration of these tools and techniques is essential for realizing the full potential of cyclic peptides as next-generation therapeutics.

## 1.10 FIGURES

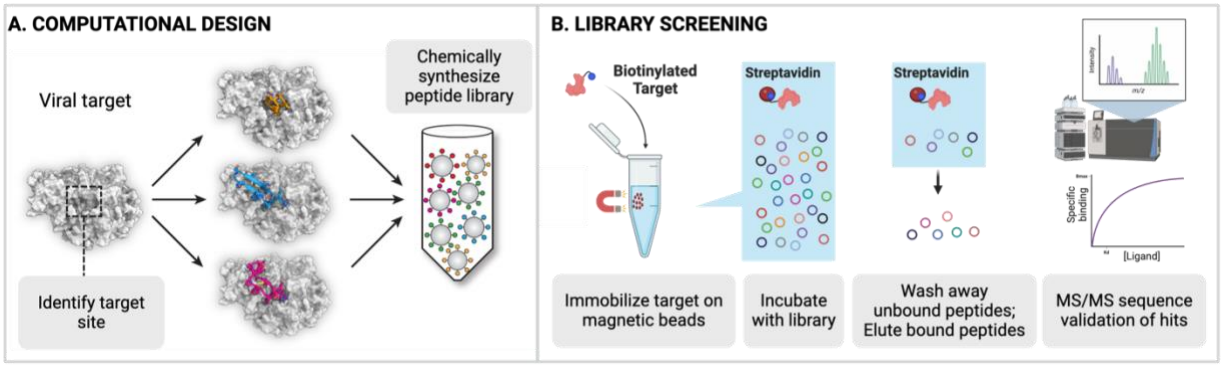


Figure 1. | Overview of Integrated Design and Screening of Cyclic Peptides.

A) Computational design against a protein target site results in selected peptides and library synthesis. B) The resulting library is incubated with the captured protein target and peptides with target interaction are retained for mass spectrometry validation.

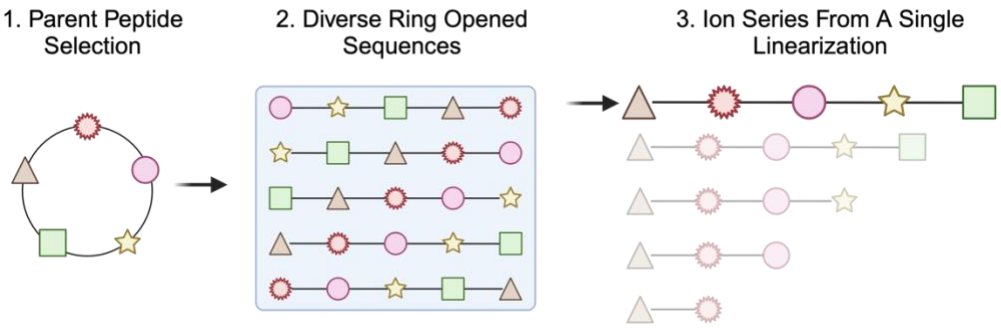


Figure 1.2 | Fragmentation Steps of Cyclic Peptides.

Parent cyclic peptides are selected, and the first fragmentation event linearizes the cyclic sequence at any peptide bond (**1**). For a five amino acid cyclic peptide, five unique linearized sequences are possible (**2**). Further fragmentation from a single linearized ion generates product fragments (**3**).

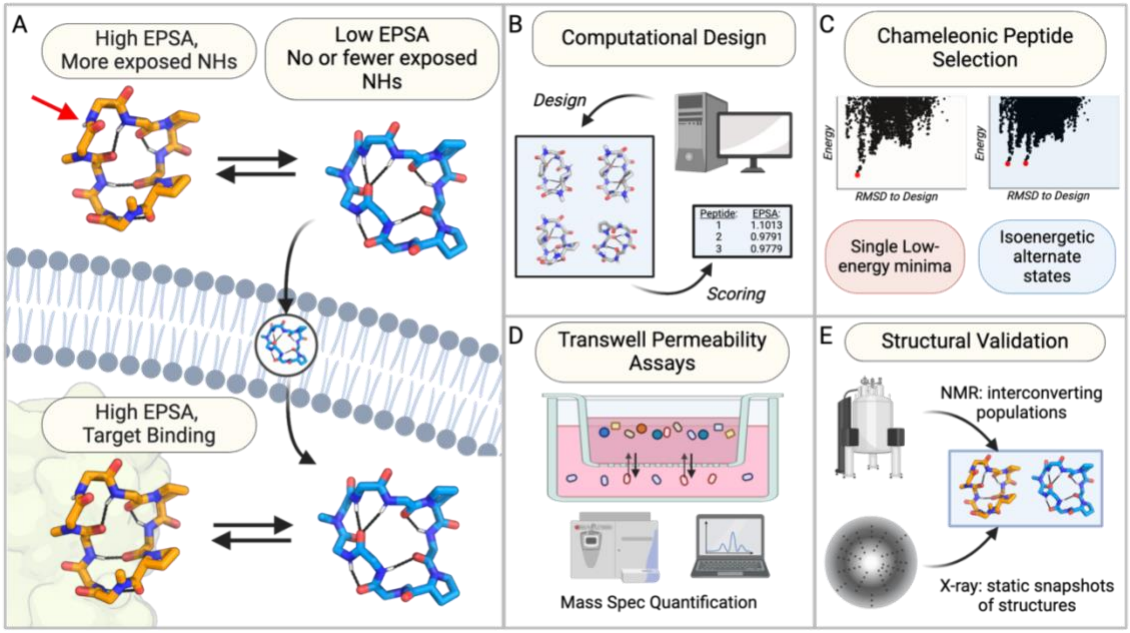


Figure 1.3 | Overview of Flexible (Chameleonic) Peptides for therapeutic development.

A) Conformational flexibility facilitates the presence of a binding competent state with more exposed polar surface area (EPSA) and a membrane traversing state without exposed backbone amide NHs. B) Computational design can be used to create cyclic peptides and score peptides with the potential for flexibility. C) Selecting peptides with competing predicted low energy states may yield more success in flexible peptide identification. D) Transwell assays can identify peptides with membrane traversal. E) Rigorous structural characterization is needed to confirm the presence of more than one isomer. Figure is reproduced with permission from Ramelot, et al. (2023)[1].

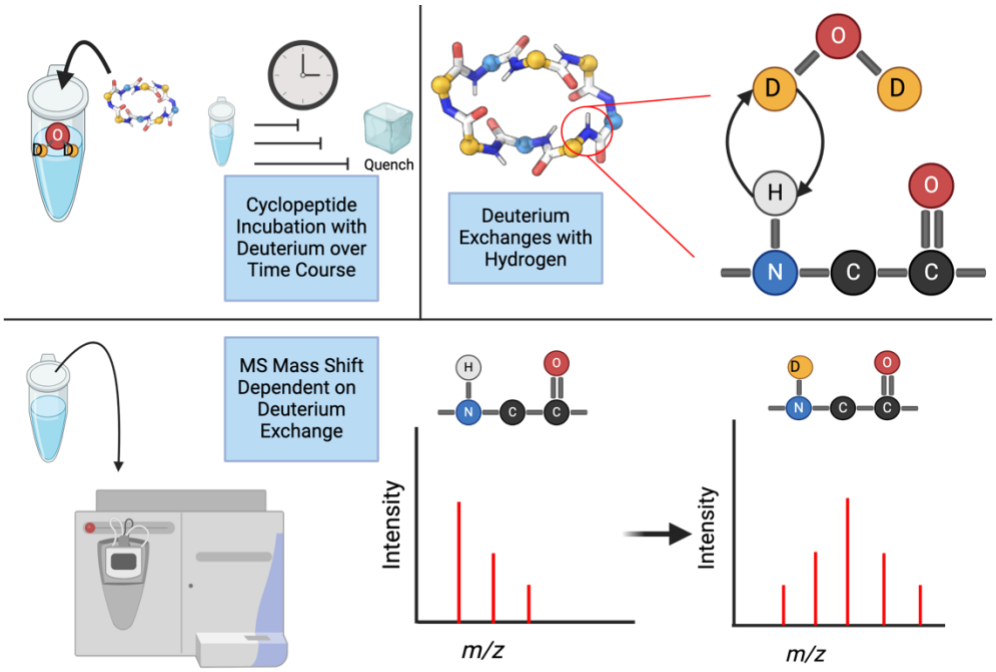


Figure 1.3 | Overview of HDX-MS for Cyclic Peptides.

A) Cyclic peptides are incubated in a deuterium buffer for exchange over a series of time points before the reaction is quenched. B) Exposed NHs will exchange with environmental deuterium. C) Mass spectrometry can measure the incorporation of deuterium over time.

## 1.11 BIBLIOGRAPHY

(1) Ramelot, T. A.; Palmer, J.; Montelione, G. T.; Bhardwaj, G. Cell-Permeable Chameleonic Peptides: Exploiting Conformational Dynamics in *de Novo* Cyclic Peptide Design. *Curr. Opin. Struct. Biol.* **2023**, *80*, 102603. https://doi.org/10.1016/j.sbi.2023.102603.

(2) Naylor, M. R.; Bockus, A. T.; Blanco, M.-J.; Lokey, R. S. Cyclic Peptide Natural Products Chart the Frontier of Oral Bioavailability in the Pursuit of Undruggable Targets. *Curr. Opin. Chem. Biol.* **2017**, *38*, 141–147. https://doi.org/10.1016/j.cbpa.2017.04.012.

(3) Martinovich, V. P.; Baradzina, K. U. Peptide Hormones in Medicine: A 100-Year History. *Russ. J. Bioorganic Chem.* **2022**, *48* (2), 221–232. https://doi.org/10.1134/S1068162022020157.

(4) Zorzi, A.; Deyle, K.; Heinis, C. Cyclic Peptide Therapeutics: Past, Present and Future. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24–29. https://doi.org/10.1016/j.cbpa.2017.02.006.

(5) Barman, P.; Joshi, S.; Sharma, S.; Preet, S.; Sharma, S.; Saini, A. Strategic Approaches to Improvise Peptide Drugs as Next Generation Therapeutics. *Int. J. Pept. Res. Ther.* **2023**, *29* (4), 61. https://doi.org/10.1007/s10989-023-10524-3.

(6) Hewitt, W. M.; Leung, S. S. F.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S. Cell-Permeable Cyclic Peptides from Synthetic Libraries Inspired by Natural Products. *J. Am. Chem. Soc.* **2015**, *137* (2), 715–721. https://doi.org/10.1021/ja508766b.

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

(8) Lee, A. C.-L.; Harris, J. L.; Khanna, K. K.; Hong, J.-H. A Comprehensive Review on Current Advances in Peptide Drug Development and Design. *Int. J. Mol. Sci.* **2019**, *20* (10). https://doi.org/10.3390/ijms20102383.

(9) Lam, K. S. Affinity Selection and Sequencing. *Nat. Chem. Biol.* **2019**, *15* (4), 320–321. https://doi.org/10.1038/s41589-019-0253-2.

(10) Saklaine, M. *The Secondary Structure of Protein: A Short Review*. EduGonist. https://www.edugonist.com/the-secondary-structure-of-protein-a-short-review/ (accessed 2024-08-06).

(11) Bhardwaj, G.; Mulligan, V. K.; Bahl, C. D.; Gilmore, J. M.; Harvey, P. J.; Cheneval, O.; Buchko, G. W.; Pulavarti, S. V. S. R. K.; Kaas, Q.; Eletsky, A.; Huang, P.-S.; Johnsen, W. A.; Greisen, P. J.; Rocklin, G. J.; Song, Y.; Linsky, T. W.; Watkins, A.; Rettie, S. A.; Xu, X.; Carter, L. P.; Bonneau, R.; Olson, J. M.; Coutsias, E.; Correnti, C. E.; Szyperski, T.; Craik, D. J.; Baker, D. Accurate de Novo Design of Hyperstable Constrained Peptides. *Nature* **2016**, *538* (7625), 329–335. https://doi.org/10.1038/nature19791.

(12) Mulligan, V. K.; Kang, C. S.; Sawaya, M. R.; Rettie, S.; Li, X.; Antselovich, I.; Craven, T. W.; Watkins, A. M.; Labonte, J. W.; DiMaio, F.; Yeates, T. O.; Baker, D. Computational Design of Mixed Chirality Peptide Macrocycles with Internal Symmetry. *Protein Sci.* **2020**, *29* (12), 2433–2445. https://doi.org/10.1002/pro.3974.

(13) Hosseinzadeh, P.; Bhardwaj, G.; Mulligan, V. K.; Shortridge, M. D.; Craven, T. W.; Pardo-Avila, F.; Rettie, S. A.; Kim, D. E.; Silva, D.-A.; Ibrahim, Y. M.; Webb, I. K.; Cort, J. R.; Adkins, J. N.; Varani, G.; Baker, D. Comprehensive Computational Design of Ordered Peptide Macrocycles. *Science* **2017**, *358* (6369), 1461–1466. https://doi.org/10.1126/science.aap7577.

(14) Russo, A.; Aiello, C.; Grieco, P.; Marasco, D. Targeting “Undruggable” Proteins: Design of Synthetic Cyclopeptides. *Curr. Med. Chem.* **2016**, *23* (8), 748–762.

(15) Tsomaia, N. Peptide Therapeutics: Targeting the Undruggable Space. *Eur. J. Med. Chem.* **2015**, *94*, 459–470. https://doi.org/10.1016/j.ejmech.2015.01.014.

(16) Gimeno, A.; Delgado, S.; Valverde, P.; Bertuzzi, S.; Berbís, M. A.; Echavarren, J.; Lacetera, A.; Martín-Santamaría, S.; Surolia, A.; Cañada, F. J.; Jiménez-Barbero, J.; Ardá, A. Minimizing the Entropy Penalty for Ligand Binding: Lessons from the Molecular Recognition of the Histo Blood-Group Antigens by Human Galectin-3. *Angew. Chem. Int. Ed.* **2019**, *58* (22), 7268–7272. https://doi.org/10.1002/anie.201900723.

(17) Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M. Structure of the Stapled P53 Peptide Bound to Mdm2. *J. Am. Chem. Soc.* **2012**, *134* (1), 103–106. https://doi.org/10.1021/ja2090367.

(18) Liu, M.; Pazgier, M.; Li, C.; Yuan, W.; Li, C.; Lu, W. A Left-Handed Solution to Peptide Inhibition of the P53-MDM2 Interaction. *Angew. Chem. Int. Ed.* **2010**, *49* (21), 3649–3652. https://doi.org/10.1002/anie.201000329.

(19) Pazgier, M.; Liu, M.; Zou, G.; Yuan, W.; Li, C.; Li, C.; Li, J.; Monbo, J.; Zella, D.; Tarasov, S. G.; Lu, W. Structural Basis for High-Affinity Peptide Inhibition of

P53 Interactions with MDM2 and MDMX. *Proc. Natl. Acad. Sci.* **2009**, *106* (12), 4665–4670. https://doi.org/10.1073/pnas.0900947106.

(20) Phan, J.; Li, Z.; Kasprzak, A.; Li, B.; Sebti, S.; Guida, W.; Schönbrunn, E.; Chen, J. Structure-Based Design of High Affinity Peptides Inhibiting the Interaction of P53 with MDM2 and MDMX. *J. Biol. Chem.* **2010**, *285* (3), 2174–2183. https://doi.org/10.1074/jbc.M109.073056.

(21) Mulligan, V. K.; Workman, S.; Sun, T.; Rettie, S.; Li, X.; Worrall, L. J.; Craven, T. W.; King, D. T.; Hosseinzadeh, P.; Watkins, A. M.; Renfrew, P. D.; Guffy, S.; Labonte, J. W.; Moretti, R.; Bonneau, R.; Strynadka, N. C. J.; Baker, D. Computationally Designed Peptide Macrocycle Inhibitors of New Delhi Metallo-β-Lactamase 1. *Proc. Natl. Acad. Sci.* **2021**, *118* (12), e2012800118. https://doi.org/10.1073/pnas.2012800118.

(22) Hosseinzadeh, P.; Watson, P. R.; Craven, T. W.; Li, X.; Rettie, S.; Pardo-Avila, F.; Bera, A. K.; Mulligan, V. K.; Lu, P.; Ford, A. S.; Weitzner, B. D.; Stewart, L. J.; Moyer, A. P.; Di Piazza, M.; Whalen, J. G.; Greisen, P. J.; Christianson, D. W.; Baker, D. Anchor Extension: A Structure-Guided Approach to Design Cyclic Peptides Targeting Enzyme Active Sites. *Nat. Commun.* **2021**, *12* (1), 3384. https://doi.org/10.1038/s41467-021-23609-8.

(23) Yaginuma, K.; Aoki, W.; Miura, N.; Ohtani, Y.; Aburaya, S.; Kogawa, M.; Nishikawa, Y.; Hosokawa, M.; Takeyama, H.; Ueda, M. High-Throughput Identification of Peptide Agonists against GPCRs by Co-Culture of Mammalian Reporter Cells and Peptide-Secreting Yeast Cells Using Droplet Microfluidics. *Sci. Rep.* **2019**, *9* (1), 10920. https://doi.org/10.1038/s41598-019-47388-x.

(24) Bashiruddin, N. K.; Suga, H. Construction and Screening of Vast Libraries of Natural Product-like Macrocyclic Peptides Using in Vitro Display Technologies. *Curr. Opin. Chem. Biol.* **2015**, *24*, 131–138. https://doi.org/10.1016/j.cbpa.2014.11.011.

(25) Josephson, K.; Ricardo, A.; Szostak, J. W. mRNA Display: From Basic Principles to Macrocycle Drug Discovery. *Drug Discov. Today* **2014**, *19* (4), 388–399. https://doi.org/10.1016/j.drudis.2013.10.011.

(26) Norman, A.; Franck, C.; Christie, M.; Hawkins, P. M. E.; Patel, K.; Ashhurst, A. S.; Aggarwal, A.; Low, J. K. K.; Siddiquee, R.; Ashley, C. L.; Steain, M.; Triccas, J. A.; Turville, S.; Mackay, J. P.; Passioura, T.; Payne, R. J. Discovery of Cyclic Peptide Ligands to the SARS-CoV-2 Spike Protein Using mRNA Display. *ACS Cent. Sci.* **2021**, *7* (6), 1001–1008. https://doi.org/10.1021/acscentsci.0c01708.

(27) *Proprietary Drug Discovery Platform - System(PDPS) ｜PEPTIDREAM INC ペプチドリーム株式会社*. PEPTIDREAM INC. https://www.peptidream.com/en/science/pdps/ (accessed 2024-08-06).

(28) Reid, P. C.; Goto, Y.; Katoh, T.; Suga, H. Charging of tRNAs Using Ribozymes and Selection of Cyclic Peptides Containing Thioethers. In *Ribosome Display and Related Technologies: Methods and Protocols*; Douthwaite, J. A., Jackson, R. H., Eds.; Springer: New York, NY, 2012; pp 335–348. https://doi.org/10.1007/978-1-61779-379-0_19.

(29) Hayashi, K.; Uehara, S.; Yamamoto, S.; Cary, D. R.; Nishikawa, J.; Ueda, T.; Ozasa, H.; Mihara, K.; Yoshimura, N.; Kawai, T.; Ono, T.; Yamamoto, S.; Fumoto, M.; Mikamiyama, H. Macrocyclic Peptides as a Novel Class of NNMT Inhibitors: A SAR

Study Aimed at Inhibitory Activity in the Cell. *ACS Med. Chem. Lett.* **2021**, *12* (7), 1093–1101. https://doi.org/10.1021/acsmedchemlett.1c00134.

(30) Janda, K. D. Tagged versus Untagged Libraries: Methods for the Generation and Screening of Combinatorial Chemical Libraries. *Proc. Natl. Acad. Sci.* **1994**, *91* (23), 10779–10785. https://doi.org/10.1073/pnas.91.23.10779.

(31) Gao, Y.; Amar, S.; Pahwa, S.; Fields, G.; Kodadek, T. Rapid Lead Discovery Through Iterative Screening of One Bead One Compound Libraries. *ACS Comb Sci* **2015**, 11.

(32) Hickey, J. L.; Sindhikara, D.; Zultanski, S. L.; Schultz, D. M. Beyond 20 in the 21st Century: Prospects and Challenges of Non-Canonical Amino Acids in Peptide Drug Discovery. *ACS Med. Chem. Lett.* **2023**, *14* (5), 557–565. https://doi.org/10.1021/acsmedchemlett.3c00037.

(33) Quartararo, A. J.; Gates, Z. P.; Somsen, B. A.; Hartrampf, N.; Ye, X.; Shimada, A.; Kajihara, Y.; Ottmann, C.; Pentelute, B. L. Ultra-Large Chemical Libraries for the Discovery of High-Affinity Peptide Binders. *Nat. Commun.* **2020**, *11* (1), 3183. https://doi.org/10.1038/s41467-020-16920-3.

(34) Zhang, G.; Li, C.; Quartararo, A. J.; Loas, A.; Pentelute, B. L. Automated Affinity Selection for Rapid Discovery of Peptide Binders. *Chem. Sci.* **2021**, *12* (32), 10817–10824. https://doi.org/10.1039/D1SC02587B.

(35) Qin, S.; Meng, M.; Yang, D.; Bai, W.; Lu, Y.; Peng, Y.; Song, G.; Wu, Y.; Zhou, Q.; Zhao, S.; Huang, X.; McCorvy, J. D.; Cai, X.; Dai, A.; Roth, B. L.; Hanson, M. A.; Liu, Z.-J.; Wang, M.-W.; Stevens, R. C.; Shui, W. High-Throughput Identification of G Protein-Coupled Receptor Modulators through Affinity Mass Spectrometry Screening

†Electronic Supplementary Information (ESI) Available. See DOI: 10.1039/C7sc04698g. *Chem. Sci.* **2018**, *9* (12), 3192–3199. https://doi.org/10.1039/c7sc04698g.

(36) Muchiri, R. N.; van Breemen, R. B. Affinity Selection–Mass Spectrometry for the Discovery of Pharmacologically Active Compounds from Combinatorial Libraries and Natural Products. *J. Mass Spectrom.* **2021**, *56* (5), e4647. https://doi.org/10.1002/jms.4647.

(37) Eng, J. K.; McCormack, A. L.; Yates, J. R. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (11), 976–989. https://doi.org/10.1016/1044-0305(94)80016-2.

(38) Ngoka, L. C. M.; Gross, M. L. Multistep Tandem Mass Spectrometry for Sequencing Cyclic Peptides in an Ion-Trap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (8), 732–746. https://doi.org/10.1016/S1044-0305(99)00049-5.

(39) Muth, T.; Renard, B. Y. Evaluating de Novo Sequencing in Proteomics: Already an Accurate Alternative to Database-Driven Peptide Identification? *Brief. Bioinform.* **2018**, *19* (5), 954–970. https://doi.org/10.1093/bib/bbx033.

(40) Boragine, D. M.; Huang, W.; Su, L. H.; Palzkill, T. Deep Sequencing of a Systematic Peptide Library Reveals Conformationally-Constrained Protein Interface Peptides That Disrupt a Protein-Protein Interaction. *ChemBioChem* **2022**, *23* (3), e202100504. https://doi.org/10.1002/cbic.202100504.

(41) Eng, J. K.; Fischer, B.; Grossmann, J.; MacCoss, M. J. A Fast SEQUEST Cross Correlation Algorithm. *J. Proteome Res.* **2008**, *7* (10), 4598–4602. https://doi.org/10.1021/pr800420s.

(42) Eng, J. K.; Jahan, T. A.; Hoopmann, M. R. Comet: An Open-Source MS/MS Sequence Database Search Tool. *PROTEOMICS* **2013**, *13* (1), 22–24. https://doi.org/10.1002/pmic.201200439.

(43) Eng, J. K.; Deutsch, E. W. Extending Comet for Global Amino Acid Variant and Post-Translational Modification Analysis Using the PSI Extended FASTA Format. *PROTEOMICS* **2020**, *20* (21–22), 1900362. https://doi.org/10.1002/pmic.201900362.

(44) Roland D. Kersten; Roland D. Kersten; Kersten, R. D.; Yu Liang Yang; Yang, Y.-L.; Yuquan Xu; Xu, Y.; Peter Cimermančič; Cimermancic, P.; Sang Jip Nam; Nam, S.-J.; Nam, S. J.; William Fenical; Fenical, W.; Michael A. Fischbach; Fischbach, M. A.; Bradley S. Moore; Moore, B. S.; Pieter C. Dorrestein; Dorrestein, P. C. A Mass Spectrometry–Guided Genome Mining Approach for Natural Product Peptidogenomics. *Nat. Chem. Biol.* **2011**, *7* (11), 794–802. https://doi.org/10.1038/nchembio.684.

(45) Behsaz, B.; Mohimani, H.; Gurevich, A.; Prjibelski, A.; Fisher, M.; Vargas, F.; Smarr, L.; Dorrestein, P. C.; Mylne, J. S.; Pevzner, P. A. De Novo Peptide Sequencing Reveals Many Cyclopeptides in the Human Gut and Other Environments. *Cell Syst.* **2020**, *10* (1), 99-108.e5. https://doi.org/10.1016/j.cels.2019.11.007.

(46) Novák, J.; Lemr, K.; Schug, K. A.; Havlíček, V. CycloBranch: De Novo Sequencing of Nonribosomal Peptides from Accurate Product Ion Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (10), 1780–1786. https://doi.org/10.1007/s13361-015-1211-1.

(47) Kavan, D.; Kuzma, M.; Lemr, K.; Schug, K. A.; Havlicek, V. CYCLONE--a Utility for de Novo Sequencing of Microbial Cyclic Peptides. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (8), 1177–1184. https://doi.org/10.1007/s13361-013-0652-7.

(48) Mohimani, H.; Liu, W.-T.; Kersten, R. D.; Moore, B. S.; Dorrestein, P. C.; Pevzner, P. A. NRPquest: Coupling Mass Spectrometry and Genome Mining for Nonribosomal Peptide Discovery. *J. Nat. Prod.* **2014**, *77* (8), 1902–1909. https://doi.org/10.1021/np500370c.

(49) Townsend, C.; Furukawa, A.; Schwochert, J.; Pye, C. R.; Edmondson, Q.; Lokey, R. S. CycLS: Accurate, Whole-Library Sequencing of Cyclic Peptides Using Tandem Mass Spectrometry. *Bioorg. Med. Chem.* **2018**, *26* (6), 1232–1238. https://doi.org/10.1016/j.bmc.2018.01.027.

(50) Syrovatkina, V.; Alegre, K. O.; Dey, R.; Huang, X.-Y. Regulation, Signaling, and Physiological Functions of G-Proteins. *J. Mol. Biol.* **2016**, *428* (19), 3850–3868. https://doi.org/10.1016/j.jmb.2016.08.002.

(51) Sarramegna, V.; Talmont, F.; Demange, P.; Milon, A. Heterologous Expression of G-Protein-Coupled Receptors: Comparison of Expression Systems from the Standpointof Large-Scale Production and Purification. *Cell. Mol. Life Sci. CMLS* **2003**, *60* (8), 1529–1546. https://doi.org/10.1007/s00018-003-3168-7.

(52) Ho, T. T.; Nguyen, J. T.; Liu, J.; Stanczak, P.; Thompson, A. A.; Yan, Y. G.; Chen, J.; Allerston, C. K.; Dillard, C. L.; Xu, H.; Shoger, N. J.; Cameron, J. S.; Massari, M. E.; Aertgeerts, K. Method for Rapid Optimization of Recombinant GPCR Protein Expression and Stability Using Virus-like Particles. *Protein Expr. Purif.* **2017**, *133*, 41–49. https://doi.org/10.1016/j.pep.2017.03.002.

(53) Nooraei, S.; Bahrulolum, H.; Hoseini, Z. S.; Katalani, C.; Hajizade, A.; Easton, A. J.; Ahmadian, G. Virus-like Particles: Preparation, Immunogenicity and Their

Roles as Nanovaccines and Drug Nanocarriers. *J. Nanobiotechnology* **2021**, *19* (1), 59. https://doi.org/10.1186/s12951-021-00806-7.

(54) Yin, D.; Zhong, Y.; Ling, S.; Lu, S.; Wang, X.; Jiang, Z.; Wang, J.; Dai, Y.; Tian, X.; Huang, Q.; Wang, X.; Chen, J.; Li, Z.; Li, Y.; Xu, Z.; Jiang, H.; Wu, Y.; Shi, Y.; Wang, Q.; Xu, J.; Hong, W.; Xue, H.; Yang, H.; Zhang, Y.; Da, L.; Han, Z.; Tao, S.; Dong, R.; Ying, T.; Hong, J.; Cai, Y. Dendritic-Cell-Targeting Virus-like Particles as Potent mRNA Vaccine Carriers. *Nat. Biomed. Eng.* **2024**. https://doi.org/10.1038/s41551-024-01208-4.

(55) Zhang, Y.-N.; Auclair, S.; Zhu, J. Virus-like Nanoparticle Vaccines for Inducing Long-Lasting Immunity against Infectious Diseases. *Natl. Sci. Rev.* **2024**, *11* (4), nwae032. https://doi.org/10.1093/nsr/nwae032.

(56) Huang, Z.; He, L.; Chen, Z.; Chen, L. Targeting Drugs to APJ Receptor: From Signaling to Pathophysiological Effects. *J. Cell. Physiol.* **2019**, *234* (1), 61–74. https://doi.org/10.1002/jcp.27047.

(57) Jiang, Y.; Yan, M.; Wang, C.; Wang, Q.; Chen, X.; Zhang, R.; Wan, L.; Ji, B.; Dong, B.; Wang, H.; Chen, J. The Effects of Apelin and Elabela Ligands on Apelin Receptor Distinct Signaling Profiles. *Front. Pharmacol.* **2021**, *12*, 630548. https://doi.org/10.3389/fphar.2021.630548.

(58) Ma, Y.; Yue, Y.; Ma, Y.; Zhang, Q.; Zhou, Q.; Song, Y.; Shen, Y.; Li, X.; Ma, X.; Li, C.; Hanson, M. A.; Han, G. W.; Sickmier, E. A.; Swaminath, G.; Zhao, S.; Stevens, R. C.; Hu, L. A.; Zhong, W.; Zhang, M.; Xu, F. Structural Basis for Apelin Control of the Human Apelin Receptor. *Struct. Lond. Engl. 1993* **2017**, *25* (6), 858-866.e4. https://doi.org/10.1016/j.str.2017.04.008.

(59) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997**, *23* (1), 3–25. https://doi.org/10.1016/S0169-409X(96)00423-1.

(60) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **2014**, *21* (9), 1115–1142. https://doi.org/10.1016/j.chembiol.2014.08.013.

(61) Ji, X.; Nielsen, A. L.; Heinis, C. Cyclic Peptides for Drug Development. *Angew. Chem. Int. Ed.* **2024**, *63* (3), e202308251. https://doi.org/10.1002/anie.202308251.

(62) Demiselle, J.; Fage, N.; Radermacher, P.; Asfar, P. Vasopressin and Its Analogues in Shock States: A Review. *Ann. Intensive Care* **2020**, *10* (1), 9. https://doi.org/10.1186/s13613-020-0628-2.

(63) Lam, K. H. B.; Le Blanc, J. C. Y.; Campbell, J. L. Separating Isomers, Conformers, and Analogues of Cyclosporin Using Differential Mobility Spectroscopy, Mass Spectrometry, and Hydrogen–Deuterium Exchange. *Anal. Chem.* **2020**, *92* (16), 11053–11061. https://doi.org/10.1021/acs.analchem.0c00191.

(64) Ahlbach, C. L.; Lexa, K. W.; Bockus, A. T.; Valerie Chen; Crews, P.; Jacobson, M. P.; Lokey, R. S. Beyond Cyclosporine A: Conformation-Dependent Passive Membrane Permeabilities of Cyclic Peptide Natural Products. *Future Med. Chem.* **2015**, *7* (16), 2121–2130. https://doi.org/10.4155/fmc.15.78.

(65) Heal, J. W.; Wells, S. A.; Blindauer, C. A.; Freedman, R. B.; Römer, R. A. Characterization of Folding Cores in the Cyclophilin A-Cyclosporin A Complex. *Biophys. J.* **2015**, *108* (7), 1739–1746. https://doi.org/10.1016/j.bpj.2015.02.017.

(66) Wang, C. K.; Swedberg, J. E.; Harvey, P. J.; Kaas, Q.; Craik, D. J. Conformational Flexibility Is a Determinant of Permeability for Cyclosporin. *J. Phys. Chem. B* **2018**, *122* (8), 2261–2276. https://doi.org/10.1021/acs.jpcb.7b12419.

(67) Bhardwaj, G.; O'Connor, J.; Rettie, S.; Huang, Y.-H.; Ramelot, T. A.; Mulligan, V. K.; Alpkilic, G. G.; Palmer, J.; Bera, A. K.; Bick, M. J.; Di Piazza, M.; Li, X.; Hosseinzadeh, P.; Craven, T. W.; Tejero, R.; Lauko, A.; Choi, R.; Glynn, C.; Dong, L.; Griffin, R.; Van Voorhis, W. C.; Rodriguez, J.; Stewart, L.; Montelione, G. T.; Craik, D.; Baker, D. Accurate de Novo Design of Membrane-Traversing Macrocycles. *Cell* **2022**, *185* (19), 3520-3532.e26. https://doi.org/10.1016/j.cell.2022.07.019.

(68) Garcia Jimenez, D.; Vallaro, M.; Rossi Sebastiano, M.; Apprato, G.; D'Agostini, G.; Rossetti, P.; Ermondi, G.; Caron, G. Chamelogk: A Chromatographic Chameleonicity Quantifier to Design Orally Bioavailable Beyond-Rule-of-5 Drugs. *J. Med. Chem.* **2023**, *66* (15), 10681–10693. https://doi.org/10.1021/acs.jmedchem.3c00823.

(69) Whitty, A.; Zhong, M.; Viarengo, L.; Beglov, D.; Hall, D. R.; Vajda, S. Quantifying the Chameleonic Properties of Macrocycles and Other High-Molecular-Weight Drugs. *Drug Discov. Today* **2016**, *21* (5), 712–717. https://doi.org/10.1016/j.drudis.2016.02.005.

(70) Watson, M. J.; Harkewicz, R.; Hodge, E. A.; Vorauer, C.; Palmer, J.; Lee, K. K.; Guttman, M. Simple Platform for Automating Decoupled LC–MS Analysis of

Hydrogen/Deuterium Exchange Samples. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (2), 597–600. https://doi.org/10.1021/jasms.0c00341.

(71) Murphree, T. A.; Vorauer, C.; Brzoska, M.; Guttman, M. Imidazolium Compounds as Internal Exchange Reporters for Hydrogen/Deuterium Exchange by Mass Spectrometry. *Anal. Chem.* **2020**, *92* (14), 9830–9837. https://doi.org/10.1021/acs.analchem.0c01328.

(72) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993**, *17* (1), 75–86. https://doi.org/10.1002/prot.340170110.

(73) Connelly, G. P.; Bai, Y.; Jeng, M. F.; Englander, S. W. Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins* **1993**, *17* (1), 87–92. https://doi.org/10.1002/prot.340170111.

(74) Nguyen, D.; Mayne, L.; Phillips, M. C.; Walter Englander, S. Reference Parameters for Protein Hydrogen Exchange Rates. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (9), 1936–1939. https://doi.org/10.1007/s13361-018-2021-z.

(75) Wang, C. K.; Northfield, S. E.; Colless, B.; Chaousis, S.; Hamernig, I.; Lohman, R.-J.; Nielsen, D. S.; Schroeder, C. I.; Liras, S.; Price, D. A.; Fairlie, D. P.; Craik, D. J. Rational Design and Synthesis of an Orally Bioavailable Peptide Guided by NMR Amide Temperature Coefficients. *Proc. Natl. Acad. Sci.* **2014**, *111* (49), 17504–17509. https://doi.org/10.1073/pnas.1417611111.

(76) Lee, D.; Lee, S.; Choi, J.; Song, Y.-K.; Kim, M. J.; Shin, D.-S.; Bae, M. A.; Kim, Y.-C.; Park, C.-J.; Lee, K.-R.; Choi, J.-H.; Seo, J. Interplay among Conformation, Intramolecular Hydrogen Bonds, and Chameleonicity in the Membrane Permeability and

Cyclophilin A Binding of Macrocyclic Peptide Cyclosporin O Derivatives. *J. Med. Chem.* **2021**, *64* (12), 8272–8286. https://doi.org/10.1021/acs.jmedchem.1c00211.

(77) Hyung, S.-J.; Feng, X.; Che, Y.; Stroh, J.; Shapiro, M. Detection of Conformation Types of Cyclosporin Retaining Intramolecular Hydrogen Bonds by Mass Spectrometry. *Anal. Bioanal. Chem.* **2014**, *406*. https://doi.org/10.1007/s00216-014-8023-1.

# CHAPTER 2.

# Identification of Cyclic Peptides Using $MS^n$ Sequencing for Affinity Selection Mass Spectrometry

## 2.1 INTRODUCTION

Cyclic peptides are an increasingly important class of therapeutic molecules with unique attributes and advantages. Their circular structure, typically formed by linking the head and tail amino acid positions, enhances pharmacokinetic properties like metabolic stability by eliminating polar termini groups[1,2]. Size-wise, cyclic peptides occupy a space between the established modalities of small molecules and large protein biologics, capturing the structural control and synthetic capabilities of small molecules while blending them with the target specificity of large biologics. This combination allows macrocyclic peptides to target proteins that are otherwise considered “undruggable” by small molecules and inaccessible to large proteins and antibodies[3–5]. The therapeutic potential of cyclic peptides is underscored by the widespread use of natural product macrocycles like antibacterial gramicidins and immunosuppressive cyclosporines[6]. However, despite their promise, drug discovery with cyclic peptides remains a significant challenge. The complexity of these molecules and the difficulties in efficiently

identifying and optimizing them mean there is still considerable space for development in this field, especially in how cyclic peptides are screened and sequenced.

Leveraging high-throughput screening (HTS) has become essential in modern drug discovery, enabling the rapid evaluation of vast compound libraries for potential therapeutic targets. However, when it comes to cyclic peptides, existing HTS techniques fall short. While display systems such as mRNA and yeast have been used to express and cyclize peptides, these methods often rely on alternative cyclization strategies, like disulfide bond formation, which tend to produce inferior pharmacokinetic properties to head-to-tail cyclization[7–10]. In contrast, affinity selection mass spectrometry (ASMS) employs synthetic libraries of compounds and utilizes direct MS identification of retained binding peptides. ASMS has proven very successful for linear peptides, allowing for the identification of interacting peptides from large, random libraries[11,12]. Unfortunately, ASMS is much less effective for cyclic peptides due to the additional complexity in cyclic peptide mass spectra. The lack of termini in cyclic peptides complicates direct sequencing, encouraging researchers to develop workaround strategies such as barcoding with linear peptide strings or chemically linearizing peptides after screening[9]. While these approaches offer some solutions, they all impose significant design restrictions and require additional synthesis and modification steps, limiting the ability to fully explore the chemical space of cyclic peptides. As a result, more advanced methods that allow direct MS fragmentation and sequence identification of cyclic peptides are necessary to overcome these limitations.

Historical attempts to directly sequence cyclic peptides have shown success using multistage $MS^n$ fragmentation. After generating linearized ions during the first

fragmentation event, subsequent events can identify and isolate unique peptide fragments, providing the opportunity to sequence a simplified fragmented portion of the peptide. This approach was first proposed in 1999 by Gross et al.[13], and the first computer program for cyclic peptide sequencing using this method was developed in 2003 by Redman et al.[14]. This program utilized multistage $MS^n$ data collection to run a database search format that creates theoretical spectra from provided sequences and matches them with experimental counterparts, scoring them based on similarity. Database searching is an ideal approach for ASMS: it is faster than *de novo* sequencing, allows for the restriction and introduction of modifications and non-natural amino acids as needed, and, due to the synthetic nature of a screening library, the sequence library is always known[15]. After this period of early development, interest in cyclic peptides shifted toward identifying non-ribosomal cyclic peptides through proteomics, leading to the creation of several *de novo* cyclic peptide sequencing programs that have very limited application to sequencing synthetic head-to-tail cyclic peptides[16–19].

While no new database-search program for cyclic peptides has been developed since, a notable contribution by Lokey et al. in 2018 introduced CycLS, a cyclic peptide sequencing program designed for HTS that does not use an $MS^n$ approach[20]. Instead, CycLS sequences multiple partial peptide fragments in a single $MS^2$ scan to rebuild the entire sequence. This method demonstrated both the feasibility of and the demand for direct cyclic peptide sequencing in HTS, highlighting the need for a modern cyclic peptide sequencing program.

However, cyclic peptide sequencing at higher fragmentation levels ($MS^3$ and above) suffers from spontaneous recyclization and rearrangement of fragment ions. Previous efforts to characterize cyclic peptides at the $MS^3$ level showed that linear precursor ions rapidly cyclize, with the resulting $MS^3$ fragmentation spectra dictated by the preferences of the cyclic isomer[21–23]. This leads to multiple fragmenting linear structures in $MS^3$ spectra, mirroring the ion complexity of multiple-ring openings observed in $MS^2$ spectra. Driven by proton localization on reactive backbone-amide-protonated positions, cyclic isomers of linear ions are more energetically favorable, with cyclization-reopening transition states also having a smaller energetic barrier than typical peptide bond fragmentation, resulting in fast interconversion between cyclic and linear isomers[23,24]. This hypothesis was supported by the identification of fragment ions that cannot be explained by primary sequence fragmentation but match cyclization and reopening fragmentation[25]. Further research has demonstrated that cyclic reopening locations are dominated by amino acid proton affinity, just like in parent cyclic peptides. This results in a high tendency to ring-reopen with a proline at the N-terminal position, which has the highest proton affinity among amino acid backbones and restrictive sterics[26]. These tendencies lend themselves to favored fragmentation pathways that can be exploited for sequence identification. Optimizing the strength of a few featured fragmented pathways will allow us to obtain accurate sequencing even for cyclic peptides.

Despite the challenges of recyclization and rearrangement, $MS^3$ fragmentation data can supplement cyclic peptide sequencing by providing additional support for $MS^2$ assignment. $MS^3$ data has been shown to eliminate ambiguity in $MS^2$ sequence

assignment for cyclic peptides and should do the same for computational sequencing[13,21,27]. Building upon mass spectrometry methods designed for database-matching programs and $MS^n$ fragmentation, our investigation aims to develop automated sequence identification of cyclic peptides. Rather than creating a new program, our approach involves expanding an existing tool, Comet, to leverage its cross-correlation scoring algorithm (reported as xcorr). Comet is fast, free, and efficient, and has been well-proven as a tool in proteomics for linear peptide identification[28–31]. Additionally, it is amenable to custom amino acid mass weights and modifications, which is essential for cyclic peptides and the use of noncanonical amino acids in screening. To enhance macrocyclic peptide sequencing beyond previous iterations, we adopt an $MS^n$ fragmentation approach, incorporating sequencing information from each level of fragmentation to maximize sequence identifications. This strategy represents a novel effort to refine and extend the capabilities of cyclic peptide sequencing.

Here, we seek to overcome the inherent complexity of MS-based cyclic peptide identification and enable unrestricted solution-phase high-throughput screening of cyclic peptides. By building a platform for direct identification of macrocyclic peptides via mass spectrometry, we aim to replicate the success shown in linear peptides. Doing so will allow researchers to rapidly evaluate thousands of compounds for target interaction, including compounds with custom non-canonical amino acids and modifications. This advancement should enable quicker identification of therapeutically relevant hit compounds and provide an avenue for lead compound optimization.

## 2.2 MATERIAL AND METHODS

*Materials*

HPLC-MS reagents were acquired from Fisher Scientific at 99% purity or better (Hampton, NH).

*Peptide synthesis*

Cyclic peptides were either purchased from Wuxi AppTec (Philadelphia, PA) or synthesized in house per Bhardwaj, et al. 2022[32]. Individual peptides were synthesized and pooled at an equimolar 200 nM concentration.

*High-Performance Liquid Chromatography – Mass Spectrometry (HPLC-MS)*

Peptide pools were prepared in 30% acetonitrile solutions in Waters total recovery MS vials. Samples were auto-injected by a Waters M-class HPLC onto a ThermoFisher Orbitrap Ascend**.** Peptides were separated using an ACQUITY UPLC Protein BEH C4 Column, 300 Å, 1.7 µm, 2.1 mm x 50 mm**.** A general gradient was achieved by increasing the amount of mobile phase B (0.1% FA in 100% ACN) in mobile phase A (0.1% FA in water) from 0 to 100 over 90 minutes at a flow rate of 0.3 mL/min. The column was heated to 60°C. All data was collected in positive mode.

MS acquisition settings were optimized for library screening conditions. $MS^1$ was collected in the orbitrap at 120000 resolution and 60% RF lens. Data dependent $MS^2$ was triggered in the orbitrap with a resolution of 60000 if the $MS^1$ scan met an intensity threshold of $2e^4$ and the precursor mass fell within the range of peptide masses in the library. Dynamic exclusion ignored a mass for forty seconds after two events within thirty seconds. $MS^2$ spectra were generated from CID at 35% fixed collision energy

across an isolation window of 1.6 m/z. Data dependent $MS^3$ was triggered in the ion trap with a normal scan rate if the $MS^2$ scan met an intensity threshold of $2.5e^3$ and the m/z shift matched the mass difference of an amino acid in the library within one Thomson. CID at fixed 30% collision energy produced $MS^3$ spectra. The method was set up to collect three $MS^2$ scans per $MS^1$ scan, and two $MS^3$ scans per $MS^2$ scan.

*MS spectral processing & comet settings*

ThermoFisher .raw MS files were converted to .mzXML files using ReAdW specifically to ensure the inclusion of the “filterLine” descriptive line in the final output, which is essential for Comet filtering of MS level. A custom program was written to drop nonproductive scans, defined as $MS^1$ and $MS^2$ scans which do not directly lead to $MS^3$ scans. This eliminates unnecessary search space by cutting up to 30% of scans in some tests. The cyclic peptide version of Comet necessitated a few specific inclusions in the params file, including the option for cyclic_peptide_search set to 1. The cyclic_peptide_NL setting was set to 28.0101, the exact mass of a CO loss. This will ignore neutral losses of a CO group, which we found to be a common loss during cyclic peptide linearization. The number of results reported per scan was set high enough to guarantee that sufficient coverage of similarly high scoring sequences is reported. For this work, we set num_output_lines to 500, but that number could be smaller for libraries with less mass degeneracy. Lastly, the decoy_search was set to 0 (off), as the cyclic peptide sequence permutations may overlap with the internally generated decoy sequences. Preprocessing in the params file filtered for a minimum of 5 distinct features per scan.  Further parameters were optimized for the MS method, a full list of settings, along with scripts and data, can be found at https://github.com/palmerjt/ASMS.

## 2.3 RESULTS

*Development of an $MS^3$ data acquisition method*

We first sought to establish an MS acquisition protocol that could generate fragmentation spectra that is clean and consistent for computational sequencing. We selected an eight amino acid peptide, 8.5, with the sequence PPAJZJVA as our model peptide (**Figure 1A**). This peptide includes canonical L- and D-amino acids with and without N-methylated backbone modifications. A common modification used to promote permeability and structural stability, N-methylation was included to explore fragmentation patterns and test non-canonical sequencing capabilities. The letter code “Z” was used to define N-methylated phenylalanine (161 Da), “B” was used to define N-methylated alanine (85 Da), and “J” was used to define N-methylated leucine (127 Da). Peptide 8.5 has a mass of 850.5 Da and ionizes as a singly charged protonated molecule with a m/z ratio of 851.5. We optimized MS scan parameters while infusing peptide 8.5 into an electrospray source at a flow rate of 5 uL/min using an external syringe pump.

In an effort to simplify spectra and keep large, informative product fragments, we compared the two most common fragmentation types: collision induced dissociation (CID) and higher-energy collisional dissociation (HCD) (**Figure 1B & 1C**). Though both methods generated fragmentation of the cyclic ring, HCD tended to distribute ring opening events more evenly throughout the peptide (**Figure 1C**). In contrast, backbone proton affinity differences amongst amino acids drove CID fragmentation towards ion

abundancies that reflect established trends[26,27]. Consistent with previous studies, we noted that preferential fragmentation N-terminal to an N-methylated amino acid is comparable to the same phenomenon in prolines[33]. Observing high m/z product ions and reliable fragmentation patterns, we decided to move forwards with CID as our primary fragmentation energy.

Within CID fragmentation, we found that higher collision energies and longer activation time led to a wider distribution of ring opening pathways. During initial tests it was generally observed that keeping collision energy as low and as fast as possible was essential for reducing the complexity of product ions in the $MS^2$ spectra. This is in support of previously described fragmentation approaches[33]. Despite pursuing low collision energy, our peptides required a relatively high energy environment (35 CE) to fragment due to the increased structural stability of cyclic peptides, the lack of termini, and the exclusion of amino acids with charged sidechains. We commonly observed strong formation of an immonium ion corresponding to a neutral loss of a carbonyl group (C=O) from C-terminal oxazolone group[34].

$MS^2$ spectra of peptide 8.5 were annotated with each of the possible linearized ion series, and it was found that the bulk of ion current was the result of three main linearized ion series (**Figure 1A & 1B**). Though b and y nomenclature are commonplace for peptide fragments, it is worth noting that y ions for cyclic peptides do not have a C-terminus. This results in identical b and y ions that can correspond to the same m/z peak despite being from different competing linearized peptides. We attributed 34% of total ion current to the linear peptide ZJVAPPAJ, which made it the highest contributor to the mass spectra. Due to overlapping b and y ions with

ZJVAPPAJ, VAPPAJZJ also accounted for a high portion of total ion current. Though we cannot rule out VAPPAJZJ, its contribution as a significant linearized ion series seems less likely due to several low intensity product ions and the unlikelihood that peptide fragmentation would favor breaking N-terminal to a valine.

To generate informative $MS^3$ spectra, we preferentially selected an *i-1* fragment (the primary sequence less one residue) as a precursor (**Figure 1D**). We observed that selecting the highest m/z fragment ion led to selection of neutral loss fragments (-28?) and co-isolation of multiple linear sequences in $MS^3$ (**Figure 1E**). Peptide 8.5 had a strong m/z peak corresponding to loss of an N-methylated leucine with a mass of 127 ("J"). Selecting the *i-1* $b_7$ peak of peptide 8.5 simplified the $MS^3$ spectra by preventing the linearized ion series PPAJZJVA from being co-isolated. We attributed 42% of the total ion current to the linearized sequence ZJVAPPAJ, with some minor contribution from VAPPAJZJ, as they can both lose a C-terminal J amino acid (**Figure 1D**). When optimizing $MS^3$ fragmentation, high fragmentation energy was required to achieve full fragmentation. Just like $MS^2$ fragmentation, a collision energy of 35% was used. This is consistent with the proposed recyclization of linearized precursor ions[23,25,35].

*Identification and minimization of sequence rearrangement*

The presence of unexplained peaks in our $MS^3$ spectra for peptide 8.5 and the recyclization propensity of cyclic peptides at higher $MS^n$ levels led us to investigate the impact of sequence rearrangement on $MS^3$ spectra. We screened a library of synthetic cyclic peptides for rearrangement using the optimized method above and identified peptide 10.j1 as a peptide with notable $MS^3$ spectral complexity. Peptide 10.j1 was

infused into an electrospray source at a flow rate of 5 uL/min using an external syringe pump while high quality $MS^2$, $MS^3$, and $MS^4$ data was collected.

Peptide 10.j1 has a known sequence of PVIPBJPLLI and $MS^n$ data was annotated with this information, where the custom amino acid “B” corresponds to an N-methylated alanine. $MS^2$ fragmentation resulted in a high intensity peak at 942.6, which corresponds to an *i-1* loss of leucine or isoleucine (**Figure 2A**). This $b_9$ was selected for $MS^3$ analysis, however due to the presence of four leucines/isoleucines in the parent sequence, four linearized ions had to be considered during $MS^3$ annotation (PBJPLLIPV, PVIPBJPLL, IPVIPBJPL, and LIPVIPBJP). Despite already having four ion series, we could not account for all of the peaks using just these product ions. We annotated the $MS^3$ spectra with several internal ion fragments but still identified one m/z peak, 506, which could not be explained with any naturally occurring sequence of amino acids (**Figure 2B**). We propose that this 506 peak is the result of the linearized sequence PBJLLIPV recyclizing and opening in a different spot for $MS^3$ fragmentation, resulting in a wholly new sequence: PLLIPVPBJ. To test this, we isolated 506 for $MS^4$ fragmentation and assigned fragments to the resulting spectra. $MS^4$ of 506 revealed many competing linear fragment ions, suggesting that the putative 506 peptide recyclized and ring opened at the $MS^4$ level as well (**Figure 2C**). We observed several mass differences corresponding to losses of P, V, B, and J amino acids and no peaks that offered refuting evidence for the composition of 506. All product ions were consistent with the assertion that peak 506 was composed of the amino acids PVPBJ, a unique sequence not found in our parent peptide.

We attempted to optimize MS parameters to abolish the presence of rearranged peaks in 10.j1. As previously reported, low fragmentation energy contributed to reducing sequence arrangement, though the decrease in rearranged ion intensity was accompanied by a decrease in overall ion intensity for lower m/z fragments. We did not find a feasible way to occlude rearranged peaks wholly. In general, it was difficult to find evidence of sequence rearrangement amongst the screened synthetic peptides, and more exhaustive tests are required to understand the impact of rearrangements and why they occur.

*Testing the sequencing capability of Cyclic Comet*

Despite the potential complexity imposed by rearrangements at the $MS^3$ level, we tested whether $MS^3$ spectra holds utility for sequencing confidence. We sought to automate cyclic peptide sequencing using the developed MS collection methods. Comet, a successful linear peptide sequencing tool, is ill-equipped to handle the complexity of cyclic peptide spectra. We made three changes to expand Comet for cyclic peptides (now referred to as Cyclic Comet). First, Cyclic Comet was adjusted to calculate peptide and amino acid weight without termini (**Figure 3A**). Secondly, Cyclic Comet had to expand its database search space by permuting all the linear sequences possible from the provided library of cyclic sequences (**Figure 3B**). This increases the number of sequences to the product of the number of peptides multiplied by the amino acid size of the library (i.e. a library of 100 10-mers will now have 1000 sequences instead of 100). Lastly, we added a filtering step to Cyclic Comet to separate $MS^2$ and $MS^3$ level spectra (**Figure 3C**). Since Comet treats each scan as an independent peptide to be sequenced, there is no existing framework for passing information

between $MS^2$ and $MS^3$ sequence. Collectively, these changes enable Cyclic Comet to appropriately evaluate $MS^n$ data of cyclic peptides.

Comet's primary sequence scoring metric, xcorr, is subject to variation depending on spectral quality, library size, and library redundancy. As a result, direct comparison between Cyclic Comet tests is impossible. To facilitate discussion and establish thresholds, we compared extracted ions to the whole population and generated a percentile rank and normalized difference metric that corresponds to the standard deviation of the population. We define this normalized difference metric as:

$$(xcorr_{max} - xcorr_{next})/\sigma_{scan}$$

This metric adjusts the separation between top scoring and next best scoring sequences based on the shape of the data distribution. The more promiscuous the scoring of the scan, the larger the standard deviation and the lower the normalized difference. This metric can be interpreted as an estimate in sequence assignment confidence.

We replaced our model peptide 8.5 with five new peptides and obtained high quality $MS^2$ and $MS^3$ fragmentation spectra for each peptide using the optimized MS acquisition settings. To test the ability of Cyclic Comet to identify the correct sequence from these scans, we generated decoy libraries to search against. For each peptide, four random amino acid positions were mutated through a ten amino acid library. The resulting combinatorial libraries had a complexity of $10^4$, or 10,000 sequences featuring a common motif and four variable positions. Our library of ten amino acids was

composed of hydrophobic amino acids and their N-methylated analogues: A, V, P, L, F, M, I, B (N-methylated A), J (N-methylated L/I), and Z (N-methylated F). The resulting libraries include multiple mass redundancies and potential sequence rearrangements of the true peptide sequence and should prove an adequate challenge for Cyclic Comet.

We analyzed the top two $MS^2$ / $MS^3$ scan pairs from each sequence in order to ascertain sequencing differences and compare performance (**Table 1**). The top two sequence assignment generally agreed, with some even scoring identically. Of the five peptides, four scored a sequence with the correct amino acid composition as the top match, and three had the correct composition and sequence order. For the two peptides where Comet was unable correctly identify the sequence, one (LPJAJPLA) did not list the correct sequence at all, and the other (APPVAJJV) was wrong by a single amino acid rearrangement. Equally as important as accurate sequence identification was confidence in the sequence identification. We observed a strong correlation between sequence assignment normalized difference and accuracy. LPJAJPLA had poor normalized differences for both $MS^2$ and $MS^3$ scans. VPPAJJAVP, meanwhile, had the strongest separation and accurately assigned sequence in all scans. The peptide with accurate composition but incorrect sequence order, APPVAJJV, has strong separation in $MS^2$ scans but poor $MS^3$ separation. Overall, the normalized difference threshold for a confident sequence assignment appears to be about one (**Table 1**). Above one, nine of eleven sequences were assigned correctly, with only two wrong assignments coming on $MS^2$ scans of a peptide with identical amino acid composition. Conversely, all nine incorrect sequence assignments had normalized differences below one. Cyclic Comet did an adequate job sequencing cyclic peptides. Inaccurate sequence assignments are

possible due to both the complexity of our mock libraries and imperfect fragment ion spectra. False positive sequence assignments are possible within Cyclic Comet, but this scoring has shown utility in filtering out incorrect assignments.

The case study of these five peptides illustrates the benefits of including $MS^3$ scans in cyclic peptide sequencing. $MS^3$ sequencing confidence may have a stronger correlation with accuracy than $MS^2$ scans, as $MS^3$ spectra is typically cleaner and easier to interpret. The peptide PPIVPPBL provides an important example of how $MS^3$ scans can reduce ambiguity in sequence assignment. Both the correct sequence (PPIVPPBL) and a minor rearrangement (LPLVPPBP) had xcorrs of 3.37. $MS^3$ sequencing was consistent between the two, and the sequenced fragment (PPBLPPI) was able to identify the PPBLP motif in dispute, thereby removing ambiguity in sequence assignment. Additionally, $MS^3$ sequencing proved strong resolving power in separating accurate and inaccurate sequence assignments by having poor confidence in the sequence rearranged misidentification of APPVAJJV. Between $MS^3$'s ability to reduce ambiguity and its strong correlation between normalized difference and sequencing accuracy, we propose incorporating $MS^3$ as a supplemental data point for interpreting $MS^2$ suggested sequences. $MS^3$ scans broadly agree with $MS^2$ scans, and the data do not suggest that an accurate $MS^2$ sequence can be at odds with $MS^3$ sequence assignment, despite the known caveats of the data.

### *Mock ASMS Whole Library Sequencing*

To test Cyclic Comet's capabilities in conditions similar to ASMS, we created a pool of cyclic peptides. We combined our five test peptides with 35 additional synthetic peptides in an equimolar pool at 200 nM and injected them onto a mass spectrometer

over a 90-minute C18 separation gradient. A longer gradient was used to separate the cyclic peptides and the full $MS^3$ data collection method was employed to mimic ASMS data collection settings. The resulting mass spectra was transformed to .mzXML using ReAdW in preparation for Cyclic Comet sequencing. The library pool represented a variety of peptide sizes from 8 to 11 amino acids but were all made from the same limited set of hydrophobic amino acids.

A corresponding library of dummy sequences was created, totaling 8,114 sequences made from scrambling the correct peptide sequences. Cyclic Comet was used to sequence all scans against all peptides in the library, which led to 46,692 individual sequence assignments across 3,842 scans ($MS^2$ and $MS^3$ combined). $MS^2$ has a higher range of xcorr scores due to the size of matching fragments and number of fragments present but both scan levels had similar distribution profiles, with the skewness of 1.42 for $MS^2$ and 1.48 for $MS^3$ (**Figure 4A**). Both populations saw significant populations extend towards higher xcorr values, lending themselves a subpopulation of high confidence correctly sequenced scans. Normalized difference, as with the model peptide libraries, showed efficacy in separating correct from incorrect scans (**Figure 4B**). As a testament to Cyclic Comet's accuracy, most of the top sequences from each scan were correct sequence IDs. Those that were not correct did not score well by normalized difference in either $MS^2$ or $MS^3$, as Normalized Difference above 2 was 100% accurate in identifying scans with correct sequences (**Figure 4B**).

The utility of $MS^3$ scans as supplemental filtering criteria for $MS^2$ sequence assignment was explored. $MS^3$ and $MS^2$ sequences were paired, and if reliable $MS^3$ sequences were not in the $MS^2$ assigned sequence, the pairing was tossed (**Figure**

**4C**). This had significant impact on reducing the number of sequence assignments. Of the 46,692 sequence assignments, 6,361 (13.6%) of them assigned a real peptide sequence. After filtering by $MS^2/MS^3$ sequence assignment, there were only 1,935 total sequence assignments, and 1,865 (96.4%) were to real peptides (**Figure 4C**). Though the number of assignments to real peptides was reduced, the bulk of those cut were low scoring and likely inaccurate.

Using the criteria that sequence assignments had to have agreeing $MS^2/MS^3$ sequence assignment and a normalized difference threshold of one, we were able to filter and sequence 47 peptide sequences, including 36 correct sequence identifications of real peptides. This left four peptides unidentified, which did not appear to trigger $MS^2$ due to either low intensity of co-elution with stronger peptide signals. These 47 peptides that passed our criteria had a range of $MS^2$ and $MS^3$ xcorrs that was not wholly different than the entire population (**Figure 5**). We believe that Cyclic Comet is accurate at sequence identification within a single scan but loses some accuracy when attempting to identify correct sequence assignment from many scans. The inclusion of 11 other peptide sequences that meet our final criteria is testament to this. Most of these sequences are likely the result of nonpeptidic mass fragments with little competition for sequence assignment, resulting in high normalized difference and paired agreement.

## 2.4 DISCUSSION

In this chapter, we have detailed the development of an $MS^n$ data collection and computational sequencing pipeline for cyclic peptides. This pipeline is geared towards enabling screening of libraries of synthetic cyclic peptides without design restrictions. Utilizing an $MS^n$ approach, we have shown that high quality $MS^2$ and $MS^3$ cyclic peptide

spectra can be sequenced. We expanded Comet for cyclic peptides and demonstrated the accuracy of Cyclic Comet against complex screening libraries. The utility of sequencing $MS^n$ data proved to make Cyclic Comet more robust, and we were able to identify the correct peptide sequence from tests. We developed an effective package that can identify notable features from mass spectra, clean spectra, and elucidate cyclic peptide sequences to drive drug discovery.

The advancement of mass spectrometers has significantly improved the speed and efficiency of mass spectra generation. Modern mass spectrometers, with their enhanced speed and multitasking capabilities, now allow for the collection of additional layers of fragmentation spectra without compromising data quality or scan rate. In this work, we demonstrated that one $MS^2$ scan in the orbitrap and two $MS^3$ scans in the ion trap can be completed within one second and is therefore compatible with online LC-MS workflows. Collecting $MS^3$ data comes at minimal cost and offers the potential for deeper insights into cyclic peptide sequencing. We have optimized our $MS^n$ data collection around this concept, focusing on generating as much high-quality $MS^3$ spectra as possible**.** We utilize fast rates of dynamic exclusion to reach peptides that may be at lower intensity thresholds, and we use data dependent mass triggering for $MS^3$ scans. As instrumentation capabilities continue to advance, it is crucial that our assays evolve to take full advantage of these improvements.

Comet was selected as a peptide sequencing program for development because of its ability to keep up with rapid and accurate sequencing, its xcorr scoring algorithm, and it is free, open source, and customizable. Comet's designed application is for proteomics, which necessitates speed in identifying thousands of peptides to rebuild

proteins and monitor changes. This is wonderfully applicable to high throughput screening because synthetic libraries can be very large and mass spectra must be searched against each sequence. In addition, Comet's internal preprocessing can drastically reduce the number of spectra needed to search. Filtering data for peptidic spectra is especially important for cyclic peptide ASMS, as the hydrophobicity of cyclic peptides commonly means that they ionize as a +1 charge state and cannot be easily separated from background +1 charge ions. Comet was particularly attractive because its xcorr scoring algorithm relies on matching theoretical peaks to peaks in the experimental spectra, not the other way around. This prevents Comet from penalizing overly complex mass spectra that are the norm for cyclic peptides because it is ok if there are peaks in the experimental spectra that go unmatched. This means that noisier spectra may be more promiscuous than more limited spectra, but it really means that sequence assignment is not punished for not matching every peak in a spectrum. This is fundamentally a different question from the proteomics background of Comet, where we are no longer asking for high accuracy across a broad number of mass spectra but in fact asking for confidence in single spectra sequencing. Lastly, Comet is free and has excellent support for customization. Just as was done here with N-methylated amino acids, it is easy to define non-canonicals amino acids or variable modifications. Comet takes .mzXML files, which makes it approachable from any mass spectrometer or background. Comet provided so many advantages that we thought it easier to expand Comet for cyclic peptides rather than build our own sequencing tool.

Optimizing Cyclic Comet for ASMS-like applications requires reducing the search space, when possible, to speed analysis. To supplement Comet's preprocessing which

filters based off of spectral content, we wrote a program to preprocess based off of the relationship between scans. Nonpeptidic features that fall within the mass range for $MS^2$ triggering typically fail to trigger $MS^3$, which is dependent on a mass loss equivalent to an amino acid. These $MS^2$ scans typically score poorly and rarely conflate results but can take up valuable computational time being scored against the peptide library sequences. For many scans this preprocessing step is redundant, as triggered $MS^2$ spectra for nonpeptidic features commonly go hand in hand with low intensity or low number of fragments and are dropped by Comet's preprocessing. Nonetheless, our additional tool is a fast and robust way to clean .mzXML files before Cyclic Comet analysis.

While xcorr is a value that shifts based on several factors (library size, fragmentation quality, etc.), a consistent trend is that $MS^3$ scans do not score as well as $MS^2$ scans. $MS^3$ scans match smaller sequences and use fewer peaks to do so, and therefore do not have as high of an xcorr. Normalized difference between top and next best scores is a way to minimize the impact of these differences and facilitate comparison. This is a double-edged sword, as we can be more confidence in $MS^3$ sequencing results if the large normalized difference is present, but if not, we cannot be sure if there is an exceptionally similar sequence in the library that scores well or if the sequence identified is incorrect.

An expected challenge with this approach was the cyclization and rearrangement of linearized $MS^3$ precursor ions which introduces mass peaks that do not match the possible linearized sequences. At this time, more work is needed to understand the full scope of sequence rearrangements on peptide identifications. There is support for the

diminished impact of rearranged sequences when using computational sequencing[36]. In our observed peptide 10.j1 we were able to identify peaks with notable intensity that were the result of sequence rearrangement. In this peptide, however, the rearranged sequence was one of five linearized ions contributing to the spectrum. Though mass degeneracy makes it impossible to distinguish the impact of each linearized fragment, we expect that the contribution of the rearranged sequence is minor compared to the other non-rearranged linearized fragments. The presence of the rearranged peaks complicates this spectrum, but so too do the co-isolated original linearized fragments. Another important consideration drives our understanding of sequence rearrangement: $MS^2$ fragmentation and ring opening is driven by the same factors that dictate ring reopening in $MS^3$. The propensity of peptide bonds to break at preferential positions results in ring reopening consistently between $MS^2$ and $MS^3$ levels. Cyclic peptides finally provide an advantage, as linearized sequences selected for $MS^3$ fragmentation should already have positioned the most favorable N-terminus amino acid at the ring breaking point. Ring recyclization and reopening is likely to break N-terminal to the same favored amino acid, which is why we consistently see prolines or N-methylated amino acids as the N-terminus. We utilized CID fragmentation over HCD to maximize the impact of these preferred pathways. As with peptide 10.j1, rearrangement may occur when the ejected amino acid leaves a new C-terminal amino acid that does not like to break C-terminal, like N-methylated leucine in the case of 10.j1. In that case, it becomes much more favorable to break a peptide bond elsewhere in the ring. This additional interpretation gives us guidelines for looking for rearrangement in other peptide sequences. If the ejected amino acid from the $MS^2$ scan leaves an amino acid

at the C-terminus that does not like to fragment there, closer investigation of that spectra may be needed.

Cyclic Comet was unable to accurately identify the correct sequence of all five model peptides. The strong correlation between poor normalized difference and sequencing performance suggests that Cyclic Comet was not wholly at fault for being unable to sequence all cyclic peptides. Poor normalized difference could indicate incredibly complex spectra that is capable of matching a variety of fragments and sequences. It could also indicate poor spectral quality in general and our data collection method could require further optimization. With our current outcome, it is likely that mass degeneracy and competition within the mock library led to many high scoring sequence assignments. Two of the five peptides tested had incorrect sequence assignments. However, it is notable that the peptide with higher normalized difference was correctly identified in terms of its amino acid composition, suggesting that Cyclic Comet's scoring system can provide valuable information even at lower normalized differences. The complexity of cyclic peptide spectra leads to difficulty discriminating sequence at the di- and tri-peptide level. Prolines, in particular, tend to fragment only as dipeptides and thus the ability to discern sequence at this level of accuracy can be difficult. Despite this challenge, our findings support the idea that Cyclic Comet sequences hold potential utility even when the normalized difference scores are suboptimal. At the very least, these sequences serve as a useful starting point for manual data annotation. Increasing resolution to filter through incorrect but high scoring sequence assignments is a main development focus moving forwards.

Normalized difference and $MS^2/MS^3$ sequence pairing have shown themselves to be efficacious tools in filtering out incorrect sequence assignments within a scan. Of the forty peptides included in the mock ASMS experiment, 36 of them had validated $MS^2/MS^3$ sequence pairings. In a less exhaustive library, Cyclic Comet was able to correctly identify all of the peptides contained in the scan. While Cyclic Comet is accurate in identifying sequences within individual scans, its performance may diminish when integrating data across multiple scans. The reduction from 6,361 “correct” sequence assignments to 1,935 after applying $MS^2/MS^3$ pairing criteria could mean that true hits are being discarded to ensure accuracy. However, the fact that 96.4% of the remaining assignments corresponded to real peptides is encouraging. This trade-off between data retention and accuracy is an important area for future development.

The appearance of eleven incorrect sequences among the final identifications suggests that additional filtering is needed. Even with these filtering criteria, Cyclic Comet may be susceptible to non-peptidic mass fragments that produce high xcorr values. Due to the large number of scans within a given MS data set and the size of ASMS libraries, high scoring incorrect sequences are likely. The identification of these false positives highlights an area of need for Cyclic Comet. In Chapter Three, we introduce a mass feature enrichment program to separately identify relevant MS scans so that Cyclic Comet filtering is not so heavily needed.

## 2.5 CONCLUSION

We have attempted to address a gap in the field of cyclic peptide high throughput screening by automating an affinity selection mass spectrometry method. In a novel

application, we paired $MS^3$ data collection with Comet database searching peptide sequencing. Cyclic peptide fragmentation is inherently complex and difficult to interpret. To simplify this, we attempted to exploit preferential fragmentation pathways and gather supplemental $MS^3$ data. Using our model peptides, we identified an efficient data dependent fragmentation approach for isolating *i-1* length product ions for $MS^3$. Peptides have the ability to recyclize and ring open for $MS^3$ spectra and we identified this phenomenon within our own synthetic cyclic peptides. Despite this, our data suggests that $MS^3$ fragmentation holds value, and our sequencing tests support this claim. $MS^3$ sequencing appears to be more accurate when a reliable sequence is found. To this end, we suggest pairing $MS^2$ and $MS^3$ sequence assignments to assure agreement. This pipeline was shown to be a feasible option for ASMS screening of cyclic peptides, but further optimization is needed. In Chapter Three, we will begin to show application of this work for binder discovery.

All data and scripts used in this work are available at https://github.com/palmerjt/ASMS.

## 2.6 FIGURES

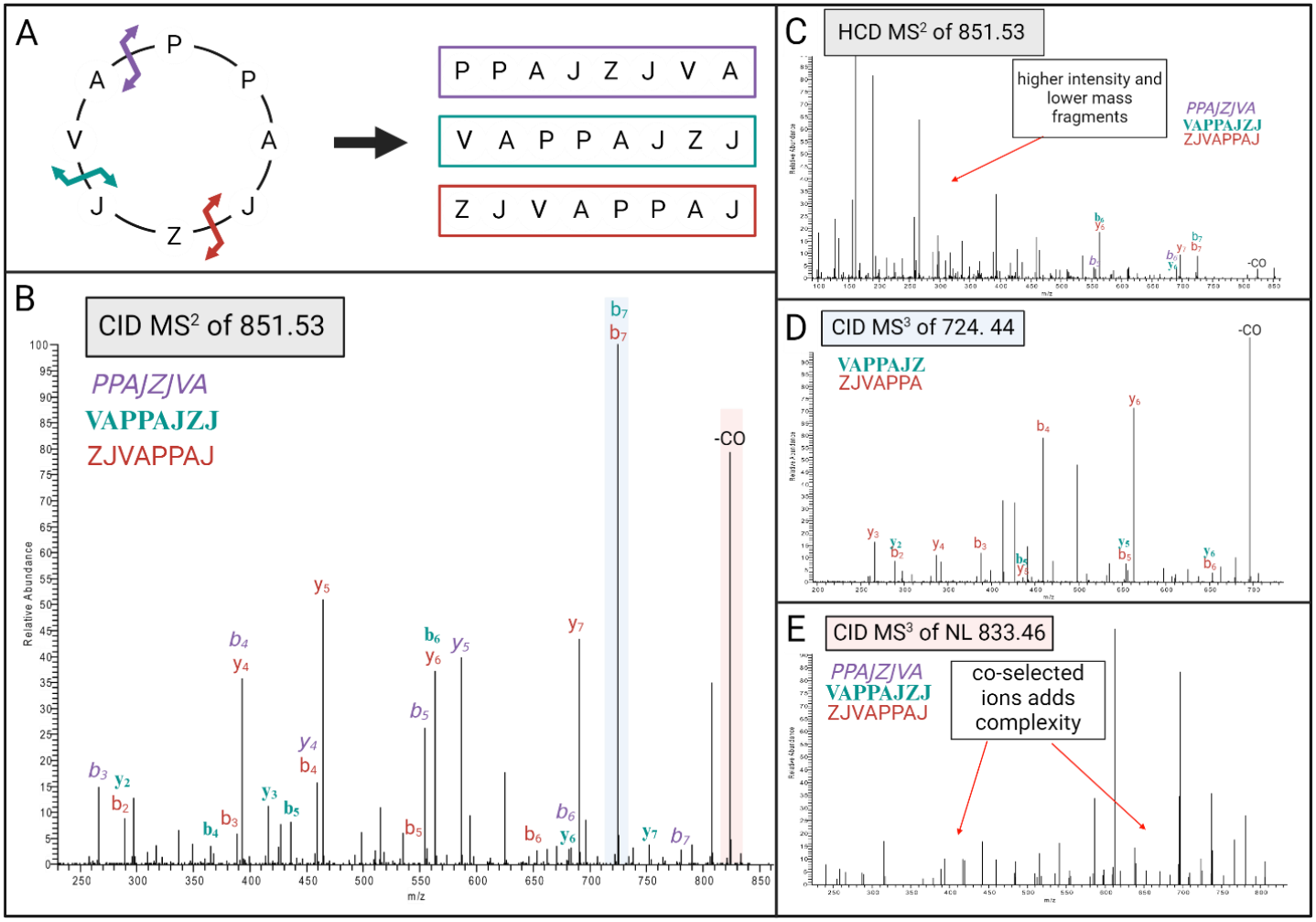


Figure 1 | Optimization of $MS^3$ Acquisition Protocol for Cyclic Peptides

**A)** Ring opening fragmentation sites and resulting linearized ion series present in $MS2^2$ spectra. **B)** $MS^2$ fragmentation of the full cyclic peptide, annotated with b and y ions from each linearized series. CID fragmentation has higher intensity, high m/z fragments than **C)** HCD $MS^2$ fragmentation of the full cyclic peptide. **D)** CID $MS^3$ of an *i-1* product ion leads to cleaner resulting fragmentation than **E)** CID $MS^3$ fragmentation of a neutral loss product ion.

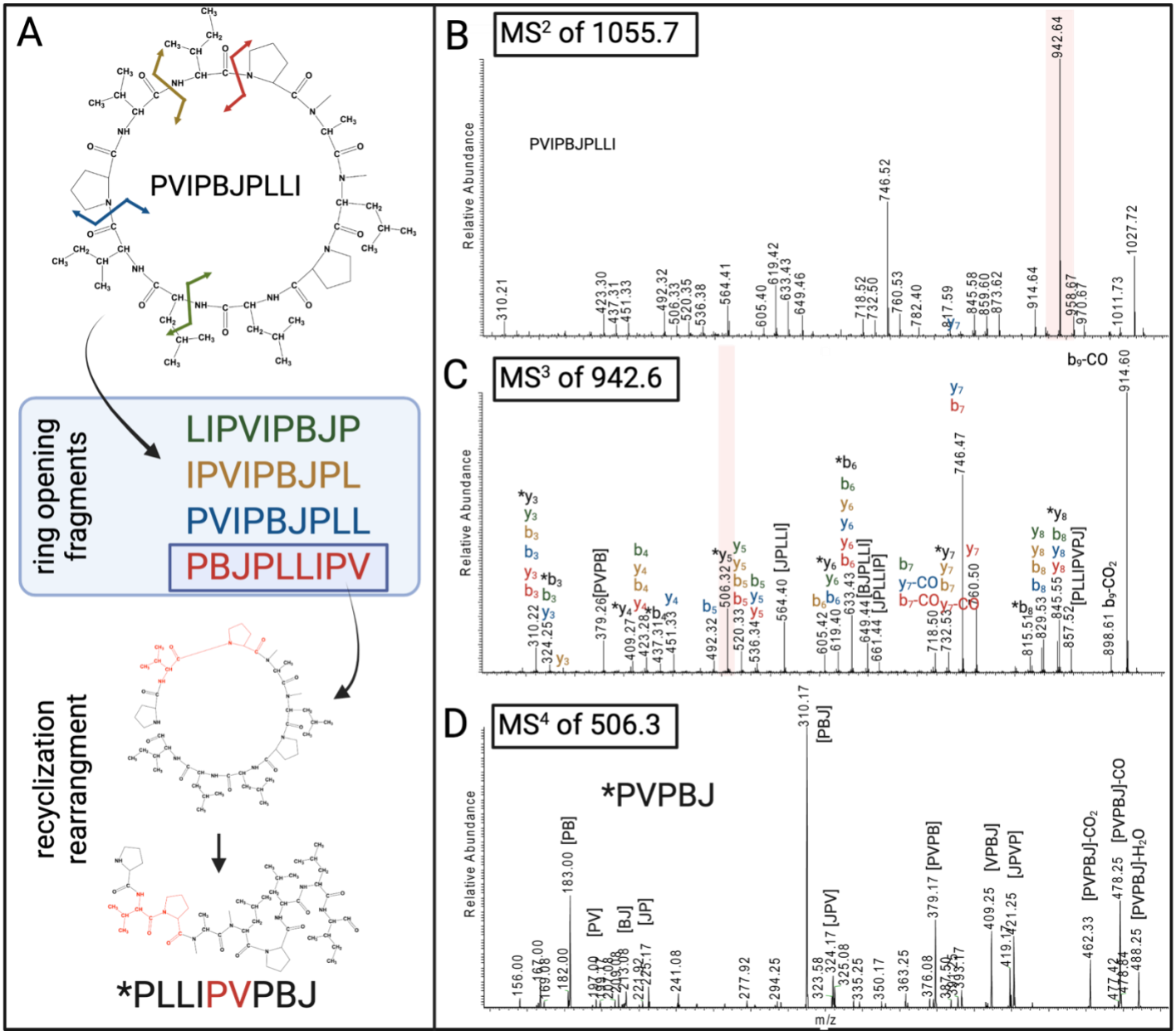


Figure 2.2 | Identification of Sequence Rearrangement Series in Cyclic Peptides

**A)** Peptide structure and ring opening sites are color coded with the resulting linear sequence. Proposed recyclization rearrangement is proposed. **B)** $MS^2$ fragmentation highlighting loss of a single leucine/isoleucine. **C)** $MS^3$ fragmentation is

annotated for identification of sequence rearrangement at the $MS^3$ level. **D)** $MS^4$ spectra of the prosed rearranged fragment.

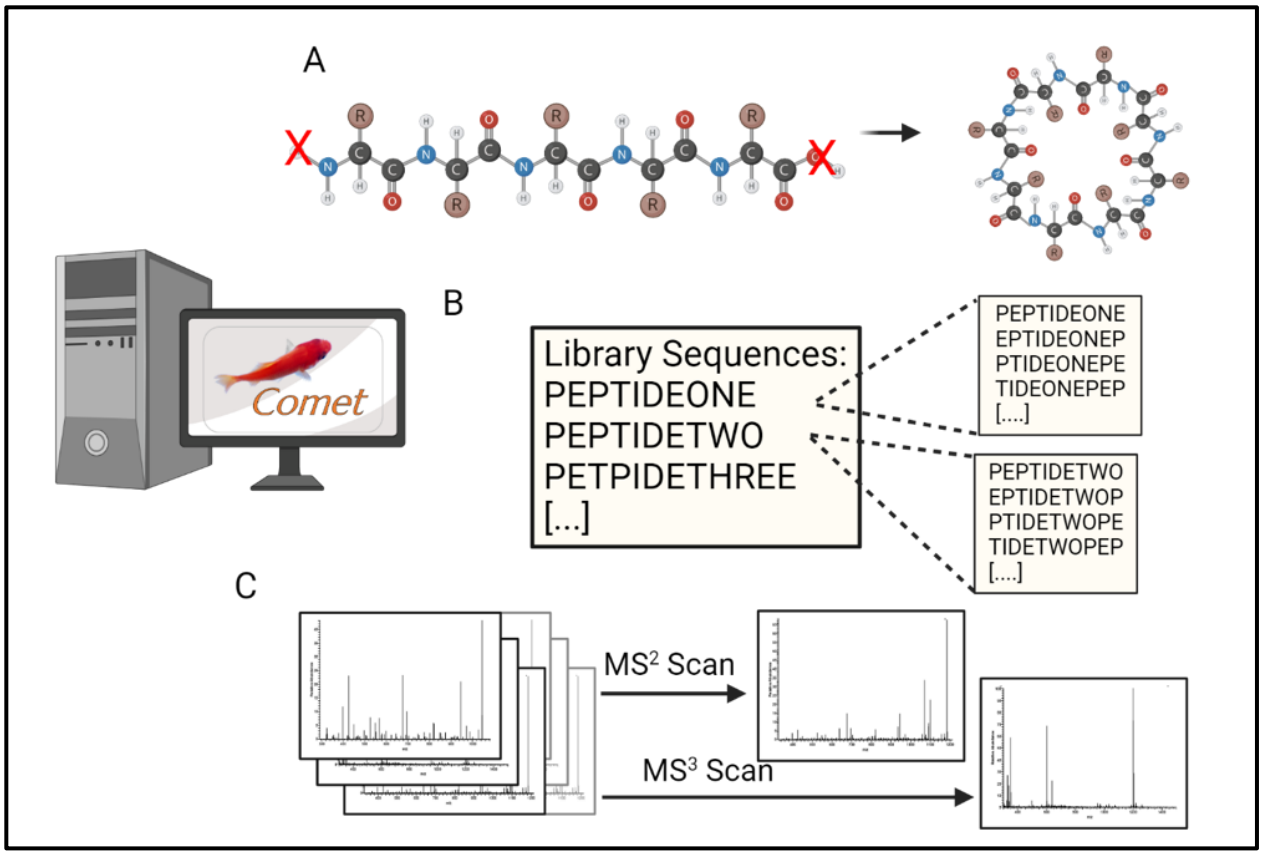


Figure 2.3. | Summary of changes made to Comet to enable Cyclic Comet

**A)** Peptide termini were removed from consideration when calculating peptide mass. **B)** Peptide sequences were cyclically permuted to encompass the full sequence space. **C)** Mass spectra is filtered by MS level for sequence matching.

| Peptide | | Scan Sequence | Xcorr | Next best | Norm. diff |
|---|---|---|---|---|---|
| PPIVPPBL | MS2 | PPIVPPBL | 3.37 | 3.11 | 0.34351811 |
| | MS3 | PPBLPPI | 1.64 | 1.22 | 1.44 |
| | MS2 | LPLVPPBP | 3.37 | 3.07 | 0.39248139 |
| | MS3 | PPBLPPI | 1.39 | 1.14 | 0.86 |
| VJBBVPPI | MS2 | VJBBVPPI | 3.12 | 2.54 | 1.09162479 |
| | MS3 | JBBVPPI | 1.55 | 0.75 | 4.54 |
| | MS2 | VJBBVPPI | 3.12 | 2.54 | 1.09162479 |
| | MS3 | BVPPIV | 1.20 | 0.79 | 2.27 |
| APPVAJJV | MS2 | AVPPAJJV | 3.37 | 2.66 | 1.41174378 |
| | MS3 | JVAPPAV | 1.24 | 1.00 | 0.97 |
| | MS2 | AVPPAJJV | 3.36 | 2.75 | 1.21820632 |
| | MS3 | JVAPPAV | 1.29 | 1.05 | 0.97 |
| VPPAJJAVP | MS2 | VPPAJJAVP | 3.94 | 2.81 | 1.48182239 |
| | MS3 | JJAVPV | 1.10 | 0.41 | 3.52 |
| | MS2 | VPPAJJAVP | 3.94 | 2.81 | 1.48182239 |
| | MS3 | PPAJJAV | 1.24 | 0.13 | 5.64 |
| LPJAJPLA | MS2 | LPPVJVLA | 3.96 | 3.65 | 0.47867108 |
| | MS3 | LJPLA | 1.03 | 0.93 | 0.40 |
| | MS2 | JPVBJPLA | 3.91 | 3.80 | 0.16174703 |
| | MS3 | LJPLA | 1.01 | 0.83 | 0.71 |

Table 2.1 | Cyclic Comet Results of $MS^n$ Sequencing Against Complex Libraries

Cyclic Comet was asked to identify five peptides from their $MS^2$ or $MS^3$ spectra. Two pairs of spectra were tested against a dummy library of 10,000 peptides. The identified scan sequence is reported and colored by accuracy, blue for correct sequence and composition, yellow for correct composition and incorrect sequence, and orange for incorrect composition. Cyclic Comet reported xcorr is reported for the assigned sequence as well as the next best assignment. The difference is normalized by the

standard deviations of the population of assignments for the scan and reported as a confidence statistic.

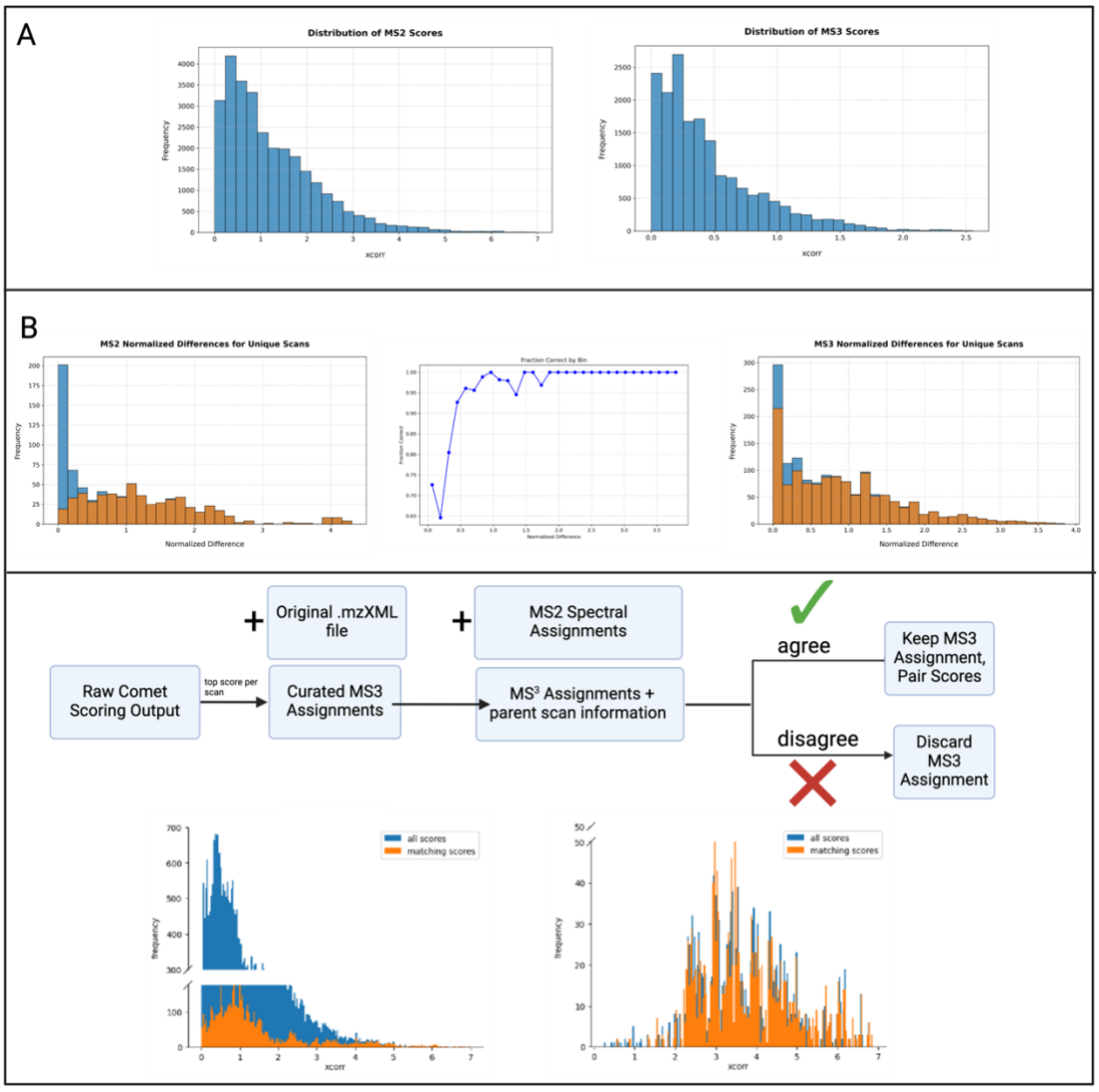


Figure 2.4 | Whole Library Sequencing of Mock ASMS Pool

40 peptides were pooled and treated as the results of a mock ASMS experiment. **A**) $MS^2$ xcorr and $MS^3$ xcorr are plotted for the plotted for the population, leading to standard deviations of 1.02 for $MS^2$ and 0.41 for $MS^3$. **B**) Normalized difference for each scan was plotted, along with the aggregate portion correct for each bin (middle). At

higher normalized differences, only correct assignments remain. **C**) A new filtering criterion, paired MS2/$MS^3$ sequences, significantly reduces the number of matches. The filtering logic is shown along with the impact on this search result.

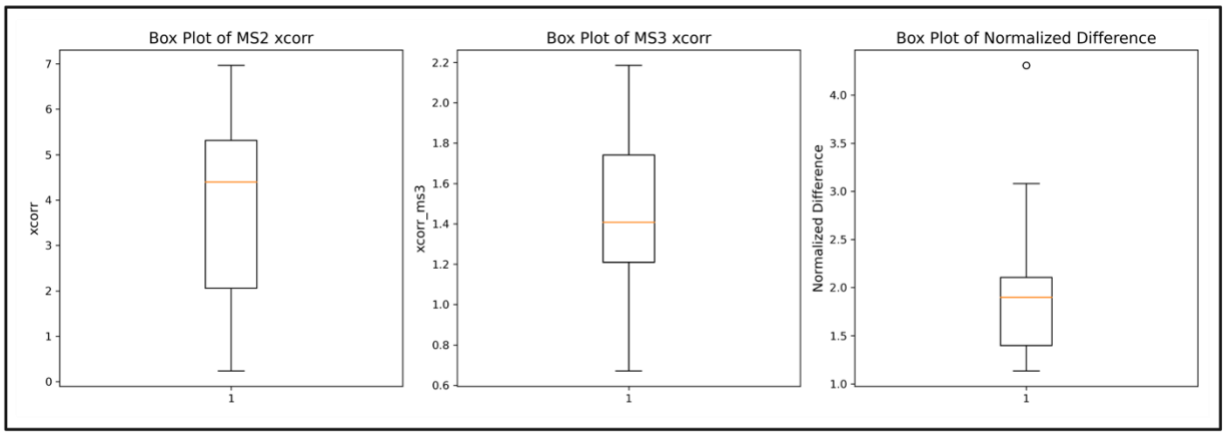


Figure 5 | Distribution of Scoring Metrics of Final Peptide Sequence Assignments

The three main evaluative metrics, $MS^2$ xcorr, $MS^3$ xcorr, and normalized difference were plotted for the sequence assignments that met final filtering criteria. Average xcorr scores were significantly higher than the general population.

## 2.7 BIBLIOGRAPHY

(1) Zorzi, A.; Deyle, K.; Heinis, C. Cyclic Peptide Therapeutics: Past, Present and Future. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24–29. https://doi.org/10.1016/j.cbpa.2017.02.006.

(2) Costa, L.; Sousa, E.; Fernandes, C. Cyclic Peptides in Pipeline: What Future for These Great Molecules? *Pharmaceuticals* **2023**, *16* (7), 996. https://doi.org/10.3390/ph16070996.

(3) Naylor, M. R.; Bockus, A. T.; Blanco, M.-J.; Lokey, R. S. Cyclic Peptide Natural Products Chart the Frontier of Oral Bioavailability in the Pursuit of Undruggable Targets. *Curr. Opin. Chem. Biol.* **2017**, *38*, 141–147. https://doi.org/10.1016/j.cbpa.2017.04.012.

(4) Russo, A.; Aiello, C.; Grieco, P.; Marasco, D. Targeting “Undruggable” Proteins: Design of Synthetic Cyclopeptides. *Curr. Med. Chem.* **2016**, *23* (8), 748–762.

(5) Tsomaia, N. Peptide Therapeutics: Targeting the Undruggable Space. *Eur. J. Med. Chem.* **2015**, *94*, 459–470. https://doi.org/10.1016/j.ejmech.2015.01.014.

(6) Drucker, D. J. Advances in Oral Peptide Therapeutics. *Nat. Rev. Drug Discov.* **2020**, *19* (4), 277–289. https://doi.org/10.1038/s41573-019-0053-0.

(7) Bashiruddin, N. K.; Suga, H. Construction and Screening of Vast Libraries of Natural Product-like Macrocyclic Peptides Using in Vitro Display Technologies. *Curr. Opin. Chem. Biol.* **2015**, *24*, 131–138. https://doi.org/10.1016/j.cbpa.2014.11.011.

(8) Chen, Y.; Duan, C.; Chen, K.; Sun, S.; Zhang, D.; Meng, X. Screening Technology of Cyclic Peptide Library Based on Gene Encoding. *Med. Drug Discov.* **2022**, *16*, 100145. https://doi.org/10.1016/j.medidd.2022.100145.

(9) Janda, K. D. Tagged versus Untagged Libraries: Methods for the Generation and Screening of Combinatorial Chemical Libraries. *Proc. Natl. Acad. Sci.* **1994**, *91* (23), 10779–10785. https://doi.org/10.1073/pnas.91.23.10779.

(10) Xiao, Q.; Pei, D. High-Throughput Synthesis and Screening of Cyclic Peptide Antibiotics. *J. Med. Chem.* **2007**, *50* (13), 3132–3137. https://doi.org/10.1021/jm070282e.

(11) Muchiri, R. N.; van Breemen, R. B. Affinity Selection–Mass Spectrometry for the Discovery of Pharmacologically Active Compounds from Combinatorial Libraries and Natural Products. *J. Mass Spectrom.* **2021**, *56* (5), e4647. https://doi.org/10.1002/jms.4647.

(12) Quartararo, A. J.; Gates, Z. P.; Somsen, B. A.; Hartrampf, N.; Ye, X.; Shimada, A.; Kajihara, Y.; Ottmann, C.; Pentelute, B. L. Ultra-Large Chemical Libraries for the Discovery of High-Affinity Peptide Binders. *Nat. Commun.* **2020**, *11* (1), 3183. https://doi.org/10.1038/s41467-020-16920-3.

(13) Ngoka, L. C. M.; Gross, M. L. Multistep Tandem Mass Spectrometry for Sequencing Cyclic Peptides in an Ion-Trap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (8), 732–746. https://doi.org/10.1016/S1044-0305(99)00049-5.

(14) Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. Automated Mass Spectrometric Sequence Determination of Cyclic Peptide Library Members. *J. Comb. Chem.* **2003**, *5* (1), 33–40. https://doi.org/10.1021/cc0200639.

(15) Kenneth Verheggen; Verheggen, K.; Helge Ræder; Ræder, H.; Frode S. Berven; Berven, F. S.; Martens, L.; Lennart Martens; Martens, L.; Harald Barsnes; Barsnes, H.; Marc Vaudel; Vaudel, M. Anatomy and Evolution of Database Search Engines-a Central Component of Mass Spectrometry Based Proteomic Workflows. *Mass Spectrom. Rev.* **2020**, *39* (3), 292–306. https://doi.org/10.1002/mas.21543.

(16) Kavan, D.; Kuzma, M.; Lemr, K.; Schug, K. A.; Havlicek, V. CYCLONE--a Utility for de Novo Sequencing of Microbial Cyclic Peptides. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (8), 1177–1184. https://doi.org/10.1007/s13361-013-0652-7.

(17) Ibrahim, A.; Yang, L.; Johnston, C.; Liu, X.; Ma, B.; Magarvey, N. A. Dereplicating Nonribosomal Peptides Using an Informatic Search Algorithm for Natural Products (iSNAP) Discovery. *Proc. Natl. Acad. Sci.* **2012**, *109* (47), 19196–19201. https://doi.org/10.1073/pnas.1206376109.

(18) Liu, W.-T.; Ng, J.; Meluzzi, D.; Bandeira, N.; Gutierrez, M.; Simmons, T. L.; Schultz, A. W.; Linington, R. G.; Moore, B. S.; Gerwick, W. H.; Pevzner, P. A.; Dorrestein, P. C. Interpretation of Tandem Mass Spectra Obtained from Cyclic Nonribosomal Peptides. *Anal. Chem.* **2009**, *81* (11), 4200–4209. https://doi.org/10.1021/ac900114t.

(19) Mohimani, H.; Liu, W.-T.; Kersten, R. D.; Moore, B. S.; Dorrestein, P. C.; Pevzner, P. A. NRPquest: Coupling Mass Spectrometry and Genome Mining for Nonribosomal Peptide Discovery. *J. Nat. Prod.* **2014**, *77* (8), 1902–1909. https://doi.org/10.1021/np500370c.

(20) Townsend, C.; Furukawa, A.; Schwochert, J.; Pye, C. R.; Edmondson, Q.; Lokey, R. S. CycLS: Accurate, Whole-Library Sequencing of Cyclic Peptides Using

Tandem Mass Spectrometry. *Bioorg. Med. Chem.* **2018**, *26* (6), 1232–1238. https://doi.org/10.1016/j.bmc.2018.01.027.

(21) Jesús Yagüe; Yagüe, J.; Alberto Paradela; Paradela, A.; Manuel Ramos; Ramos, M.; Manuel Ramos; Samuel Ogueta; Ogueta, S.; Anabel Marina; Marina, A.; Fernando Barahona; Barahona, F.; José A. Łópez de Castro; de Castro, J. A. L.; Jesús Vázquez; Vázquez, J. Peptide Rearrangement during Quadrupole Ion Trap Fragmentation: Added Complexity to MS/MS Spectra. *Anal. Chem.* **2003**, *75* (6), 1524–1535. https://doi.org/10.1021/ac026280d.

(22) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. Novel Peptide Dissociation: Gas-Phase Intramolecular Rearrangement of Internal Amino Acid Residues. *J. Am. Chem. Soc.* **1997**, *119* (24), 5481–5488. https://doi.org/10.1021/ja9640758.

(23) Saminathan, I. S.; Wang, X. S.; Guo, Y.; Krakovska, O.; Voisin, S.; Hopkinson, A. C.; Siu, K. W. M. The Extent and Effects of Peptide Sequence Scrambling Via Formation of Macrocyclic *b* Ions in Model Proteins. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (12), 2085–2094. https://doi.org/10.1016/j.jasms.2010.09.001.

(24) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406. https://doi.org/10.1002/1096-9888(200012)35:12<1399::AID-JMS86>3.0.CO;2-R.

(25) Tasoglu, C.; Arslanoglu, A.; Yalcin, T. Gas Phase Fragmentation Behavior of Proline in Macrocyclic B7 Ions. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (8), 1576–1583. https://doi.org/10.1021/jasms.3c00049.

(26) Béla Paizs; Paizs, B.; Sándor Suhai; Suhai, S. Fragmentation Pathways of Protonated Peptides. *Mass Spectrom. Rev.* **2005**, *24* (4), 508–548. https://doi.org/10.1002/mas.20024.

(27) Ramachandran, S.; Thomas, T. Characterization of the Fragmentation Pattern of Peptide from Tandem Mass Spectra. *Mass Spectrom. Lett.* **2019**, *10* (2), 50–55. https://doi.org/10.5478/MSL.2019.10.2.50.

(28) Eng, J. K.; Deutsch, E. W. Extending Comet for Global Amino Acid Variant and Post-Translational Modification Analysis Using the PSI Extended FASTA Format. *PROTEOMICS* **2020**, *20* (21–22), 1900362. https://doi.org/10.1002/pmic.201900362.

(29) Eng, J. K.; Jahan, T. A.; Hoopmann, M. R. Comet: An Open-Source MS/MS Sequence Database Search Tool. *PROTEOMICS* **2013**, *13* (1), 22–24. https://doi.org/10.1002/pmic.201200439.

(30) Eng, J. K.; Hoopmann, M. R.; Jahan, T. A.; Egertson, J. D.; Noble, W. S.; MacCoss, M. J. A Deeper Look into Comet—Implementation and Features. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (11), 1865–1874. https://doi.org/10.1007/s13361-015-1179-x.

(31) Eng, J. K.; Fischer, B.; Grossmann, J.; MacCoss, M. J. A Fast SEQUEST Cross Correlation Algorithm. *J. Proteome Res.* **2008**, *7* (10), 4598–4602. https://doi.org/10.1021/pr800420s.

(32) Bhardwaj, G.; O'Connor, J.; Rettie, S.; Huang, Y.-H.; Ramelot, T. A.; Mulligan, V. K.; Alpkilic, G. G.; Palmer, J.; Bera, A. K.; Bick, M. J.; Di Piazza, M.; Li, X.; Hosseinzadeh, P.; Craven, T. W.; Tejero, R.; Lauko, A.; Choi, R.; Glynn, C.; Dong, L.; Griffin, R.; Van Voorhis, W. C.; Rodriguez, J.; Stewart, L.; Montelione, G. T.; Craik, D.;

Baker, D. Accurate de Novo Design of Membrane-Traversing Macrocycles. *Cell* **2022**, *185* (19), 3520-3532.e26. https://doi.org/10.1016/j.cell.2022.07.019.

(33) Vaisar, T.; Urban, J. Gas-Phase Fragmentation of Protonated Mono-N-Methylated Peptides. Analogy with Solution-Phase Acid-Catalyzed Hydrolysis. *J. Mass Spectrom.* **1998**, *33* (6), 505–524. https://doi.org/10.1002/(SICI)1096-9888(199806)33:6<505::AID-JMS662>3.0.CO;2-1.

(34) Eckart, K.; Holthausen, M. C.; Koch, W.; Spiess, J. Mass Spectrometric and Quantum Mechanical Analysis of Gas-Phase Formation, Structure, and Decomposition of Various B2 Ions and Their Specifically Deuterated Analogs. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (10), 1002–1011. https://doi.org/10.1016/S1044-0305(98)00076-2.

(35) Alex G. Harrison; Harrison, A. G. Peptide Sequence Scrambling Through Cyclization of B5 Ions. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1776–1780. https://doi.org/10.1016/j.jasms.2008.06.025.

(36) Goloborodko, A. A.; Gorshkov, M. V.; Good, D. M.; Zubarev, R. A. Sequence Scrambling in Shotgun Proteomics Is Negligible. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (7), 1121–1124. https://doi.org/10.1007/s13361-011-0130-z.

(37) Yaginuma, K.; Aoki, W.; Miura, N.; Ohtani, Y.; Aburaya, S.; Kogawa, M.; Nishikawa, Y.; Hosokawa, M.; Takeyama, H.; Ueda, M. High-Throughput Identification of Peptide Agonists against GPCRs by Co-Culture of Mammalian Reporter Cells and Peptide-Secreting Yeast Cells Using Droplet Microfluidics. *Sci. Rep.* **2019**, *9* (1), 10920. https://doi.org/10.1038/s41598-019-47388-x.

# CHAPTER 3.

# Development of Affinity Selection Mass Spectrometry for Cyclic Peptides Against Complex Targets

Portions of this work were completed in collaboration with Alex Wiley (Totah Group, UW). More details can be found in Chapter Five of Wiley, A. M. Elucidating the Role of Two Inotropic Agents in Cardiac Hypertrophy: The Ceramides and Apelins. Doctoral Dissertation, University of Washington, UW Campus Repository, 2024[1].

Portions of the text in this chapter have been modified and reproduced with permissions from:

Bruce, A.; Adebomi, V.; Czabala, P.; **Palmer, J**.; McFadden, W. M.; Lorson, Z. C.; Slack, R. L.; Bhardwaj, G.; Sarafianos, S. G.; Raj, M. A Tag-Free Platform for Synthesis and Screening of Cyclic Peptide Libraries. *Angewandte Chemie International Edition* **2024**, *63* (21), e202320045. https://doi.org/10.1002/anie.202320045[2]

## 3.1 INTRODUCTION

Affinity selection mass spectrometry (AS-MS) is emerging as a powerful and efficient method in drug discovery. AS-MS involves incubating a mixture of compounds

with a target protein, followed by separating bound and unbound compounds through affinity capture techniques. Mass spectrometry then analyzes the bound compounds, providing detailed information on their identity and binding characteristics. This method has been particularly successful in peptide drug discovery, where mass spectrometry identification pairs well with fragmentation sequencing[3,4].

AS-MS offers several distinct advantages over traditional screening techniques. Its high throughput capability allows researchers to screen vast libraries of synthetic compounds in a single experiment, streamlining the hit identification process while minimizing resource consumption. Other screening methods, like surface plasmon resonance (SPR) or fluorescence polarization (FP) have their own advantages but do not enable the screening of large libraries of compounds. In contrast, AS-MS requires minimal protein material and is fast, highly sensitive, and capable of detecting low-abundance binding compounds that other techniques might overlook[5]. Rigorous AS-MS experiments can also provide information on binding affinity and stoichiometry, offering an understanding of the interaction strength and specificity. Additionally, AS-MS is adaptable in both protein target and assay requirements, allowing the use of a variety of affinity selection techniques such as size-exclusion chromatography or magnetic bead-based separation. This adaptability makes AS-MS invaluable in the early stages of drug discovery, particularly for challenging targets that require precise and robust screening methods.

AS-MS has recently become more popular and better developed. Though the technique has existed for a while, the Pentelute group has pioneered development of AS-MS for linear peptides over several studies[4,6]. Using the model protein 12ca5, which

recognizes a known amino acid motif, the Pentelute group has been able to optimize assays, push the limits of library size, and automate some of their discovery pipeline[4,6]. These principles have been applied to a variety of difficult and complex protein targets. They were optimized against like human papillomavirus 16 early protein 6 (HPV16 E6) and peptidyl-prolyl cis–trans isomerase NIMA-interacting 1 protein (Pin1) in order to find covalent peptide binders. In this example, where the wash steps were specifically tuned to rigorously wash away everything but the covalent binders and a special elution step was needed, the authors ended up calling this the ReAct AS-MS[7].

While AS-MS has seen a recent expansion of applications and methodologies, its use in G protein-coupled receptor (GPCR) screening remains an enticing challenge. GPCRs are notorious for their inability to fold outside the membrane and flexibility within the membrane, yet they remain a privileged therapeutic protein class[8–10]. Approximately 30–40% of all approved small molecule drugs target GPCRs but a large portion of GPCRs remain unmodulated[11]. Traditional GPCR screening techniques have been unable to fill the gap. Cellular activity assays are able to directly measure downstream signaling markers and inform on compound agonism and antagonism but require cell care and complex data interpretation which can make them slow and inconclusive. Alternative approaches will frequently attempt to solubilize the GPCR by making mutations to the protein but this can be fraught with method development and testing[12]. As a result, researchers are turning increasingly towards membrane-mimetics like nanodiscs for structure stabilization[13]. Some related work has been attempted using AS-MS to study GPCRs. One group has used vacuum filtration and lysed cell membranes to specifically retain GPCRs in their native state. These authors optimized their own AS-

MS protocol but had difficulty selectively retaining the protein and encountered buffer incompatibilities with mass spectrometry[14]. The inherent variability and difficulty in studying GPCR activity assays, juxtaposed with the need for properly folded purified recombinant protein in ligand interaction studies, underscores why GPCRs are notoriously difficult targets for conventional drug discovery approaches. These challenges necessitate innovative assay designs and screening strategies that can accommodate the unique structural and functional demands of GPCRs.

To address these challenges, virus-like particles (VLPs) can be used to provide properly folded protein in a simple buffer condition[15–17]. We selected the apelin receptor (APJ) as a model GPCR for our initial studies because it is a difficult target that is currently being explored in drug design due to its therapeutic potential[18–20]. Representative of its class, APJ is difficult to study yet may play an important role reducing hypertension as a treatment for heart failure. APJ and its endogenous ligands represent an important physiological signaling system imperative to cardiovascular homeostasis, yet it has no approved drug[21]. Heavy investment led to the development of two strong clinical candidates, MM07 and AMG 986, that eventually failed to progress due to a lack of clinically meaningful effects[22,23]. Both compounds were APJ agonists, encouraging specific activation of APJ towards reducing hypertension. Further investigations into an APJ agonist will require similarly heavy investment. APJ is difficult to study in terms of functional screening, and without mutations it cannot be reliably extracted from the cell membrane without altering protein folding. This makes APJ an excellent model GPCR for our study. Additionally, we included ADORA2A in our study as a well-studied GPCR that can be utilized as a control and benchmark for other

techniques. In terms of GPCRs, ADORA2A represents a validated drug target that is better understood and easier to work with. The two GPCRs, APJ and ADORA2A, provide divergent yet accessible test cases for experimental evaluation.

The structural and functional complexity of GPCRs presents significant challenges, necessitating innovative approaches like AS-MS to increase the speed of drug discovery against GPCRs. We hypothesize that VLPs represent a portable and tunable system for maintaining GPCR structural stability and folding and can be utilized in an AS-MS-based HTS technique for identifying therapeutics against GPCRs. The generation of GPCR-VLPs avoids the complexities of purifying GPCRs from their native bilayer and expands the current capabilities of AS-MS with correctly folded GPCRs in simple buffer systems. In this chapter, we discuss the development of a reliable and reproducible method for generating and purifying GPCR-VLPs. We demonstrate ligand specificity and selectivity with two different GPCR systems. We conclude this work by validating positive control binding enrichment when GPCR-VLPs are co-incubated with a complex mixture of thousands of peptides.

## 3.2 MATERIALS AND METHODS

### *Materials*

The detergent Tween-20 and salt guanidine hydrochloride were purchased from ThermoFisher (Waltham, MA). Bovine serum albumin was purchased from Millipore Sigma (Bellevue, WA). Chromatography solvents were purchased from Fisher Scientific (Hampton, NH). Peptide synthesis solvents including Piperidine, Trifluoroacetic Acid,

PyAOP, and N,N-Diisopropyletheylamine were purchased from Millipore Sigma (Bellevue, WA) with the exception of N,N-Dimethylformamide which was purchased from Fisher Scientific (Hampton, NH). Cl-TCP(Cl) ProTide resin was used for all synthesis and was purchased from CEM (Matthews, NC). Fmoc protected amino acids were purchased from P3 biosystems (Louisville, KY) or Chem-Impex (Wood Dale, IL).

*Synthesis of peptide library*

The peptide library was prepared using standard SPPS split and pool synthesis with Fmoc protected amino acids. An original scaffold of QSY[dPhe][dThr]WK[Wai][FGE][FBE]NVI was synthesized with the following combinatorial substitutions: dPro at the 5$^{th}$ position, Pro at the 6$^{th}$ position, FED, FCF, FEJ, and FCE at the 8$^{th}$ position, FCF, FCC, HAA, FEH, FDI at the 9$^{th}$ position, FDJ, FAB, FEI, BCD, WAG, and FDM at the 10$^{th}$ position, and Pro at the 12$^{th}$ position. Standard synthesis protocol was followed as detailed in Standard practices for Fmoc-based solid-phase peptide synthesis in the Nowick laboratory (Version 1.6.1) by Kreutzer and Salveson (2018).

*Development and purification of SNAP-GPCR proteins in VLPs*

The development of SNAP-APJ along with its co-expression with HIV-1 Gag-mCherry to generate GPCR-virus like particles will be described in detail in Chapter Five of Wiley, A. M. Elucidating the Role of Two Inotropic Agents in Cardiac Hypertrophy: The Ceramides and Apelins. Doctoral Dissertation, University of Washington, UW Campus Repository, 2024.[1] Extensive characterization and imagining using electron microscopy is found there as well.

*A magnetic bead-based protein capture*

After purification, each batch of Virus-Like Particles (VLPs) was subjected to protein normalization using a small-volume protein assay (Bio-Rad Laboratories, #500012). This step ensures that consistent protein amounts are used across experiments, which is critical for reproducibility and accurate data interpretation. The normalized VLP samples were then introduced to 40 µL of pre-equilibrated SNAP-Capture Magnetic Beads in low protein binding microcentrifuge tubes. The beads serve as a solid support to capture and hold the VLPs, facilitating downstream analysis. These samples were gently mixed on a rotating tube mixer for a minimum of 1 hour at 4°C to ensure adequate interaction between the VLPs and the beads, promoting efficient binding. The amount of protein added per experiment ranged between 2.75 µg and 5.5 µg, contingent on batch yield and the required signal intensity for mass spectrometry analysis. Typically, higher protein content correlated with improved MS signal quality, ensuring more reliable detection of interactions.

Blocking buffer was prepped for use throughout as 1 mg/mL BSA and 0.02% Tween 20 in PBS. Post-incubation with the magnetic beads, the samples were isolated using a magnetic rack and subjected to three washes with 500 µL of blocking buffer in order to remove nonspecific interactions and contaminants. These washes utilized Tween 20 as a detergent to stabilize the VLPs and protein and BSA to fill hydrophobic nonspecific binding patches. For a full magnetic bead-based screening protocol, see Appendix A.

*Target library incubation and elution*

After blocking, magnetic bead-VLP samples were incubated with the respective test conditions. All test solutions were prepared in blocking buffer, and each 100 µL incubation consisted of 1 µM known binding partner along with equivalent concentrations of the test compounds. These incubations were performed with light agitation on the rotating tube mixer for 1 hour at 4°C. Post incubation, samples were rapidly washed three times with 200 µL of ice-cold PBS to wash away non-specific binding. The bound peptides were eluted twice with 25 µL of 6 M guanidine hydrochloride while undergoing constant agitation.

*Size Exclusion Chromatography Affinity Selection Mass Spectrometry (SEC ASMS)*

0.13 nmol of 12ca5 protein was dissolved in 100 µL of PBS at pH 7.4 and incubated with a peptide library prepared at a concentration of approximately 400 nM per library member. To prepare the library, 24 mg of the peptide library was dissolved in 5 mL of 5% DMSO in PBS at pH 7.5. The mixture was incubated at room temperature for 30 minutes, allowing the peptide ligands to interact with 12ac5. This incubation step was performed under gentle agitation to facilitate binding while maintaining the integrity of the VLPs.

Following the incubation, the mixture was loaded onto a High-Performance Size Exclusion Chromatography (HPSEC) column for separation of the protein-ligand complexes from unbound components. The HPSEC was performed using a 3 µm

Agilent SEC column 7.8 x 150 mm. The chromatography was conducted with an isocratic mobile phase of PBS at pH 7.5, with a flow rate of 1 mL/min.

The peptide library mixed with the target protein (100 µL of the mixture) was injected into the HPLC system. The run was carried out for 15 minutes, with the target protein-binder complexes eluting at approximately 4.155 minutes. The separation was monitored using UV detection at 214 nm to track the elution of the protein-ligand complexes.

Fractions corresponding to the expected elution time of the complexes were collected and pooled for further analysis. The SEC method effectively separated the bound complexes, which were then eluted using 6 M guanidine hydrochloride to disrupt the protein-ligand interactions, allowing for the release of the bound ligands. The collected samples were subsequently prepared for mass spectrometry analysis, ensuring high specificity and sensitivity in detecting the target interactions. After the run, the SEC column was thoroughly cleaned with deionized water to remove any residual contaminants and prepare it for subsequent runs. For a full size exclusion chromatography screening protocol, see Appendix B.

*High-Performance Liquid Chromatography – Mass Spectrometry (HPLC-MS)*

The eluted samples were either immediately chilled to 8°C in a Waters UPLC M-class autosampler or stored at 4°C for up to 24 hours before analysis. C18 prep spin columns were cut to avoid sample loss and replaced with directly dilution at 1:1 cold buffer and guanidine hydrochloride (GnHCl) before proceeding to HPLC-MS analysis.

Liquid chromatography (LC) methods were adjusted to include a diversion of specific parts of the run to prevent salts from reaching the detector.

Separation was performed using a heated Waters BEH C18 Column, 130Å, 2.5 µm, 2.1 mm x 50 mm (Waters, #186006029) for apelin separation, while an InfinityLab Poroshell 120 HILIC, 2.1 x 150 mm, 4 µm was utilized for adenosine separation. The mobile phases consisted of Optima water with 0.1% formic acid (Buffer A) and acetonitrile with 0.1% formic acid (Buffer B). A standard gradient of 0-100% B over 15 minutes was employed for apelin separation, while adenosine was separated using an isocratic gradient of 100% B.

For assays using known controls, MRM methods were developed on a Waters Xevo TQ-S-2 with MassLynx 4.1to quantify apelin, adenosine, and angiotensin II. Precursor and fragment ions for each target were identified and these transitions were programmed into the instrument. Collision energies and other parameters were automatically optimized by the instrument for each compound. Using an MRM method gave sensitive readings at low abundance and kept spectra clean from potentially messy screening samples.

Mass spectrometry data was collected, and peak integration was performed using either a Waters Xevo TQ-S-2 with MassLynx 4.1 or a Thermo Fisher Orbitrap Ascend Tribrid with Xcalibur V4.1. Compound retention was quantified using mass spectra. Multiple Reaction Monitoring (MRM) methods were developed on the Waters Xevo for both adenosine and pyr-AP-13, offering increased sensitivity and dynamic

range. Injecting dilution gradients of each compound yielded linear response curves to assess MS signal response.

Data from the mock library experiment was acquired using the ThermoFisher Tribrid Orbitrap Ascend in Data Independent Acquisition (DIA) mode, allowing comprehensive mass tracking at the $MS^1$ level. To minimize injection variability, internal standards were incorporated into the quench buffer. Initially, angiotensin II was selected as an internal standard due to its similar characteristics to apelin, but its poor retention under HILIC conditions and its role as a test compound necessitated the inclusion of a secondary standard. The in-house peptide 8.11, composed primarily of hydrophobic D-amino acids ([dVal][dLeu]A[dAla][dVal]P[dPro]I) with a mass of 802.5, was chosen. Peptide 8.11 was selected for its consistent retention in both C18 and HILIC chromatography and its lack of interaction with the target protein or other compounds of interest, ensuring accurate normalization across experimental runs.

*Automated analysis of mock ASMS*

Automated data processing of the 12ca5 and APJ library screens were performed as detailed in Chapter 1.

## 3.3 RESULTS

*Development of in-house ASMS protocol using 12ca5 & CyClick chemistry*

To establish AS-MS within our laboratory, we adapted protocols for cyclic peptide screening and collaborated with the Raj group at Emory University (Atlanta, GA) to screen a cyclic peptide library against the model protein 12ca5 (**Figure 1**). 12ca5 has an extensive history in ASMS development as a model protein and recognizes a known DXXDY motif[4,6]. A library of 10,000 peptides was synthesized around the DXXDY motif for the purpose of screening against 12ca5 by the Raj group and was shipped to the University of Washington for screening and analysis.

The cyclic peptides were incubated with 12ca5 and then subjected to size-exclusion chromatography to separate unbound peptides from those bound to the protein (**Figure 1A**). The bound peptides were eluted with a 6M GnHCl wash and analyzed by mass spectrometry. A C18 prep spin column helped to desalt our eluted fraction and to prevent any leftover target protein from contaminating our sample. Peptide mass features had their enrichment calculated in order to determine which masses needed sequencing (**Figure 1B**). Sequencing of the identified mass feature was performed automatically with Cyclic Comet and validated manually as in **Figure 1C**. In addition, half of the screened sample was retained to be chemically linearized. $MS^2$ fragmentation was used to sequence the linearized and derivatized peptides. Only one peptide had agreeing results from Cyclic Comet and the chemically linearized pool, peptide XDAGPYDAYA. Lastly, the identified peptide was resynthesized in the Raj group and their binding was confirmed using microscale thermophoresis[2]. In order to confirm that ASMS screening identified a peptide with specific target affinity and didn’t

exploit the common DXXDY motif, a peptide that did not show enrichment was selected from microscale thermophoresis as well. There was no binding observed for the random negative control (**Figure 1D**).

*GPCR-VLP validation using controlled ASMS*

We aimed to apply ASMS to GPCRs, a challenging yet valuable class of therapeutic targets. Collaborating with the Totah group, we produced virus-like particles (VLPs) displaying GPCRs on their surface (**Figure 2**). HEK293T cells were used to overexpress the HIV Gag polyprotein and snap-tagged APJ. The Gag polyproteins induced the budding of small membrane particles that included the overexpressed APJ GPCR. After media collection and a three-centrifugation step procedure, we confirmed the presence of APJ through western blot analysis. Coomassie staining further validated the presence of expected protein bands in both APJ- and ADORA2A-VLP sample (**Figure 2A**). Cryogenic electron micrography visualization confirmed the formation of virus-like particles (**Figure 2B**). These images showed the VLPs formed consistent spherical shapes. Also visible is the lattice structure that Gag polyprotein creates upon formation of the VLP. GPCRs on the VLP surface were captured using covalent magnetic beads, optimized with a specific washing protocol to retain binding-competent proteins. Tagged VLPS were retained specifically by the magnetic support through an irreversible covalent SNAP linker (**Figure 2C**).

Structural validation of APJ in VLPs was demonstrated through ASMS screening. Ligand incubation and pulldown assays aided development of ASMS protocols and provided support for the binding competence of APJ. In a series of stepwise experiments, we demonstrated specific binding interactions capable of enriching [pyr]-

Apelin-13 (pAp13) in eluted solutions. We showed first that APJ was able to enrich pAp13 from using an ASMS-like bead-based pulldown assay (**Figure 3A**). Second, we demonstrated specific retention of pAp13 by APJ-VLPs, with no retention of the nonspecific control angiotensin II (AngII), confirming the binding specificity of APJ in our assay (**Figure 3B**).

Lastly, we developed a second GPCR-VLP using the protein ADORA2A (A2A). We cloned and expressed A2A into cells in the same manner as APJ and confirmed its viability with ASMS screens. Total protein levels with A2A- and APJ-VLP samples were normalized, and equivalent amounts were incubated with SNAP-beads. Equimolar amounts of adenosine and PA13 in excess of the protein were co-incubated with A2A and APJ, and VLP-negative samples. A2A was used as a negative control in our tests, and we were able to demonstrate that pAp13 was enriched by APJ at a greater magnitude than A2A (**Figure 3C**). A2A was separately confirmed as being binding competent by screening with its own endogenous ligand, adenosine. APJ served as the negative control for A2A tests (**Figure 4**). Both A2A and APJ showed increased retention of the other's endogenous ligand when compared to the VLP-negative control, indicating a base level of nonspecific interaction with the VLP membrane and surface proteins present. This is intuitive, as the VLPs introduce a large and hydrophobic surface for nonspecific retention. This base level of nonspecific interaction increases the difficulty of identifying significant differences in binding retention, as evidenced by PA13's enrichment factor in APJ falling below the level of statistical significance (**Figure 3**). It is worth noting that the typically better behaved GPCR A2A retained its binding partner with a greater significance than APJ (**Figure 3 & 4**).

*Mock ASMS using GPCR-VLPs*

Encouraged by the retention of co-incubated PA13, we advanced our co-incubation to feature PA13 and a mock screening library (“dummy library”). This dummy library was utilized to mimic high throughput screening conditions and further investigate nonspecific retention of compounds. PA13 was spiked into a library of 2,688 13 amino acid peptides at equimolar concentration, giving a final library of 2,689 peptides with one positive control for APJ. As a last check, MS raw signal counts of PA13 and library members were compared, and similar ionization efficiency was observed for all compounds. We then incubated this test library with APJ- and A2A-VLPs, with A2A serving as our specific interaction negative control. A full ASMS experiment was ran with GPCR-VLPs using the library incubation in lieu of singular controls (**Figure 6A**).

In contrast to previous experiments, mass spectrometry data collection was gathered in a data independent manner to track the presence of all peptides. To identify peptidic mass features within our dataset, we utilized a custom-made enrichment calculator. This program was tasked with calculating the peak area of each mass in the library across all MS files and finding the enrichment between sample and control runs. Sample scans that fragmented the identified mass features were isolated and sequenced with Cyclic Comet. This method was integrated into the Cyclic Comet search for a fully automated analysis pipeline. By providing the MS .mzXML files and the library sequences, the automated pipeline is capable of identifying enriched features and isolating those features from the total population of Cyclic Comet sequencing for interpretation (**Figure 5**).

A binding enrichment ratio was calculated by dividing the enrichment of the peptide in the APJ sample by the enrichment in the A2A sample (**Figure 6C**). Each of these values were normalized with internal standards and final peak area ratio (PAR) was used to calculate enrichment. This facilitated interpretation for specific retention by APJ. Due to the low complexity of this library and unique mass of pAp13, Comet was not needed for analysis. We were able to identify pAp13 by parent mass alone and no other compounds were significantly retained for Cyclic Comet sequencing. We found that most peptides from the test library were retained slightly more in A2A samples than APJ samples, with an average binding enrichment ratio around 0.8 (**Figure 6B**). pAp13 had a binding enrichment ratio greater than 2, which is often used to separate specific and nonspecific retention [24]. No other test peptides had an enrichment factor above 1.5. (**Figure 6B**). We conclude that not only was our positive control pAp13 enriched in APJ samples compared to A2A samples, but that our wash steps were effective in limiting the nonspecific carryover of the other library members into the final MS sample. This work demonstrates the feasibility of our GPCR-VLP-ASMS approach in a library-based screening setting.

## 3.4 DISCUSSION

In this study, we have demonstrated the development and implementation of a novel high-throughput screening platform using Affinity Selection Mass Spectrometry (AS-MS) tailored for G protein-coupled receptors (GPCRs). Our approach leverages Virus-Like Particles (VLPs) to maintain the proper folding of GPCRs, eliminating the

need for extensive buffer optimization and validating proper GPCR folding, which are often significant challenges in traditional GPCR studies. By tagging VLPs with SNAP-tags and capturing them on solid-support magnetic beads, we have established a robust and adaptable system that enhances the speed and efficiency of AS-MS for GPCRs, a class of proteins traditionally difficult to work with.

Our methodology relies on the fundamental principle of AS-MS: the affinity interaction between a target protein and a compound of interest, isolating active compounds for MS analysis. The identification of active compounds is typically quantified by an enrichment factor, representing the ratio of the compound in the sample compared to a control. Through this study, we used this metric to benchmark the performance of our assays, focusing on two distinct GPCRs: A2A and APJ. These receptors provided divergent test cases—A2A as a well-established drug target and APJ as a more challenging receptor with significant therapeutic potential.

Initial attempts to screen against GPCRs using AS-MS without the VLP platform were met with limited success, primarily due to the inability to maintain proper GPCR folding outside their native membrane environment. The aggressive use of detergents and other stabilizing agents proved inadequate, often resulting in destabilized proteins and poor AS-MS results. Additionally, these stabilizing agents were incompatible with mass spectrometry, complicating sample preparation and leading to potential analyte loss. These challenges led us to adopt the VLP approach, which provided a more stable and effective platform for AS-MS experimentation.

Using the VLP platform, we demonstrated enriched signal detection for the natural binding partner PA13 with APJ-VLPs compared to controls. This enrichment was observed even when APJ-VLPs were co-incubated with a library of 2,688 peptides, showcasing the specificity and selectivity of our approach. The SNAP-tag system enabled the covalent capture of GPCR-VLPs on magnetic beads, providing a solid support structure that was crucial for the downstream AS-MS analysis. We further validated our system by co-incubating APJ-VLPs with native peptide ligand PA13 and angiotensin II (AngII), observing specific enrichment of PA13, which confirmed the proper folding and functional activity of APJ within the VLPs.

Our results also highlight the importance of minimizing nonspecific interactions to enhance the accuracy of AS-MS. By introducing a control GPCR (A2A) and normalizing protein levels, we were able to compare the enrichment of binding partners and further refine our assay's specificity. Although the well-characterized A2A showed better retention of its binding partner compared to APJ, which did not reach statistical significance in some cases, the data nonetheless support the potential of our platform for high-throughput screening.

To rigorously test our optimized protocol, we challenged it with a mock screening library (dummy library) spiked with the positive control PA13. The successful enrichment of PA13 in APJ samples, coupled with the effective exclusion of nonspecific peptides, demonstrated the feasibility of our GPCR-VLP-AS-MS approach in a library-based screening setting. This experiment underscores the capability of our method to

differentiate specific interactions even in the presence of complex mixtures, a key requirement for practical drug discovery applications.

The development of an automated data analysis pipeline should greatly speed the ability to sequence and identify peptides from ASMS experiments. The key development in this chapter, the use of the enrichment calculator, is an approach used in ASMS screening to identify enriched compounds for analysis. Though not demonstrated here, this has the added benefit of reducing the need for the Cyclic Comet-based filtering that struggled in Chapter Two. The only downside of this additional step is the added calculation time. Enrichment calculation is a slower step than Cyclic Comet sequencing, and we highly recommend parallelizing chunks of peptide sequencing for enrichment calculation rather than running everything at once.

Enrichment calculation serves as an important step in making sure that peptides are retained at a level that demonstrates target interaction. However, it is a known caveat of ASMS that it can be insensitive to weak binding compounds. Previous studies have benchmarked ASMS and found that binders with high micromolar affinity will be lost during the various wash and prep steps[4,6]. Combined with the naturally faster on/off rates of GPCR binders, we sought to minimize loss by keeping our samples and washes cooled and as quick as possible. pAp13 retention struggled at times despite its high affinity of 0.33 nM for APJ, which reflects the challenge of screening against APJ. Added to this complexity is the inherent variability introduced by the use of VLPs, as not all VLPs in a batch may contain the target SNAP-GPCR. This variability was mitigated through the use of additional control samples and optimized wash steps, yet it remains

a consideration for future assay development. Moreover, the enrichment factors observed suggest potential nonspecific sequestration of compounds within the VLPs and solution proteins, an area that could benefit from further refinement.

## 3.5 CONCLUSION

In conclusion, we have developed a platform for high throughput screening of compounds for target interaction against GPCRs by standardizing GPCR production in VLPs. We have demonstrated that we can link GPCR expressing VLPs to solid support beads and observe enriched binding with positive controls in both simple and complex solutions. By presenting the potential of GPCR-VLPs as a screening tool for identifying and shaping new therapeutics, we hope to accelerate and unlock compound screening against GPCRs at a level previously unreachable. The integration of GPCR-VLPs in ASMS will make targeting specific GPCRs quicker and easier and will greatly improve the capability of drug screening today, as well as advance the development of GPCR-targeted therapeutics. Continued refinement and testing with diverse libraries and GPCRs will likely further enhance the utility and precision of our GPCR-VLP-AS-MS approach, ultimately contributing to more efficient and targeted drug discovery efforts.

## 3.6 FIGURES

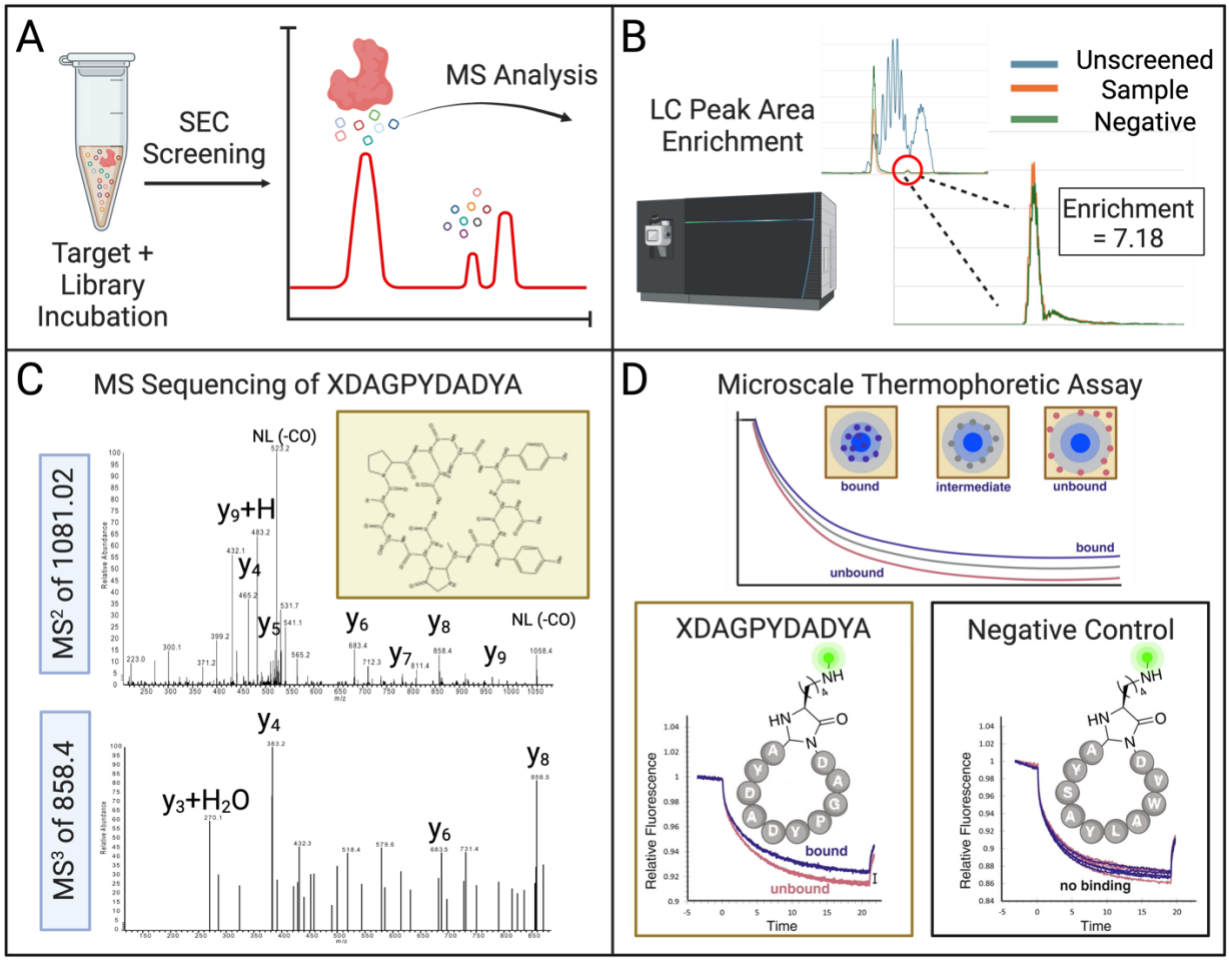


Figure 3.4 | SEC-ASMS of Library Against Model Protein 12ca5.

The model protein 12ca5 was used to develop the SEC-ASMS workflow and discover a novel cyclic peptide binder, peptide XDAGPYDAYA. **A**) The peptide library is incubated with the target protein and non-interacting peptides are separated from the protein fraction by size exclusion chromatography. **B**) MS analysis uses an enrichment search to ID mass features that are retained by the protein. This mass, 1081.02, was enriched 7 times over the blank. **C**) $MS^2$ and $MS^3$ fragmentation is used to obtain the peptide sequence for peptide 1081.02. Spectra is annotated with y ions for the matching sequence. X denotes the custom CyClick amino acid. **D**) Confirmation of binding

interaction with the target protein using microscale thermophoresis. The identified peptide is compared to a negative control from the library. Parts of this figures is adopted with permission from Bruce et al. (2024)[2].

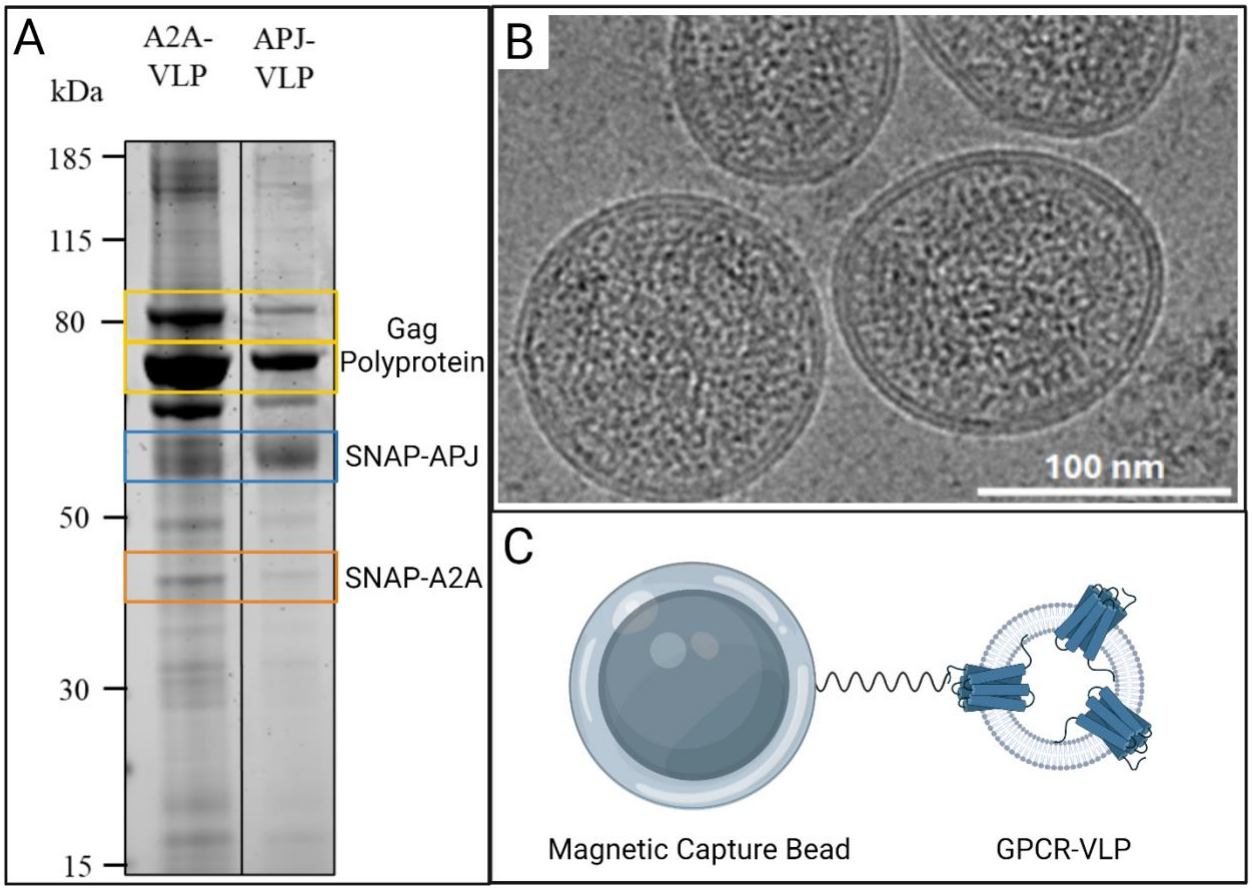


Figure 3.5 | Characterization of GPCR-VLPs for ASMS

**A**) Coomassie staining was used to show that a significant amount of protein present in expressed VLP samples were GPCR protein. Also notable is the high amount of Gag polyprotein, which is present in two oligomerization states. **B**) Virus-like particles are imaged by cryogenic electron micrography. The VLPs were sized as expected and formed consistent spherical shapes. **C**) Diagram of the GPCR-VLP being tagged to the solid support bead. Not to scale. Beads are much larger than VLPs, and we would expect that multiple VLPs dot the surface of a single bead. The bead-GPCR connection is formed via covalent linkage using New England Biolab's SNAP tag (Ipswich, MA). This data is adopted with permission from Wiley (2024)[1].

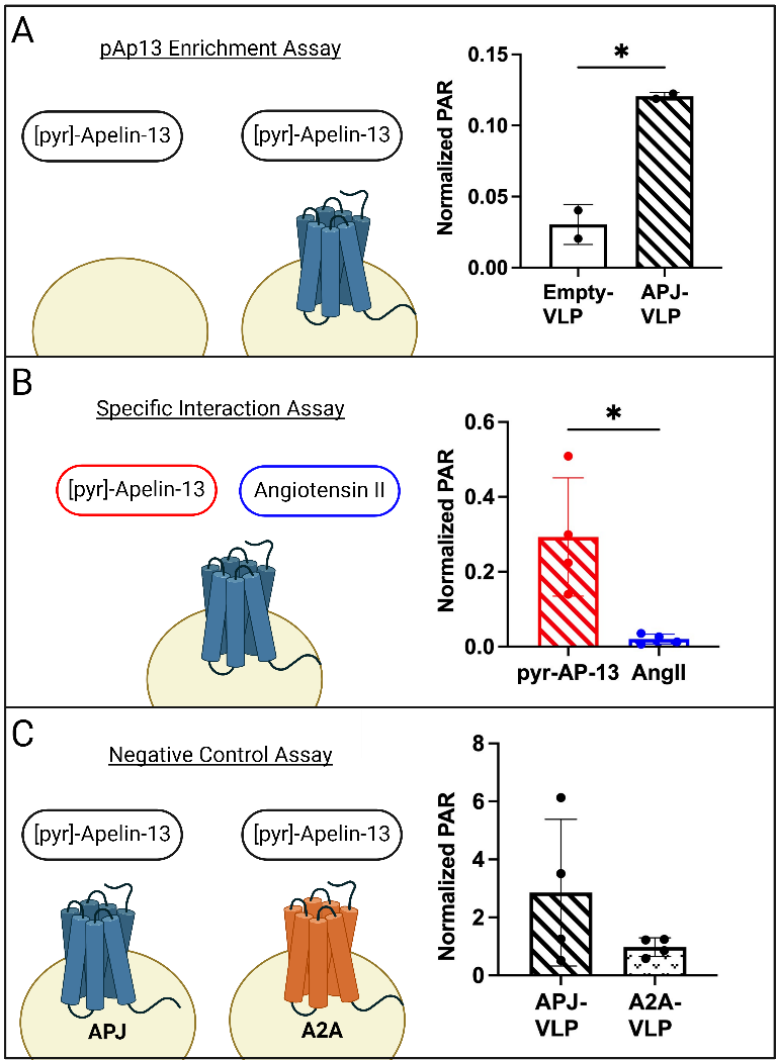


Figure 3.6 | Validation of APJ Structure Using ASMS

A series of ASMS experiments were performed to sequentially to determine if APJ was present in a binding-competent state. **A**) pAp13 was enriched using the VLP-GPCRs against a control of just beads. **B**) Specific retention of pAp13 was shown by enriching pAp13 more than Angiotensin II, a peptide which should have no interaction with APJ. **C**) A second GPCR, Adora2A, was used to strengthen our negative controls. pAp13 was significantly enriched from a co-incubation with adenosine in the APJ sample but not the A2A sample. This work done in collaboration with Wiley, 2024[1].

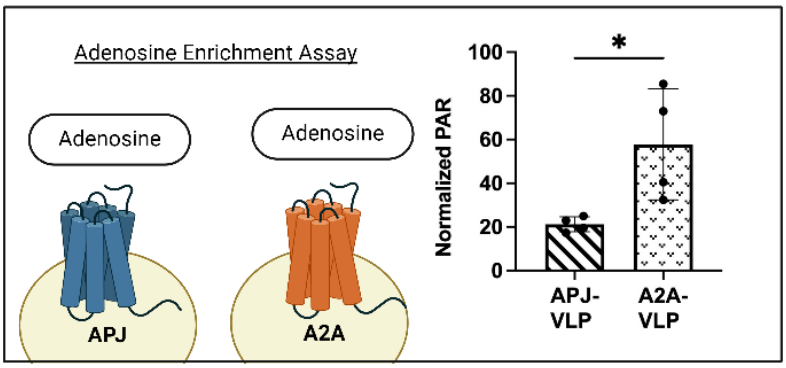


Figure 3.7 | Adenosine Enrichment Using Adora2A-VLPs

ASMS was used to test A2A for binding competency. Adenosine, A2A's native small molecule ligand, was enriched from a co-incubation with pAp13 in A2A but not in APJ. This work done in conjunction with Wiley, 2024[1].

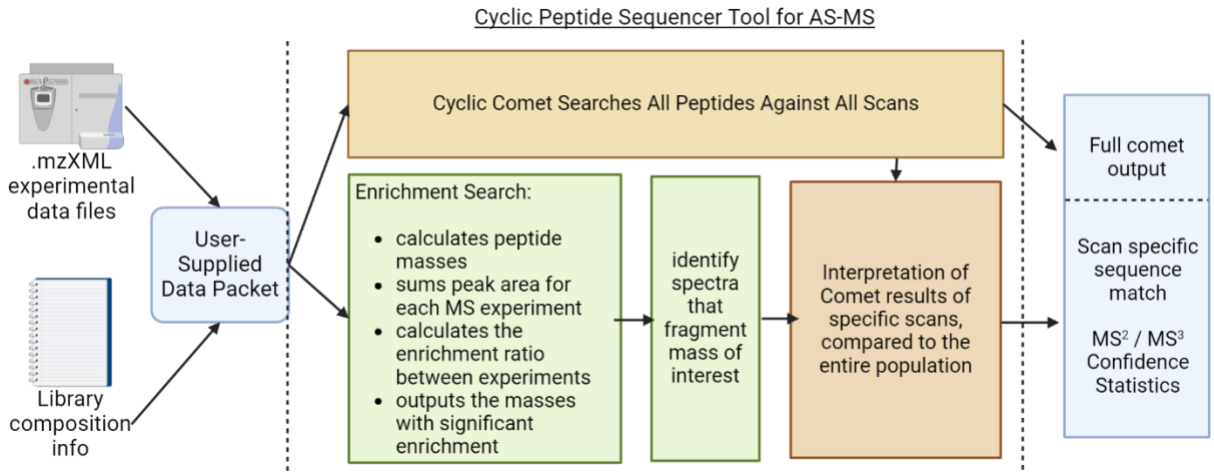


Figure 3.5 | Pipeline of Automated Data Processing for Cyclic Peptide ASMS

User provided experimental .mzXML files and library composition information is combined and input as a user-supplied data packet. While Cyclic Comet searches all sequences against all scans, a search finds enriched mass features and extracts related sequences scans for the output.

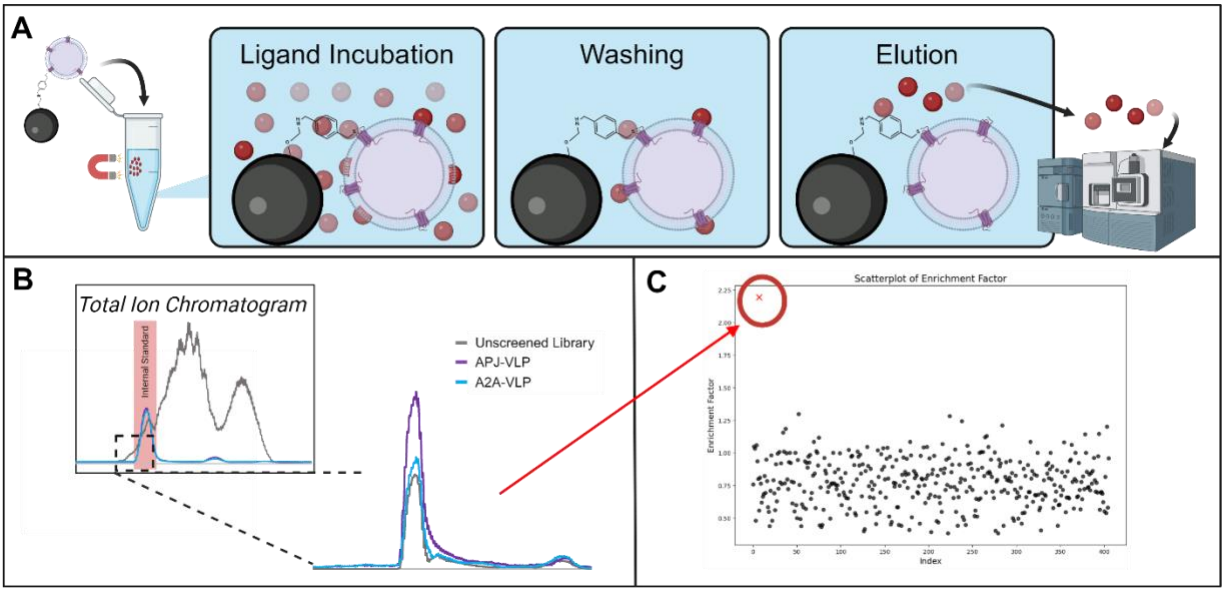


Figure 3.6. | Mock GPCR-VLP-ASMS Screening and Analysis

**A**) Generalized workflow for bead-based ASMS using GPCR-VLPs, including incubating the library with the target protein, stringent washing to remove nonspecific binding peptides, and eluting the remaining peptides with salt or acid. The eluted sample is then ready for MS analysis. **B**) Each mass within the library has their peak area recorded and the enrichment for each mass between sample and control runs is calculated. The positive control pAp13 is identified here as being positively enriched. **C**) The enrichment factor of all observed mass features is plotted. pAp13 has an enrichment of about two, while most of the other library peptides are below one. This work done in conjunction with Wiley, 2024[1].

## 3.7 BIBLIOGRAPHY

(1) Wiley, A. M. Elucidating the Role of Two Inotropic Agents in Cardiac Hypertrophy: The Ceramides and Apelins. Doctoral Dissertation, University of Washington, UW Campus Repository, 2024.

(2) Bruce, A.; Adebomi, V.; Czabala, P.; Palmer, J.; McFadden, W. M.; Lorson, Z. C.; Slack, R. L.; Bhardwaj, G.; Sarafianos, S. G.; Raj, M. A Tag-Free Platform for Synthesis and Screening of Cyclic Peptide Libraries. *Angewandte Chemie International Edition* **2024**, *63* (21), e202320045. https://doi.org/10.1002/anie.202320045.

(3) Lam, K. S. Affinity Selection and Sequencing. *Nat Chem Biol* **2019**, *15* (4), 320–321. https://doi.org/10.1038/s41589-019-0253-2.

(4) Zhang, G.; Li, C.; Quartararo, A. J.; Loas, A.; Pentelute, B. L. Automated Affinity Selection for Rapid Discovery of Peptide Binders. *Chem. Sci.* **2021**, *12* (32), 10817–10824. https://doi.org/10.1039/D1SC02587B.

(5) Muchiri, R. N.; van Breemen, R. B. Affinity Selection–Mass Spectrometry for the Discovery of Pharmacologically Active Compounds from Combinatorial Libraries and Natural Products. *Journal of Mass Spectrometry* **2021**, *56* (5), e4647. https://doi.org/10.1002/jms.4647.

(6) Quartararo, A. J.; Gates, Z. P.; Somsen, B. A.; Hartrampf, N.; Ye, X.; Shimada, A.; Kajihara, Y.; Ottmann, C.; Pentelute, B. L. Ultra-Large Chemical Libraries for the Discovery of High-Affinity Peptide Binders. *Nat Commun* **2020**, *11* (1), 3183. https://doi.org/10.1038/s41467-020-16920-3.

(7) Zhang, P.; Ye, X.; Wang, J. C. K.; Baddock, H. T.; Jensvold, Z.; Foe, I.; Loas, A.; Eaton, D. L.; Hao, Q.; Nile, A. H.; Pentelute, B. L. Reversibly Reactive Affinity Selection Mass Spectrometry Enables Maturation of Covalent Peptides. ChemRxiv December 22, 2023. https://doi.org/10.26434/chemrxiv-2023-2ts17.

(8) Guo, S.; Zhao, T.; Yun, Y.; Xie, X. Recent Progress in Assays for GPCR Drug Discovery. *Am J Physiol Cell Physiol* **2022**, *323* (2), C583–C594. https://doi.org/10.1152/ajpcell.00464.2021.

(9) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat Rev Drug Discov* **2017**, *16* (12), 829–842. https://doi.org/10.1038/nrd.2017.178.

(10) Zhang, R.; Xie, X. Tools for GPCR Drug Discovery. *Acta Pharmacol Sin* **2012**, *33* (3), 372–384. https://doi.org/10.1038/aps.2011.173.

(11) Sriram, K.; Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol Pharmacol* **2018**, *93* (4), 251–258. https://doi.org/10.1124/mol.117.111062.

(12) Laeremans, T.; Sands, Z. A.; Claes, P.; De Blieck, A.; De Cesco, S.; Triest, S.; Busch, A.; Felix, D.; Kumar, A.; Jaakola, V.-P.; Menet, C. Accelerating GPCR Drug Discovery With Conformation-Stabilizing VHHs. *Front. Mol. Biosci.* **2022**, *9*, 863099. https://doi.org/10.3389/fmolb.2022.863099.

(13) Köck, Z.; Ermel, U.; Martin, J.; Morgner, N.; Frangakis, A. S.; Dötsch, V.; Hilger, D.; Bernhard, F. Biochemical Characterization of Cell-Free Synthesized Human B1 Adrenergic Receptor Cotranslationally Inserted into Nanodiscs. *Journal of Molecular Biology* **2022**, *434* (16), 167687. https://doi.org/10.1016/j.jmb.2022.167687.

(14) Yaginuma, K.; Aoki, W.; Miura, N.; Ohtani, Y.; Aburaya, S.; Kogawa, M.; Nishikawa, Y.; Hosokawa, M.; Takeyama, H.; Ueda, M. High-Throughput Identification of Peptide Agonists against GPCRs by Co-Culture of Mammalian Reporter Cells and Peptide-Secreting Yeast Cells Using Droplet Microfluidics. *Sci Rep* **2019**, *9* (1), 10920. https://doi.org/10.1038/s41598-019-47388-x.

(15) Shirbaghaee, Z.; Bolhassani, A. Different Applications of Virus-like Particles in Biology and Medicine: Vaccination and Delivery Systems. *Biopolymers* **2016**, *105* (3), 113–132. https://doi.org/10.1002/bip.22759.

(16) Zeltins, A. Construction and Characterization of Virus-Like Particles: A Review. *Mol Biotechnol* **2013**, *53* (1), 92–107. https://doi.org/10.1007/s12033-012-9598-4.

(17) Yin, D.; Zhong, Y.; Ling, S.; Lu, S.; Wang, X.; Jiang, Z.; Wang, J.; Dai, Y.; Tian, X.; Huang, Q.; Wang, X.; Chen, J.; Li, Z.; Li, Y.; Xu, Z.; Jiang, H.; Wu, Y.; Shi, Y.; Wang, Q.; Xu, J.; Hong, W.; Xue, H.; Yang, H.; Zhang, Y.; Da, L.; Han, Z.; Tao, S.; Dong, R.; Ying, T.; Hong, J.; Cai, Y. Dendritic-Cell-Targeting Virus-like Particles as Potent mRNA Vaccine Carriers. *Nat. Biomed. Eng* **2024**. https://doi.org/10.1038/s41551-024-01208-4.

(18) Ma, Y.; Yue, Y.; Ma, Y.; Zhang, Q.; Zhou, Q.; Song, Y.; Shen, Y.; Li, X.; Ma, X.; Li, C.; Hanson, M. A.; Han, G. W.; Sickmier, E. A.; Swaminath, G.; Zhao, S.; Stevens, R. C.; Hu, L. A.; Zhong, W.; Zhang, M.; Xu, F. Structural Basis for Apelin Control of the Human Apelin Receptor. *Structure* **2017**, *25* (6), 858-866.e4. https://doi.org/10.1016/j.str.2017.04.008.

(19) Yue, Y.; Liu, L.; Wu, L.-J.; Wu, Y.; Wang, L.; Li, F.; Liu, J.; Han, G.-W.; Chen, B.; Lin, X.; Brouillette, R. L.; Breault, É.; Longpré, J.-M.; Shi, S.; Lei, H.; Sarret, P.; Stevens, R. C.; Hanson, M. A.; Xu, F. Structural Insight into Apelin Receptor-G Protein Stoichiometry. *Nat Struct Mol Biol* **2022**, *29* (7), 688–697. https://doi.org/10.1038/s41594-022-00797-5.

(20) Shin, K.; Kenward, C.; Rainey, J. K. Apelinergic System Structure and Function. *Comprehensive Physiology* **2018**, *8* (1), 407–450. https://doi.org/10.1002/cphy.c170028.

(21) Ma, Y.; Yue, Y.; Ma, Y.; Zhang, Q.; Zhou, Q.; Song, Y.; Shen, Y.; Li, X.; Ma, X.; Li, C.; Hanson, M. A.; Han, G. W.; Sickmier, E. A.; Swaminath, G.; Zhao, S.; Stevens, R. C.; Hu, L. A.; Zhong, W.; Zhang, M.; Xu, F. Structural Basis for Apelin Control of the Human Apelin Receptor. *Structure* **2017**, *25* (6), 858-866.e4. https://doi.org/10.1016/j.str.2017.04.008.

(22) Winkle, P.; Goldsmith, S.; Koren, M. J.; Lepage, S.; Hellawell, J.; Trivedi, A.; Tsirtsonis, K.; Abbasi, S. A.; Kaufman, A.; Troughton, R.; Voors, A.; Hulot, J.-S.; Donal, E.; Kazemi, N.; Neutel, J. A First-in-Human Study of AMG 986, a Novel Apelin Receptor Agonist, in Healthy Subjects and Heart Failure Patients. *Cardiovasc Drugs Ther* **2023**, *37* (4), 743–755. https://doi.org/10.1007/s10557-022-07328-w.

(23) Yang, P.; Read, C.; Kuc, R. E.; Nyimanu, D.; Williams, T. L.; Crosby, A.; Buonincontri, G.; Southwood, M.; Sawiak, S. J.; Glen, R. C.; Morrell, N. W.; Davenport, A. P.; Maguire, J. J. A Novel Cyclic Biased Agonist of the Apelin Receptor, MM07, Is Disease Modifying in the Rat Monocrotaline Model of Pulmonary Arterial Hypertension. *Br J Pharmacol* **2019**, *176* (9), 1206–1221. https://doi.org/10.1111/bph.14603.

(24) Qin, S.; Meng, M.; Yang, D.; Bai, W.; Lu, Y.; Peng, Y.; Song, G.; Wu, Y.; Zhou, Q.; Zhao, S.; Huang, X.; McCorvy, J. D.; Cai, X.; Dai, A.; Roth, B. L.; Hanson, M. A.; Liu, Z.-J.; Wang, M.-W.; Stevens, R. C.; Shui, W. High-Throughput Identification of G Protein-Coupled Receptor Modulators through Affinity Mass Spectrometry Screening. *Chem. Sci.* **2018**, *9* (12), 3192–3199. https://doi.org/10.1039/C7SC04698G.

## 3.8 APPENDICES OF PROTOCOLS FOR GENERALIZED ASMS

*Appendix A: Magnetic Bead-Based Screening of Cyclic Peptide Libraries (Direct Capture Protocol)*

#### 3.8.1.1 Materials:

- MyOne Streptavidin T1 Dynabeads (10 mg/mL; 1 mg; 0.13 nmol IgG binding capacity)
- Biotinylated protein target
- Blocking buffer: 1 mg/mL BSA, 0.02% Tween 20, 1X PBS or 10% FBS, 0.02% Tween 20, 1X PBS
- 6M Guanidine Hydrochloride in 1X PBS
- C18 ZipTips
- Peptide Retention Time Calibration Standard (PRTC)
- NanoLC/MS setup

#### 3.8.1.2 Step 1: Immobilize Biotinylated Protein Target on Magnetic Beads

1. Vortex the bottle of Dynabeads to achieve a homogeneous suspension.
2. Transfer 100 µL (1 mg) of beads to glass vials.
3. Wash the beads three times with blocking buffer (1 mg/mL BSA, 0.02% Tween 20, 1X PBS). Use a magnetic rack to separate the beads from the supernatant.
4. Treat the beads with biotinylated protein target (~1.2 to 1.5 equivalents of the protein).
5. Agitate the solution on a rotating mixer for 15 minutes at ambient temperature or 1 hour at 4 °C.
6. Remove the supernatant and wash the beads three times with blocking buffer to remove any unbound protein.

#### 3.8.1.3 Step 2: Affinity Selection from In-Solution Library (or Known Binders)

1. Prepare the peptide library in blocking buffer (final volume: 1 mL). Library concentrations can vary depending on size, typically at either 100 fmol/member, 1 pmol/member, or 10 pmol/member.
2. Incubate the prepared library with the protein-loaded beads on a rotating mixer for 1 hour at 4 °C.
3. After incubation, transfer the beads to the magnetic separation rack and carefully remove the supernatant.
4. Wash the beads three times with 1 mL of chilled PBS. Perform these washes rapidly to minimize nonspecific binding.

#### 3.8.1.4 Step 3: Elution of Bound Peptides

1. Elute the bound peptides by adding 2 x 100 µL of 6 M Guanidine Hydrochloride prepared in 1X PBS.
2. Concentrate the eluates using C18 ZipTips, followed by lyophilization.
3. Resuspend the lyophilized peptides in 6 µL of water containing 0.1% formic acid. Submit 5 µL of the sample for nanoLC/MS analysis.

*Alternative Elution with Peptide Retention Time Calibration Standard (PRTC):*

- For control experiments, elute the peptides using 2 x 150 µL of the elution buffer containing 1 fmol/µL of Peptide Retention Time Calibration Standard.

*Appendix B: Size Exclusion Chromatography (SEC) Affinity Selection Mass Spectrometry (ASMS) Protocol for Cyclic Peptides*

#### 3.8.1.5 Materials and Reagents:

- Peptide library (prepared in DMSO or appropriate solvent)
- Target protein (e.g., 12ca5 or other target of interest)
- SEC column (e.g., Agilent SEC Column, 7.8 x 150 mm, 3 µm particle size)
- HPLC system with UV detector
- Nano-LC system with appropriate C18 column for desalting (e.g., PepMap RSLC C18 column)
- Mass spectrometer (e.g., Orbitrap Fusion Lumos Tribrid MS)
- Mobile phases for SEC (e.g., PBS, pH 7.5)
- Elution buffer for peptide recovery (e.g., 6 M guanidine hydrochloride, 200 mM phosphate buffer, pH 7.0)
- Internal standard for MS quantitation (e.g., Peptide Retention Time Calibration Standard, PRTC)
- Solvents for nano-LC (e.g., 0.1% formic acid in water and acetonitrile)
- BSA, Tween 20 for blocking buffers
- Centrifuge tubes, pipettes, and other standard lab equipment

#### 3.8.1.6 SEC Affinity Selection:

1. **Sample Preparation:**
    - Dissolve the peptide library in 5 mL of 5% DMSO in PBS (pH 7.5) to achieve a concentration of ~400 nM per library member.
    - Dissolve the target protein (e.g., 12ca5) in PBS (pH 7.4) and add to the peptide library to achieve the desired protein concentration.
2. **Incubation:**
    - Incubate the peptide library with the target protein at room temperature for 30 minutes to allow binding interactions.
3. **Size Exclusion Chromatography:**
    - Load 100 µL of the incubated mixture onto the SEC column.
    - Run the SEC using an isocratic mobile phase of PBS buffer (pH 7.5) at a flow rate of 1 mL/min.
    - Collect fractions at a UV absorbance of 214 nm, noting the elution time of the protein-binder complexes (typically around 4.155 min).
4. **Fraction Collection:**

- Collect the fraction containing the target protein-binder complexes.
- Perform additional washes with PBS if needed, and pool the collected fractions.

#### 3.8.1.7 Peptide Elution and Linearization:

5. **Elution of Peptides:**
   - Treat the collected SEC fraction with 2 x 150 µL of elution buffer (6 M guanidine hydrochloride, 200 mM phosphate, pH 7.0).
   - Ensure complete peptide recovery by keeping all eluted material for downstream analysis.
6. **Sample Preparation for MS:**
   - Prepare reference samples by diluting peptide library stock solutions into the elution buffer to mimic the concentration in eluted fractions.
   - Add an internal standard (e.g., PRTC) to each sample for normalization in MS analysis.

#### 3.8.1.8 Nano-LC and Mass Spectrometry Analysis:

7. **Nano-LC-MS Setup:**
   - Set up the nano-LC system with a C18 desalting column and analytical column (e.g., PepMap RSLC C18 column).
   - Load 5 µL of each sample, including eluates and reference samples, onto the system.
   - Run a gradient with solvents A (0.1% formic acid in water) and B (80% acetonitrile, 20% water, 0.1% formic acid), typically ramping from 1% B to 61% B over 40-60 minutes.
8. **MS Detection:**
   - Positive ion mode with a spray voltage of 2200 V for MS detection.
   - Set the Orbitrap for primary MS detection with parameters such as 120,000 resolution, scan range of 200-1400 m/z, and AGC target of 1 x 10^6.

#### 3.8.1.9 Data Analysis:

9. **Data Processing:**

- Extract ion chromatograms (EICs) for peptides of interest using the appropriate mass range.
- Identify and integrate the peaks corresponding to the peptide signals.

10. **Quantitation:**

- Compare the MS counts for each peptide in the affinity-selected samples to those in the reference samples.
- Normalize these counts using the internal standard to account for run-to-run variability.

11. **Dose-Response Curve Generation:**

- Generate a dose-response curve by plotting the MS detector counts against the known concentrations of peptides in the reference samples.
- Verify linearity over the sample loading range.

12. **Enrichment Calculation:**

- Calculate the recovery of each peptide as the ratio of counts in the affinity-selected sample to the reference sample.

# CHAPTER 4.

# Hydrogen-Deuterium Exchange to Characterize the Impact of Conformational Dynamics in Designed Structured Cyclic Peptides

## 4.1 INTRODUCTION

Cyclic peptides are an exciting modality in drug discovery, offering a unique blend of the pharmacokinetic properties of small molecule therapeutics and the high target specificity of large biologics[1,2]. Leveraging the improved accuracy and speed of computational cyclic peptide structure prediction, researchers can confidently design peptides that adopt the predicted structure[3–5]. Unlike larger proteins, cyclic peptide design cannot rely on secondary structure, such as alpha helices and beta sheets, and must focus on residue-level structural constraints and chemical modifications to control structure[3]. We have previously demonstrated that designing cyclic peptides with conformation-inducing amino acids like proline and a stable internal network of H-bonds enables conformationally stable peptides without the need for structural supports like stapling or a distinct secondary motif[6].

Despite this success, experimental exploration of synthetic cyclic peptide structure is needed. Traditional approaches like nuclear magnetic resonance spectroscopy (NMR) and X-ray crystallography have been used extensively in the initial

characterization and validation work of cyclic peptides to illustrate the accuracy of the computationally predicted structure[6,7]. Both NMR and X-ray crystallography are the gold standard for obtaining structural information, yet each consumes large amounts of time and sample for the analysis of a single peptide. Moving forwards, there is a need for a tractable experimental approach for rapidly characterizing cyclic peptide structure. Hydrogen deuterium exchange mass spectrometry (HDX-MS) distinguishes itself as a supplemental structural tool to NMR and X-ray crystallography and is well suited to examine the native conformation of cyclic peptides. Unlike NMR and X-ray crystallography, HDX-MS can be performed in medium throughput in a physiological environment, enabling evaluation of pools of peptides in parallel[8–10]. HDX-MS measures the uptake of deuterium in amide backbones and is capable of differentiating between accessible and protected amides by their distinct exchange kinetics[10,11]. This method is ideal for evaluating the structure of cyclic peptides whose conformational stability was designed around internal hydrogen bonds. Unlike with larger proteins, where solvent accessibility contributes to the exchange kinetics, participation in H-bonds is the main protector of amide hydrogens in cyclic peptides. HDX-MS, therefore, provides insight into the folded-ness of cyclic peptides and can determine if the H-bonding is in alignment with the designed structure.

Though structural stability is typically sought in cyclic peptide design, there is a precedent for conformational flexibility between stable alternate structures providing considerable advantages[12–14]. It has been hypothesized that cyclic peptides could possess dual-function capability across two or more conformations, wherein a peptide can exist in a permeability-competent conformation or a binding-competent

conformation. This dual-function capability is inspired by the natural product cyclosporine. Cyclosporine exhibits a remarkable conformational flexibility that is central to its biological function, allowing it to both passively penetrate cell membranes and bind to its target protein with high affinity[10,15]. As a result, much work has been done to characterize this attribute. HDX at a single time point has been used to distinguish the differences in flexibility between cyclosporine A and its analogues more than once in an effort to elucidate the impact of flexibility on permeability [10,16]. In an analogous but more intensive approach, NMR has been used to measure the accessibility of specific backbone positions along cyclosporine[17]. The prospect of capturing this dual-function capability in designed cyclic peptides represents significant potential in therapeutic design and offers a roadmap to balancing bioavailability and target specificity.

Traditional design of cyclic peptides for passive permeability requires accurate and stable structural control**.** Cyclic peptides have very little margin for error in permeability and must optimize known permeability determinants such as polar surface area and available H-bond donors[18,19]. This structural precision is essential to prevent interactions with polar solvents like water and instead favor insertion into greasy membranes. As such, membrane-traversing cyclic peptides prioritize amino acids with uncharged and hydrophobic sidechains. In contrast, cell-penetrating peptides often employ cationic amino acids like arginine and lysine to cause cellular disruption and promote permeability[2,20]. Although this method can be effective, cationic cell penetrating peptides can be associated with toxicity, and all-hydrophobic cyclic peptides typically possess more favorable pharmacokinetic properties[21–23]**.** In addition to amino acid composition, limiting hydrogen bonding networks with environmental water is crucial for

maintaining hydrophobicity. A polar solvation network prevents association with the cell membrane, and peptides incur an energetic penalty for shedding water molecules upon entry into the cell membrane. This challenge can be mitigated by obscuring backbone amide positions, either by forming internal H-bonds or by N-methyl modification. Both routes offer distinct advantages: internal H-bonds encourage structural stability, while N-methylation increases the overall hydrophobicity of the peptide[19,24,25]. Consequently, both strategies are common modifications used to promote permeability. These design principles are essential to permeable peptide design, but only effective if the cyclic peptide exists in the designed state. As such the impact of structural movement on permeability is relatively unknown. HDX-MS stands to fill in this gap in understanding.

Here, we report that HDX-MS measures the backbone accessibility of designed cyclic peptides. We relate this data to the stability of the designed structure and use its medium-throughput functionality to screen for peptides with flexibility between conformational isomers. Cyclic peptides with more than one stable conformation can provide dual-function capabilities, and we use HDX-MS to suggest new compounds with this flexibility. Furthermore, HDX-MS is used to illustrate the relationship between structural stability and permeability. We measure permeability and stability for a large pool of cyclic peptides. HDX-MS is an experimental technique that can contribute unique insights to permeability and structural stability.

## 4.2 MATERIALS AND METHODS

*Materials*

$D_2O$ was purchased at 99.99% purity from Cambridge Isotope Labs (#DLM-4-10). MS total recovery vials were acquired from Waters and paired with magnetic screw caps (PAL #09151907). Solvents like water, acetonitrile, dimethyl sulfoxide, and methanol were acquired from ThermoFisher at optima-grade purity.

*Peptide design and synthesis*

Rosetta's GenKIC was used to cyclize a string of glycines head-to-tail and design each glycine position into a full amino acid based on backbone angle. The set of amino acids available for design was limited to the L- and D-versions of hydrophobic amino acids to promote permeability. Cyclic peptide designs were modified by methylating the backbone amine position of each amino acid not participating in an internal hydrogen bond so that every position is satisfied. Peptides sizes span from 8 to 12 amino acids and contain between 2 and 6 N-methylated positions. Each cyclic peptide has at least 2 positions with exchangeable protons that are participating in internal H-bonds in the designed state. Designs were filtered for their predicted permeability and stability, and the top candidates with unique structures were individually synthesized using SPPS or purchased from WuXi AppTec as in Bhardwaj, et al. 2022[6]. The resulting library of synthetic cyclic peptides was generated from pooling individual peptides at equimolar concentrations.

*HDX-MS experimental procedure*

Individually synthesized peptides were dissolved and pooled in DMSO at a concentration of 0.015 mg/mL before evaluation by HDX-LC-MS in a series of experiments**.** Internal exchange reporters were included in the $D_2O$ mixture to measure exchange differences between runs[26]**.** When molecular weight was not able to uniquely distinguish peptides, the identity of resolved LC peaks was assigned by matching LC elution times to pure peptide samples. Due to the hydrophobic nature of the designed cyclic peptides, a final concentration of 20% acetonitrile and 10% DMSO were included in the exchange solution for solubility. Peptides were incubated in an exchange $D_2O$ buffer with 20% ACN and 10% DMSO, 10 mM Na Citrate and 150 mM NaCl at pH 6.0. Bradykinin standards were included in the deuterium solution to estimate back exchange**.** Peptide controls indicated there was a maximum deuteration of 79.83% of theoretical max. Undeuterated samples were prepped by incubating the test peptides with an identical solution replacing $D_2O$ with $H_2O$**.** Fully deuterated samples were exchanged for 24 hours at 60 degrees Celsius. Deuterium exchange was halted after the determined incubation period by 1:1 dilution in ice-cold quench buffer containing 8 M urea and 200 mM TCEP with 0.2% formic acid. This drops the pH to 2.5 and prevents further exchange. Quenched solutions were flash frozen in ethanol cooled by dry ice and then stored in a freezer at -80°C until ready for analysis.

Sample handling and LC separation were performed by a custom robot and cold-storage system that keeps samples frozen until injection and cold until data collection as in Watson, et al. 2021[27]**.** Data was collected on a Thermo Orbitrap Elite or Waters

Synapt G2-Si. Peptides were separated using an ACQUITY UPLC Protein BEH C4 Column, 300 Å, 1.7 µm, 2.1 mm x 50 mm in water with a mobile phase of 50/50 ACN/MeOH. A 15-minute gradient with a 10-minute elution window was utilized where the %B was initialized at 45% until 1 minute and raised to 70% B by 10 minutes before climbing to 95% B by 11 minutes and holding for three minutes. Tandem MS was collected to guarantee cyclic peptide ID in the case of any mass degeneracy that could not be solved by matching LC retention time to a pure sample**.** In contrast to traditional HDX-MS, no proteolytic digestion was performed on the peptides.

*Data analysis*

Initial analysis of the mass spectrometry raw data was performed using HDExaminer v3, which allowed us to extract the mass of each peptide and average the distribution of related mass peaks to generate a deuterium uptake curve over the time course. Data was further processed in HX Express for bimodal analysis[28]**.** Post HX Express, each unimodal or bimodal distribution was processed by fitting the uptake to an exponential curve and comparing the fitted rate against the theoretically calculated rate. Raw values were translated to account for the increased rate of exchange in organic solvent using previously established rate adjustments for organic buffer inclusion[26]. The theoretical rate of deuterium uptake was calculated using proton affinity exchange rates for amino acids[29,30]. Final protection factor values are the calculated log slowing factor (log(kc/kex))[31]**.** To account for the lack of termini and accurately reflect the neighboring effect of all amino acids, peptide sequences were doubled, and a final average of rates was made from an internal single repeat of the sequence, less the

positions with N-methylated modifications that could not exchange. The final result of data analysis is a single number protection factor, which represents the slowing of deuterium uptake as compared to the theoretical.

*Thermal shift experiments*

Peptide pools used for HDX-MS experiments were subjected to HPLC-MS using the same column and a rapid LC gradient, which ran from 0% to 100% mobile phase over 6 minutes. The mobile phase was consistent with HDX-MS experiments and was a 50/50 blend of ACN/MeOH. Peptide pools were injected from a 30% ACN solution at comparable concentration to HDX-MS experiments. Samples were repeatedly injected at varying LC column temperatures. The LC column was heated stepwise at every 10 degrees from 25 degrees to 75 degrees Celsius, and extracted LC traces for each peptide were plotted.

In vitro *permeability assays*

Permeability was measured using standard parallel artificial membrane permeability assays per Bhardwaj, et al. 2022[6].

## 4.3 RESULTS

*Development of HDX experiment*

Optimizing HDX-MS for hydrophobic cyclic peptides necessitated the inclusion of organic solvent for compound solubility, which accelerated the intrinsic exchange rate[32].

To counter the accelerated rate and be able to adequately sample exchange time points, we used sodium citrate to buffer the solution to pH 6.0 and slow exchange. We found that our peptides were too hydrophobic for standard C18 LC separation at 0°C and switched to C4 to reduce compound retention. Switching from 100% ACN mobile phase to 50/50 ACN/MeOH also helped with reducing the retention of peptides on C4 columns. These changes allowed us to observe full deuterium incorporation in CSA across a range of time points.

With the aim of examining conformational stability, we selected 77 designed cyclic peptides to test. This pool includes 18, 18, 24, and 17 peptides from the size range of 8 through 11 amino acids, respectively. This pool was evaluated for deuterium uptake with our optimized H/D Exchange method for cyclic peptides (**Figure 1**). We did not find any correlation between protection factor and other descriptors, such as size, mass, amino acid composition, hydrophobicity, or any computationally calculated metrics (**Supplementary Table 1**). For the majority of peptides, we were able to observe a gradual exchange across timepoints. The average protection factor was 1.76, but peptides with weaker internal H-bond networks had protection factors as low as 0.85, which corresponds to a 7-fold slowdown in deuterium incorporation. The best peptide with strong internal H-bonding slowed deuterium incorporation by a factor of 5500 and had a protection factor of 3.74 (good designs) (**Figure 1**)**.** While these protection factors are consistent and comparable within our data pool, comparison to external studies may be misleading. The inclusion of organic co-solvent in the exchange buffer was controlled using a universal correction, but the full impact on peptide amide intrinsic exchange rates may vary[26].

Though we cannot directly observe conformational structure with HDX-MS, we have previously established with X-ray crystallography that peptides with folded structure adopt their computationally predicted structure within an angstrom. Out of the 77 cyclic peptides tested, 18 of them have previously had their structure confirmed with X-ray crystallography, with an average backbone RMSD of 0.58 Angstroms[6]**.** This set of 18 cyclic peptides spanned a representative range of protection factors, from 0.92 to 3.74, with an average of 2.02. HDX-MS supports the accuracy of our computational models by demonstrating that all 77 of the synthetic cyclic peptides tested showed some level of delayed deuterium exchange, indicating an internal H-bonding network **(Supplementary Table 1 and Supplementary Figure 1).** This is in contrast to unstructured linear peptide controls, which deuterate fully within the first timepoint under these same exchange conditions. Peptide termini are known to accelerate exchange, therefore peptide cyclization should naturally slow deuterium uptake. This was controlled for by calculating protection factor from an internal peptide sequence with multiple repeats so as to exclude termini. Given the relationship between HDX-MS protection and structural stability, we propose using HDX-MS for rapid structure verification where the presence of highly protected internal H-bonds is indicative of a well-ordered structure.

*Liquid chromatography peak isomers*

For some cyclic peptides, we were able to observe more than one LC peak or an unnaturally broad distribution in the low temperature chromatography of our HDX-MS experiments, potentially indicating the presence of more than one isomer or conformer.

We hypothesize that atypical chromatographic peak shape is the result of either synthesis-derived chiral epimerization[33] or conformational differences and set out to elucidate the relationship by increasing temperature during liquid chromatography. CSA, a cyclic peptide with known conformational isomers, was again used as a model peptide. We found that CSA was able to collapse a broad LC distribution into a single narrow peak at higher temperatures, which is consistent with previously observed proline isomerization-driven conformational isomers **(Figure 2A)**[10,17,34]**.** We tested the rest of our HDX-MS set of designed cyclic peptides and found that most chromatographic irregularities were the result of synthesis-derived epimers and were unable to equilibrate between the two peaks, but a second group of eight peptides displayed collapsing LC profiles at higher temperatures (**Figure 3**). These peptides were able to collapse either partial shoulders or wholly distinct peaks into a narrow LC profile, which suggests the ability to sample multiple conformational isomers at the low temperature employed for HDX-MS.

Using HDX-MS to separately integrate the deuterium uptake of different isomers facilitated structural comparison and assignment. For cyclic peptides that had epimers present, HDX-MS allows the assignment of correctly and incorrectly structured epimers based on the relative deuterium protection (**Supplementary Figure 2**). The proper expected structure, with its designed H-bonding core, is expected to be the most protected. Therefore, we can assign the better-protected peak as the correctly structured epimer. We can do the same for cyclic peptides with conformational isomers, though the differences in protection factor are less stark. CSA's two uptake curves suggest one "closed" isomer with great protection and one "open" isomer with poor

protection **(Figure 2A).** We can identify the open isomer as being more weakly retained and eluting at the front of the LC peak, while the closed isomer is retained for longer. This interpretation is supported by previous literature which has suggested that CSA has a binding-competent open state and a permeability-competent closed state. CSA was joined by designed peptide 9.16 as a known model peptide with alternate conformational isomers. 9.16 had been previously confirmed as possessing conformational flexibility after structural NMR studies in a range of solvents, both polar and organic**.** Unlike CSA, peptide 9.16 showed minimal peak collapse given the increased temperature in the system and had minimal difference in deuterium uptake across the LC distribution (**Figure 2B**). Though 9.16 was shown to be conformationally flexible by NMR, we may not have been able to capture the same dynamics with HDX-MS.

In addition to CSA and 9.16, we suggest 6 of the 77 tested cyclic peptides as new peptides with observable conformational states (**Figure 3**). For each of the two 8-mers, three 9-mers, and one 10-mer, we experimentally observed clearly differential deuterium uptake and demonstrated interconversion between the states. Though we were able to demonstrate LC peak collapse at increased temperatures, peak tailing was evident in some traces due to secondary interactions and the population in the tail was no different from the main population. The isomers in peptides 8.4, 9.6, 9.8, and 10.23 had differences in deuterium incorporation at the first time point, indicating that there were immediate differences in the solvent accessibility of exchangeable positions. Peptides 8.10 and 9.4 had little difference between isomers at early time points, but deuteration differences across the middle timepoints suggests a less rigid structure in

one isomer. Two additional peptides, both 11 amino acids, had unnatural LC profiles that collapsed under high temperature but did not display significant deuterium uptake differences. The six peptides identified from this screen are now candidates for further verification using NMR.

In search of more support for conformational isomers, we utilized computational prediction of the structure-energy landscape of the putative conformationally flexible peptides (**Figure 4**). In brief, energy landscapes depict the possible conformations of a cyclic peptide and the corresponding energies of each state. For single state cyclic peptides, a desirable landscape takes on a funnel-like shape and predicts a single low energy structure. In previous work, we noted that energy landscapes with many isoenergetic minima are potentially indicative of conformational flexibility. For each of these six peptides, there was no clear distinct funnel-ness towards a low energy structure. This supports our previous assertion and allows us to suggest structures for the conformers identified by HDX-MS. Synthesizing peptides with this feature can be used to encourage conformational flexibility, but experimental characterization with assays like HDX-MS is the only current method to screen for flexible peptides with the conformational flexibility to adopt more than one conformer.

*Impact on permeability*

Computationally designed membrane traversing cyclic peptides are structured for optimal permeability. We sought to explore the relationship between structural rigidity and passive permeability using HDX-MS. Two cyclic peptides, 8.2.2 and 8.4.2, were selected as stable, permeable model peptides and were mutated to test both

permeability tolerance and conformational flexibility **(Figure 5A & C).** Libraries were generated by mutating a single amino acid position to each of the amino acids originally used in design, and each library was tested using PAMPA and HDX-MS (**Supplementary Table 2**). Computational prediction of each mutated sequence does not suggest a major alternate structure (backbone RMSD all <1 angstrom) from the original design (**Supplementary Table 2**).

We observed a few global trends in the conformational stability of single site mutagenesis. As expected, any attempt to introduce an amino acid with a charged sidechain significantly decreased permeability, and we conclude that the impact of the charged sidechain is far greater than the change made to the structural stability of the peptide **(Supplementary Table 2).** We therefore limit our analysis to hydrophobic amino acids. Glycine's lack of a sidechain grants it freedom of movement that enables alternative conformations without penalty, and substitution from any residue to glycine significantly reduced permeability **(Figure 5).** This tracks well with the known destabilizing effect that glycine has on secondary structures. Surprisingly, substitution from any residue to tyrosine also significantly reduced permeability in every case but one, likely the result of tyrosine's sidechain hydrogen bonding with water.

Substituting the three leucines in 8.2.2 to other hydrophobic amino acids revealed a positive correlation between structural stability and permeability (**Figure 5B**). Broadly, each of these grouped amino acid mutations was translated from the original sequence based off changes in hydrophobicity, and differences within each group were the result of structural differences. Grouping by amino acid type, we observe that

increased structural stability leads to increased permeability. This relationship is most apparent for alanine substitutions but is present for phenylalanine as well. Substitutions at the first position to either phenylalanine or tyrosine resulted in higher-than-expected permeability for its structural stability. This could be due to the orientation of the peptide's sidechain, which is on the opposite side of the peptide from positions 8 and 2 and could allow it to extend into space and "anchor" into the membrane.

Amino acid substitution in 8.4.2 showed a global positive correlation between stability and permeability, once glycines and tyrosines are accounted for. Interestingly, substitution at the second position from 2-aminoisobutyric acid to leucine was more permeable than substitution to phenylalanine, despite the increased hydrophobicity of phenylalanine. This agrees with the difference in HDX protection factor between the two mutants, but there are likely additional factors, like size, that impact permeability **(Figure 5B).** The third position follows a more predictable trend, where a phenylalanine substitution has greater permeability than alanine. Substituting isoleucine to either phenylalanine or valine at the fourth position did not significantly impact stability or permeability, providing evidence that some amino acid positions can be tuned for target specificity without losing permeability.

We attempted to translate our findings from mutagenesis studies to the general pool of synthetic cyclic peptides. The 77 peptides tested for conformational flexibility were paired with permeability data **(Figure 6).** Thirty-four of the 77 tested peptides met the criteria for significant apparent permeability (Papp) in PAMPA assays (Papp > 1 × 10−6 cm/s), corresponding to 10, 8, 8, and 8 permeable peptides of amino acid lengths

of 8 to 11. We found that of the 15 peptides below a threshold for “very poor” permeability (Papp < 1e-7), 12 of them had a protection factor of less than 1.3, and that 16 of the 19 peptides with a protection factor worse than 1.3 were considered non-permeable by PAMPA threshold (Papp < 1e-6) **(Figure 2).** There is also a size bias for this group, with only one 8 amino acid peptide, and 10 and 11 amino acid sizes having 6 and 5 members each. This size bias plays a role in this group’s poor permeability and structural stability, as larger peptides can be more flexible and difficult to stabilize without a rigid secondary structure. Notably, the 8 amino acid peptide is one of the exceptions to this class of poorly protected peptides which has good permeability. At higher levels of protection factor, we did not find any strong correlation as 31 of the remaining 58 peptides had good permeability and protection factor showed no preference towards permeability **(Figure 6).** Within the full scope of the data, we did not find a clear universal correlation between deuterium protection factor and cyclic peptide permeability, unlike what was observed in the controlled mutational studies.

## 4.4 DISCUSSION

We demonstrate that Hydrogen-Deuterium exchange can effectively measure the folding stability of designed cyclic peptides. Protection from deuterium exchange provides evidence of an internal Hydrogen bonding network that would suggest a cyclic peptide exists in its predicted conformation. We can apply this interpretation to identify epimer contaminants after peptide synthesis. Moreover, this technique can identify conformationally flexible peptides by finding separate deuterium kinetics within a single

peptide. HDX-MS provides a medium-throughput screening assay to identify flexible peptides at a much faster discovery rate than existing techniques like NMR. Protection factors can also provide an additional layer of information for the optimization of permeable cyclic peptides. Interestingly, the protection factors derived from HDX-MS were not predicted by any computational metric (including energy contributions from H-bonds), demonstrating that HDX-MS is a unique and valuable measure of conformational stability that presently cannot be effectively computationally modeled.

HDX-MS may be most valuable as a screening assay for cyclic peptides with flexibility between conformational isomers. Conformationally flexible peptides have the potential to be high-value therapeutic compounds and are currently difficult to intentionally design. The utility of this assay is supported by the agreement between our CSA findings and previous studies. Our results match crystal structures that have been solved showing CSA in different states, an open conformation while binding to a target protein and a closed conformation while buried in a hydrophobic environment. Driven by a single isomerization, CSA removes four intramolecular H-bonds between its closed and open states and fully solvates with environmental water. HDX-MS is an ideal tool to monitor the structural perturbations that result in alternate states like conformational state switching and proline isomerization. Proline isomers will be stable over a reasonable timescale for experimental measurement (on the order of minutes)**.** At room temperature (25°C), the rate of proline isomerization generally falls in the range of $3\times10^{-2}$ to $4\times10^{-5}$ $s^{-1}$, which corresponds to a stable isomer with a half-life of anywhere from 23 seconds to a few hours[35]**.** These isomers can be further stabilized by the conformational constraints of cyclic peptides, and the separation of conformational isomers is enhanced

by the 0°C LC environment during HDX-MS, which slows interconversion approximately 20-fold**.** As evidenced by CSA, cyclic peptide conformational state switching is largely driven by proline or N-methyl cis/trans isomerization and HDX-MS can screen for its presence.

Our HDX-MS results of designed cyclic peptide 9.16 did not agree with the previously observed NMR conformational states[10]. NMR studies of 9.16 in $CDCl_3$ observed a nearly even ratio of the designed all-trans state and a secondary conformation driven by a trans-to-cis proline isomerization[6]**.** Despite observing an unnaturally broad LC-resolved peak for 9.16 in our HDX experiments, integrating separate sides of the peak did not yield different uptake curves. It is possible that the solvent conditions of HDX-MS did not permit conformational flexibility in 9.16, or that 9.16's rate of isomerization was so fast that the exchanging populations were observed as an average **(Figure 2B).** Though the inclusion of organic solvent in the HDX-MS buffer may encourage and stabilize "closed" conformations, it is possible that further buffer optimization may result in stronger conformational flexibility. Comparative studies between two buffers could also provide insight, though not without added complexity. Further characterization is needed to understand how conformationally flexible peptides function in a physiological environment.

The six newly identified flexible cyclic peptides have the potential to be valuable model peptides for the development of therapeutic compounds. Peptides 8.4, 9.6, 9.8, and 10.23 display a significant amount of exchange at the earliest time point. From this, we can infer that the open isomer of these peptides has amide backbone positions

exposed. Having a mix of hydrophobic and polar interactions is essential for specific target interaction, and rapid deuteration suggests that polar backbones are exposed. Beyond this, peptides 9.6 and 10.23 show LC peak collapse at lower temperatures which indicates a lower energetic barrier between states. Conversely, peptides 9.4 and 8.10 resist peak collapse until higher temperatures, implying a higher energetic barrier between conformational isomers. Lower energetic barriers might be preferrable for permeability, as peptides will be able to sample alternate states faster when coming into contact with a cell membrane**.** These experimental measures can reveal quite a bit about a peptide's pharmacokinetic properties.

Computationally predicted energy landscapes of our designed cyclic peptides can provide putative isomeric structures. While the designed state of peptide 9.4 contains four H-bonds, a likely alternate state is separated by two isomerization events but maintains three internal H-bonds. Alternate states separated by isomerization that maintain or reform H-bonds commonly appear in the landscapes of our designed structures, which may reflect the energetic incentive a cyclic peptide has to create a hydrophobic pocket at its core. Experimentally, this ability to maintain or reform internal H-bonds rapidly within isomerized alternate states reduces the magnitude of difference in the deuterium uptake kinetics between configurational isomers. Of the experimentally identified flexible peptides, peptides 9.6 and 9.8 stand out with good permeability values and large deuterium uptake differences between their conformational states. This profile, which is more like CSA, should give them a favorable conformation for target interaction and solvency in addition to the permeability granted by the design state. Computationally predicted structures for 9.6 suggest that a single isomerization

between an Nme-dLeu and a Val is the most likely conformational shift and results in exposing 3 H-bond donors. While energy landscapes are an inelegant way to assign potential structures to the identified conformational isomers, it can provide insight into which amino acid is isomerizing.

HDX-MS measured structural stability provides an additional layer of interpretation and prediction for permeability. In controlled comparisons, we found that structural stability was positively correlated with permeability. Single site mutagenesis libraries revealed that hydrophobicity did not dictate permeability outright and that a cyclic peptide needs to be stable in its designed structure to be permeable. Mutagenesis libraries paired with HDX-MS to study permeability could prove a valuable tool for lead compound maturation. When we attempted to compare the structural stability of many different cyclic peptides and their permeability, the relationship was too complex to elucidate any trend. Size, polar surface area (PSA), and H-bond donors are traditional predictors for cyclic peptide permeability and combining them with protection factor may provide a more accurate model of permeability. HDX-MS derived protection factor is a unique and valuable tool for evaluating therapeutic peptides.

The single site mutagenesis libraries with HDX-MS afforded direct evaluation of the impact of structural stability on permeability. Though we initially included substitutions with charged amino acids, these peptides had remarkably poor permeability. Universally, glycine substitutions abolished permeability. This effect extends in magnitude beyond what would be expected from the reduction in hydrophobicity and clearly illustrates the importance of structure for permeability. More

surprising is the reduction in permeability incurred by tyrosine substitutions. Though commonly understood as possessing a hydrophobic sidechain, the phenolic hydroxyl group can act as a hydrogen bond donor to form solvation networks with water and prevent permeability**.** Design of membrane traversing passive permeability relies on the structural stability of the designed state. As a result, mutations that introduced structural instability reduced permeability.

In applying this study to our population of designed cyclic peptides, we found minimal relation between protection factor and permeability. The strongest inference we can draw is that protection factors below a threshold of 1.3 were associated with a decreased likelihood of permeability. Below this value, cyclic peptides may not be structurally stable enough to retain internal H-bonding against environmental solvation. Conversely, higher protection factors indicate a designed state that is stable for long enough to facilitate permeation.

There is clearly a complex relationship between protection factor and permeability. In the simplest model, hydrogen bonding with water in the environment irreversibly breaks the structure and prevents permeability. If this were the case, there would be a simple positive correlation between protection factor and permeability. A more sporadic relationship suggests that deuteration can occur without full solvation and without irreversibly breaking the designed state. Cyclic peptides may be able to briefly break internal H-bonds and associate transiently with water long enough to exchange deuterium before reforming the internal H-bond. These transient interactions likely do not break the designed state, and in this case protection factor may indicate

internal H-bond strength and correlate to the fraction of time the peptide is transiently associated with water.

In cases of peptides identified with stable alternate states, permeability and protection factor have a less impactful relationship. Once again, cyclosporine provides a model for permeability. Cyclosporine has been modeled to shed its solvation network upon coming into contact with a hydrophobic cell membrane[10]. This allows cyclosporine to recover from a solvated state and achieve permeability. The energetic barrier must be high enough to provide stability and allow interconversion but low enough to be overcome in the appropriate environments. The six new peptides identified in this work offer the opportunity to be scaffolds for future work exploring the limits of flexibility and design maturation towards a biological target. In tandem with the mutational tolerance explored by single site mutagenesis, we anticipate that there is some allowance for sequence optimization for properties beyond permeability. We expect that the high number of conformationally flexible cyclic peptides in our synthetic pool is an unintended side effect of frequent N-methylation and proline inclusion, both of which were intended to drive structure and permeability but also introduce isomerization sites.

## 4.5 CONCLUSION

We have developed and optimized an HDX-MS protocol for evaluating structured cyclic peptides in medium-throughput. This technique was used to measure and identify an internal H-bonding network in 77 designed cyclic peptides, supporting the accuracy

of designed states in solution. We found that poor structural stability (protection factor < 1.3) largely translated to poor permeability. Beyond this threshold, protection factor and permeability have a complex relationship that warrants further experimental characterization. Six of the measured peptides had stable alternate conformations that allowed for differential deuterium uptake kinetics, showing potential for HDX-MS as a discovery tool for conformationally flexible peptides. Future investigations into these compounds may provide guidance for next-generation dual-state design of cyclic peptides. Single site mutagenesis studies like the ones performed here provided an increased ability to draw conclusions from controlled comparisons. Conformationally flexible peptides would benefit from similar studies to test mutational tolerance and identify motifs that aid flexibility. Understanding the specific structural features that promote passive permeability can guide the rational design of new cyclic peptides with improved pharmacokinetic properties. In one experiment, we can gain insight into dozens of peptides and their folded-ness, conformational stability, and relationship with permeability. This should allow us to validate the current generation of single-state designed cyclic peptide while identifying clues that can lead to the development of next-generation conformationally flexible peptides.

## 4.6 FIGURES

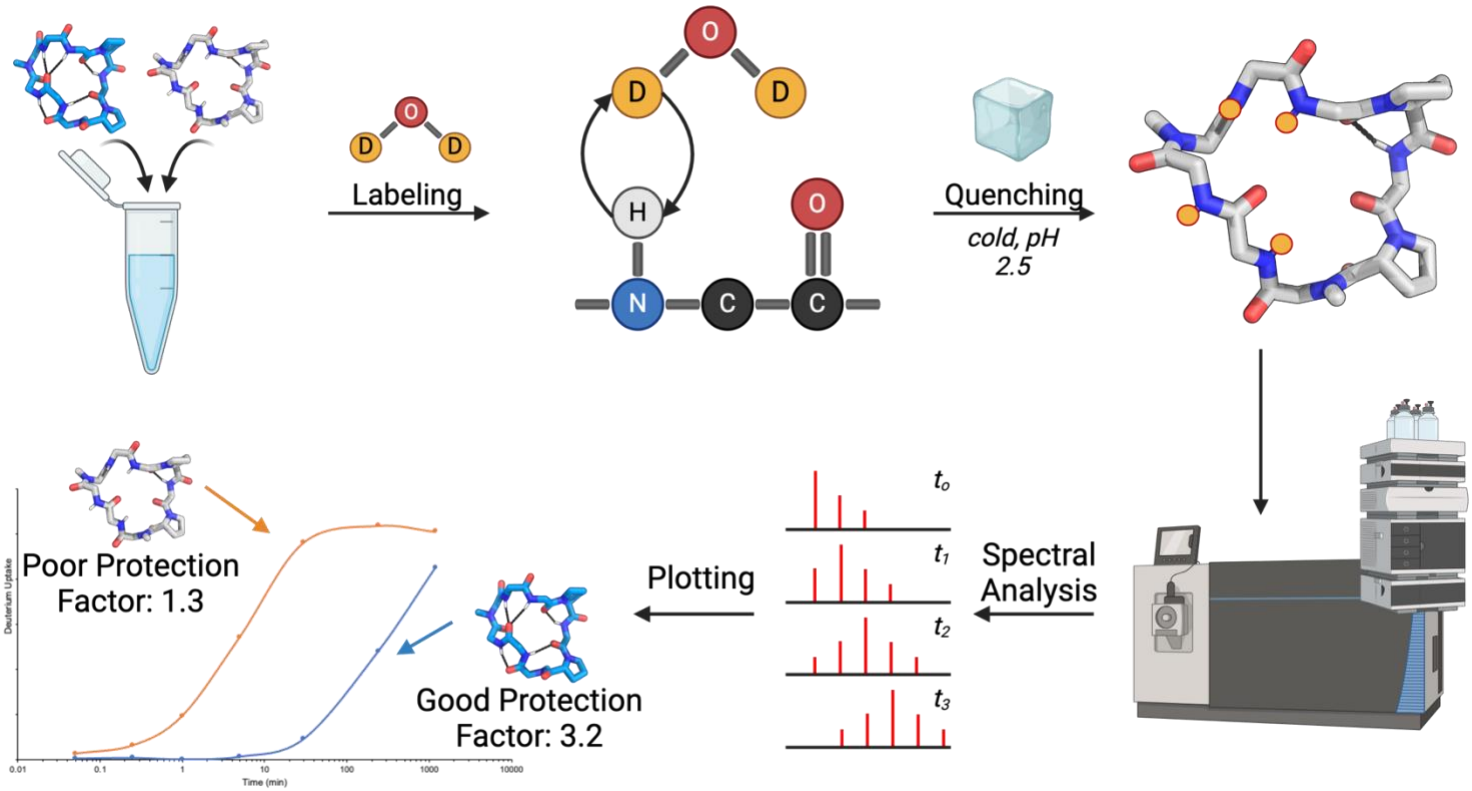


**Figure 4.1 | Overview of H/D Exchange for Cyclic Peptides**

Cyclic peptides in a pool are deuterated to their time point before being quenched to stop exchange. Data collection is performed using cold HPLC-MS. Data analysis begins with centroiding mass envelopes and plotting the increase in mass over time. Finally, this uptake plot is compared to a theoretical to generate a protection factor estimate of deuterium incorporation slowdown.

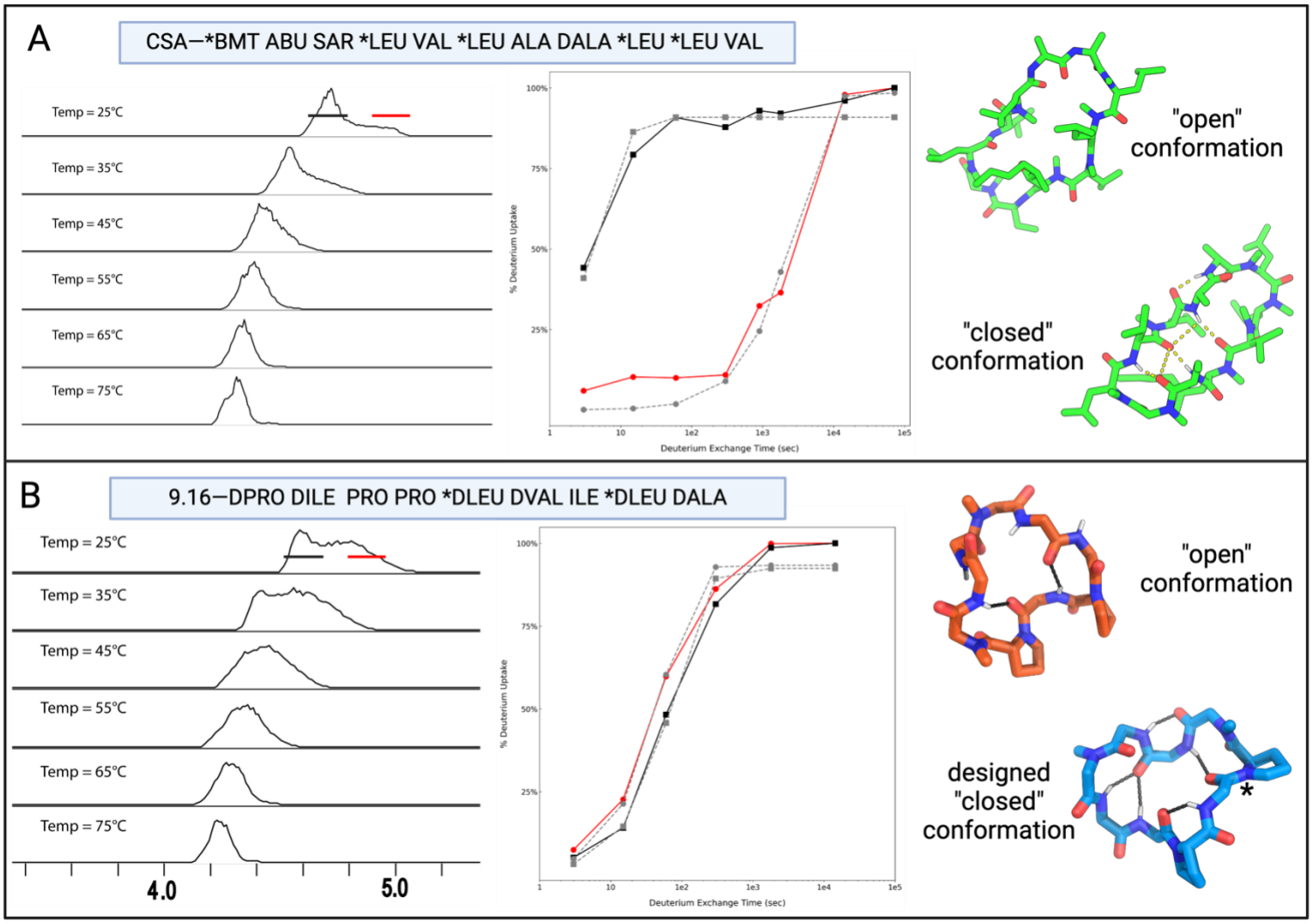


**Figure 4.2 | Experimental Characterization of Known Conformationally Flexible Peptides**

**A)** The natural product undecapeptide cyclosporine has collapsing peaks under high heat (left) and two distinct deuterium uptake rates (middle). These experimental results are the result of the peptide existing in two known separate conformational states (right). **B)** The designed synthetic peptide was previously identified using NMR. Despite having characteristic peak collapse under heat (left), the HDX uptake profiles are very similar (middle). The two identified conformational structures of 9.16 are shown on the right.

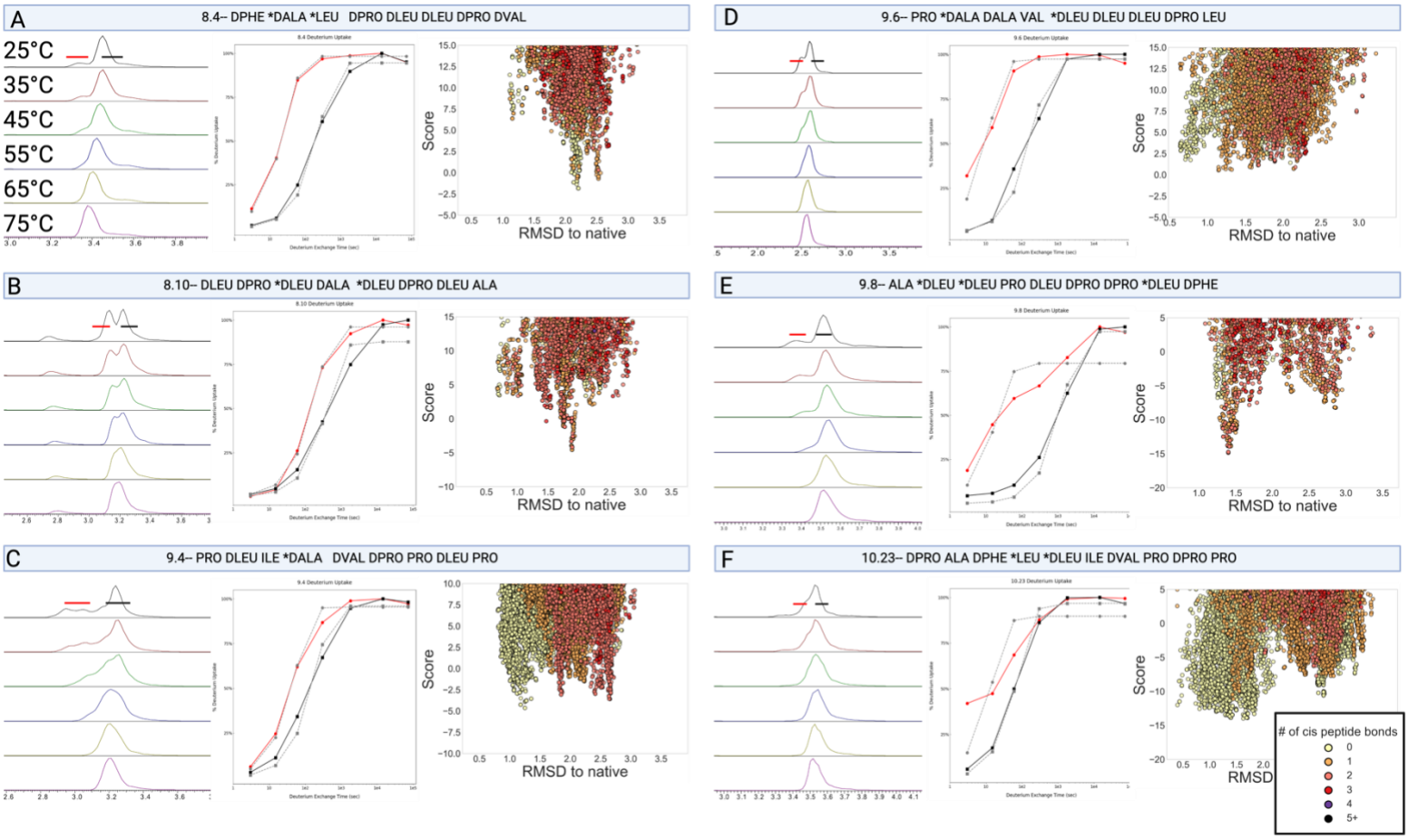


**Figure 4.3 | Experimental Characterization of Putative Conformationally Flexible Peptides**

Six peptides have experimental profiles that fit conformationally flexible peptides. In **A – F)**, peptide sequence is shown along with collapsing LC profiles under increased heat (left). Separate HDX integration of deuterium over time shows structural differences between the conformational isomers (middle). Computational modeling of the landscape of conformational isomers reveals possible isoenergetic states, and the isomers are color coded by number of cis peptide bonds to monitor trans/cis isomerization (right).

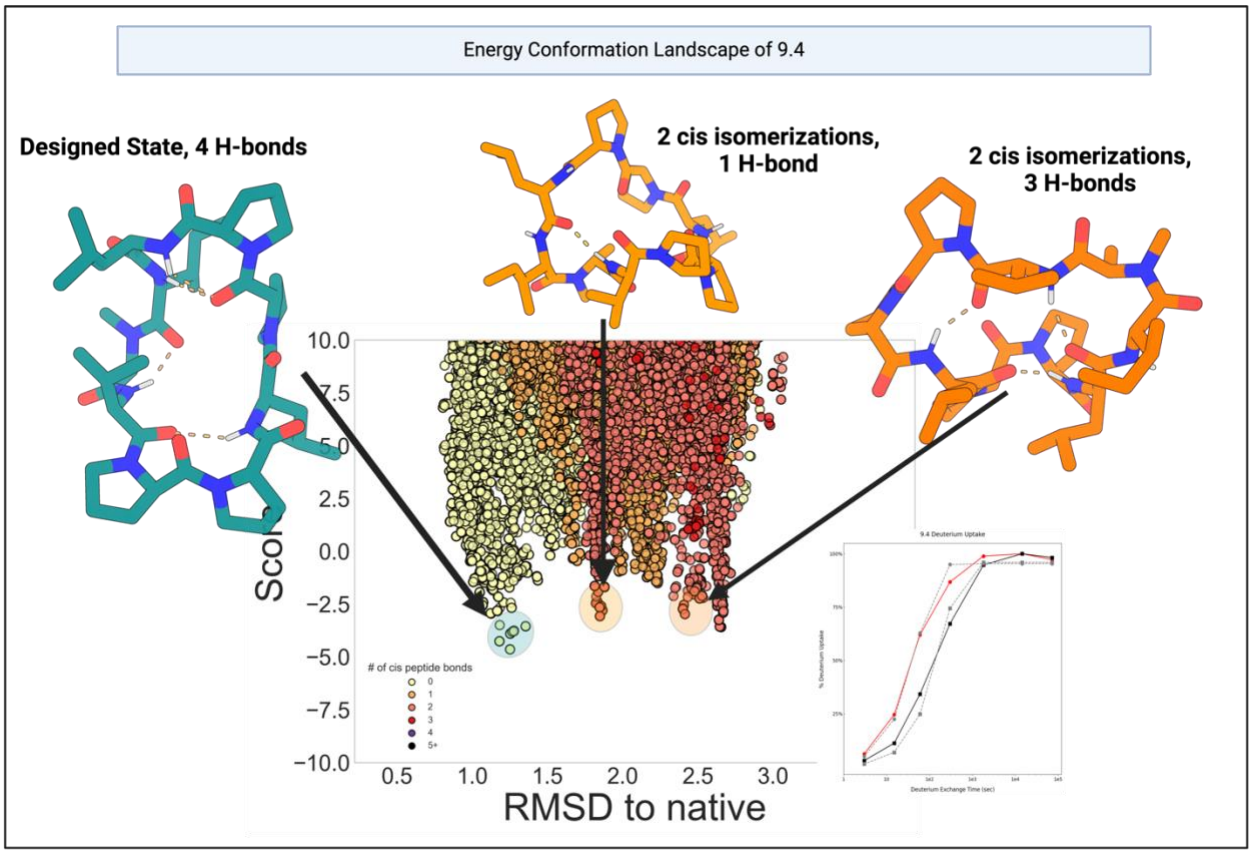


**Figure 4.4 | Conformational Landscape of Peptide 9.4.**

The putative conformationally flexible cyclic peptide 9.4 has a broad distribution of low energy conformations. We can identify the dominant structure and suggest possible alternative structures for the conformational isomer observed in our HDX study (inset). The two major alternate structures proposed are separated by two proline trans to cis isomerization events but restabilize with internal H-bonds.

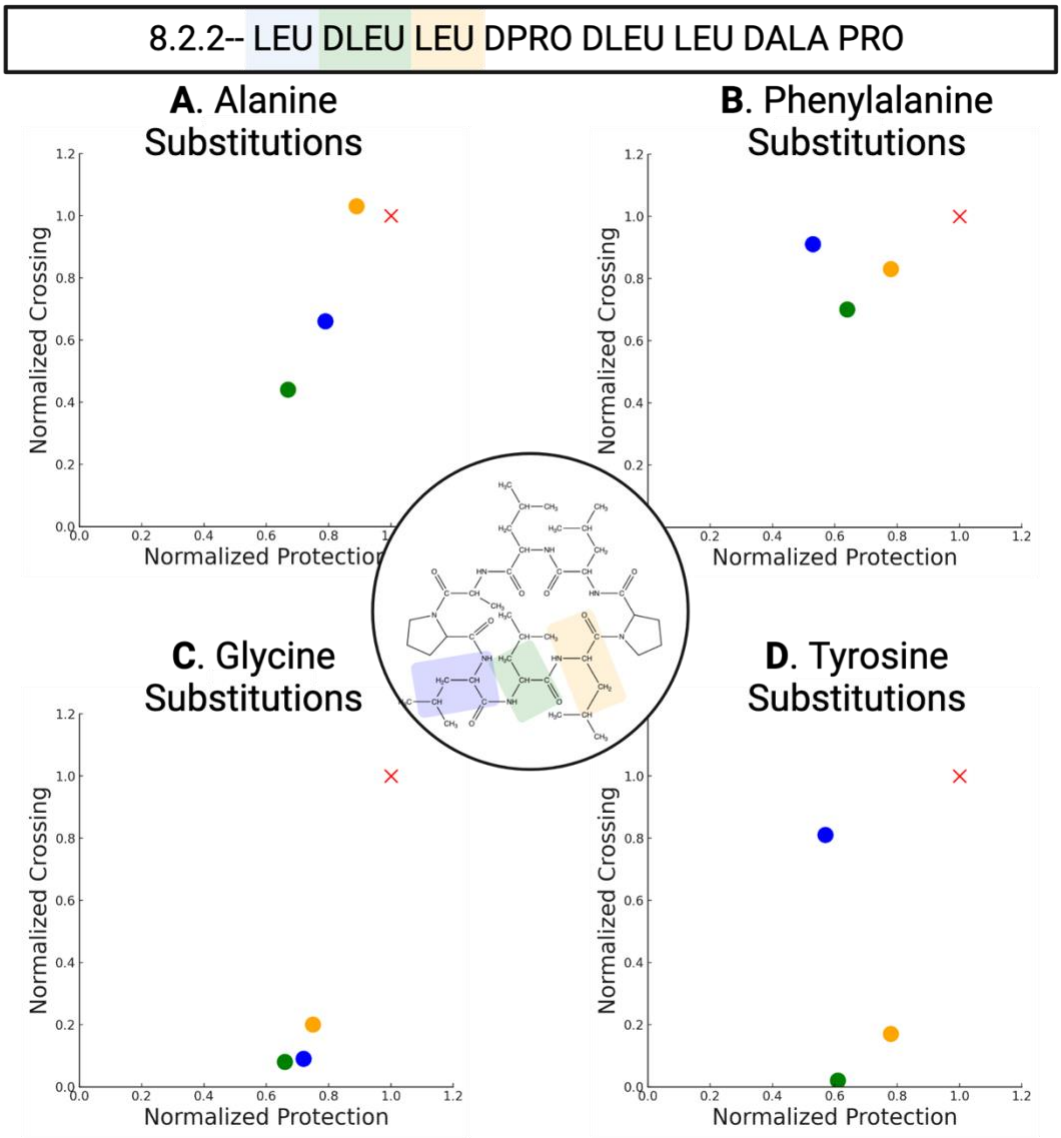


**Figure 4.5 | Impact of Single Amino Acid Mutations on Scaffold 8.2.2**

The permeable peptide 8.2.2 was used to observe the impact of single site mutations on structural stability and permeability. Three leucines in sequential amino acid positions were mutated to four other hydrophobic amino acids, alanine (**A),** phenylalanine (**B**), glycine (**C**), and tyrosine (**D**). A positive correlation between stability and permeability emerged. The structure of 8.2.2 is shown inset in the middle.

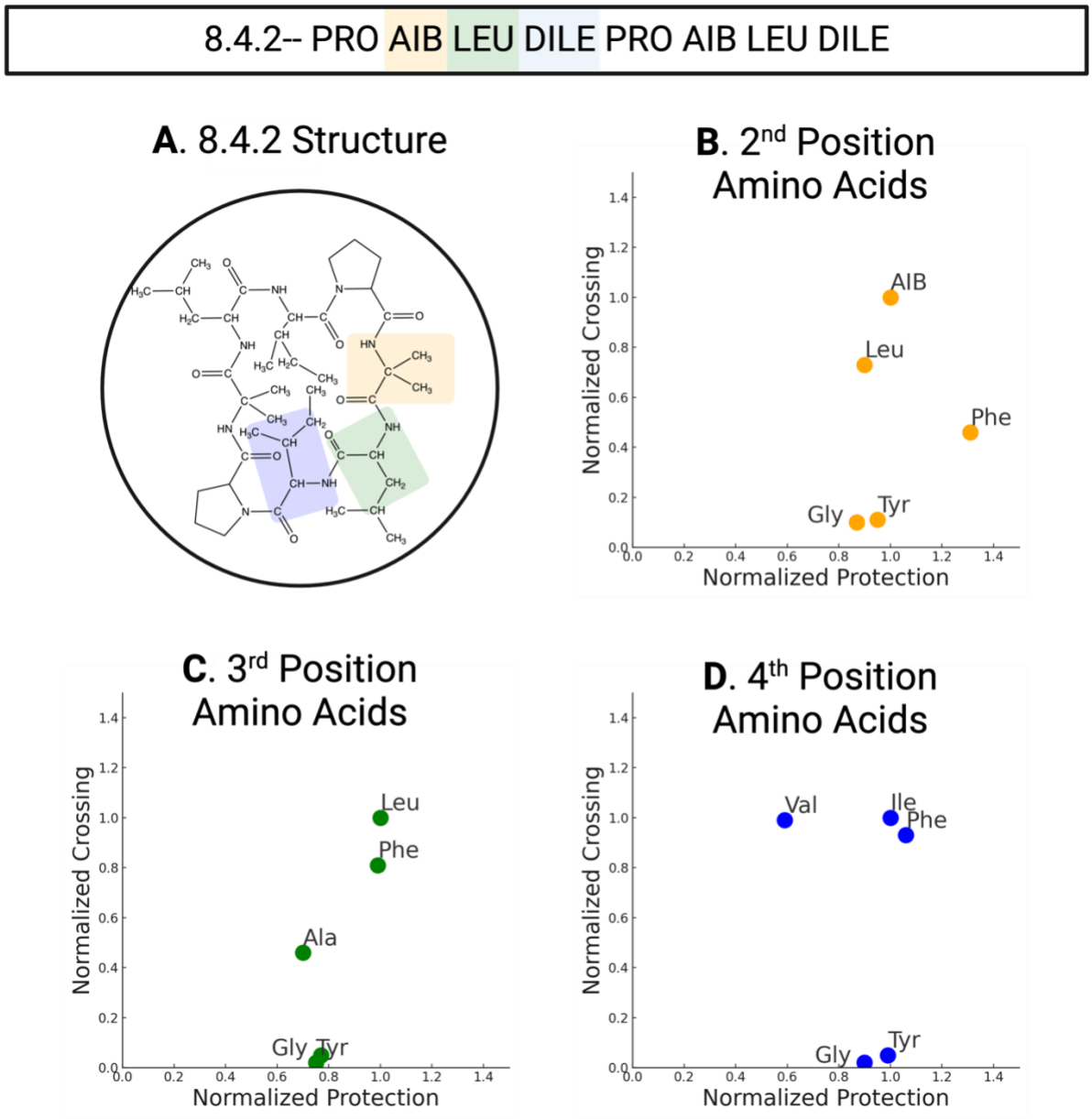


**Figure 4.6 | Impact of Single Amino Acid Mutations on Scaffold 8.4.2**

The permeable peptide 8.4.2 was used to observe the impact of single site mutations on structural stability and permeability. **A**) Structure of 8.4.2 with mutated positions highlighted. Three sequential amino acids within peptide 8.4.2 were mutated to four other hydrophobic amino acids, as labeled above the data points (**B-D**). Glycine and Tyrosine universally impeded permeability while other amino acids were tolerable.

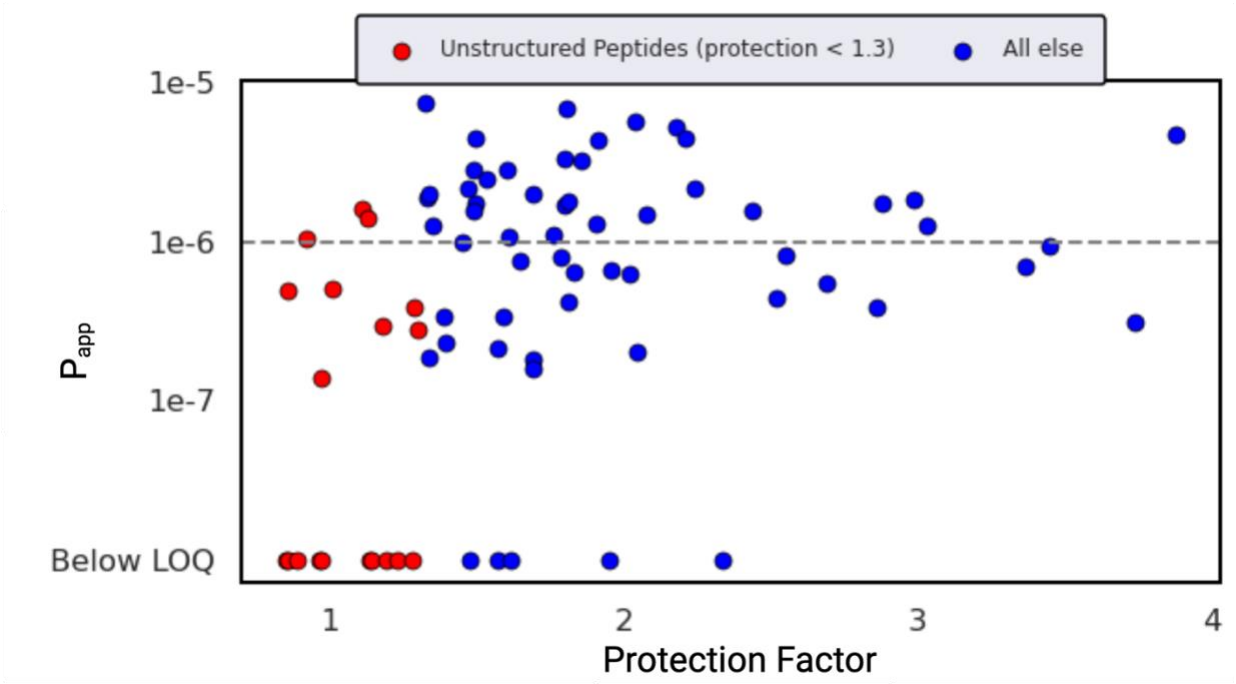


**Figure 4.7 | Protection Factor vs. Apparent Permeability**

A pool of 77 designed synthetic peptides were tested for their structural stability and their permeability. Structural stability by protection factor did not show a correlation with permeability as measure by the PAMPA assay. We noted that below a protection factor of 1.3, obtaining permeability is more difficult than above 1.3 protection factor.

**Supplementary Table 1 | RMSD of Tested Designed Synthetic Peptides**

| Peptide | CA RMSD | RMSD backbone | RMSD all heavy atoms | Papp | Protection Factor |
|---|---|---|---|---|---|
| 8.17 | 1.25 | 1.7 | 3.14 | 5.23E-06 | 2.18 |
| 8.9 | 0.7 | 1.04 | 1.65 | 1.00E-07 | 1.95 |
| 10.22 | 0.75 | 0.82 | 1.58 | 3.88E-07 | 1.29 |
| 8.6 | 1.04 | 1.35 | 1.56 | 2.03E-06 | 1.69 |
| 10.23 | 0.38 | 0.41 | 1.28 | 1.00E-07 | 1.57 |
| 8.3 | 0.56 | 0.57 | 1.23 | 1.04E-06 | 0.92 |
| 8.14 | 0.46 | 0.5 | 1.17 | 2.35E-07 | 1.40 |
| 11.1 | 0.43 | 0.43 | 1.11 | 2.81E-06 | 1.61 |
| 8.15 | 0.34 | 0.42 | 0.98 | 4.32E-06 | 1.92 |
| 810 | 0.22 | 0.35 | 0.96 | 8.24E-07 | 2.55 |
| 11.3 | 0.4 | 0.43 | 0.93 | 1.00E-07 | 1.23 |
| 11.4 | 0.51 | 0.54 | 0.93 | 1.00E-07 | 1.62 |
| 9.8 | 0.39 | 0.33 | 0.88 | 1.76E-06 | 2.88 |
| 8.2 | 0.2 | 0.24 | 0.86 | 9.50E-07 | 3.45 |
| 8.1 | 0.2 | 0.21 | 0.85 | 2.16E-06 | 2.25 |
| 10.31 | 0.47 | 0.45 | 0.84 | 1.00E-07 | 2.44 |
| 8.12 | 0.29 | 0.48 | 0.79 | 3.15E-07 | 3.74 |
| 8.5 | 0.11 | 0.24 | 0.55 | 1.08E-06 | 1.61 |

A subset of peptides tested in this work have had their structures reported in our previous work, Bhardwaj et al. (2022)[6]. For all tested peptides, high accuracy to the designed state was observed. Moreover, each peptide had substantial structural stability when tested for protection factor.

## Supplementary Table 2 | Summary of Tested Library of Cyclic Peptides

| Peptide | Hydrophobicity (RT) | Mass | Papp | HDX |
|---|---|---|---|---|
| 8001 | 52.1434 | 842.6 | 0.00000216 | 2.245619 |
| 8002 | 52.6196 | 876.6 | 0.00000095 | 3.446556 |
| 8003 | 51.96 | 816.5 | 0.00000104 | 0.916152 |
| 8004 | 56.88 | 878.5 | 6.26E-07 | 2.020231 |
| 8005 | 54.44 | 850.5 | 0.00000108 | 1.612336 |
| 8006 | 60.94 | 872.6 | 0.00000203 | 1.692508 |
| 8007 | 50.16 | 788.5 | 2.15E-07 | 1.56967 |
| 8008 | 58.66 | 872.6 | 0.00000128 | 3.025019 |
| 8009 | 46.28 | 798.5 | 0.0000001 | 1.949447 |
| 8010 | 53.36 | 816.5 | 8.24E-07 | 2.547985 |
| 8011 | 50.92 | 802.5 | 5.44E-07 | 2.691664 |
| 8012 | 57.54 | 888.5 | 3.15E-07 | 3.737756 |
| 8014 | 51.191 | 814.4 | 2.35E-07 | 1.399122 |
| 8015 | 51.6672 | 816.5 | 0.00000432 | 1.91503 |
| 8016 | 50.2386 | 788.5 | 0.00000247 | 1.535428 |
| 8017 | 52.6196 | 868.5 | 0.00000523 | 2.181288 |
| 9004 | 55.44 | 911.6 | 4.21E-07 | 1.809597 |
| 9005 | 57.42 | 947.5 | 3.41E-07 | 1.58979 |
| 9006 | 46.6 | 915.6 | 0.00000327 | 0.837003 |
| 9007 | 56.74 | 963.6 | 0.00000174 | 1.503807 |
| 9008 | 57.72 | 1003.6 | 0.00000176 | 2.883928 |
| 9009 | 53.572 | 943.6 | 0.00000128 | 1.354967 |
| 9010 | 52.1434 | 945.6 | 0.0000001 | 0.971328 |
| 9011 | 49.7624 | 885.5 | 2.94E-07 | 1.179768 |
| 9012 | 51.191 | 897.5 | 1.89E-07 | 1.34128 |
| 9013 | 47.3814 | 885.6 | 1.41E-07 | 0.972268 |
| 9014 | 49.2862 | 943.5 | 8.69E-08 | 1.146129 |
| 9015 | 52.1434 | 961.5 | 0.00000219 | 1.473108 |
| 9016 | 55.3 | 941.6 | 0.00000199 | 1.335726 |
| 9017 | 48.78 | 897.6 | 0.0000001 | 2.337671 |
| 9018 | 49.2862 | 947.6 | 0.00000157 | |
| 9019 | 52.2 | 885.5 | 6.64E-07 | 1.956951 |
| 9020 | 53.572 | 915.6 | 0.0000016 | 1.139117 |
| 9021 | 44.98 | 937.5 | 5.05E-07 | 1.007074 |
| 10001 | 55.12 | 1060.6 | 7.56E-07 | 1.64575 |
| 10002 | 56.2 | 1074.6 | 8.08E-07 | 1.79488 |
| 10003 | 53.94 | 1078.6 | 0.00000049 | 0.864015 |
| 10004 | 52.9 | 1010.6 | 1.84E-07 | 1.685788 |
| 10005 | 56.56 | 1032.6 | 4.41E-07 | 2.52433 |

| | | | | |
|---|---|---|---|---|
| 10006 | 59.44 | 1088.7 | 0.00000478 | 3.883746 |
| 10007 | 58.84 | 1046.6 | 0.0000019 | 1.332946 |
| 10008 | 62.04 | 1060.6 | 0.00000186 | 2.986176 |
| 10009 | 57.42 | 1158.7 | 0.0000001 | 1.283857 |
| 10010 | 60.88 | 1102.7 | 0.00000451 | 1.494856 |
| 10011 | 55.44 | 1026.6 | 2.79E-07 | 1.301135 |
| 10012 | 51.4 | 1176.7 | 0.0000001 | 0.848624 |
| 10013 | 56.96 | 1046.6 | 6.54E-07 | 1.214872 |
| 10014 | 54.26 | 1032.6 | 0.0000001 | 1.14 |
| 10015 | 60.36 | 1040.7 | 0.00000148 | 2.075235 |
| 10016 | 49.74 | 1016.6 | 0.00000034 | 1.393196 |
| 10017 | 61.54 | 1068.7 | 0.00000569 | 2.040308 |
| 10018 | | 1100.7 | 0.00000071 | |
| 10019 | 52.9 | 1012.6 | 1.61E-07 | 1.693643 |
| 10020 | 59.82 | 1148.67 | 0.00000333 | 1.139957 |
| 10021 | | 1064.6 | 4.87E-07 | |
| 10022 | 57.14 | 1054.7 | 3.88E-07 | 1.2884 |
| 10023 | 58.66 | 1072.6 | 0.0000001 | 1.572332 |
| 10028 | 57.3816 | 1057.4 | 0.0000001 | 1.192896 |
| 10031 | 55.4768 | 1131.4918 | 0.0000001 | 2.440007 |
| 11001 | 54.5244 | 1153.7 | 0.00000281 | 1.606073 |
| 11002 | 60.46 | 1187.7 | 0.0000017 | 1.80345 |
| 11003 | 53.0958 | 1143.7 | 0.0000001 | 1.231907 |
| 11004 | 52.2 | 1109.7 | 0.0000001 | 1.619845 |
| 11006 | 60.16 | 1139.7 | 0.00000179 | 1.812423 |
| 11007 | 60.06 | 1199.7 | 0.00000281 | 1.489062 |
| 11008 | 50.2386 | 1130.42 | 0.0000001 | 0.964433 |
| 11009 | 55.0006 | 1137.7 | 0.00000143 | 1.12959 |
| 11010 | 61.38 | 1211.66 | 0.00000445 | 2.212324 |
| 11011 | 61.02 | 1167.6 | 0.00000129 | 1.914972 |
| 11012 | 57.42 | 1157.6 | 9.93E-07 | 1.450554 |
| 11013 | 53.572 | 1172.5 | 0.0000001 | 0.861878 |
| 11014 | 55.953 | 1224.6 | 0.0000001 | 0.892797 |
| 11015 | 58.92 | 1109.7 | 3.87E-07 | 2.860566 |
| 11016 | 54.0482 | 1187.7 | 0.0000011 | 1.765321 |
| 11017 | 52.1434 | 1185.7 | 2.06E-07 | 2.044377 |
| 11018 | 58.8102 | 1272.7 | 0.0000001 | 1.479692 |

A pool of 77 designed cyclic peptides were tested for permeability using the PAMPA assay and for protection factor using HDX-MS.

**Supplementary Figure 1 | H/D Uptake Plots for All Tested Peptides**

Each peptide was individually evaluated for H/D uptake using HDExaminer software and measures of incorporated deuterium were exported and replotted using matplotlib. The library of peptide represents a wide variety of uptake curves, from immediate deuteration to extended protection. Data in blue, fitted single exponent in orange.

## Supplementary Table 2 | SSM HDX Results

| Scaffold | Position | Amino Acid | Protection Factor | Normalized PF | % Crossing | Normalized Crossing |
|---|---|---|---|---|---|---|
| 8.2.2 | 1 | Ser | 2.12 | 1.17 | 0.27 | 0.03 |
| 8.2.2 | 1 | Leu | 1.81 | 1.00 | 8.04 | 1.00 |
| 8.2.2 | 1 | Asp | 1.54 | 0.85 | 0.21 | 0.03 |
| 8.2.2 | 1 | Ala | 1.43 | 0.79 | 5.32 | 0.66 |
| 8.2.2 | 1 | Gly | 1.30 | 0.72 | 0.75 | 0.09 |
| 8.2.2 | 1 | Arg | 1.26 | 0.70 | 0.06 | 0.01 |
| 8.2.2 | 1 | Tyr | 1.03 | 0.57 | 6.52 | 0.81 |
| 8.2.2 | 1 | Phe | 0.97 | 0.53 | 7.33 | 0.91 |
| 8.2.2 | 2 | Gly | 1.19 | 0.66 | 0.61 | 0.08 |
| 8.2.2 | 2 | Ala | 1.20 | 0.67 | 3.51 | 0.44 |
| 8.2.2 | 2 | Ser | 1.26 | 0.70 | 0.04 | 0.01 |
| 8.2.2 | 2 | Leu | 1.81 | 1.00 | 8.04 | 1.00 |
| 8.2.2 | 2 | Asp | 1.28 | 0.71 | 2.62 | 0.33 |
| 8.2.2 | 2 | Arg | 1.13 | 0.62 | 0.05 | 0.01 |
| 8.2.2 | 2 | Phe | 1.15 | 0.64 | 5.66 | 0.70 |
| 8.2.2 | 2 | Tyr | 1.11 | 0.61 | 0.12 | 0.02 |
| 8.2.2 | 8 | Gly | 1.35 | 0.75 | 22.14 | 0.20 |
| 8.2.2 | 8 | Ala | 1.60 | 0.89 | 114.94 | 1.03 |
| 8.2.2 | 8 | Ser | 1.41 | 0.78 | 10.11 | 0.09 |
| 8.2.2 | 8 | Leu | 1.81 | 1.00 | 111.71 | 1.00 |
| 8.2.2 | 8 | Asp | 1.60 | 0.89 | 50.94 | 0.46 |
| 8.2.2 | 8 | Arg | 1.36 | 0.75 | 0.22 | 0.00 |
| 8.2.2 | 8 | Phe | 1.40 | 0.78 | 93.16 | 0.83 |
| 8.2.2 | 8 | Tyr | 1.40 | 0.78 | 18.89 | 0.17 |
| 8.4.2 | 2 | GLY | 1.14 | 0.87 | 11.11 | 0.10 |
| 8.4.2 | 2 | AIB | 1.32 | 1.00 | 106.02 | 1.00 |
| 8.4.2 | 2 | SER | 1.63 | 1.23 | 32.92 | 0.31 |
| 8.4.2 | 2 | LEU | 1.19 | 0.90 | 77.11 | 0.73 |
| 8.4.2 | 2 | dASP | 1.41 | 1.07 | 18.02 | 0.17 |
| 8.4.2 | 2 | ARG | 1.19 | 0.90 | 0.97 | 0.01 |
| 8.4.2 | 2 | PHE | 1.74 | 1.31 | 48.66 | 0.46 |
| 8.4.2 | 2 | TYR | 1.26 | 0.95 | 11.19 | 0.11 |
| 8.4.2 | 3 | dSER | 1.46 | 1.03 | 20.62 | 0.25 |
| 8.4.2 | 3 | dLEU | 1.42 | 1.00 | 83.42 | 1.00 |
| 8.4.2 | 3 | dASP | 1.56 | 1.09 | 48.86 | 0.59 |
| 8.4.2 | 3 | dALA | 1.40 | 0.99 | 38.74 | 0.46 |
| 8.4.2 | 3 | GLY | 1.52 | 1.07 | 3.98 | 0.05 |
| 8.4.2 | 3 | dARG | 1.41 | 0.99 | 0.21 | 0.00 |
| 8.4.2 | 3 | dTYR | 1.50 | 1.05 | 1.98 | 0.02 |
| 8.4.2 | 3 | dPHE | 1.40 | 0.99 | 67.80 | 0.81 |
| 8.4.2 | 4 | dAsp | 1.92 | 1.45 | 86.92 | 0.47 |
| 8.4.2 | 4 | dPhe | 1.41 | 1.06 | 172.28 | 0.93 |
| 8.4.2 | 4 | dArg | 1.33 | 1.00 | 0.73 | 0.00 |
| 8.4.2 | 4 | dIle | 1.33 | 1.00 | 186.00 | 1.00 |
| 8.4.2 | 4 | dTyr | 1.31 | 0.99 | 9.41 | 0.05 |
| 8.4.2 | 4 | dThr | 1.27 | 0.96 | 32.07 | 0.17 |
| 8.4.2 | 4 | Gly | 1.19 | 0.90 | 3.04 | 0.02 |
| 8.4.2 | 4 | dVal | 0.78 | 0.59 | 183.52 | 0.99 |

Full results of the HDX and PAMPA studied of mutational studies of scaffold 8.2.2 and 8.4.2. Throughout the scaffolds and positions, charged amino acids abolished permeability, leading to only the hydrophobic amino acids being considered.

## 4.7 BIBLIOGRAPHY

(1) Costa, L.; Sousa, E.; Fernandes, C. Cyclic Peptides in Pipeline: What Future for These Great Molecules? *Pharmaceuticals (Basel)* **2023**, *16* (7), 996. https://doi.org/10.3390/ph16070996.

(2) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov Today* **2015**, *20* (1), 122–128. https://doi.org/10.1016/j.drudis.2014.10.003.

(3) Bhardwaj, G.; Mulligan, V. K.; Bahl, C. D.; Gilmore, J. M.; Harvey, P. J.; Cheneval, O.; Buchko, G. W.; Pulavarti, S. V. S. R. K.; Kaas, Q.; Eletsky, A.; Huang, P.-S.; Johnsen, W. A.; Greisen, P. J.; Rocklin, G. J.; Song, Y.; Linsky, T. W.; Watkins, A.; Rettie, S. A.; Xu, X.; Carter, L. P.; Bonneau, R.; Olson, J. M.; Coutsias, E.; Correnti, C. E.; Szyperski, T.; Craik, D. J.; Baker, D. Accurate de Novo Design of Hyperstable Constrained Peptides. *Nature* **2016**, *538* (7625), 329–335. https://doi.org/10.1038/nature19791.

(4) Hosseinzadeh, P.; Bhardwaj, G.; Mulligan, V. K.; Shortridge, M. D.; Craven, T. W.; Pardo-Avila, F.; Rettie, S. A.; Kim, D. E.; Silva, D.-A.; Ibrahim, Y. M.; Webb, I. K.; Cort, J. R.; Adkins, J. N.; Varani, G.; Baker, D. Comprehensive Computational Design of Ordered Peptide Macrocycles. *Science* **2017**, *358* (6369), 1461–1466. https://doi.org/10.1126/science.aap7577.

(5) Mulligan, V. K.; Workman, S.; Sun, T.; Rettie, S.; Li, X.; Worrall, L. J.; Craven, T. W.; King, D. T.; Hosseinzadeh, P.; Watkins, A. M.; Renfrew, P. D.; Guffy, S.; Labonte, J. W.; Moretti, R.; Bonneau, R.; Strynadka, N. C. J.; Baker, D. Computationally Designed Peptide Macrocycle Inhibitors of New Delhi Metallo-β-Lactamase 1.

*Proceedings of the National Academy of Sciences* **2021**, *118* (12), e2012800118. https://doi.org/10.1073/pnas.2012800118.

(6) Bhardwaj, G.; O'Connor, J.; Rettie, S.; Huang, Y.-H.; Ramelot, T. A.; Mulligan, V. K.; Alpkilic, G. G.; Palmer, J.; Bera, A. K.; Bick, M. J.; Di Piazza, M.; Li, X.; Hosseinzadeh, P.; Craven, T. W.; Tejero, R.; Lauko, A.; Choi, R.; Glynn, C.; Dong, L.; Griffin, R.; Van Voorhis, W. C.; Rodriguez, J.; Stewart, L.; Montelione, G. T.; Craik, D.; Baker, D. Accurate de Novo Design of Membrane-Traversing Macrocycles. *Cell* **2022**, *185* (19), 3520-3532.e26. https://doi.org/10.1016/j.cell.2022.07.019.

(7) Craik, D. J.; Kaas, Q.; Wang, C. K. A Practical Guide to Structural Aspects of Macrocycles (NMR, X-Ray, and Modeling). In *Practical Medicinal Chemistry with Macrocycles*; Marsault, E., Peterson, M. L., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2017; pp 25–57. https://doi.org/10.1002/9781119092599.ch2.

(8) Guttman, M.; Lee, K. K. Isotope Labeling of Biomolecules: Structural Analysis of Viruses by HDX-MS. *Meth. Enzymol.* **2016**, *566*, 405–426. https://doi.org/10.1016/bs.mie.2015.05.021.

(9) Ständer, S.; R. Grauslund, L.; Scarselli, M.; Norais, N.; Rand, K. Epitope Mapping of Polyclonal Antibodies by Hydrogen–Deuterium Exchange Mass Spectrometry (HDX-MS). *Anal. Chem.* **2021**, *93* (34), 11669–11678. https://doi.org/10.1021/acs.analchem.1c00696.

(10) Lam, K. H. B.; Le Blanc, J. C. Y.; Campbell, J. L. Separating Isomers, Conformers, and Analogues of Cyclosporin Using Differential Mobility Spectroscopy, Mass Spectrometry, and Hydrogen–Deuterium Exchange. *Anal. Chem.* **2020**, *92* (16), 11053–11061. https://doi.org/10.1021/acs.analchem.0c00191.

(11) Englander, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical Perspective. *J Am Soc Mass Spectrom* **2006**, *17* (11), 1481–1489. https://doi.org/10.1016/j.jasms.2006.06.006.

(12) Linker, S. M.; Schellhaas, C.; Kamenik, A. S.; Veldhuizen, M. M.; Waibl, F.; Roth, H.-J.; Fouché, M.; Rodde, S.; Riniker, S. Lessons for Oral Bioavailability: How Conformationally Flexible Cyclic Peptides Enter and Cross Lipid Membranes. *J. Med. Chem.* **2023**. https://doi.org/10.1021/acs.jmedchem.2c01837.

(13) Rezai, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P. Conformational Flexibility, Internal Hydrogen Bonding, and Passive Membrane Permeability: Successful in Silico Prediction of the Relative Permeabilities of Cyclic Peptides. *J. Am. Chem. Soc.* **2006**, *128* (43), 14073–14080. https://doi.org/10.1021/ja063076p.

(14) Wang, C. K.; Swedberg, J. E.; Harvey, P. J.; Kaas, Q.; Craik, D. J. Conformational Flexibility Is a Determinant of Permeability for Cyclosporin. *J. Phys. Chem. B* **2018**, *122* (8), 2261–2276. https://doi.org/10.1021/acs.jpcb.7b12419.

(15) Ahlbach, C. L.; Lexa, K. W.; Bockus, A. T.; Valerie Chen; Crews, P.; Jacobson, M. P.; Lokey, R. S. Beyond Cyclosporine A: Conformation-Dependent Passive Membrane Permeabilities of Cyclic Peptide Natural Products. *Future Med Chem* **2015**, *7* (16), 2121–2130. https://doi.org/10.4155/fmc.15.78.

(16) Hyung, S.-J.; Feng, X.; Che, Y.; Stroh, J.; Shapiro, M. Detection of Conformation Types of Cyclosporin Retaining Intramolecular Hydrogen Bonds by Mass Spectrometry. *Analytical and bioanalytical chemistry* **2014**, *406*. https://doi.org/10.1007/s00216-014-8023-1.

(17) Lee, D.; Lee, S.; Choi, J.; Song, Y.-K.; Kim, M. J.; Shin, D.-S.; Bae, M. A.; Kim, Y.-C.; Park, C.-J.; Lee, K.-R.; Choi, J.-H.; Seo, J. Interplay among Conformation, Intramolecular Hydrogen Bonds, and Chameleonicity in the Membrane Permeability and Cyclophilin A Binding of Macrocyclic Peptide Cyclosporin O Derivatives. *J Med Chem* **2021**, *64* (12), 8272–8286. https://doi.org/10.1021/acs.jmedchem.1c00211.

(18) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chemistry & Biology* **2014**, *21* (9), 1115–1142. https://doi.org/10.1016/j.chembiol.2014.08.013.

(19) Garcia Jimenez, D.; Vallaro, M.; Rossi Sebastiano, M.; Apprato, G.; D'Agostini, G.; Rossetti, P.; Ermondi, G.; Caron, G. Chamelogk: A Chromatographic Chameleonicity Quantifier to Design Orally Bioavailable Beyond-Rule-of-5 Drugs. *J. Med. Chem.* **2023**, *66* (15), 10681–10693. https://doi.org/10.1021/acs.jmedchem.3c00823.

(20) Dougherty, P. G.; Sahni, A.; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem Rev* **2019**, *119* (17), 10241–10287. https://doi.org/10.1021/acs.chemrev.9b00008.

(21) Saar, K.; Lindgren, M.; Hansen, M.; Eiríksdóttir, E.; Jiang, Y.; Rosenthal-Aizman, K.; Sassian, M.; Langel, Ü. Cell-Penetrating Peptides: A Comparative Membrane Toxicity Study. *Analytical Biochemistry* **2005**, *345* (1), 55–65. https://doi.org/10.1016/j.ab.2005.07.033.

(22) Aguilera, T. A.; Timmers, M. M.; Olson, E. S.; Jiang, T.; Tsien, R. Y. Systemic in Vivo Distribution of Activatable Cell Penetrating Peptides Is Superior to Cell

Penetrating Peptides. *Integr Biol (Camb)* **2009**, *1* (5–6), 371–381. https://doi.org/10.1039/b904878b.

(23) Lafarga, V.; Sirozh, O.; Díaz-López, I.; Galarreta, A.; Hisaoka, M.; Zarzuela, E.; Boskovic, J.; Jovanovic, B.; Fernandez-Leiro, R.; Muñoz, J.; Stoecklin, G.; Ventoso, I.; Fernandez-Capetillo, O. Widespread Displacement of DNA- and RNA-binding Factors Underlies Toxicity of Arginine-rich Cell-penetrating Peptides. *The EMBO Journal* **2021**, *40* (13), e103311. https://doi.org/10.15252/embj.2019103311.

(24) Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. N-Methylation of Peptides: A New Perspective in Medicinal Chemistry. *Acc. Chem. Res.* **2008**, *41* (10), 1331–1342. https://doi.org/10.1021/ar8000603.

(25) Nguyen, A. M. T.; Brettell, S.; Douanne, N.; Duquette, C.; Corbeil, A.; Fajardo, E. F.; Olivier, M.; Fernandez-Prada, C.; Lubell, W. D. Influence of N-Methylation and Conformation on Almiramide Anti-Leishmanial Activity. *Molecules* **2021**, *26* (12), 3606. https://doi.org/10.3390/molecules26123606.

(26) Murphree, T.; Vorauer, C.; Brzoska, M.; Guttman, M. Imidazolium Compounds as Internal Exchange Reporters for Hydrogen/Deuterium Exchange by Mass Spectrometry. *Anal Chem* **2020**, *92* (14), 9830–9837. https://doi.org/10.1021/acs.analchem.0c01328.

(27) Watson, M. J.; Harkewicz, R.; Hodge, E. A.; Vorauer, C.; Palmer, J.; Lee, K. K.; Guttman, M. Simple Platform for Automating Decoupled LC–MS Analysis of Hydrogen/Deuterium Exchange Samples. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (2), 597–600. https://doi.org/10.1021/jasms.0c00341.

(28) Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of Overlapped and Noisy Hydrogen/Deuterium Exchange Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (12), 1906–1912. https://doi.org/10.1007/s13361-013-0727-5.

(29) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993**, *17* (1), 75–86. https://doi.org/10.1002/prot.340170110.

(30) Nguyen, D.; Mayne, L.; Phillips, M. C.; Walter Englander, S. Reference Parameters for Protein Hydrogen Exchange Rates. *J Am Soc Mass Spectrom* **2018**, *29* (9), 1936–1939. https://doi.org/10.1007/s13361-018-2021-z.

(31) Walters, B. T. Empirical Method To Accurately Determine Peptide-Averaged Protection Factors from Hydrogen Exchange MS Data. *Anal. Chem.* **2017**, *89* (2), 1049–1053. https://doi.org/10.1021/acs.analchem.6b03908.

(32) James, E. I.; Murphree, T. A.; Vorauer, C.; Engen, J. R.; Guttman, M. Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chem. Rev.* **2022**, *122* (8), 7562–7623. https://doi.org/10.1021/acs.chemrev.1c00279.

(33) Duengo, S.; Muhajir, M. I.; Hidayat, A. T.; Musa, W. J. A.; Maharani, R. Epimerisation in Peptide Synthesis. *Molecules* **2023**, *28* (24), 8017. https://doi.org/10.3390/molecules28248017.

(34) Ono, S.; Naylor, M. R.; Townsend, C. E.; Okumura, C.; Okada, O.; Lee, H.-W.; Lokey, R. S. Cyclosporin A: Conformational Complexity and Chameleonicity. *J Chem Inf Model* **2021**, *61* (11), 5601–5613. https://doi.org/10.1021/acs.jcim.1c00771.

(35) Silzel, J. W.; Murphree, T. A.; Paranji, R. K.; Guttman, M. M.; Julian, R. R. Probing the Stability of Proline Cis/Trans Isomers in the Gas Phase with Ultraviolet Photodissociation. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (9), 1974–1980. https://doi.org/10.1021/jasms.0c00242.

# CHAPTER 5.

# Perspectives and Future Directions

## 5.1 SUMMARY OF WORKS PRESENTED

Translating excitement into progress can prove difficult. For cyclic peptides, years of anticipation of a breakthrough therapeutic drug has fueled investment and research with minimal success. Not without payoff, that investment has resulted in the number FDA approved peptide drugs increasing in every decade[1]. The most recent success story comes in the form of Voclosporin, an orally bioavailable head-to-tail cyclized peptide that targets an intracellular protein[2]. It is the only FDA approved cyclic peptide drug in the last decade and a half[3]. Despite being a reason for optimism, voclosporin is a natural product analogue of cyclosporine and the clinical success of synthetic designed cyclic peptides remains elusive. The majority of indications for cyclic peptides remain antibiotics, antifungals, or gut-health. Cyclic peptides have yet to reach their potential as therapeutics.

We have identified two large barriers to the development of cyclic peptide therapeutics—identifying peptides with target interaction and understanding how to make structured cyclic peptides designed for permeability. Alone, accomplishing either task is feasible. Designing cyclic peptides for target interaction is now met with anticipated success. Likewise, guiding principles for design of structurally stable

permeable peptides with optimized pharmacokinetic profile are relatively well understood. Unfortunately, these goals are often diametrically opposed. Only by pushing the limits of experimental characterization and discovery can these goals be combined.

In Chapter 1 we briefly introduce cyclic peptides, the importance of their structure prediction and modeling, and the difficulties that exist in engineering cyclic peptide therapeutics. Fundamental chemical underpinnings of the processes in this work—peptide fragmentation, target interaction, and structural determination are highlighted. The foundational instrumentation of this work, the mass spectrometer, gets some introduction as a powerful tool with diverse applications. Indeed, several different applications and types of mass spectrometers are described in this work for their use in characterizing cyclic peptides.

In Chapter 2 the mass spectrometer is used to probe the driving principles of peptide fragmentation. Understanding these principles and the resulting mass spectra generated by cyclic peptides facilitated development of a computational program to assist in sequencing cyclic peptides. Despite spectral complexity, we demonstrated accuracy in computational sequencing of cyclic peptides. In a novel application, we demonstrated that $MS^3$ spectra can enhance cyclic peptide sequencing and pair with $MS^2$ spectra to make a powerful tool. This platform enables the use of affinity selection mass spectrometry (ASMS) techniques to screen cyclic peptides in high throughput for target interaction. In Chapter 3 the ASMS platform is extended to complex environments. G-protein coupled receptors (GPCRs) are key drivers of biological functions within the body and have historically been high-impact drug targets. We

solved challenges with maintaining GPCRs in their functional state in ASMS experiments by sequestering GPCRs in HIV-inspired virus-like particles (VLPs). These self-assembling GPCR-VLPs were amenable to ASMS protocols. Even in this complexed environment, mass spectrometers showed the capability to identify target interaction between peptide and GPCR target.

Finally, in Chapter 4, we investigated the underpinnings of structural flexibility and its impact on permeability using HDX-MS. HDX-MS was explored as a powerful technique for evaluating cyclic peptide stability. We unveiled the different relationships structural stability can have as a driver of permeability. Notably, HDX-MS revealed the structural flexibility of the natural product cyclosporine A. This flexibility served as a template for identification of new cyclic peptides that possess the ability to switch conformational states, likely as the result of proline isomerization.

In summary, we have built and extended a platform for high throughput screening of cyclic peptides facilitated by computational sequencing. We demonstrated that tandem MS sequencing is possible for cyclic peptides and have a computational pipeline in place to aid these experiments. We showed that ASMS is possible against membrane-bound protein like GPCRs and have installed a portable and tunable platform for pursuing these proteins as therapeutic targets. We also characterized one of the lesser appreciated aspects of permeability using HDX-MS: structural stability. We showed that structural stability can be necessary for the permeability of some peptides, and that flexibility can be necessary for some others.

## 5.2 FUTURE DIRECTIONS

*Integration with machine learning tools*

Structure prediction of cyclic peptides, the key to intentional design towards any target, has gotten better and faster with deep learning tools. *In silico,* it is now easier to make more peptide designs, find more diverse structures, and assess binding interactions with greater accuracy than ever before.

Machine learning approaches to drug discovery may prove more successful than high throughput screening. Advances in deep learning design of cyclic peptide have unlocked the ability to create cyclic peptides with a high success rate of strong interaction with a target protein. Small batch *in silico* design of cyclic peptides may surpass high throughput screening as an efficacious tool for drug discovery. ASMS is an excellent tool for affinity maturation of lead compounds identified via computational prediction. As of now, deep learning models do not have a rule set to predict the interactions and presence of noncanonical amino acids and peptide modifications. Incorporation of these nonstandard building blocks into lead hit peptides can drive affinity beyond the original compound or maintain target affinity while introducing better pharmacokinetic properties. Affinity maturation experiments offer the inclusion of a positive control in the form of the lead compound, which can make assay development and testing much easier. Direct comparisons between the strength of the original compound and discovered variants will ease development as well. Lead peptides will need affinity maturation and sequence mutational studies to identify key binding interactions, and ASMS provides a compromise-free platform to do so.

Several new programs have taken advantage of the extensive public library of peptide fragmentation data to train models capable of sequencing peptides[4–6]. While these tools have shown adequate accuracy and speed, the training sets are typically limited to naturally occurring linear peptides made of canonical amino acids. This restricts their application to proteomics, mostly. The interest in therapeutic cyclic peptides raises the possibility of training a new model for cyclic peptide sequencing or retraining an existing model for cyclic peptide capability. The amount of cyclic peptide fragmentation data currently available is the largest hurdle to accomplishing this. In that regard, data generation faces two problems: obtaining a large number of diverse cyclic peptides and creating accurately annotated spectra. An ambitious project could scale the amount of cyclic peptide fragmentation data by synthesizing libraries of cyclic peptides and annotating it with Cyclic Comet or CycLS. This would effectively replace Cyclic Comet and CycLS with a deep learning model, but its accuracy and capabilities might be limited by the accuracy and capabilities of Cyclic Comet and CycLS. It is worth considering if a deep learning trained model for sequencing would represent an improvement on these existing techniques or not.

In addition to affinity maturation, ASMS provides a high throughput approach to screening head-to-tail cyclization replacements for alternative cyclization strategies like disulfide bonds. Display-based high throughput screening strategies like yeast display are effective at identifying lead compounds with target interaction, but the cyclization strategies will have to be replaced by head-to-tail cyclization to make a feasible drug candidate. ASMS can easily screen through two and three amino acid motifs to identify

peptides that cyclize and retain their target interaction. This adjustment would be the first step in maturation of a disulfide-cyclized lead compound.

## 5.3 OPPORTUNITIES FOR IMPROVING THE PRESENT WORK

### *Opportunities for Improving Cyclic Comet*

Cyclic Comet sequencing treats each scan as a standalone experiment and is ignorant of information from surrounding scans and related sequence identifications. The tool described here to pair $MS^2$ and $MS^3$ scans is completed post-processing, after Cyclic Comet's metrics have been established. As a result, there is more that can be done to perfectly integrate the data between linked scans. The opportunity exists to pass sequence information between $MS^3$ and $MS^2$, using composition and sequence motif information to filter identifications that may otherwise appear reliable. Relatedly, though the post processing score has some influence from compositional amino acid agreement, using multiple sequence assignments from a single scan to assign a cohesive sequence offers the opportunity to increase sequence assignment confidence. This is already done in a program like CycLS, which employs multiple sequence assignments to rebuild the whole. Adopting a similar approach should reveal more opportunities to evaluate a sequence match.

### *Exploring GPCRs as Targets for Cyclic Peptides*

ASMS was expanded to GPCRs with the aim of screening cyclic peptides against the GPCR APJ. To aid this, APJ was the focus of computational design. APJ has been previously targeted by cyclic peptide design with some success, demonstrating the feasibility of designing peptides against this target[7–9]. As a membrane protein, APJ

presents its own set of problems specific to computational design of cyclic peptides. It has a pocket that restricts sizes, it is a large protein that demands computationally intensive structure prediction, and it has hydrophobic surfaces meant for membrane insertion. Nonetheless, the next step in our project is to design and optimize cyclic peptide binders computationally before synthesis and testing. Pairing this approach with our newly developed GPCR-VLPs ASMS remains a feasible option, and APJ remains an attractive drug target.

Efforts to screen using GPCR-VLPs should focus on maximizing the ability to express and collect large numbers of VLPs. At the scale used in this paper, each ASMS experiment represented a full round of cell growth, VLP expression, VLP generation, collection, and purification before any screening could even start. Future researchers should endeavor to optimize their expression first, as having enough VLPs to run multiple ASMS experiments from a single expression batch will both reduce the amount of investment needed for ASMS experiments and reduce the amount of variability between experiments. Seriously consideration should be given to the necessity for the magnetic bead vs. other retention systems. The size of the VLP gives it an advantage for retention through size filters that other protein only ASMS systems do not have. Rapid equilibrium dialysis is a screening technique that we tried first with native GPCR in a stabilizing buffer. Though we failed to demonstrate enriched binding with native protein, equilibrium dialysis may hold value using VLPs. The porous membrane separating the wells has size cutoffs which can easily separate peptide from VLP. This would have the added benefit of reducing the complexity of the system.

*Mutational Tolerance of Flexible Peptides*

Though the common-sense approach is to find a lead compound and optimize for permeability later, computational advancement offers the opportunity to generate thousands of high-quality peptide designs and filter for permeability directly. I suggest a large design run and in addition to filtering by target interaction metrics, filter by the metrics that typically predict permeability. HDX-MS was used to study the mutational tolerance of cyclic peptide scaffolds whose permeability was driven by structural stability. HDX-MS was also used to suggest new cyclic peptides with sustainable conformational flexibility. Repeating the mutational tolerance studies using conformationally flexible scaffolds will give crucial insights into the maturation of flexible cyclic peptides. It is unclear what impact amino acid mutations will have on conformational flexibility and what impact that flexibility change will have on permeability. Identifying trends in mutations could be important for encouraging flexibility in future peptide design. Likewise, finding positions where mutational tolerance is high is important for optimizing flexible scaffolds for target interaction.

Conformationally flexible peptides will be difficult to model for standard pharmacokinetic properties. The two distinct structures will possess separate profiles, and the interpretation will be further muddied by the interconversion of the profiles. Experimental characterization of these compounds will be particularly important. Some work has already been put into using metabolic fractions to characterize the stability of cyclic peptides in our group, and conformationally flexible peptides are prime candidates for testing.

## 5.4 Concluding Remarks

In summary, these works detail the development of tools designed to ease and enable future cyclic peptide drug discovery and development. Substantial efforts have been made to facilitate cyclic peptide drug discovery, though much of the ASMS detailed here is translatable to other modalities. We have designed and developed a platform that resulted in the discovery of novel cyclic peptide binders. HDX-MS was used to study cyclic peptides in a unique manner and proved to be an incredibly insightful assay. This experimental work has established platforms for discovery and characterization of cyclic peptides. Their continued use and development should aid meaningful discovery efforts for clinically relevant problems.

## 5.5 BIBLIOGRAPHY

(1) Costa, L.; Sousa, E.; Fernandes, C. Cyclic Peptides in Pipeline: What Future for These Great Molecules? *Pharmaceuticals* **2023**, *16* (7), 996. https://doi.org/10.3390/ph16070996.

(2) Rovin, B. H.; Teng, Y. K. O.; Ginzler, E. M.; Arriens, C.; Caster, D. J.; Romero-Diaz, J.; Gibson, K.; Kaplan, J.; Lisk, L.; Navarra, S.; Parikh, S. V.; Randhawa, S.; Solomons, N.; Huizinga, R. B. Efficacy and Safety of Voclosporin versus Placebo for Lupus Nephritis (AURORA 1): A Double-Blind, Randomised, Multicentre, Placebo-Controlled, Phase 3 Trial. *Lancet Lond. Engl.* **2021**, *397* (10289), 2070–2080. https://doi.org/10.1016/S0140-6736(21)00578-X.

(3) Zhang, H.; Chen, S. Cyclic Peptide Drugs Approved in the Last Two Decades (2001–2021). *RSC Chem. Biol.* **2022**, *3* (1), 18–31. https://doi.org/10.1039/D1CB00154J.

**(4)** Gagneur, J. Deep Learning-Driven Fragment Ion Series Classification Enables Highly Precise and Sensitive de Novo Peptide Sequencing. *Nat. Commun.* **2024**, *15* (1), 151. https://doi.org/10.1038/s41467-023-44323-7.

(5) Tran, N. H.; Qiao, R.; Xin, L.; Chen, X.; Liu, C.; Zhang, X.; Shan, B.; Ghodsi, A.; Li, M. Deep Learning Enables de Novo Peptide Sequencing from Data-Independent-Acquisition Mass Spectrometry. *Nat. Methods* **2019**, *16* (1), 63–66. https://doi.org/10.1038/s41592-018-0260-3.

(6) Tran, N. H.; Zhang, X.; Xin, L.; Shan, B.; Li, M. De Novo Peptide Sequencing by Deep Learning. *Proc. Natl. Acad. Sci.* **2017**, *114* (31), 8247–8252. https://doi.org/10.1073/pnas.1705691114.

(7) Chapman, F. A.; Maguire, J. J.; Newby, D. E.; Davenport, A. P.; Dhaun, N. Targeting the Apelin System for the Treatment of Cardiovascular Diseases. *Cardiovasc. Res.* **2023**, *119* (17), 2683–2696. https://doi.org/10.1093/cvr/cvad171.

(8) Johnson, J. A.; Kim, S.-H.; Jiang, J.; Phillips, M.; Schumacher, W. A.; Bostwick, J. S.; Gargalovic, P. S.; Onorato, J. M.; Luk, C. E.; Generaux, C.; He, Y.; Chen, X.-Q.; Xu, C.; Galella, M. A.; Wang, T.; Gordon, D. A.; Wexler, R. R.; Finlay, H. J. Discovery of a Hydroxypyridinone APJ Receptor Agonist as a Clinical Candidate. *J. Med. Chem.* **2021**, *64* (6), 3086–3099. https://doi.org/10.1021/acs.jmedchem.0c01878.

(9) Ma, Y.; Yue, Y.; Ma, Y.; Zhang, Q.; Zhou, Q.; Song, Y.; Shen, Y.; Li, X.; Ma, X.; Li, C.; Hanson, M. A.; Han, G. W.; Sickmier, E. A.; Swaminath, G.; Zhao, S.; Stevens, R. C.; Hu, L. A.; Zhong, W.; Zhang, M.; Xu, F. Structural Basis for Apelin Control of the Human Apelin Receptor. *Struct. Lond. Engl. 1993* **2017**, *25* (6), 858-866.e4. https://doi.org/10.1016/j.str.2017.04.008.

# VITA

Jonathan Palmer hails from Hopkins, Minnesota where he attended Hopkins High School. While studying biochemistry in undergrad, Jonathan was a student athlete and member of the men's soccer team. Jonathan worked as a student researcher on stability and identification of commercial oils and helped develop a colorimetric infection sensing plastic for wound care while studying at Queen's University Belfast. Jonathan joined the University of Washington's medicinal chemistry graduate program directly after graduating in 2019. As a member of the Guttman and Bhardwaj research groups, he studied cyclic peptides and their application as therapeutics using high-throughput mass spectrometry-based approaches.