A Step Toward Achieving Peptide Based Therapeutics I. Analysis of the Effect of Olefin Position and Linker Length in Hydrocarbon Stapled Peptides II. Development and Structural Studies of Direct Peptide Inhibitors of Ras

A step toward achieving peptide based therapeutics

I. Analysis of the effect of olefin position and linker length in hydrocarbon stapled peptides
II. Development and structural studies of direct peptide inhibitors of Ras

A dissertation presented
by

So Youn Shim

to

The *Department of Chemistry and Chemical Biology*

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Chemistry

Harvard University
Cambridge, Massachusetts

May 2015

Professor Gregory L. Verdine So Youn Shim

A step toward achieving peptide based therapeutics

I. Analysis of the effect of olefin position and linker length in hydrocarbon stapled peptides

II. Development and structural studies of direct peptide inhibitors of Ras

Abstract

Traditionally the two major classes of drugs, small molecules and biologics, have been extensively used in drug discovery. However, it is difficult to inhibit protein-protein interactions with small molecules, and targets of biologics are limited to extracellular proteins. Considering such limitations, there has been a call for a new therapeutic modality to target 'undruggable' proteins. Structurally-stabilized peptide therapeutics has emerged as a new class that could overcome the limitations of the previously established drug classes while retaining their benefits. Among several strategies that have been employed to constrain the conformation of peptides to enhance their pharmacological properties, the all-hydrocarbon stapling system has been particularly successful. This system combines the helix-stabilizing effects of alpha methylation with peptide macrocyclization, which confers improvement not only in helix stabilization but also in protease resistance and cell permeability. The first part of my dissertation focuses on exploring this system further by shortening the established hydrocarbon bridge length as well as moving the position of the olefin along the cross-link, and looking at the effects of these attributes on the helical propensity of the peptides by circular dichroism.

In cases where there is no structural information or obvious binding partner that has an alpha-helical component that could be used for rational design of a stapled peptide inhibitor, screening a naïve library of stabilized scaffolds using yeast cell-surface display would be an alternative starting point. Oncogenic K-Ras is one such case; the structures of Ras with its effectors have shown that they interact with Ras via beta-sheet complementation. The second part of my dissertation will describe efforts to develop peptide inhibitors that target Ras by using screening and directed evolution of libraries of an avian pancreatic polypeptide-based scaffold by yeast cell-surface display. We have identified lead peptides that disrupt the binding of effectors to Ras, and the structures of these peptides bound to Ras have been solved and analyzed, leading us to improve the lead peptides further by rational design.

Table of Contents

Acknowledgements ..... **vii**

List of figures ..... **xi**

List of Tables ..... **xv**

A step toward achieving peptide based therapeutics **1**

Part I. Analysis of the effect of olefin position and linker length in hydrocarbon stapled peptides **6**

Chapter 1. Introduction **6**

Chapter 2. Shorter hydrocarbon bridge in *i, i*+4 stapled peptides **8**

Chapter 3. 6-carbon *i, i*+3 stapled peptides **19**

Chapter 4. 4-carbon *i, i*+3 stapled peptides **25**

Chapter 5. Conclusions and Future directions **36**

Chapter 6. Experimental Method **39**

Part II. Development and structural studies of direct peptide inhibitors of Ras **43**

Chapter 1. Ras protein **43**

1-1. What is Ras? 43

1-2. Ras and tumor formation 47

1-3. Strategies for Ras inhibition 49

Chapter 2. Directed evolution via yeast surface display **50**

2-1. Yeast surface display 50

2-2. Scaffolds 52

2-3. Library generation & initial screens 56

2-4. Screening with aPP libraries 64

Chapter 3. RDA (Ras binding Dimeric aPP) peptides 69

3-1. Structural analysis 69

3-2. Stabilization of state 1 in Ras by RDA peptides 79

3-3. *i, i+4+4+4* stitched peptide 84

Chapter 4. Heterodimeric Ras inhibitor 88

4-1. Homodimer vs. heterodimer 88

4-2. X-ray crystallization studies 92

Chapter 4. Conclusions and future directions 99

Chapter 5. Experimental Method 100

5-1. Yeast passage 100

5-2. MACS 101

5-3. FACS 102

5-4. Hit sequencing 102

5-5. Yeast transformation 103

5-6. RDA peptides expression and purification 104

5-7. Heterodimer peptides expression and purification 107

5-8. Ras protein transformation, expression & purification 109

5-9. X-ray crystallography studies of the Ras/peptide complex 113

Acknowledgements

During my years in graduate school, I have met a number of people who have guided and encouraged me throughout this journey. It is hard to imagine the past six years without these people, and I would like to thank them before I begin my thesis.

First of all, I would like to sincerely thank my advisor, Prof. Gregory Verdine. I remember the first time I met him as my curriculum advisor; I was imagining an authoritative and stern professor, which was my prejudice of what professors will be like at that time, but instead I was pleasantly surprised to find myself speaking with a friendly and caring mentor. This first impression of him continued throughout the years as I rotated and joined his lab. Often times when I stumbled or failed in work, I was afraid that it might make him angry, but on the contrary he patiently waited for me to learn through my mistakes and kept encouraging me to find new solutions. Looking back it was his belief in me that helped me stay strong and motivated during my grad school years. In addition, his passion and enthusiasm toward drug discovery and science in general has led me to open my eyes more widely and think more deeply.

I thank my Graduate Advising Committee members, Prof. David Liu and Prof. Alan Saghatelian, for giving me numerous feedback and advice during the annual committee meetings. I would also like to thank Prof. Loren Walensky for joining my Dissertation Advisory Committee, and it is a great honor to me to have one of the pioneers in the stapled peptide field as a member of my committee.

I always tell the people around me about how happy I feel to be in my lab and with my labmates. I feel blessed to be in a community where people truly care about one another, and eager to help out other lab members. I would first like to thank Dr. Young-Woo Kim who was my mentor when I first joined the lab. He taught me how to conduct chemistry experiments from peptide synthesis to ring closing metathesis reaction. If Dr. Kim is my mentor in chemistry, Dr. John McGee, who I collaborated with on the Ras project, would be my mentor in biological experiments. I would like to thank them both for teaching me and guiding me through my projects. I would also like to thank Dr. Matt Lee, who first introduced me to the world of X-ray crystallography, and Lan Wang and Dr. Minyun Zhou who helped me and gave me advice when I began conducting crystallography experiments on my own and was refining and solving x-ray structures. Especially Lan was always available whenever I had a question related to X-ray crystallography. Thanks to her shared knowledge and experience, I was able to solve the structures shown in this thesis. I am grateful for Dr. Kazu Hayashi for helping me in NMR experiments to solve the double bond geometry of the 4-carbon *i,i*+3 stapled peptide isomers. My special thanks, without a doubt, go to Dr. Jerry Hilinski and Dr. Yvonne Nagel, both of whom I clearly remember meeting for the first time. Jerry had showed the kindness of ordering a new vacuum manifold when I began my rotation in the lab, so that I could have my own to use for peptide synthesis, and his role as Big Brother continued even afterward. Not only I, but also other people in the lab consult with him for various experimental or scientific questions, and he will always make time for discussions. I was especially grateful that he read the first draft of my thesis from the beginning to the end, not only giving me helpful comments but also correcting my grammatical errors and awkward English, which was really helpful as English is a second language to me. The first time I met Yvonne I could see from her eyes how excited she was to

join the lab, and I could sense from her smile that she was a good person. My thoughts were right; Yvonne is now not only my role model as a female scientist, but also a true friend, who genuinely cares about me and tries her best to help me out with difficult challenges. I also want to express my gratitude toward Lydia Carmosino and Betsy Donovan, our wonderful lab administrators, who make sure things go smoothly in lab, so that we can focus on research. Especially Lydia has helped me a number of times. For instance, she had gotten an induction cooktop and cooking tools that can be used with it for the new group room kitchen, when she learned that some lab members, including myself, often cooked in the old group room kitchen, because hotplates were no longer allowed after the renovation. This was something she did of her own accord, which clearly shows how she cares about the members of the lab. Once again, I would like to thank everyone in the Verdine Group, for both the professional and friendly lab environment.

Outside lab, there were numerous friends who I would like to thank. Although it would be hard to write all of them here, I would like to point out a few people: Nicolai Jakobsen for supporting me and helping me out, especially during the time period when I was preparing for my dissertation, so that I could focus on writing in ease, Dr. Seung-Yeon Kang, who I spent a lot of time together with during grad school years, and my ELP family and study mates, Dr. Sukeun Jeong, Dr. Jiho Choi, Seonmi Park, Dr. Hyunsung Park, Dongwoo Lee, Hyungsuk Tak, Changwook Min. Not only are they my friends, but also like a family to me.

Last but not the least, I would like to thank my mom, my dad, and my brother, and also my grandparents for their love and support. Whenever I felt hardship or difficulties, it was with their encouragement and trust in many aspects of life that I could stand up again and keep on going.

List of figures

Figure 1. Metathesis catalysts used in RCM reaction 10

Figure 2. Comparison of different catalysts for RCM reaction for S_5-S_5 peptide 11

Figure 3. Comparison of different catalysts used for S_3-S_5 and S_5-S_3 metathesis reaction, at 65 ^{o}C. 12

Figure 4. Temperature dependent RCM reaction 13

Figure 5. Catalyst dependent RCM reaction 14

Figure 6. Circular dichroism spectra analysis of the 6-carbon *i, i*+4 stapled peptide. 15

Figure 7. RCM of *i, i*+4 peptides with GI catalyst at room temperature. 17

Figure 8. RCM of *i,i*+4 peptides with GI catalyst at 65 ^{o}C 17

Figure 9. Circular dichroism of 6-carbon and 7-carbon *i, i*+4 stapled peptides 18

Figure 10. Impact of metathesis conditions on reaction yield and product distribution for 6-carbon *i,i*+3 peptides 22

Figure 11. Circular dichroism spectra 23

Figure 12. 4-carbon bridge in *i, i*+4 and *i, i*+3 formats. 26

Figure 13. Circular dichroism analysis of 4-carbon *i, i*+3 stapled peptides. 27

Figure 14. Left handed 3_{10}-helix shown in the EGF-A domain of the LDL receptor 28

Figure 15. Stabilization of left handed 3_{10}-helix. 30

Figure 16. stabilization of right handed 3_{10}-helix from a peptide composed of D-amino acids 31

Figure 17. Different solvent system used for separation of two building block isomers 34

Figure 18. NMR analysis for determining double bond geometry of 4-carbon *i, i*+3 stapled peptide 35

Figure 19. Circular dichroism of the 4-carbon *i,i*+3 stapled peptide isomers show that the cis-isomer is stabilizing the 3_{10}-helix. 36

Figure 20. GTP-bound Ras 44

Figure 21. GDP-bound Ras. 45
Figure 22. Loaded-spring mechanism of Ras. 45
Figure 23. Ras families binding to their effectors. 46
Figure 24 Yeast surface display. 51
Figure 25. aPP scaffold for yeast display. 53
Figure 26. Beta hairpin scaffold for yeast display. 54
Figure 27. PI3K for yeast display scaffold. 54
Figure 28. Scyllatoxin scaffold for yeast display. 55
Figure 29. Plasmid library for yeast surface display system 56
Figure 30. GTP and its analogue. 57
Figure 31. Negative controls in FACS sorting in yeast surface display experiment 58
Figure 32. Screening of aPP library by FACS 59
Figure 33. Screening of beta hairpin library by FACS 59
Figure 34. Screening of PI3K library by FACS 60
Figure 35. Screening of scyllatoxin library by FACS 60
Figure 36. Control experiments. 61
Figure 37. PI3K library control experiments 62
Figure 38. Scyllatoxin library control experiments 63
Figure 39. Beta hairpin library control experiments 63
Figure 40. aPP library control experiments 64
Figure 41. Control FACS experiments for the 223 library 66
Figure 42. aPP library hit sequences. 68

Figure 43. Fluorescence polarization (FP) measurements of the RDA peptides binding to K-Ras (G12V) 70

Figure 44. K-Ras (G12V) GppNHp in complex with Ral GDS and RDA peptides 71

Figure 45. His 21 of RDA1 first monomer 73

Figure 46. His21 of RDA1 second monomer. 73

Figure 47. Asn29 of the RDA1 first monomer. 74

Figure 48. Tyr 32 of the RDA1 first monomer 74

Figure 49. Trytophan residues in the RDA peptides 75

Figure 50. Ser16 in the RDA first (left) and second (right)monomer 76

Figure 51. Glu18 of the RDA1 first monomer 76

Figure 52. Glu 22 of the RDA1 first monomer 77

Figure 53. Arg 26 of the RDA1 first monomer 77

Figure 54. Tyr 23 and Cys 10 residues in the RDA1 peptides 78

Figure 55. Ala33Arg mutation in RDA2 79

Figure 56. Wild type Ras in different states 80

Figure 57. RDA bound Ras aligned with either state 2 Ras or state 1 Ras 81

Figure 58. RDA bound Ras. 82

Figure 59. FP measurement of 226-1 peptides binding to KRas G12V D38P double mutant. 83

Figure 60. RDAI bound Ras G12V D38P double mutant compared with RDAI bound Ras G12V mutant. 84

Figure 61. Stitched RDA peptides. 85

Figure 62. Stitched RDA peptides. 87

Figure 63. SDS-PAGE of 24G dimer, Ras, and the Ras/24G complex 92

Figure 64. Crystallization conditions for 24G bound K-Ras (G12V) GppNHp complex. 94

Figure 65. Analysis of the 24G/K-Ras complex.. 95

Figure 66. Comparison of the 24G/K-Ras and 32P/K-Ras crystals in the same condition..... 96

Figure 67. Heteroseeding. 97

Figure 68. Comparison of 32P (yellow) and 24G (gray) peptide..... 99

Figure 69. FPLC for buffer exchange of the peptides for TEV cleavage. 106

Figure 70. After purification of the RDA peptides by HPLC. 106

Figure 71. After HPLC purification of the heterodimer..... 109

Figure 72. SDS-PAGE analysis of Ras TEV cleavage. 111

Figure 73 Elution time for GDP (1st column), GMPPNP (2nd column) and nucleotide loaded on Ras (3rd column). 112

List of Tables

Table 1. Shorter linker *i, i*+4 stapled peptide sequences. 9

Table 2. 6-carbon & 7-carbon *i, i*+4 peptides 16

Table 3. Analysis of the circular dichroism results for the 6-carbon and 7-carbon *i, i*+4 stapled peptides 19

Table 4. Shorter linker *i, i*+3 peptides 20

Table 5. Percent alpha-helicity of the peptides at 20 ^{o}C 24

Table 6. 4-carbon *i, i*+3 peptide sequences 26

Table 7. 4-carbon *i, i*+3 stapled peptide sequences for left-handed 3_{10}-helix stabilization 29

Table 8. Part of the EGF-A domain from LDL receptor sequence for left handed 3_{10}-helix stabilization. 32

Table 9. Frequency of Ras isoform mutations in selected human cancers. 48

Table 10. Residues in RDA peptides that are important in binding to Ras and/or dimer stabilization 79

Table 11. Stitched RDA peptides. 86

Table 12. Initial template design for the heterodimeric RDA peptides. 89

Table 13. Library design to achieve improvement in affinity of the Ras binding peptides 90

Table 14. Library design to achieve improvement in cell-permeability of the Ras binding peptides 90

Table 15. 24G peptide sequence. 91

Table 16. Library design and the resulting 32P peptide sequence 92

Table 17. Residue interactions in 32P and 24G peptides 98

A step toward achieving peptide based therapeutics

Traditionally, small molecules have been used heavily as drug modalities. They can be optimized to pass through membranes, so there isn't much geographic limitation. However, with their small sizes, they have limited surface area. According to Lipinski's rule of five for attaining oral bioavailability, the general guideline of the molecular weight cut off for small molecules has been limited to 500. Unfortunately, as molecular weight and therefore compound size is restricted, the pool of the possible targets also becomes smaller. Empirically small molecules work by binding to hydrophobic pockets of the targets, which confer high affinity and specificity. Therefore the use of small molecules as therapeutics is limited to target proteins that can accommodate them by having hydrophobic pockets. To make the problems worse, hydrophobic pockets that are already in the proteins are there for a reason; they are already optimized for binding to cofactors or substrates against which small molecules would have to compete. In many cases, protein-protein interactions (PPIs) have large contact area with an extensive interface. However with the area being a mostly flat surface, there is no hydrophobic pocket that a small molecule can target. On the other hand, for biologics, they have unlimited targeting power regardless of the protein surface. However, such large molecules are limited in that they are not able to penetrate cell membranes; they could only be used to target extracellular proteins.

It has been estimated that small molecules can be used for only 10 % of prospective target proteins, and biologics for another 10 %[1] . Since there might be some overlap between those two, more than 80% of targets are thought to be "undruggable". This calls for a new therapeutic modality that can overcome the limitations of small molecules and biologics while retaining their

benefits. This is why the idea of a cell-penetrating 'mini-protein' is appealing. In a given PPI, there is generally an interaction domain in the protein that is crucial for the binding, while the remaining part of the protein is there to support the overall structure. Therefore, the idea here is to mimic the hot-spot binding region to disrupt the target PPI. However, in reality it is not as simple as just stripping down the protein to its business end. When the peptide is taken out from its context, it tends to lose its structure since there is an entropy cost of going from disordered to ordered structure. The problems coming from peptides being unfolded are that they show low binding affinities to targets, get rapidly degraded by proteases due to their exposed amide bonds, and have poor cell permeability due in part to the partial charges that derive from amide bonds. Therefore, there is a need to constrain the conformation so that the peptide can retain its active structure.

In order to preserve the recognition ability of the secondary structure by retaining its ordered structure, macrocyclic ring formation has long been established as a particularly effective strategy for the conformational reinforcement of peptides [2,3] [4] . Many types of bond-forming reactions have been used to close macrocyclic rings in peptides, but the powerful ring-forming ability of ring closing metathesis (RCM) [5] has proven particularly useful in generating conformationally stabilized peptides. The olefin metathesis here refers to metal-catalyzed redistribution of carbon-carbon double bonds [6] . Following the groundbreaking work of Miller and Grubbs [7] on the conformational stabilization of beta-turn structures via RCM, this reaction has been employed in ring-formation processes leading to a variety of enforced peptide conformations [8–12] . Peptide macrocyclization is among the most effective strategies for promoting the formation of helical structures in peptides [13] ; here again, RCM-based systems

have demonstrated value in stabilizing both 3_{10}-[8,9] and alpha-helical[11] conformations. Although both these helical forms are well represented in protein structures, alpha-helices are considerably more common and of much greater average length than 3_{10}-helices. Considering that alpha-helices are more prevalent than any other single motif in mediating protein-macromolecule interactions[13], this particular helical form has attracted recent attention as one potential avenue toward generating next-generation therapeutics. Thus, there have been numerous attempts in trying to stabilize the α-helix, such as employing salt bridges[14], lactams[15] and disulfide bridges[16].

Of the dozen or so systems reported to stabilize alpha-helical structures in peptides, one has gained particular interest owing to its superior helix-stabilizing power and unprecedented ability to confer robust cell penetration upon the peptide bearing it. This system, the all-hydrocarbon 'staple'[10,17], is the first to combine the helix-stabilizing effects of macrocyclization of the peptides[8], with the equally powerful effects of alpha-methylation, first described by Karle et al.[18].

One of the most widely employed staples, $S_{i,i+4}S(8)$, is introduced by incorporation of two units of the alpha-methyl, alpha-alkenyl amino acid S_5 (The unnatural amino acids are denoted in the form of X_y, where X indicates the stereochemistry of the alpha-carbon and Y the number of carbons in the alkene chain) into a peptide at the relative positions of *i* and *i*+4, followed by RCM to close an 8-carbon all-hydrocarbon macrocyclic ring spanning one turn of a helix.

In a spate of published studies, this and a closely related *i,i*+7 system, which covers two turns of an alpha-helical backbone, have been shown to provide significant improvements in alpha-helical character, target-binding affinity, proteolytic resistance, and serum half-life. Most distinctively, stapling has been found frequently to confer robust cell penetration by an active mechanism involving endocytic vesicle formation and trafficking [19–22] . This hydrocarbon stapling system has been successfully deployed by our laboratory and others against a variety of biological targets and stands apart from all other helix-stabilization systems as the sole example that has been shown effective in animal models of a human disease driven by an intracellular target [23–28] .

The ability of stapled peptides to access intracellular targets that engage in protein-protein interactions, targets typically considered intractable, underlies the tremendous interest in these agents as potential next-generation therapeutics having a distinctive and uniquely powerful targeting range. The first part of my thesis focuses on the extension of these efforts by exploring staples having shorter linker lengths and/or alternative positions of the olefin along the hydrocarbon cross-link. The content from the chapter within this part that describes the 6-carbon *i, i*+3 stapled peptides has been published : Shim, S. Y., Kim, Y.-W. W. & Verdine, G. L. A new *i, i*+3 peptide stapling system for α-helix stabilization. *Chem Biol Drug Des* **82,** 635–42 (2013).

High-resolution structures of PPIs are important in employing and rationally designing stapled peptides. However, there are cases where there are no obvious binding partners that have an alpha-helical component that could be used as a starting point for this approach. For example, the oncogenic K-Ras protein does not have any known alpha-helical interactor; a crystal structure of

the binding face of Ras and Raf shows that they are interacting via beta sheet complementation. In such cases, we cannot rationally design a helical structure as an inhibitor. Instead, we have approached the problem by first screening a yeast cell surface display library of a stabilized scaffold using directed evolution, and then improving the affinity of the lead peptide further by rational design based on the structural analysis, which will be described in the second part of my thesis. This project was collaborated with Dr. John McGee and Dr. Matt Lee; Dr. John McGee designed the libraries for yeast display system and did screening for the heterodimer peptide, as well as most biochemical experiments to determine the binding of the hit peptides to Ras. With the hit peptides I expressed from the bacterial expression, Dr. Matt Lee did crystallization experiment and solved the structures for the RDA 1 peptides bound Ras as well as RDA 2 peptides bound Ras.

Part I. Analysis of the effect of olefin position and linker length in hydrocarbon stapled peptides

Chapter 1. Introduction

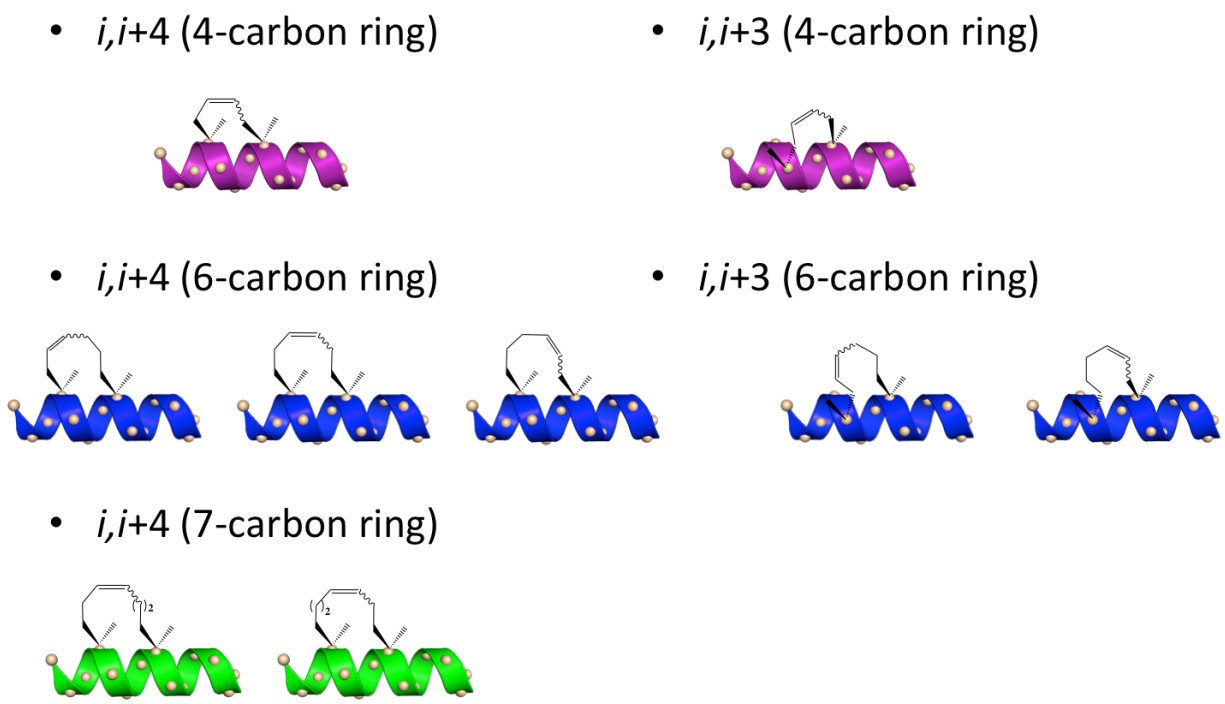
It has been determined in the published studies from our lab that the extent of alpha-helix stabilization and nucleation by staples that encompass a single helical turn (*i, i*+4 and *i, i*+3) and that possess an 8-carbon macrocyclic bridge (i.e., formed using a combination of S_5 and R_5) is highly dependent both on the stereochemistry of the alpha-methyl, alpha-alkenyl amino acids used to form the staple and also on the position of the staple within a given peptide sequence. Whereas two units of S_5 gave the most helix-stabilizing staple in the *i, i*+4 configuration ($S_{i,i+4}S(8)$ staple), the *i, i*+3 configuration was most optimal with an N-terminal R_5 residue and a C-terminal S_5 residue ($R_{i,i+3}S(8)$ staple) [29,30] .

Although the conventional 8-atom *i, i*+4 system continues to enjoy considerable success and scope of applicability [17,23,24,28,31] , we have occasionally observed troublesome aggregation in peptides bearing this staple and have reasoned that this problem might be mitigated by diminishing the hydrocarbon content of the macrocyclic bridge. Additional motivation for a broader investigation of variant hydrocarbon staples comes from several recent studies in which the hydrocarbon portion of the staple has been found to engage the target directly through hydrocarbon contacts with residues lining the peptide-binding pocket of the target [32,33] [27] . These findings have suggested that in certain instances, the macrocyclic bridge should itself be the subject of structure-activity investigations to maximize target-binding affinity.

In the *i, i*+4 system, it has been previously shown that contraction of the hydrocarbon bridge from 8 to 6 atoms leads to a complete failure of RCM, such that helicity in the product could never be assessed [10] . However, we decided to revisit and further investigate the effects on alpha-helicity of having short linkers of 6 to 7 carbons in the *i, i*+4 system, this time trying high temperature and/or different catalysts to see if that would help overcome the energy barrier for RCM and yield stapled products. This effort is described in chapter 2, where we learned not only that it is possible to yield stapled products with the 7-carbon bridge, but also that it retains alpha helicity comparable to that of the 8-carbon bridge.

In the case of the *i, i*+3 staple configuration, preliminary modeling studies suggested that a 6-atom hydrocarbon bridge bearing a cis-olefin could be accommodated without significant disruption of the helix. (P.S. Kutchukian and G.L. Verdine, unpublished results). The studies in chapter 3 were therefore undertaken to explore the synthetic feasibility and biophysical effects of contracting the hydrocarbon bridge in *i, i*+3 staple peptides from 8 atoms to 6. In chapter 4, we describe studies to reduce the length of the bridge to 4 atoms, which was found to stabilize another secondary structure, the 3_{10}-helix.

The overall scheme of the shorter linker stapled peptides examined in this thesis is shown below.



Scheme 1. The *i, i*+3 and *i, i*+4 shorter linker hydrocarbon stapled peptides examined in this study.

Chapter 2. Shorter hydrocarbon bridge in *i,i*+4 stapled peptides

It has been shown that for *i, i*+4 stapled peptides, an 8-atom hydrocarbon bridge results in the most alpha-helix stabilization. When the hydrocarbon linker was contracted down to a 6-atom bridge, it was reported that the metathesis reaction failed [10] . Here, we decided to revisit the possibility of obtaining a shorter hydrocarbon bridge for the *i, i*+4 stapling system that still stabilizes alpha-helicity, in the hope of adding another possible linker to our stapling system tool box.

Table 1. Shorter linker *i, i*+4 stapled peptide sequences. The sequences are N-terminally acetylated and C-terminally amidated.

RNaseA peptide	Substrate Sequence														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	E	W	A	E	T	A	A	A	K	F	L	A	A	H	A
B	E	W	A	E	T	A	A	S5	K	F	L	S5	A	H	A
C	E	W	A	E	T	A	A	S3	K	F	L	S3	A	H	A
D	E	W	A	E	T	A	A	S3	K	F	L	S5	A	H	A
E	E	W	A	E	T	A	A	S5	K	F	L	S3	A	H	A

We utilized the C-peptide sequence of RNaseA [34] for the purpose of direct comparison with previous studies, since how readily the stapling system works and the resulting helical content is sequence dependent due to factors such as the tendency of the original sequence to form an alpha-helix [22] **(Table 1)**. The RNaseA sequence has been used in a number of studies due to its partial helicity in water, which enables us to observe either an increase or decrease in helicity that comes from its modification. Along with the wild type sequence, either S_3 or S_5 unnatural amino acids were incorporated in *i, i*+4 phasing to make all possible pairs (S_3-S_3, S_3-S_5, S_5-S_3 or S_5-S_5). Although the typical reaction conditions from the literature would be to use Grubbs first generation catalyst (GI catalyst) at room temperature [19], in order to maximize conversion to the stapled products we also tested elevated temperatures up to 65 ^{o}C and different catalysts, namely Grubbs second generation catalyst (GII catalyst), Grubbs third generation catalyst (GIII catalyst) and Hoveyda-Grubbs second generation catalyst (HGII catalyst) at room temperature (**Figure 1**).

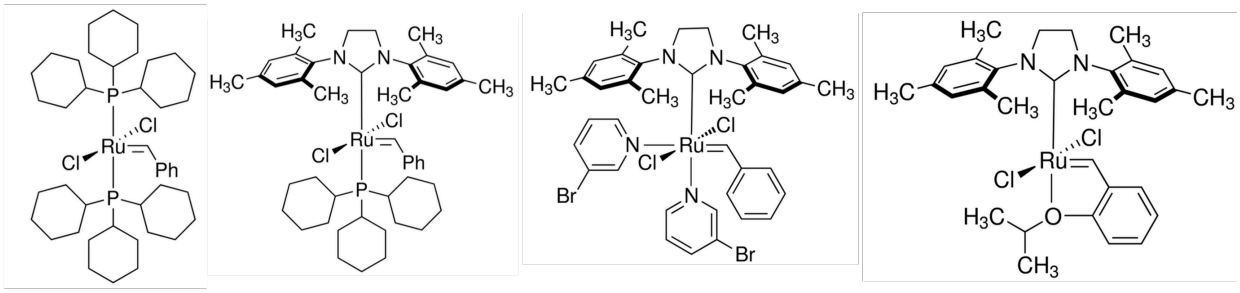


Figure 1. Metathesis catalysts used in RCM reaction

Such ruthenium catalysts are known to be useful for olefin metathesis reaction due to their stability towards air and moisture as well as compatibility with broad range of functional groups [35] . Unlike the GI catalyst where both neutral ligands are phosphines, the GII have one of the ligands replaced with a N-heterocyclic carbene with a saturated backbone. It was shown to have improved catalytic activity than GI catalyst [36] , although the initiation rate is lower [37] . HGII catalyst that has substitution of the second phosphine ligand for a bidentate alkylidene, has higher thermal stability and tolerance against oxygen and moisture compared to its phosphine-containing analogue GII, albeit with decreased initiation rate [38] .

First we tested ring closing metathesis reaction for the S_5- S_5 pair **B** with the different catalysts. The reactions were monitored by LC/MS after cleaving the peptides off the resin. As shown in the graph in **Figure 2**, the GI catalyst showed the best RCM performance, showing almost 90% product conversion after 2 hours, while the other catalysts barely achieved 50% product conversion after 4 hours. Therefore, for the conventional stapling system, GI catalyst was confirmed as the best choice for the RCM reaction.

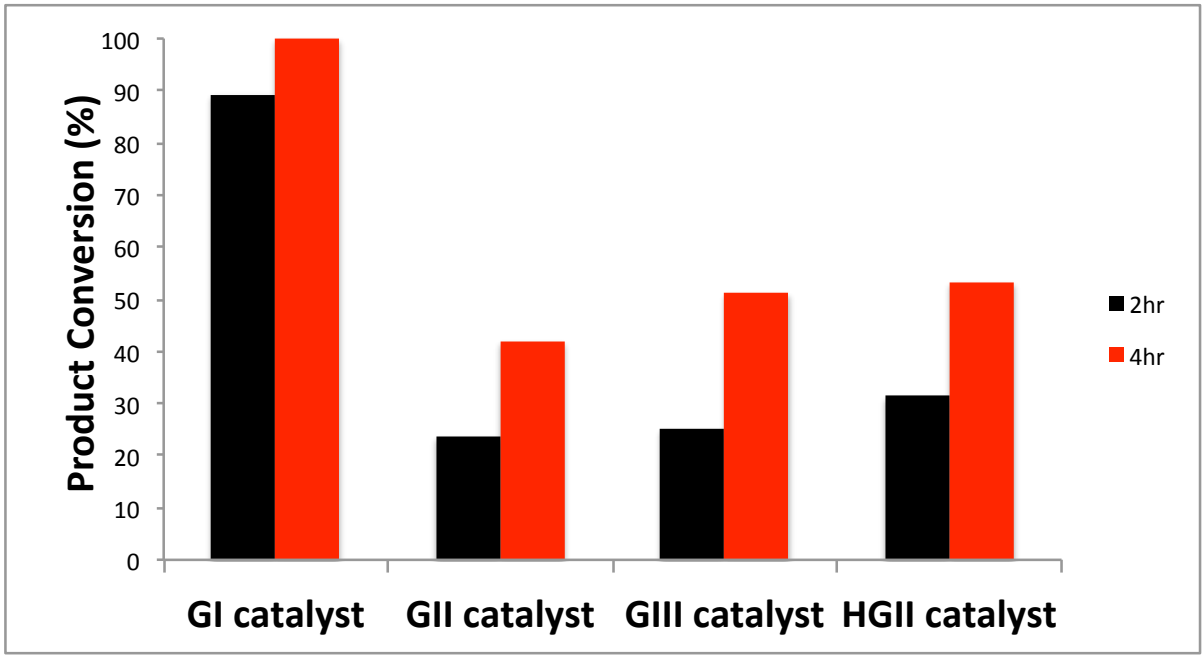


Figure 2. Comparison of different catalysts for RCM reaction for S_5-S_5 peptide. Reaction condition: 30mole % of each catalyst, room temperature, 2 hr for two times

For the S_3-S_3 pair **C**, no RCM reaction was observed under any conditions. This result is probably because the 4-carbon linker is too short to connect the two positions in an *i, i*+4 format. For the S_5-S_3 pair **E**, no reaction occurred at room temperature, and even at higher temperature, only with HGII catalyst did the RCM reaction proceed at all (~9% product conversion). Interestingly, while **E** exhibited a very sluggish reaction, the S_3-S_5 pair **D** formed stapled product readily when certain catalysts were used at elevated temperature. **(Figure 3)**

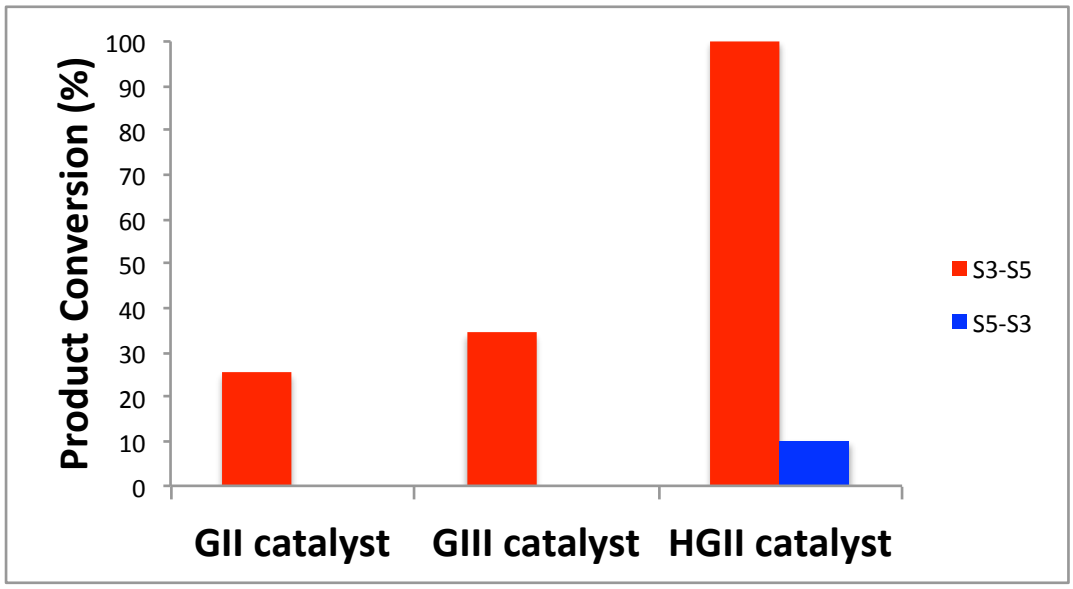


Figure 3. Comparison of different catalysts used for S_3-S_5 and S_5-S_3 metathesis reaction, at 65 ^{o}C.

When GI catalyst was used at room temperature, no product was seen after 2 hours of RCM reaction, and only about 14 % product conversion was shown after 4 hours of RCM reaction. However, when the temperature was elevated to 65^{o}C, ~34 % product conversion was seen after 2 hours (**Figure 4**), and the reaction could be pushed to around 80 % after a total of 6 hours of RCM reaction.

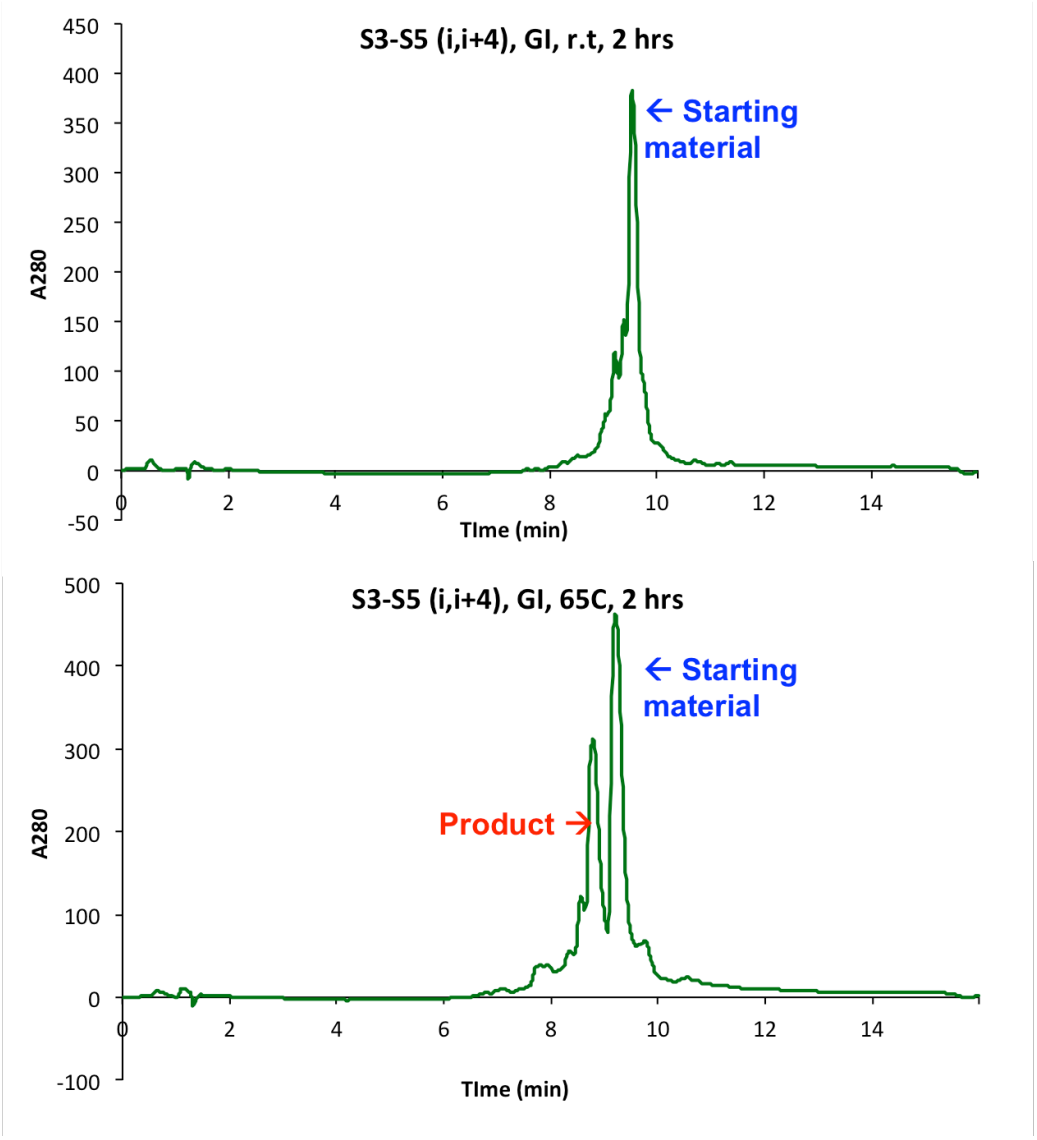


Figure 4. Temperature dependent RCM reaction. Although no stapled product is seen after 2 hours of RCM reaction for S_3-S_5 pair (*i, i*+4) at room temperature, with elevated temperature ~34% product conversion is observed.

In the same pattern, when GII, GIII or HGII catalysts were used at room temperature, the reaction did not proceed for **D**. However, at high temperature, product conversions of 25%, 34%, and >95%, respectively, were shown after 2x2 hours (**Figure 5**). This clearly shows that the HGII catalyst promotes the reaction better than the other catalysts when used at the higher temperature of 65 ^{o}C.

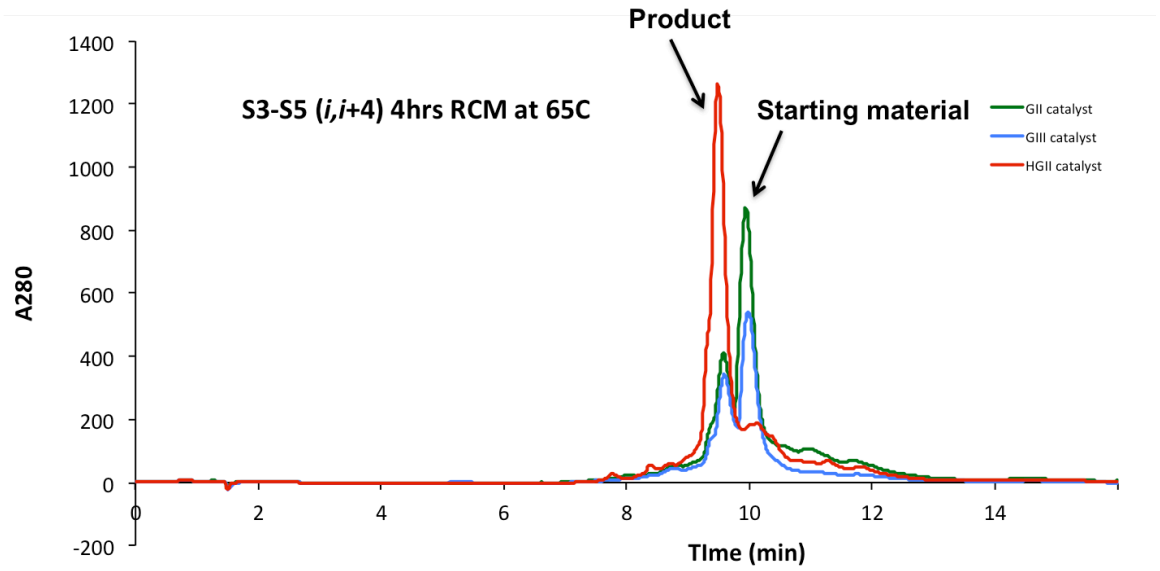


Figure 5. Catalyst dependent RCM reaction. In the RCM reaction for S_3-S_5 pair(*i, i*+4) at 65 ^{o}C, HGII catalyst gives the best yield among the catalysts.

The difference between the two sequences **D** and **E** only comes from the swap of the S_3 and S_5 in the *i, i*+4 positions, which would result in a different olefin position within the hydrocarbon bridge in the stapled product; the stereochemistry and total length of the bridge is the same. Therefore, it indicates that the regiochemistry of the linker is important in terms of the RCM reaction.

In order to assess the average alpha-helical content of the peptides, circular dichroism analysis was performed.

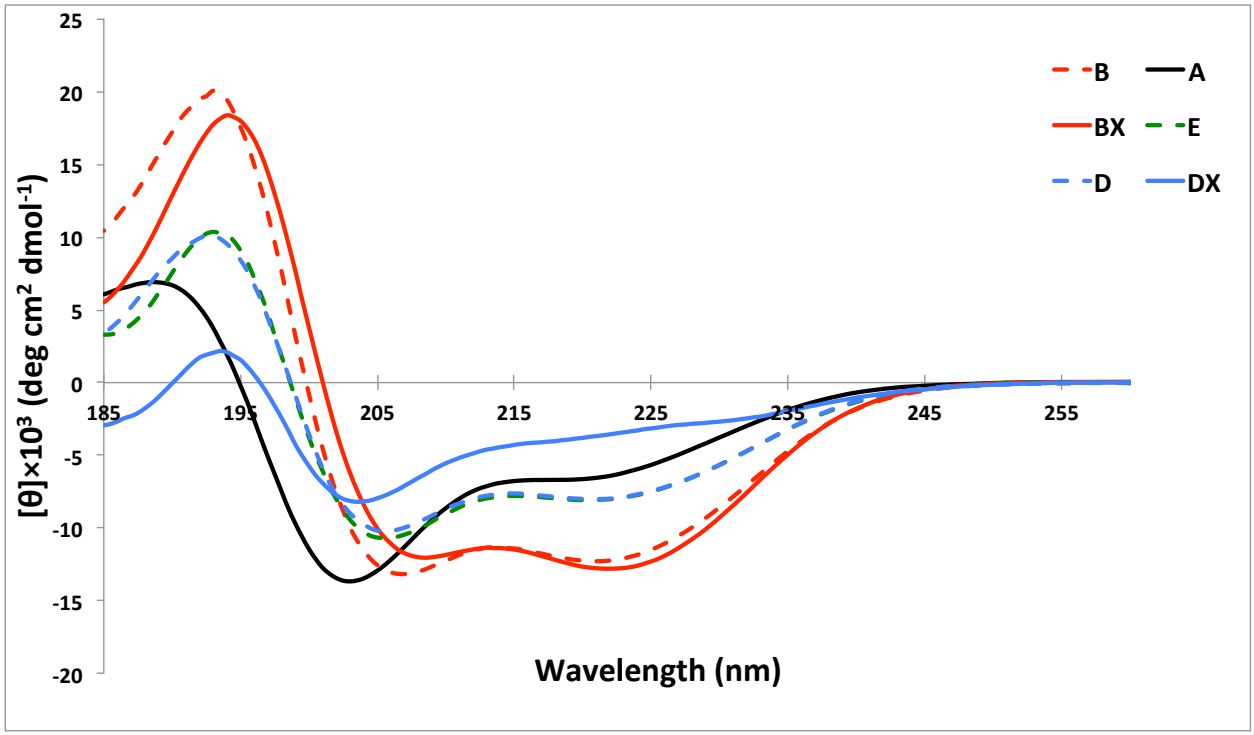


Figure 6. Circular dichroism spectra analysis of the 6-carbon *i, i*+4 stapled peptide.

As shown in **Figure 6**, the wild type RNaseA sequence **A** exists as a random coil, but when alpha-methyl, alpha-pentenyl amino acids are incorporated into the helix in an *i, i*+4 phase (**B**) it becomes alpha helical, and more so after metathesis to produce the macrocyclic bridge (**BX**) (all stapled products are denoted by **X**). When the alpha-methyl, alpha-propenyl amino acid is introduced into the *i* or *i*+4 position instead (**D** & **E**), helicity is enhanced compared to the wild type sequence but the stabilization effect is diminished to a high degree compared to the S_5-S_5 pair **B**. Interestingly, this stabilization effect decreases further when the macrocyclic bridge is formed. The S_3-S_5 stapled peptide pair **DX** shows a weaker negative maximum around 203-205 nm. This overall destabillization might come from the fact that the linker length is too short for the peptide to retain a normal alpha-helical conformation, and such strain might be reversing the helix stabilization effect that came from the introduction of the alpha-methyl, alpha-alkenyl amino acids as shown in **D** & **E**.

The **DX** and **EX** peptides differ only in the position of the olefin within the hydrocarbon bridge, closer to N-terminus or C-terminus of the peptide, respectively. To further examine the effect of the location of the olefin, we decided to look at the S_4-S_4 pair, which would have the olefin in the middle of the hydrocarbon bridge. In addition we prepared peptides that contained S_4-S_5 and S_5-S_4 and would form a 7-atom hydrocarbon bridge to examine their RCM kinetics as well as the helicity of the resulting peptides.

Table 2. 6-carbon & 7-carbon *i, i*+4 peptides

RNaseA peptide	Substrate Sequence														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
D	E	W	A	E	T	A	A	S3	K	F	L	S5	A	H	A
F	E	W	A	E	T	A	A	S4	K	F	L	S4	A	H	A
G	E	W	A	E	T	A	A	S4	K	F	L	S5	A	H	A
H	E	W	A	E	T	A	A	S5	K	F	L	S4	A	H	A

As shown in **Figure 7**, when the GI catalyst was used at room temperature, the S_4-S_4 pair **F** showed much faster RCM reaction than its S_3-S_5 counterpart **D**. For both of these peptides, the reaction seems to increase when the temperature was increased to 65 ^{o}C. As for the S_4- S_5 and S_5-S_4 pairs, as seen for the S_3- S_5 vs. S_5- S_3 pairs, again we observed a prominent difference in the RCM kinetics depending on the position of the olefin. The substrate that has the olefin closer to the N-terminus, **G**, shows over 50% product conversion after 2 hours, and completes within 6 hours. On the other hand, the substrate **H** that has the olefin closer to the C-terminus only shows ~42% product conversion even after 6 hours of reaction.

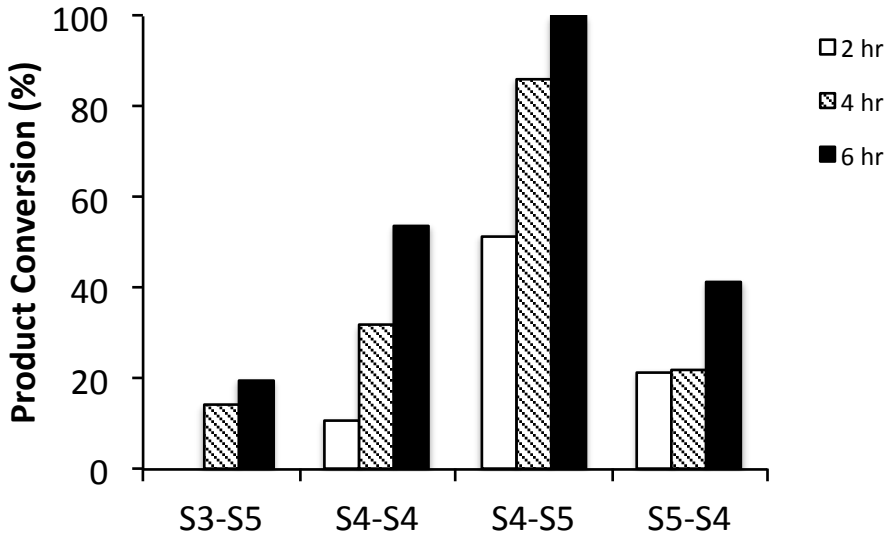


Figure 7. RCM of *i, i*+4 peptides with GI catalyst at room temperature.

The rates of metathesis for both **G** and **H** increase when high temperature is used, and even **H** reaches over 95% product conversion within 6 hours. (**Figure 8**)

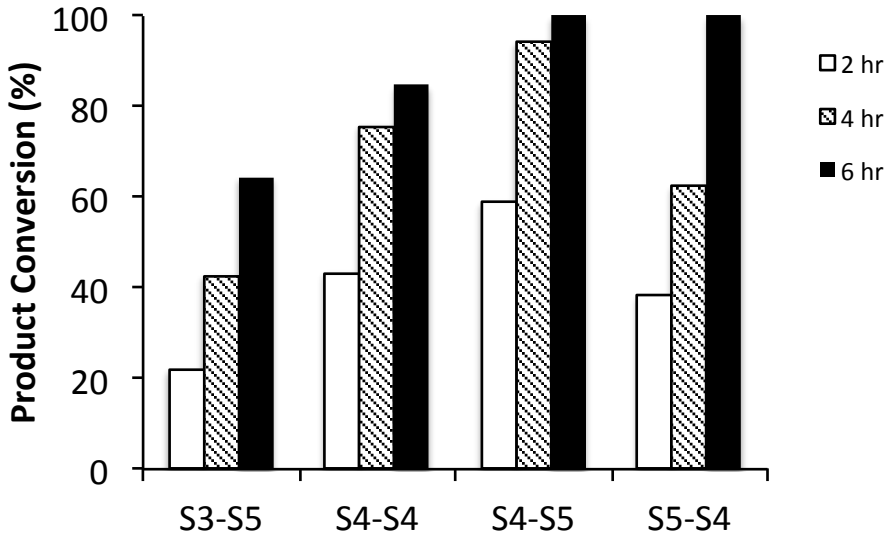


Figure 8. RCM of *i, i*+4 peptides with GI catalyst at 65 ^{o}C

The trend of having different properties attributable to the olefin position in the hydrocarbon bridge continues in terms of the alpha helical content. In **Figure 9**, we see spectra ranging from a

random coil for unmodified peptide **A** to a nearly ideal alpha helical profile for the S_5-S_5 stapled peptide **BX**. Previously we observed that the S_3-S_5 stapled product **DX** does not adopt an alpha helical structure, but in contrast, **FX**, which differs from **DX** only in the position of the olefin, is much more alpha helical. By employing a slightly longer linker as in **GX** and **HX**, the helicity increases even further. Again, compared to S_5-S_4 stapled product **HX**, the S_4-S_5 stapled product **GX** shows stronger alpha helicity, which approaches the helicity of the conventional *i, i*+4 stapled peptide **BX**. (**Table 3**) The effect of regiochemical variation in the 8-carbon system has been studied in the recent paper from Kim group [39] , where it compared S_3-S_7 and S_7-S_3 pairs with regard to macrocyclic ring closure and alpha-helicity. There again, it showed that while one pair, S_3-S_7, readily undergoes RCM at higher temperature with the product showing substantial alpha helicity (49%), the S_7-S_3 pair does not form stapled product. These results emphasize the importance of screening all these attributes with any new stapling system.

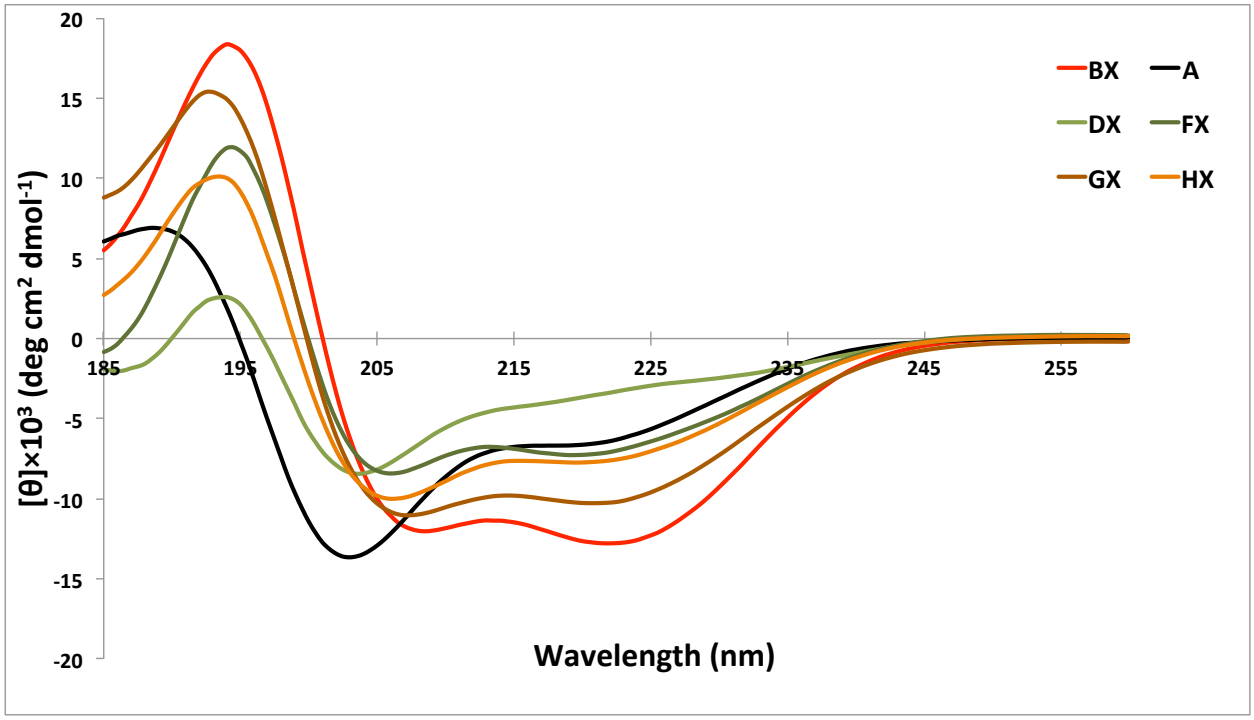


Figure 9. Circular dichroism of 6-carbon and 7-carbon *i, i*+4 stapled peptides

Table 3. Analysis of the circular dichroism results for the 6-carbon and 7-carbon *i, i*+4 stapled peptides.

Peptide Sequence	π→π* wavelength (nm)	π→π* magnitude	n→π* magnitude	R value	alpha-helicity
WT (A)	203.0	-13.65	-6.41	0.47	24.6
S3-S5 (DX)	203.5	-8.14	-3.56	0.44	13.6
S4-S4 (FX)	206.0	-8.42	-7.09	0.84	27.2
S4-S5 (GX)	207.0	-11.03	-10.24	0.93	39.2
S5-S4 (HX)	206.0	-10.01	-7.61	0.76	29.2
S5-S5 (BX)	208.5	-12.03	-12.78	1.06	49

Note that as shown in the previous study by Shafmeister *et al.* [10] , the 8-carbon *i, i*+4 stapled peptide **BX** here does not seem to give a very high percent alpha-helicity. However, this might be sequence and position dependent, since when the staple position was changed from **B** (position 8 and 12) to **R** (position 4 and 8), we see the percent helicity for it increased to ~69 % as will be shown in the next chapter. Although we have used the same stapling position as in the previous study for the sake of direct comparison, it would be worthwhile to repeat the experiment in a different position of the same sequence or in a different peptide sequence to see if the pattern stays the same. This would address the feasibility of incorporating the S_4- S_5 pair for stapling peptides instead of S_5-S_5 if desired.

Chapter 3. 6-carbon *i,i*+3 stapled peptides

In parallel to the investigation of 6-carbon and 7-carbon *i, i*+4 stapled peptides in the previous chapter, we investigated the feasibility of forming a 6-carbon *i, i*+3 staple by screening all possible combinations of stereochemical configurations at the point of staple attachment to the peptide main chain. The hydrocarbon bridge was contracted from 8 to 6 atoms in an asymmetric fashion by shortening one tether by two methylenes while leaving the other unchanged. This

allowed us to analyze simultaneously the effect of translocating the double bond within the hydrocarbon bridge of the stapled product. Thus, we employed the four unnatural amino acids, alpha-methyl, alpha-pentenylglycine (R_5 or S_5) and alpha-methyl, alpha-allylglycine (R_3 or S_3), to construct a panel of peptide substrates, based on the RNase A model sequence, that sampled all possible permutations that would furnish a 6-carbon macrocyclic bridge anchored at positions 4 and 7. To facilitate comparison with known stapling systems, we also prepared the corresponding unmodified peptides and the modified but unmetathesized peptides as well as the $S_{i,i+4}S(8)$- stapled and $R_{i,i+3}S(8)$-stapled versions. (**Table 4**)

Table 4. Shorter linker *i, i*+3 peptides

RNaseA peptide	Substrate Sequence														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	E	W	A	E	T	A	A	A	K	F	L	A	A	H	A
I	E	W	A	R3	T	A	S5	A	K	F	L	A	A	H	A
J	E	W	A	R5	T	A	S3	A	K	F	L	A	A	H	A
K	E	W	A	R3	T	A	R5	A	K	F	L	A	A	H	A
L	E	W	A	R5	T	A	R3	A	K	F	L	A	A	H	A
M	E	W	A	S3	T	A	R5	A	K	F	L	A	A	H	A
N	E	W	A	S5	T	A	R3	A	K	F	L	A	A	H	A
O	E	W	A	S3	T	A	S5	A	K	F	L	A	A	H	A
P	E	W	A	S5	T	A	S3	A	K	F	L	A	A	H	A
Q	E	W	A	S5	T	A	A	S5	K	F	L	A	A	H	A
R	E	W	A	R5	T	A	S5	A	K	F	L	A	A	H	A

The majority of the permuted peptides (**K**-**P**) failed to provide any appreciable amount of stapled product in our initial screen; these were not pursued further. However, two of the substrates, namely **I** and **J**, did undergo metathesis to an appreciable extent, yielding products **IX** and **JX**.

Substrates **I** and **J** possess the same stereochemical configuration (R at position-4 and S at position-7) but they swap olefinic tethers between a 3- and 5- carbon version, such that the metathesis products **IX** and **JX** differ from each other only in the position of the double bond within the macrocyclic hydrocarbon bridge. Despite the seeming subtlety of the structural differences in **I** and **J** and their respective metathesis products, the RCM reactions with these substrates showed marked differences (**Figure 10**). Specifically, RCM on substrate **I**, having R_3 at position *i* and S_5 at *i*+3, provided an eightfold higher yield of product than with substrate **J**, which has R_5 at position *i* and S3 at *i*+3. After the RCM reaction was allowed to proceed for an additional 2 hours using fresh catalyst, the difference became even more obvious. In an attempt to promote greater product formation in RCM, especially with substrate **J**, we next explored the use of the G-II catalyst and HGII catalyst. We additionally examined the effect of raising the reaction temperature from 25 ^{o}C to 65 ^{o}C with all three catalysts. At 25 ^{o}C, both GII and HGII proved to be less efficient than GI for RCM of both substrates **I** and **J**. When the RCM reactions were performed at 65 ^{o}C, however, significantly higher conversion was observed using all three types of catalyst. Interestingly, HGII, which was nearly inactive at 25^{o}C, exhibited higher percent conversions at 65 ^{o}C (92% and 71% for **I** and **J**, respectively, after 2 hours of RCM reaction) than G-I (87% and 43% for **I** and **J**, respectively, after 2 hours of RCM reaction). These results are reminiscent of those reported by Arora and co-workers in their hydrogen bond surrogate alpha-helix-stabilization system [40] .

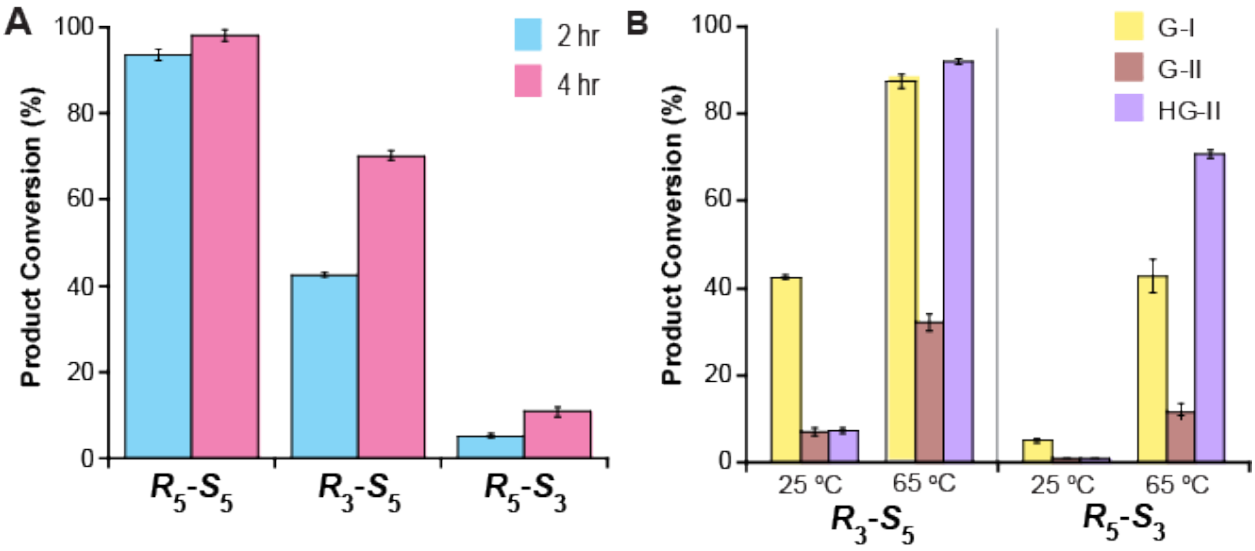


Figure 10. Impact of metathesis conditions on reaction yield and product distribution for 6-carbon *i,i*+3 peptides. Adapted from shim *et al*.[41]

Ring closing metathesis of substrate **I** actually yielded two cross-linked products (**IX1** and **IX2**) in relative ratios that depended on the particular metathesis catalyst employed. GI showed a strong preference to form the later-eluting product (ratio of **IX1** to **IX2** =1/5.8) whereas GII and HGII gave a less pronounced inverse selectivity (ratio of **IX1** to **IX2**=1/0.8-0.9). Transfer hydrogenation of stapled peptides **IX1** and **IX2** yields a single-reduced product, **IX3**, as judged by LC/MS analysis, consistent with the notion that **IX1** and **IX2** are double-bond isomers. In contrast to the situation described above for substrate **I**, the ratio of olefin isomers in the stapled products formed with substrate **J** was not strongly influenced by the catalyst (ratio of **JX1** to **JX2**=1/0.6-0.9) Again, transfer hydrogenation of **JX1** and **JX2** provided a single product, **JX3**. Reduced products **IX3** and **JX3** had indistinguishable retention times on HPLC, consistent with the expectation of these being the same molecule. This is furthermore consistent with the notion that **IX1**, **IX2**, **JX1** and **JX2** differ only in the location and/or configuration of the double bond within the hydrocarbon bridge; this difference is extinguished upon reduction of the double bond.

Incorporation of an $R_{i,i+3}S(8)$ staple into a peptide has been shown to enhance alpha-helical content as efficiently as the $S_{i,i+4}S(8)$ counterpart [30]. To determine the conformational preferences of the peptides constrained by the $R_{i,i+3}S(6)$ staple, far ultraviolet circular dichroism (CD) spectra and the percent alpha-helicity of the peptides were measured (**Figure 11, Table 5**).

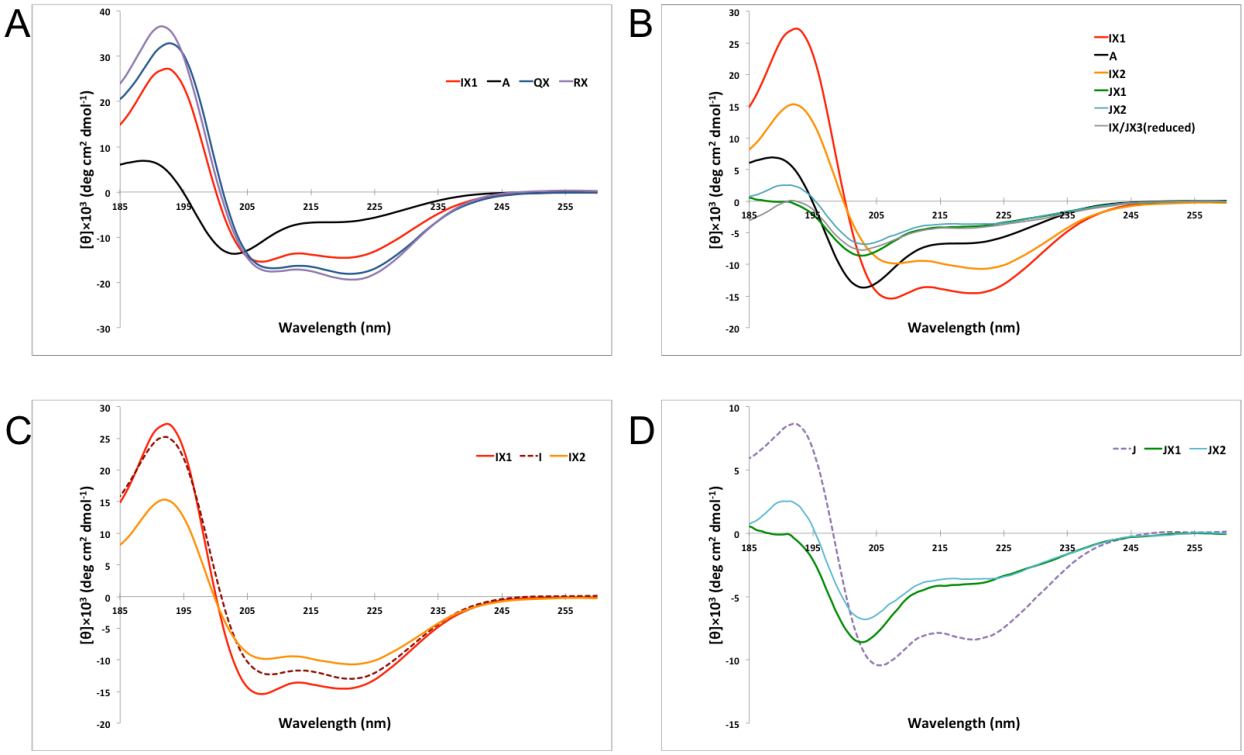


Figure 11. Circular dichroism spectra. A: CD spectra of the $R_{i,i+3}S(6)$ stapled peptide having the greatest extent of alpha-helix induction and comparison with its unmodified counterpart and the previously reported versions having an 8-methylene bridge, $S_{i,i+4}S(8)$ $R_{i,i+3}S(8)$. B: CD spectra of $R_{i,i+3}S(6)$-stapled peptides in comparison with their unmodified and reduced counterparts. C: CD spectra of $R_{i,i+3}S(6)$ stapled peptides **IX1** and **IX2** in comparison with their unmetathesized counterpart **I**. D: CD spectra of $R_{i,i+3}S(6)$ stapled peptides **JX1** and **JX2** in comparison with their unmetathesized counterpart **J**. (adapted from Shim *et al.*[41])

Table 5. Percent alpha-helicity of the peptides at 20 °C .

Peptide	% alpha-helicity
I	50
IX1	55
IX2	41
J	32
JX1	15
JX2	14
IX/JX3	16
A	25
QX	69
RX	74

IX1, the major stapled product of **I**, exhibited a respectable increase in alpha-helical content (55%) as compared to the corresponding unmodified peptide **A** (25%) The stapled peptide bearing an isomeric olefin, **IX2**, showed only a modest staple-dependent improvement of helicity (41%). Unexpectedly, although respective members of the **IX** and **JX** series differ only in the position of the olefin within the 6-carbon macrocyclic chain, the two displayed dramatically different alpha-helix-induction properties, with the **IX** series being far superior. As the presence of the alpha-methyl groups is intrinsically helix stabilizing [18] , the effects of olefin geometry and regiochemistry in the **IX** and **JX** peptides on helix stability can be ascribed to the macrocyclic bridge, certain structures of which are less compatible with the alpha-helix framework than others. To test the effect of macrocyclic ring closure, we measured the CD spectra of peptides bearing the alpha, alpha-disubstituted amino acids but having no macrocyclic bridge; the synthesis of these peptides simply omitted the olefin metathesis step. In the **I** series, macrocyclic ring closure slightly destabilized the helix in the case of **IX2** double-bond isomer, and slightly stabilized helical structure in the other isomer, **IX1**. In the case of the **J** series, closure of the macrocyclic hydrocarbon ring actually destabilized helical structure in both isomeric products. Furthermore, hydrogenation of the olefin to the product bearing a fully saturated staple, **IX3** (same as **JX3**), resulted in a significant decrease in helical character (16% alpha-helicity). It appears that the particular olefin placement within the macrocyclic ring in the **JX** series, and also

the saturated 6-atom macrocycle in **IX**/**JX3**, has an especially powerful helix-destabilizing effect. These results indicate that the presence, stereochemistry and regiochemistry of the olefin moiety in the $R_{i,i+3}S(6)$ staple are all critical for effective alpha-helix stabilization. Olefin geometry is also important with the $S_{i,i+4}S(8)$ staple, but this system is almost completely unaffected by saturation of the bridge.

Once again, the RCM yields in the $R_{i,i+3}S(6)$ system mirror the extent of aqueous helix stability in the products, suggesting that the RCM reactions are templated by helix formation. Substrate **I** undergoes much more efficient RCM than **J** under our standard conditions, and indeed, **I** yields products having greater helix stability than **J** in aqueous solution. In addition, the most helical $R_{i,i+3}S(6)$ -stapled peptide **IX1** is less helical than the corresponding *i,i*+3 and *i,i*+4 cross-linked peptides having an 8-carbon bridge, and indeed, this 6-carbon system is formed less efficiently via RCM than 8-carbon variants. Although the thermodynamic driving force for helix formation in DCE is substantially different from that in aqueous solution, both are dependent upon the formation of main-chain hydrogen-bonding interactions that could be weakened by introduction of a torsionally strained macrocyclic ring. Such torsional strain would both decrease the rate of RCM in DCE and diminish helix stability in water. The 6-carbon system has fewer rotatable bonds and thereby a lower capacity to avoid repulsive torsional interactions than the 8-atom system, explaining the greater sensitivity in the former to structural features of the macrocyclic bridge.

Chapter 4. 4-carbon *i, i*+3 stapled peptides

We investigated the effect of having 4-atom hydrocarbon bridge in *i, i*+3 stapled peptides. Here again, we utilized the RNaseA sequence so as to be able to compare with the previous stapled

peptides. At the same *i, i*+3 position as in **I-P** and **R**, we placed an R_3 and S_3 pair in substrate **S**, and an S_3 and R_3 pair in substrate **T**.

Table 6. 4-carbon *i, i*+3 peptide sequences

RNaseA peptide	Substrate Sequence														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	E	W	A	E	T	A	A	A	K	F	L	A	A	H	A
S	E	W	A	R3	T	A	S3	A	K	F	L	A	A	H	A
T	E	W	A	S3	T	A	R3	A	K	F	L	A	A	H	A
U		W	A	R3	T	A	S3	A	K						

For **S**, no RCM reaction occurred when treated with GI catalyst at room temperature; however, elevation of the temperature to 65 ^{o}C resulted in over 90% conversion after three 2-hour treatments. **T** did not undergo metathesis even at high temperature. Recall that when S_3-S_3 was introduced in an *i,i*+4 format (substrate **C**), no metathesis reaction happened even at high temperature. This might be due to the fact that the distance between position *i* and *i*+3 that bridges a 0.83 turn of the helix is shorter than that of *i* and *i*+4 that bridges a 1.11 turn, making it possible to accommodate a 4-carbon bridge. (**Figure 12**)

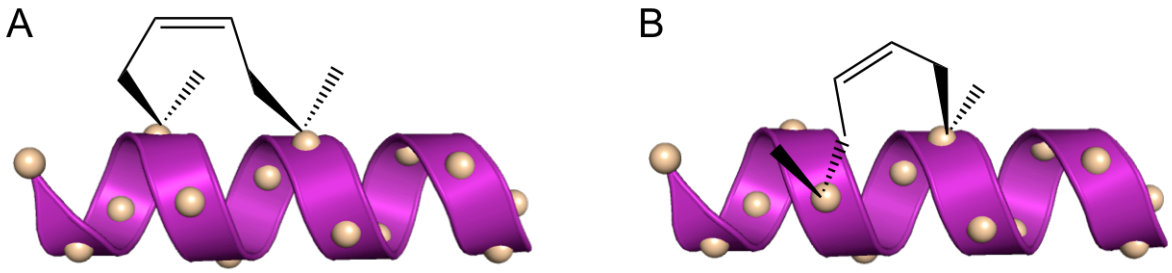


Figure 12. 4-carbon bridge in *i,i*+4 and *i,i*+3 formats.

Interestingly, the CD spectra of **SX** showed a strong negative maximum around 205 nm and a weaker (60-75% intense) negative maximum at 222-232 nm. (**Figure 13**) According to the

literature, such a spectrum represents a right handed 3_{10}-helix conformation [9,42,43] . Compared to an alpha-helix (phi(Φ)=-63, psi(Ψ)=-42) that has *i,i*+4 hydrogen bonds, a 3_{10}-helix (Φ=-57, Ψ=-30) has *i, i*+3 hydrogen bonds, having 10 atoms in the pseudo-ring formed by an intramolecular *i, i*+3 hydrogen bond. It is more tightly wound than an alpha helix, thus exhibiting a slightly narrower and longer shape [44] . This 3_{10} helix-like conformation is lost when the olefin is reduced. When the sequence was shortened to 8 amino acids from the previous 15 amino acid as in **U**, the stapled product **UX** shows a more intense 3_{10}-helix-like conformation, probably due to the fact that the 3_{10}-helix conformation does not propagate well throughout the length of the longer peptide.

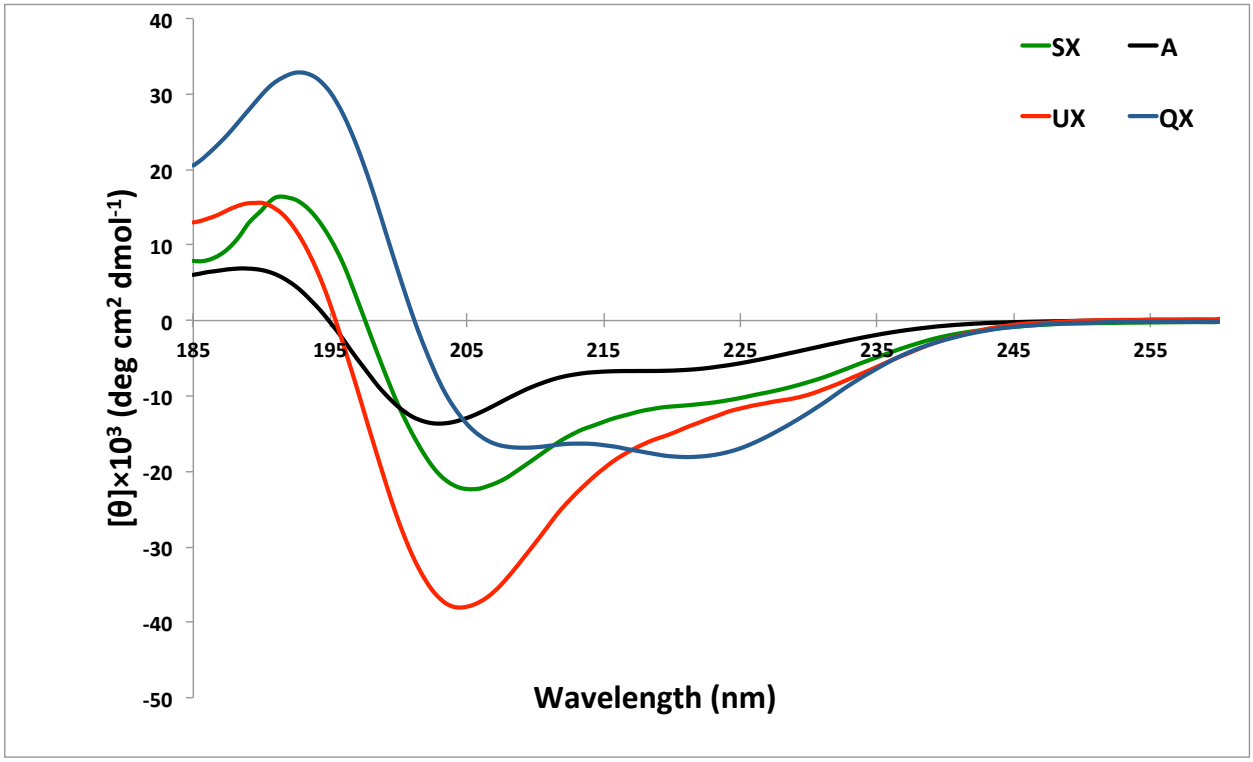


Figure 13. Circular dichroism analysis of 4-carbon *i, i*+3 stapled peptides. 3_{10}-helix conformation is shown.

In related work, Dr. Young-Woo Kim from the Verdine group has shown that when 4-carbon *i, i*+3 staple is put in a peptide sequence consisting of achiral alpha-aminoisobutyric acid (Aib) and

glycine(sequence: UGR_3GGS_3GU and UGS_3GGR_3GU), it is not only possible to stabilize a 3_{10}-helix conformation, but also to revert the handedness of the helix by inverting the stereochemistry from the R_3-S_3 to S_3-R_3 combination.

The left-handed helix is a rare motif, constituting only a small percentage of the secondary structure pool. However, in a number of cases they have been shown to have a functional or structural importance in the protein [45] . Although in the case of the RNaseA sequence **T** with the S_3-R_3 pair the RCM reaction did not work at all, it might be because there is an innate tendency to form a right handed helix. Whereas peptides consisting of L-amino acids preferentially form right-handed helices due to the steric hindrance between side-chain atoms and the main-chain carbonyl group, with certain amino acids there seems to be more tolerance and a lower energy barrier to form a left-handed helix. Thus, we decided to apply this system to an example of a left-handed 3_{10}-helix sequence from nature, to look at the possibility of stabilizing the conformation with a 4-carbon S_3-R_3 staple.

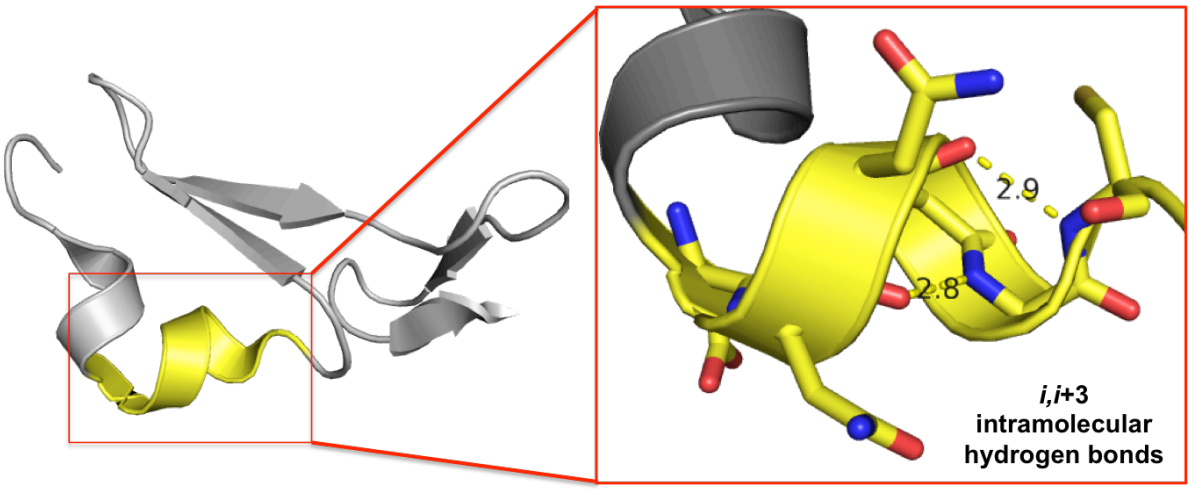


Figure 14. Left handed 3_{10}-helix shown in the EGF-A domain of the LDL receptor. (PDB: 3BPS)

Table 7. 4-carbon *i,i*+3 stapled peptide sequences for left-handed 3_{10}-helix stabilization

LDLR peptide	Substrate Sequence								
	1	2	3	4	5	6	7	8	9
V	W	L	D	N	N	G	G	S	S
W	W	L	D	R3	N	G	S3	S	S
Y	W	L	D	S3	N	G	R3	S	S

We synthesized a model sequence that is derived from the EGF-A domain of LDL receptor that is shown to adopt a left-handed 3_{10}-helix in the context of the folded protein (**Figure 14, Table 7**). Along with the wild type sequence **V**, all four possible stereochemical combinations were tested, including the R_3-S_3 (**W**) and the S_3-R_3 (**Y**) pairs. (**Table 7**) Although the RCM reaction proceeded for all four sequences, only **YX** stabilized the left-handed 3_{10}-helix. There were two stapled product isomers for **YX**, and only one of them showed the ability to stabilize the 3_{10} helical conformation. In addition, the ability of the macrocyclic ring to stabilize the 3_{10}-helix collapsed when the olefin was hydrogenated. (**Figure 15**)

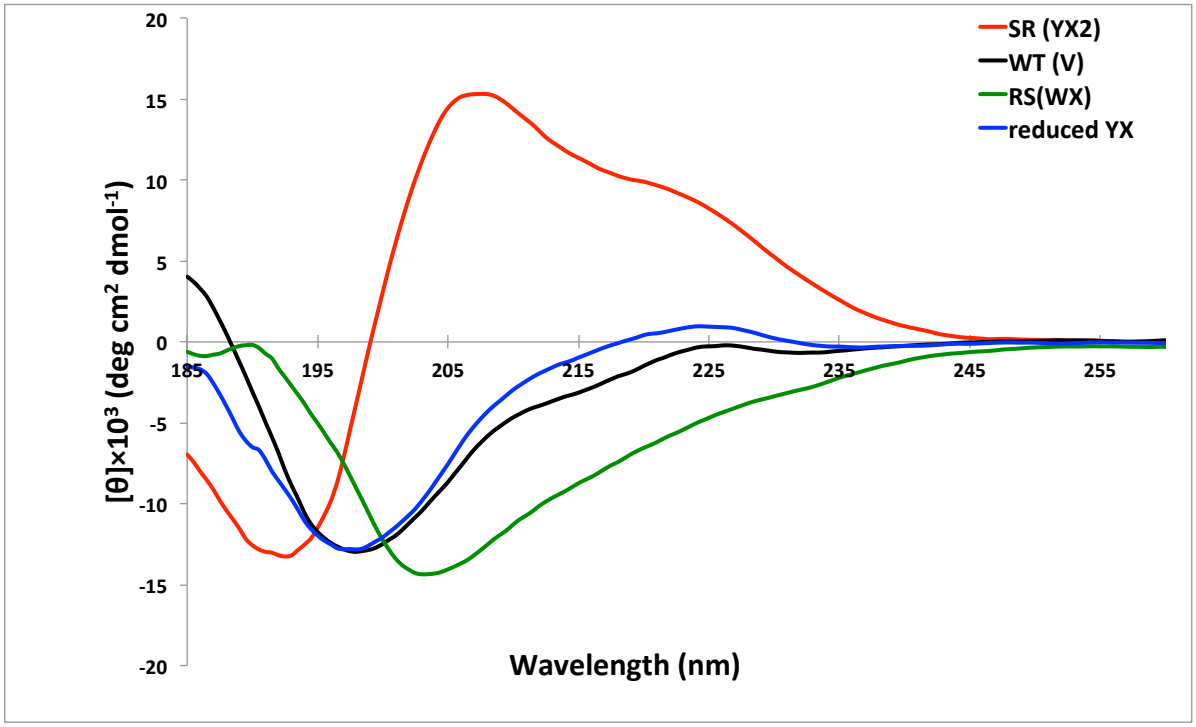


Figure 15. Stabilization of left handed 3_{10}-helix.

These results proved our hypothesis that a 4-carbon macrocylic bridge with the S-R stereochemical combination could be used to stabilize a left-handed 3_{10}-helix for a naturally occurring left-handed sequence. We also made an observation that when we incorporate D-amino acids instead of L-amino acids and reverse the stereochemistry of the unnatural amino acids as well, the CD measurement shows mirror images of those of corresponding peptides. (**Figure 16**) Here, we were able to stabilize a right-handed 3_{10}-helix conformation out of peptide sequence composed of D-amino acids.

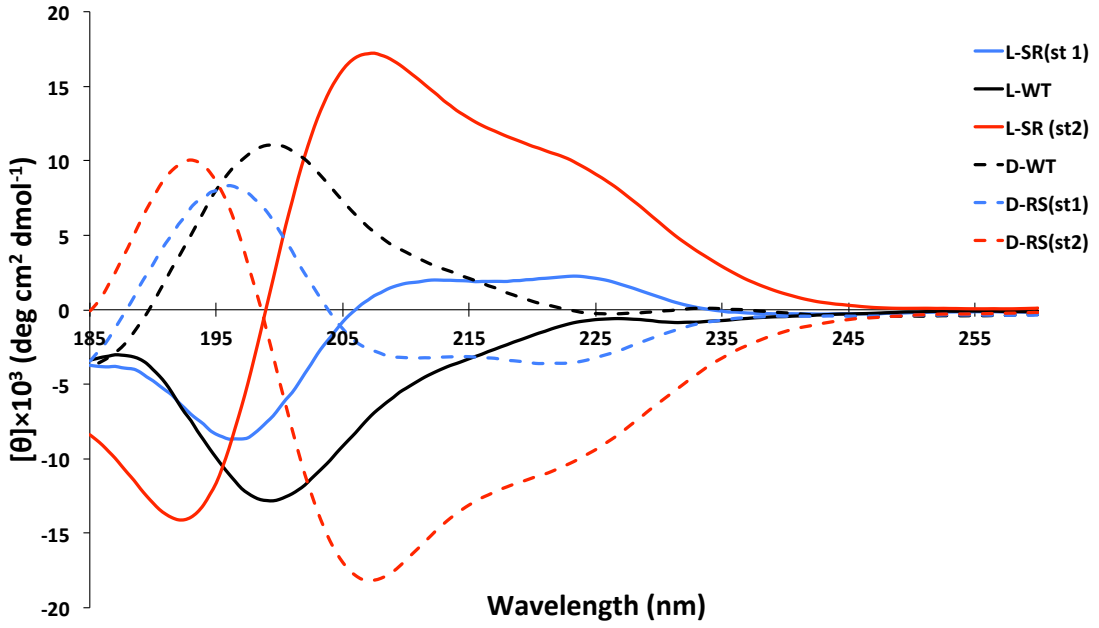


Figure 16. Stabilization of right handed 3_{10}-helix from a peptide composed of D-amino acids.

Attempt to form a mini-protein of EGF-A domain of LDLreceptor.

Following the success in stabilizing a left-handed 3_{10}-helix as shown above, we continued on to synthesize a part of the EGF-A domain of LDL receptor in order to produce a structurally-stabilized mini-protein that binds to PCSK9. The LDL-receptor is a plasma membrane glycoprotein that removes cholesterol-rich LDL particles from the plasma. Its level is known to be regulated by PCSK9, which is a 72-kDa protease that is expressed highly in liver. Gain of function mutations in PCSK9 result in reduced LDL receptor levels in liver and increased susceptibility to coronary heart disease, while loss of function mutations confer protection from coronary heart disease. When the PCSK9 protein binds to LDL receptor, it inhibits the recycling of the receptor, leading to its degradation. Therefore, PCSK9 is expected to be a good target for treating hypercholesterolemia.

Therefore, we decided synthesize of mini-protein that mimics the EGF-A domain of the LDL receptor to bind PCSK9 and to eventually prevent the degradation of the LDL receptor. In the design of the mini-protein, a portion of the EGF-A domain of the LDL receptor that interacts with PCSK9 was used. (**Table 8**) First, the peptides were synthesized starting from the C-terminus up to 8th residue where the stapling position *i* is, and then the RCM reaction was performed using HGII catalyst for 5 hours at 65 ^{o}C (2+2+1 hours, changing into fresh catalyst each time). Afterward, the remaining amino acids were coupled to complete the full sequence. This strategy was used because the rate of product conversion was faster when the stapling position was at the N-terminus of the 19 amino acid sequence than in the middle of the full 26 amino acid sequence. This might be due to the fact that the peptide consists of L-amino acids, which have the tendency to form a right-handed helix. In addition, the ring closing metathesis reaction is performed in dichloroethane, an organic solvent that aids right-handed helix formation. Thus, the ring closing metathesis reaction to form a left-handed helix would have to compete with these opposing factors that favor right-handed helix formation.

Table 8. Part of the EGF-A domain from LDL receptor sequence for left handed 3_{10}-helix stabilization.

LDLR sequence	Substrate Sequence																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
WT	G	T	N	E	C	L	D	N	N	G	G	C	S	H	V	C	N	D	L	K	I	G	Y	E	C	L
Stapled	G	T	N	E	C	L	D	S3	N	G	R3	C	S	H	V	C	N	D	L	K	I	G	Y	E	C	L

Again, there were two RCM products, which probably correspond to the double bond isomers. The final peptide products were rather insoluble; in most solvents the peptides precipitated, and only when ammonium hydroxide was used was it possible to dissolve the peptide. Since the mini-protein consists of 2 disulfide bonds with a 1-3, 2-4 configuration, we had to find an

optimal refolding buffer. After screening a number of conditions, the final refolding buffer conditions were set to 50 mM Tris buffer pH 8.5, 10 mM $CaCl_2$, 1mM reduced glutathione (GSH), and 50 uM oxidized glutathione (GSSG). These conditions yielded 2-3 distinct refolded peptide products. However, after the multi-step procedure to produce the final refolded product, the overall yield was very low. Unfortunately, when tested in a cell-based assay to see if the peptides are inhibiting the PCSK9 activity at Pfizer, these peptide products did not seem to show binding to PCSK9. It is possible that this may be, at least in part, due to the peptides being very insoluble and behaving poorly during the assay.

One method we tried to increase the overall yield of the final product was to make a building block of the stapled peptide fragment. We reasoned that once the olefin geometry that gives the left-handed 3_{10}-helix stabilization was determined, it would be possible to improve the overall yield of the synthetic mini-protein by separating the two double bond isomers of the building block and incorporating the desired fragment into the peptide sequence.

A building block of S_3NGR_3CS was synthesized with an Fmoc group on the N-terminus and a COOH group at the C-terminus. After ring closing metathesis, two stapled products appeared that were inseparable with the conventional HPLC solvent system of acetonitrile and water. However, when the solvent system was switched to 20-50% methanol in acetonitrile and water, the isomer peaks could be separated. (**Figure 17**)

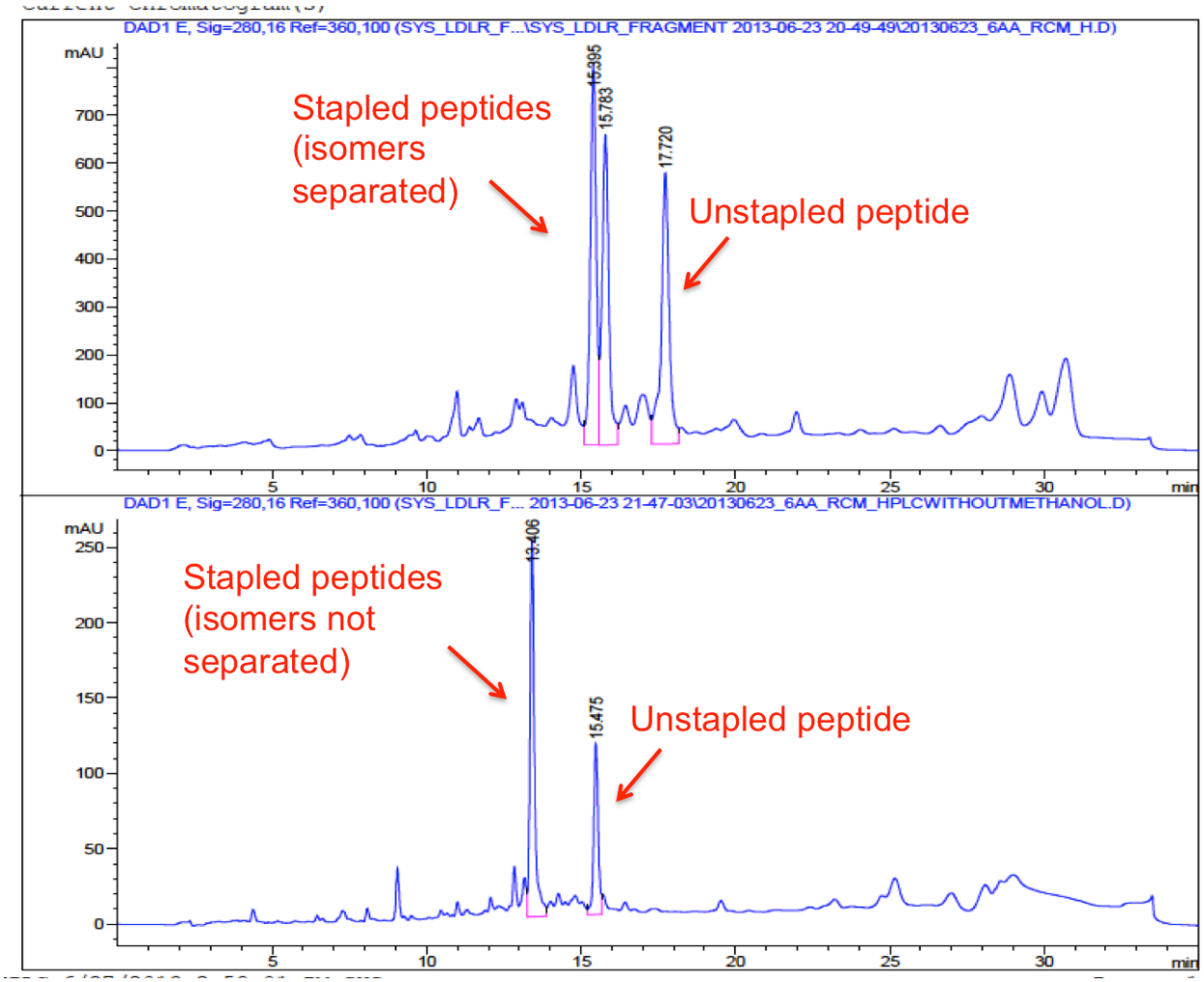


Figure 17. Different solvent system used for separation of two building block isomers. First row, solvent A: water(0.1% TFA) solvent B: 20% MeOH in acetonitrile(0.1% TFA). Second row, solvent A: water(0.1% TFA) solvent B: acetonitrile (0.1% TFA)

The NMR analysis of the two double bond isomer building blocks was performed with the help of Dr. Kazu Hayashi. Each isomer peak was dissolved in DMSO-d6. Although the first peak shows broadening of the olefinic signals, the second peak shows J1 of ~15Hz for both olefinic protons, which indicates that it is a trans isomer (**Figure 18**). Thus, we could infer that the first peak was the cis isomer, and the second one trans.

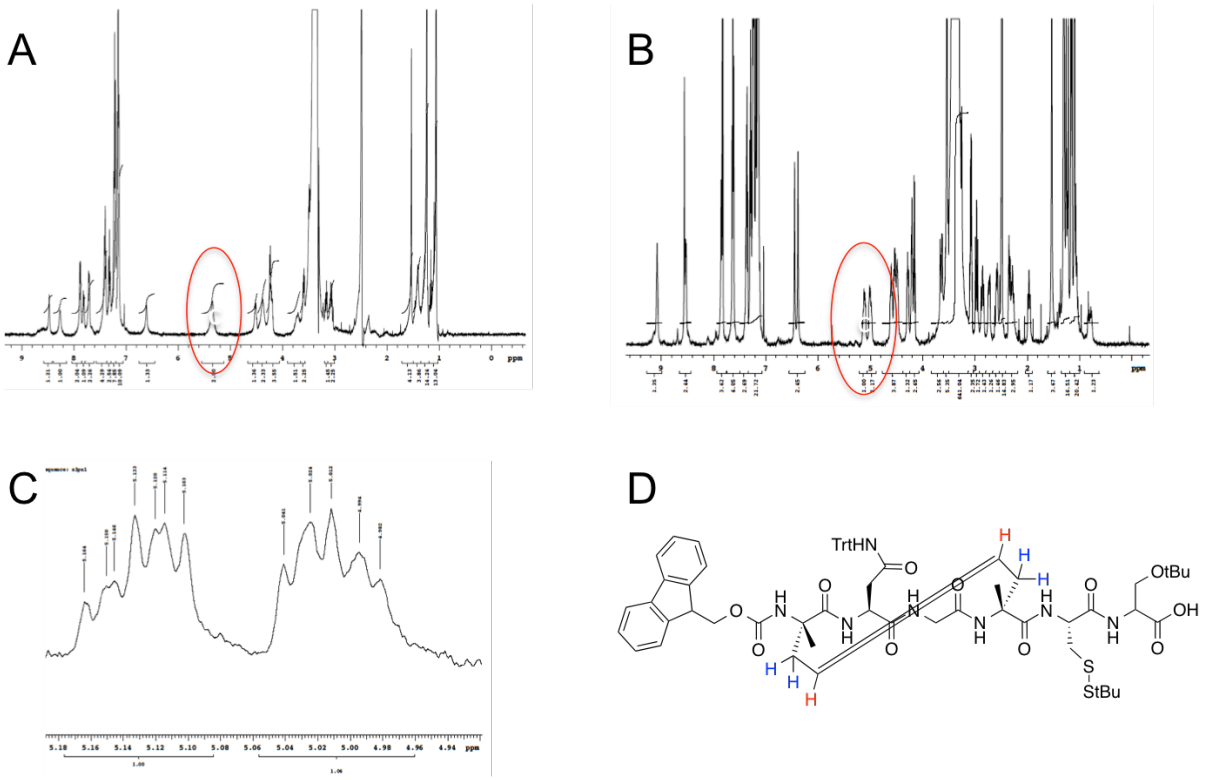


Figure 18. NMR analysis for determining double bond geometry of 4-carbon *i, i*+3 stapled peptide. A: Proton NMR of 1st peak B: Proton NMR of 2nd peak C: The olefinic peaks from the 2nd peak shows to be trans(E) isomer D: Peptide building block studied in NMR

Although it was the second peak **YX2** that had shown a left-handed 3_{10}-helix stabilization, the peptides used for NMR analysis were protected peptide fragments that were purified using a non-standard solvent system. We therefore coupled the building block (unstapled and two stapled peptides) back onto the resin, and elongated to make the same sequence as in **V**, **W** and **Y**. Interestingly, the retention time reversed, i.e the building block that had eluted second produced a full peptide that eluted earlier than the other isomer in the acetonitrile/water solvent system. When these peptides were analyzed by CD, the peptide containing the 1st building block with cis (Z) isomer, now eluting later, showed the left-handed 3_{10}-helix stabilization (**Figure 19**). Therefore, we could conclude that the 4-carbon *i, i*+3 staple that stabilizes the 3_{10}-helix probably has a Z olefin, which makes sense since there might be too much ring strain for in the case of the E isomer.

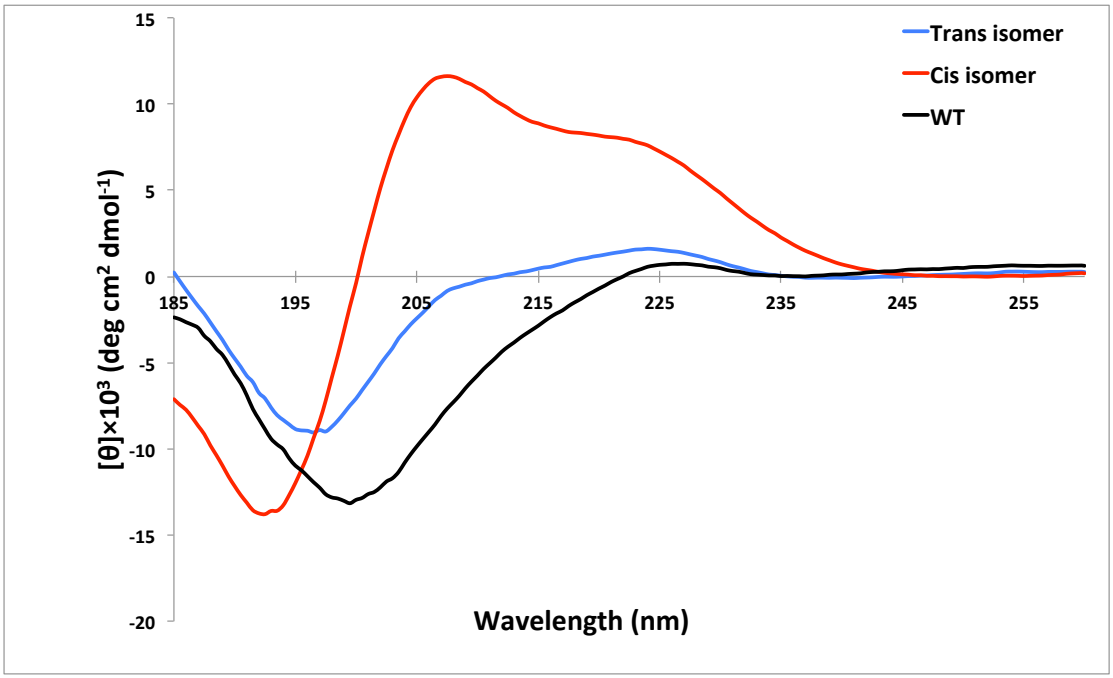


Figure 19. Circular dichroism of the 4-carbon *i, i*+3 stapled peptide isomers show that the cis-isomer is stabilizing the 3_{10}-helix.

Chapter 5. Conclusions and Future directions

In these studies, we investigated a 6-carbon and 7-carbon *i, i*+4 cross-linking system as well as a 4-carbon and 6-carbon *i, i*+3 cross-linking system. Although previously it has been shown that the RCM reaction for 6-carbon *i, i*+4 is not possible, we learned that it is possible to make the reaction to happen by optimizing the reaction conditions such as the temperature and the catalysts. Although we find the 6 and 7-carbon *i, i*+4 cross-linking systems and the 6-carbon *i, i*+3 cross-linking system to be less effective at stabilizing alpha-helices than the 8-carbon stapling system, the most favorable 7-carbon *i, i*+4 staple and the 6-carbon *i, i*+3 staple can nevertheless induce a considerable degree of alpha-helical character. It has been shown previously that inducing the maximum alpha-helicity that a peptide sequence can acquire does not always directly relate to the best biochemical or biological assay outcomes [46] , suggesting that the alternative stapling systems warrant further investigation in ligand discovery projects.

The impetus for this study was provided by the recent reports in both the *i, i*+4 and *i, i*+7 systems that the atoms of the staple itself can engage residues on the surface of the target, and by the observation that staple phasings of *i, i*+4 and *i, i*+7 are not always suitable candidates for substitution. Therefore, the alternative stapling systems were evaluated with respect to RCM kinetics and the conformational stability of the constrained products. Previously it has been shown that there is a good general correlation between the yield of peptide in RCM reaction and the extent of helix induction in the stapled product [10,30] . This is not unexpected, as the RCM reactions are carried out in DCE, a non-polar solvent, and such low-dielectric solvents are known to strongly promote alpha-helix formation [47] . This templating effect is presumably a key-driving force in the RCM reaction, effectively preorganizing the RCM substrate in a conformation that increases the local concentration of the reacting olefins relative to that in an unstructured or alternatively structured peptide. This templating effect, together with kinetic reversibility of RCM, also guarantees that torsionally strained cross-link configurations will not be formed in high yields. Thus, the yield of product in the RCM reaction provides a rough preliminary indicator of helix stability in the product. Another motivation for this study came from the hypothesis that the shorter carbon linkers would alleviate the aggregation problem, and this could be examined by comparing the light scattering signals of the stapled peptides with different linker lengths.

We also found that the presence of an olefin, and its position and geometry, are all critical factors in achieving maximal alpha-helix stabilization of a peptide using the 6-carbon and 7-carbon *i, i*+4 cross-link and 6-carbon *i, i*+3 cross-link. Interestingly, it was apparent in a number of cases

that not only the length and stereochemistry, but also the position, of the olefin within the bridge affected the RCM kinetics as well as the helical induction. When the olefin of the staple was closer to the N-terminus than to the C-terminus, the ring formed more readily and the resulting macrocycle was more alpha-helical. In addition, when the olefin would end up in the middle, the substrate showed even better performance in the RCM reaction as well as higher alpha-helicity. One possible reason to explain this phenomenon might be that steric hindrance of the catalyst with the peptide makes it harder to form a bridge when the linker is shorter (i.e. allylglycine). This would explain why the staple forms better when the linker is symmetric. The reason why the 3-5 or 4-5 pair shows better RCM reaction than 5-3 or 5-4 pair, might come from the possibility of a Ru-Carbene complex from the allylic sidechain or butenyl sidechain forming a stable chelate with the carbonyl oxygen within the same residue, as suggested in a similar study where S_3-S_7 and S_7-S_3 in *i, i*+4 phasing was investigated [39] Such an interaction would form an intermediate pseudo 6 and 7-membered ring, respectively, which would make it harder for the RCM reaction to occur when the complex is facing the opposite direction as in the case where allylglycine or butenyl glycine is at the latter position.

In addition, we have shown that by using the 4-carbon *i, i*+3 stapling system, it is possible to stabilize a right-handed 3_{10}-helix, and to similarly stabilize a left-handed 3_{10} helix when the stereochemistry of the cross-linking amino acids are reversed. Previously, stabilization of a 3_{10}-helix has been observed using ring-closing metathesis reaction [9] or click chemistry [48] , however in both studies the sequences contained the non-proteinogenic amino acid Aib which induces a 3_{10}-helical conformation when present in a large fraction in a peptide [18] .

Although additional experiments examining proteolytic degradation or cell permeability should be performed to complete the story, we hope these initial studies on different stapling systems would aid in finding the sweet spot for the most optimized binding between the stapled peptides and the target protein. In addition to the conventional way of screening staple positions, we could also try removing one or two carbon atoms from the linker in case the bridge itself forms a contact with the target. In addition, the cis isomer of the 4-carbon *i, i*+3 stapling system can be used if the critical binding interaction relies on a 3_{10}-helical format. By reversing the stereochemistry of the unnatural amino acids from the R, S combination to the S,R combination, we have proved the possibility of stabilizing a left-handed 3_{10}-helix for naturally occurring left-handed helices. This also suggests that it might be possible to stabilize a left-handed alpha-helix in an *i, i*+4 format as well for certain sequences.

Chapter 6. Experimental Method

Peptide synthesis

All peptides were synthesized using Fmoc chemistry on Rink Amide MBHA resin with a loading capacity of 0.50 mmol/g [19] . The resin was swelled with NMP for 15 min before use. The Fmoc protecting group was removed by treatment with 25% piperidine in NMP (2 x 10min) Natural amino acids were coupled for 45 min using HCTU as the activating agent (4.75 equiv.), five equivalents of Fmoc-protected amino acid, and 10 equivalent of DIPEA in NMP. For the coupling of olefin-bearing unnatural amino acids, a reaction time of 2 h was used with Fmoc-protected amino acid (3.0 equiv), HCTU (2.85 equiv.), and DIPEA (6.0 equiv.). These unnatural

amino acids were bought from Okeanos Tech. After each coupling or deprotection reaction, the resin was washed with DCM (1x1min), NMP (1x1min), DCM (1x1min), and NMP (1x1min).

Metathesis reaction

Ring-closing metathesis of the fully elongated, fully protected, resin-bound peptides was performed using 20 mol % of either Grubbs first- or second-generation catalyst or Hoveyda-Grubbs second-genetration catalyst in degassed DCE for 2h at 25 ^{o}C or 65 ^{o}C. For the reactions at 65 ^{o}C, the reaction vessels were shaken in a temperature-controlled environment. The reactions were monitored by liquid chromatography/mass spectrometry (LC/MS) after cleavage of the peptides from a resin aliquot. For the second round of metathesis, the resin was treated with fresh catalyst for an additional 2 h. After each round of metathesis, the reaction solution was drained and the resin was washed first with DCM (5x2min) and then with NMP (3x2min). After the Fmoc group was removed using 25% piperidine in NMP, the N-terminal amino group was treated with 30 equivalents of acetic anhydride (102 g/mol, 1.08g/ml) and 60 equivalents of DIPEA (129.25 g/mol, 0.742 g/ml) in NMP for 45 min at room temperature. The synthesis resin was washed with DCM (1x1min), NMP (1x1min), DCM (3x1min), and methanol (3x2min) and dried in vacuo overnight.

Olefin hydrogenation

The hydrogenation of the olefin in the hydrocarbon bridge was performed by adding 200 ul of 0.7 M TIBSH(2,4,6 tri-isopropyl benzenesulphonyl hydrazide) and 1.4 M piperidine in NMP to the peptides on resin. The reaction tube was sealed tightly and put in 55 ^{o}C incubator. The

progress of the reaction was monitored by doing test cleavage and injecting the product into LC/MS after 2 hours. The same procedure was repeated until the olefin was completely reduced.

Peptide purification

The peptides were deprotected and cleaved from the resin by treatment with a mixture of TFA/TIS/water (95/2.5/2.5) for 2 h. In the case where cysteine was present in the peptide, a cocktail of TFA/TIS/EDT/water was used instead. After the majority of volatile material was removed under a gentle stream of nitrogen in a fume hood, the peptide products were precipitated by addition of a 1:1 mixture of n-pentane and diethyl ether. The precipitate was suspended by vigorous vortexing and collected by centrifugation. The pellet was allowed to air-dry until the traces of liquid ether were no longer visible, the products were dissolved in a 1:1 mixture of acetonitrile and water, and the resin was removed by filtration. The products were purified by reverse-phse HPLC using a Zorbax C18 column (Agilent, 5 um, 9.4x250 mm) and characterized by LC/MS using a Zorbax C18 column (Agilent, 3.5 um, 2.1 x 150 mm) and amino acid analysis.

Circular Dichroism

Peptides were dissolved in 20 mM potassium phosphate buffer (pH 6.5), and the concentrations as determined by absorbance spectroscopys (the extinction coefficient, λ_{280}, for tryptophan =5960/M/cm) [49] were adjusted to 70-110 uM. Circular dichroism spectra were collected on a Jasco J-710 spectropolarimeter (Japan Spectroscopic Corporation, Tokyo, Japan) equipped with a temperature controller using the following standard measurement parameters: 0.5 nm step

resolution, 20 nm/seconds speed, 10 accumulations, 1 second response, 1 nm bandwidth, 0.1 cm path length. All spectra were converted to a uniform scale of molar ellipticity after background subtraction. Curves shown are smoothed with standard parameters.

Building Block synthesis

Acid-labile 2-chlorotrityl chloride resin with a loading capacity of 1.2 mmol/ g was used for building block synthesis. For the first amino acid residue attachment, 3 equivalents of fmoc-amino acid mixed with 7.5 equivalent of DIPEA in dry DCM was used. The coupling reaction went for 30-60 min at room temperature, and afterward the resin was washed with DCM/MeOH/DIPEA (80:15:5) three times, 15 min each. The subsequent couplings are the same as in conventional Fmoc SPPS. For cleavage reaction, mixture of HFIP:DCM (1:4) was used to obtain the protected peptide fragments. The purified building block was put back onto the resin in the same way, but this time using about 1 equivalent of the building block mixed with 2.5 equivalent of DIPEA in dry DCM, and capping the remaining linkers on the resin with MeOH.

NMR

^{1}H NMR spectra were measured in DMSO-d6. Mass spectral data were obtained at the Harvard Mass Spectrometry Facility.

Part II. Development and structural studies of direct peptide inhibitors of Ras

Chapter 1. Ras protein

1-1. What is Ras?

Ras is a molecular switch that cycles between on and off depending on the guanine nucleotide binding mode. When the bound GDP dissociates and exchanges for GTP which has a higher concentration in the cell, Ras switches on. This reaction is mediated by guanine nucleotide exchange factors (GEF) and modulates various intracellular signaling pathways regulating cell growth, differentiation and proliferation via effector proteins [50,51] . Conversely, the intrinsic GTPase reaction hydrolyzes GTP to GDP, which is accelerated by about five orders of magnitude by association with a GTPase activating protein (GAP) [52] [53,54] , which causes Ras to switch off.

The guanine nucleotide binds and hydrolyzes in the 20-kD G domain of Ras. It consists of 5 alpha-helices and 6 beta-strands, and the guanine nucleotide binding site shows conserved sequence elements, such as the N/TKXD motif that interacts with the nucleotide base or the phosphate-binding loop (P-loop) that binds to the phosphates of the guanine nucleotide [55] . The high affinity binding of guanine nucleotide to Ras is due to the interaction of the phosphates with the P loop and Mg^{2+} ion. To promote nucleotide exchange GEFs interact with the switch I and

switch II regions by inserting residues close to the P loop and Mg^{2+} binding area and creating a structural change that inhibits the binding of phosphates and Mg^{2+}. The disturbance of the P loop seems to be the major reason for the drastically decreased nucleotide affinity [56]

The differences between the structures of GTP-bound Ras and GDP-bound Ras lie in the two switch regions, switch I (residues 32-38) and switch II (residue 59-67). The crystal structures of H-Ras show that Ras contacts the γ-phosphate via the main chain NH groups of Thr35 in switch I and Gly60 in switch II [57] (**Figure 20**).

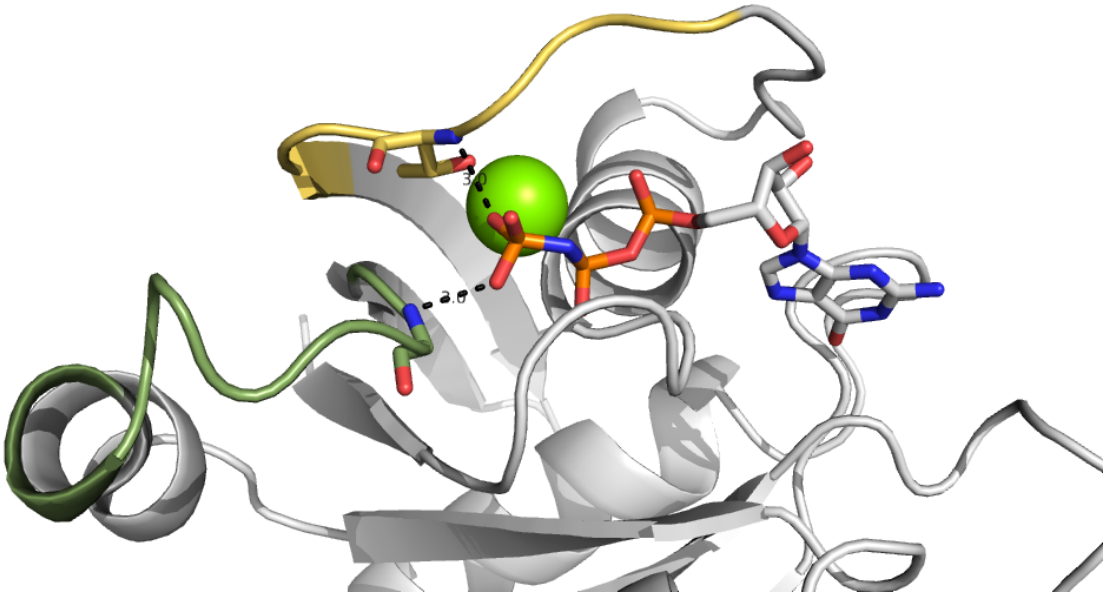


Figure 20. GTP-bound Ras. The Ras domain is shown in cartoon representation, and the nucleotide and magnesium ion are shown in ball-and-stick representation. Threonine 35 in switch I (yellow) and glycine 60 in switch II (green) are shown in sticks. The main chain NH groups of Thr35 and Gly60 are interacting with the γ -phosphate of the GTP analogue. (PDB 5P21)

When GTP hydrolyzes, the gamma phosphate is released and induces increased flexibility in the switch regions [55] , thus resulting in allosteric conformational changes [58] (**Figure 21**).

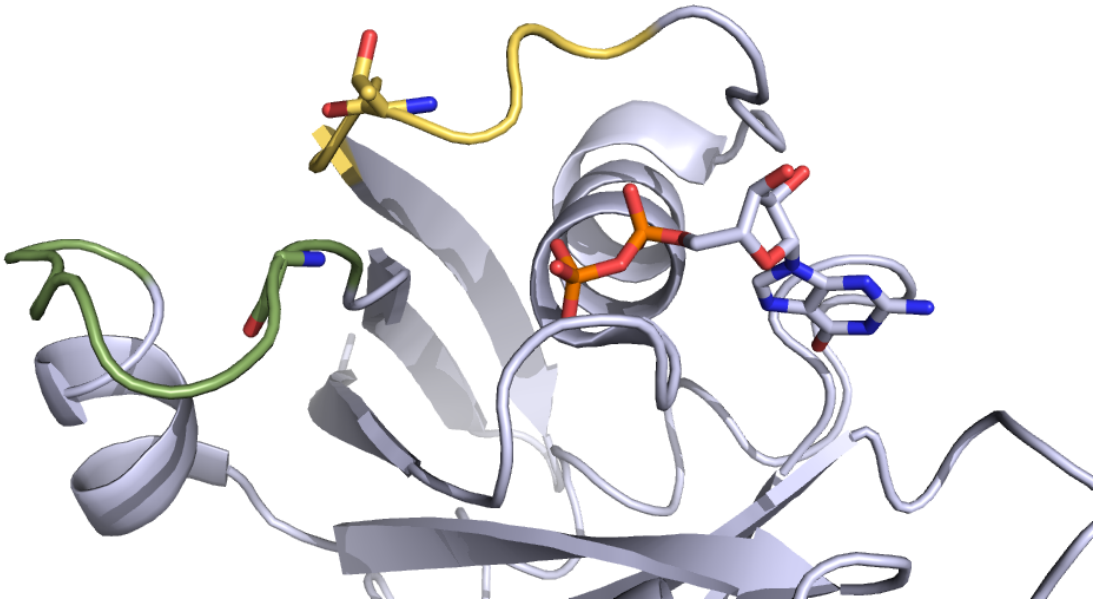


Figure 21. GDP-bound Ras. The Ras domain is shown in cartoon representation, and the nucleotide in ball-and-stick representation. Threonine 35 in switch I (yellow) and glycine 60 in switch II (green) are shown in sticks. Unlike in the GTP-bound form, no interaction is shown between Thr35 and Gly60 with guanine nucleotide. (PDB 4Q21)

This can be described as a 'loaded spring mechanism', where the release of the γ-phosphate from the hydrolysis of GTP results in the two switch regions relaxing into a GDP-specific conformation [55] (**Figure 22**).

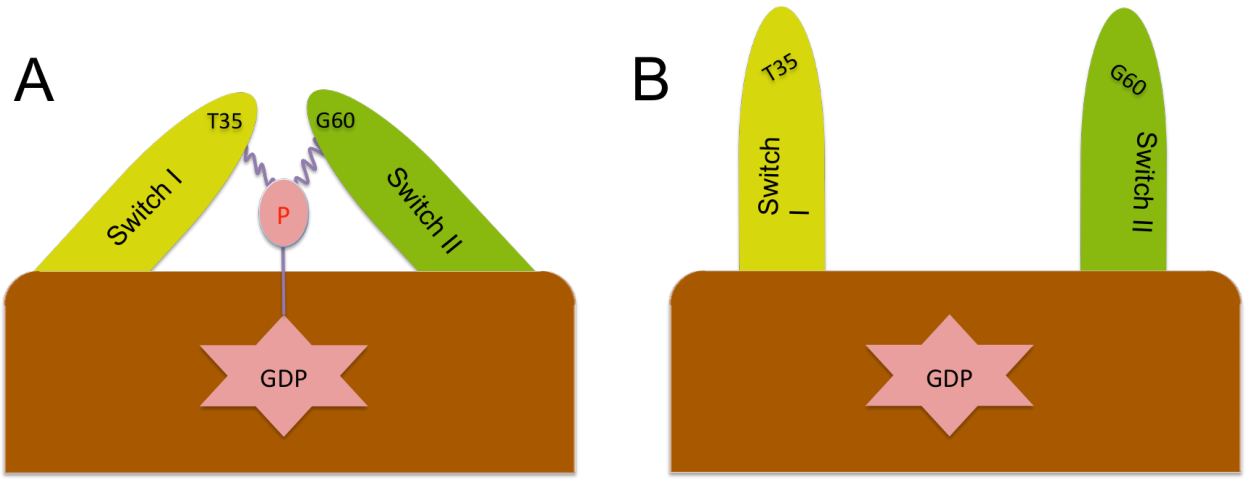


Figure 22. Loaded-spring mechanism of Ras. A. Schematic diagram of Ras in GTP form. B. Schematic diagram of Ras in GDP form. With the release of γ-phosphate from GTP, the interactions with switch I and switch II that were present in GTP-bound Ras are absent, conferring flexibility in the switch regions.

Effectors for Ras interact more tightly when Ras is in the GTP-bound form rather than the GDP-bound form; such nucleotide-dependent recruitment of effectors in a major signal transduction mechanism. As described above, in the GDP-bound state, flexibility of the switch I region increases significantly, which is shown to be related to the decreased affinity to effectors [59] . One well-known effector that presents a characteristic Ras-binding domain (RBD) is the Raf protein kinase, which binds to Ras with 3.5nM affinity [60] by forming an interprotein beta sheet between the two molecules [61] . Ral-GDS is another example [62] that has ubiquitin-like domain that binds to Ras in a way similar to that of Raf. While these effectors are shown to bind mostly to switch I region of Ras, another well-known Ras effector PI3K is shown to make critical interactions with switch II region of Ras as well, which produces structural changes that might affect its phospholipid substrate binding and catalytic activity. (**Figure 23**)

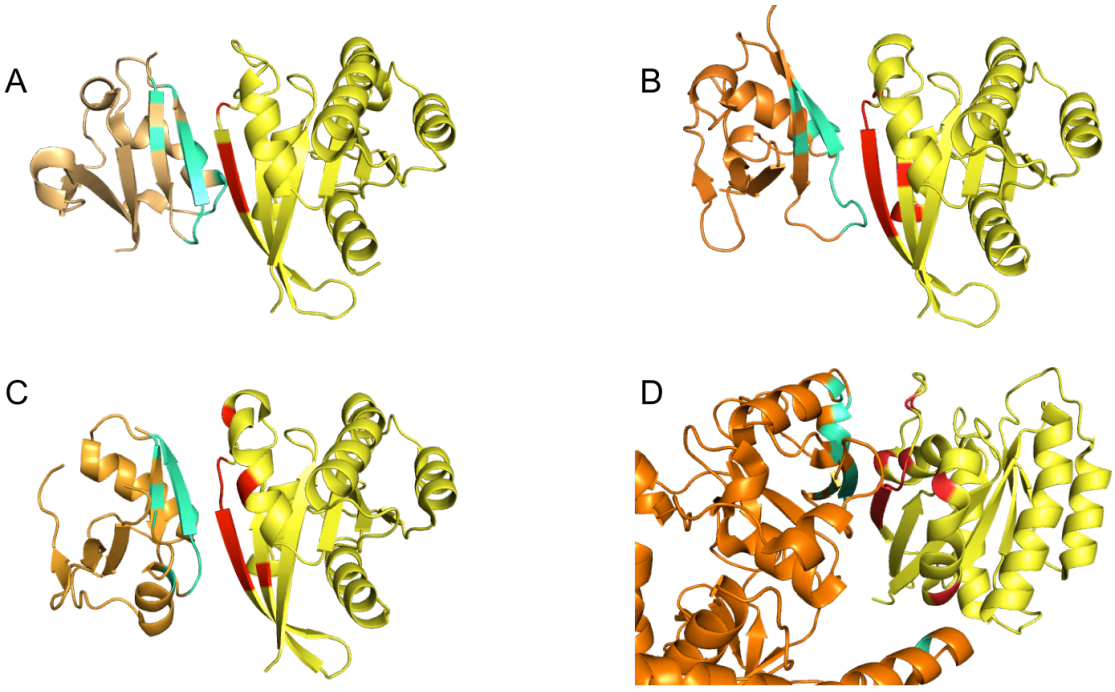


Figure 23. Ras families binding to their effectors. Ras is on the right, colored in yellow, and its effectors on the left, in orange-related colors. The residues within 4A of Ras are colored in cyan, and the residues within 4A of the effectors are colored in red. A: Rap1A-RasRBD complex (PDB:1C1Y) B: Ras-RalGDS complex (PDB: 1LFD) C: Ras-Byr2RBD complex (PDB:1K8R) D: Ras-PI3K complex(PDB: 1HE8). The effectors are shown to be interacting with Ras via beta-sheet complementation with its switch I region. Interactions with switch II are also shown in C & D.

1-2. Ras and tumor formation

The Ras genes are known to be the most frequently mutated oncogenes in human cancer [63] (**Table 9**). There are four isoforms of Ras: Kirsten rat sarcoma virus (K-Ras) 4A, K-Ras 4B, Harvey rat sarcoma virus (H-Ras) and Neuroblastoma rat sarcoma virus (N-Ras). These Ras proteins are highly conserved, especially in their effector binding regions where they are actually identical. They mostly differ only in the C-terminal 25 amino acids, the hypervariable region that is responsible for anchoring Ras into the cell membrane [64] . K-Ras-4A and K-Ras 4B are actually encoded by the same gene, and only differ in their fourth exons. Although the isoforms are highly similar, there are actually different frequencies of mutations of K-Ras, N-Ras and H-Ras in human cancer (**Table 9**). Each isoform is correlated specifically with certain types of cancer; K-Ras mutation is frequent in pancreatic, colon, small intestinal and lung cancer, H-Ras in urinary tract cancer, and N-Ras in skin cancer. However, among these K-Ras is known to be the major Ras gene mutated in about 30% of human cancer, with poor prognosis and chemoresistance [65] . Although it is not known for sure why K-Ras mutation correlates more highly with cancer than the other isoforms, it might be due to the differential regulation of the gene expression during development rather than the properties of the proteins themselves [66] .

Table 9. Frequency of Ras isoform mutations in selected human cancers. Data adapted from Stephen *et al* [63] .

Frequency of Ras Isoform Mutations in Selected Human Cancers

Primary Tissue	KRAS (%)	HRAS (%)	NRAS (%)	Total (%)
Pancreas	71	0	<1	71
Colon	35	1	6	42
Small intestine	35	0	<1	35
Biliary tract	26	0	2	28
Endometrium	17	<1	5	22
Lung	19	<1	1	20
Skin (melanoma)	1	1	18	20
Cervix	8	9	2	19
Urinary tract	5	10	1	16

Missense mutations at codons 12, 13 or 61 in Ras are known to promote oncogenesis by inhibiting GAP-mediated GTP hydrolysis. The active site of Ras shows a conserved glutamine residue Q61 near the γ-phosphate. Since it plays an important role in stabilizing the transition state of the hydrolysis reaction by interacting with a water molecule that bridges one of the gamma-phosphate oxygen atoms to OH group of Y32 [67] , mutation of this catalytic residue is known to be oncogenic. Another important residue is G12 in the P-loop, where a codon substitution sterically blocks GAP mediated GTP hydrolysis. This inhibition of GTP hydrolysis leads to a large proportion of stabilized GTP-bound mutant Ras proteins, resulting in constitutive activation of the Ras signaling pathway.

1-3. Strategies for Ras inhibition

While the significance of targeting K-Ras for cancer therapeutics is evident, to date there have been no reported high affinity inhibitors of Ras. One previous attempt to inhibit Ras activation was to use a farnesyltransferase inhibitor to impede posttranslational modification Ras and prevent insertion into the inner leaflet of the cell membrane [68] . However, it was found that farnesyltransferase is not required for insertion of K-Ras and N-Ras into the membrane; geranylgeranyltransferase is alternatively capable of modifying Ras for membrane insertion [69] . Another failed strategy for Ras inhibition was the production of a Raf inhibitor that led to paradoxical activation of Raf kinase [70] . Another possible targeting strategy for Ras would be to block the GTP binding site or disrupt nucleotide exchange. Since there is a high level of GTP in cells and Ras binds GTP with picomolar affinity, however, any strategy targeting the GTP binding site would be very difficult.

The Shokat group recently developed a covalent catalytic site-inhibitor for Ras that undergoes through electrophilic attack to covalently modify the non-natural cysteine present in K-Ras G12C. By engaging a binding pocket beneath the switch II region, the inhibitor induces a preference for Ras to bind GDP and adopt its inactive state [71] . Importantly, this work has shown that the Ras proteins exist in in dynamic and flexible states, revealing the potential for new therapeutic strategies that exploit this malleability. The Walensky group produced a stapled SOS1 helix that directly inhibited the interaction between K-Ras and SOS1, leading to inhibition of nucleotide association and a decrease in K-Ras signaling [72] .

Here we describe our ongoing attempts to develop peptides that directly block the recruitment of effector proteins to the Ras effector domain [73] as a means to suppress Ras signaling in cancer cells.

Chapter 2. Directed evolution via yeast surface display

2-1. Yeast surface display

The surface of the Ras effector domain does not have an obvious hydrophobic pocket that could be targeted by small molecules, and as shown in figure this domain interacts with effector proteins via beta-sheet complementation. Therefore with limitations in using rational design for discovering inhibitors of this protein-protein interaction, we performed directed evolution of peptides using yeast surface display.

Yeast surface display system has been developed Wittrup lab as a powerful means for expressing 10^4~10^5 copies of a heterologous protein on the cell wall of yeast *Saccharomyces cerevisiae* using the a-agglutinin system. [74] [75] It has been used for many interesting applications such as affinity maturation through repeated rounds of error-prone PCR and fluorescence activated cell sorting (FACS), screening cDNA libraries [76] or engineering proteins for improved production and stability [77,78] . Compared to phage display, yeast display has advantages in that it utilizes eukaryotic secretory machinery that enables post-translational processes such as oxidative folding or N-glycosylation. In addition, the ability to use multiple fluorescent channels in the

FACS step of yeast surface display makes multiparameter sorting possible. For example, with a two-label system we can check both the binding level and the expression level by flow cytometry, enabling normalization of binding to expression level and minimizing artifacts that might come from host-expression-bias [79] . Therefore, we expect to obtain proteins with both increased affinity and stability by using iterative rounds of random mutagenesis and directed evolution.

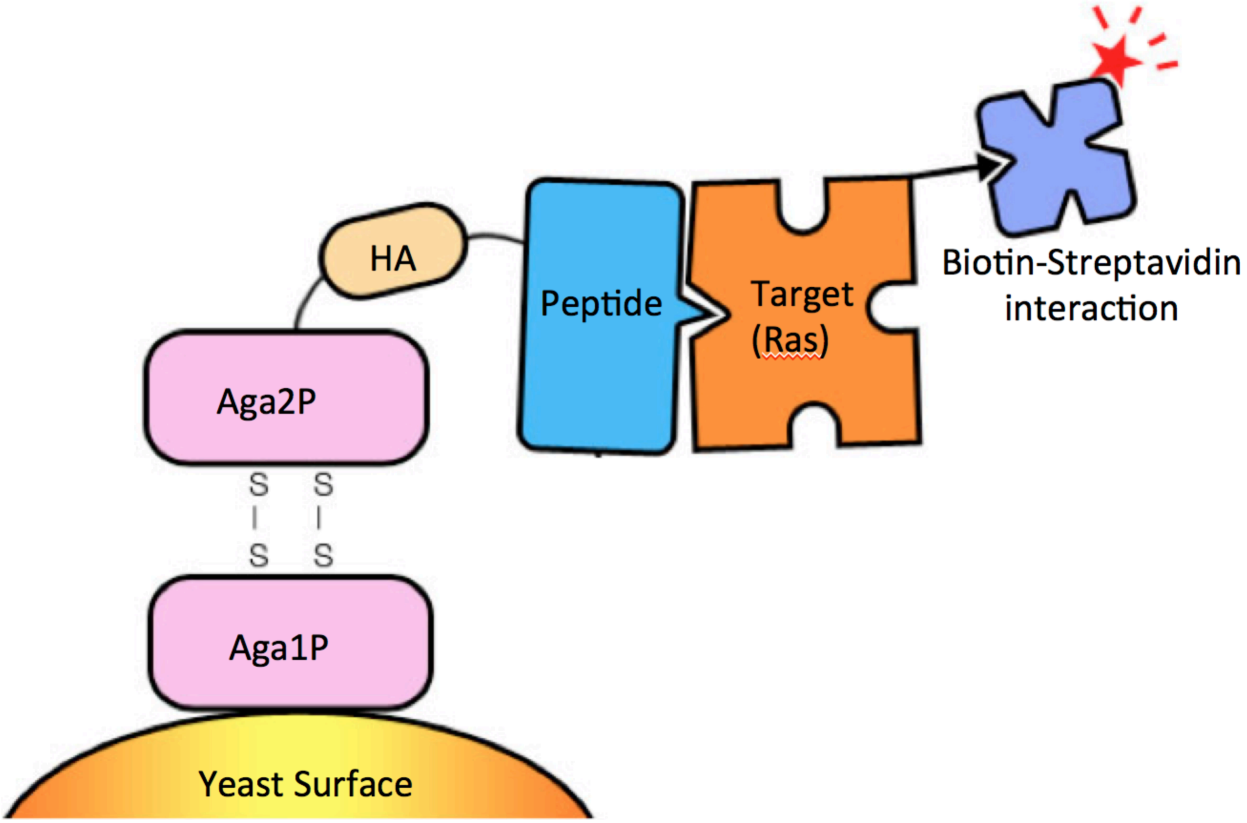


Figure 24 Yeast surface display. The binding of Aga2P-fused peptide (blue) to Ras protein (orange) is detected using secondary reagent such as streptavidin conjugated to fluorophore, which would bind to the biotinylated Ras. Expression of the peptides is detected through immunofluorescence labeling of hemagglutinin (HA) epitope tag (Beige). Adapted from Chao et al [80] .

As shown in **Figure 24**, Aga1P is anchored to the cell wall, and Aga2P is linked to Aga1P via disulfide bonds. The C-terminus of Aga2P can then be fused to the N-terminus of a protein of interest through a flexible linker such as $(GGGGS)_n$. The diversity in the library constructs is

generated by PCR. Using this system, it is possible to express and screen genetically encoded combinatorial libraries of peptides against a soluble ligand in a rapid and quantitative way using FACS. By limiting the concentration of the soluble target, it is possible to enrich for mutants with high affinity. In addition, with the genotype-phenotype linkage it is straightforward to identify initial binders and to further optimize the sequence with successive rounds of directed evolution.

2-2. Scaffolds

We initially began the yeast surface display experiment using four different scaffolds: the avian pancreatic polypeptide sequence, a short beta hairpin sequence, the PI3K knuckle region and the scyllatoxin sequence. These initial libraries for screening were designed by Dr. John McGee in Verdine group.

A. Avian pancreatic polypeptide scaffold

Avian Pancreatic Polypeptide (aPP) is a well established natural folded polypeptide that contains a polyproline II motif and an α-helical region. The interaction between these two helices stabilizes the fold, and it is possible to modify the exposed face of the α -helix to achieve interaction with a target [81,82] . Thus, we used mutagenesis to randomize residues in the α –helix residues that face outward and have no structural role, and three arginines were incorporated into the polyproline helix to enhance cell permeability (**Figure 25**). This arginine mutant of aPP was also shown to be more thermally stable than the original aPP protein, which might help to improve the quality of the library.

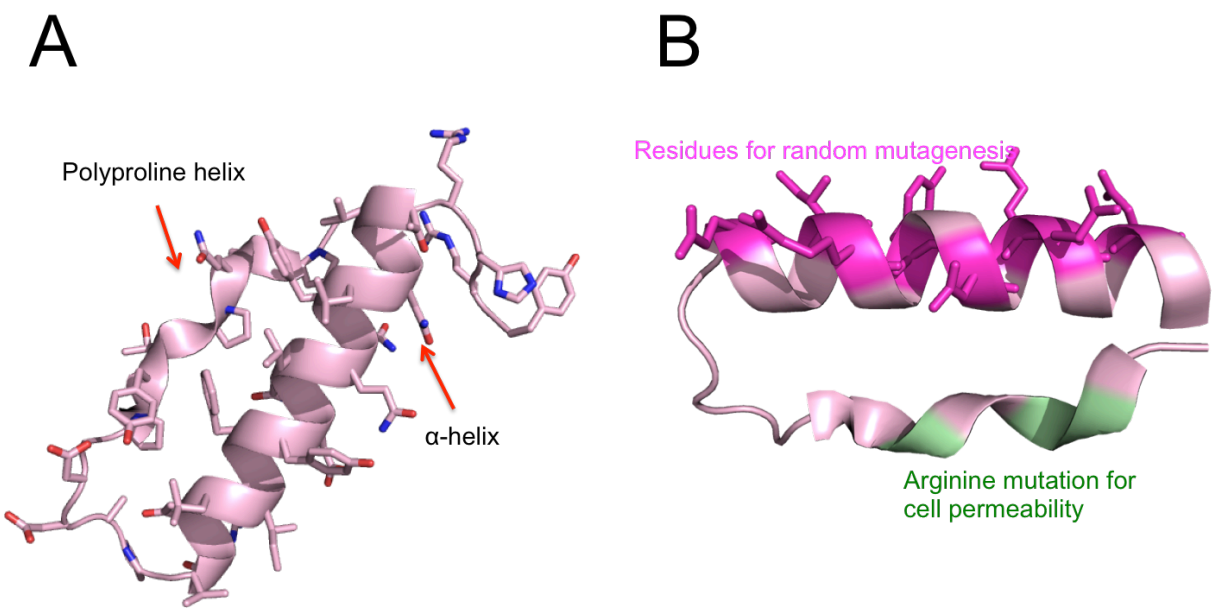


Figure 25. aPP scaffold for yeast display. A: Stabilized aPP with the hydrophobic residues in the core. B: the template used for yeast display. Residues in magenta are those that underwent random mutagensis, and residues in green have been mutated to arginine for possible aid in cell permeability (PDB: 1PPT)

B. Beta hairpin scaffold

A Beta hairpin template was also used with the intention of taking advantage of the tendency of Ras to bind partners through beta sheet complementation. The conformation of this scaffold was constrained by a disulfide bond and cross-strand pair residues that packed together to produce a structural stabilization effect [83,84] . The residues on one face of the beta hairpin were mutated to optimize packing, and the residues on the other face were randomized for target interaction (**Figure 26**).

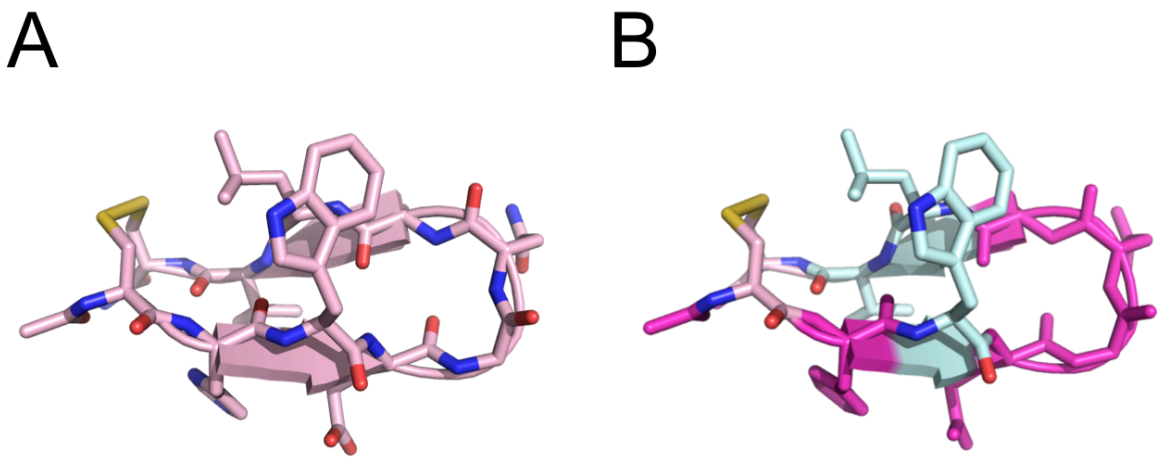


Figure 26. Beta hairpin scaffold for yeast display. A: Beta hairpin structure in cartoon representation. B: The residue pair in cyan color packs together, which is important for stabilization of the structure, along with the disulfide bond. Therefore, those residues were mutated to amino acids that would contribute to hydrophobic interaction. The residues in magenta color went through randomized mutation. (PDB: 1N0C)

C. PI3K scaffold

The Ras binding domain of PI3K, a well-known Ras effector, consists of a beta strand followed by an alpha helix that exhibits most of the interactions between Ras and PI3K [85] . The PI3K knuckle library was formed from random mutagenesis using error-prone PCR. (**Figure 27**)

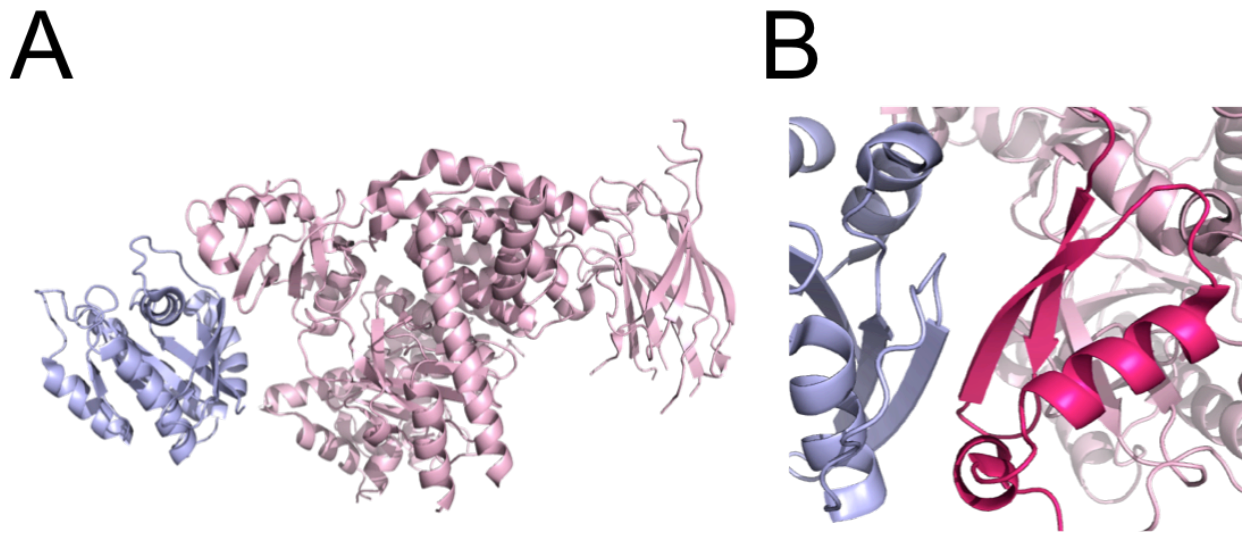


Figure 27. PI3K for yeast display scaffold. A: Interaction between Ras(light purple)and PI3K(light pink). B: The binding domain of the PI3K and Ras has been enlarged. The Ras binding domain residues of PI3K that were used as a scaffold for yeast display are colored in hot pink. (PDB:1HE8)

D. Scyllatoxin scaffold

The scyllatoxin peptide shows a structure similar to that of the abovementioned PI3K scaffold, with a N-terminal helix followed by an antiparallel beta-strand. The fold seems to be stabilized by three disulfide bonds [86,87] . In the scyllatoxin scaffold, residues from each face of the scaffold were mutated for binding to Ras. (**Figure 28**)

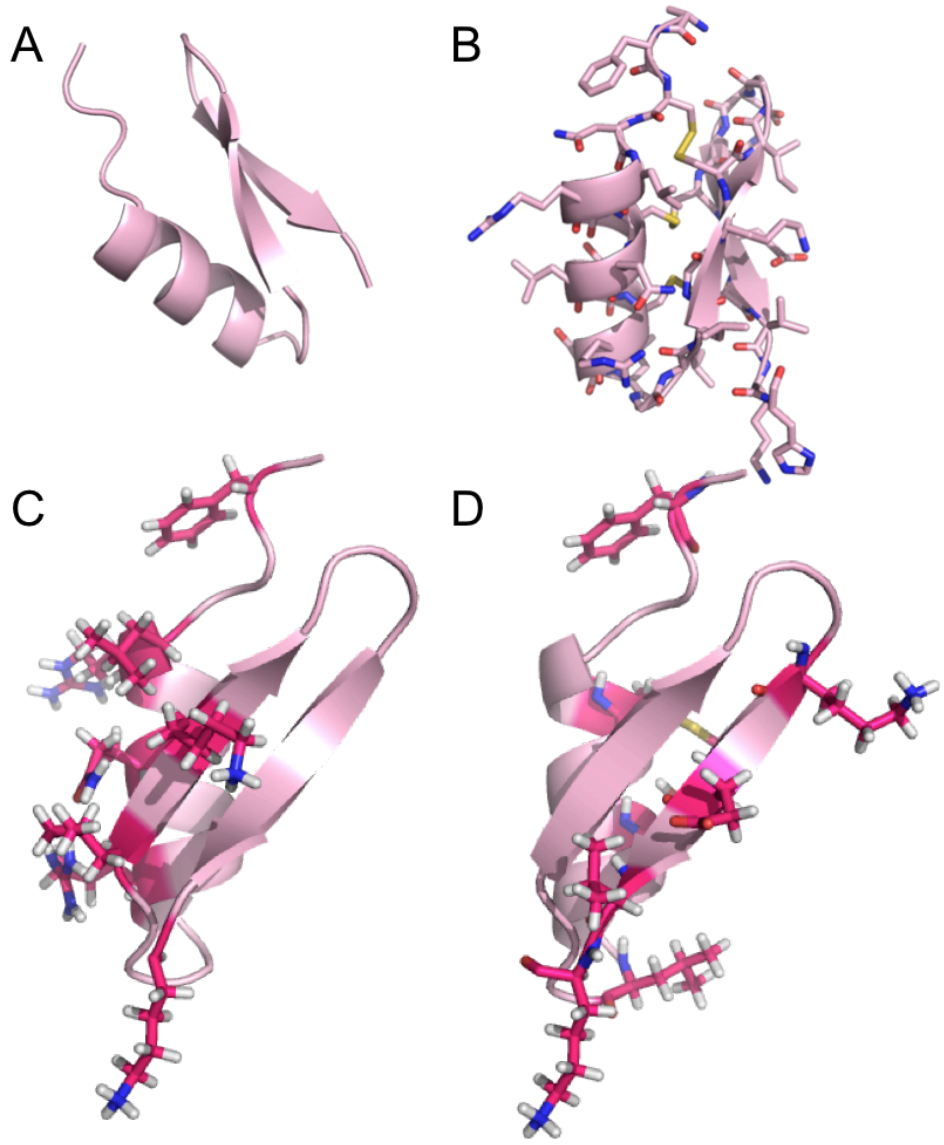


Figure 28. Scyllatoxin scaffold for yeast display. A: The scyllatoxin sequence consists of an alpha helix followed by a beta-hairpin. B. There are three disulfide bonds in the structure that seem to aid in the stabilization of the scaffold. C,D: Two opposite faces of the scaffold were mutated for binding to Ras. (PDB:1SCY)

2-3. Library generation & initial screens

The DNA library containing the scaffold sequence of interest was linked to the linearized pCTCON2 vectors via homologous recombination and transformed into yeast strain EBY100[88] . (**Figure 29**) The proteins were fused to the C-terminus of Aga2P, the expression of which was under the control of galactose-inducible promoter.

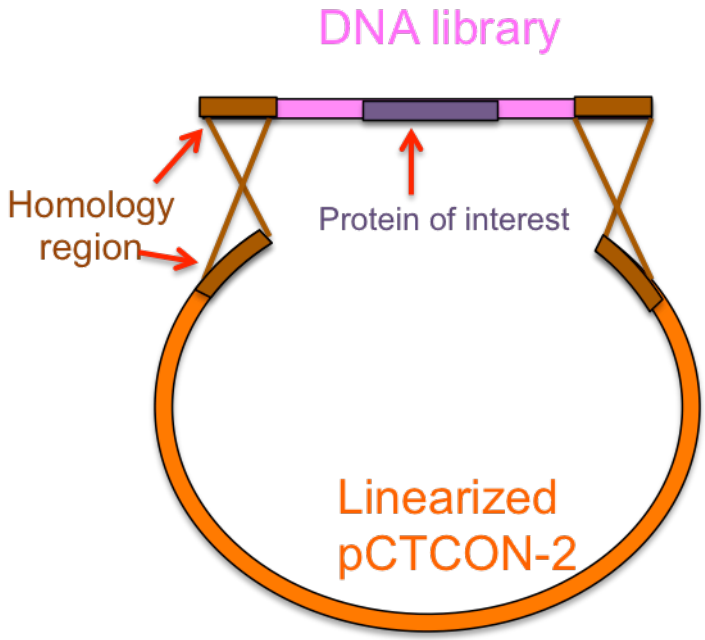


Figure 29. Plasmid library for yeast surface display system. The linearized DNA library fuses with linearized pCTCON-2 vector via homologous recombination.

The transformed yeast cells were grown in SDCAA media, where they used glucose as a carbon source to repress the expression of the heterologous protein in case of toxicity to the cells. For induction of the Aga2P protein fusion on the cell surface, galactose-containing media was used.

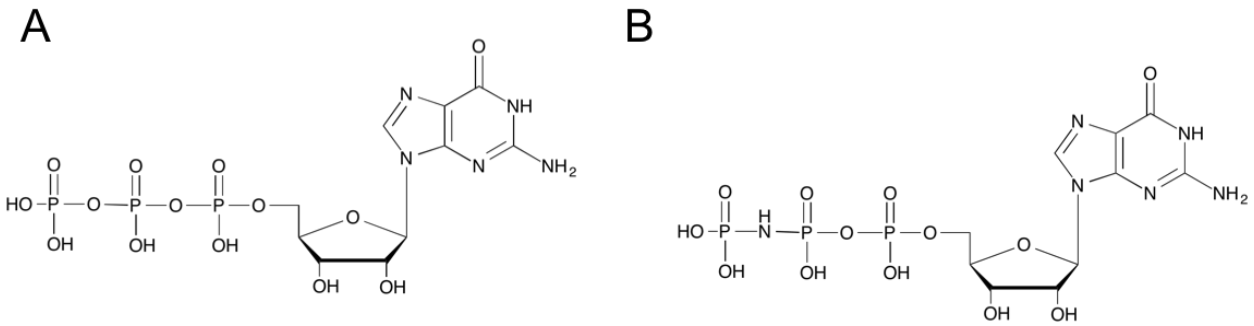


Figure 30. GTP and its analogue. A: Guanosine triphosphate B: 5'-Guanylyl imidodiphosphate(GppNHp). Also called GMP-PNP

The yeast libraries were screened for binding to the Ras G12V mutant with GppNHP, a non-hydrolyzable GTP analogue where one of the oxygen atoms is replaced by an amine. (**Figure 30**) This Ras protein was biotinylated so that it could be detected with PE fluorescence-labeled avidin. Streptavidin-allophycocyanin (SA-APC) and neutravidin-phycoerythrin (NA-PE) were used in alternation in order to prevent the enrichment of peptides that bind to them. The HA epitope tag at the N-terminal of the peptides was labeled with FITC to measure the expression level of the peptides on the yeast cells. The size of each initial library was approximately 10^7-10^8. In order to maximize the probability of isolating rare clones, any yeast display library should be oversampled by 10-fold.[75] As typical FACS sorting rates are 10^3~10^4 cells/sec, we performed the initial screening for Ras binders using magnetic-activated cell sorting (MACS), which is capable of sorting 10^9 cells at a time. This was intended to ensure that all unique clones in the library would be screened. After the initial MACS screen using 3uM Ras, subsequent sorting was performed by FACS. Note that background fluorescence that is equivalent to that of the negative control (**Figure 31**) is shown in all FACS data. This negative peak might come from the cells in which the secretory apparatus is saturated and blocked with misfolded protein.[89] Fortunately, it is known that this unlabeled cell population does not seem to affect the library screening.[75]

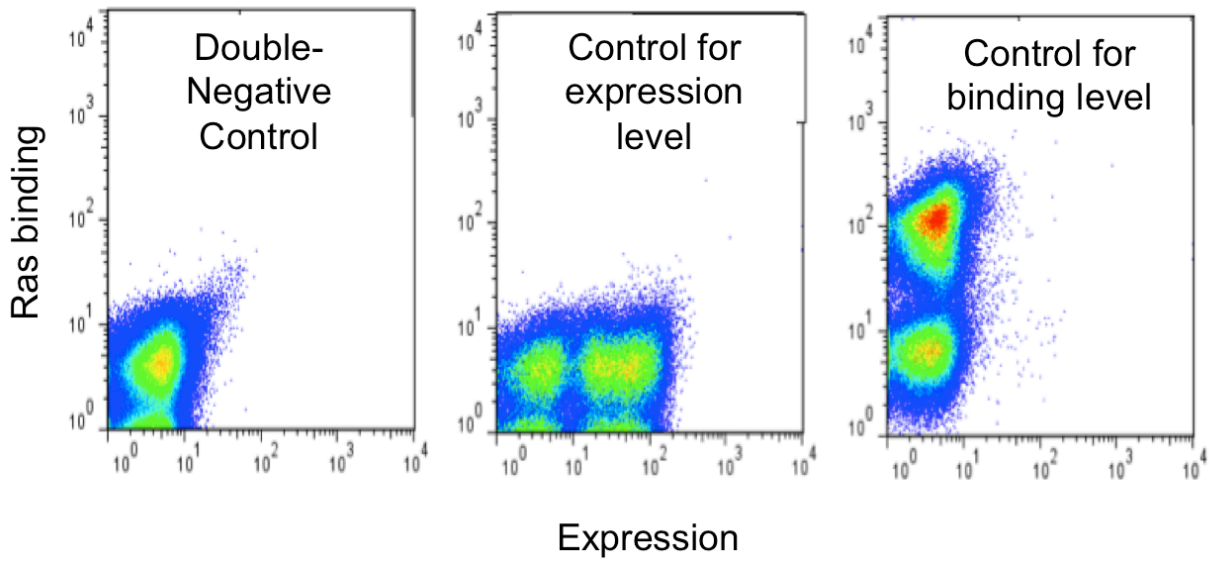


Figure 31. Negative controls in FACS sorting in yeast surface display experiment. Double negative control, control for surface expression level and control for binding level are shown from left to right.

In the first few rounds of FACS, sorting gates were set generously (~top 5%), whereas a more stringent gate (top 0.1~1%) was used for the following sorts. For the PI3K library, the protein expression level was expected to be low due to the instability of the peptide. Thus, for the first screening, the yeasts showing expression of the protein were initially collected using the FLAG-tag at the C-terminal of the protein and then analyzed for binding to Ras. After about total of seven rounds of sorting for each library, (**Figures 32-35**) most libraries began to show some enrichment.

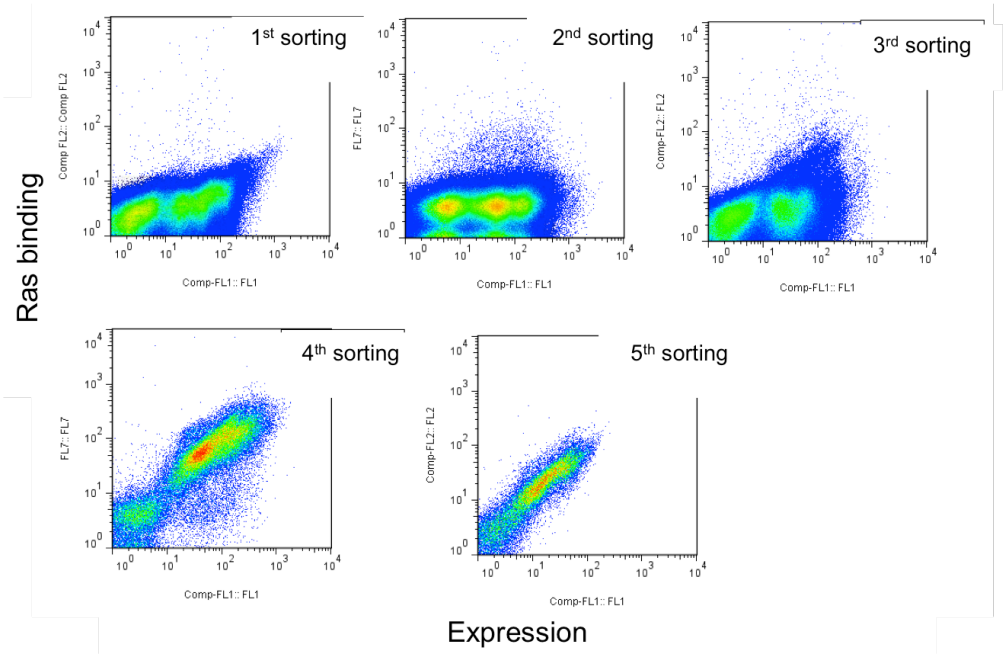


Figure 32. Screening of aPP library by FACS

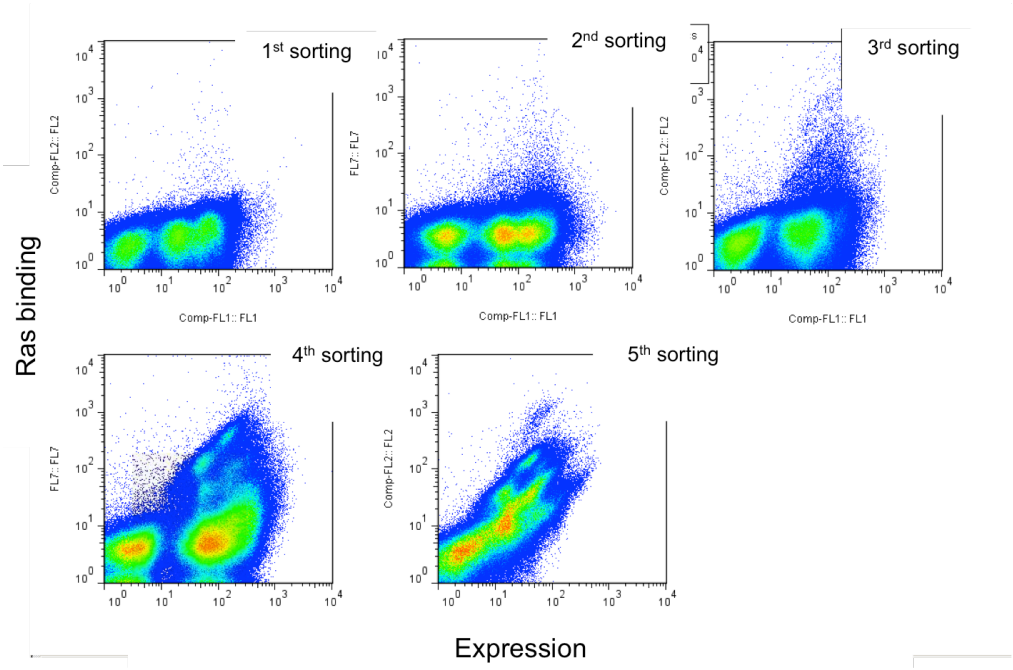


Figure 33. Screening of beta hairpin library by FACS

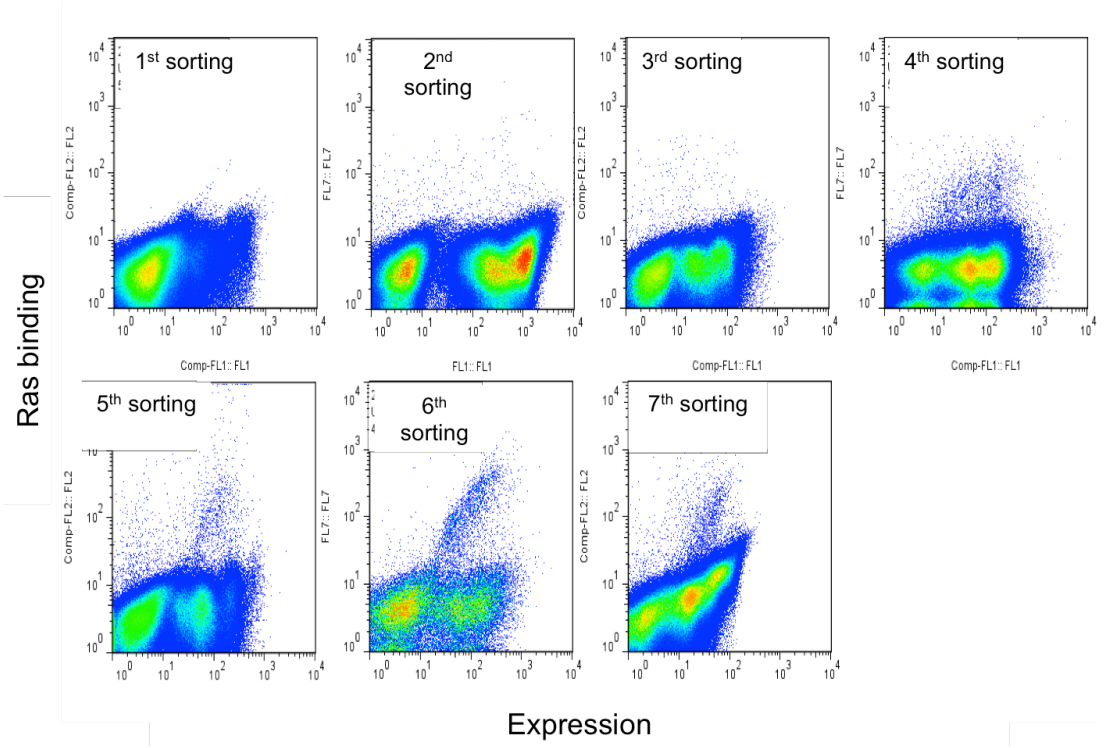


Figure 34. Screening of PI3K library by FACS

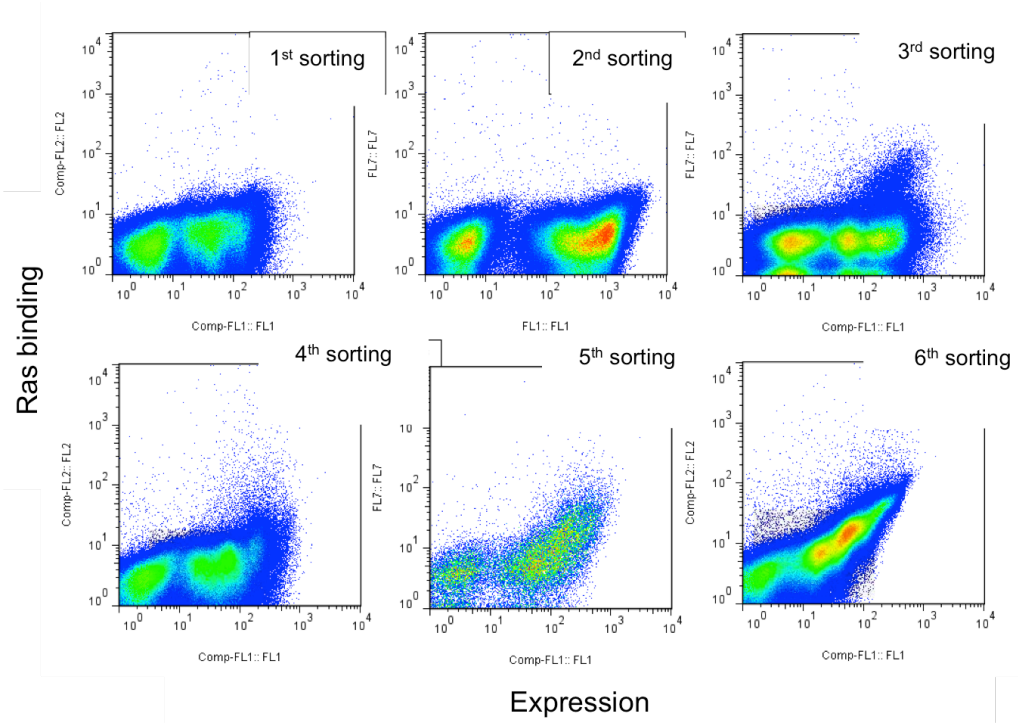


Figure 35. Screening of scyllatoxin library by FACS

To ensure that the enrichment was for Ras binders, a test FACS was performed with the same conditions but without the Ras protein. If the enrichment still existed even in the no Ras control, it indicated that the peptides might be secondary reagent binders. Another control FACS experiment was performed by incubating the cells with K-Ras that had been pre-mixed with a large excess of Raf, to see if the enriched peptides were competing with Raf for the binding site on Ras. If this were the case, it was expected that there would be a decrease in the level of binding signal. (**Figure 36**)

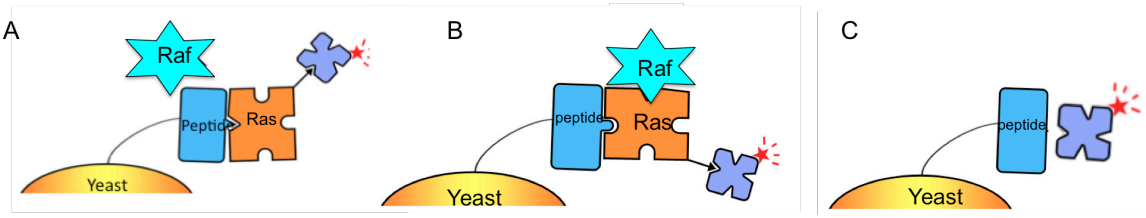


Figure 36. Control experiments. A: Peptide will compete with Raf if the binding site for Ras overlaps. B: If the peptide is binding to another region of Ras, the Raf will stay and there won't be decrease in the binding signal. C: if the binding signal does not diminish in the absence of Ras, it suggests that the peptide has been evolved as a secondary reagent binder.

From the results of the control experiments, we observed that the weak binding signal that was shown in the PI3K library (**Figure 37**) and Scyllatoxin library (**Figure 38**) stayed the same in the Ras-Raf preincubation control and the no Ras control experiment, which suggested the absence of real Ras binders. In the beta hairpin library (**Figure 39**), although the signal diminished in the no Ras control experiment, the binding level seemed to stay the same in Ras-Raf preincubation control experiment. This indicated that even if the enrichment was for peptides that bind to Ras, they might not be able to compete with the effectors. On the other hand, for the aPP library (**Figure 40**), the signals clearly diminished in both control experiments, which implied

enrichment for Ras binding peptides that can compete with Raf for binding to Ras. Therefore, we decided to pursue further the hits that came from the aPP library.

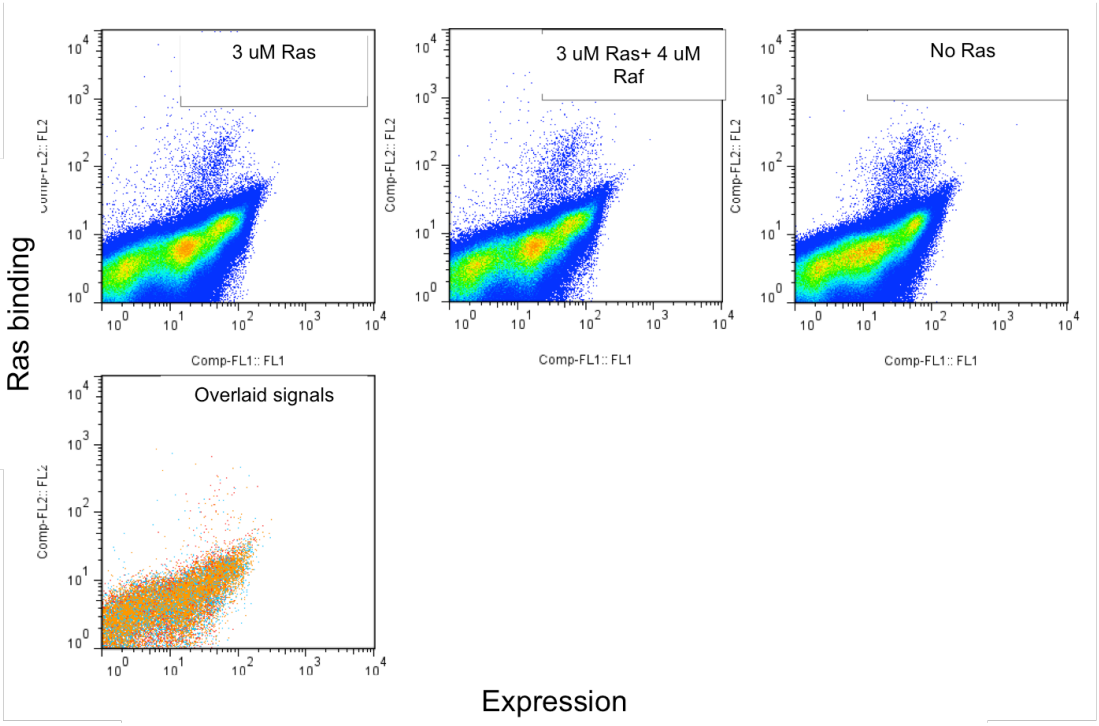


Figure 37. PI3K library control experiments

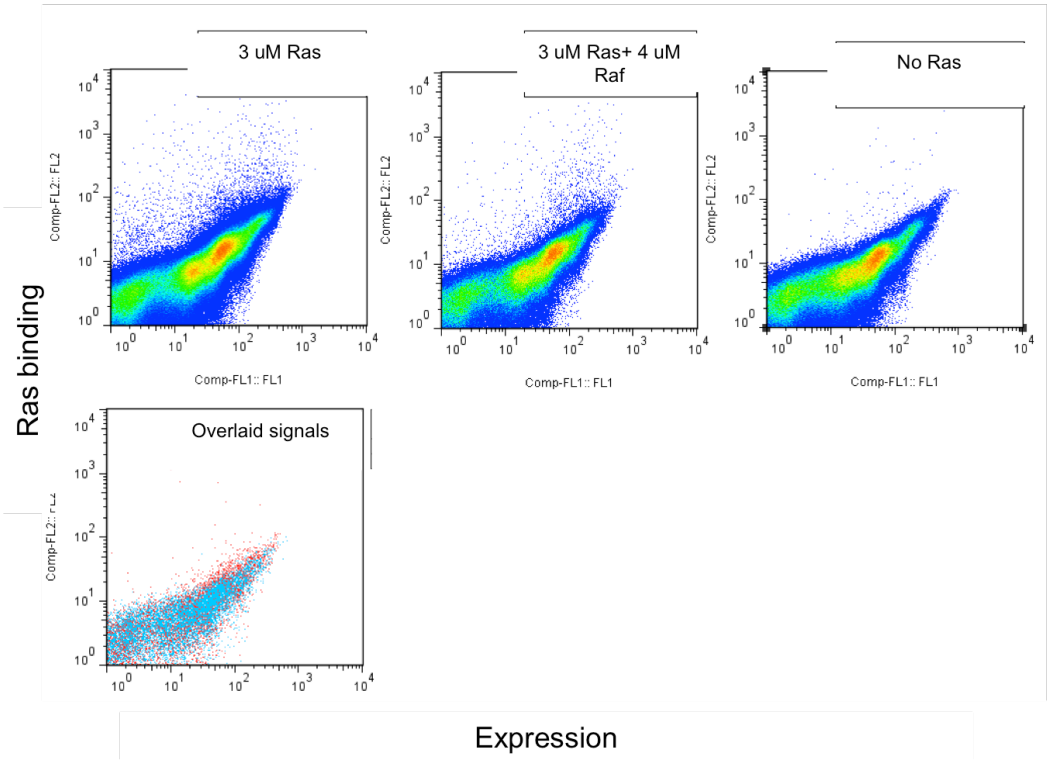


Figure 38. Scyllatoxin library control experiments

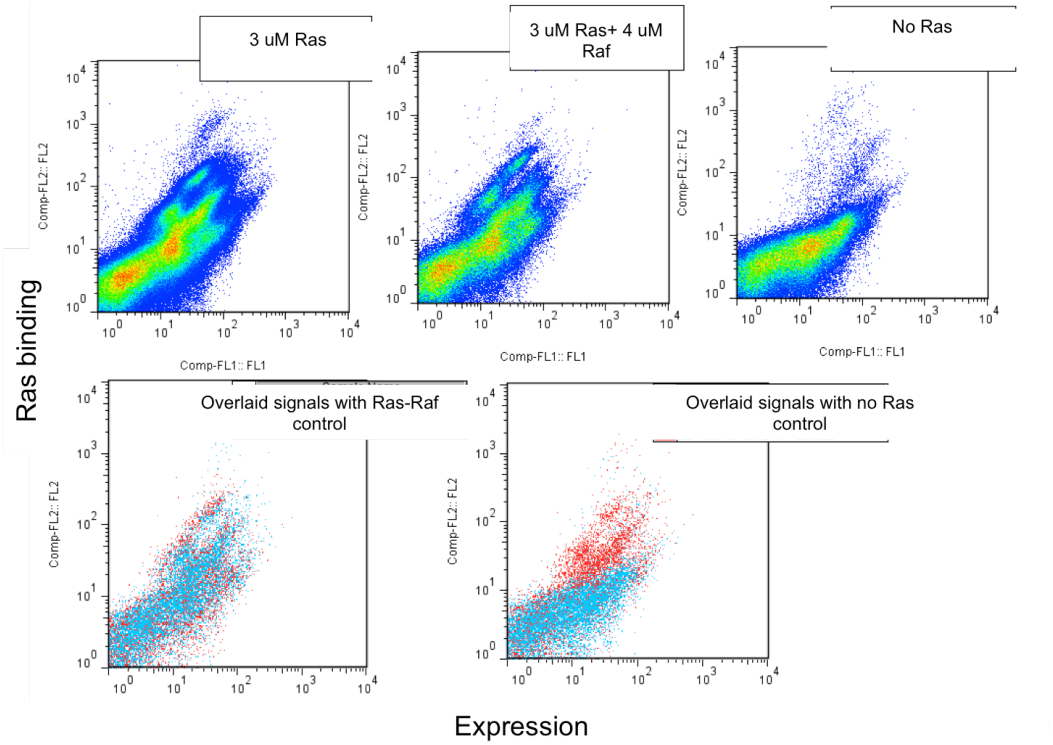


Figure 39. Beta hairpin library control experiments

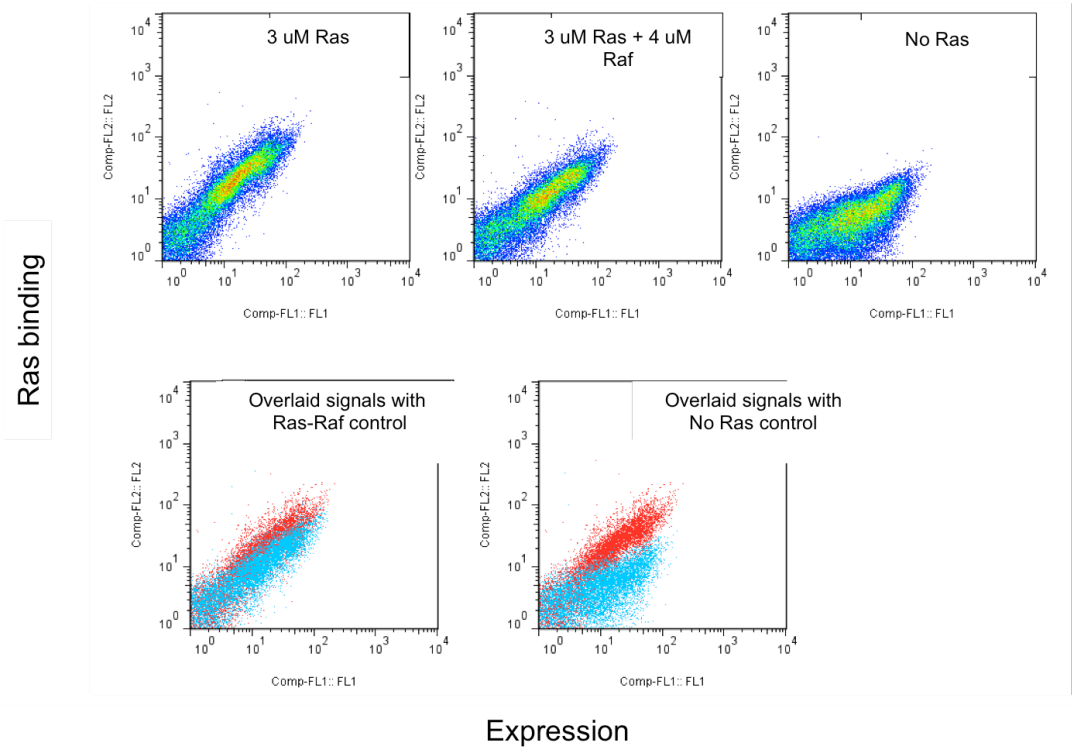


Figure 40. aPP library control experiments

2-4. Screening with aPP libraries

One of the benefits of the yeast display system is that with the genotype-phenotype linkage, it is easy to figure out the sequence of the binders. By sequencing the hits we got from the first aPP library (221 library), we saw that some residues had already strongly converged to specific amino acids. (**Figure 42**) For instance, position 21, 24, 28, 29 and 32 that were randomly mutated in the 221 library, has converged to single amino acid residue histidine, tryptophan, tryptophan, asparagine and tyrosine, respectively.

Using the hit sequences as another starting point, error-prone PCR was performed to introduce mutations into the sequence to generate a new library. The error-prone PCR method is often used to introduce high numbers of mutations into DNA sequences to create a combinatorial library based on a single gene [90] . The resulting library was named the 223 library. For the 223 library, the Ras concentration used for MACS was decreased to 100-250 nM compared to the 3 uM used for the 221 library screening. For subsequent FACS sorts, 10-50 nM Ras was used depending on the robustness of the enrichment. Additionally, 5% (v/v) human serum and 100 ug/ml salmon sperm single stranded DNA (ssDNA) were used as blocking agents to prevent non-specific binders that might come from hydrophobic interactions.

After a total of 7 rounds of sorting, we performed control experiments. In **Figure 41 A**, where the screening and sorting was done in 50 nM Ras, the enrichment of the peptides clearly decreased when there was no Ras. However, when the screening and sorting was done in 5nM Ras, enrichment was observed even in the absence of Ras as shown in **Figure 41 B**. This observation clearly indicated that the enrichment was resulting from the selection of secondary reagent binders. Therefore, even for the same library, it was important to carefully choose the concentration of the target protein in order to ensure selection for target binders. In **Figure 41 C**, a 300nM Ras-5 uM Raf preincubation experiment showed a shift downward of the enrichment, implying that the enriched peptides were competing for the effector binding domain. **Figure 41 D** shows a similar experiment in which GFP-labeled Raf was used. Since the GFP is detected in the same fluorescence channel as FITC, we chose the GFP fluorescence channel and APC channel to look at the binding levels to Raf and Ras, respectively. Again, we saw that the binding

to Ras diminished upon Raf preincubation, and the lack of a GFP signal confirmed that the peptides were not binding to Raf.

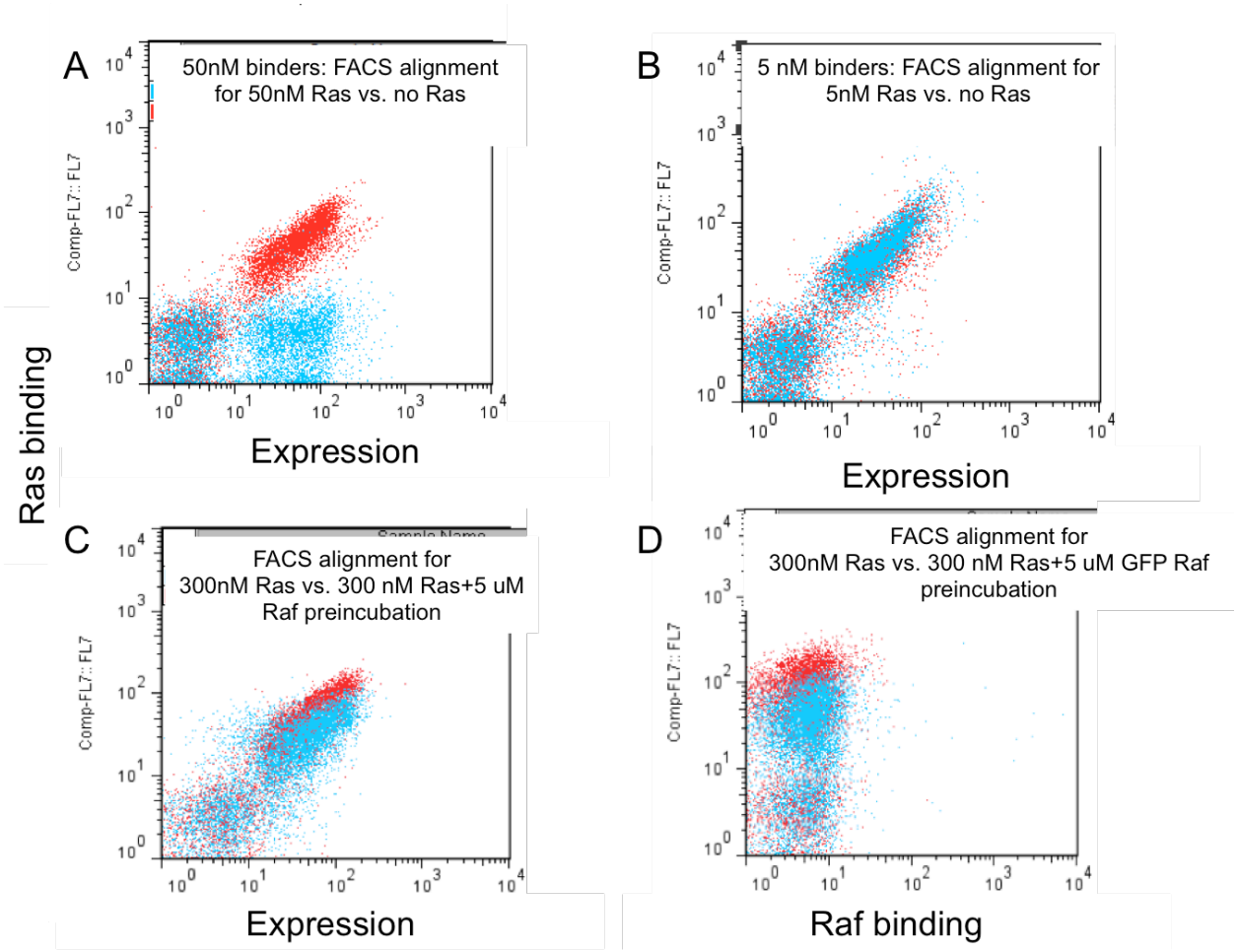


Figure 41. Control FACS experiments for the 223 library. red plots: FACS experiments with yeast cells incubated with designated concentration of Ras. Blue plots: no Ras (A & B) or 300 nM Ras+5 uM Raf preincubation (C & D). Note that in D, GFP-Raf was used, thus we looked at the Ras binding signal and Raf binding signal in FACS.

The hits from the 223 library were mini prepped and sequenced, and the results are summarized in **Figure 42**. Here we see that the position 16 and 18 has converged to serine and glutamate respectively, and new residues such as arginine for position 26 and alanine for the position 33 appeared.

Using each of these sequences as templates, we conducted another error-prone PCR to put mutations into the hit sequences and make new libraries for the next round of screening and sorting. The 224A library is based on the 223-1 sequence, while the 224B library is based on the 223-2 sequence.

For the screening of the 224 library, we dropped the Ras concentration down to 25 nM for the initial screening, and further went down to 5 nM. After 7 to 10 rounds of sorting, we saw that the 224A library converged to one sequence, 224A-1, and the 224B library converged mostly to 224B-1 although there were a few 224B-2 sequences that have PRR insertion in the beginning of the sequences. Four more subsequent screening rounds with the 224B library with a lower Ras concentration of 5 nM yielded only 224B-1, suggesting that 224B-1 might be a stronger binder than 224B-2. In 224A-1, there were Y10C and F23Y mutations that weren't seen in the previous sequences. In 224B-1, we again observed the Y10C mutation, and although 223-2 sequence was the template for 224B library, we can see that the residues in position 22 and 26 converged to those that were already evolved in the 223-1 sequence, suggesting that those residues contribute more to Ras binding.

We set up two strategies for the next libraries. The first one, 225 library, was to combine and shuffle all the different amino acids that came out in the 224 library hit sequences without further mutation. In case the cysteine mutation in the 10^{th} residue was coming from an artifact such as forming disulfide bond with one of the cysteine residues in Ras, we used tyrosine instead of cysteine at that position in the library. For the second one, 226 library, we used error-prone PCR in 225 library to introduce mutations into the sequence.

After three rounds of screening at 100 nM Ras, the 225 library mostly converged to one sequence, 225-1. After five rounds of sorting beginning at 100 nM Ras and tapering to 20 nM Ras, the 226 library showed mostly peptides with amino acid combinations that were already shown in the previous library hits. The dominant residues were the same ones as in 225-1, such as glutamate at position 22, tyrosine at position 23 and isoleucine or valine at the position 17. A dramatic change was revealed in position 10, where again cysteine emerged dominantly over the original residue tyrosine even though it was not in the original amino acid mixtures. The results of the evolution will be explained further with the structures of the complex of Ras (G12V) GppNHp and the peptides in the following section.

Clone	Sequence																															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Original aPP				G	P	S	Q	P	T	Y	P	G	D	D	A	P	V	E	D	L	I	R	F	Y	D	N	L	Q	Q	Y	L	N	V	V	A
Library template				G	P	R	R	P	R	Y	P	G	D	D	A	P	V	X	D	L	X	X	F	X	A	X	L	X	X	Y	L	X	V	V	A
221-1				G	P	R	R	P	R	Y	P	G	D	D	A	P	V	E	D	L	H	A	F	W	A	A	L	W	N	Y	L	Y	V	V	A
221-2				G	P	R	R	P	R	Y	P	G	D	D	A	P	V	G	D	L	H	E	F	W	A	Q	L	W	N	Y	L	Y	V	V	A
221-3				G	P	R	R	P	K	Y	P	G	D	D	A	P	V	E	D	L	H	W	F	W	A	A	L	W	N	Y	L	Y	V	V	A
223-1				G	P	R	R	P	R	Y	P	G	D	D	A	S	V	E	D	L	H	E	F	W	A	R	L	W	N	Y	L	Y	A	V	A
223-2				G	P	R	R	P	R	Y	P	G	D	D	A	S	I	E	D	L	H	A	F	W	A	A	L	W	N	Y	L	Y	A	V	A
224A-1				G	P	R	R	P	R	C	P	G	D	D	A	S	V	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
224B-1				G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	H	E	F	W	A	R	L	W	N	Y	L	Y	A	V	A
224B-2	G	P	R	R	P	R	R	P	R	Y	P	G	D	D	A	S	I	E	D	L	H	A	F	W	A	A	L	W	N	Y	L	Y	A	V	A
225-1				G	P	R	R	P	R	Y	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
225-2				G	P	R	R	P	R	Y	P	G	D	D	A	S	I	E	D	L	H	A	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
225-3				G	P	R	R	P	R	Y	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	P	V	A
226-1				G	P	R	R	P	R	Y	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
226-2				G	P	R	R	P	R	C	P	G	D	D	A	S	V	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
226-3				G	P	R	R	P	R	Y	P	G	D	D	A	S	V	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
226-4				G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	H	E	F	W	A	R	L	W	N	Y	L	Y	A	V	A
226-5				G	P	R	R	P	R	C	P	G	D	D	A	S	T	E	D	L	H	A	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
226-6				G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
226-7				G	P	R	R	P	R	C	P	G	D	D	A	S	V	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	A	V	A
226-8				G	P	R	R	P	R	C	P	G	D	D	A	S	V	E	D	L	H	E	Y	W	A	K	L	W	N	Y	L	Y	A	V	A

Figure 42. aPP library hit sequences. X: site of directed mutagenesis. Red color used to show the residues that came out from the directed evolution screens, and the ones that have converged to a strong consensus have been highlighted. The green color indicates the residues that have been shuffled from a set of amino acids that came out from previous rounds. (This particular numbering system was used to be consistent with that of the structural data)

Chapter 3. RDA (Ras binding Dimeric aPP) peptides

3-1. Structural analysis

We decided to use the most dominant sequence from the evolution, 226-6, as well as 225-1 that has the same sequence except for cysteine at position 10 instead of tyrosine, for binding affinity studies and structural studies. The initial FP and pull-down experiments using 225-1 performed by John McGee verified that the peptides bind to Ras with mid-low nM affinity *in vitro* and also that they are capable of binding to endogenous Ras from cancer cell lysates and competing with Raf for Ras binding.

For use in biochemical studies and structural studies, the hit peptides were prepared by bacterial expression, which was less time consuming and cheaper than SPPS. The following sequences were expressed and purified:

RDA1: GSGGPRRPRCPGDDASIEDLHEYWARLWNYLYAVA (226-6)

RDA2: GSGGPRRPRCPGDDASIEDLHEYWARLWNYLYRVA

The peptides were named "Ras binding Dimeric aPP" (RDA) peptides, as the structures that were solved later showed that the high-affinity Ras-interacting species was a covalent dimer linked via the cysteines that arose at the 10^{th} position during the directed evolution. The RDA2 peptide is another hit peptide that was evolved in a yeast display experiment by John McGee to favor binding Ras-GTP over Ras-GDP. The GSG linker at the beginning of the sequences was put in order to facilitate proteolysis at the TEV site that comes directly before it.

Briefly, the RDA peptides were subcloned into a pET expression vector with an N-terminal His_6 tag followed by TEV cleavage site. Then the construct was transformed into BL21 Rosetta cells, and the peptide expression was induced with IPTG. In order to increase the yield, molecular chaperone was co-expressed using an arabinose-inducible promoter. Afterward, the cells were harvested, lysed by sonication, and then the His_6 tag was used to purify the peptides from the cell lysate using either Nickel or Cobalt immobilized metal affinity chromatography. Then the tag was cleaved off by TEV overnight. After TEV cleavage, the peptide was desalted and purified by reverse-phase HPLC, then lyophilized. As a final step, the peptides were dissolved in DMSO and quantified. FP experiments (**Figure 43**) showed that the K_d of the RDA peptides for K-Ras (G12V) are in low nM range (~8-17 nM).

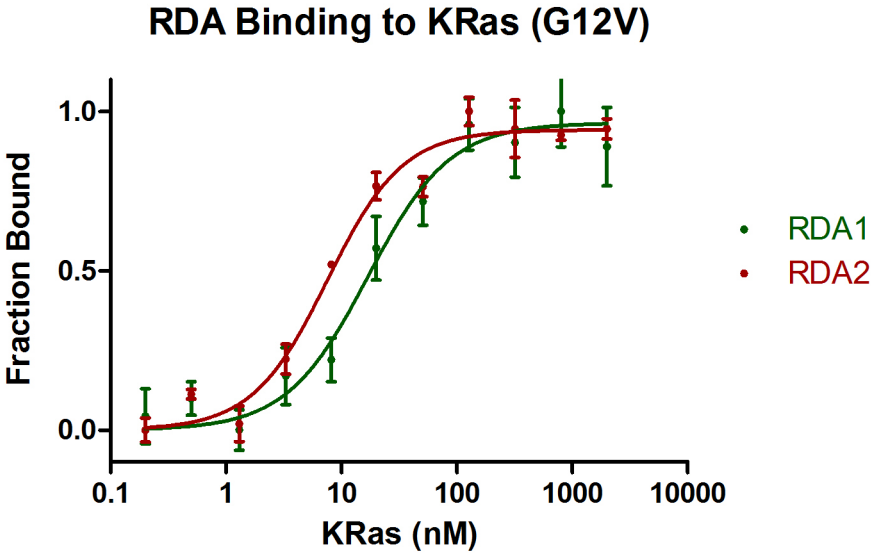


Figure 43. Fluorescence polarization (FP) measurements of the RDA peptides binding to K-Ras (G12V). The Kd for RDA1 for Ras is 17 nM with hill coefficient of 1.2, and the Kd of RDA2 is 8 nM with hill coefficient of 1.3. The FP experiment was conducted by John McGee.

The RDA peptides and K-Ras (G12) GppNHp complex was crystallized and the x-ray structure was solved by Dr. Matt Lee in the Verdine group. As expected, the overall structure shows that the RDA peptides bind to Ras in the region that overlaps with the effector domain (**Figure 44**), which explains the competition between Raf and the RDA peptides for the binding site in Ras from the yeast display experiments. While RalGDS or Raf interacts with mostly switch I region of Ras, the RDA peptides seem to bind to both switch I and switch II regions.

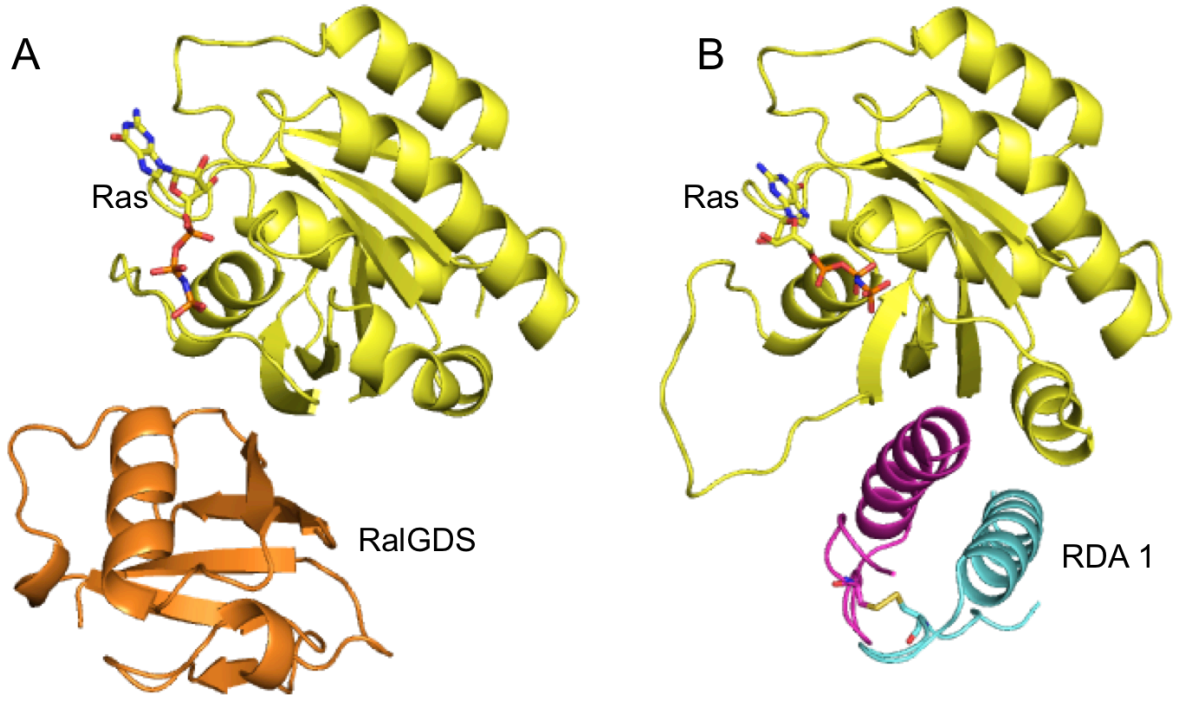


Figure 44. K-Ras (G12V) GppNHp in complex with Ral GDS (PDB:1LFD) and RDA peptides. A: Ras protein (yellow) binding to its effector RalGDS (orange). B: Ras protein binding to RDA1 peptides. The RDA1 peptides are in a homodimeric form, the first monomer (pink) and the second monomer (blue) linked by covalent disulfide bond (represented in stick form)

In addition, the peptides bind to Ras in a homodimeric form, covalently linked by a disulfide bond. This explains why the cysteines were continually found in position 10 during the directed evolution, as the covalent bond would stabilize the dimer and increase the affinity to Ras. The observation that the evolved RDA peptides dimerize is not entirely unexpected, as the original

aPP peptide that the RDA peptides are derived from have previously been reported to form head-to-tail homodimers [91] and visualized as such in x-ray crystal structures [81].

Looking into the details of the structure, we can see that from the very first evolution round, the peptides have acquired residues that are important both for their binding to Ras as well as forming a stable dimer. The residue hits from the 221 library, His21, Asn29 and Tyr32 are important for binding to Ras while Trp24 and Trp28 are important for both binding to Ras and stabilizing the dimer via π stacking interactions. First of all, we see that the nitrogen atom in the imidazole of His21 from the first RDA monomer is forming salt bridges with the carboxylates of the Glu3 of Ras. In addition, the carbonyl group of His21 is also forming an indirect water mediated interaction with Glu3 of Ras (**Figure 45**). The imidazole group of His21 on the second monomer is also interacting with Ras, by forming hydrogen bonds with the OH groups of the side chains of Gln 70 and Tyr71 of Ras (**Figure 46**). The side chain of Asn 29 from the first RDA monomer forms hydrogen bonds with the carbonyl group of Ile55 of Ras and backbone amide nitrogen of Asp57(**Figure 47**). The backbone carbonyl from Tyr32 of the first RDA monomer interacts with guanidinium group of Arg68 of Ras, and OH group of the side chain of Tyr32 also indirectly interacts via water with carbonyl group of Ala59 of Ras (**Figure 48**).

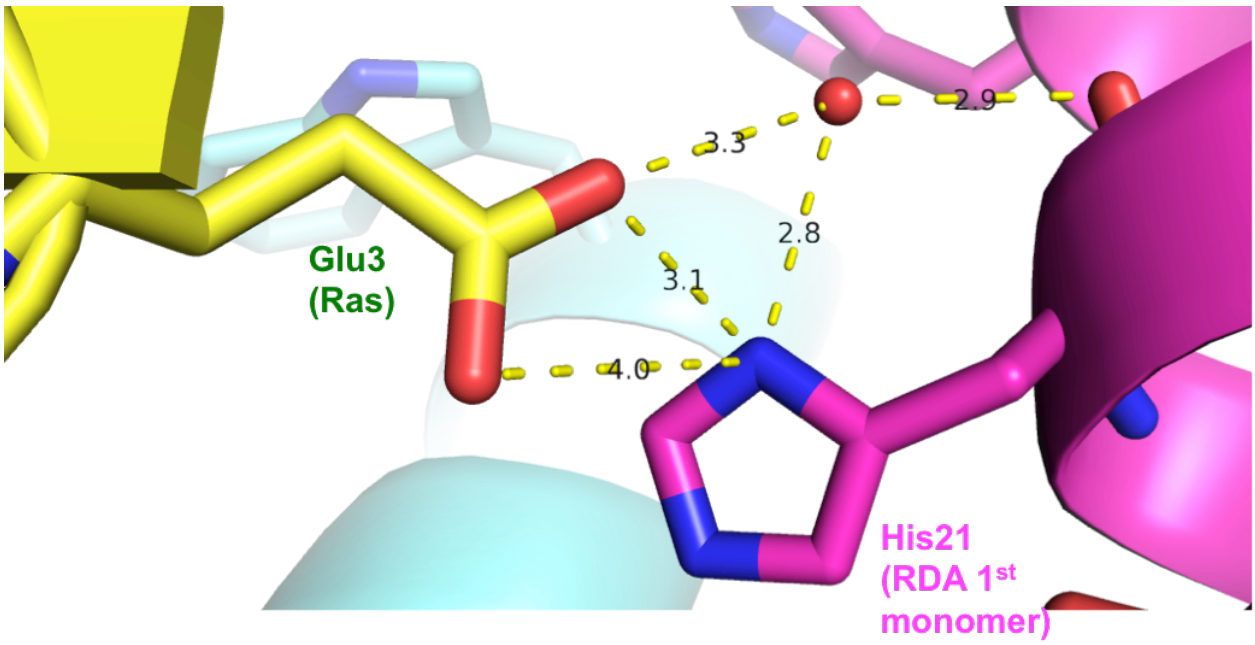


Figure 45. His 21 of the RDA1 first monomer

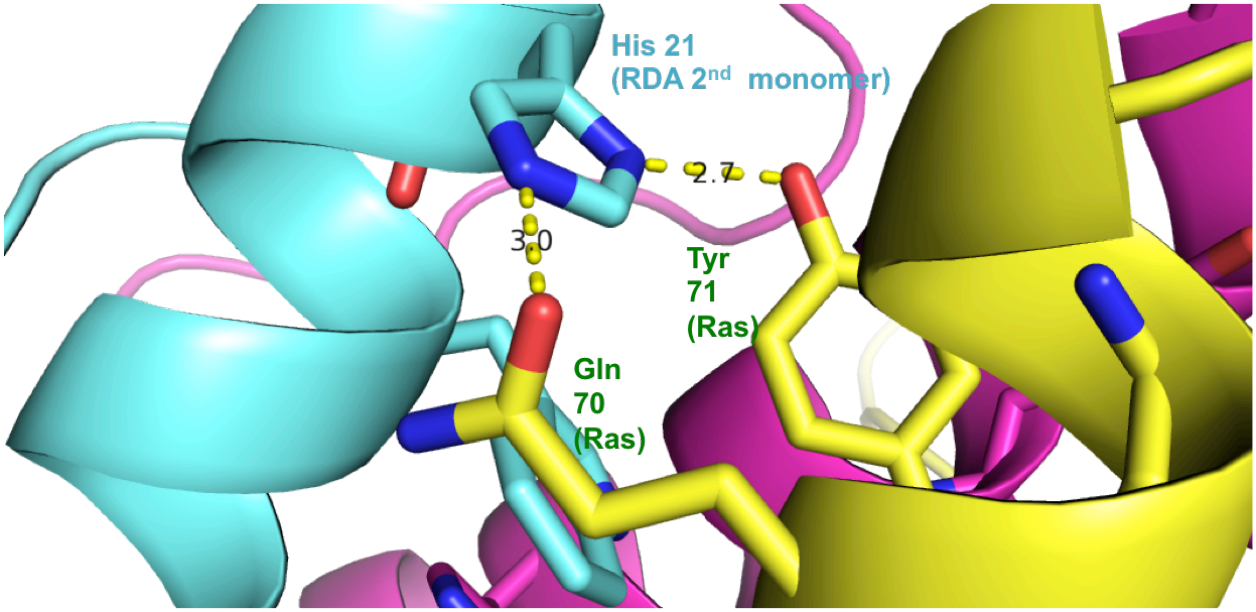


Figure 46. His21 of the RDA1 second monomer.

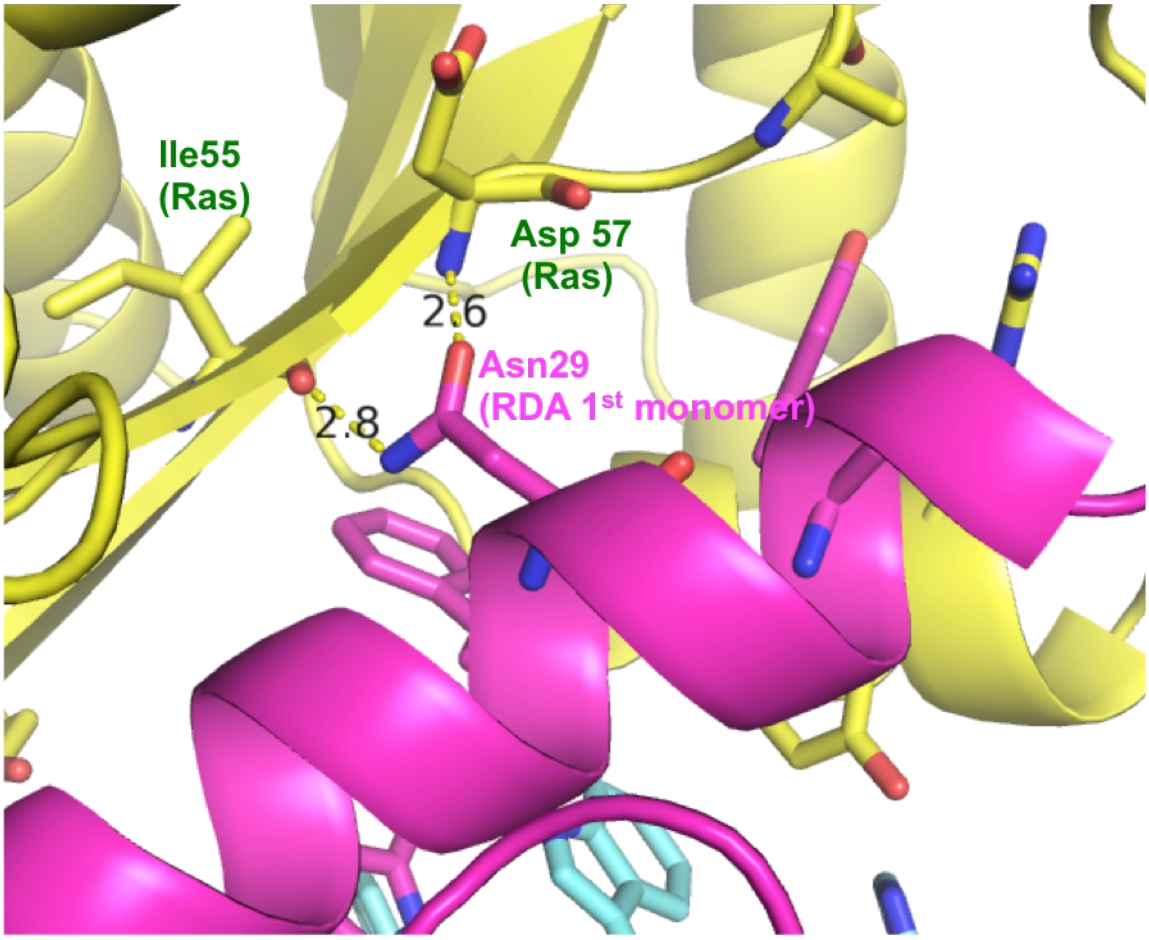


Figure 47. Asn29 of the RDA1 first monomer

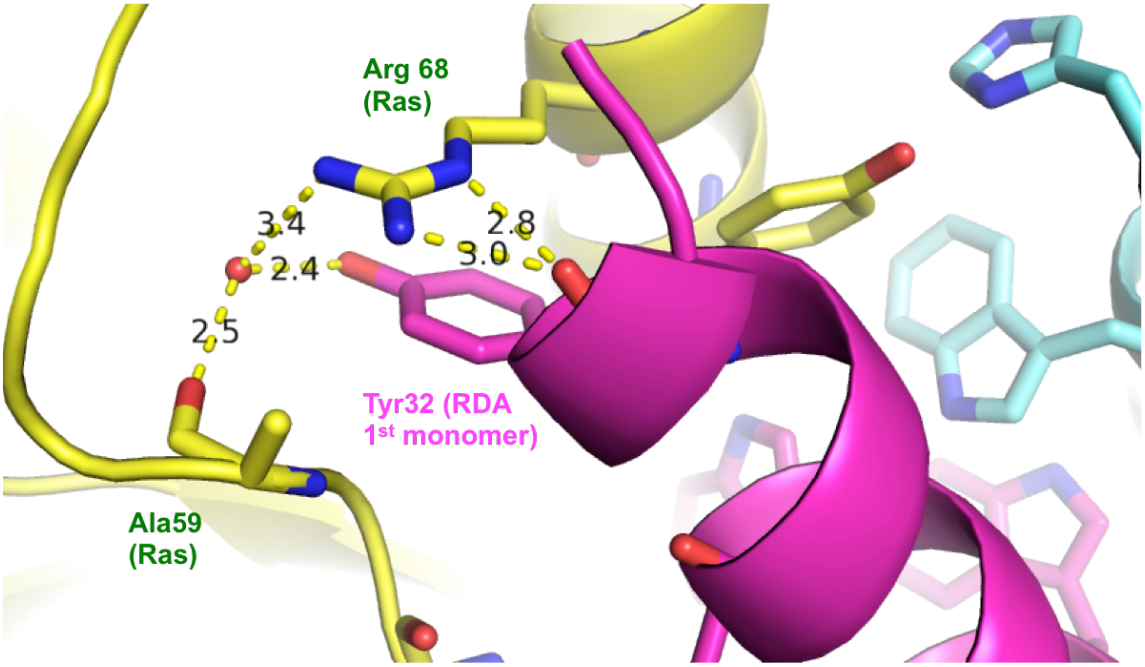


Figure 48. Tyr 32 of the RDA1 first monomer

The packing of the four tryptophan residues from the 24th and 28th positions in both monomers seems to contribute to the binding via hydrophobic interactions and also stabilizes the dimer by forming the core of the two peptides (**Figure 49**).

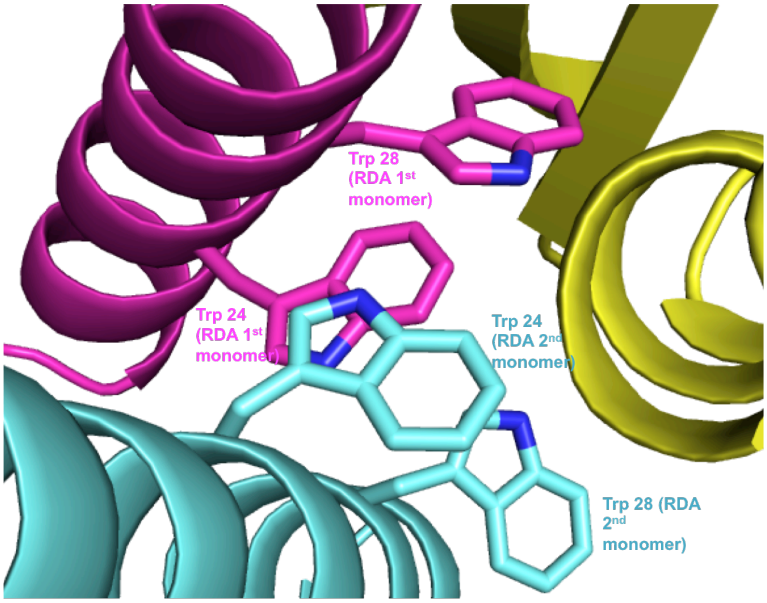


Figure 49. Trytophan residues in the RDA peptides

In hits from the 223 library, the hydroxyl group of serine side chain of Ser16 might be aiding in the stabilization of the monomer structure by forming a hydrogen bond with the NH group of Asp19, which is in *i, i*+3 position from it, and also by forming hydrogen bonds with the carboxylates of Asp19. These interactions are seen in both the first monomer and the second monomer, possibility contributing in helix stabilization (**Figure 50**). The carboxylates of Glu18 from the first RDA monomer forms a salt bridge with the guanidinium sidechain of Arg41 from Ras (**Figure 51**). It was hard to tell from the structure why evolution favored Ala33, since it was neither binding to Ras or contributing to dimer stabilization.

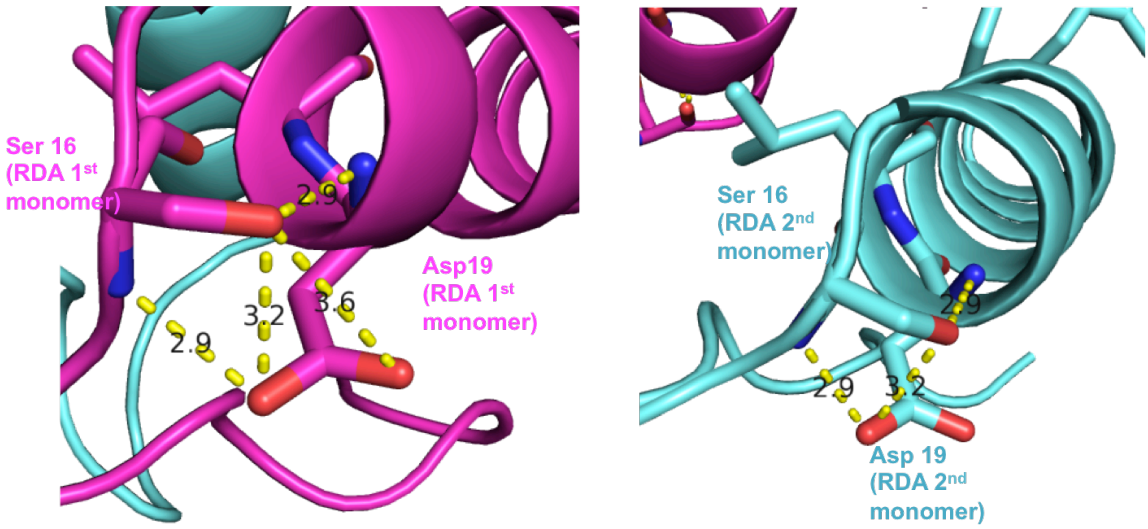


Figure 50. Ser16 in RDA first (left) and second (right) monomer

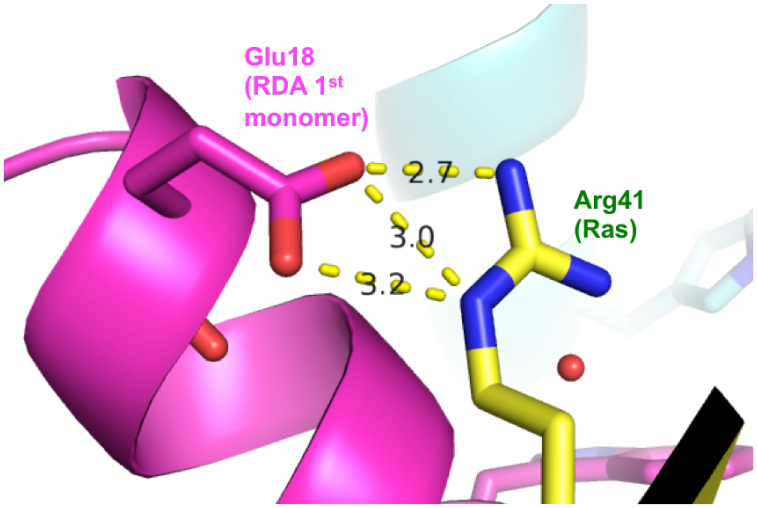


Figure 51. Glu18 of the RDA1 first monomer

Throughout the evolutions, glutamate seemed to dominate over alanine in position 22, which might be due to its hydrogen bond interaction with OH group of Ser39 of Ras, as well as water-mediated interaction with backbone amide nitrogen of Tyr40 of Ras (**Figure 52**). The position 26 of RDA peptides was converged to arginine, the guanidinium of which seems to be forming a

salt bridge with carboxylate of Glu 37 of Ras and a hydrogen bond with the main chain carbonyl group of Asp 38 of Ras (**Figure 53**).

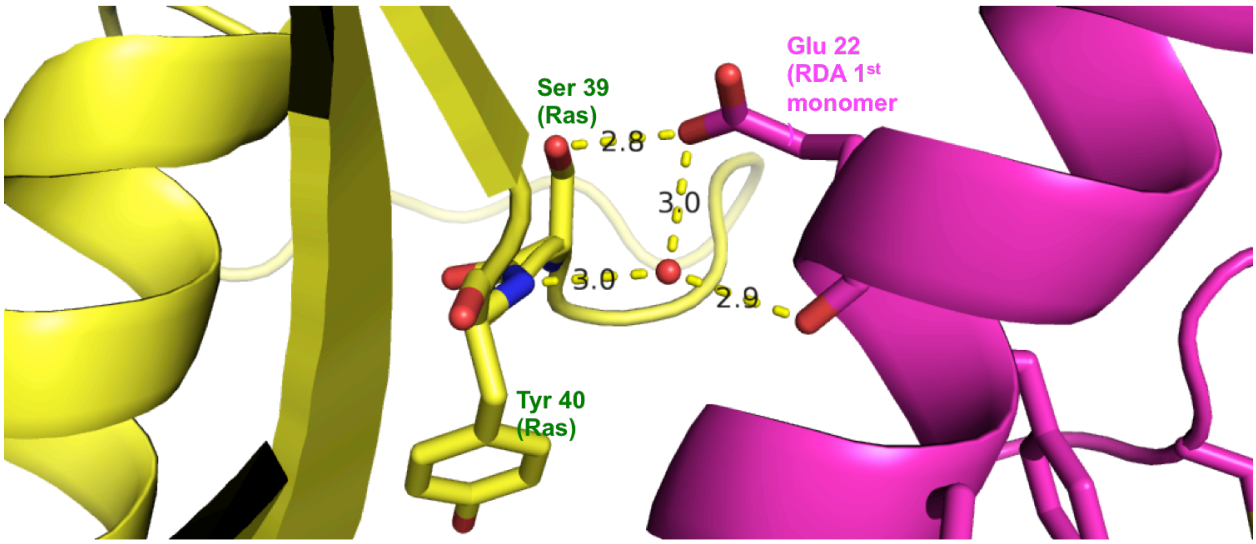


Figure 52. Glu 22 of the RDA1 first monomer

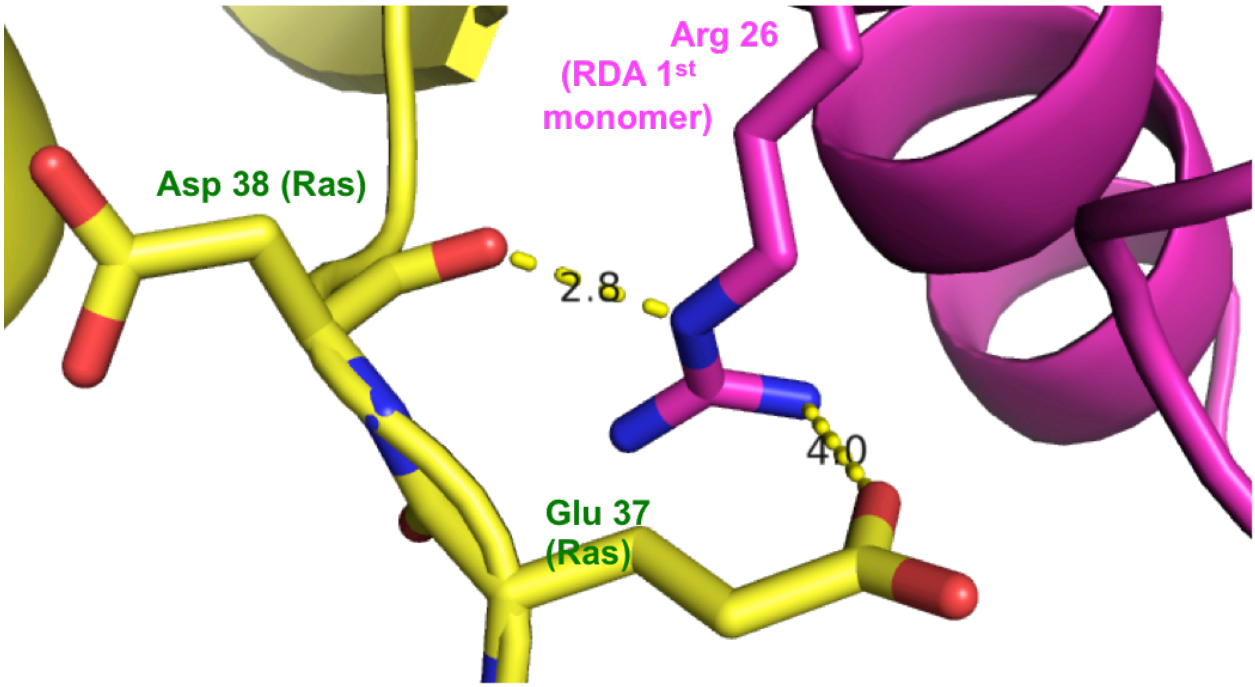


Figure 53. Arg 26 of RDA1 first monomer

As for the hits from the 224 library, both Cys10 and Tyr23 seem to be contributing to the dimer stabilization along with the tryptophan residues on the other side (**Figure 54**). The PRR extension that had emerged in the hits might also be due to the contribution to the stabilization of the structure rather than Ras binding, although the SAR for it was not investigated further.

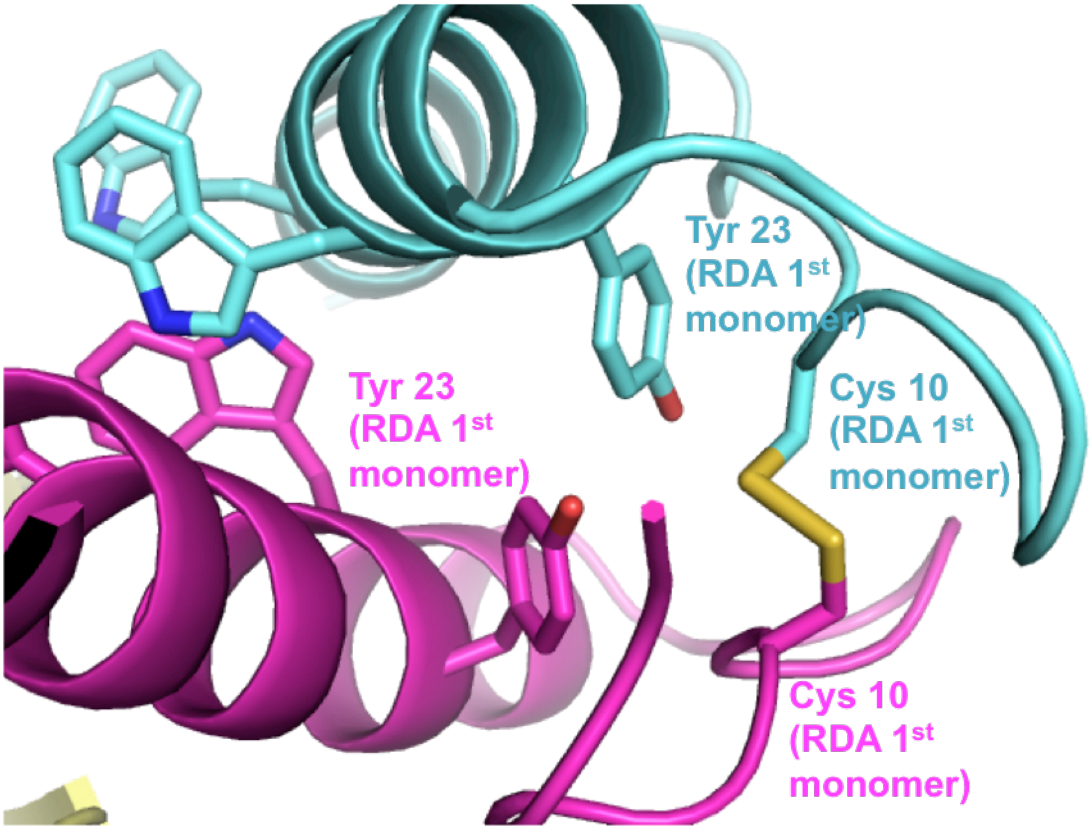


Figure 54. Tyr 23 and Cys 10 residues in the RDA1 peptides

The structure of the RDA2 bound Ras complex showed that the arginine that was mutated from alanine in position 33 has water-mediated interaction with the γ-phosphate of the GTP analogue as well as an extra hydrogen bond with the backbone carbonyl group of Asp 57 or Ras (**Figure 55**). This explains the higher affinity and selectivity for GTP-bound Ras over GDP-bound Ras of the RDA2 peptides.

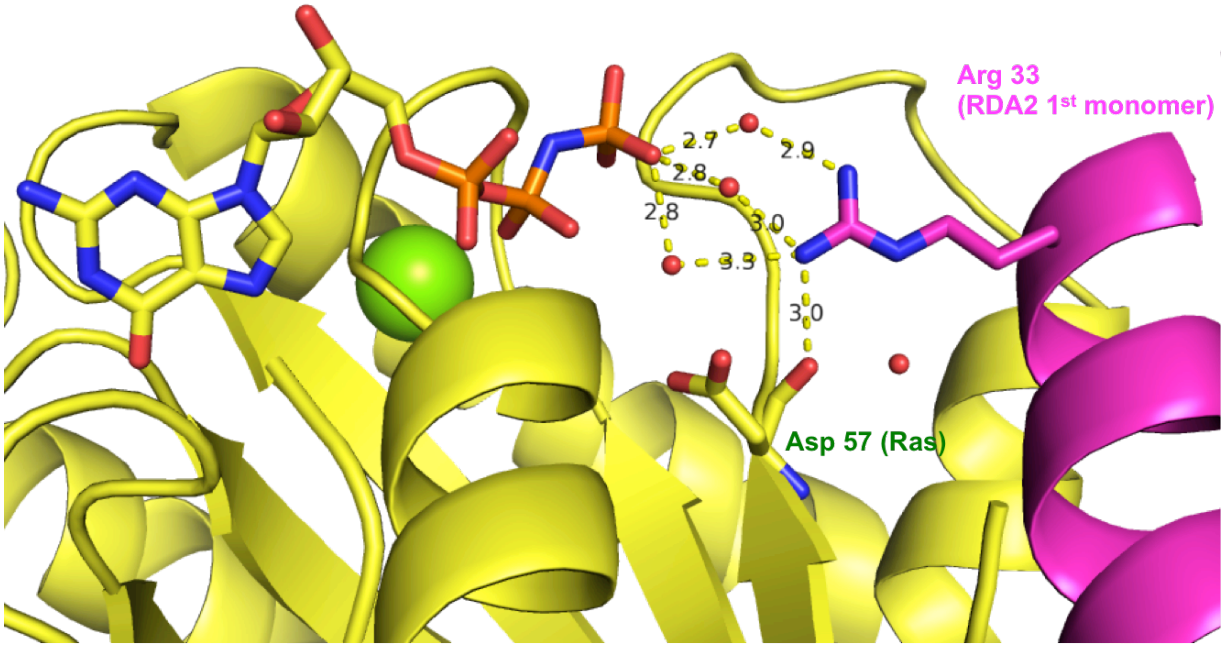


Figure 55. Ala33Arg mutation in RDA2. Arginine interaction with gamma-phosphate of GTP analogue, GppNHP is shown.

In summary, the residues that are important for binding and/or dimer interaction are highlighted as below.

Table 10. Residues in RDA peptides that are important in binding to Ras and/or dimer stabilization. The residues important for Ras binding are highlighted in yellow, for dimer interaction in blue, and for both in magenta.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1st monomer	G	S	G	G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	A
2nd monomer	G	S	G	G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	A

3-2. Stabilization of state 1 in Ras by RDA peptides

It has been shown previously that Ras in complex with GTP exists in dynamic equilibrium between two different conformations called state 1 and state 2 [92,93] . In state 1, Ras adopts an inactive conformation with reduced binding to effectors, whereas in state 2 Ras adopts an active conformation. Effector binding is expected to shift the equilibrium toward state 2 [92,94,95] . In state

2, Thr35 of switch I forms an interaction with the gamma phosphate of the GTP analogue, which closes the conformation and shields the GTP from the solvent [93] . In state 1, however, Thr35 loses its contact with the gamma phosphate and switch I adopts an open structure revealing two surface pockets: the first one surrounded by two switch regions, and the second one by the switch I preceding residues and GppNHp [96] . (**Figure 56**)

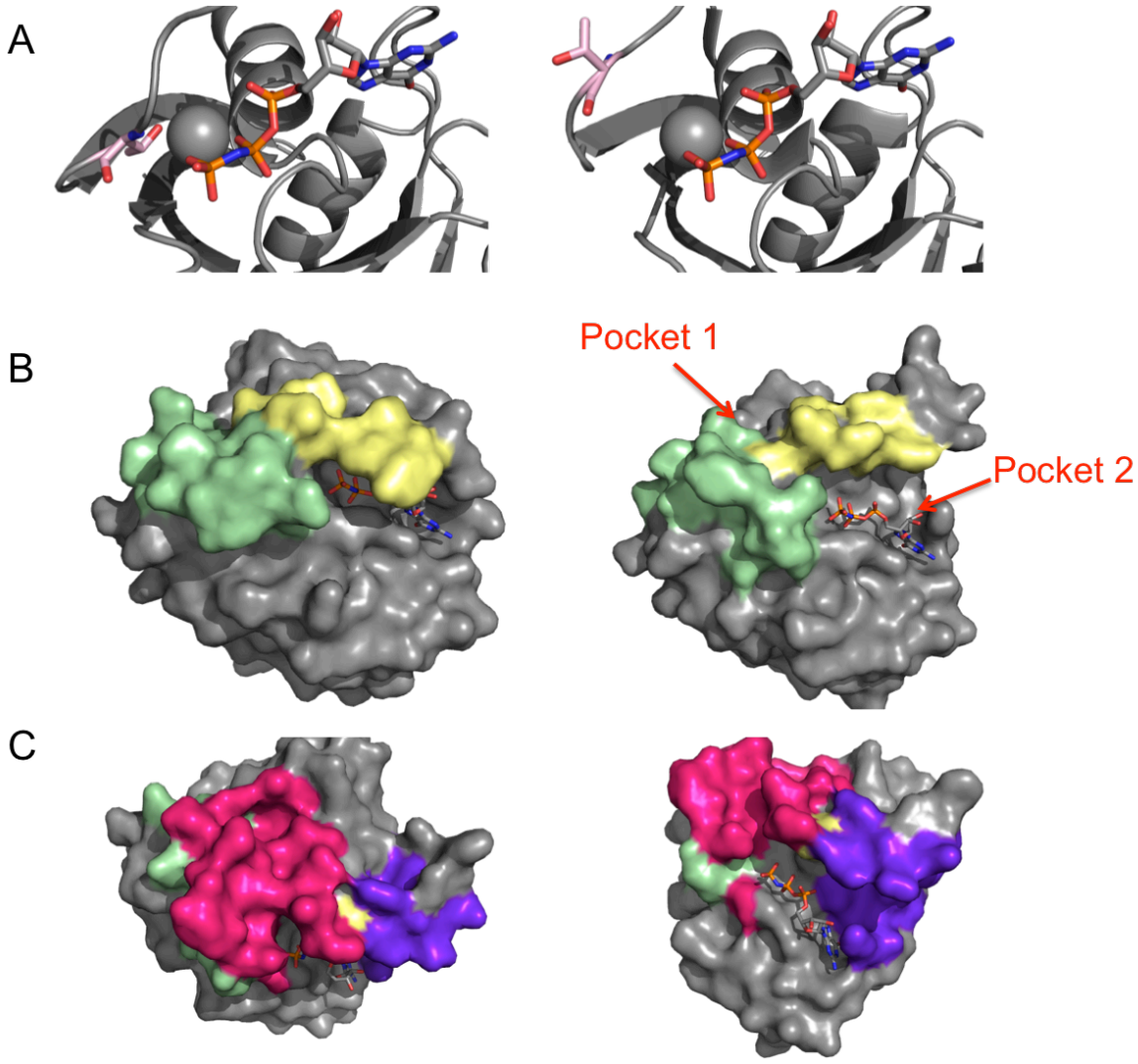


Figure 56. Wild type Ras in different states. A: Cartoon representation of Ras in state 2(left) and state 1(right). The GppNHp is shown in ball and sticks model. Thr 35 has been colored in light pink, showing either the presence or absence of its interaction with the nucleotide phosphate. B: while Ras in state 2(left)shows a closed conformation, the state 1(right) shows an open conformation, with two possible binding pockets. C: Pocket 1(left) shown in pink, pocket 2(right) shown in purple (PDB: 4EFL for the state 1 Ras, 5P21 for state 2 Ras)

When we look at the structure of Ras that has RDA bound to it, we can see that the conformation is quite similar to the inactive conformation, state 1, in both the switch I and switch II regions. (**Figure 57**) Thr35, which is an indicator of the Ras state, also shows an absence of direct or indirect contact with the γ –phosphate.

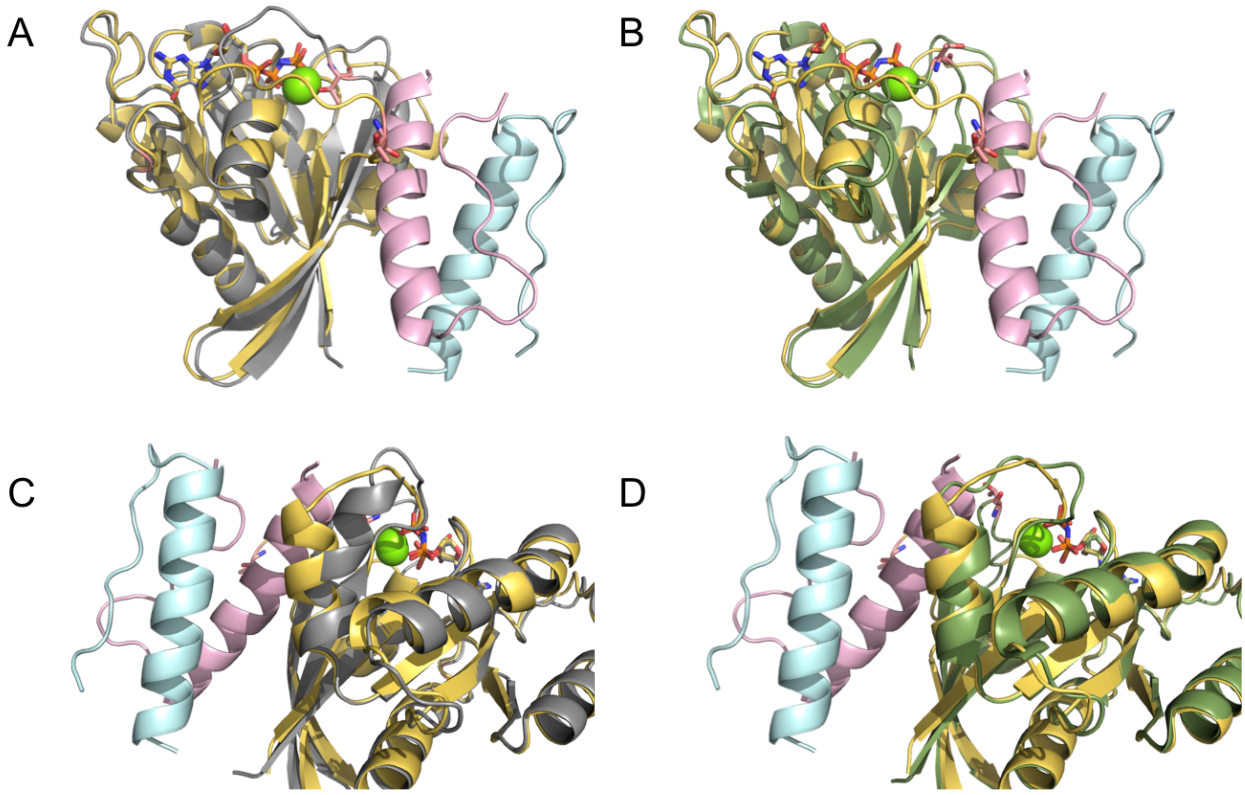


Figure 57. RDA bound Ras aligned with either state 2 Ras or state 1 Ras. The RDA bound Ras is colored in yellow, state 2 Ras in gray (left), and state 1 Ras in green (right). The primary helix of the RDA peptides is in light pink and the secondary helix in light blue. A & B: Switch I region shows that the RDA bound Ras closely mimics the Ras in state 1, and the beginning of the beta strand residue 37-41 is now in destabilized conformation. C&D: Switch II region shows that the RDA bound Ras is in close conformation with Ras in state 1.

The RDA peptides that are bound to Ras are positioned along pocket 1, causing the switches to flip outward. (**Figure 58**) This observation suggests that state 1, which shows a previously unobserved binding pocket, has the potential to be used as a targetable structure in the design of Ras inhibitors.

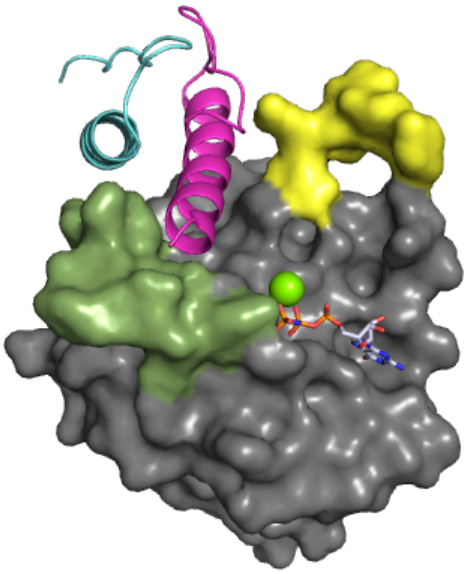


Figure 58. RDA bound Ras. Ras is depicted in surface representation, and the RDA peptides in cartoon representation. The switch I region is in yellow, and switch II in green. We see here that with the binding of RDA peptides the switch region opens up, resembling the Ras state 1.

As an extension of this idea, an FP experiment conducted by John McGee shows that the binding affinity of the 226-1 peptides toward K-Ras G12V D38P double mutant improved to K_d of 14 nM compared to 81 nM for K-Ras G12V (**Figure 59**). This D38P mutation was introduced in order to destabilize the conformation of the second beta strand of Ras, which normally resides from residue 38 to residue 46. In the RDA bound Ras, this second beta strand was pushed away losing half its secondary structure. In order to mimic this conformation, proline was introduced since it is known to disrupt secondary structures owing to restriction in its allowable dihedral angles and an inability to complete the H-bonding network.

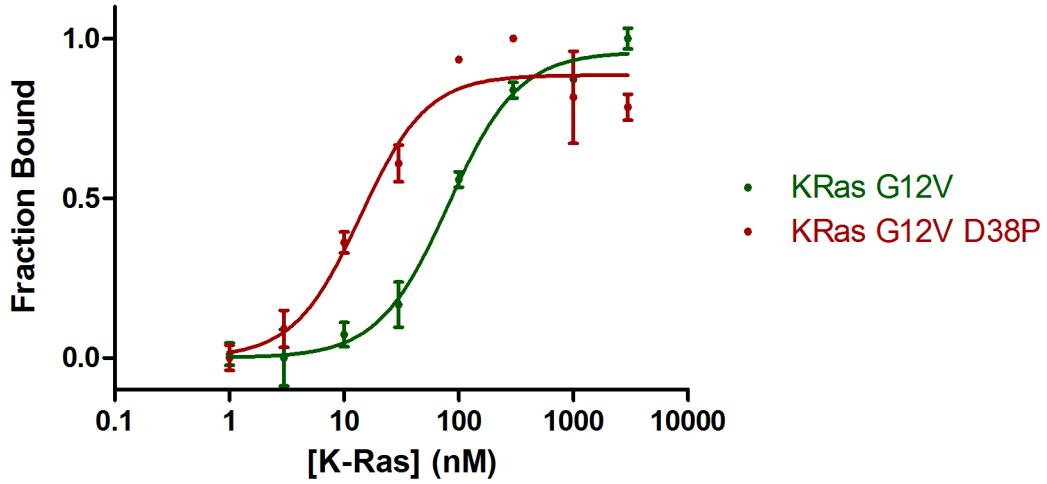


Figure 59. FP measurement of 226-1 peptides binding to KRas G12V D38P double mutant.

In order to confirm that the improved binding affinity came from Ras being stabilized in the conformation of the RDA bound form as we hypothesized, the RDAI bound Ras G12V, D38P double mutant complex was crystallized in the condition of 0.2-0.35M calcium chloride, 20-24% PEG 3350. A crystal that was formed under the condition of 21% (w/v) PEG 3350 and 0.25M calcium chloride using the hanging drop vapor diffusion method was soaked in cryoprotectant of 0.3M calcium chloride, 22% PEG 3350 and 16% PEG 400 before being flash-frozen in liquid nitrogen for x-ray analysis. The crystal structure showed four molecules in the asymmetric unit in the space group of P 1 2_1 1, with a resolution of 1.91 Å and unit cell dimensions of a=40.55, b=232.5, c=48.21, α=90, β=90.16, γ=90. This structure was nearly identical with the RDAI bound Ras (G12V) complex.

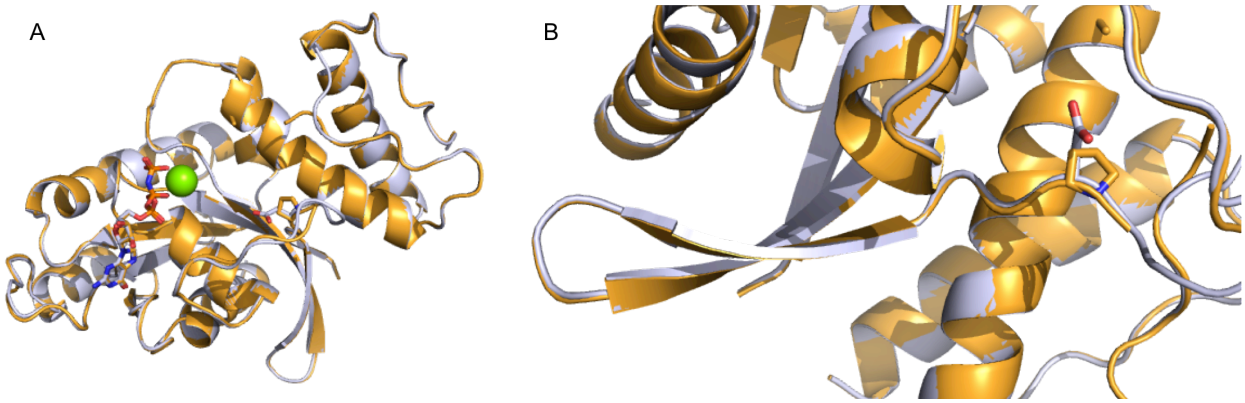


Figure 60. RDAI bound Ras G12V D38P double mutant compared with RDAI bound Ras G12V mutant.

We are currently in the process of crystallizing and solving the structures of the Ras G12V D38P double mutant to see if the introduction of proline also destabilizes the beta sheet formation in the absence of the RDA peptides. If the apo structure adopts an open state of Ras as we expect, we plan to use this mutant to screen for additional more Ras inhibitor candidates.

3-3. *i, i+4+4+4* stitched peptide

There have been several attempts in the Verdine group to look at these promising Ras binders for activity in cells. However, they have shown poor results in cell-based activity assays, which might be attributable to the fact that the RDA peptides are approximately 70 amino acids and may be too large for adequate cell penetration. By analyzing the structure of the K-Ras/RDA1 complex, we could see that all the residues that were interacting with the Ras protein seemed to be on the alpha-helices, with most of the interactions originating from the first helix. As expected, the polyproline helix seemed to be stabilizing the dimer. Therefore, we hypothesized that both the secondary alpha helix and the polyproline helix might be dispensable for interaction with Ras if we could stabilize the primary alpha-helix enough so that it could retain its structure

independently. We reasoned that this might help to enhance the cell permeability of the peptides while still retaining high-affinity binding to Ras.

Previous RDA minimization efforts in the Verdine group involved strategies such as putting a staple on the helix of the monomer, linking the two helices at the termini, or forming a non-covalent dimer of the two helices. However, none of these efforts gave promising results when the compounds were evaluated for binding to Ras. In an extension of this effort, I tried making a *i, i*+4+4+4 stitched peptides on both RDA1 and RDA2 peptides (**Figure 61**). Stitched peptides have been shown to give strong alpha helicity along with high thermal stability[97] , making this strategy appealing for the purposes of this project.

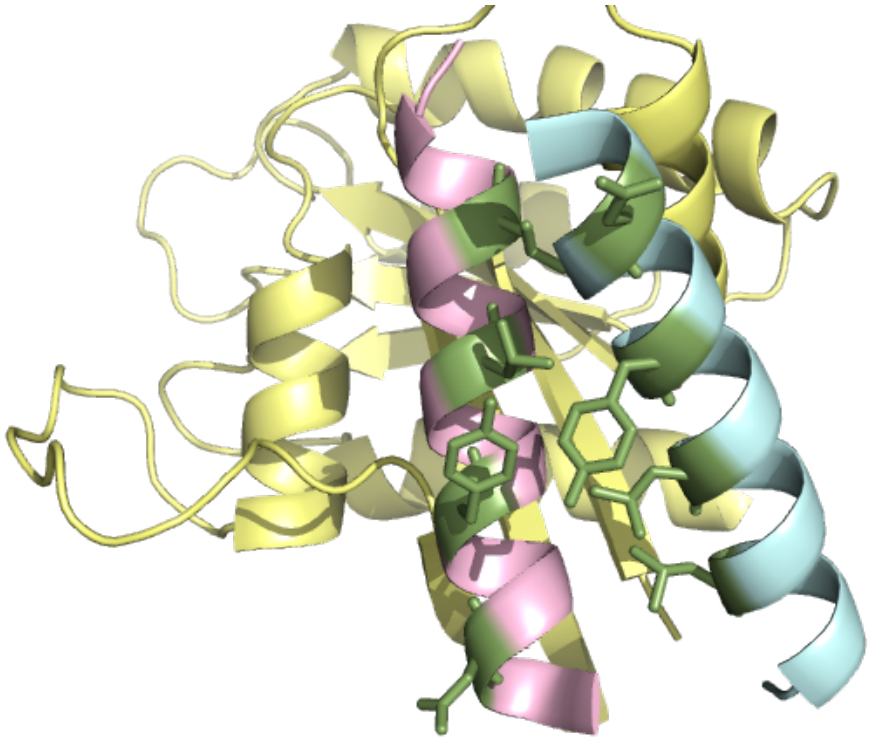


Figure 61. Stitched RDA peptides. The RDA peptides have been truncated retaining only the primary alpha-helix, and the residues in green have been stitched in order to enforce the alpha-helicity.

The positions for non-natural amino acid incorporation were chosen in order to avoid the residues that were important for binding (**Table 11**).

Table 11. Stitched RDA peptides. Stitched RDAI-1, 2 are based on RDA1 peptides, while stitched RDAII-1, 2 are based on RDA2 peptides that have A33R mutation.

Peptide	Sequence
Stitched RDAI-1	I E *S5* L H E *B5* W A R *B5* W N Y *S5* Y A V A
Stitched RDAI-2	I E *R5* L H E *B5* W A R *B5* W N Y *R5* Y A V A
Stitched RDAII-1	I E *S5* L H E *B5* W A R *B5* W N Y *S5* Y R V A
Stitched RDAII-2	I E *R5* L H E *B5* W A R *B5* W N Y *R5* Y R V A

Upon metathesis of the substrate peptides shown in the table above, staples with S5-S5, R5-R5, and S5-S5 stereochemical combinations were expected for stitched RDAI-1 and stitched RDAII-1, while R5-R5, S5-S5, and R5-R5 stereochemical combinations were expected for stitched RDAI-2 and stitched RDAII-2.

The RCM reaction was performed for with 30 mole % GI catalyst twice for 2 hours, with each treatment utilizing a freshly prepared solution in dichloroethane. Here we observed efficient RCM by almost 100% product conversion after only one 2 hour treatment (**Figure 62**). The single product peak showed a decrease in mass of 84 relative to the starting material, as expected from the loss of ethylenes from the metathesis reactions.

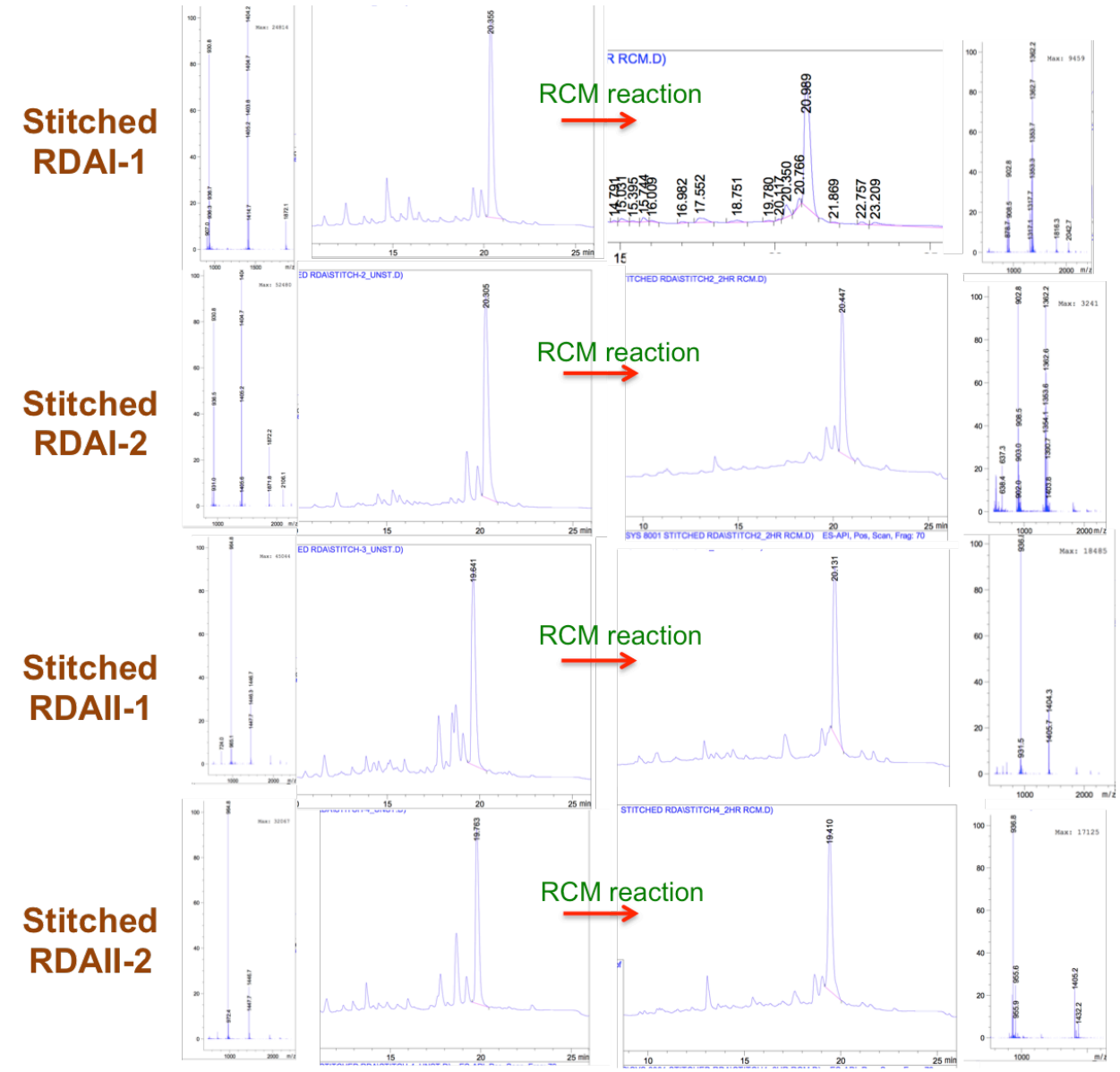


Figure 62. Stitched RDA peptides. 10% acetonitrile (0min) -100% acetonitirile (20min) run in LC/MS

However, the products had low solubility in water or buffer, as could be predicted from the observation that their LC/MS retention times corresponded to very high percentages of acetonitrile. In addition, the stitched peptides did not show detectable binding to Ras up to 10 uM in a fluorescence polarization binding experiment and had only moderate helical content (40~55 % helicity) when analyzed by circular dichroism. These experimental results suggest that the polyproline helix might be more crucial for structural stabilization than we assumed.

Chapter 4. Heterodimeric Ras inhibitor

4-1. Homodimer vs. heterodimer

The RDA peptides are a symmetrical dimer, but it is apparent in the co-crystal structures described above that rarely do residues from both monomers interact with Ras. In reality, for the most part only residues from the primary helix seem to be involved in the binding. Therefore, although the RDA peptides are symmetrical in structure, they interact asymmetrically with Ras. This suggests that there is an opportunity to optimize the dimer for interaction with Ras if we can break its symmetry. Essentially, the residue at a given position on one monomer might not be optimal for the corresponding position on the other monomer, and the more favorable amino acid mutation might be suppressed if the corresponding residue from the other monomer is not compatible with it.

It has been shown that two different proteins can be expressed from a single yeast display vector using a separate promoter and that the heterodimer can be stabilized through disulfide bonds [98] . Furthermore, by using the yeast Gal 1-10 bidirectional promoter, it is possible to express two open reading frames from one plasmid by galactose-induction [99] . This bidirectional Gal 1-10 promoter enables simultaneous expression of two different genes, providing the opportunity to evolve heterodimeric RDA peptides.

Our idea was to fuse one monomer with Aga2p to be anchored on the yeast cell surface (anchored peptide), and another monomer to be produced as a soluble fragment (free peptide). These peptides were coexpressed in the yeast cells for yeast display experiment. In order to ensure selecting for heterodimer over homodimer, we planned to modify the anchored peptide

such that it would not bind to Ras as a homodimer. Thus, the homodimers of the anchored peptide on the cell surface would not be able to bind to Ras, while the homodimer of the free peptide would not be screened by FACS since it would not be attached to the cells anymore. Therefore, the discovery of hits would depend on two steps: (1) the heterodimer formation of the anchored peptide with the free peptide and (2) recognition of Ras by the heterodimer.

In pursuit of this idea, Dr. John McGee from the Verdine group performed a yeast display experiment using the dual promoter system to express the two peptides simultaneously. The construct was designed with mutations in the non-binding anchored peptide in positions 25, 29 and 32, in order to disrupt binding to Ras. The A25N mutation would give a steric bump, while N29A and Y32A were intended to abrogate the binding by loss of important interaction between the peptides and Ras. In addition, in order to aid the dimerization of the two different peptides, Ile17 on the primary helix and Val34 on the secondary helix were mutated to form a putative bump-hole (**Table 12**).

Table 12. Initial template design for the heterodimeric RDA peptides. The blue residues are mutations introduced to disrupt binding to Ras, and the green residues are put in for putative bump-hole.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Free peptide (1st monomer)	G	S	G	G	P	R	R	P	R	C	P	G	D	D	A	S	X	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	A
Anchored peptide (2nd monomer)	G	S	G	G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	H	E	Y	W	N	R	L	W	A	Y	L	A	R	X	A

Using this design, we wanted to achieve two goals by evolving a heterodimer RDA: improving binding affinity and enhancing cell permeability.

In the library designed to achieve improvement in affinity of the peptide to Ras, the C-terminus of the primary helix was mutated and elongated by one more residue to form another interaction with Ras. As for the secondary helix, mutations were introduced into positions 21, 25 and 28, which appear to face Ras in the crystal structures of the RDA homodimers (**Table 13**).

Table 13. Library design to achieve improvement in affinity of the Ras binding peptides

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Free peptide (1st monomer)	G	S	G	G	P	R	R	P	R	C	P	G	D	D	A	S	X	E	D	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	X	X
Anchored peptide (2nd monomer)	G	S	G	G	P	R	R	P	R	C	P	G	D	D	A	S	I	E	D	L	X	E	Y	W	X	R	L	X	A	Y	L	A	R	X	A	

For improvement in cell permeability, anionic residues such as glutamates or aspartates were removed, avoiding those that seemed important for either binding to Ras or dimer formation. In this library, positions 14 and 19 of the primary helix and positions 14, 18, 19 of the secondary helix were mutated (**Table 14**).

Table 14. Library design to achieve improvement in cell-permeability of the Ras binding peptides

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Free peptide (1st monomer)	G	S	G	G	P	R	R	P	R	C	P	G	D	X	A	S	X	E	X	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	A
Anchored peptide (2nd monomer)	G	S	G	G	P	R	R	P	R	C	P	G	D	X	A	S	I	X	X	L	H	A	Y	W	N	R	L	W	A	Y	L	A	R	X	A

The directed evolution results from the libraries showed us that for the bump-hole mutation, leucine was strongly favored in position 17 of the primary helix, while for its pair either valine or phenylalanine was selected. For the affinity mutations, it did not seem like there was a preference for specific amino acids for c-terminal residues in the primary helix, but in the secondary helix, histidine was favored over tyrosine at position 21, and glutamine was favored at

position 25. Interestingly, Tyrosine seemed to be favored over the tryptophan that originally occupied position 28. Lastly, for surface charge mutations, Gln, Asn or Lys replaced the anionic residues in the parent peptide, shifting the overall charge of the peptide from -2 to +5. This increase in the formal charge was expected to be helpful in enhancing cell permeability. The results from the two libraries were combined and the resulting heterodimer was named 24G (**Table 15**), which proved to be a ~20 pM binder to Ras according to the surface plasmon resonance assays performed by John McGee.

Table 15. 24G peptide sequence. Residues from evolution for improved affinity are in red, for improved cell permeability in blue, and for putative bump-hole are in green.

24G Peptide	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1st monomer			S	G	P	R	R	P	R	C	P	G	D	Q	A	S	L	E	E	L	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	A	H
2nd monomer		G	G	G	P	R	R	P	R	C	P	G	D	N	A	S	I	K	Q	L	H	A	Y	W	Q	R	L	Y	A	Y	L	A	A	V	A	

Another heterodimer pair based on 24G peptide template was evolved in a yeast display experiment using 100 nM Ras in order to focus on forming a stronger heterodimer. Leu20 of the primary helix was mutated to either tyrosine or phenylalanine, and the residues on the secondary helix that would be surrounding them were mutated so that the amino acid combination that best accommodates the designed bump would be evolved. The hit heterodimer from this round of evolution was named the 32P peptide (**Table 16**).

Table 16. Library design and the resulting 32P peptide sequence

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Free peptide (1st monomer)			S	G	P	R	R	P	R	C	P	G	D	Q	A	S	X	E	E	Y/F	H	E	X	W	A	R	L	W	N	Y	L	Y	R	V	A	H
Anchored peptide (2nd monomer)		G	G	G	P	R	R	P	R	C	P	G	D	N	A	S	I	K	Q	L	H	A	X	W	Q	R	X	Y	A	X	L	A	A	V	A	
↓																																				
32P Peptide	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1st monomer			S	G	P	R	R	P	R	C	P	G	D	Q	A	S	N	E	E	Y/F	H	E	Y	W	A	R	L	W	N	Y	L	Y	R	V	A	H
2nd monomer		G	G	G	P	R	R	P	R	C	P	G	D	N	A	S	I	K	Q	L	H	A	F	W	Q	R	L	Y	A	F	L	A	A	V	A	

4-2. X-ray crystallization studies

In order to investigate how the 24G peptide gained stronger affinity toward Ras, we proceeded to form a complex between this peptide and Ras and undertake crystallization trials for x-ray structural studies.

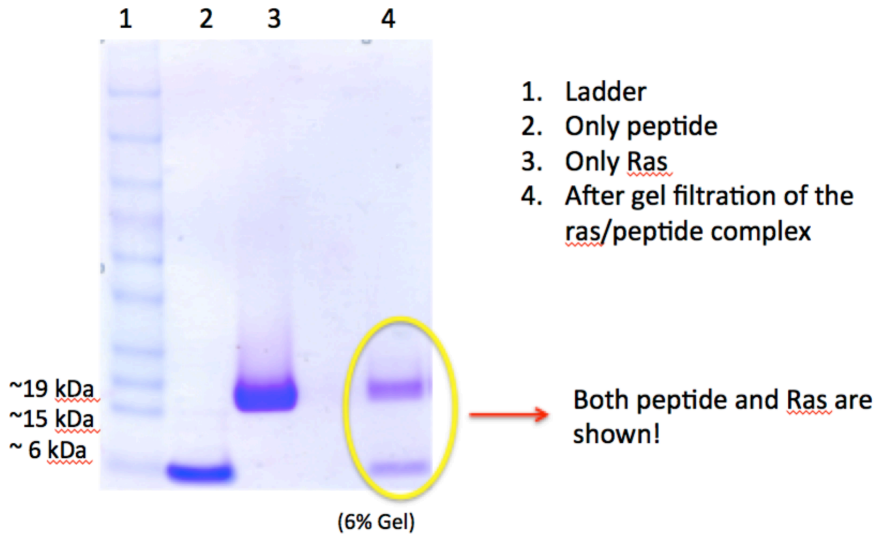


Figure 63. SDS-PAGE of 24G dimer, Ras, and the Ras/24G complex

We separately expressed Ras and the 24G peptide in bacteria. For the 24G peptide, we opted to introduce selenocysteine at the positions occupied by cysteine, expecting that this would provide

a non-reducible variant that would simplify the complex formation. After incubating Ras and the 24G peptide for ~2 hours (**Figure 63**), the complex was purified by gel filtration and subsequently screened initially against ~500 crystallization conditions. We analyzed the initial crystal hits under UV light to confirm the existence of peptide, as the peptides have three tryptophan residues in the sequences whereas Ras has none. For the crystals that glowed under UV, we expanded and optimized the crystallization conditions to get bigger and higher quality crystals. One of the conditions that was expanded and eventually yielded a crystal that diffracted upon x-ray irradiation was from the screening solution of 0.2M ammonium sulfate, 0.1 M HEPES pH 7.5, 25% (w/v) PEG 3350. In the expansion trays, reservoir solution in the range of 0.15M-0.25M ammonium sulfate, 0.1 M HEPES pH 7.5, 23-27% PEG solution was screened by mixing 0.5 ul of the reservoir solution with 0.5 ul of 0.44 mM K-Ras/24G complex (11.9 mg/ml in solution of 20 mM Tris pH 7.0, 50 mM NaCl, 5 mM $MgCl_2$) at 20°C in hanging drop vapor diffusion method. The crystal that grew under the condition of 0.2M ammonium sulfate, 26% PEG 3350, 0.1M HEPES pH 7.5 was picked and soaked in a cryoprotectant solution containing 0.2 M ammonium sulfate, 26% PEG 3350, 0.1M HEPES pH 7.5, and 14 % PEG 400 just before flash-freezing the crystal in liquid nitrogen. In the crystal structure, there were two molecules in the asymmetric unit with a space group of P 1 2_1 1 and unit cell dimensions of a=38.079, b=82.900, c=70.416, α=90, β=104.65, and γ = 90. The resolution was 2.57 Å, so we optimized the condition further by varying the ratio of the protein and the reservoir solution in the crystallization drops. The crystal that grew with the ratio of 2:3 (0.33 ul protein: 0.5 ul reservoir solution) in the same reservoir solution diffracted even better, producing a 1.72 Å dataset in the same space group (P 1 2_1 1) with unit cell dimensions of a=38.15, b=86.84, c=71.04, α=90, β=103.6, and γ = 90.

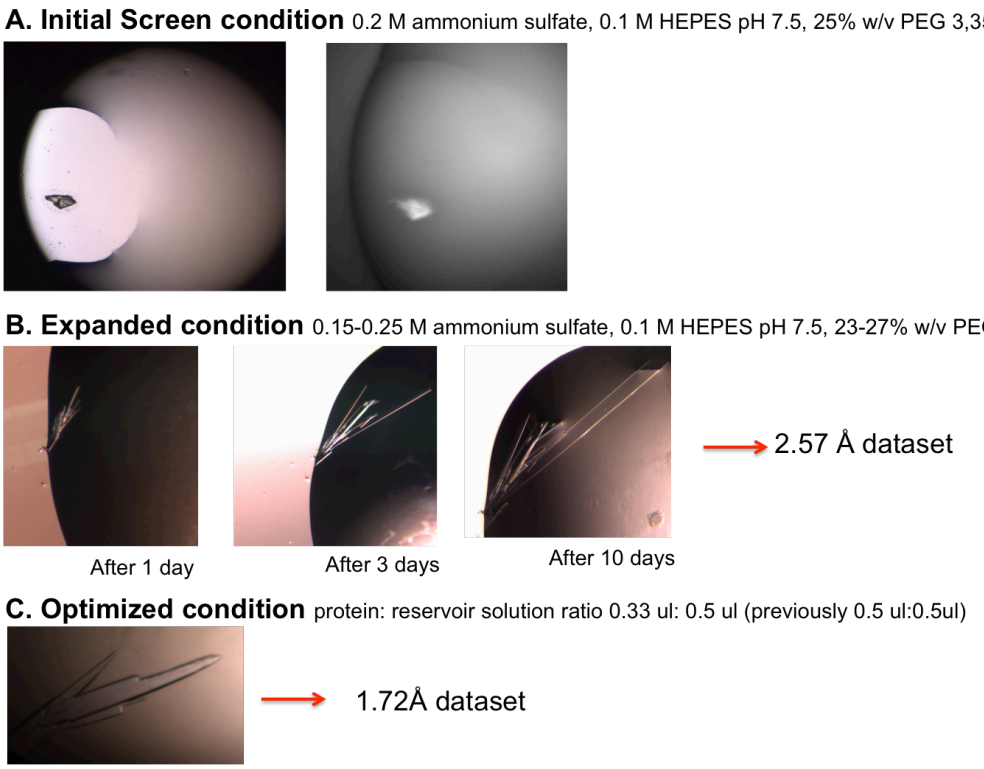


Figure 64. Crystallization conditions for 24G bound K-Ras (G12V) GppNHp complex

In the structure of the 24G peptide bound to K-Ras (G12V), we observed new interactions between Ras and the secondary helix. (**Figure 65**) As shown in the previous RDA bound K-Ras (G12V) complex, the imidazole group of His21 forms hydrogen bonds with side chains of Gln70 and Tyr71 of Ras, which explains why the evolution favored histidine. The newly evolved residue Gln25 was observed to form a direct hydrogen bond with Gln70 of Ras, and the newly evolved Tyr28 indirectly interacts with the backbone of Thr74 of Ras via an intermediary water molecule and might be having van der waals interaction with side chain of Lys5 of Ras. These mutations overall seem to help the peptide dimer to bind even more strongly to Ras protein by positioning the secondary helix toward the switch II region of Ras, explaining the improved Kd measured in the biochemical assay.

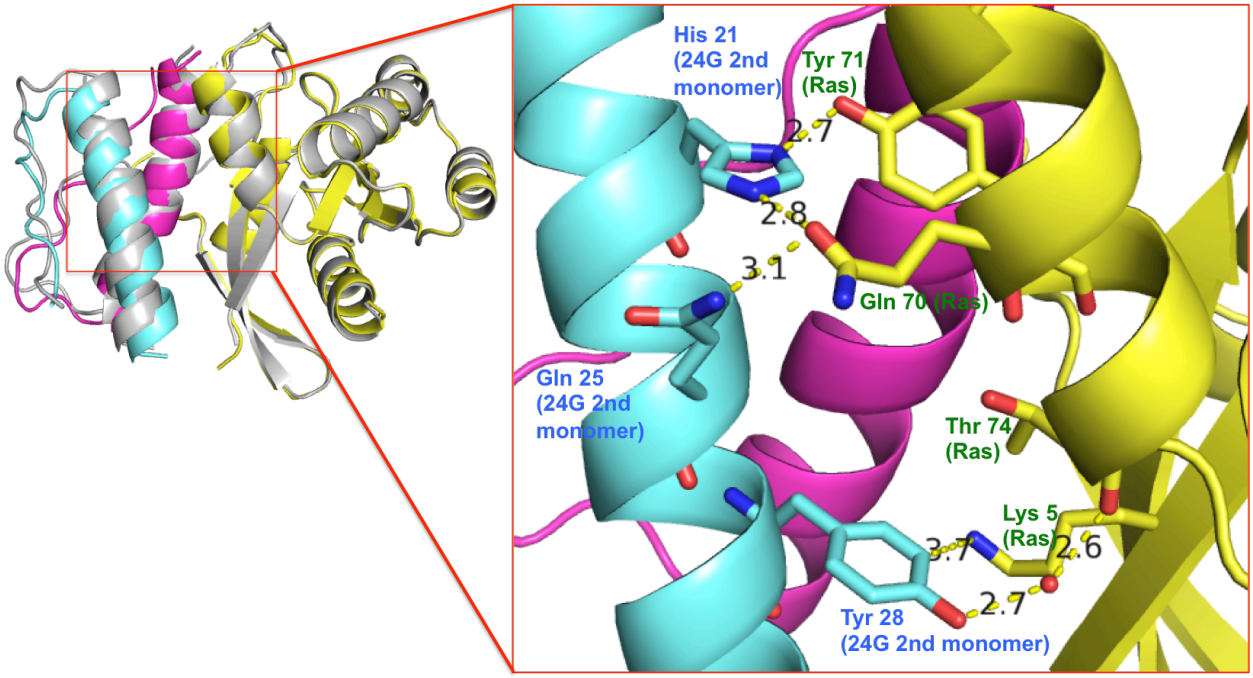


Figure 65. Analysis of the 24G/K-Ras complex. On the left shows the 24G/K-Ras complex overlaid with RDAI/K-Ras complex(gray). On the right shows the interactions from the second monomer of the 24G peptides with Ras.

In initial screening for crystallization of the 32P/Ras complex, the drop mixture of 0.5 ul of 0.37 mM K-Ras/32P complex(10.0 mg/ml in solution of 20 mM Tris pH 7.0, 50 mM NaCl, 5 mM $MgCl_2$) with 0.5 ul of 0.15-0.25 M ammonium formate, 18-21% PEG3350 yielded small, grain-like crystals. Unfortunately the crystals did not diffract at all, and varying the ratio of the drop mixture of the reservoir condition did not seem to help. Looking back at the 24G/Ras complex in the same conditions, we see that in the initial screen the crystals were also very small and took a few days to grow. However, in the larger scale screening, the complex yielded a bigger crystal that diffracted to a 1.98 Å resolution. Since there are only a few residues mutated in the 32P peptide compared to 24G peptide, and since the overall surface and the backbone should be similar, we expected the crystal packing to be similar as well. Therefore, we decided to try heteroseeding the crystals from 24G/Ras complex to grow 32P/Ras complex crystals (**Figure 66**).

The K-Ras/32P crystals began to grow within a few hours once the drop was heteroseeded with 24G/K-Ras complex crystals (**Figure 67**). A crystal was picked from 0.25M ammonium formate, 15% PEG 3350 condition, and was soaked in the reservoir solution supplemented with 20% glycerol prior to being flash frozen in liquid nitrogen. The structure was determined to be in the space group of P 2_1 2_1 2_1, unit cell dimension of a=70.97, b=73.22, c=81.04, α=90, β=90, γ = 90, with a 1.92 Å resolution.

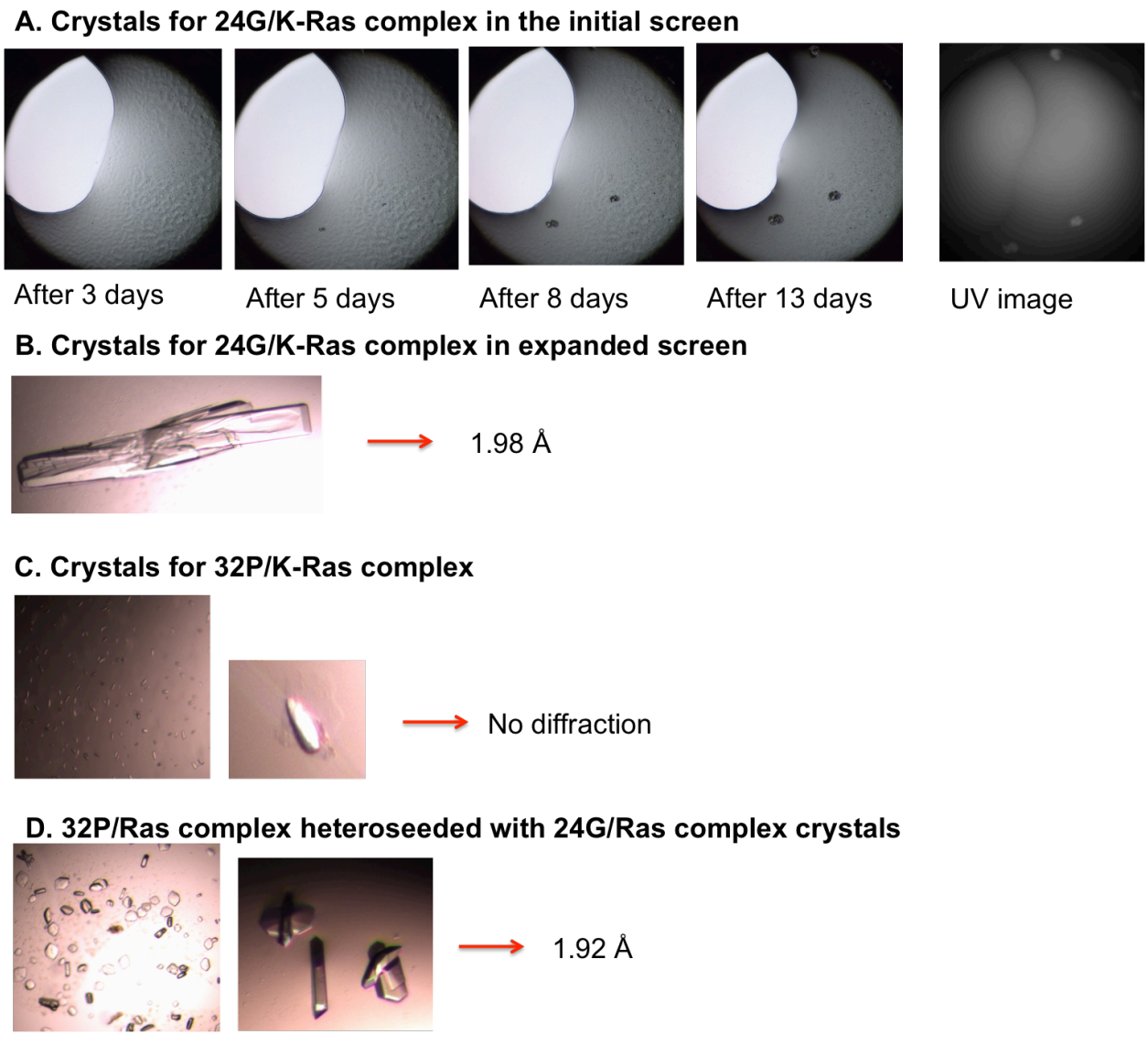


Figure 66. Comparison of the 24G/K-Ras and 32P/K-Ras crystals in the same condition

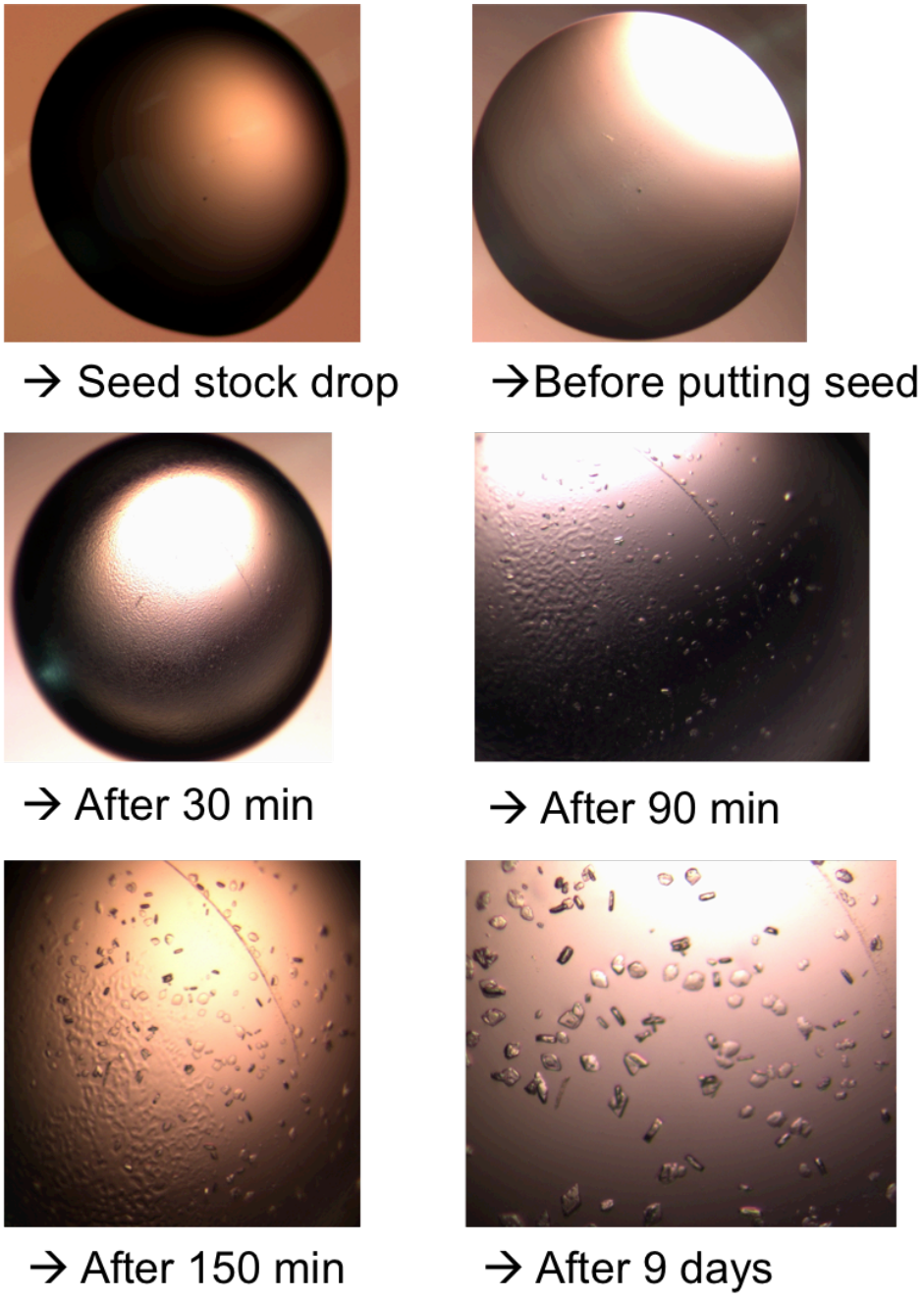


Figure 67. Heteroseeding. The crystals of the 32P-Ras complex began to grow when heteroseeded with the 24G-Ras microcrystals.

In the overall crystal structures of 24G and 32P bound to Ras, the peptides look quite similar to each other (**Figure 68**). When we look at the residues that were evolved via the yeast display experiment, compared to the ones in 24G those in 32P seem to be engaged in more intermonomer packing (**Table 17**). This suggests that the heterodimers might be packing better against each other, although it is hard tell for sure without direct comparison of the corresponding homodimer structures.

Table 17. Residue interactions in 32P and 24G peptides

	32P Peptide residues	Residues from the other helix within 4Å	24G Peptide residues	Residues from the other helix within 4Å
1st	Asn17	Phe30	Leu17	Tyr30
	Phe20	Phe23, Leu27	Leu20	Tyr30
	Tyr23	Phe23, Leu27	Tyr23	Tyr23, Leu27
2nd	Phe23	Phe20, Tyr23	Tyr23	Tyr23
	Leu27	Phe20, Tyr23	Leu27	Tyr23
	Phe30	Asn17	Tyr30	Leu17, Leu20

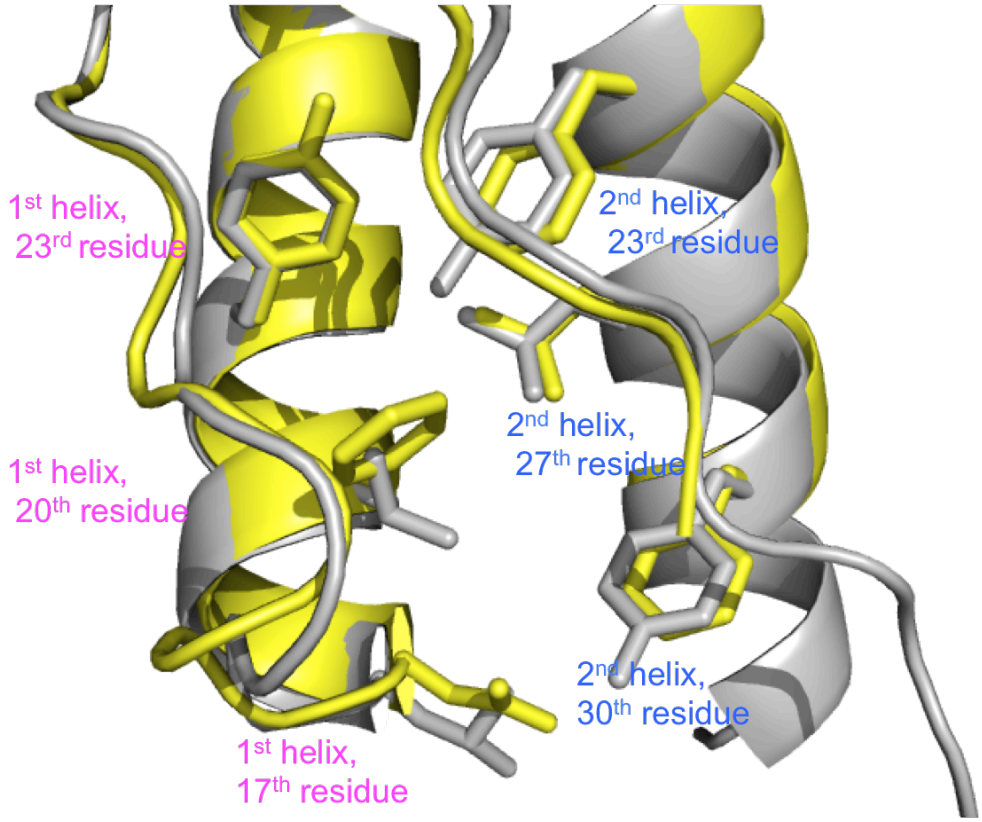


Figure 68. Comparison of 32P (yellow) and 24G (gray) peptide

Chapter 4. Conclusions and future directions

In part II of this thesis, we have shown to evolve peptide binders to Ras, which was previously thought to be one of the most intractable targets. With the difficulty of using rational design in the beginning, we have started with a naïve screening of a libraries of a number of peptide scaffolds, one of which eventually yielded peptide binders to Ras with K_d of low-mid nM. By solving the structure of the hit peptides bound to Ras and analyzing it, we were able to further optimize the peptides for improved affinity. The initial homodimer binders were evolved to a heterodimer binders, which yielded even higher affinity to Ras from having extra contact residues in the secondary monomer that were previously not possible due to the constraints

associated with the symmetry of the homodimer. The binding of the hit peptides to Ras was assessed in a number of biochemical experiments such as fluorescence polarization, surface plasmon resonance (biacore) and immunoprecipitation of cancer cell lysates. Currently in Verdine group we are trying to improve the cell permeability as well as endosomal escape of the peptides by optimizing the sequence.

The crystal structures of the hit peptides bound to Ras suggest that the peptides might be stabilizing state 1 conformation of Ras, which is a weak binding state for effectors. The fact that Ras exists in a dynamic equilibrium between these different states, where one of which would abrogate the binding of effector to Ras, gave us a new targeting strategy. By fixing the equilibrium to state 1 by introducing mutation into Ras, we speculate that we might be able to screen for more potential inhibitors that would fit into the binding pockets that exist in that particular state.

Chapter 5. Experimental Method

5-1. Yeast passage

Typically yeast cell was passaged using SDCAA media a day before, and then induced with SGCAA media 20-24 hours prior to sorting. For antibiotics Kanamycin and ampicillin was used.

- SDCAA Media: 8.56 g of sodium phosphate monobasic ($NaH_2PO_4H_2O$), 5.4 g of sodium phosphate dibasic (Na_2HPO_4), 20 g of 20 g dextrose, 6.7 g of Difco yeast nitrogen base, 5 g of bacto casamino acids were dissolved and mixed well in 1000 ml of deionized H_2O, and filter sterilized.

- SGCAA Media: 8.56 g of sodium phosphate monobasic ($NaH_2PO_4H_2O$), 5.4 g of sodium phosphate dibasic (Na_2HPO_4), 20 g of 20 g galactose, 6.7 g of Difco yeast nitrogen base, 5 g of bacto casamino acids were dissolved and mixed well in 1000 ml of deionized H_2O, and filter sterilized.

5-2. MACS

A day after the induction, the density of the yeast was measured by OD_{600}. Generally absorbance of 1 indicates about 10^7 cells/ml. For MACS experiment, typically 1 or $2x10^9$ cells were collected and pelleted. The pellet was washed and resuspended with MACS buffer to the final binding volume. In the cold room, labeled K-Ras G12V mutant protein was added to a desired binding concentration, and incubated for about 30-60 minutes at room temperature. The sample was then washed and pelleted with MACS buffer, and either streptavidin beads or anti-biotin beads were put and incubated for 10-30 minutes. Afterward, autoMACS machine was used to collect the labeled cells. The collected cells were grown in SDCAA media with Kan/Amp antibiotics at 30 ^oC. The number of the sorted cells was determined by plating and counting the colonies. (a colony represents ~16,000 eluted cells)

- MACS buffer: 150 mM NaCl, 5mM KCl, 15 mM sodium phosphate, 0.5 % (w/v) BSA, 5mM $MgCl_2$, adjusted to pH 7.4~7.5

5-3. FACS

After about 20 hours of induction, desired amount of cells were washed, pelleted and resuspended with FACS buffer to the final binding volume. Labeled K-Ras G12V mutant protein was added, pelleted and resuspended with either Alexa 488 labeled anti-HA antibody or FITC-labeled anti-FLAG antibody and SA-APC or NA-PE to the appropriate labeling volume and incubated at 4 ^{o}C for 10 minutes, protected from light. SA-APC and NA-PE was alternated to prevent enrichment of secondary binder. The labeled cells were once again washed and pelleted, and resuspended in FACS buffer immediately prior to sorting. The sorting was done using the FACS machine at the Bauer core at Harvard University.

- FACS buffer: 150 mM NaCl, 5mM KCl, 15 mM sodium phosphate, 0.1 % (w/v) BSA, 5mM $MgCl_2$, adjusted to pH 7.4~7.5

5-4. Hit sequencing

After sorting, the enriched cells were passaged overnight, and mechanically lysed, and then Qiagen miniprep kit was used to acquire the plasmid DNA. The yeast display vector was then transformed into bacteria (XL-10 gold cells) for colony PCR. The PCR product was used for sequencing the hits and to use them as templates for error-prone PCR to prepare new yeast display library for directed evolution screening.

5-5. Yeast transformation

S. cerevisiae strain EBY100 was streaked on a YPD plate and grown for 2-3days at 30 ^{o}C. The evening before transformation, a single colony was picked and inoculated in YPD media, and was shaken overnight. Then 110 ml of pre-warmed YPD to an OD_{600} of 0.1 was inoculated and shaken until the OD_{600} reached 1.40. Then, 1.0 ml of Tris-DTT was added and continued to shake for exactly 15 minutes. The cells were then pelleted, washed and resuspended in cold E buffer.

To prepare the DNA for transformation, appropriate amount of digested and purified pCTCON2 vector was mixed with appropriate amount of DNA insert (typically 1:4 ratio was used). The mixture was precipitated using PelletPaint. After thoroughly drying the pellet, it was redissolved in milliQ water and kept at 4 ^{o}C until transformation.

Appropriate amount of cells were mixed with the DNA and incubated for 10 minutes, and electroporation was performed for transformation at 0.54kV, 25 μF in a genePulser II electroporator. Afterward, they were shaken in YPD media for 1 hour at 30 ^{o}C, pelleted and resuspended in SDCAA media with kanamycin and ampicillin antibiotics. A small portion of the cell suspension was taken out and plated to determine the transformation efficiency and the library size.

- YPD media: 20g of dextrose, 20 g of bacto peptone, and 10 g of bactoyeast extract, mixed into 1000 ml deionized water and sterilized by filtration
- Tris-DTT: 2.5M DTT in 1M Tris pH 8.0

- E-Buffer: 10 mM Tris, 270 mM sucrose, 1mM $MgCl_2$, pH 7.5

5-6. RDA peptides expression and purification

Peptide expression (for RDA1 and RDA2 peptides)

RDA1 : GSGGPRRPRCPGDDASIEDLHEYWARLWNYLYAVA (Dimer)

RDA2 : GSGGPRRPRCPGDDASIEDLHEYWARLWNYLYRVA (Dimer)

A small amount of RDA peptide transformed BL21 cells were scraped into LB media with Kanamycin and Ampicillin antibiotics and shaken at 37 ^{o}C overnight. The next day the overnight culture was distributed equally to twelve 1L of LB media with Kanamycin and Ampicillin antibiotics. The cells were shaken at 37 ^{o}C, 215 rpm until they reached an OD_{600} of 0.7, which typically takes around 5-7 hours. Once the cells reached the OD, 500 mM IPTG and 2 ml of 10% w/v arabinose was added to each 1L flask, and were shaken for 4-5 more hours at 30 ^{o}C. Note that arabinose was used here for the induction of dnaK, dnaJ, grpE, GroEL, and GroES chaperones to aid the expression of the peptides. Afterward the cells were pelleted for 20 minutes at 4200 rpm, and the media was poured off. About 40 ml of His buffer 1S was used to resuspend the pellet of 2L amount of cell culture, and the mixture was snap-frozen in liquid nitrogen for several minutes, and stored at -80 ^{o}C.

- His Buffer 1S: 50 mM sodium phosphate pH ~7.5, 1000 mM NaCl, 10 mM imidazole

Peptide purification

The pellets were thawed and one tablet of EDTA-free protease inhibitor was mixed per 4 liters of culture. The cells were sonicated with a tip sonicator using 6 cycles of 10 s on, 15 s off at 6.5 power level. Once finished, the lysate was centrifuged at 15,000 rpm for 35 minutes in the high-speed centrifuge. After the centrifuge is done, decant and filter the supernatant through 1.2 um membrane, and then mixed and incubated with cobalt resin for 30-40 minutes. Typically ~2ml bed volume of cobalt resin was used per 2 liters of culture. Afterward, the flow-through was drained, and then the column was washed with His Buffer 1S. Then Hisbuffer 2S was used to elute the protein. The existence of protein in the eluted buffer was verified by adding 5 ul of liquid from the column tip to 95 ul Bradford reagent. EDTA was added to the final concentration of 0.5 mM to the eluate, and then the sample was transferred to a centriprep YM-3 and centrifuged at 3200 rpm until it concentrated down to about 2ml per 2 liters of cell culture. Once the eluate has finished concentrating, it was filtered through 0.45 um Tuffryn membrane and buffer exchanged into TEV cleavage buffer using a pair of tandem 5 ml HiTrap desalting columns. The peptide fractions were collected, with care not to include the salt peak (**Figure 69**).

TEV was added to the sample, and the cleavage reaction went overnight at room temperature. Three connected C-18 sep-pak columns were used to desalt the solution, and the peptide was purified by HPLC (**Figure 70**). The collected peptides went through speed vac to evaporate the acetonitrile, then lyophilized.

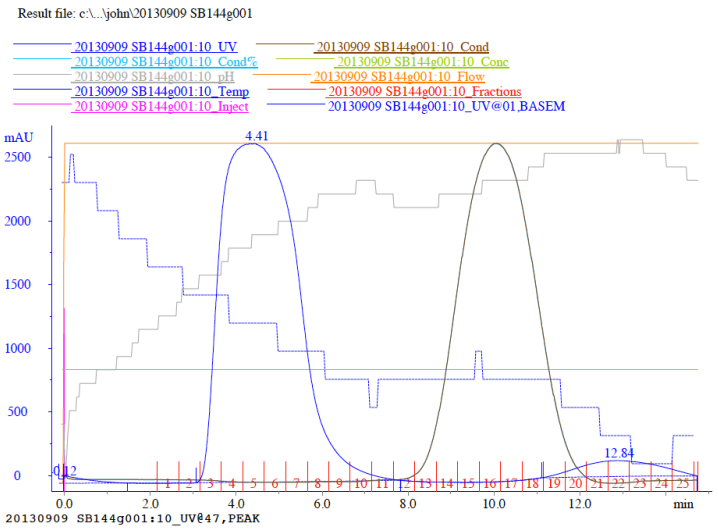


Figure 69. FPLC for buffer exchange of the peptides for TEV cleavage. The peak at ~4 min shows the peptide peak, while the peak at ~10 min shows salt peak.

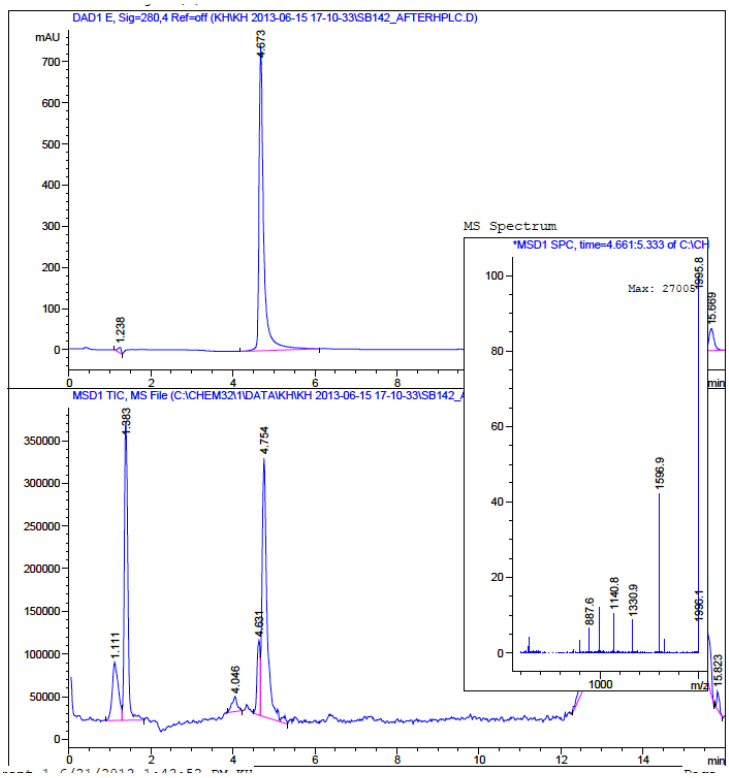


Figure 70. After purification of the RDA peptides by HPLC. The first column shows the UV spectrum, and the second column the mass spectrum.

5-7. Heterodimer peptides expression and purification

Peptide expression (for 24G and 32P peptides)

24 peptide: SGPRRPR**Se**PGDQASLEELHEYWARLWNYLYRVAH

G peptide: GGGPRRPR**Se**PGDNASIKQLHAYWQRLYAYLAAVA

32 peptide: SGPRRPR**Se**PGDQASLEELHEYWARLWNYLYRVAH

P peptide: GGGPRRPR**Se**PGDNASIKQLHAYWQRLYAYLAAVA

Note that in these peptides we have incorporated selenylcysteines (SeCys) instead of cysteines. The constructs are in Cys auxotrophic E. coli cell line BL21 CysE. The cells were grown overnight in LB media, supplemented with Kanamycin, Carbenicillin and cysteines. The next day the cells were pelleted and resuspended and grown in two flasks of 500 ml DLMM media as previously described [100,101] , with Kanamycin, carbenicillin and cysteines added. The cells were shaken at 37 ^{o}C, 215 rpm until they reached an OD_{600} of ~2.2, and then induced with 200 mM IPTG and shaken for 15 more minutes. Afterward, the cells were pelleted and resuspended in a second batch of DLMM with IPTG and selenylcysteines. The cells were grown overnight at 25 ^{o}C. The next day, I pelleted the cells and froze them slowly at -80 ^{o}C.

Peptide purification (for 24G and 32P peptides)

The pellets were slowly thawed, and then ~25 ml Lysis buffer B-Per+lysozyme(3mg/50ml)+universal nuclease for cell lysis was used to resuspend each pellet from ~500 ml media. After rocking the solution for ~10 min at room temperature, salt was put to a final concentration of 500 mM, and imidazole to ~10 mM. (For 25 ml, 2.5 ml of 5 M NaCl was

used, and 0.125 ml of 2 M imidazole was used) After rocking for 5 more minutes, the solution was centrifuged for ~25 min at ~15,000 rpm. About 2 ml bed volume of cobalt resin was prepared for each column. They were washed with milliQ water and His Buffer 1S. After the centrifuge, the lysate was clarified through 1.2 um membrane into the column, and they were mixed with the resin for 15-30 minutes. Afterward, the lysate was drained through the column by gravity, and then column was rinsed with His Buffer 1S until the color of the Bradford reagent didn't turn blue with the solution coming out from the column (95 ul of Bradford reagent+ 5ul of sample). Then Hisbuffer2S was addedto elute the protein, and EDTA was added to the eluate to the final concentration of 0.5 mM. It was transferred to a centriprep YM-3 and centrifuged at 3200 rpm until it concentrated down to 2~4 ml. The concentrated eluate was filtered through 0.45 um Tuffryn membrane and buffer exchanged into TEV cleavage buffer using a pair of HiTrap Columns. TEV was added to cleave the peptides overnight at room temperature. After overnight TEV cleavage, the peptides were desalted using sep-pak column, eluting at 75 acetonitrile in water, 0.1% TFA. Afterward the heterodimer peak was purified by HPLC (**Figure 71**) and lyophilized.

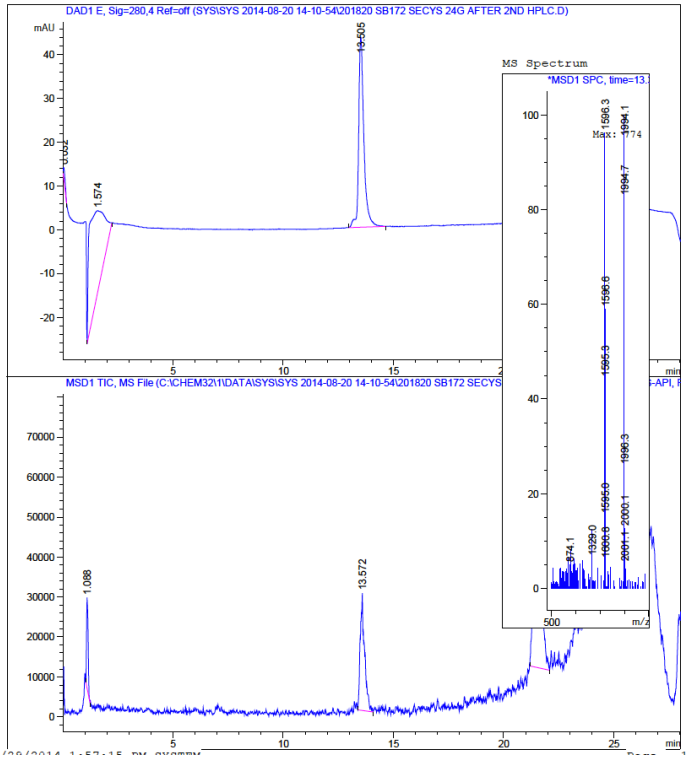


Figure 71. After HPLC purification of the heterodimer

5-8. Ras protein transformation, expression & purification

Ras protein Transformation

Verdine group plasmid stock 2638 (pET28b, human k-ras G12V core domain, hexahistidine tag+GGGG) was transformed into BL21 DE3 cells for expression of recombinant G12V k-Ras protein. . ~1 ul of the plasmid was mixed into competent cells taken out from -80 ^{o}C stock, and let it sit for ice for 30 minutes. The sample was heat shocked for 45 seconds at 42 ^{o}C, and then rested on ice for 2 min. Under sterile technique, 500 ul of SOC media was put in and incubated in the shaker for 1 hour at 37 ^{o}C. Afterward, ~6 ml of LB media with Kanamycin antibiotic was put in and incubated at 37 ^{o}C overnight.

The next day, for transformed BL21 DE3 cells, the cells were pelleted and then resuspended in LB media with 15% glycerol. They were then snap-frozen in -80 ^{o}C.

Ras protein expression

For expression, some cells from the K-Ras transformed cell stock was scarped out, and they were grown overnight in LB media with Kanamycin and chloramphenicol as antibiotics. The next day, they were equally distributed into twelve 1-L cell flasks, and grown until they reached the OD of ~1.0, and then 1 mM IPTG was put in for induction. Afterward, they were grown at 30 ^{o}C overnight (12~16 hours). The next day, the cells were pelleted 2L each, resuspended in Ras lysis buffer, snap-frozen and stored in -80 ^{o}C.

- Ras Lysis: 20 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole

Ras protein purification

A 2L cell pellet from -80 ^{o}C freezer was taken out and thawed. Then it was sonicated (3 min, 3 sec on, 8 sec off, 6.5) and ultracentrifuged for 20 minutes. The cell lysate was incubated with ~5 ml regenerated Ni resin for about 30 min. After the incubation the column was washed with ~70ml buffer and then the protein was eluted using ~30 ml elution buffer. I adjusted the amount of buffer needed by checking the presence or absence of protein using 5 ul of the eluant with 95 ul Bradford reagent. The eluate was concentrated to about ~2 ml, and then buffer exchanged to TEV protease cleavage buffer using 2 tandem 5ml HiTrap desalting columns. The collected fractions were pooled and ~1mg of TEV was added at 4 ^{o}C, overnight (**Figure 72**). The next morning the tagless protein was concentrated down to less than 2ml, and then buffer exchanged

into Goode Exchange buffer. The correct fractions were collected and concentrated down to less than 2ml. Noticeably the sample that was yellow to brownish color are shown to be transparent in this step. For nucleotide exchange, 4X of the desired nucleotide (GMP-PNP) along with CIP at ~2.5U/mg protein was added to the sample at room temperature for 1 hour. After the exchange, the sample was spinned down and the supernatant buffer exchanged into Ras GF buffer using Superdex 200 Prep column. The fractions containing the Ras protein was collected and concentrated to a desired concentration, then saved in -80 ^{o}C.

- Ni-NTA Elution buffer: 20 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole
- TEV protease cleavage buffer: 20 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT
- Goode exchange buffer: 32 mM Tris pH 8.0, 200 mM NH_4SO_4
- Ras GF Buffer: 20 mM Tris pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$

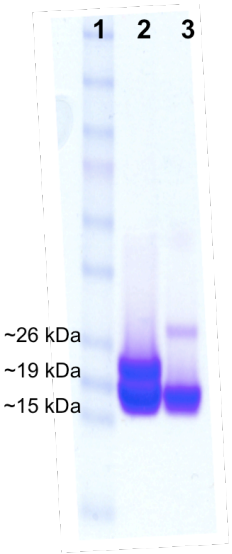


Figure 72. SDS-PAGE analysis of Ras TEV cleavage. 1: ladder, 2: Ras before TEV cleavage, 3: Ras after TEV cleavage

Ras Nucleotide loading check

~2nmol of Ras proteins were raised to 150 ul with ultrapure water, and then mixed with 150 ul of Nucleotide loading buffer. The nucleotide loading was checked by analyzing the sample by reverse-phase HPLC using a 20 minutes isocratic run in Nucleotide loading buffer for HPLC. The elution time was compared with those of GMPPNP and GDP. (**Figure 73**)

- Nucleotide loading buffer: 100 mM Potassium phosphate pH 6.5, 10 mM tetrabutylammonium bromide
- Nucleotide loading buffer for HPLC: 100 mM Potassium phosphate pH 6.5, 10 mM tetrabutylammonium bromide, 7.5% acetonitrile

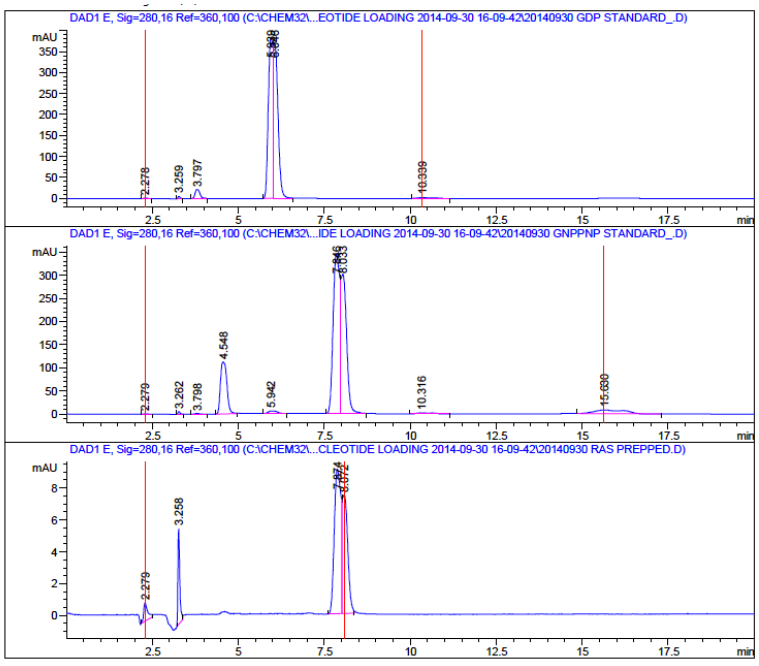


Figure 73 Elution time for GDP (1st column), GMPPNP (2nd column) and nucleotide loaded on Ras (3rd column). Here the elution time for Ras nucleotide matches that of the GMP-PNP

5-9. X-ray crystallography studies of the Ras/peptide complex

Complex Formation

Lyophilized peptides (RDA1, RDA2, 24G and 32P) were dissolved in DMSO and quantified, and were made to a starting concentration of 1 mM in DMSO. 100 nmol of peptides and Ras protein were diluted in ~5ml Ras buffer each. Then they were gradually mixed together. The mixture was incubated at 4^{o}C for 2 hours. Then it was centrifuged for 5 min at 4500 g, and the supernatant was concentrated down to 2 mL for purification by gel filtration using superdex 75 column. The fractions for the peak of the complex was pooled and concentrated.

- Ras buffer: 20 mM Bis-Tris pH 7.0, 50 mM NaCl, 5 mM $MgCl_2$

Crystal condition screening and X-ray diffraction

Typically, initial screens were conducted using 5 different kinds of commercially available 96 reagent crystallization screen kits: Crystal Screen 1, Crystal Screen 2, JCSG+ Suite, PEGs Suite, PEGs II Suite, Index screen. 96-well plates in a sitting drop vapor diffusion were used, and crystallization drops were made by mixing 0.2 ul complex with 0.2 ul of the screening solution. The crystals that grew were checked with UV images to see if they contain the peptides. While the peptides have tryptophan fluorescence while Ras proteins do not have any tryptophan residues that would glow in UV. This helps to exclude salt crystals or only Ras protein crystals. Afterward, the conditions that gave the crystals were expanded and optimized using 0.5 ul of peptides mixed with 0.5 ul of the screening solution, or with varying ratios, using the hanging drop vapor diffusion method at 20 ^{o}C.

For heteroseeding, first the K-Ras/24G crystals were collected using a loop, and were washed using the reservoir solution a few times. They were transferred to fresh drops 2-3 times for the washing. Then ~50 ul of drops containing the crystals were collected into the sead bead tube(Hampton Research) which was pre-chilled on a bucket of ice. The seed bead tube was vortexed for three minutes, stopping every 30 seconds to cool the tube on ice. The seed stock was diluted 10 fold and 100 fold to determine the optimal seeding concentration. 0.5 ul of seed stock was mixed into the K-Ras/32P crystallizaion drops that were equilibrated in advance for a day. The remaining crystal seed stocks were snap frozen in liquid nitrogen and saved in -80 ^{o}C freezer.

X-ray diffraction data were collected at the beamline of the Argonne Photon Source. HKL2000 program suites were used to process all diffraction data. The structures were determined by molecular replacement with Phaser either using the previously determined K-Ras structure (PDB 3GFT) or K-Ras/RDA1 complex complex that was solved as a search model. The structure was refined using COOT and Phenix Refinement.

References

1. Verdine, G. L. & Walensky, L. D. The challenge of drugging undruggable targets in cancer: lessons learned from targeting BCL-2 family members. *Clin. Cancer Res.* **13,** 7264–70 (2007).

2. Rizo, J. & Gierasch, L. M. Constrained peptides: models of bioactive peptides and protein substructures. *Annu. Rev. Biochem.* **61,** 387–416 (1992).

3. Fairlie, D. P., Abbenante, G. & March, D. R. Macrocyclic peptidomimetics-forcing peptides into bioactive conformations. *Curr. Med. Chem.* **2,** 654–686 (1995).

4. McGeary, RP & Fairlie, DP. Macrocyclic peptidomimetics: potential for drug development. *Curr. Opin. Drug Discovery Dev.* **1,** 208–17 (1998).

5. Fu, G. & Grubbs, R. The application of catalytic ring-closing olefin metathesis to the synthesis of unsaturated oxygen heterocycles. *J. Am. Chem. Soc.* **114,** 5426–5427 (1992).

6. Calderon, N. Olefin metathesis reaction. *Acc. Chem. Res.* (1972).

7. Miller, S. & Grubbs, R. Synthesis of Conformationally Restricted Amino Acids and Peptides Employing Olefin Metathesis. *J. Am. Chem. Soc.* **117,** 5855–5856 (1995).

8. Blackwell, H. E. & Grubbs, R. H. Highly efficient synthesis of covalently cross linked peptide helices by ring closing metathesis. *Angew. Chem., Int. Ed.* **37,** 3281–3284 (1998).

9. Boal, AK, Guryanov, I, Moretto, A & Crisma, M. Facile and e-selective intramolecular ring-closing metathesis reactions in 3_{10}-helical peptides: a 3D structural study. *J. Am. Chem. Soc.* (2007).

10. Schafmeister, C., Po, J. & Verdine, G. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.* (2000).

11. Chapman, R. N., Dimartino, G. & Arora, P. S. A highly stable short alpha-helix constrained by a main-chain hydrogen-bond surrogate. *J. Am. Chem. Soc.* **126,** 12252–3 (2004).

12. Lierop, ávan B. & RoyáJackson, W. Regioselective formation of interlocked dicarba bridges in naturally occurring cyclic peptide toxins using olefin metathesis. *Chem. Comm.* (2009).

13. Henchey, L. K., Jochim, A. L. & Arora, P. S. Contemporary strategies for the stabilization of peptides in the alpha-helical conformation. *Curr Opin Chem Biol* **12,** 692–7 (2008).

14. Scholtz, J. M., Qian, H., Robbins, V. H. & Baldwin, R. L. The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry* **32,** 9668–76 (1993).

15. Phelan, J., Skelton, N., Braisted, A. & McDowell, R. A General Method for Constraining Short Peptides to an α-Helical Conformation. *J. Am. Chem. Soc.* **119,** 455–460 (1997).

16. Jackson, D., King, D., Chmielewski, J., Singh, S. & Schultz, P. General approach to the synthesis of short .alpha.-helical peptides. *J. Am. Chem. Soc.* **113,** 9391–9392 (1991).

17. Walensky, L. D. *et al.* Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* **305,** 1466–70 (2004).

18. Karle, IL & Balaram, P. Structural characteristics of. alpha.-helical peptide molecules containing Aib residues. *Biochemistry* (1990).

19. Kim, Y.-W. W., Grossmann, T. N. & Verdine, G. L. Synthesis of all-hydrocarbon stapled α-helical peptides by ring-closing olefin metathesis. *Nat Protoc* **6,** 761–71 (2011).

20. Verdine, G. & Hilinski, G. All-hydrocarbon stapled peptides as Synthetic Cell-Accessible Mini-Proteins. *Drug Discov Today Technol* **9,** e1–e70 (2012).

21. Verdine, G.L. & Hilinski, G.J. Stapled peptides for intracellular drug targets. *Meth. Enzymol.* **503**, 3-33 (2012).

22. Walensky, L. D. & Bird, G. H. Hydrocarbon-stapled peptides: principles, practice, and progress. *J. Med. Chem.* **57,** 6275–88 (2014).

23. Walensky, L. D. *et al.* A stapled BID BH3 helix directly binds and activates BAX. *Mol. Cell* **24,** 199–210 (2006).

24. Danial, N. N. *et al.* Dual role of proapoptotic BAD in insulin secretion and beta cell survival. *Nat. Med.* **14,** 144–53 (2008).

25. Gavathiotis, E. *et al.* BAX activation is initiated at a novel interaction site. *Nature* **455,** 1076–81 (2008).

26. Moellering, R. *et al.* Direct inhibition of the NOTCH transcription factor complex. *Nature* **462,** 182–188 (2009).

27. Stewart, M. L., Fire, E., Keating, A. E. & Walensky, L. D. The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nat. Chem. Biol.* **6,** 595–601 (2010).

28. Bhattacharya, S., Zhang, H., Debnath, A. K. & Cowburn, D. Solution structure of a hydrocarbon stapled peptide inhibitor in complex with monomeric C-terminal domain of HIV-1 capsid. *J. Biol. Chem.* **283,** 16274–8 (2008).

29. Kim, Y.-W. W. & Verdine, G. L. Stereochemical effects of all-hydrocarbon tethers in i,i+4 stapled peptides. *Bioorg. Med. Chem. Lett.* **19,** 2533–6 (2009).

30. Kim, Y.-W., Kutchukian, P. & Verdine, G. Introduction of All-Hydrocarbon *i , i* +3 Staples into α-Helices via Ring-Closing Olefin Metathesis. *Organic Letters* **12,** 3046–3049 (2010).

31. Zhang, H. *et al.* A cell-penetrating helical peptide as a potential HIV-1 inhibitor. *J. Mol. Biol.* **378,** 565–80 (2008).

32. Baek, S. *et al.* Structure of the stapled p53 peptide bound to Mdm2. *J. Am. Chem. Soc.* **134,** 103–6 (2012).

33. Phillips, C. *et al.* Design and structure of stapled peptides binding to estrogen receptors. *J. Am. Chem. Soc.* **133,** 9696–9 (2011).

34. Bierzynski, A., Kim, P. S. & Baldwin, R. L. A salt bridge stabilizes the helix formed by isolated C-peptide of RNase A. *Proc. Natl. Acad. Sci. U.S.A.* **79,** 2470–4 (1982).

35. Trnka, T. M. & Grubbs, R. H. The development of L_2X_2Ru=CHR olefin metathesis catalysts: an organometallic success story. *Acc. Chem. Res.* **34,** 18–29 (2001).

36. Scholl, M., Ding, S., Lee, C. W. & Grubbs, R. H. Synthesis and activity of a new generation of ruthenium-based olefin metathesis catalysts coordinated with 1,3-dimesityl-4,5-dihydroimidazol-2-ylidene ligands. *Org. Lett.* **1,** 953–6 (1999).

37. Getty, K., Delgado-Jaime, M. U. & Kennepohl, P. An electronic rationale for observed initiation rates in ruthenium-mediated olefin metathesis: charge donation in phosphine and N-heterocyclic carbene ligands. *J. Am. Chem. Soc.* **129,** 15774–6 (2007).

38. Garber, S., Kingsbury, J., Gray, B. & Hoveyda, A. Efficient and Recyclable Monomeric and Dendritic Ru-Based Metathesis Catalysts. *J. Am. Chem. Soc.* (2000).

39. Pham, T. K., Yoo, J. & Kim, Y.-W. Comparison of Oct-2-enyl and Oct-4-enyl Staples for Their Formation and Helix Stabilizing Effects. *Bull. Korean Chem. Soc* **34,** 2641 (2013).

40. Dimartino, G., Wang, D., Chapman, R. N. & Arora, P. S. Solid-phase synthesis of hydrogen-bond surrogate-derived alpha-helices. *Org. Lett.* **7,** 2389–92 (2005).

41. Shim, S. Y., Kim, Y.-W. W. & Verdine, G. L. A new i, i + 3 peptide stapling system for α-helix stabilization. *Chem Biol Drug Des* **82,** 635–42 (2013).

42. Manning, MC & Woody, RW. Theoretical CD studies of polypeptide helices: examination of important electronic and geometric factors. *Biopolymers* (1991).

43. Toniolo, C, Polese, A & Formaggio, F. Circular dichroism spectrum of a peptide 3_{10}-helix. *J. Am. Chem. Soc.* **118**, 2744-2745 (1996).

44. Vieira-Pires, RS & Morais-Cabral, JH. 3_{10} helices in channels and other membrane proteins. *J. Gen. Physiol.* **136**, 585-592 (2010).

45. Novotny, M. & Kleywegt, G. J. A survey of left-handed helices in protein structures. *J. Mol. Biol.* **347,** 231–41 (2005).

46. Bird, GH, Madani, N & Perry, AF. Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14093-14098 (2010).

47. Urnes, P. & Doty, P. Optical rotation and the conformation of polypeptides and proteins. *Adv. Protein Chem.* **16,** 401–544 (1961).

48. Jacobsen, *Ø. et al.* Stapling of a 3_{10}-helix with click chemistry. *J. Org. Chem.* **76,** 1228–38 (2011).

49. Du, H, Fuh, R., Li, J., Corkan, L. & Lindsey, J. PhotochemCAD: A Computer-Aided Design and Research Tool in Photochemistry. *Photochem. Photobiol.* (1998).

50. Barbacid, M. Ras genes. *Annu. Rev. Biochem.* **56,** 779–827 (1987).

51. Lowy, D. R. & Willumsen, B. M. Function and regulation of ras. *Annu. Rev. Biochem.* **62,** 851–91 (1993).

52. Scheffzek, K, Ahmadian, MR, Kabsch, W & Wiesmüller, L. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* 277, 333-338 (1997).

53. McCormick, F. Ras GTPase activating protein: signal transmitter and signal terminator. *Cell* **56,** 5–8 (1989).

54. Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* **348,** 125–32 (1990).

55. Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294,** 1299–304 (2001).

56. John, J, Sohmen, R, Feuerstein, J & Linke, R. Kinetics of interaction of nucleotides with nucleotide-free H-ras p21. *Biochemistry* (1990).

57. Pai, E. F. *et al.* Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 A resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **9,** 2351–9 (1990).

58. Milburn, M. V. *et al.* Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247,** 939–45 (1990).

59. Filchtinski, D. *et al.* What makes Ras an efficient molecular switch: a computational, biophysical, and structural study of Ras-GDP interactions with mutants of Raf. *J. Mol. Biol.* **399,** 422–35 (2010).

60. Minato, T, Wang, J, Akasaka, K & Okada, T. Quantitative analysis of mutually competitive binding of human Raf-1 and yeast adenylyl cyclase to Ras proteins. *J. Biol. Chem.* **269**, 20845-20851 (1994).

61. Nassar, N. *et al.* The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature* **375,** 554–60 (1995).

62. Huang, L., Hofer, F., Martin, G. S. & Kim, S. H. Structural basis for the interaction of Ras with RalGDS. *Nat. Struct. Biol.* **5,** 422–6 (1998).

63. Stephen, A. G., Esposito, D., Bagni, R. K. & McCormick, F. Dragging ras back in the ring. *Cancer Cell* **25,** 272–81 (2014).

64. Willumsen, BM, Christensen, A & Hubbert, NL. The p21 ras C-terminus is required for transformation and membrane association. *Nature* **310**, 583-6 (1984).

65. Prior, I. A., Lewis, P. D. & Mattos, C. A comprehensive survey of Ras mutations in cancer. *Cancer Res.* **72,** 2457–67 (2012).

66. Potenza, N. *et al.* Replacement of K-Ras with H-Ras supports normal embryonic development despite inducing cardiovascular pathology in adult mice. *EMBO Rep.* **6,** 432–7 (2005).

67. Buhrman, G., Holzapfel, G., Fetics, S. & Mattos, C. Allosteric modulation of Ras positions Q61 for a direct role in catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 4931–6 (2010).

68. Kohl, N. E. *et al.* Inhibition of farnesyltransferase induces regression of mammary and salivary carcinomas in ras transgenic mice. *Nat. Med.* **1,** 792–7 (1995).

69. James, G. L., Goldstein, J. L. & Brown, M. S. Polylysine and CVIM sequences of K-RasB dictate specificity of prenylation and confer resistance to benzodiazepine peptidomimetic in vitro. *J. Biol. Chem.* **270,** 6221–6 (1995).

70. Lito, P., Rosen, N. & Solit, D. B. Tumor adaptation and resistance to RAF inhibitors. *Nat. Med.* **19,** 1401–9 (2013).

71. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503,** 548–51 (2013).

72. Leshchiner, E. S. *et al.* Direct inhibition of oncogenic KRAS by hydrocarbon-stapled SOS1 helices. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 1761–6 (2015).

73. Herrmann, C., Martin, G. A. & Wittinghofer, A. Quantitative analysis of the complex between p21ras and the Ras-binding domain of the human Raf-1 protein kinase. *J. Biol. Chem.* **270,** 2901–5 (1995).

74. Boder, E.T. & Wittrup, K.D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotech.* **15**, 553-7 (1997).

75. Boder, E. T. & Wittrup, K. D. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Meth. Enzymol.* **328,** 430–44 (2000).

76. Wadle, A. *et al.* Serological identification of breast cancer-related antigens from a Saccharomyces cerevisiae surface display library. *Int. J. Cancer* **117,** 104–13 (2005).

77. Kieke, M. C. *et al.* Selection of functional T cell receptor mutants from a yeast surface-display library. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 5651–6 (1999).

78. Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M. & Wittrup, K. D. Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J. Mol. Biol.* **292,** 949–56 (1999).

79. Bowley, D. R., Labrijn, A. F., Zwick, M. B. & Burton, D. R. Antigen selection from an HIV-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage. *Protein Eng. Des. Sel.* **20,** 81–90 (2007).

80. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nat Protoc* **1,** 755–68 (2006).

81. Blundell, T. L., Pitts, J. E., Tickle, I. J., Wood, S. P. & Wu, C. W. X-ray analysis (1. 4-A resolution) of avian pancreatic polypeptide: Small globular protein hormone. *Proc. Natl. Acad. Sci. U.S.A.* **78,** 4175–9 (1981).

82. Zondlo, NJ & Schepartz, A. Highly specific DNA recognition by a designed miniature protein. *J. Am. Chem. Soc.* **121**, 6938-9 (1999).

83. Russell, S. J., Blandl, T., Skelton, N. J. & Cochran, A. G. Stability of cyclic beta-hairpins: asymmetric contributions from side chains of a hydrogen-bonded cross-strand residue pair. *J. Am. Chem. Soc.* **125,** 388–95 (2003).

84. Cochran, A. G. *et al.* A minimal peptide scaffold for beta-turn display: optimizing a strand position in disulfide-cyclized beta-hairpins. *J. Am. Chem. Soc.* **123,** 625–32 (2001).

85. Pacold, M. E. *et al.* Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. *Cell* **103,** 931–43 (2000).

86. Pagel, M. D. & Wemmer, D. E. Solution structure of a core peptide derived from scyllatoxin. *Proteins* **18,** 205–15 (1994).

87. Martins, J. C., Van de Ven, F. J. & Borremans, F. A. Determination of the three-dimensional solution structure of scyllatoxin by 1H nuclear magnetic resonance. *J. Mol. Biol.* **253,** 590–603 (1995).

88. Swers, J. S., Kellogg, B. A. & Wittrup, D. K. Shuffled antibody libraries created by in vivo homologous recombination and yeast surface display. *Nucleic Acids Res. re* **32,** e36–e36 (2004).

89. Robinson, AS & Wittrup, KD. Constitutive overexpression of secreted heterologous proteins decreases extractable heavy chain binding protein and protein disulfide isomerase levels in Saccharomyces cerevisiae. *Biotechnol. Prog.* **11**, 171-177 (1995).

90. McCullum, EO, Williams, B. & Zhang, J. Random mutagenesis by error-prone PCR. *Methods in Mol. Biol.* **634** 103-111(2010).

91. Hodges, A. M. & Schepartz, A. Engineering a monomeric miniature protein. *J. Am. Chem. Soc.* **129,** 11024–5 (2007).

92. Geyer, M. *et al.* Conformational transitions in p21ras and in its complexes with the effector protein Raf-RBD and the GTPase activating protein GAP. *Biochemistry* **35,** 10308–20 (1996).

93. Spoerner, M., Herrmann, C., Vetter, I. R., Kalbitzer, H. R. & Wittinghofer, A. Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 4944–9 (2001).

94. Linnemann, T, Geyer, M, Jaitner, BK & Block, C. Thermodynamic and kinetic characterization of the interaction between the Ras binding domain of AF6 and members of the Ras subfamily. *J. Biol. Chem.* **274**, 13556-6 (1999).

95. Geyer, M., Herrmann, C., Wohlgemuth, S., Wittinghofer, A. & Kalbitzer, H. R. Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling. *Nat. Struct. Biol.* **4,** 694–9 (1997).

96. Muraoka, S. *et al.* Crystal structures of the state 1 conformations of the GTP-bound H-Ras protein and its oncogenic G12V and Q61L mutants. *FEBS Lett.* **586,** 1715–8 (2012).

97. Hilinski, G. J. *et al.* Stitched α-helical peptides via bis ring-closing metathesis. *J. Am. Chem. Soc.* **136,** 12314–22 (2014).

98. Van den Beucken, T. *et al.* Affinity maturation of Fab antibody fragments by fluorescent-activated cell sorting of yeast-displayed libraries. *FEBS Lett.* **546,** 288–94 (2003).

99. Boder, E. T., Bill, J. R., Nields, A. W., Marrack, P. C. & Kappler, J. W. Yeast surface display of a noncovalent MHC class II heterodimer complexed with antigenic peptide. *Biotechnol. Bioeng.* **92,** 485–91 (2005).

100. Strub, M. P. *et al.* Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* **11,** 1359–67 (2003).

101. Salgado, P. S., Taylor, J. D., Cota, E. & Matthews, S. J. Extending the usability of the phasing power of diselenide bonds: SeCys SAD phasing of CsgC using a non-auxotrophic strain. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 8–13 (2011).