UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mono and Bicyclic Peptide as Probes for Diverse Biological Applications

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Rohit Chaudhuri

September 2021

Dissertation Committee:
Dr. Min Xue, Chairperson
Dr. Joseph Genereux
Dr. Wenwan Zhong

The Dissertation of Rohit Chaudhuri is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

I feel extremely honored and fortunate to be able to be in this moment at the doorstep of earning my PhD degree and I have a heartfelt gratitude to a long list of people who have contributed to my destiny. First and foremost, I cannot be more thankful to my PhD advisor, Dr. Min Xue for everything. He has been a wonderful mentor and more importantly, a great person who has trusted me more than I trusted myself. The research projects of our lab, the enjoyable learning environment, man-management were all brilliantly organized by Dr. Xue. He helped me believe academia can be enjoyable.

I am also thankful to my fellow PhD lab-mates- Priyanka Sarkar, Siwen Wang, Nicole Perkins, Zhili Guo, Fei Ji, Nathan Guevara, Boxi Chen for the friendship which contributed immensely to my research. It was a great pleasure to work in collaboration with our two postdocs, Dr. Zhonghan Li and Dr. Shiqun Shao. Nothing would have been possible without their constant support and guidance. The importance of having a wonderful camaraderie, closely associated with the professional environment can never be underestimated. My advisor along with my lab-mates, together made my graduate life, an absolute delight.

I am thankful to my undergraduate advisor, Dr. Bhubaneswar Mandal, for the encouragement to pursue a research career in distant United States and instilling in me the liking towards peptide chemistry. I had begun appreciating the work ethics of a researcher, the invaluable importance of a sound temperament since my days in my

undergraduate research lab. My then lab members, especially my senior Dr. Jyoti Chandra, encouraged me a lot. Being at the Department of Chemistry at Indian Institute of Technology, has been particularly, inspiring experience.

I am very thankful to my PhD defense committee members, Dr. Joseph Genereux, Dr. Wenwan Zhong for valuable inputs during my candidacy examination. Their advice has been of immense help to prepare for a journey of a researcher. I am very thankful to Prof. Jason Cheng, for providing me an opportunity to collaborate in research with his lab. I am also thankful to Fatimah Abouhajar, from Dr. Cheng's lab for participating in a very interesting collaborative project.

I am also thankful to all the professors of Analytical Chemistry for organizing very insightful seminars and patiently listening to my presentations which were a part of my PhD curriculum. I am very thankful to the Professors and Department of Chemistry, UCR for providing me with the facilities for seamless endeavor in research activities. I am thankful to all the Staff from Department of Chemistry, UCR, for aiding during need. I am very thankful to Dr. Kevin Simpson, coordinator for General Chemistry teaching labs, for efficient organization of teaching schedules, which made my life as a TA, particularly smooth and enjoyable.

My childhood friends back in India, have been very instrumental in creating a perfect balance between life and work especially on particularly hectic days. I could not be more

thankful to my group of friends for the relief, they provided in all these years. I am also thankful to my roommates and other friends in Riverside, California for the occasional breaks in these years. My cousins have also been inspirational during my journey. Finally, no achievement in life is ever possible without the support of family. The love, sacrifices and blessings of my parents have shaped my journey throughout and more importantly, grew me as an individual capable of facing challenges and appreciating them irrespective of the result. The mental strength, the sincerity and diligence required for a life of a researcher can only be developed over several years, and my parents have been very instrumental in the process. A big part of my family also involves Suchitra, my partner. She has been supportive, a pillar of my strength and a constant influence throughout the ups and downs of my life. The bond with her, her understanding of my approach to life, made me navigate through these five years seamlessly. She propelled me into never stop believing in my dreams and continue to have faith in my ambitions. Her simplicity in living and uniqueness in enjoying the joys of life have always helped me overcome and appreciate the slightly tougher phases of life.

Dedication

To the relentless COVID-19 warriors across the globe

ABSTRACT OF THE DISSERTATION

Mono and Bicyclic Peptide as Probes for Diverse Biological Applications

by

Rohit Chaudhuri

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, September 2021
Dr. Min Xue, Chairperson

Macrocyclic peptides are proven binders against structureless proteins. The complexity incorporated in multiple cyclization within a peptide backbone demands demonstration in multifaceted binding activity. We introduce the applications of cyclic peptide and an outline of the various research projects we aim to develop around them. As a starting point, we demonstrate the scope of our bicyclic peptides as highly specific fluorescent reporter probes for MAPK/ERK signaling pathway, often aberrantly involved in cancer. Bicyclic peptides can be designed to be highly specific binders for homologous ERK1 and 2 and can signal its phosphorylation. A library of significant diversity can assist in discovering such probes. The transcription factor c-Myc is a proto-oncogene which is overexpressed in most known cancers. c-Myc is commonly considered “undruggable” due to its intrinsically disordered structure. We had previously discovered a bicyclic peptide that could effectively inhibit the proliferation of a c-Myc-dependent cell line. Its drawbacks as a therapeutic were addressed through a series of medicinal chemistry modifications. We also address cytotoxicity concerns, applicability across multiple cancer cell lines, and mechanistic

details of suppressed proliferation through multiple optimization processes. Additionally, we explored the potential of peptides as miniature enzyme and the structural considerations necessary for designing such molecules.

Table of Contents

Chapter 1: Introduction

1.1 Peptides as Protein-protein interactions inhibitors……..... 1

1.2 Usage of peptide libraries discover capable binder peptides……... 2

1.3 Targeting intrinsically disordered proteins using peptide library..... 3

1.4 Tagged cyclic peptides as specific probes……..... 4

1.5 Summary …….. 4

1.6 References……. 5

Chapter 2: Bicyclic peptides as fluorescent reporter probes for ERK activities

2.1 Introduction……. 7

2.2 Experimental……... 8

2.3 Results and Discussion…… 14

2.4 Conclusion……... 24

2.5 References……... 25

Chapter 3: Optimization of bicyclic peptide-based inhibitor for c-Myc and investigating its core biological impacts

3.1 Introduction……. 26

3.2 Experimental……... 28

3.3 Results and Discussion…… 42

3.4 Conclusion……... 58

3.5 References……... 58

Chapter 4: The exploration for short cyclic peptides with enzymatic activities

4.1 Introduction……. 60

4.2 Experimental…… 61

4.3 Results and Discussion…… 66

4.4 Conclusion……... 72

4.5 References…… 73

Chapter 5: Concluding remarks…….. 74

List of Figures

Figure 2.1 Schematic representation of screening bicyclic peptide library. An alkyne handle is conjugated to the library and is screened against dye-tagged proteins. The unreacted ones are reacted with an azide-containing epitope with a biotin tag. The hit beads undergo enzyme amplification through a biotin-streptavidin specific reaction and a corresponding Alkaline phosphatase reacts with BCIP-NBT to form purple-colored beads due to precipitation. The hit beads are then linearized 16

Figure 2.2 a) Epitope 1- G-F-L-T-E-Y-V-A-T-R-W-Y-R and Epitope 2- Y-L-E-Q-Y-Y-DP-T-D-P-T-D-E-P-V-A-E. Epitope 1 has 100% sequence match with both ERK 1 and ERK 2, while Epitope 2 has 100% match with ERK 1 and 87% with ERK 2. **b)** Biotin-PEG5 linker attached to both these epitopes at their N-terminals to aid the screening process.....17

Figure 2.3 Bicyclic library structure, used for screening. The table shows the hits obtained from the library with the respective amino acids at the designated positions 18

Figure 2.4 a) Rhodamine (red)-Linker and **b)** Fluorescein (blue)-Linker combinations are conjugated to hit peptides and epitopes respectively, at their N-termini as tags for FRET assay.....19

Figure 2.5 FRET binding curves for rhodamine tagged B22 against fluorescein tagged Epitope 1 and rhodamine tagged C14, A16 and D23 against Fluorescein tagged Epitope 2, these are the better binders among the hits obtained from screening 20

Figure 2.6 Validation. **a)** Crystal structure of ERK1 from PDB. **b)** FP assay showing the binding between ERK1 and Rhodamine-tagged C14. **c)** FP assay showing the binding between ERK1 and Rhodamine-tagged B22. **d)** Crystal structure of ERK1 from PDB. **e)** FP assay with showing absolutely no binding between ERK1 and Rhodamine-tagged C14. **f)** FP assay with showing absolutely no binding between ERK1 and Rhodamine-tagged C14.....21

Figure 2.7 B22 (left) and C14 (right) are the best binders, proven through FRET binding data against the epitopes and the FP binding data against the proteins 22

Figure 2.8 We have a pair of probes in Cy5 tagged B22 and Rhodamine tagged C14, phosphorylated ERK1, can have only probe bound to it, resulting in absence of FRET. The unphosphorylated/ inactive ERK1 will have both probes bound to it, resulting in FRET phenomenon and an increased signal.....23

Figure 2.9. a) Serum starved U87 cells were stimulated with EGF, followed by treatment MEK inhibitor trametinib. **b)** Changes in FRET signal in the lysates that were collected

right after treatment with EGF and 12 0 inhibitor. **c)** An increase in active MEK, leads to a decrease in the FRET signal 24

Chapter 3

Figure 3.1 MALDI-TOF mass spectrum for NT-A1. Expected mass from calculation 1328.66. Mass found 1328.63.....35

Figure 3.2 MALDI-TOF mass spectrum for NT-A2. Expected mass from calculation-1300.63. Mass found 1300.6315.....36

Figure 3.3 MALDI-TOF mass spectrum for NT-A3. Expected mass from calculation 1284.67. Mass found 1285.5879.....36

Figure 3.4 MALDI-TOF mass spectrum for NT-A4. Expected mass from calculation 1243.59. Mass found 1243.437.....37

Figure 3.5 MALDI-TOF mass spectrum for NT-A5. Expected mass from calculation 1243.59. Mass found 1243.5734.....37

Figure 3.6 MALDI-TOF mass spectrum for NT-A7. Expected mass from calculation 1300.65. Mass found 1300.7915.....38

Figure 3.7 MALDI-TOF mass spectrum for NT-A8. Expected mass from calculation 1300.65. Mass found 1300.5098.....38

Figure 3.8 MALDI-TOF mass spectrum for NT-A9. Expected mass from calculation 1300.65. Mass found 1300.5098.....39

Figure 3.9 MALDI-TOF mass spectrum for NT-A9. Expected mass from calculation 1342.67. Mass found 1342.7187.....39

Figure 3.10 Figure 3.10 MALDI-TOF mass spectrum for NT-A10. Expected mass from calculation 1327.71. Mass found 1328.647 40

Figure 3.11 a) Bicyclic peptide library structure **b)** Alanine scanning variants of NT-A1 **c)** Competitive FRET studies efficiencies **d)** Table summarizing better binders (green) and worse binders (red) than NT-A1 41

Figure 3.12 a) Three variants of NT-A1 aimed at optimization- NT-A7, NT-A8 and NT-A9 and their respective Competitive FRET studies. **b)** Table summarizing the performance of the variants.....44

Figure 3.13 Three variants of NT-A1 developed through chopping and changing amino acids. NT-A7 with arginine changed to lysine, NT-A10 with glutamic acid changed to arginine from NT-A10, and NT-A14 with a shorter ring sized version of NTA10.....45

Figure 3.14 ELISA experiments probing binding of NT-A7 and NT-A10 against the recombinant c-Myc protein.....46

Figure 3.15 a) NT-A14 **b)** ELISA experiment probing the binding between NT-A14 and recombinant protein c-Myc c Cell viability of U87 cells after treatment with NT-A14 entrapped in liposomes 48

Figure 3.16 Cell viability data of WM115 (top) and MCF-7 (bottom) cells after treating them with NT-A14, the positive controls (green bar) showing cells treated with media only, and (blue bar) showing cells treated with liposomes 49

Figure 3.17 RNA-seq analysis of U87 cells treated with the NT-A14. **a)** Principal component analysis result of the RNA-seq datasets. **b)** Volcano plot showing the overall landscape of the differentially expressed genes in the NT-A14 24 h treatment versus the control and the 24 h treatment samples. (Cutoff criteria: $p < 0.01$, fold 1,5) **c)** Volcano plot showing the overall landscape of differentially expressed genes as the result of 24h treatment (left) and 48 h treatment (right). **d)** Cutoff criteria: $p < 0.01$, fold change > 1 (left) and fold change >1.5 (right) 50

Figure 3.18 a) Gene sets enriched in the 24 h liposome treated sample vs NT-A14 treatment for 24 h. Enrichment plot of MYC_TARGETS_V1 after **b)** 24 h NT-A14 treatment and **c)** 48 h NT-A14 treatment. **d)** 20 most contributing genes after 24 h treatment **e)** T-test plots moderately high p (**) values between 0.01 and 0.001 for increased Myc expression levels 52

Figure 3.19 a) Gene sets enriched in the NT-A14 treatment after 24 h vs liposome treated sample. Enrichment plots for TNFA_SIGNALING_VIA_NFKB after **b)** 24 h treatment and **c)** 48 h treatment.....53

Figure 3.20 mRNA levels for a) FOS b) JUN c) FOS B d) JUN B. T-test was performed to assess the statistical significance of upregulation. *: $p < 0.05$. **: $p < 0.01$. ***: $p < 0.001$.....54

Figure 3.21 Gene sets enriched after **a)** 48 h liposome treatment vs 48 h NT-A14 and **b)** 48 h NT-A14 vs 48 h liposome treatment.....55

Chapter 4

Figure 4.1 The monocyclic peptide library synthesized on TentaGel (TG) beads, used for screening. X_1 -X_5 represents the five randomized amino acids added through five cycles of iterative split and pool library synthesis. Pra (propargyl glycine), Az_4 (azido lysine), M (methionine) are added to the entire set of beads for specific purposes 66

Figure 4.2 MS-MS spectra of hit sequence PHLYR, solving the spectra using the b and y ions provides us with the information of the hit sequence.....69

Figure 4.3 cy (P-H-LY-R) on TentaGel beads are capable of instantly reducing MTT solution (left) and NBT solution (right) into their respective formazans, characterized by intensely purple-colored beads 71

Chapter 1: Introduction

Short peptides (typically ten amino acids), which belong to the newest class of pharmaceutical drugs, have an intermediate molecular weight between small molecule drugs (<500 Da) and biologics (>5000 Da)[1,2] . Small molecules often suffer from a lack of specificity and are often limited by the requirement of protein-binding pockets at the targets. Biologics bring in selectivity and potency but lack the membrane permeability, stability, and bioavailability of small molecules. Short peptides bring together the advantages of both the traditional groups of therapeutics[1-4].

The potency and selectivity of peptide-based drugs can be improved by altering the amino acids, which we will touch upon in the later chapters. However, the greatest challenge remains in conferring stability to the peptide, particularly, to combat proteolytic cleavages and reduce the entropic penalty on binding to targets [1,5]. To address those challenges, cyclic peptides have been introduced, with inspirations from nature. Cyclization enables peptides to have a rigid backbone with the required complexity and be resistant to degradation[1]. Cyclic peptides can engage with targets with a significantly lesser loss of entropy than its linear version [6,7,8]. An aspect that makes them useful against targets without binding pockets. Cyclization can also facilitate cell penetrability [5,8].

1.1 Peptides as Protein-protein interactions inhibitors

Proteome-wide maps display thousands of protein-protein interactions (PPI)s that govern the three essential components of cellular life- growth, differentiation, and survival mediated through processes like transcription, translation, trafficking, energy production, protein folding, cytokinesis, and signaling. PPIs exist between proteins of various sizes,

and their interfaces usually span 1100–2000 $Å^2$. A few examples span up to 4500 $Å^2$ and fewer exist with as low as 600 $A°^2$ The combined free energy of Van der Waals interactions and desolvation stabilizes the PPIs [9,10]. Electrostatic forces and hydrogen bonds also play a significant role in forming the complexes [10]. PPIs often involve aromatic amino acids such as tyrosine and tryptophan, which engage in a pi-pi interaction[10,12].

PPIs have been extensively studied due to their applications in several diseases [8,13]. Targeting PPIs with small molecules have been prevalent over quite some time. However, the success rate has been extremely low due to the lack of hydrophobic pockets in most cases, and hence, PPIs have been generally considered "undruggable" [13, 14].

1.2 Usage of peptide libraries discover capable binder peptides

OBOC (one-bead-one-compound) peptide libraries can be synthesized by split-and-pool techniques [15,16]. Solid-phase peptide synthesis protocols are very well adaptable to this technique. The harsh reactants used in the peptide synthesis are easily washed away from the chemically inert solid support. Unnatural amino acids can be incorporated in the library, and sidechain mediated cyclization techniques can be introduced to form the complex macrocyclic library of peptides, with each bead carrying a unique sequence [16]. Another advantage lies in the ease with which the hits can be physically isolated after incubating them against a particular target. Cyclized peptides can be washed, linearized, detached from the solid support, and analyzed using Tandem Mass Spectrometry [16, 17].

The identified hit sequences can go through validation, optimization, and subsequent in-vitro experiments, leading to drug discovery.

In-situ click chemistry can be used for protein capture on the surface beads. Epitopes can be modified to have a click handle in the form of an azide, which can bind to an alkyne handle attached to the library[16-18]. Specific binding between the library and the epitope would enable the click reaction, which would not be entropically favorable otherwise[16,18]. This method helps to create a very specific interaction analogous to that between a protein and an antibody sans the high cost and other physicochemical disadvantages of antibodies.

1.3 Targeting intrinsically disordered proteins using peptide library

Intrinsically disordered proteins bind with a partner to compensate for the loss in ordering entropic penalty. Binding to these kinds of proteins, particularly, can be achieved by a complex cyclic peptide. Target-based screening against proteins does not require extensive knowledge of the protein structure. The diversity of the library plays a critical role in the success of the screening process[19]. Once again, a carefully designed macrocyclic library of substantial diversity can unearth a perfectly capable binder that may be modified and developed into a potential therapeutic, which we synthesize after adding even more complexity to the previously described peptide library synthesis.

1.4 Tagged cyclic peptides as specific probes

Several cascading pathways are often found to have aberrant signaling mechanisms in cancer. These signaling processes are dynamic. Well-known techniques like mass spectrometry and immunolabelling [20,21] are not suitable because they provide snapshot measurements. The specificity and potency of macrocyclic peptides can be used as fluorescent probes when they are tagged with a fluorophore. Our previous successes with macrocyclic peptides inspire another similarly designed screening process [22]. Unlike the previous quests with PPI inhibitors, we aim to probe the de-regulated PPIs as a part of the signaling process. Utilizing a pair of fluorescence probes provides a high-intensity FRET signal, which is possible only when both the probes bind to the target. Only one of the probes binding to the protein would indicate possible phosphorylation, a critical step in the cascading signaling pathway.

Our designed probes can be utilized for single-cell imaging experiments, which can provide information about cancer cell heterogeneity. Combining microscopy and statistical analysis gives us an insight into the reaction of each cell towards stress. Additionally, these probes can be easily removed by cell fixation.

1.5 Summary

In summary, cyclic peptides are versatile and capable tools as therapeutics, owing to their structure and functionality. In the following chapters, we will explore several facets of cyclic peptides and demonstrate their applications. In Chapter 2, we will explore the scope of designing a library of two million bicyclic peptides on polymer resin beads. We will demonstrate an effective method to utilize this diversity to discover specific

signaling probes against ERK proteins, which can be used to monitor their phosphorylation process. In Chapter 3, we will explore the process of optimizing a previously discovered bicyclic peptide by altering the amino acids in a way to optimize its binding against a disordered protein. Simultaneously, we will shed light on the various implications of such binding. In Chapter 3, we will explore the ability of a peptide to function as a miniature enzyme and the roadblocks that can arise while executing such a study.

1.6 References

1. Guillen Schlippe, Y,V.; Hartman, M.C.T.; Josephson, K; Szostak, J.W. *J. Am. Chem. Soc.* **2012**, *134*, 10469.

2. Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2014,** *114*, 901-926.

3. Dougherty, P.G.; Sahni, A; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019**, *119,* 10241-10287.

4. Morimotp, J.; Hayashi, Y.; Suga, H. Discovery of Macrocyclic peptides Armed with a Mechanism-Based Wargead: Isoform-Selective Inhibition of Human Deacetylase SIRT2. *Angew. Chem., Int. Ed.* **2012**, *51*, 3423-3427.

5. Bechara, C.; Sagan, S.; Cell-Penetrating Peptides: 20 Years Later, Where Do We Stand? FEBS Lett. **2013**, *587*, 1693- 1702.

6. Rennie, Y. K.; McIntyre, P. J.; Akindele, T.; Bayliss, R; Jamieson, A.G. A TPX2 Proteomimetic Has Enhanced Affinity for Aurora-A Due to Hydrocarbon Stapling of a Helix. *ACS Chem. Biol.* **2016**, *11*, 3383-3390.

7. Ogura, K.; Okamura, H. Conformational change of Sos-derived proline-rich peptide upon binding Grb2 N-terminal SH3 domain probed by NMR. *Sci. Rep.* **2013**, *3*, 2913.

8. Robertson, N. S.; Spring, D. R. Using peptidomimetics and constrained peptides as valuable tools for inhibiting protein-protein interactions. *Molecules* **2018**, *23*, 959.

9. Bourgeas, R.; Basse, M. J. ; Morelli, X.; Roche, P. Atomic analysis of protein-protein interfaces with known inhibitors : the 2P2I database. *PLoS One,* **2010**, *5*, e9598.

10. Thompson, A. D.; Dugan, A.; Gestwicki, J.E.; Mapp, A.K. Fine-Tuning Multi-Protein Complexes Using Small Molecules. *ACS Chem. Biol.* 2012, 7, 1311.

11. Moreira, I. S.; Fernandes, P. A.; Ramos, M.J. Hot spots- A review of protein-protein interface determinant amino-acid residues. Proteins 2007, 68, 803.

12. Fernandez, A; Scheraga, H.A. Insufficiently dehydrated hydrogen bonds as determinants of protein interactions. *Proc. Natl. Acad. Sci.* **2003**, *100*, 113-118.

13. Yan, C.; Higgins, P.J. Drugging the undruggable. Transcription therapy for cancer. *Biophys. Acta Rev Cancer* **2013**, *1835*, 76-35.

14. Zhang, G.; Andersen, J.; Gerona-Navarro, G. Peptidomimetics targeting protein-protein interactions for therapeutic development. *Protein Pept let.* **2018**, *25*, 1076-1089.

15. Lam, K.S; Lehman, A, L.; Song, A.; Doan, N.; Enstrom, A. Maxwell, J.; Liu, R. *Methods Enzymol.* **2003**, *369*, 298-322.

16. Das, S.; *et al.* A General Synthetic Approach for Designing Epitope Targeted Macrocyclic Peptide Ligands. *Angew Chem., Int. Ed.* **2015**, *54*, 13219-13224.

17. Matsueda, G. R.; Haber, E.; Margolies, M. N. *Biochemistry* **1981**, *20(9)*, 2571-2580.

18. Mocharla, V.P.; Colasson, B.; Lee, L.V.; Roper, S.; Sharpless, K.B.; Wong, -H. C.; Kolb, H.C.; *Angew Chem., Int. Ed.* **2005**, *44*, 116-120.

19. Modell, A.E.; Blosser, S.L.; Arora, P. S.; Systematic targeting of protein-protein interactions. *Trends Pharmacol. Sci.* **2016**, *37*, 702-713.

20. Ong, S.E.; Mann, M. *Nat. Chem. Biol.* 2005, *1*, 252.

21. Haab, B.B. *Proteomics*, 2003, *3*, 2116.

22. Li, Z.; Shao, S.; Ren, X.; Sun, J.; Guo, Z.; Wang, S.; Song, M. M.; Chang, C. A.; Xue, M. Construction of a sequenceable protein mimetic peptide library with a true 3D diversifiable chemical space. *J Am Chem Soc* **2018**, *140*, 14552–14556.

Chapter 2: Bicyclic peptides as fluorescent reporter probes for ERK activities

2.1 Introduction

Mitogen-Activated Protein Kinase 3 (MAPK3/ERK1) and mitogen-activated protein kinase 1 (MAPK1/ERK2) are serine/threonine kinases that form an integral part of cascading signaling networks responsible for cell cycle, including differentiation, proliferation, cell cycle progression, and survival [1-5]. Aberrant signaling of both these pathways is involved in different types of cancers, which prompted the development of many inhibitors targeting ERK and other pathway members, such as Ras, Raf, and MEK. Drugs targeting ERK pathways have been long studied, yet durable therapeutic responses are often difficult to achieve, partly due to the heterogeneity of tumor cells [6,7,8].

There are sequence homologies for both isoforms of ERK, which has always posed significant challenges in developing highly specific inhibitors to a particular isoform[8,9]. Despite their structural similarities, their biological functions show significant differences. Lowering levels of ERK1 expression alone has shown phenotypical changes that are not observed when knocking down ERK 2. Conversely, silencing ERK2 led to some biological impacts that are not related to ERK1 expression. Hence, probes that specifically target each ERK isoform are necessary[10,11].

In this chapter, we devise a methodology for specifically probing the phosphorylation of ERK isoforms. Target-based high throughput screening of bicyclic peptide library was used to identify sequences that can bind to selected epitopes, which are carefully chosen to suit our purpose. The binding was subsequently validated using fluorescence readout

assays like FRET (Fluorescence Resonance Energy Transfer) and FP (Fluorescence Polarization).

2.2 Experimental

Materials

Fmoc-protected natural amino acids were purchased from Anaspec (Fremont, CA). Fmoc-L-propargylglycine and Fmoc-L-azidolysine were purchased from Combi-blocks (San Diego, CA). Rink amide MBHA resin was purchased from Aapptec (Louisville, KY), and TentaGel S-NH_2 resin was obtained from Rapp Polymere GmbH (Tübingen, German), with loading capacities of 0.678 and 0.28 mmol/g respectively. Fmoc-PEG5-OH (98.68%) was obtained from BroadPharm (San Diego, CA). Fmoc-Glu (OAll)-OH was purchased from Novabiochem (Latvia). Hoveyda-Grubbs Catalyst™ 2nd Generation (97%) was purchased from Sigma-Aldrich. Tetrakis(triphenylphosphine) palladium (0) (99%) was purchased from STREM CHEMICALS (Newburyport, MA). Diisopropylethylamine (DIEA, 99.5%) was purchased from ACROS (Germany) and piperidine (99%) was purchased from Alfa Aesar (Ward Hill, MA). Trifluoro-acetic acid (TFA, 99%) was obtained from Oakwood Chemical (Estill, SC). Phenyl-silane (>97%) and Phenyl Isothiocyanate (>98%) were purchased from TCI (Japan). Tris base, Tween 20, sodium phosphate dibasic anhydrous (Na_2HPO_4, 99.6%), sodium chloride, sodium phosphate monobasic monohydrate (NaH_2PO_4, 99.4%), bovine serum albumin (BSA), acetonitrile (CH_3CN), ethyl acetate, dichloromethane (DCM), N,N′-dimethylformamide (DMF), and ascorbic acid were all purchased from Thermo Fisher Scientific (Waltham,

MA). 2-(1H-benzotriazole-1-yl)- 1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 99.6%) was purchased from Chem-Impex (Wood Dale, IL) and Phenyl isothiocyanate (PhNCS), triisopropylsilane (TIPS) from TCI (Portland, OR). From Sigma-Aldrich (St. Louis, MO), α-cyano-4- hydroxycinnamic acid (CHCA), cuprous iodide (CuI) and Rhodamine B were purchased. 5(6)-carboxy-fluorescein, cyanogen bromide (CNBr) were obtained from ACROS (Pittsburg, PA). U87 glioblastoma cells were obtained from ATCC (the American Type Culture Collection)

Synthesis of the linear version of peptide library

2g of Tentagel S NH_2 resin was used to synthesize the library. The resin beads were swelled overnight in 20 mL of DMF. A mixture of ANP (1.25 g, 2.9 mmol), DIEA (1.2 mL, 7.25 mmol), HBTU (1.06 g, 2.78 mmol) was dissolved in 20 mL DMF and the solution was added to the beads. The whole mixture was shaken at room temperature for 2 hours. The beads were then washed with DMF (5 min × 3), and Fmoc protecting group was removed with 20 mL of 20% piperidine/DMF solution (10 min x 3), followed by washing with DMF (20 mL x 5 min x 3). For the next coupling, a mixture of Fmoc-L-glutamic acid 5-allyl ester, DIEA, HBTU in DMF (in equal molar amounts as described for ANP) was added to the beads and shaken for about 2 hours at room temperature. The beads were then washed again with DMF. The Fmoc protecting group was again cleaved off by incubating the beads with 20% piperidine/DMF solution (20 mL x 5min x 3), followed by washing with DMF, five times.

For the iterative cycles of peptide synthesis, the beads were split into 18 equal parts for the coupling of 18 different L-amino acids Ala, Arg, Asn, Asp, Gln, Glu, Gly, His, Ile, Leu, Lys, Phe, Pro, Ser, Thr, Trp, Tyr, and Val), each dissolved in a mixture of DIEA, HBTU in DMF. The beads with the mixture were then left on a rotator for 2 hours at room temperature. The beads were combined and then washed with DMF for 5 times. The beads were then mixed thoroughly to have a homogenous mixture collected from the eighteen different sets. The split and mix process was repeated for another four more cycles. In the case of coupling one amino acid to all the beads like Fmoc-L- glutamic acid 5-allyl ester, Fmoc-propargylglycine, and Fmoc-azidolysine-OH, the beads were placed into one reaction vessel, and the same coupling method as ANP was applied.

First cyclization reaction using click reaction (CuAAC)

The beads were incubated with 20 ml of 20% lutidine/ DMF solution containing CuI (5 equiv.) and ascorbic acid (15 equiv.) at room temperature overnight. The beads were washed with a solution of sodium diethyldithiocarbamate (5% w/v), DIEA (5% v/v) in DMF, (5 min x 5 times) to get rid of the copper catalyst.

Ring-closing metathesis (RCM)

The beads were then collected and dried. They were placed in a 100 ml round bottom flask with 0.12 mmol of 2nd generation Hoveyda-Grubbs catalyst (HG-II). 3 mL of Anhydrous 1,2-dichloroethane was added to the mixture was stirred gently at 70°C, maintaining an inert atmosphere by flowing argon through the system. 0.12 mmol of

Grubbs catalyst (HG-II) was added to the reaction and left for 16 hours. The beads were then washed with DMF followed by incubation with 0.2 M tris(hydroxymethyl)phosphine in isopropanol for about 12 hours at 80 °C. After 12 hours, the beads were washed with a mixture of DMF/H2O (50:50) and later with DMF, five times each.

FRET Binding Assay

For the FRET binding assay, the hit peptides were tagged with rhodamine and the epitopes with fluorescein, through a linker, i.e. 4-(Dimethylamino) butanoic acid conjugated in between, in both cases. The fluorophore-linker combination was conjugated to the peptides using Solid Phase Peptide Synthesis protocols. A concentration gradient was prepared with several different concentrations of Rhodamine-tagged peptide (in PBS buffer), and they were incubated with 100 nM fluorescein-labeled epitope (in PBS buffer). 60 μL of this mixture was pipetted in a Greiner black 384-well microplate in triplicates for each sample, and was left for incubation for about 30 minutes. Next, the fluorescence intensity was measured using Synergy H1 multi-mode microplate reader and the average FRET signal from triplicates was considered for calculation and analysis. An excitation of 480 nm and an emission wavelength of 600nm was used. The average of the three fluorescence intensity readings were calculated along with standard deviations. In addition, the samples for background correction were made where PBS buffer solution was incubated with the same varying concentration of Rhodamine tagged peptide, with similar calculations of average and standard deviations.

The difference between the average fluorescence intensity readings from the reaction sample and average intensity from the corresponding background sample was calculated for each concentration of Rhodamine tagged peptide. These differences were plotted against the concentration of Rhodamine tagged peptide in Origin 2021 software. The standard deviations of each concentration were used to plot the error bars. The scattered plot was fitted using the Hill1 function available in Origin Pro 2021 software. The adjusted R-square and reduced Chi-square values were used to assess the quality of fitting.

FP Binding Assay

For the FP assay, the concentration of rhodamine tagged hit peptides was kept constant while changing the concentration of the recombinant ERK proteins. The fixed concentration of the rhodamine-tagged peptide was kept at 100 nM in all samples. From the peptide-protein mixture, 65 μL solution was pipetted in a Greiner 384 black microwell plate. Once the mixtures were added, the plate was incubated for about 30 minutes at room temperature. After 30 minutes, fluorescence polarization was measured using Synergy H1 multi-mode microplate reader. For each concentration, all the samples were prepared individually and in triplicates. The average and standard deviations of polarization of the triplicates were taken and plotted against the ERK concentrations. After that, the data were fitted using the Hill1 function in Origin Pro 2021 software. The adjusted R-square and reduced Chi-square values were used to assess the quality of fitting.

High-Performance Liquid Chromatography (HPLC)

Preparative Reversed-Phase (RP) HPLC- A Thermo Ultimate 3000BX HPLC instrument was used. A Phenomenex C18 reversed-phase preparative column (Kinetex 5 μm EVO, 250 × 21.2 mm^2) was employed for preparative HPLC. A multi-wavelength UV−vis detector was used to monitor the absorbance at 215, 280, 480, and 560 nm. Flow gradients of 0−100% acetonitrile (with 0.1% TFA) in water (with 0.1% TFA) were utilized with a flow rate of 15 mL/min.

Analytical HPLC- A Thermo Ultimate 3000SD HPLC instrument, with a Phenomenex C18 reversed-phase analytical column (Kinetex 2.6 μm EVO, 250 × 4.6 mm^2) was used for Analytical HPLC. A gradient of 0−100% acetonitrile (with 0.1% TFA) in water (with 0.1% TFA) was employed with a flow rate of 1.2 mL/min. A UV−vis detector was used to monitor the absorbance at 215, 280 and/or 560 nm. The bicyclic peptides were all purified at purity levels >95%.

Cell Culture

U87 were cultured in a culture media, constituting DMEM (Dulbecco's modified Eagle's medium), 10% heat-inactivated FBS (Fetal Bovine Serum) (v/v) and 100 U/mL penicillin/streptomycin, under 5% CO_2 in a 37°C incubator. Cells were passaged every time they reached a confluency of 80-90%.

In vitro activity assay

A black GRENIER 384 well plate was incubated with 1% BSA in PBST for about an hour and washed with PBS (3 times). Afterwards, inactive ERK1/2 with a final

concentration of 200 nM and active MEK with concentrations 100, 200, and 400 nM were added to the plate. 10 μM ATP was added to the solution, and the plate was left for 45 minutes at room temperature. The FRET pair, Cy5-B22, and Rhod-C14, was added to the solution, and the fluorescence intensity was measured using the plate reader. All the solutions were made in HEPES buffer. The intensities were measured in triplicates, and the average values were used in the bar chart plots. Standard deviations from the three readings were used as errors, and the bar graphs were plotted using Origin Pro 2021 software.

2.3 Results & discussion

Epitope targeted bicyclic peptide library screening

The two isoforms of ERK, ERK 1, and ERK 2, are homologous, and it is extremely difficult to design a method to target one specific isoform. For the discovery of probes for these proteins, we selected our first epitope, Epitope 1, shared by both ERK1, G199-R211, and ERK2, G182-R194 (G-F-L-T-E-Y-V-A-T-R-W-Y-R) and another one, Epitope 2, on ERK1, Y327-E341, (Y-L-E-Q-Y-Y-D-P-T-D-E-P-V-A-E) with 87% sequence homology with ERK2.

A bicyclic peptide library promises a lot of complexity in its structure to bind to targets. In this case, the targets are structureless epitopes. The backbone rigidity and three-dimensional distribution of the functional groups of the amino acids are beneficial in achieving binding activity. Our group's previous successes with the extremely

structureless c-Myc epitope only inspired us to target our ERK epitopes in a similar endeavor[12].

The library was synthesized using L-amino acids, with each member of the library having five randomized modular amino acids added through the five iterative cycles of split and pool library synthesis described previously. Six other amino acids were common in all sequences. These were added to the entire library without splitting them. Azidolysine, a proline, two L-glutamic acid 5-allyl esters, another modular amino acid, and then an alkyne handle. Two cyclization processes, copper-catalyzed azide−alkyne cycloaddition (CuAAC) and Ru-catalyzed ring-closing metathesis (RCM) reactions were utilized to generate the bicyclic structures. The unique topology of the library components gave us access to a three-dimensional chemical space which is advantageous for generating hits for these challenging epitopes.

Validation using fluorescence readout assays-FRET

Epitope 1 and Epitope 2, tagged with fluorescein, a FRET donor, were used to estimate binding against the bicyclic peptide, which was tagged with rhodamine, a FRET acceptor. An effective binding was expected to bring the two counterparts together within 10 nm, exhibiting a suddenly increased fluorescence signal at the effective concentration. Out of the 17 hits, B22 proved to be a good binder against Epitope 1, while A16, C14, and D23 bound well against Epitope 2. Fluorescence Polarization assays validated rhodamine tagged B22's efficient binding against both ERK 1 and ERK 2, while rhodamine tagged C14 bound against only ERK1.

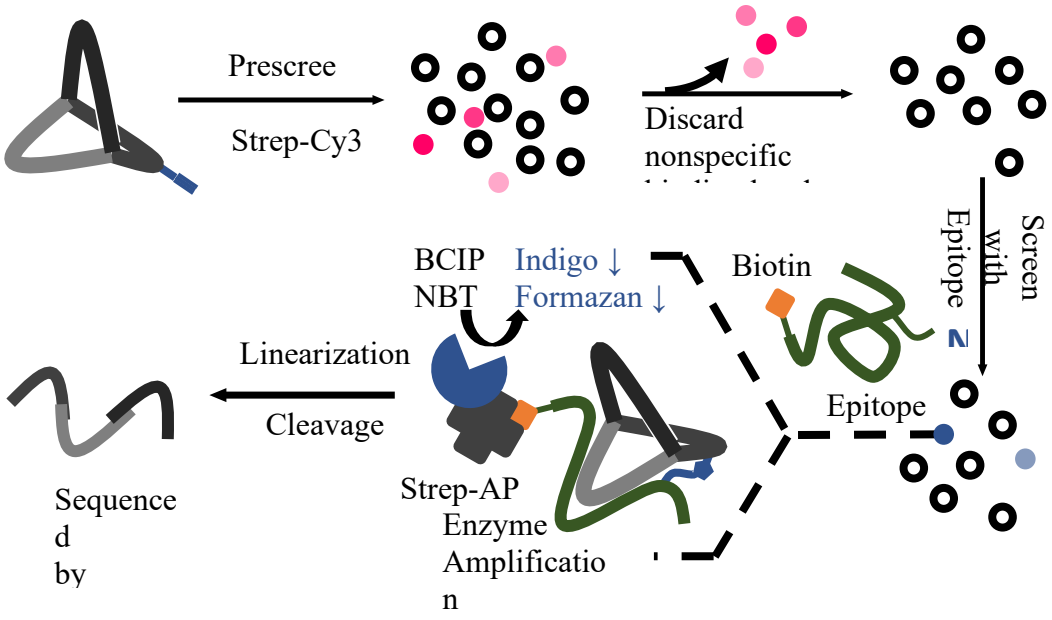
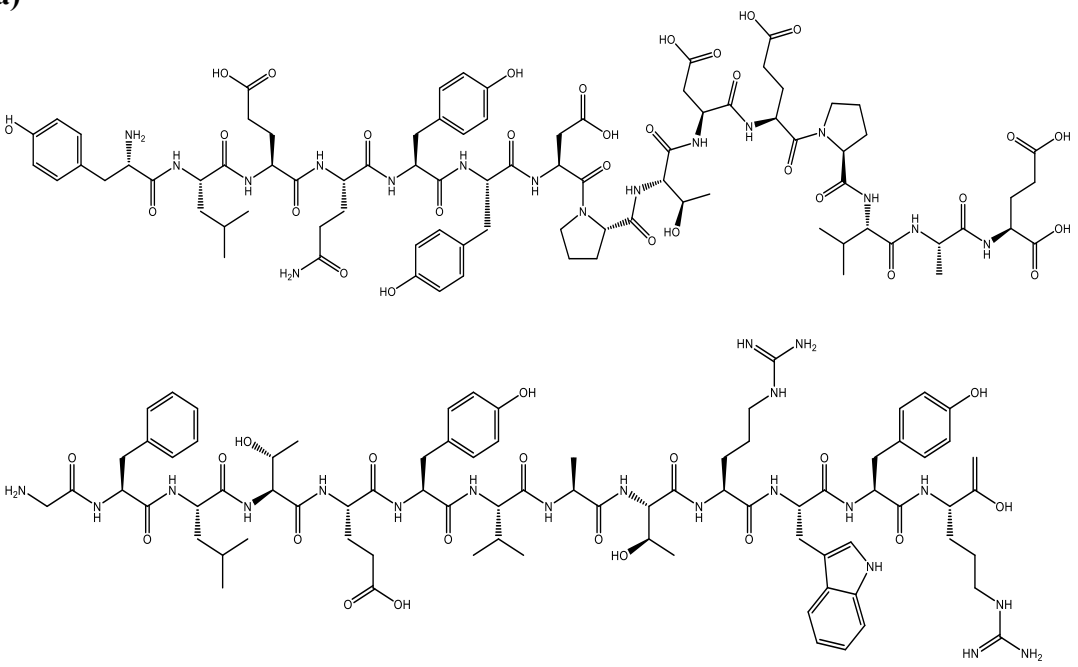


Figure 2.1 Schematic representation of screening bicyclic peptide library. An alkyne handle is conjugated to the library and is screened against dye-tagged proteins. The unreacted ones are reacted with an azide-containing epitope with a biotin tag. The hit beads undergo enzyme amplification through a biotin-streptavidin-specific reaction, and a corresponding Alkaline phosphatase reacts with BCIP-NBT to form purple-colored beads due to precipitation. The hit beads are then linearized

a)



b)

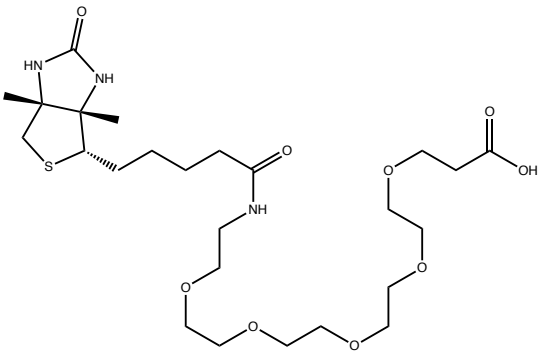
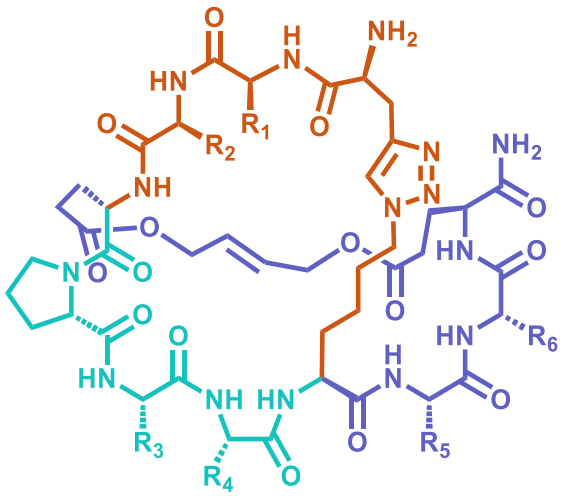


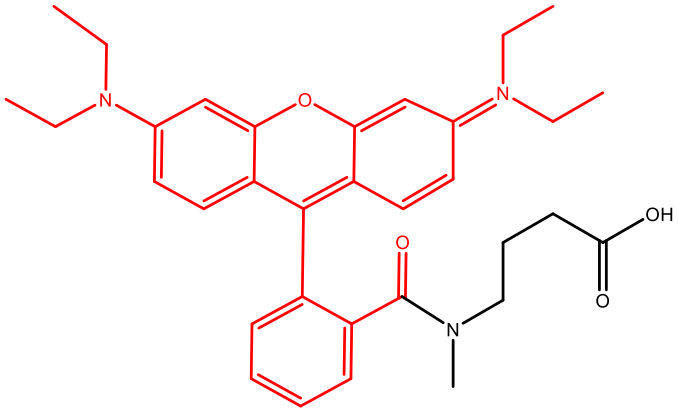
Figure 2.2 a) Epitope 1- G-F-L-T-E-Y-V-A-T-R-W-Y-R and Epitope 2- Y-L-E-Q-Y-Y-DP-T-D-P-T-D-E-P-V-A-E. Epitope 1 has 100% sequence match with both ERK 1 and ERK 2, while Epitope 2 has 100% match with ERK 1 and 87% with ERK 2. **b)** Biotin-PEG5 linker attached to both these epitopes at their N-terminals to aid the screening process.



	R_1	R_2	R_3	R_4	R_5
A3	R	A	I/L	G	E
C1	S	E	E	D	E
C3	E	D	D	A	S
D13	G	R	L	R	R
D15	T	T	D	A	D
A16	R	D	D	D	S
B11	P	S	G	D	N
C22	S	H	S	N	T
F13	L	T	E	Q	Q
B17	N	P	T	T	I/L
C19	W	D	D	N	V
B14	D	A	D	D	R
D23	N	L	P	T	D
B22	P	Q	D	D	S
E6	P	D	A	I/L	E
C14	D	D	I/L	I/L	D
C2	P	D	G	I/L	Y

Figure 2.3 Bicyclic library structure, used for screening. The table shows the hits obtained from the library with the respective amino acids at the designated positions.

a)



b)

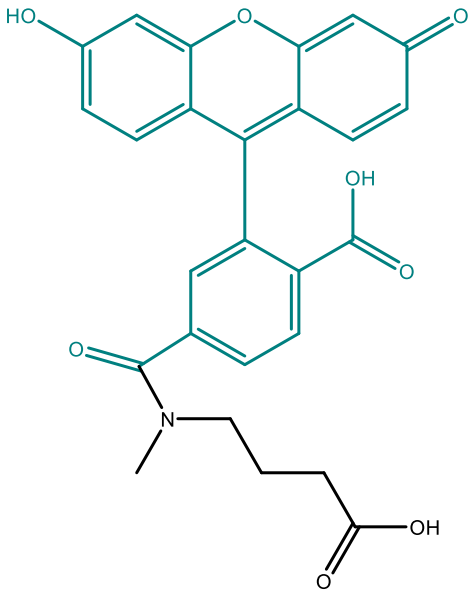


Figure 2.4 a) Rhodamine (red)-Linker and **b)** Fluorescein (blue)-Linker combinations are conjugated to hit peptides and epitopes respectively, at their N-termini as tags for FRET assays.

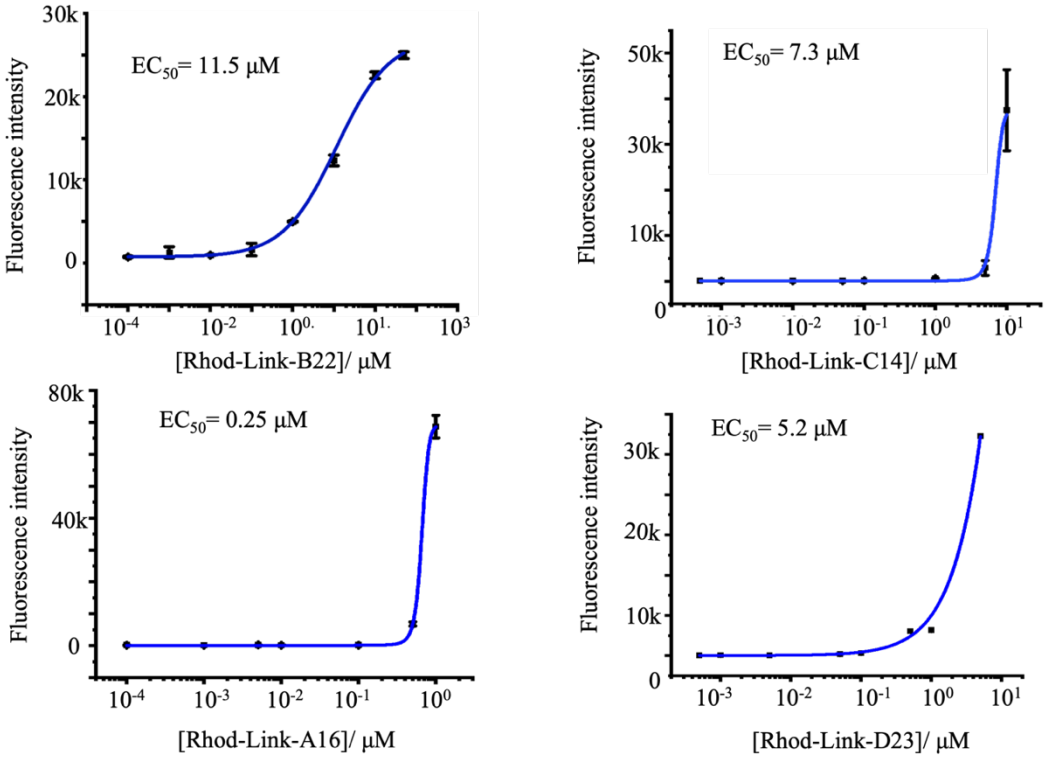


Figure 2.5 FRET binding curves for rhodamine-tagged B22 against fluorescein-tagged Epitope 1 and rhodamine-tagged C14, A16, and D23 against fluorescein-tagged Epitope 2, these are the better binders among the hits obtained from screening.

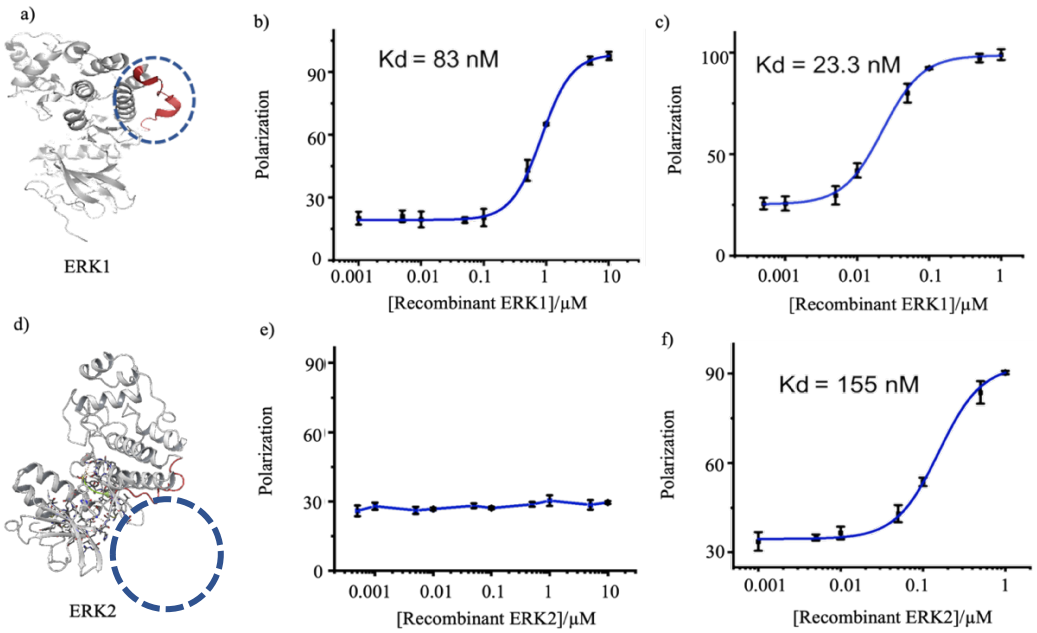


Figure 2.6 Validation. **a)** Crystal structure of ERK1 from PDB. **b)** FP assay showing the binding between ERK1 and Rhodamine-tagged C14. **c)** FP assay showing the binding between ERK1 and Rhodamine-tagged B22. **d)** Crystal structure of ERK1 from PDB. **e)** FP assay with showing absolutely no binding between ERK1 and Rhodamine-tagged C14. **f)** FP assay with showing absolutely no binding between ERK1 and Rhodamine-tagged B22.

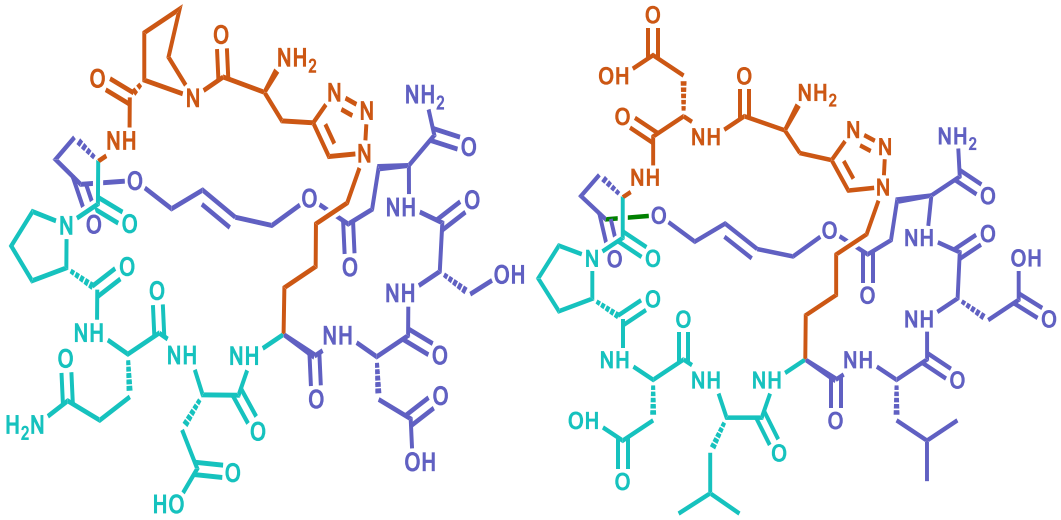


Figure 2.7 B22 (left) and C14 (right) are the best binders, proven through FRET binding data against the epitopes and the FP binding data against the proteins.

FRET-based reporting process to probe lysates

The newly discovered reporter pair's activity was tested by incubating them with cell lysates collected from different conditions. B22 was conjugated to Cy5 and C14 to Rhodamine. This pair showed a lowered FRET intensity compared to the control, while lysates from trametinib-treated cells showed increased FRET signal. These results were consistent with our expectations where the FRET pair reported the amount of the unphosphorylated (inactivated) ERK1 (Figure 2.8, 2.9a, 2.10b).

***In vitro* kinase assay**

An in-vitro assay was performed utilizing the full-length inactive recombinant ERK1 protein, the immediate upstream protein (MEK), and ATP. MEK phosphorylates both Tyr and Thr sites on ERK with the help of ATP. We treated the inactive ERK1 protein with active MEK and ATP. The MEK concentration was varied, while ATP and inactive ERK1 concentrations were kept constant. This setup was expected to generate more copies of active ERK1. Since our probe recognized the unphosphorylated ERK1, we expected to observe reduced FRET signals, which was validated by our experiments (Figure 2.9 c).

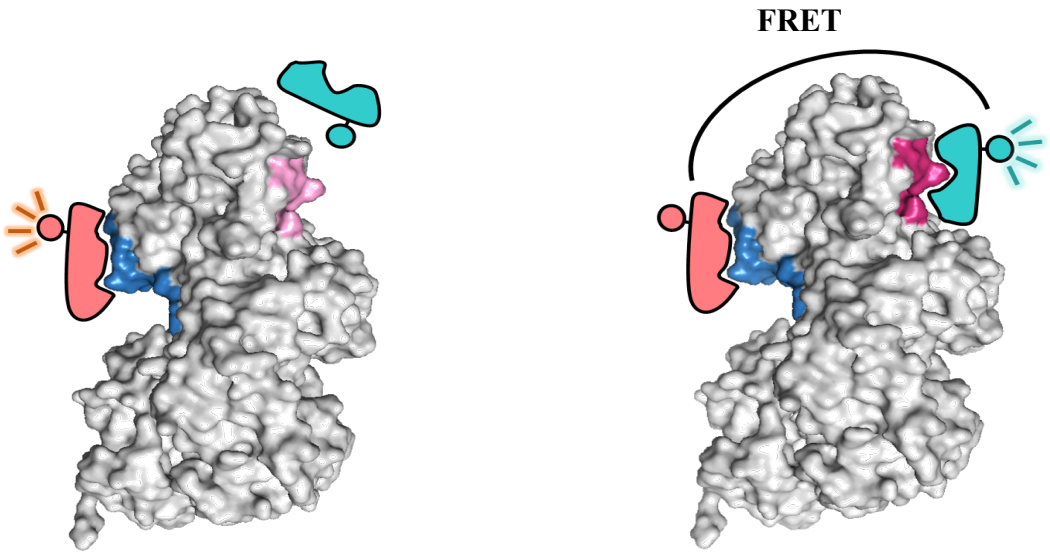
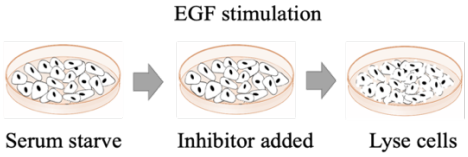
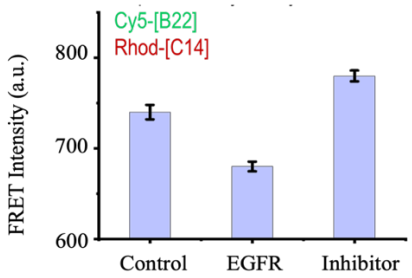


Figure 2.8 We have a pair of probes in Cy5 tagged B22 and Rhodamine tagged C14, phosphorylated ERK1, can have only probe bound to it, resulting in absence of FRET. The unphosphorylated/ inactive ERK1 will have both probes bound to it, resulting in FRET phenomenon and an increased signal.

a)



b)



c)

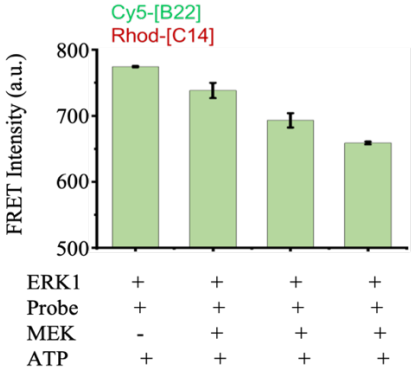


Figure 2.9 a) Serum starved U87 cells were stimulated with EGF, followed by treatment MEK inhibitor trametinib. **b)** Changes in FRET signal in the lysates that were collected right after treatment with EGF and 12 0 inhibitor. **c)** An increase in active MEK, leads to a decrease in the FRET signal.

2.4 Conclusion

A pair of specific fluorescent probes for ERK1 isoform were discovered through epitope-targeted library screening processes. These ERK sensors can specifically translate the level of unphosphorylated ERK1 protein to FRET signal, which can be used to study the isoform-specific roles of ERK signaling.

2.5 References

1. Widmann, C; Gibson, S.; Jarpe, M.B.; Johnson, G.L. Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. Physiol Rev **1999**,*79*,143–180.

2. Lewis, T.S.; Shapiro, P.S.;Ahn, N.G. Signal transduction through MAP kinase cascades. *Adv. Cancer Res.* **1998**, *74*, 49–139.

3. Pearson G.; Robinson F.; Beers Gibson T.; Xu B.E.; Karandikar, M; Berman, K. *et al*. Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. *Endocr. Rev.* **2001**, *22*, 153–183.

4. Meloche, S.; J. Pouyssegur. The ERK1/2 mitogen-activated protein kinase pathway as a master regulator of the G1- to S-phase transition. *Oncogene* **2007**, *26*, 3227-3239.

5. Lefloch, R.; Pouyssegur, J.; Lenormand, P. Total ERK1/2 activity regulates cell proliferation. *Cell Cycle* **2009***, 8,* 705–711.

6. Flaherty, K.T.; Infante, J.R.; Daud, A.; Gonzalez, R.; Kefford, R.F.; Sosman, J.; *et al.* Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations . *N Engl J Med* **2012**, *367*, 1694 – 703.

7. Flaherty, K.T.; Hodi, F.S.; Fisher, D.E. From genes to drugs: targeted strategies for melanoma. *Nat Rev Cancer* **2012**,*12*, 349 – 61.

8. Crews, CM.; Alessandrini, A.; Erikson, RL. The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product. *Science* **1992,** 258, 478-80

9. Welsh, S. J.; Corrie P. G. Management of BRAF and MEK inhibitor toxicities in patients with metastatic melanoma. *Therapeut. Adv. Med. Oncol.* 7, **2015,** 122–136.

10. Fremin, C.; Ezan, F.; Boisselier, P.; Bessard, A.; Pages, G.; Pouyssegur, J.; Baffet, G. ERK2 but not ERK1 plays a key role in hepatocyte replication: An RNAi-mediated ERK2 knockdown approach in wild-type and ERK1 null hepatocytes. *Hepatology* **2007**, *45*, 1035–1045.

11. Maillet, M.; Purcell, N. H.; Sargent, M. A.; York, A. J.; Bueno, O. F.; Molkentin, J. D. DUSP6 (MKP3) null mice show enhanced ERK1/2 phosphorylation at baseline and increased myocyte proliferation in the heart affecting disease susceptibility. *J Biol Chem* **2007**, 7, 31246–31255.

12. Li, Z.; Shao, S.; Ren, X.; Sun, J.; Guo, Z.; Wang, S.; Song, M. M.; Chang, C. A.; Xue, M. Construction of a sequenceable protein mimetic peptide library with a true 3D diversifiable chemical space. *J Am Chem Soc* **2018**, *140*, 14552–14556.

Chapter 3: Optimization of bicyclic peptide-based inhibitor for c-Myc and investigating its core biological impacts

3.1 Introduction-

The MYC family proteins are master regulators of gene transcription, and their activities are abnormally high in more than 75% of known human cancers[1,2,3]. Increased MYC activities can promote metabolic reprogramming, accelerate protein synthesis, and inhibit apoptosis, all of which directly facilitate oncogenesis and tumor progression[4,5]. Moreover, increasing evidence has shown that upregulated MYC also contributes to drug-resistance development as well as tumor metastasis. Because of its prominent role in cancer, is the third most common putatively actionable alteration across all cancer types. Numerous studies have established that inhibiting MYC activities can halt tumor progression in many cancer types. Consequently, MYC is one of the most sought-after drug targets in cancer therapeutics for decades. Three decades of drug development efforts have led to 30 reported MYC inhibition strategies, but a clinically viable one remains elusive [6,7,8,9]. Particularly, very few of them can achieve direct MYC inhibition- a preferred strategy that promises higher therapeutic efficiency and specificity. The goal is difficult to meet for three reasons. First, MYC is an intrinsically disorderd protein that lacks apparent binding pockets[1]. Therefore, conventional small molecule-based approaches cannot generate hits easily. Second, MYC has a very short lifetime (20-30 min), which makes covalent modification strategies useless, as maintaining an effective intracellular drug concentration is difficult. Third, MYC acts in the nucleus, which calls

for drugs which can pass through the nuclear envelope, thus excluding antibody-based therapeutics[9,10,11]. To date, there still exists a pressing need to develop MYC inhibition. Inspired by recent developments in cyclic peptide-based protein-protein interaction inhibitors, we envisioned that we could address the unmet need using rigid cyclic peptides, i.e., binding to MYC to interfere with its transcription activities. We further reasoned that multicyclic peptides could compensate for the entropic cost required for binding to the intrinsically disordered MYC. In our previous work, we developed a unique bicyclic peptide library that presented modular functional groups in a true 3-D manner, which provided access to a 3D -diversifiable chemical space[13]. Using this library, we identified a bicyclic peptide (NT-A1) that could bind to MYC. We then demonstrated that this peptide was able to disrupt MYC function and elicit cell-death at a micromolar concentration in a glioblastoma cell line. Despite the ability to inhibit MYC in glioblastoma cells, NT-A1's binding to MYC was not strong enough. We observed a low- μM binding affinity, which resulted in *in vitro* IC50 values around 1 μM. To develop an effective MYC inhibitor, we shall improve the binding affinity and achieve low nM-level Kd. In this chapter, we sought to study the structure-activity relationship of NT-A1, and perform medicinal chemistry modifications to improve its binding affinity towards MYC. In addition, we study how cells respond to MYC inhibition and investigate potential routes of resistance development.

3.2 Experimental

Materials

Rink amide MBHA resin with a loading capacity of 0.678 mmol/g was purchased from Aapptec (Louisville, KY), and all Fmoc-protected natural amino acids were purchased from Anaspec (Fremont, CA). Fmoc-L-propargylglycine and Fmoc-L-azidolysine were purchased from Combi-blocks (San Diego, CA). TentaGel S-NH_2 resin with a loading capacity of 02.8 mmol/g was obtained from Rapp Polymere GmbH (Tübingen, German). Tris base, Tween 20, sodium phosphate dibasic anhydrous (Na_2HPO_4, 99.6%), sodium chloride, sodium phosphate monobasic monohydrate (NaH_2PO_4, 99.4%), sodium dodecyl sulfate (SDS), bovine serum albumin (BSA), acetonitrile (CH3CN),), ethyl acetate (EA), dichloromethane (DCM), N,N′-dimethylformamide (DMF), and ascorbic acid was purchased from Thermo Fisher Scientific (Waltham, MA). 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 99.6%) was obtained from Chem-Impex (Wood Dale, IL) and Phenyl isothiocyanate (PhNCS), triisopropylsilane (TIPS) from TCI (Portland, OR). From Sigma-Aldrich (St. Louis, MO), α-cyano-4-hydroxycinnamic acid (CHCA) cuprous iodide (CuI), Rhodamine B, and cuprous iodide (CuI) were purchased. Piperidine was obtained from Alfa Aesar (Ward Hill, MA) and 5(6)-carboxyfluorescein, cyanogen bromide (CNBr) from ACROS (Pittsburg, PA).

Solid-phase peptide synthesis (SPPS) of the linear precursor of bicyclic peptide

The linearized precursor peptides were synthesized on Rink Amide MBHA resin mediated by the CSBio CS336S peptide synthesizer (Menlo Park, CA). An automated

version of the traditional SPPS coupling process was used. The Fmoc group on the resin was first removed by 20% piperidine/DMF solution. Fmoc protected L-amino acids were weighed out and placed in the amino acid holder of the peptide synthesizer. DIEA (0.8 M in DMF) and HATU (0.4 M in DMF). The process of De-Fmoc and amino acid were alternately followed with beads washed by DMF after each respective step of peptide synthesis.

Click reaction

To construct the first cyclization, Fmoc-propargylglycine-OH (Pra) and Fmoc-azidolysine-OH (Az4) were previously added at the N and C terminals, respectively. A Cu-catalyzed click reaction was used for cyclization. The beads were incubated with a solution of CuI (5 equivalents, 0.1291g) and L-ascorbic acid (15 equivalents, 0.357g) in 10 mL 20% Lutidine/ DMF (v/v) overnight at room temperature overnight. After cyclization, the beads were washed with sodium diethyldithiocarbamate (5% w/v) and DIEA (5% v/v) in DMF.

Ring Closing Metathesis reaction

The beads were collected after the click wash, and they were washed thoroughly with DMF (5 times), followed by DCM (5 times). The beads were dried under vacuum and transferred to a 100 mL round-bottom flask equipped with a stirring bar. Hoveyda-Grubbs 2nd generation catalyst was added (0.025 mmol) along with 3 mL of Dichloroethane. The round-bottom flask was sealed, and the mixture was stirred

overnight at 70° C in an inert atmosphere with a constant flow of argon. The next day the catalyst was added (0.025 mmol) to the mixture and the reaction was further carried out for another 12 h. Afterwards, the beads were collected was washed with DMF (5 times), followed by washing with isopropanol (5 times). The beads were then incubated overnight with tris(hydroxymethyl)phosphine (0.2 M in isopropanol, 20 mL) at 80 °C. Then the resin was washed with (50% DMF in H_2O 20 mL×3), DMF (20 mL×3). The beads were cleaved off their Fmoc protection using 20% Piperidine/DMF.

Cleavage and purification

After deprotection of the Fmoc, cleavage was performed using a mixture of trifluoroacetic acid, tri-isopropyl silane and water (95:2.5:2.5 v/v, 20 mL) for 2 hours. The beads were then washed with DMF (20 mL×3) and were ready to use.

Preparative Reversed-Phase (RP) HPLC and Analytical HPLC- A Thermo Ultimate 3000BX HPLC instrument was used for RP-HPLC with a preparative column (Kinetex 5 µm EVO, 250 × 21.2 mm^2) and a multi-wavelength UV−vis detector monitoring the absorbance at 215, 280, 480, and 560 nm. A flow gradient of 0−100% acetonitrile (with 0.1% TFA) in water (with 0.1% TFA) were utilized with a flow rate of 15 mL/min. Preparative HPLC was used to synthesize, all the free and tagged versions of the peptides. They were further purified using Analytical HPLC.

A Thermo Ultimate 3000SD HPLC instrument, with a Phenomenex C18 reversed-phase analytical column (Kinetex 2.6 µm EVO, 250 × 4.6 mm^2) was used for Analytical HPLC.

A gradient of 0−100% acetonitrile (with 0.1% TFA) in water (with 0.1% TFA) was employed with a flow rate of 1.2 mL/min. An UV−vis detector was used to monitor the absorbance at 215, 280 or 560 nm. The bicyclic peptides were all purified at purity levels >95%.

Competitive FRET assay

For the competitive FRET assay, the concentration of rhodamine-conjugated NT-A1 and fluorescein-labeled epitope were kept constant while changing the concentration of competitor ligand. From the mixture, 65 µL solution was pipetted in a black Greiner 384 microwell plate and left for incubation at room temperature for about 30 minutes. After 30 minutes of incubation, the FRET signal was measured using Synergy H1 multimode microplate reader. Samples were set up in triplicates and prepared individually for each concentration. The fluorescence intensity was measured with an excitation wavelength of 480 nm and an emission wavelength of 600 nm. The average of these triplicates was calculated and analyzed. The standard deviations were used to plot the error bars. The scatter plot of fluorescence intensity vs. concentration of competitor ligand was fitted in Origin Pro 2021 using Hill 1 function. The adjusted R-square and reduced Chi-square values were used to assess the quality of fitting.

ELISA

A 96-well plate (Nunc Immobilizer Nickel-Chelate, Thermo Fisher) was incubated with c-Myc protein, 10 µg/mL, 100 µL/well, 9 pM 6His (His-His-His-His-His-His) (100

μL/well) and varying concentration of Biotin-PEG-NT-A7,10, 14 in 0.1% BSA-PBST, Streptavidin-HRP in 0.1% BSA-PBST (1:200) (v/v) successively, in three replicates. Each incubation period was an hour each, and after each period, the solution was dumped, followed by washing 3 times with PBST. Chemiluminescent substrate (SuperSignal West Pico Plus Chemiluminescent Substrate, Thermo Scientific) was added to the plate (100 μL/well). The luminescence density of each well was measured. The average luminescence was plotted against concentrations of Biotin-Peg5 tagged ligands and fitted using Hill1 function in Origin Pro 2021. The adjusted R-square and reduced Chi-square values were used to assess the quality of fitting.

Cell culture

U87 cells were cultured in DMEM media containing 10% FBS and 1% PS, at 37°C with 5% CO_2

Liposome preparation

200 μL each of DOTAP and DOPE solution (10 mg/mL of each in chloroform) were mixed and dried under vacuum to generate a film. 1 mL of 20 mM HEPES (for empty liposome control), or a 100 μM peptide solution (in 20 mM HEPES) was added and sonicated for a few minutes to create a final concentration of the liposomes of 4 mg/mL. The liposomes were extruded at 60 °C using a polycarbonate membrane with a pore size of 0.2 μm on a mini extruder (Avanti Polar Lipids, Inc.) and dialyzed against a 20 mM HEPES solution overnight using a Tube-O-Dialyzer (G-Biosciences).

Liposome loading capacity determination

200 μL of the prepared liposome was dried under vacuum, diluted with 50 μL of DMSO and it was analyzed on a reverse-phase analytical HPLC. The loading capacities were calculated was then determined by comparing the peak area of the peptide in liposome solution to that of a standard solution in HEPES buffer**.**

Cell proliferation test

U87 cells were seeded in a 96-well plate at 4 k/well. Afterwards, the cells were treated with varying concentrations of peptide in four replicates encapsulated in liposomes in a solution of culture media to make 200 μL in each well. The cells were treated for 48 hours, with old media dumped and fresh media with peptide dissolved in it, were added every 24 hours. At the end of 48 hours, resazurin 20 μL mixed with 100 μL cell media was added, the plate was incubated at 37 °C for 3 hours. The 96-well plate was shaken, and fluorescence intensity with an excitation wavelength of 530 nm and emission wavelength 590 nm was read by the plate reader. The average and standard deviations were estimated from the replicates. All values of fluorescence readouts were normalized with respect to the fluorescence readouts from the control sample to obtain the normalized cell proliferation. The calculated normalized cell proliferation was used in bar graph analysis of cell proliferation vs concentration of peptide added in Origin Pro 2021. The cell proliferation assays with other cell lines were carried out in an exact identical manner.

Synergy Experiments

For synergy experiments, a combination of two drugs were added in the same manner as mentioned above. The cell proliferation fluorescence readouts were used to calculate the synergy score of the two drugs, using the following equation:

$S_{A,B} = I_{A,B} - (I_A + I_B - I_A \times I_B)$

Where $S_{A,B}$ is the synergy effect between drugs A and B, $I_{A,B}$ is the cell killing efficiency by using the combination of drug A and B while I_A and I_B are the cell killing efficiencies from independent doses of drug A or B, respectively.

RNA seq experiments

400k U87 cells were seeded in two 35 mm petri dishes each. 100nM of NT-A14 (entrapped in liposomes) was treated for 24 hours and 48 hours respectively in the two dishes. The set up was repeated with NT-A10, NT-A7 and empty liposomes (control). The cells were then lysed, and the RNA was extracted following using RNeasy Kit (QIAGEN) protocols. Poly-A selective RNA-seq libraries were prepared using the NEB NextUltra II kit (New England Biolabs) and sequenced on an Illumina® NextSeq 500 with 2 X 75 pair-end reads. The obtained sequencing data was preprocessed using fastp, aligned using STAR, and counted using featureCounts. The normalized expression counts lists were used as the input for principal component analysis and DESeq2. Gene set enrichment analysis was performed using the GSEA tool and the Hallmark gene sets.

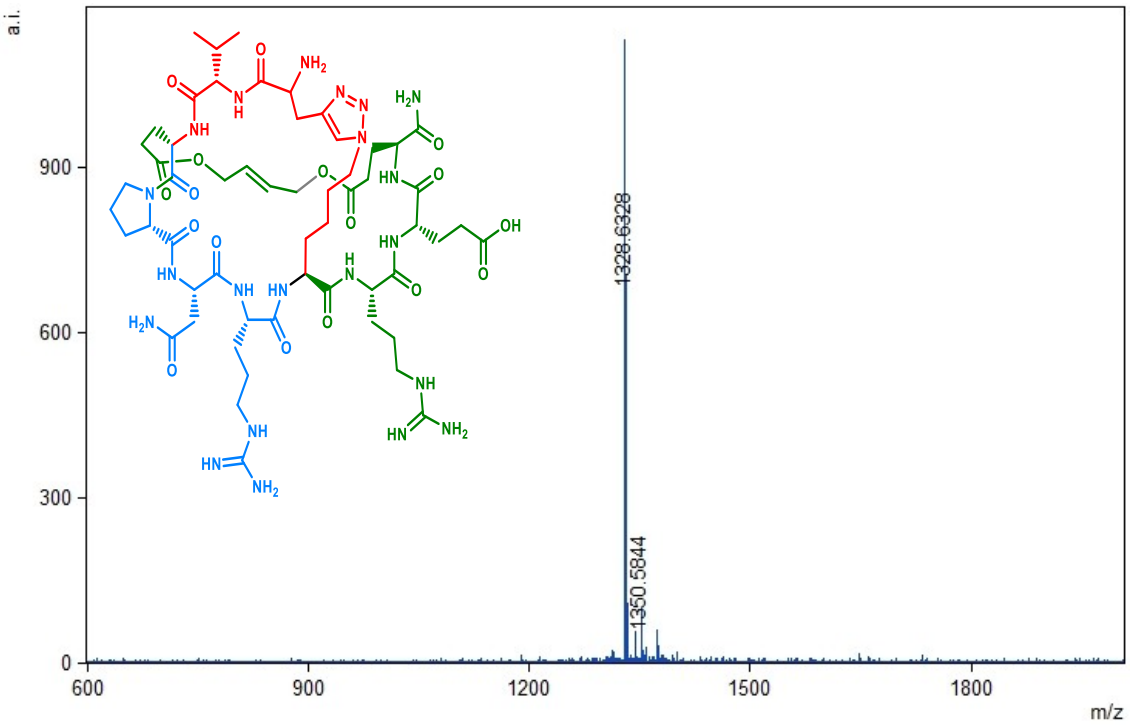


Figure 3.1 MALDI-TOF mass spectrum for NT-A1. Expected mass from calculation 1328.66. Mass found 1328.6.

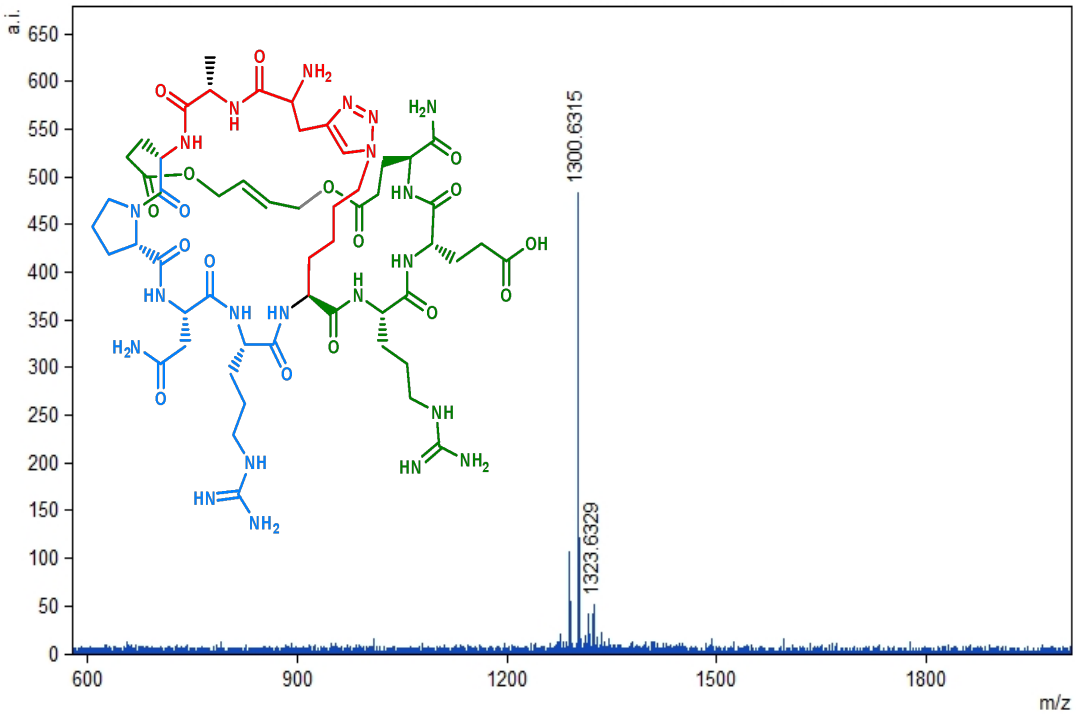


Figure 3.2 MALDI-TOF mass spectrum for NT-A2. Expected mass from calculation-1300.63. Mass found 1300.6315.

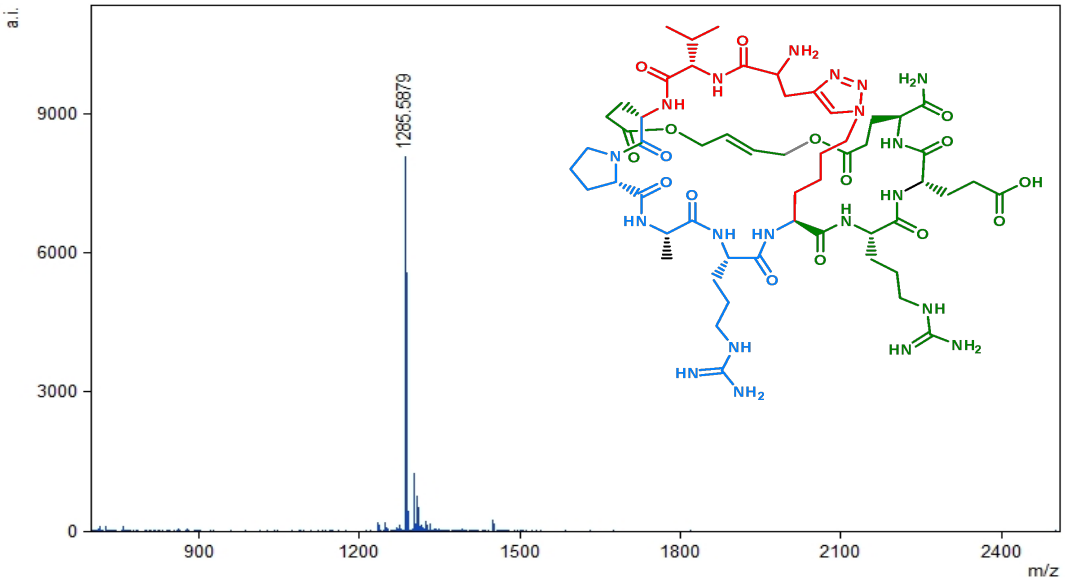


Figure 3.3 MALDI-TOF mass spectrum for NT-A3. Expected mass from calculation 1284.67. Mass found 1285.5879.

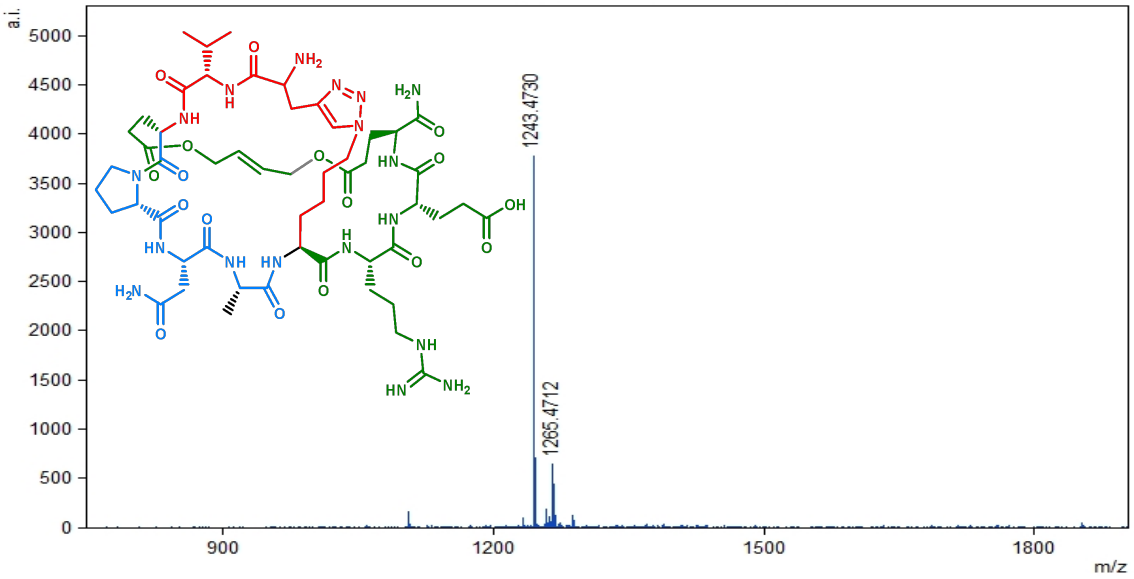


Figure 3.4 MALDI-TOF mass spectrum for NT-A4. Expected mass from calculation 1243.59. Mass found 1243.4370.

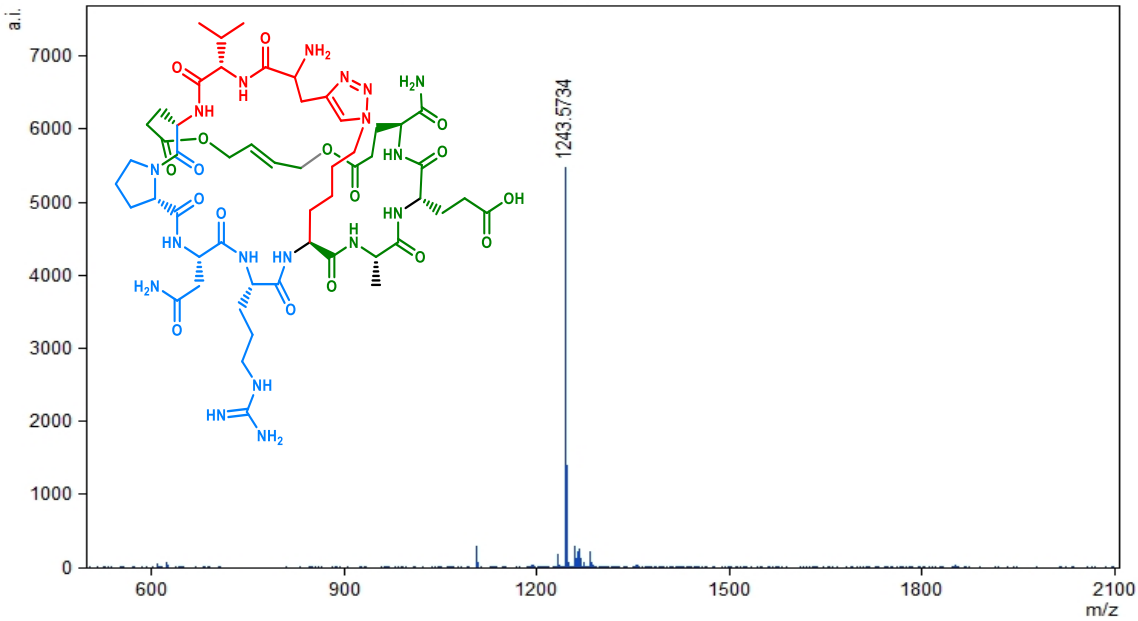


Figure 3.5 MALDI-TOF mass spectrum for NT-A5. Expected mass from calculation 1243.59. Mass found 1243.5734.

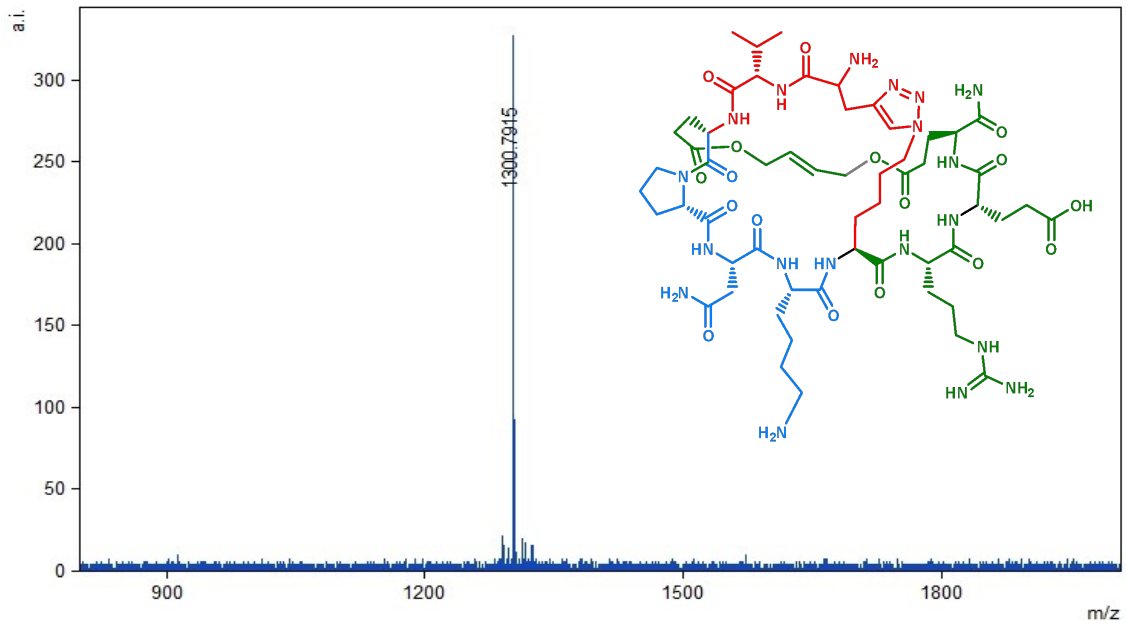


Figure 3.6 MALDI-TOF mass spectrum for NT-A7. Expected mass from calculation 1300.65. Mass found 1300.7915.

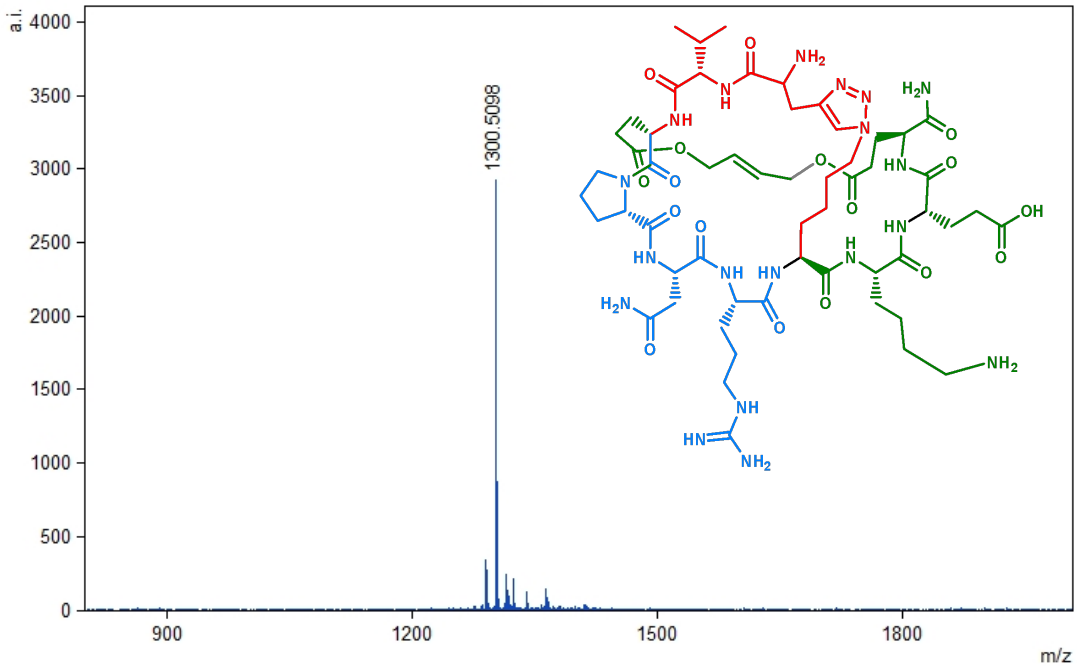


Figure 3.7 MALDI-TOF mass spectrum for NT-A8. Expected mass from calculation 1300.65. Mass found 1300.5098.

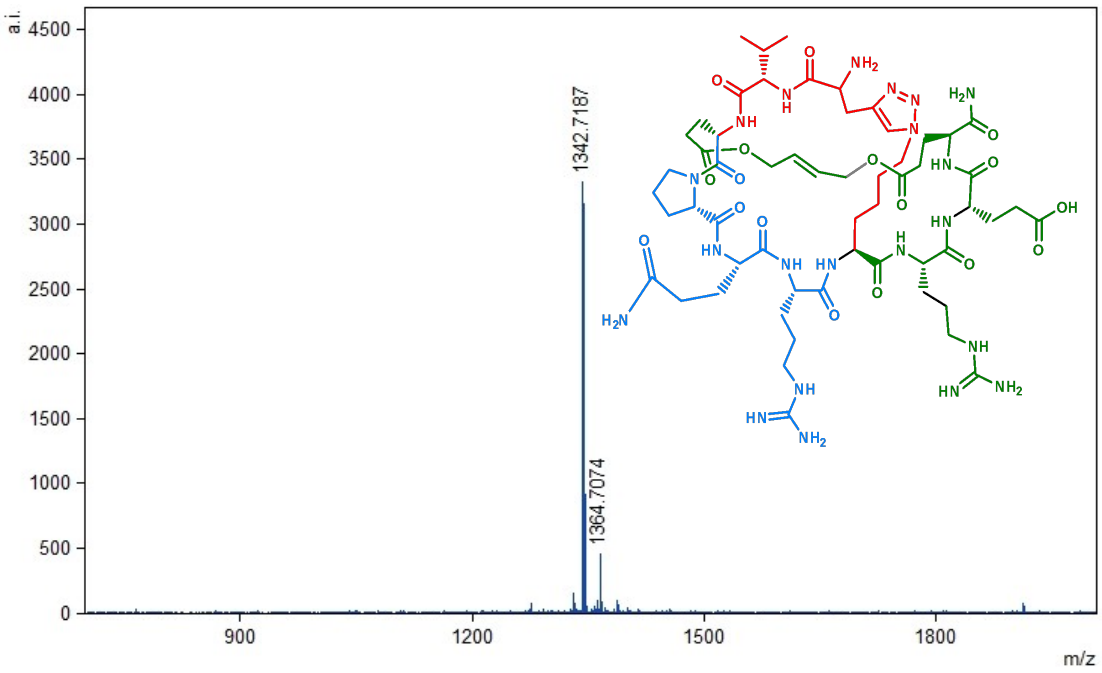


Figure 3.8 MALDI-TOF mass spectrum for NT-A9. Expected mass from calculation 1342.67. Mass found 1342.7187.

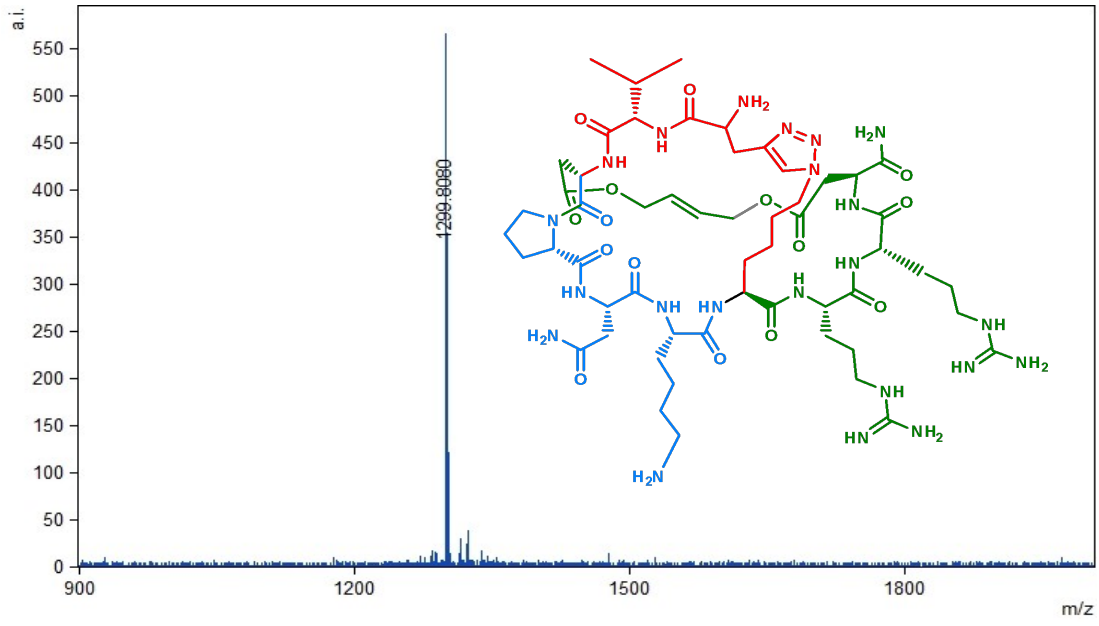


Figure 3.9 MALDI-TOF mass spectrum for NT-A14. Expected mass from calculation 1299.68. Mass found 1299.8080.

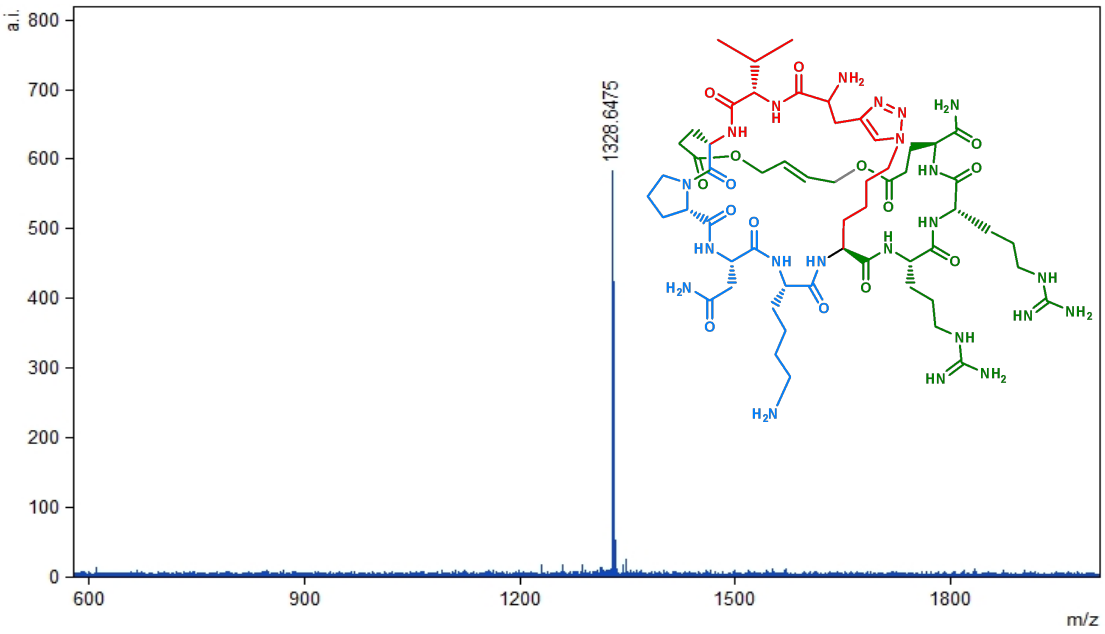
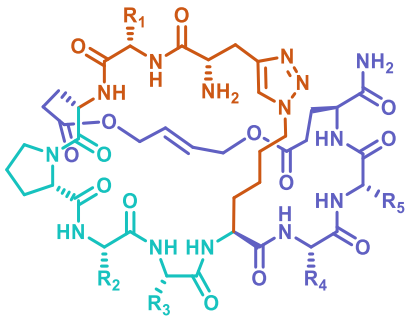


Figure 3.10 MALDI-TOF mass spectrum for NT-A10. Expected mass from calculation 1327.71. Mass found 1328.6475

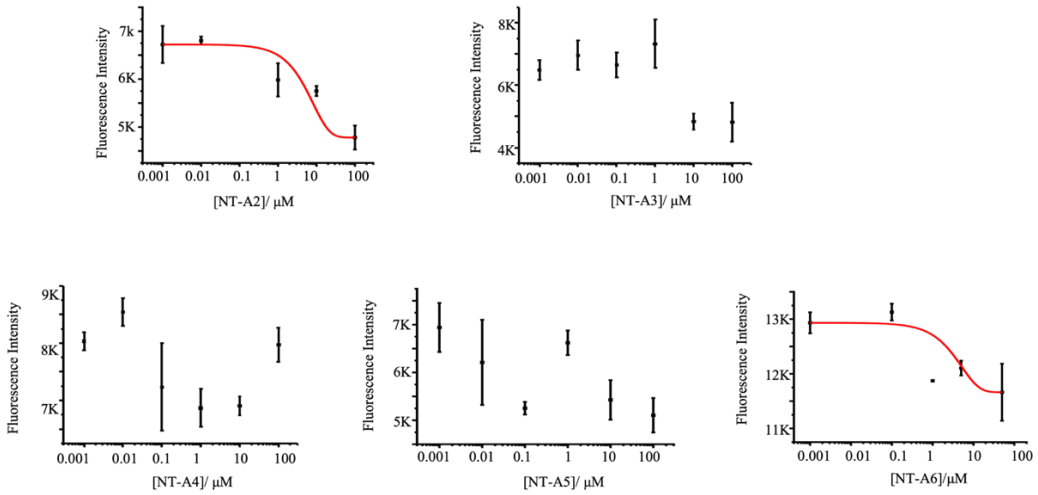
a)



b)

NT-A #	R_1	R_2	R_3	R_4	R_5
1	**V**	**N**	**R**	**R**	**E**
2	A	N	R	R	E
3	V	A	R	R	E
4	V	N	A	R	E
5	V	N	R	A	E
6	V	N	R	R	A

c)



d)

FRET Donor	FRET Acceptor	Competitor
Fluorescein-Linker-Epitope	Rhodamine-Linker-NT-A1	**NT-A2**
		NT-A3
		NT-A4
		NT-A5
		NT-A6

Figure 3.11 a) Bicyclic peptide library structure **b)** Alanine scanning variants of NT-A1 **c)** Competitive FRET studies efficiencies **d)** Table summarizing better binders (green) and worse binders (red) than NT-A1.

3.3 Results and Discussion

As a first attempt to elucidate the structure-activity relationship, an Alanine scanning campaign was performed. Here, we replace each amino acid with alanine (Figure 3.11 a,b). and assessed the binding affinity towards MYC.Figure 3.11 a) Bicyclic peptide library structure b) Chart showing sequences of Alanine scanning variants of NT-A1 c) Competitive FRET studies of each of the variants, estimating binding efficiencies d) Table summarizing better binders (green) and worse binders (red) than NT-A1.We chose to perform competitive FRET assays to evaluate the binding between our candidates and the E_{363}- R_{378} epitope of c-Myc. To perform the competitive FRET assay, varying concentrations of untagged ligand under consideration to compete with a fixed concentration of Rhodamine tagged NT-A1 for the binding sites on the Fluorescein tagged epitope which would prevent the rhodamine-tagged NT-A1 from binding to the Myc epitope and generating FRET signals. (Figure 3.11c).

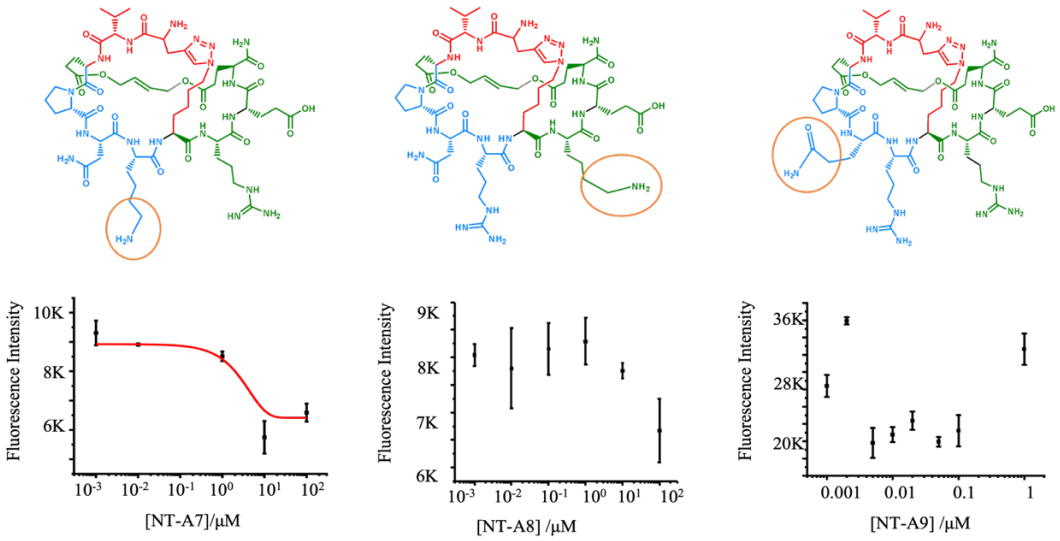
Consequently, increasing concentrations of a better binder would lead to decreasing FRET intensities. By contrast, a weaker ligand would not be able to disrupt the NT-A1 binding to c-Myc epitope. Therefore, a weaker binder would lead to a constant FRET. Our results revealed the importance of three essential amino acids, two arginines and one glutamine residue. The valine and glutamic Acid proved to be replaceable. (Figure 3.11 a).

Having realized the key components involved in the binding, we replaced each of the arginines with lysines one at a time and the asparagine with glutamine. To evaluate the effect of these modifications, we again used competitive FRET binding assays.

Competitive FRET assays revealed a better binder with lysine instead of arginine at position R_3 (with reference to Figure 3.11a), while the other positions did not result in better binding affinities (Figure 3.12a).

The binding assays needed to be validated using the recombinant c-Myc protein. Although our previous experience proved that epitope binding could easily translate well to the actual protein binding, there remains a risk that the surrounding residues of the epitope may interfere with the binding. We chose ELISA to validate the protein binding to recombinant Myc (Figure 3.14). ELISA results confirmed that NT-A7 was binding to c-Myc with a lower binding affinity (EC_{50} = 7.5 μM) as compared to NT-A1 (15.9 μM).

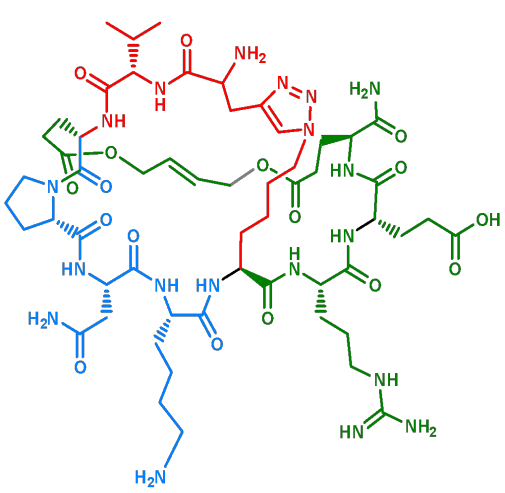
a)



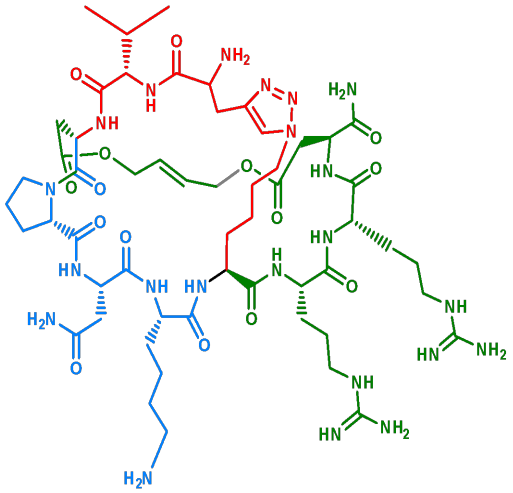
b)

FRET Donor	FRET Acceptor	Competitor
Fluorescein-Linker-Epitope	Rhodamine-Linker-NT-A1	**NT-A7**
		NT-A8
		NT-A9

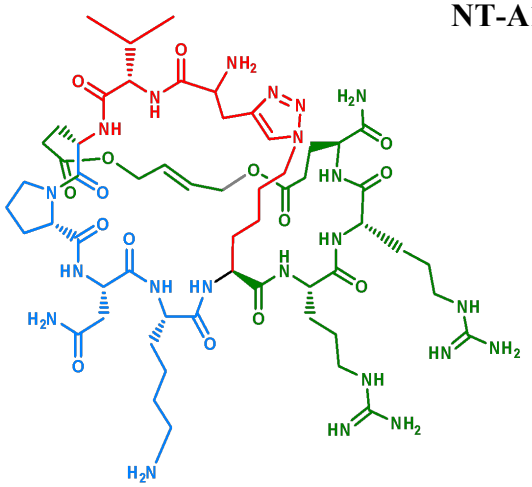
Figure 3.12 a) Three variants of NT-A1 aimed at optimization- NT-A7, NT-A8 and NT-A9 and their respective Competitive FRET studies. **b)** Table summarizing the performance of the variants.



NT-A7



NT-A14



NT- A10

Figure 3.13 Three variants of NT-A1 developed through chopping and changing amino acids. NT-A7 with arginine changed to lysine, NT-A10 with glutamic acid changed to arginine from NT-A10, and NT-A14 with a shorter ring sized version of NT-A10.

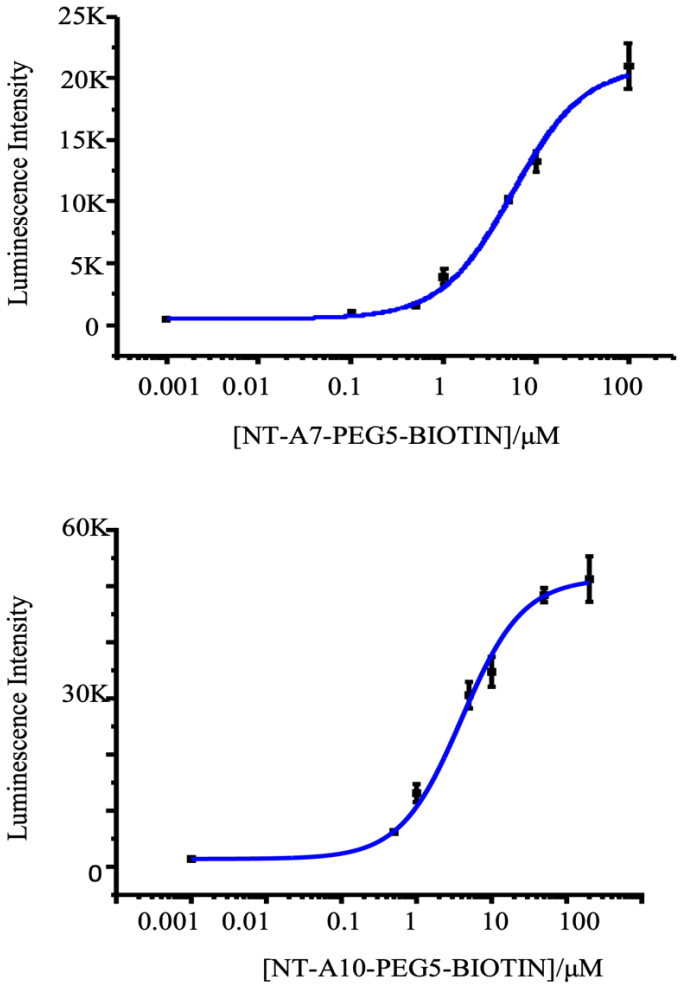


Figure 3.14 ELISA experiments probing binding of NT-A7 and NT-A10 against the recombinant c-Myc protein.

Cell proliferation tests in-vitro

With an optimized NT-A7, we sought out to test its success in affecting cell-proliferation. Previous experiences with NT-A1 led us to trust liposomes to transfer NT-A7 inside the cell, since the modifications to NT-A1 did not promise a significant improvement in autonomous cell-penetrating abilities. NT-A7 shows an improved IC50 value with in-vitro cell proliferation tests, although it was not significant, which is expected from similar ELISA results.

Since the glutamic acid at position R_1 was not useful and the nearby positively charged residues were critical to the binding, we made a bold hypothesis that replacing the glutamic acid with arginine (NT-A10) could improve the binding.

Interestingly, ELISA experiments with recombinant c-Myc showed good binding capabilities of NT-A10. Cell proliferation results showed NT-A10 desired improvement in terms IC_{50}=72 nM, while NT-A1 had 130 nM.

We then hypothesized that the ring size could be made slightly shorter to ensure a tighter grip on the target. The glutamic allyl ester pair, which constituted our second ring was replaced by an aspartic acid allyl ester pair. RCM cyclization on this pair yielded our next candidate in NT-A14 which retained all the other features of NT-A10. (Figure 3.13). Protein binding assays validated our assumptions (Figure 3.15b). Cell proliferation experiments showed a much-improved anti-proliferative effect IC_{50}=35 nM (Figure 3.15c). We carried out cell proliferation with NT-A14 in several cancer cell lines, cytotoxicity concerns have been safely considered as other cells also showed inhibition of proliferation (Figure 3.16), which highlighted the generalizability of NT-A14.

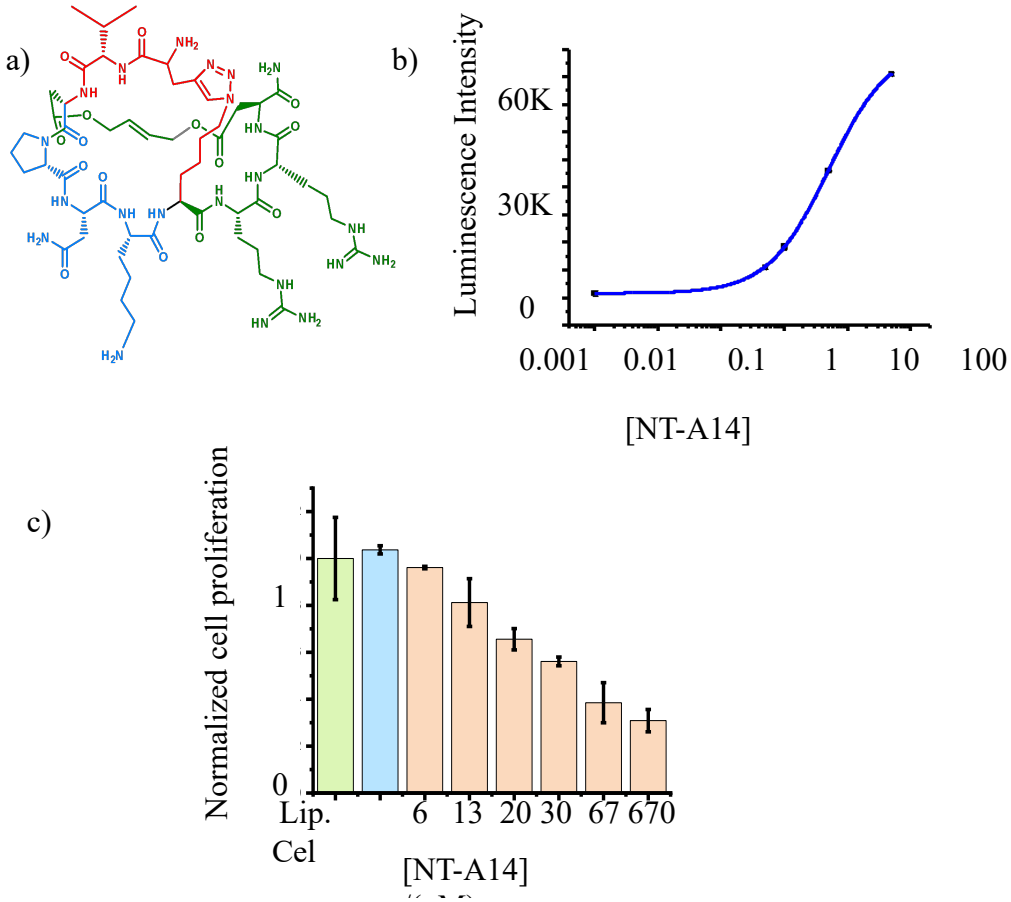
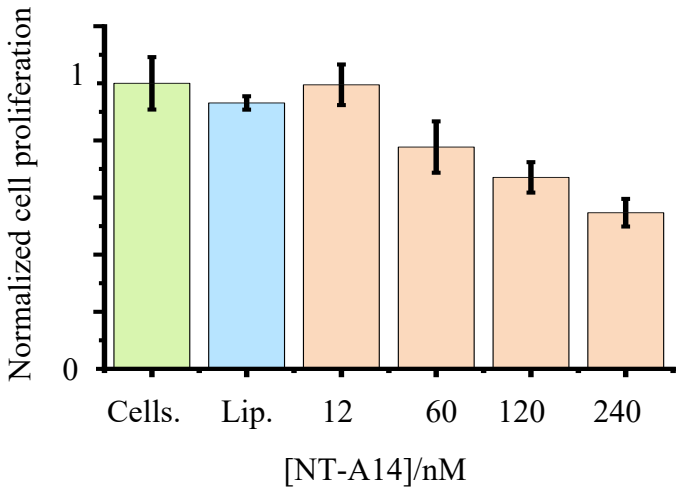


Figure 3.15 a) NT-A14 **b)** ELISA experiment probing the binding between NT-A14 and recombinant protein c-Myc **c)** Cell proliferation of U87 cells after treatment with NT-A14 entrapped in liposomes.



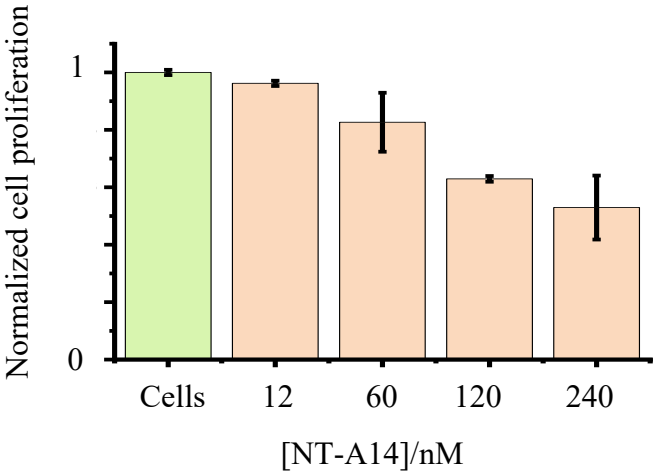


Figure 3.16 Cell proliferation data of WM115 (top) and MCF-7 (bottom) cells after treating them with NT-A14, the positive controls (green bar) showing cells treated with media only, and (blue bar) showing cells treated with liposomes.

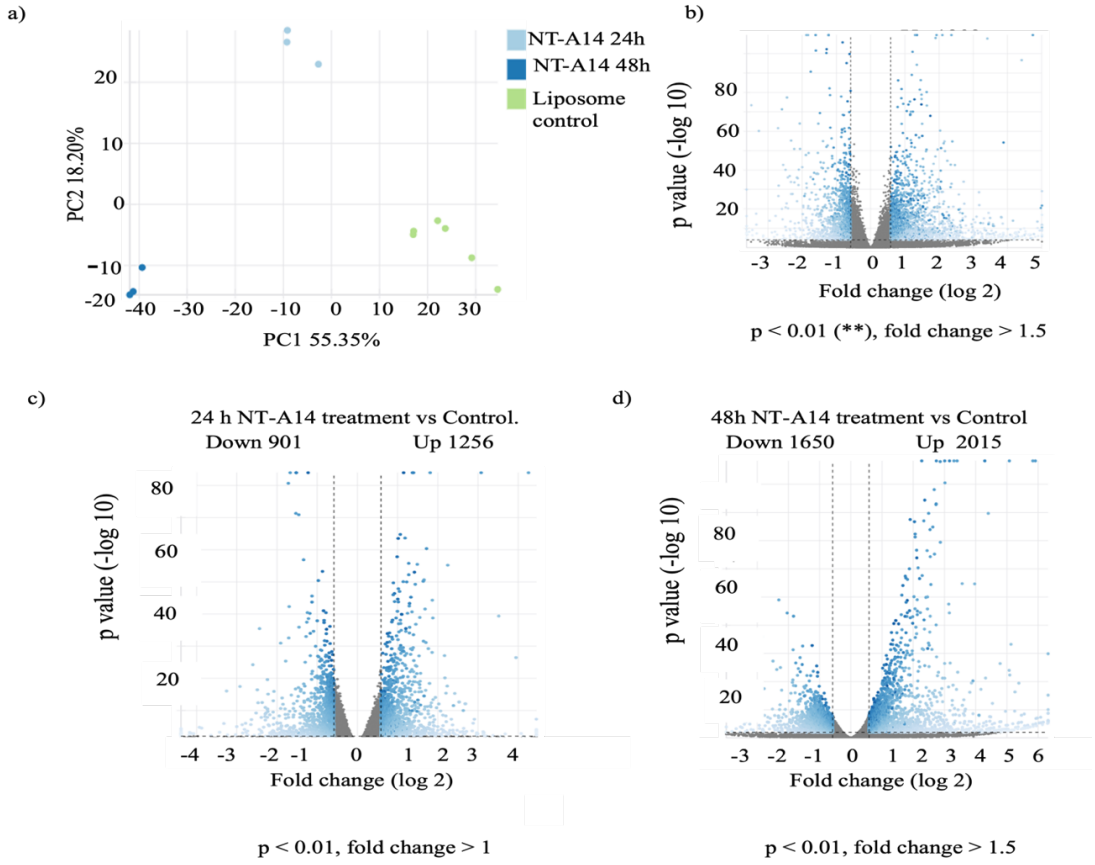
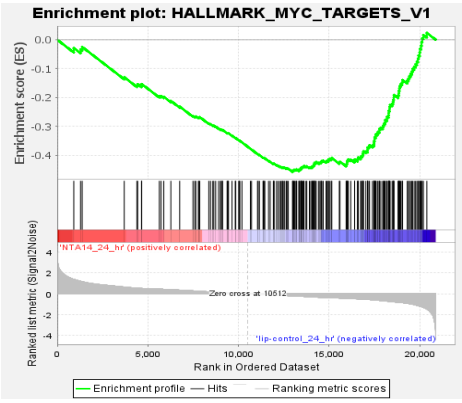


Figure 3.17 RNA-seq analysis of U87 cells treated with the NT-A14. **a)** Principal component analysis result of the RNA-seq datasets. **b)** Volcano plot showing the overall landscape of the differentially expressed genes in the NT-A14 24 h treatment versus the control and the 24 h treatment samples. (Cutoff criteria: $p < 0.01$, fold 1,5) **c)** Volcano plot showing the overall landscape of differentially expressed genes as the result of 24h treatment (left) and 48 h treatment (right). Cutoff criteria: $p < 0.01$, fold change > 1 (left) and fold change >1.5 (right).

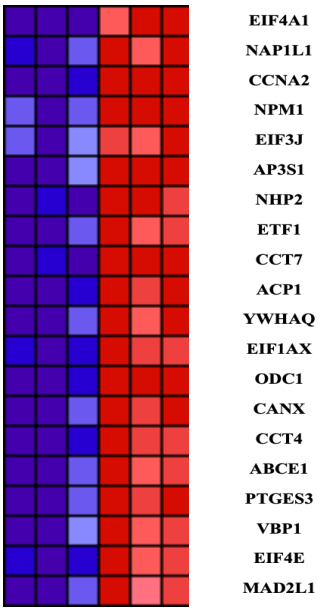
a)

Gene set	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX
HALLMARK_MYC_TARGETS_V1	193	-0.46	-2.27	0.000	0.000	0.000	7896
HALLMARK_E2F_TARGETS	195	-0.40	-1.97	0.000	0.003	0.003	6820
HALLMARK_KRAS_SIGNALING_UP	182	-0.38	-1.87	0.000	0.006	0.009	2668
HALLMARK_MTORC1_SIGNALING	195	-0.36	-1.76	0.000	0.010	0.019	5053
HALLMARK_PROTEIN_SECRETION	94	-0.36	-1.61	0.000	0.028	0.060	4015
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	191	-0.32	-1.54	0.000	0.045	0.114	2535
HALLMARK_G2M_CHECKPOINT	189	-0.30	-1.51	0.003	0.047	0.137	6447
HALLMARK_ANDROGEN_RESPONSE	94	-0.35	-1.50	0.012	0.048	0.157	3326
HALLMARK_SPERMATOGENESIS	108	-0.33	-1.48	0.008	0.049	0.181	6323

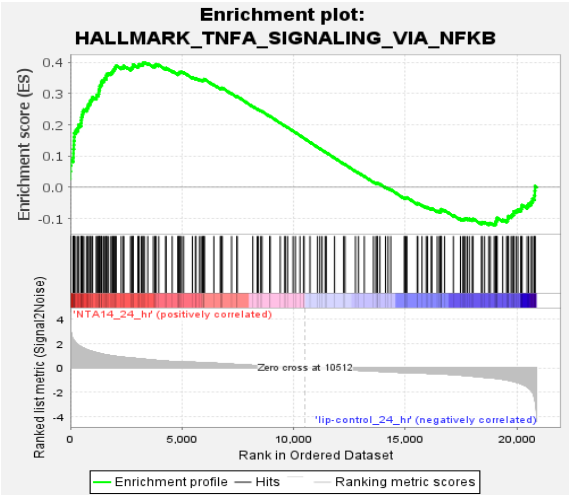
b)



c)



d)



e)

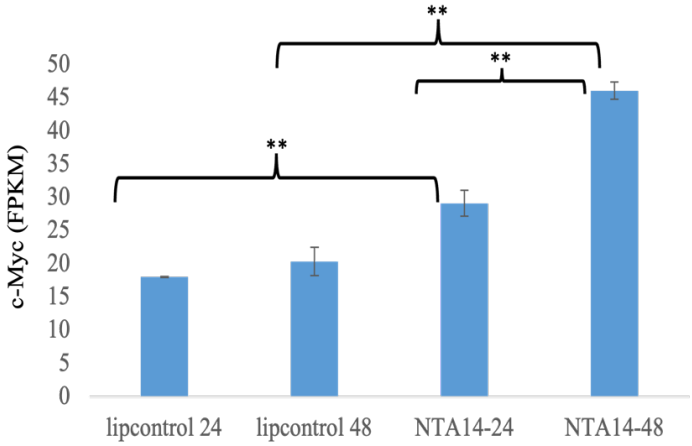
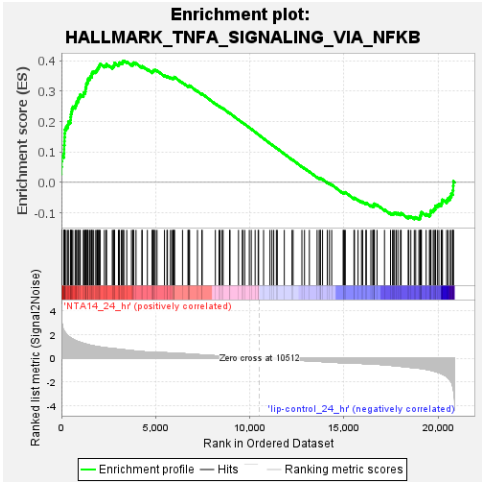


Figure 3.18 a) Gene sets enriched in the 24 h liposome treated sample vs NT-A14 treatment for 24 h. Enrichment plot of MYC_TARGETS_V1 after **b)** 24 h NT-A14 treatment and **c)** 20 most contributing genes after 24 h treatment **d)** 48 h NT-A14 treatment. **e)** T-test plots moderately high p (**) values between 0.01 and 0.001 for increased Myc expression levels.

a)

Gene Sets	Gene set	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val
HALLMARK_TNFA_SIGNALING_VIA_NFKB	193	0.40	1.82	0.000	0.003	0.009	3314
HALLMARK_CHOLESTEROL_HOMEOSTASIS	73	0.41	1.62	0.003	0.018	0.093	2889
HALLMARK_HYPOXIA	188	0.34	1.54	0.003	0.029	0.216	4225
HALLMARK_MYOGENESIS	184	0.31	1.38	0.017	0.101	0.678	4840

b)



c)

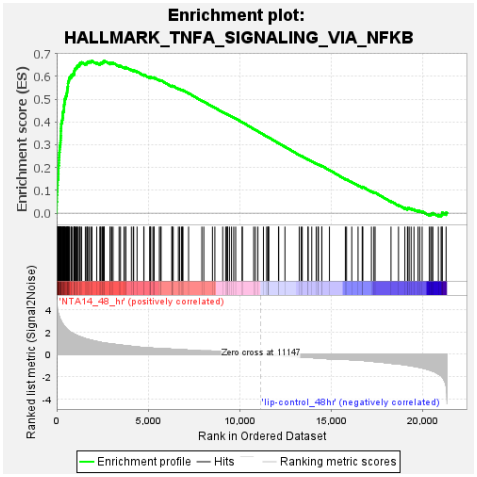


Figure 3.19 a) Gene sets enriched in the NT-A14 treatment after 24 h vs liposome treated sample. Enrichment plots for TNFA_SIGNALING_VIA_NFKB after **b)** 24 h treatment and **c)** 48 h treatment

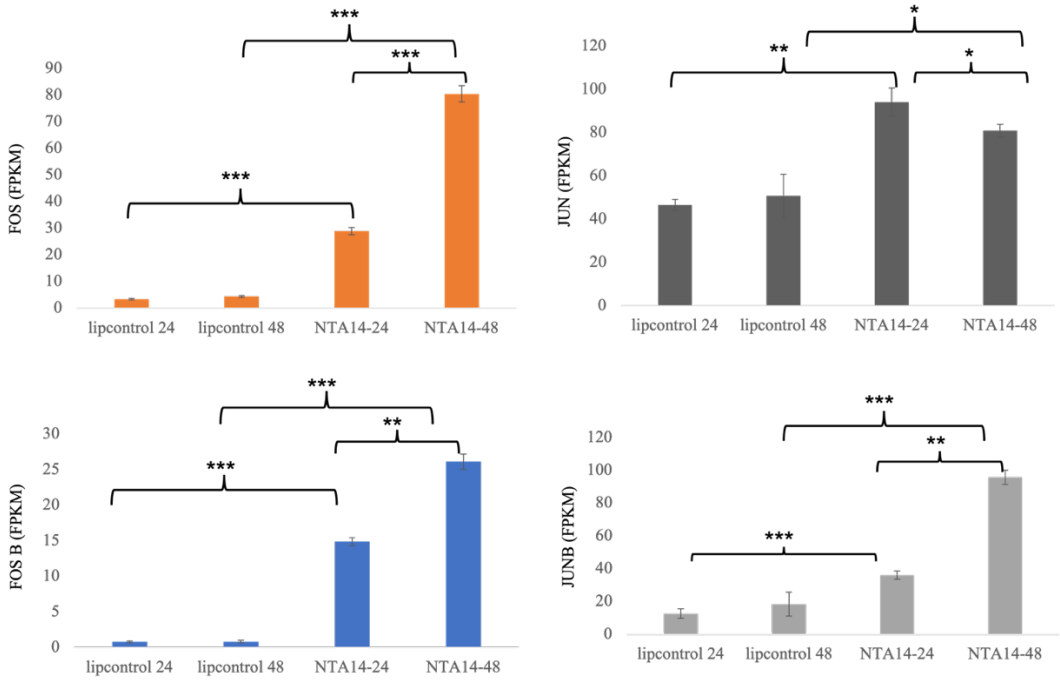


Figure 3.20 mRNA levels for a) FOS b) JUN c) FOS B d) JUN B. T-test was performed to assess the statistical significance of upregulation. *: $p < 0.05$. **: $p < 0.01$. ***: $p < 0.001$.

a)

Gene set	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX
HALLMARK_E2F_TARGETS	193	-0.42	-2.12	0.000	0.000	0.000	4487
HALLMARK_MYC_TARGETS_V1	192	-0.37	-1.81	0.000	0.005	0.009	5042
HALLMARK_G2M_CHECKPOINT	188	-0.34	-1.66	0.000	0.014	0.036	4329

b)

Gene set	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX
HALLMARK_TNFA_SIGNALING_VIA_NFKB	195	0.67	2.82	0.000	0.000	0.000	2600
HALLMARK_CHOLESTEROL_HOMEOSTASIS	71	0.57	2.09	0.000	0.000	0.000	3679
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	192	0.44	1.86	0.000	0.002	0.005	3085
HALLMARK_P53_PATHWAY	189	0.42	1.77	0.000	0.003	0.012	3468
HALLMARK_UV_RESPONSE_UP	147	0.43	1.77	0.000	0.002	0.012	4421
HALLMARK_MYOGENESIS	191	0.42	1.76	0.000	0.002	0.014	3971
HALLMARK_HYPOXIA	190	0.41	1.71	0.000	0.005	0.035	2785
HALLMARK_APOPTOSIS	156	0.41	1.67	0.001	0.007	0.055	3999
HALLMARK_TGF_BETA_SIGNALING	54	0.46	1.61	0.004	0.011	0.105	1119
HALLMARK_INFLAMMMATORY_RESPONSE	193	0.38	1.58	0.000	0.014	0.137	3407
HALLMARK_IL6_JAK_STAT3_SIGNALING	77	0.41	1.53	0.008	0.023	0.228	2915
HALLMARK_COAGULATION	130	0.36	1.46	0.019	0.044	0.429	5942
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	107	0.37	1.45	0.009	0.043	0.455	3146
HALLMARK_UV_RESPONSE_DN	135	0.35	1.43	0.013	0.053	0.547	2959
HALLMARK_IL2_STAT5_SIGNALING	188	0.34	1.42	0.012	0.053	0.660	4769
HALLMARK_KRAS_SIGNALING_DN	162	0.34	1.40	0.022	0.063	0.660	4769
HALLMARK_COMPLEMENT	188	0.31	1.32	0.039	0.124	0.891	3763
HALLMARK_ESTROGEN_RESPONSE_LATE	184	0.31	1.31	0.045	0.125	0.908	2399

Figure 3.21 Gene sets enriched after **a)** 48 h liposome treatment vs 48 h NT-A14 and **b)** 48 h NT-A14 vs 48 h liposome treatment.

To validate our proposed mechanism of action, we treated U87 cells with NT-A14-containing liposomes, extracted RNAs from the cells at 24 h and 48 h and performed transcriptome analysis by RNA-seq. The observed reads were processed and aligned using fastp, STAR and FeatureCounts to generate the gene expression of each sample. We performed the Principal component analysis to evaluate the overall variances of each sample. As shown in Figure 3.17 a, PC1 and PC2 captured 55% and 18% of the total variance, respectively. The three sample populations exhibited significant differences, with the strongest variance occurring between the control and 48-hr samples, which occupied the opposite ends of the PC1 axis. In comparison, the difference between the control and the 24 h samples was not prominent, as evidenced by their distance along the PCA axis.

Intrigued by the results above, we further examined the differentially expressed genes in the samples using Dseq2. As shown in the volcano plots (Figure 3.17 c,d), NT-A14 treatment led to profound changes in the overall transcription landscape. At 24 hours, 901 genes were downregulated, and 1256 genes were upregulated, as compared to the control sample. The effect of treatment became more prominent after 48 hours, where 1650 genes appeared to be downregulated, and 2015 genes were upregulated. These numbers echoed with the PCA analysis above and highlighted that cells were differently affected by NT-A14 treatment. Such a result also hints at the activation of potential survival pathways, which is worth exploring.

We then performed gene set enrichment analysis (GSEA) to identify critically involved pathways in response to NT-A14 treatment. Using the Hallmark gene sets, we first

compared the 24 hr NT-A14 treatment with the control. Using the cutoff criteria of $p<0.01$ and $q<0.2$, we identified nine pathways that were significantly downregulated after the treatment (Figure 3.18 a). As expected, the MYC_Targets pathway was the top enriched one, which underscored our hypothesized NT-A14 mechanism of action. As shown in Figure 3.18 b, 126 out of 192 genes participated in the core enrichment, which proved NT-A14 has indeed inhibited Myc transcription activities and greatly impacted the transcription targets. On the other hand, four gene sets appeared to be upregulated after 24 hour NT-A14 treatment (Figure 3.19a). Especially, the TNFA-NFKB signaling pathway was heavily involved. A closer analysis revealed that 65 out of 193 genes participated in the core enrichment. (Figure 3.19 b). We noticed that AP-1 family transcription factors, FOS, FOSB, JUN, JUNB were significantly upregulated (Figure 3.20 a-d). Because the FOS-JUN heterodimer closely resembles the MYC-MAX dimer, and they share many transcription targets, this finding led us to a potential survival pathway where cells upregulate the AP-1 to compensate for the loss of MYC transcription activities. In addition, we found that MYC expression levels went up after NT-A14 treatment (Figure 3.18 e), which indicated that cells also tried to increase the transcription of Myc itself to compensate the loss of its function.

Similarly, we found that MYC targets remained significantly downregulated after 48 hours of NT-A14 treatment (Figure 3.18d). However, the number of genes that were involved in the core enrichment decreased significantly compared with the 24 hour treatment (65 versus 126) (Figure 3.18d). This result indicated that the Myc inhibition effects were alleviated in these cells. On the other hand, TNFA-NFKB signaling

remained significantly upregulated (Figure 3.19) with even higher levels of FOS and JUN transcription factors (Figure 3.20). Taken together, these results supported our prior hypothesis that cells upregulated MYC as well as the AP-1 transcription factors to compensate for the MYC inhibition.

3.4 Conclusion

In summary, we optimized the anti-proliferative capabilities of our previously discovered drug NT-A1 and designed it into NT-A14 through a variety of med-chem changes. We validated the target specificity and mechanism of action by using transcriptomic analysis, and we demonstrated its therapeutic efficacy in several different cell lines.

3.5 References

1. Linchong, S; Ping, G; Small molecules remain on target for c-Myc. *eLife* **2017**, *6.*

2. Chen, H.; Liu, H.; Qing, G. Targeting Oncogenic Myc as a Strategy for Cancer Treatment. Signal Transduct. *Target. Ther.* **2018**, *3 (1)*, 1−7.

3. Levens, D. Disentangling the MYC web. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99*, 5757–5759.

4. Dang, C. V.; O'Donnell, K. A.; Zeller, K. I.; Nguyen, T.; Osthus, R. C.; Li, F. *Semin. Cancer Biol.* **2006**, *16*, 253–264

5. Hueber AO, Zornig M, Lyon D, Suda T, Nagata S, Evan GI. Requirement for the CD95 receptor-ligand pathway in c-Myc-induced apoptosis. *Science* **1997**,*278*,1305–1309.

6. Miller, D. M.; Thomas, S. D.; Islam, A.; Muench, D.; Sedoris, K. *Clin. Cancer Res.* **2012**, *18*, 5546.

7. Gustafson, W. C.; Weiss, W. A. Myc Proteins as Therapeutic Targets. *Oncogene* **2010**, *29 (9),* 1249−1259.

8. Prochownik, E. V. Expert ReV. c-Myc as a therapeutic target in cancer. *Anticancer Ther.* **2004**, *4*, 289–302.

9. Hsieh, A. L., *et al.* MYC and metabolism on the path to cancer. *Semin. Cell Dev. Biol.* **2015** ,*43*, 11–21.

10. Sun, X. X. *et al.* The nucleolar ubiquitin-specific protease USP36 deubiquitinates and stabilizes c-Myc. *Proc. Natl Acad. Sci. USA* **2015**, *112*, 3734–3739.

11. Dawson, M.; *et al*. *Nature* **2011**, *478*, 529–533.

12. Berg, T. *et al. Proc. Natl Acad. Sci. USA.* Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts. **2002**, *99*, 3830–3835.

13. Li, Z.; Shao, S.; Ren, X.; Sun, J.; Guo, Z.; Wang, S.; Song, M. M.; Chang, C. A.; Xue M. Construction of a sequenceable protein mimetic peptide library with a true 3D diversifiable chemical space. *J Am Chem Soc* 2018, *140*, 14552–14556.

Chapter 4: The exploration for short cyclic peptides with enzymatic activities

4.1 Introduction

Enzymes have an exceptional ability to affect chemo, stereo, and regioselective transformations due to their complex macro-molecular structure, which can form scaffolds and enable bond breaking and forming. This ability occurs due to its ability to form various chemical associations like hydrogen bonds, covalent bonds, and salt bridges, with the substrates and stabilizes the transition states.

A synthetic peptide library consisting of millions of peptides, each on a single resin bead can be considered to resemble a miniature protein library [1,2]. We then envisioned that short monocyclic peptides could have an "active site" and enough functional diversity to form a variety of chemical interactions with the substrate it entraps, which may lead to enzymatic activity. Indeed, peptides of various lengths have been employed as catalysts independently in some studies[3,4].

In many examples, metal-based catalysts or organo-catalysts have often been replaced by single amino acid or a short-peptide-based catalyst for a while now[6,7]. Peptide-based catalysts provide diversity as a compound as a combination of amino acids with diverse functional groups can be coupled with extremely high specificity. Several short peptides with monoamine oxidase activity, obtained via colorimetric screening of a peptide library, would be extremely helpful in the understanding of the plausible mechanisms,

which in turn would provide incentives for further improvements by changing amino acids and achieving a higher catalytic efficiency ratio, i.e., K_{cat}/K_m.

With a multi-step catalytic reaction envisaged, we consider cyclic peptides as better candidates than their linear counterparts. The cyclic ring is expected to accommodate a small molecule, a key requirement for enzyme catalysis. Because the substrate has to bind with the catalyst (the peptide) long enough to enable transformation, the binding affinity has to be strong enough to allow for this prolonged recognition. In addition, linear peptides often lack the ability of cell penetration[8]. Therefore, if we aim to target intracellular reactions, cyclic peptides would be much more useful.

Herein, we set to test if we could identify monocyclic peptide sequences with enzymatic activities from a library. These generic library components consist of eight amino acid sequences. Five out of the eight amino acids are modular sites with randomly inserted natural amino acids, which constitute the functional diversity of the peptides.

4.2 Experimental

Materials

TentaGel S-NH_2 resin and Rink amide MBHA resin was purchased from Rapp Polymere GmbH and from Aapptec (Louisville, KY) respectively. Fmoc-protected amino acids were obtained from Anaspec (Fremont, CA). Fmoc-L-propargylglycine (Pra) and FmocLys(N_3)-OH (Az4), were obtained from Chempep (Wellington, FL) and Chem-Impex (Wood Dale, IL), respectively. The coupling reagent 2-(1H-benzotriazol-1-yl)-

1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 99.6%) was purchased from Chem-Impex (Wood Dale, IL). Diisopropylethylamine (DIEA, 99.5%) was purchased from ACROS (Germany). Phenyl isothiocyanate (PhNCS) and triisopropylsilane (TIPS) were obtained from TCI (Portland, OR). Piperidine was purchased from Alfa Aesar (Ward Hill, MA). Cyanogen bromide (CNBr) was obtained from ACROS (Pittsburg, PA). Cuprous iodide (CuI), and α-cyano-4-hydroxycinnamic acid (CHCA) were obtained from Sigma-Aldrich (St. Louis, MO). Nitro blue tetrazolium (NBT) were purchased from Promega (Madison, WI). Tris base, sodium phosphate dibasic anhydrous (Na_2HPO_4, 99.6%), sodium phosphate monobasic monohydrate (NaH_2PO_4, 99.4%), sodium chloride (NaCl), ascorbic acid, N,N′-dimethylformamide (DMF), and dichloromethane (DCM) were purchased from Thermo Fisher Scientific (Waltham, MA).

Synthesis of the library

A ‘one-bead-one compound’ library using ‘split-and-pool’ method is synthesized on TentaGel beads. 2g of Tentagel beads were incubated with DMF overnight. A solution of azidolysine, DIEA and HATU were dissolved in DMF, and were shaken with the beads for 2 hours. The beads were then washed with DMF for five times. The Fmoc group from azidolysine was cleaved off by reacting the beads with 20% Piperidine/DMF solution. The beads were washed again with DMF. The beads were split into 18 different portions, each portion was used to couple against each of 18 amino acids, Gly, Glu, Ala, Asp, Val, Leu, Arg, Lys, His, Ile, Gln, Asn, Ser, Thr, Trp, Tyr, Phe, Pro. For this, we make solutions of each of these amino acids in DMF, add DIEA, and HATU. We incubate each

of the solutions incubate with its designated portion of beads for couple of hours. After this, the beads were pooled together and washed with 20 mL of DMF for five times. The amino acids have Fmoc protections at this point and they were deprotected by incubating with 20% Piperidine/DMF (v/v) (5 min x 3). The beads were washed again with DMF for five times. The beads were then mixed thoroughly to make them homogenous. The split and mix process was repeated for four more times to have five randomized amino acid sequences in the library. After that, propargylglycine was coupled to the beads in a similar manner as Azido lysine.

Click reaction to cyclize the peptide library

Copper (I) Iodide (5 equivalent) and Ascorbic Acid (15 equivalent) were dissolved in 20 mL 20% Lutidine/DMF (v/v). The beads were incubated with this solution overnight, with mild shaking. After that, the beads were washed with a 20 mL solution of 5% NaDTC (w/v), 5% DIEA (v/v), in DMF (5 min x 5) to form a complexation with the residual copper and extract it out of the beads. The beads were then washed with DMF.

Colorimetric screening to elucidate hits

We perform a colorimetric reduction assay of MTT {3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide} to elucidate peptide sequences that would react positively and perform monoamine oxidase activity. Hence, creating a redox reaction, which has the MTT getting reduced to the purple-colored formazan which forms a precipitate in an aqueous solution. The precipitate settles on the surface of beads, giving it a purple color,

which could be detected and extracted. The positive sequences would produce a purple coloration, which would validate the sequence. Enzyme efficiency tests would be carried on with the positive peptide and variation of substrates would generate a complete profile on its biocatalytic activity with further understanding of its scope of substrate compatibility and pave the pathway for intracellular applicability. The screening test was carried out with reaction times of 10 and 25 minutes and at pH 8 and 9. The reaction was not carried out at a more basic solution of pH 10, to prevent hydrolysis. The fastest reactants were obtained with 10 minutes of incubation at pH 9 in the form of several hit beads.

Linearization and sequencing

The hit beads were incubated with phenyl isothiocyanate (2.5% v/v) in pyridine: water (1:1) and heated to 50 °C for 30 min. The solution was drained, and the beads were washed with DMF×3, EtOAc ×3 and dried. Afterwards, the beads were treated with TFA and at 50 °C for 10 min and then 20% TFA in H_2O (v/v) at 80 °C for 10 min. The individual beads were treated with 10 μL of 0.5 M CNBr /0.2 M HCl solution in a microcentrifuge tube. The tubes were heated in microwave oven for 1 min and dried. The residue was dissolved in 0.5 μL (4 mg/mL CHCA in 50% acetonitrile in H_2O with 0.1% TFA) and spotted on a MALDI plate, ready for MALDI-TOF.

Mass spectrometry

A SCIEX 5800 mass spectrometer was used to obtain the MS-MS spectra using the MALDI-TOF technique.

Synthesis of hit peptide on Rink Amide resin

The hit peptide, cy (PHLYR) was synthesized with similar solid-phase synthesis protocols as the library synthesis on TentaGel beads, except the beads had to be treated with 20% Piperidine/DMF (v/v) before the start of the synthesis steps. The peptide was cleaved off the Rink amide beads using a solution of TFA, tri-isopropyl silane and water in a ratio of 95:2.5:2.5 by volume. The peptides were purified using RP-HPLC and analyzed using SCIEX 5800 mass spectrometer using a MALDI-TOF analysis.

Reverse Phase HPLC

RP-HPLC with a preparative column (Kinetex 5 μm EVO, 250 × 21.2 mm^2) and a multi-wavelength UV−vis detector monitoring the absorbance at 215, 280, 480, and 560 nm. 0−100% acetonitrile (with 0.1% TFA) in water (with 0.1% TFA) was maintained as the flow rate at 15 mL/min. The hit peptide was purified using this setup. Thermo Ultimate 3000BX HPLC was the instrument used.

4.3 Results and discussion

Design monocyclic peptide library

TentaGel, a polymer resin, is used as a support for the peptide synthesis. After attaching the common amino acids, the beads are split into 18 different fractions to have 18 different amino acids attached. Out of 20 amino acids, naturally present, 18 are used in the functional part of the library, to avoid catalytic poisoning by the sulfur-containing amino acids-Cysteine and Methionine.

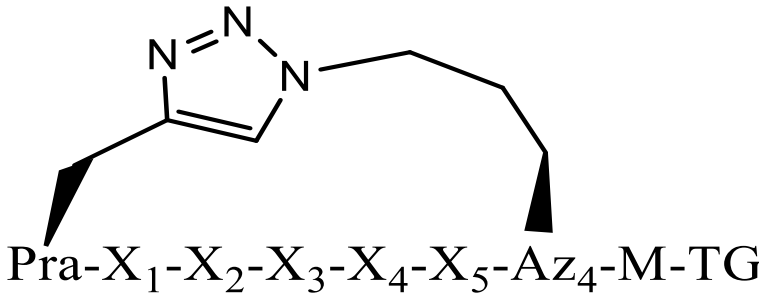


Figure 4.1 The monocyclic peptide library synthesized on TentaGel (TG) beads, used for screening. X_1 -X_5 represents the five randomized amino acids added through five cycles of iterative split and pool library synthesis. Pra (propargyl glycine), Az_4 (azido lysine), M (methionine) are added to the entire set of beads for specific purposes.

Propargyl glycine and Azido lysine, would comprise the pair responsible for CuAAC, and form a ring enclosing the five functional amino acid residues, encompassing the reaction site, i.e., active site. The alkyne and the azide groups, have the advantage of being inert to most types of reactions. Hence, the selectivity of the click reaction is ensured and unlike other forms of functional groups, these do not need to be protected.

Methionine was added to the sequence at the C-terminal, to facilitate the cleavage of peptides from the beads using CNBr.

Five iterative cycles of split and mix cycles are performed, so that the peptide beads have a variable peptide chain of 5 amino acids in the functional part of the peptide. The split-and-pool method of making a library generates 18^5 different compounds, a large pool of compounds, which is ideal for high-throughput screening. Once the library is synthesized, it can be used for experiments repeatedly. Forming disulfide linkages to develop a cyclic peptide is avoided to rule out interferences from the sulfur groups towards our screening tests. The positive beads from the screening and the respective validation tests vindicate our hypothesis and give an experimental justification.

Screening

A library can bear beads that carry peptide sequences to catalyze a reaction typically carried out by monoamine oxidase. MTT, a yellow compound, on reduction, forms a purple-colored product called formazan. These beads would be identifiable. The MTT-formazan assay is one of the commonest tests in the world of biology. The mitochondrial respiration causes reduction of MTT to formazan. Oxidation of a monoamine to imine or a ketone/aldehyde would reduce MTT into formazan. The beads carrying peptides would be the sites of reaction and the purple precipitate of formazan would settle on the beads, making them clearly identifiable. The library is screened against MTT dissolved in a buffer solution of PBS (Phosphate Buffered Saline), containing Glycine. MTT is traditionally dissolved in PBS buffer, for cell proliferation test, making it an automatic

choice as a solvent. PBS has a pH close to 7.4 (close to neutral), and it is isotonic with the cell components. Glycine is a small molecule that is highly soluble in the buffer solution and a target molecule as a substrate that is oxidized, resulting in the reduction of MTT. Under the microscope, a certain number of beads would be expected to be visible quite distinctly, and they could be collected and separated. The beads which are collected would bear the cyclized peptide. Edman degradation is performed on the beads to linearize the cyclic structure. An MS-MS data on the linear version peptide, analyzed by a MALDI-TOF-TOF mass spectrometer, would help sequence the peptide. This set of positive sequences would be termed as hit sequences. The possible hit sequences can be synthesized manually or with the help of a peptide synthesizer in a few days.

Validation

Sequencing MS-MS data yields- NH_2-Pra-P-H-L-Y-R-Az_4-Met

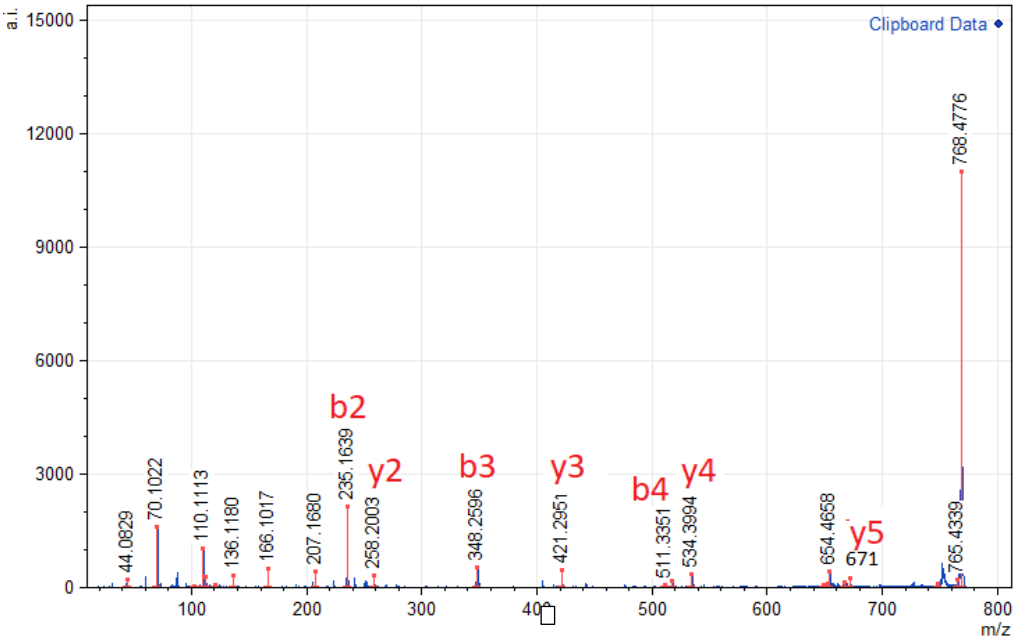


Figure 4.2 MS-MS spectra of hit sequence PHLYR, solving the spectra using the b and y ions provides us with the information of the hit sequence.

The hit peptide cy (P-H-L-Y-R) was synthesized on TentaGel beads. With all the beads carrying the same known peptide, the reaction can be carried out by incubating the beads with MTT dissolved in the same PBS-glycine buffer as before. All the beads expectedly became spontaneously dark colored due to the formation of formazan. The beads were washed with DMSO to wash away the color, and we sought to re-use them for repetitive tests. Fortunately, the beads could be re-used, revealing the reversible nature of the enzymatic action of the peptide on TentaGel beads. The experiment was repeated with

NBT (nitro-blue tetrazolium), and a similar result was obtained in the form of instant formazan precipitation.

Solution phase tests

The hit peptide cy (P-H-LY-R) was synthesized on Rink Amide to test for solution-phase reactions with the hit peptide. Since formazan dissolves in DMSO, the reaction media can be diluted with DMSO, and UV-visible absorption at 570 nm can be monitored over time to determine the kinetics of the catalyzed reaction, and it was dissolved in the PBS-glycine buffer solution. This solution would also be able to catalyze the oxidase reaction. The catalytic efficiency ratio, i.e., K_{cat}/K_m would be estimated to determine the efficiency of the enzyme. Several control experiments would be added to have a detailed picture. The colored solution, i.e., formazan in DMSO, could easily be analyzed spectro-photometrically using Beer's Law.

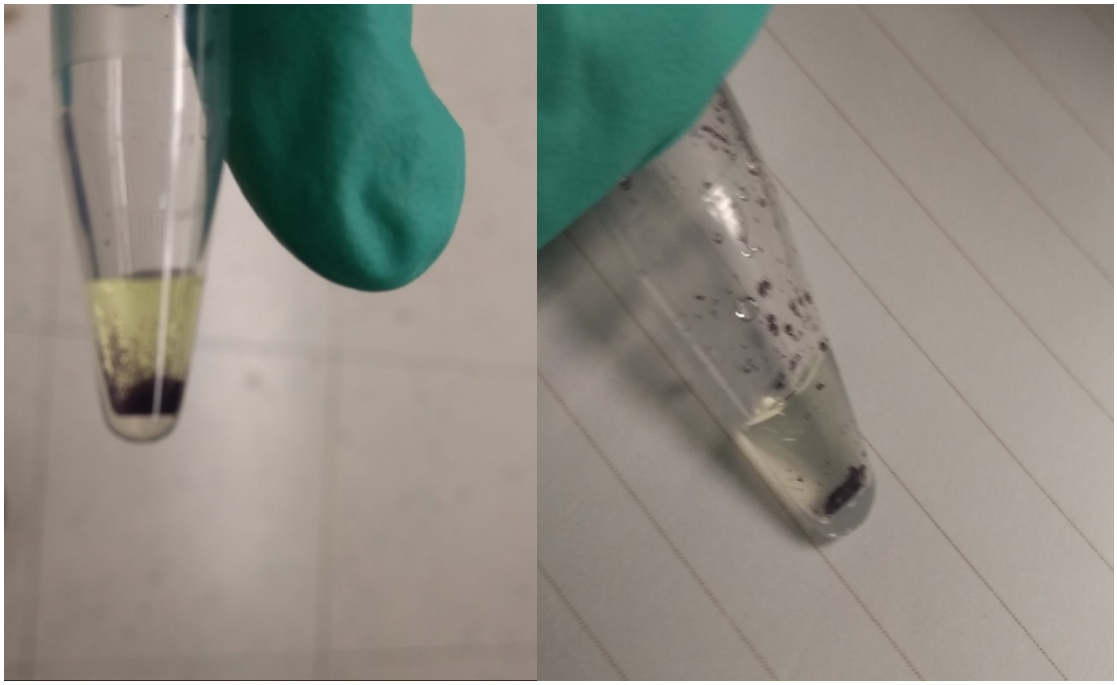


Figure 4.3 cy (P-H-LY-R) on TentaGel beads are capable of instantly reducing MTT solution (left) and NBT solution (right) into their respective formazans, characterized by intensely purple-colored beads.

The reaction was carried out in the solution phase. Unfortunately, the reduction of MTT was not evident as in the case of the reaction on beads, despite carrying out the reaction at a wide range of pH conditions. The results made us re-assess our strategy by re-examining the implications of our positive results on TentaGel beads.

Linear sequence on TentaGel beads

We synthesized the linear version of the functional part of the hit sequence, P-H-L-Y-R on TentaGel beads. We sought to repeat the experiment on the beads, and there was an instant MTT reduction at the same reaction conditions. As a negative control, we

repeated the experiment on empty TentaGel beads, and we did not observe any reduction signs. While our hypothesis regarding the requirement of a cyclized ring did not substantiate, we could rule out the possibility of residual copper catalyst from cyclization reaction taking part in the redox reaction.

4. 4 Conclusion

The absence of positive results with the hit peptide was a major roadblock to our exploration of enzymatic peptides. It further complicated our search of mechanistic details of the enzymatic action. The peptides transformed in due course of enzymatic activity on the beads could not be assessed. The ability of a peptide sequence to have enzymatic activity could be showcased in our study. However, mechanistic studies were not possible.

Our study provides us with an outlook to provide directions concerning using peptides for such a purpose. The requirement of solid support for the enzymatic reaction hints at the necessity of the peptides to have an interacting surface to react with the substrate. While key strategies of a peptide-based enzyme would prioritize the scaffold size and cyclization to prevent enzymatic degradation, we would now investigate interacting surfaces when we design a high throughput library based on target-based screening, fragment-based drug discovery, which are built on polymer resin beads or encoded on DNA. These are new aspects that should be considered when we envision the peptide engaging with the substrate of interest.

Another positive aspect was the reversibility of enzymatic action was the availability of the peptides for re-use after washing them off with DMSO. The most fundamental aspect of a catalyst was evident in the form of regenerated peptides. The beads did not react favorably to the reduction reactions at slightly acidic pH (close to 5), which clarifies the necessity of very specific amino acid side chains for the activity on a particular substrate.

4.5 References

1. Das, S.; *et al.* A General Synthetic Approach for Designing Epitope Targeted Macrocyclic Peptide Ligands. *Angew Chem., Int. Ed.* **2015**, *54*, 13219-13224.

2. Lam, K.S; Lehman, A, L.; Song, A.; Doan, N.; Enstrom, A. Maxwell, J.; Liu, R. *Methods Enzymol.* **2003**, *369*, 298-322.

3. Copeland, G. T.; Miller, S. J. *J. Am. Chem. Soc.* **2001**, *123*, 6496.

4. Jarvo, E. R.; Copeland, G. T.; Papaioannou, N.; Bonitatebus, P. J.; Miller, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 11638-11643.

5.Gro¨ger, H.; Wilken, J. The Application of L-Proline as an enzyme Mimic and Further New Asymmetric Syntheses Using small Organic Molecules as Chiral Catalysts. *Angew. Chem., Int. Ed.* **2001**, *40*, 529.

6. List, B.; Lerner, R, A.; Barbas III, C.F.; *J. Am. Chem. Soc* **2000**,*122*, 2395.

7. Dougherty, P.G.; Sahni, A; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019**, *119,* 10241-10287.

Chapter 5: Concluding remarks

We have explored the applicability of cyclic peptides in three different contexts in the previous three chapters. We started our discussion with the usage of cyclic peptides as cell penetrable, robust therapeutics. With the diversity of functional groups forming all possible kinds of interactions, it was useful to explore the best sequence as a candidate from a library. The use of such a library with three-dimensional orientations was a concept that was generalizable. Certainly, the process of discovering a highly specific binder for a specific isoform was a fitting way to introduce the feasibility. The homologous nature of ERK isoforms and the need to differentiate their signaling process differently was our goal, and we have devised a way to address it by demonstrating the binding against two epitopes. The ability to bind against the chosen epitopes was well translated to protein binding specificity as well and validated by *in vitro* activities as well. We would be interested in the applicability of these probes in the future. The use of dye-tagged specific probes develops interest in creating binding probes which can image various activities within the cell by associating itself with a protein of interest. Continuous single-cell imaging can be attempted with the dye-tagged probes that we created. A pair of probes attaching to the peptide can combine to generate FRET signals which can coordinate with the dynamic process like ERK signaling owing to differential binding of these probes during the cascading signaling processes. We can create a method to assess the tumor cell heterogeneity by analyzing the ERK probes at the single-cell level. We can extend our methodology to other signaling proteins by achieving it using the same or different screening process.

PPI inhibition was also addressed efficiently in Chapter 3. An attractive target like Myc could be specifically bound to, and we could demonstrate the implications of such direct blockage of PPT on the core biology of the cells. Our previous success in having a binder at hand proved to be useful. While chapter 2 explored the design of a bicyclic peptide library and the usage of such library to screen against protein targets, chapter 3 is a continuation of the robustness of such a process. It coherently connects the success from screening to improvise the abilities of the hit by utilizing the structural and functional diversity which we had incorporated during the screening process. We duly explored the impacts on the blockage of a key cell proliferative pathway involving PPI and the generalizability of Myc targeting. The modifications on the bicyclic peptide provided us insights into even more structural considerations while designing a peptide or a library. We discovered the survival mechanisms that the cell undertakes to combat the process of c-Myc inhibition. It is imperative to attempt a drug synergy treatment, which is touched upon, to combat the survival process and aim to achieve a better anti-proliferative activity. Identifying the alternative activation of other transcription pathways promises the ability to achieve better results. In the future, we can look forward to extending our technology against other undruggable proteins and improvise on the design of our peptide to create more cyclization and complexity id required.

In Chapter 4, we attempted a miniature enzyme, which could be the smallest to entrap a substrate on its own. We could not shed light on the mechanistic details of our hypothesized enzymatic action, but we could observe a pH-dependent enzyme activity on a surface-attached peptide, which makes us realize the importance of looking into

engaging surfaces while designing peptide libraries. In the future, we can try to address our concerns by revisiting our peptide design strategy to engulf a substrate, and a stronger and more engaging scaffold would be prioritized. We can build the most miniaturized version of proteins capable of acting as an enzyme independently, and we can also create the possibility of protein-protein interaction activation, which is a relatively newer research direction.

To conclude, we managed to show what complexity can impart on a peptide and the numerous possibilities that could be achieved by optimizing its chemically functional space. Our macrocyclic peptides were easy to build and, more importantly, easily modifiable to optimized versions. Our designs were aimed at creating a three-dimensional structure where we could make our targets interact with the amino acids collectively. Designing a screening method that is inexpensive and less time-consuming was our goal, and it was utilized by making slight modifications from time to time according to the target of interest. Overall, we managed to provide an insight into the diverse usage of cyclic peptides ranging from PPI-based therapeutics to signaling probes and miniature enzymes. The biological capabilities of a synthesized peptide can have diverse applications owing to the additional features we incorporate, taking inspiration from nature.