Solving Molecular Recognition Problems with Evolvable Peptide Motifs

Thesis by

Ryan James Austin

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2007

(Defended April 23, 2007)

Forward

When I first joined Rich's group, his was one of a relatively small number of labs using combinatorial approaches to explore the boundaries of biological chemistry. I was particularly excited to learn what *in vitro* selection experiments might teach us about natural selection and the mechanisms of molecular evolution. To this end, I have focused my research on the study of peptide molecular recognition, characterizing the relationship between peptide amino-acid sequence and specificity in an effort to better understand the plasticity of these biological polymers. One story that I believe is emerging from this research is that peptide molecular recognition is highly evolvable, requiring only a small number of sequence mutations to affect significant changes in binding specificity. This evolvability may in part explain the increased incidence of unfolded protein sequences in differentiated organisms and could be useful to investigators searching for new molecular tools and capture agents.

In my studies, I have been fortunate to collaborate with several excellent experimentalists including Dr. Terry Takahashi, Dr. William Ja, and Dr. Scott Ross, whose assistance has been instrumental to me. Other colleagues whom I would like to thank include Dr. Diasuke Umeno, Dr. Jeffrey Barrick, Dr. Jinsong Ren, Dr. Tianbing Xia, Dr. Shuwei Li, Dr. Shelley Starck, Dr. Karin Crowhurst, Dr. Adam Frankel, Dr. Christine Ueda, Anders Olson and Steven Millward. I'm grateful for the counsel of my advisory committee, which includes Professor Stephen Mayo, Professor Douglas Rees, and Professor Peter Dervan; and for the mentorship and support of my advisor, Professor 'Rich' Roberts.

Table of Contents

List of Figures

List of Tables

Chapter 1: Solving molecular recognition problems with evolvable peptide motifs

Ryan J. Austin

Abstract

Specific protein-nucleic acid and protein-protein recognition events are frequently mediated through the flexible binding surfaces of these polymers. The functional plasticity and sequence conservation of these surfaces suggests that they are highly evolvable molecular recognition sites. It may therefore be possible to develop discriminate ligands for many protein and nucleic acid targets by directed evolution of consensus ligand scaffolds or motifs. Here we review and present work on the development and use of peptide motifs to evolve high-specificity ligands toward flexible RNA-hairpin and G protein targets. The evolvabilities of these motifs and the compact arrangement of specificity-determining elements in selected sequences, demonstrate the economy of motif-based directed evolution approaches.

Introduction

Our understanding of molecular recognition in cellular processes has expanded from static models to encompass a continuum of structurally adaptive binding interfaces. Adaptive recognition mechanisms involving the coupled folding and binding of one or both binding partners figure prominently in protein-nucleic acid binding events [1], and interestingly, are being found to play a growing role in protein-protein interactions [2]. Such mechanisms range from global fold rearrangement in prions [3], to secondary structure and side-chain conformational flexibility observed in the binding surfaces of germ-line antibodies and preferred protein-protein binding 'hot spots' [4, 5]. The correlation between conformational flexibility and functional plasticity in these binding partners suggests that dynamic binding sites are inherently more evolvable than rigid folds [6, 7]. Bio-informatic analysis has revealed a growing incidence of unstructured protein domains in cell signaling and transcription regulation pathways of higher eukaryotes, suggesting that the evolvable properties of dynamic folds have facilitated the differentiation of complex signaling networks in these organisms [8].

The wealth of these flexible targets in the cell, and the evolvability of their binding specificities, presents an intriguing problem to biologists interested in developing discriminate molecular tools and sensors. If the binding specificities of these flexible targets are inherently evolvable, tools that discriminate these targets should likewise be evolvable. It may therefore be possible to generate specific ligands for many cellular targets by way of directed evolution experiments, using promiscuous peptide motifs as starting points. In a selection or 'directed evolution' experiment, a selective pressure,

such as the ability to bind a particular target, is applied to a degenerate pool of molecules. The fittest molecules are reproduced, coming to dominate the molecular pool after iterative rounds of selection, which allows for their identification. The larger the initial pool of molecules searched, the more laborious the selection process. However, if selection pools can be optimized by incorporation of conserved scaffolds or motifs, the development of discriminate ligands for the detection of or modulation of various cellular interactions could be expedited. In approaching this problem, peptide motifs offer several advantages as molecular starting points, including their mutability, modulatable chemistries, and demonstrated selectability against flexible molecular targets [9-11].

Here we present work on the development and use of promiscuous peptide motifs to evolve high-specificity ligands towards two types of flexible cellular target: the λboxB RNA hairpin, and the signal transduction Gα protein. The work indicates that conserved, dynamic binding surfaces can be targeted with significant specificities by relatively short peptides. Discrete mutations adjacent to and framed within conserved motifs of these peptides confer dramatic changes in binding specificity, underscoring the evolvability of these sequences.

RNA Targets

Targeting flexible RNA structures with the ARM-consensus motif

A number of natural protein scaffolds that bind to RNA have been characterized, including the RNA binding domain (RBD or RRM) and KH domain [12], zinc finger and arginine-rich motif (ARM) [13, 14]. Of these, the ARM presents the most concise and versatile framework for development of novel RNA binding peptides. The ARM was originally identified by Asis Das's lab in bacteriophage and is not, as the name implies, a discrete motif, but rather an arginine rich sequence [15]. Subsequently, Alan Frankel's group identified a loose consensus motif (T/R)RXXRR (where X represents any amino-acid), referred to here as the 'ARM-consensus motif', present as a sub-sequence in a great number of RNA binding peptides and proteins [16]. Directed evolution experiments investigating the HIV Rev and λ N peptide scaffolds have expanded the phylogenic pool of ARMs [17-21], and peptide sequences from these experiments along with a sampling of natural ARMs are listed in Figure 1A. These experiments, in addition to RNA aptamer selections and NMR structural analyses have demonstrated the evolvability and structural plasticity of the ARM-framework [10, 22].

Directed evolution experiments from Alan Frankel's group have presented themes relating the structural plasticity of ARM peptides to sequence elements arrayed about the ARM-consensus motif. Incorporating the (T/R)RXXRR motif at the amino-terminus of a Rev-like peptide library, the group evolved peptides with increased affinities and specificities for the Rev cognate RNA hairpin, RRE-IIB [17-19, 23]. Two types of peptide solution emerged from the selections: one maintaining the α-helical structure of

the Rev peptide, and a second adopting an extended backbone structure. These two peptide solutions employ distinct sequence elements in and around the consensus motif. The α-helical peptide sequences are enriched for acidic residues, which stabilize the secondary structure of the peptide, as well as a glutamine residue, adjacent to the motif, which makes specific contacts with a sheared G-A base pair in the RRE-IIB hairpin. In contrast, extended peptide solutions incorporate glycine and proline residues in and around the consensus motif. These residues are critical for shape-specific recognition in the evolved RSG-1.2-RRE-IIB complex [24, 25], and their importance is reiterated in various ARM-RNA complexes. A proline residue framed within the consensus motif of the HTLV-1 Rex peptide constrains a shape-specific S-like peptide fold in the Rex-RNA aptamer complex [26]. Separately, the consensus motif of HIV Tat peptide does not contain any glycine residues and adopts an unstructured fold in the HIV Tat-TAR complex [27, 28], while the BIV Tat peptide contains a glycine residue substitution in the consensus motif, allowing the peptide to adopt a β-hairpin fold in the cognate BIV Tat-TAR complex [29, 30]. In a striking example of structural plasticity, the JDV Tat peptide adopts multiple functional structures, binding as a β-hairpin to JDV- and BIV-TAR RNAs, but adopting an extended conformation in complex with the non-cognate HIV-TAR RNA [31]. The solution structures of these complexes, as well as several ARM-RNA aptamer complexes solved by Dinshaw Patel's group [26, 32], further illustrate the structural plasticity of the ARM-consensus motif. In general, ARM-RNA structural analyses support a shape-specific mechanism of molecular recognition where RNA architecture dictates peptide fold (Reviewed in [13, 14, 33, 34]).

After the Rev-RRE complex, the ARM-RNA system most studied by NMR structural analysis and peptide directed evolution is the bacteriophage N peptide-boxB hairpin complex. Similar structures of cognate λ, P22, and ϕ21 N peptide-boxB complexes, as well as the HK022 Nun-λboxB complex, have been solved by NMR [35-38]. In each of the complexes, the peptide adopts a bend that divides it into two modules; an amino-terminal helix containing the ARM-consensus motif, and a carboxy-terminal helix that makes shape-specific contacts with the boxB hairpin loop. Shape complementarity is particularly evident in the λ complex, where the N peptide adopts a 60 degree bend to accommodate π–π stacking interactions between a tryptophan residue and the boxB hairpin loop (Fig. 1B). The conclusion drawn from this shape-specific recognition has been that the carboxy-terminal helix of the N peptide dictates loop binding specificity, while the ARM-consensus motif in the amino-terminal helix functions as a conserved stem-binding module [37, 39]. However, directed evolution experiments employing a λ N peptide amino-terminal scaffold (mdaqTRRreRRa-X_{10}) have indicated that the specificity of λ N peptide is dictated by its amino-terminal helix [20, 40].

This finding contradicts the modular formulation of shape-specific binding and exposes a weakness in our present understanding of ARM-RNA molecular recognition. Despite the breadth of NMR structural information elucidated in the past 10 years, fundamental binding-specificity concepts remain unresolved. For instance, is the ARM-consensus motif to be thought of as a binding module that works in complement with other modules, or is the motif better understood as a symmetric platform upon which specificity-determining elements are arrayed? Such concepts are particularly relevant to

the design of new RNA-binding ligands using peptide scaffolds. To this effect, it remains important to identify the specificity-determining elements within the ARM-framework and the relationship of these elements to the ARM-consensus motif.

Most studies of ARM-RNA binding specificity have relied on the Gel-Mobility Shift Assay (GMSA), which presumes a kinetically stable binding interaction [41, 42]. Many ARM-RNA complexes are kinetically unstable, however, owing to the electrostatic nature of the binding interaction. Electrostatic steering forces between positively charged ARMs and negatively charged RNAs can increase the association rate between these binding partners beyond the limits of a diffusion controlled process ($> 1 \times 10^9$ $s^{-1} M^{-1}$). In such instances, even a thermodynamically stable ARM-RNA interaction with a dissociation constant of 1 nM will be kinetically unstable, having a dissociation rate of less than a second ($K_D = k_{off}/k_{on}$). Alternatives to the GMSA employing affinity electrophoresis have been developed for the study of ARM-RNA complexes and are capable of evaluating less-stable binding pairs [43], but the charge gradient employed in these gel based assays can perturb the ARM-RNA binding equilibrium. Fluorescence-based techniques are sensitive to kinetics and offer a preferable means of binding measurement to the GMSA. The efficacy of one such approach, substituting the fluorescent base 2-aminopurine (2AP) at RNA loops and bulges, has been demonstrated in several ARM-RNA interactions [20, 44-46].

Using 2AP-labeled RNA hairpins, our lab has performed a combinatorial analysis of amino-acid binding energetics in the λ N peptide (Chapter 2). Coupled binding and folding events often involve intra-molecular binding cooperativities that are invisible to traditional alanine-scanning techniques, but can be identified by combinatorial or pair-

wise mutagenesis experiments [47-51]. This type of analysis has isolated two specificity-determining residues in the λ N peptide, Gln4 and Arg8 (Fig. 1C), which recapitulate the binding specificity of the wild-type peptide when grafted into a non-specific ARM-framework. The Gln4 residue immediately adjacent to the ARM-consensus motif of λ N peptide makes specific contacts with a sheared G-A loop closing base pair of the λboxB hairpin, similar to glutamine G-A recognition in Rev-RRE. Interestingly, the Arg8 residue, framed inside the ARM-consensus motif, exhibits cooperative binding with all non-conserved elements of the λ N peptide. The binding specificity of this residue arises from shape-specific contacts within the λboxB pentaloop that stabilize the induced fit of the complex. This mode of loop recognition is illustrated by the non-cognate λ N peptide-P22boxB complex. In this unfavorable complex the λ N peptide forces the P22boxB pentaloop into a strained λboxB-like conformation, which can be fully alleviated by a simple substitution of the Arg8 residue. Based on these findings, our lab has been able to engineer discriminate peptides with orders of magnitude greater affinity for the λboxB target than the natural λ N peptide (Chapter 3). Results from λ N peptide studies support the evolvability of the ARM-framework within and around the ARM-consensus motif and add evidence to observations of peptide dictating local RNA fold [22, 25]. Though more work will be necessary to characterize the relationship between the ARM-consensus motif and specificity-determining elements within the ARM-framework, analysis of the N peptide has demonstrated that short peptide sequences containing the ARM-consensus motif are capable of discriminating RNA targets.

Returning to the problem of designing RNA-binding peptide scaffolds, it is interesting to note that the specificity-determining elements of λ N peptide are arrayed

immediately adjacent to and within the ARM-consensus motif. In the case of the Trp18 residue, which is 7 residues from the ARM-consensus motif of λ N peptide, experiments from our lab have demonstrated that shape-specific stacking interactions are critical to biological function, not binding specificity [21]. Interestingly, a two-residue λ N peptide mutant that binds λboxB with increased specificity (λ N($E_{14}R_{15}$)), adopts an unbent α-helical structure extending 4 residues to either side of the ARM-consensus motif (Chapter 4: Fig. 4.1C). NMR dynamic analysis indicates that the λ N($E_{14}R_{15}$) peptide fold becomes disordered 7 residues from the ARM-consensus motif, similar to structural observations in many ARM-RNA complexes, which show disorder distal to the consensus motif [26, 29, 30, 36, 52, 53]. Recent selection experiments offer additional evidence of specificity-determining elements adjacent to the ARM-consensus motif. In a striking example, two independent peptide selection experiments performed against the P6.1 telomerase hairpin dimer have isolated specific peptides with a core RKYXRV motif flanked by conserved 7 residue sequences [54, 55]. Separately, a study of peptide binding RNA-aptamers in Andrew Ellington's lab has demonstrated the paucity of specificity-determining elements further than 6 residues from the ARM-consensus motif of Rev [22]. These findings suggest that a relatively small peptide library, incorporating an ARM-consensus motif, flanked by randomized amino-acid heptamers (X_7-(T/R)RXXRR-X_7) will contain discriminate peptide binding solutions for a wide variety of RNA targets.

Protein Targets

Targeting a flexible Gα hot spot with an aromatic-consensus motif

Preferred protein-protein interaction surfaces, or protein binding 'hot spots', are often highly conserved across differentiated classes of proteins, exhibiting preferred physiochemical compositions and conformational dynamics that predispose these surfaces to targeting by *in vitro* selection experiments [56]. Structural analyses and selection investigations performed on the signal transduction protein Gα have identified a hot spot at the effector-binding site of this subunit. The structural plasticity and evolvability of this binding surface make it a potential target for the discrimination of Gα subunits using motif-based directed evolution experiments.

There are 19 unique Gα subunits in humans, categorized into 4 classes (i/o, q/11, s, 12/13) that describe the downstream effector coupling of the G protein [57]. Differentiation of these Gα classes correlates with increasing complexity of cellular function in higher organisms, as only two Gα subunits are expressed in the yeast *S. cerevisiae* and four Gα subunits in the nematode *C. elegans* [58]. The high degree of sequence conservation in the effector-binding site of differentiated Gα subunits is illustrated in Figure 1.2A. This conserved site, indicated with an asterisk in Figure 1.2, includes a malleable switch II (SII) structural element and a more stable α-helical element (α3). Conformational stability of SII is coupled to GTP binding within the active-site of Gα, allowing for temporal regulation of the 'SII-α3' effector-binding site [59-62].

The preferred binding character of a hydrophobic pocket in the SII-α3 cleft has been underscored by several independent *in vitro* selection experiments against Gαi1, which have targeted this site (Fig. 1.2) [63-67]. A family of selected peptides containing a conserved h(T/Y)W(W/Y)EFL "aromatic-consensus motif" (where h represents a hydrophobic amino-acid), bind specifically to Gα, but exhibit limited Gα class-binding specificity (Fig. 1.2C). Separate crystal structures of Gαi1 bound to the R6A-1 and KB-752 peptides demonstrate docking of the peptide EFL-sequence within the hydrophobic binding pocket of SII-α3 (Fig. 1.2B)[64, 68]. While the R6A-1 peptide exhibits a clear binding preference for the Gαi1 subunit over Gαs(s), it is unclear from the crystal structure how this specificity is conferred. Paul Sigler has suggested that the effector-binding site could be discriminated by specific contacts with the α3 helix, which shows some variability across Gα classes [60]. A structural survey of effector-binding to the four classes of Gα suggests a different mechanism of binding discrimination, where non-specific contacts between the effector and the SII-α3 hydrophobic binding pocket are complemented by specific interactions outside of the SII-α3 site, at the α3-β5 and α2-β4 loops of Gα [62]. Structural analysis has, however, been unable to elucidate the origins of binding specificity for Gαi-specific adenylyl cyclase effectors, which interact with the SII-α3 site of the subunit [69]. In general, the malleability of the SII-α3 binding site has complicated structure-based analyses of binding specificity, leaving the question of molecular discrimination within the SII-α3 site largely unanswered.

To explore the specificity of the SII-α3 binding site, our lab has performed a two-step directed evolution experiment (Chapter 5). In the first step, the Gαi-specific peptide R6A-1, flanked by random amino-acid hexamers, was evolved to bind both Gαi and Gαs

subunit classes. This Gαs-binding peptide (GSP), was then evolved a step further into matured GSP (mGSP) sequences, which bind specifically to Gαs. Mutagenic experiments indicate that GSP and mGSP peptides target the SII-α3 site of Gαs(s), making discriminate contacts with α3 and the α3-β5 loop of the subunit. These findings present two interesting themes. Firstly, the success of mGSP peptides at discriminating Gαs subunit targets with class and even sub-class binding specificity demonstrates that relatively short peptide sequences can distinguish very similar Gα targets. The 36 residue G protein regulatory motif (GPR or GoLoco)(Fig. 1.2C) has previously been demonstrated to discriminate Gα targets within the i/o class of subunits, but this specificity requires extensive peptide contacts with the helical-domain of Gα, necessitating a relatively long peptide sequence [70]. mGSP peptides are comparatively short and appear able to discriminate Gα targets within the highly conserved SII-α3 site. It should be noted that the tetra-decapeptide mastoparan is also capable of discriminating Gα subunit classes, albeit at much higher effective concentrations [71]. Secondly, the short mutagenic distances separating the specific mGSP peptide sequences from the non-specific aromatic-consensus motif, indicate that Gα-binding specificity can be evolved using relatively small combinatorial searches. Analysis of conserved residues in mGSP sequences highlights the importance of $K_1(R/L)_2$ and V_5R_6 specificity-determining residues. These residues are immediately adjacent and internal to the aromatic-consensus motif, suggesting that a relatively small peptide library based on a broader consensus motif, such as X_5-h(T/Y)XXEFL-X_6, will contain specific binding solutions for a variety of Gα subunits.

Conclusion

Flexible molecules involved in transcription regulation and cellular signaling networks present a growing pool of molecular targets [72]. Selection experiments against the transcription regulator λboxB RNA and the signal transduction Gα protein have demonstrated that these dynamic molecular targets can be discriminated with *in vitro* specificities by small peptides. Specific peptides evolved from the ARM-consensus motif and the Gα-binding aromatic-consensus motif contain relatively few amino-acid substitutions within and around the motif, demonstrating the evolvability of these short sequences [73, 74] (Chapter 5).

Peptide motifs are one means to reduce the sequence complexity of a selection search and can have the advantage of targeting a selection to a particular surface of a molecule, as demonstrated by peptide selections against Gαs (Chapter 5). Localized targeting by Jim Wells' group has similarly demonstrated the efficacy of this approach using small molecules [75, 76]. It may be possible to focus peptide combinatorial search complexity further by employing smart amino-acid vernaculars. A comparison of the RNA-binding arginine-rich motif (Fig. 1.1A) and the G protein-binding aromatic motif (Fig. 1.2C) illustrates the preferential amino-acid composition of nucleotide- and protein-binding peptides. Arginine is a versatile amino-acid, capable of participating in electrostatic, hydrogen bonding, and van der Waals interactions, and its importance in nucleotide-binding proteins has been well documented [77]. Separately, analyses of the amino-acid composition of protein hot spot surfaces and antibody binding sites have revealed an increased incidence of aromatic residues [78, 79]. Like arginine, aromatic

residues can participate in a variety of binding interactions, including hydrophobic contacts, π-bonds, and in the case of tryptophan and tyrosine, hydrogen bonds. It is interesting to find that these versatile residues are also critical in selected small peptide binding solutions. Independent selection experiments in our own laboratory have identified similar -R-W-R- binding motifs for the Methuselah protein and a phospho-serine peptide target [80, 81]. Separately, selection experiments using a limited 4 amino-acid vernacular, have evolved discriminate antigen-binding sites that are dominated by tyrosine residues. [82, 83]. The binding versatility of arginine and aromatic amino-acids, coupled with precise doping methods afforded by phosphoramidite codon-triplet chemistries [84], will afford a powerful complement to motif-based peptide selections.

One practical complication from using motifs and codon biases in selection experiments against single targets is that these designs can bias a selection in the wrong direction. It is often enough the case that a selection experiment walks around all of the helpful hands that investigators try to offer it [85]. So goes the saying in directed evolution labs, 'You get what you select for.' One way to counter the limited success rate of selection experiments is to increase the number of targets searched. Nucleotide targets are relatively easy to synthesize and *in vitro* expression systems offer a promising means for economical development of protein targets [86], however, the low yield of these expressions complicates naïve selection experiments. By increasing the percentage of target binding molecules in a combinatorial library, motif-based peptide selections offer a means to counter the low yields of *in vitro* expressed proteins. Such designs should likewise be useful in selection experiments against cell surfaces, where distinct receptor targets are presented at low densities [87].

References

1. Spolar, R.S., and Record, M.T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science *263*, 777-784.
2. Wright, P.E., and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol *293*, 321-331.
3. Chien, P., and Weissman, J.S. (2001). Conformational diversity in a yeast prion dictates its seeding specificity. Nature *410*, 223-227.
4. Wedemayer, G.J., Patten, P.A., Wang, L.H., Schultz, P.G., and Stevens, R.C. (1997). Structural insights into the evolution of an antibody combining site. Science *276*, 1665-1669.
5. Atwell, S., Ultsch, M., De Vos, A.M., and Wells, J.A. (1997). Structural plasticity in a remodeled protein-protein interface. Science *278*, 1125-1128.
6. Joyce, G.F. (1997). Evolutionary chemistry: getting there from here. Science *276*, 1658-1659.
7. James, L.C., and Tawfik, D.S. (2003). Conformational diversity and protein evolution-a 60-year-old hypothesis revisited. Trends Biochem Sci *28*, 361-368.
8. Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., and Obradovic, Z. (2002). Intrinsic disorder and protein function. Biochemistry *41*, 6573-6582.
9. DeLano, W.L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol *12*, 14-20.
10. Das, C., and Frankel, A.D. (2003). Sequence and structure space of RNA-binding peptides. Biopolymers *70*, 80-85.
11. Ja, W.W., and Roberts, R.W. (2005). G-protein-directed ligand discovery with peptide combinatorial libraries. Trends Biochem Sci *30*, 318-324.
12. Perez-Canadillas, J.M., and Varani, G. (2001). Recent advances in RNA-protein recognition. Curr Opin Struct Biol *11*, 53-58.
13. Draper, D.E. (1999). Themes in RNA-protein recognition. J Mol Biol *293*, 255-270.
14. Frankel, A.D. (2000). Fitting peptides into the RNA world. Curr Opin Struct Biol *10*, 332-340.
15. Lazinski, D., Grzadzielska, E., and Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell *59*, 207-218.
16. Tan, R., and Frankel, A.D. (1995). Structural variety of arginine-rich RNA-binding peptides. Proc Natl Acad Sci U S A *92*, 5282-5286.
17. Harada, K., Martin, S.S., and Frankel, A.D. (1996). Selection of RNA-binding peptides in vivo. Nature *380*, 175-179.
18. Tan, R., and Frankel, A.D. (1998). A novel glutamine-RNA interaction identified by screening libraries in mammalian cells. Proc Natl Acad Sci U S A *95*, 4247-4252.
19. Peled-Zehavi, H., Horiya, S., Das, C., Harada, K., and Frankel, A.D. (2003). Selection of RRE RNA binding peptides using a kanamycin antitermination assay. RNA *9*, 252-261.

20. Barrick, J.E., Takahashi, T.T., Ren, J., Xia, T., and Roberts, R.W. (2001). Large libraries reveal diverse solutions to an RNA recognition problem. Proc Natl Acad Sci U S A *98*, 12374-12378.
21. Xia, T., Frankel, A., Takahashi, T.T., Ren, J., and Roberts, R.W. (2003). Context and conformation dictate the function of a transcriptional switch. Nat Struct Biol *10,* 812-819
22. Bayer, T.S., Booth, L.N., Knudsen, S.M., and Ellington, A.D. (2005). Arginine-rich motifs present multiple interfaces for specific binding by RNA. RNA *11*, 1848-1857.
23. Harada, K., Martin, S.S., Tan, R., and Frankel, A.D. (1997). Molding a peptide into an RNA site by in vivo peptide evolution. Proc Natl Acad Sci U S A *94*, 11887-11892.
24. Zhang, Q., Harada, K., Cho, H.S., Frankel, A.D., and Wemmer, D.E. (2001). Structural characterization of the complex of the Rev response element RNA with a selected peptide. Chem Biol *8*, 511-520.
25. Gosser, Y., Hermann, T., Majumdar, A., Hu, W., Frederick, R., Jiang, F., Xu, W., and Patel, D.J. (2001). Peptide-triggered conformational switch in HIV-1 RRE RNA complexes. Nat Struct Biol *8*, 146-150.
26. Jiang, F., Gorin, A., Hu, W., Majumdar, A., Baskerville, S., Xu, W., Ellington, A., and Patel, D.J. (1999). Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples. Structure *7*, 1461-1472.
27. Puglisi, J.D., Chen, L., Frankel, A.D., and Williamson, J.R. (1993). Role of RNA structure in arginine recognition of TAR RNA. Proc Natl Acad Sci U S A *90*, 3680-3684.
28. Long, K.S., and Crothers, D.M. (1999). Characterization of the solution conformations of unbound and Tat peptide-bound forms of HIV-1 TAR RNA. Biochemistry *38*, 10059-10069.
29. Puglisi, J.D., Chen, L., Blanchard, S., and Frankel, A.D. (1995). Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. Science *270*, 1200-1203.
30. Ye, X., Kumar, R.A., and Patel, D.J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. Chem Biol *2*, 827-840.
31. Smith, C.A., Calabro, V., and Frankel, A.D. (2000). An RNA-binding chameleon. Mol Cell *6*, 1067-1076.
32. Ye, X., Gorin, A., Frederick, R., Hu, W., Majumdar, A., Xu, W., McLendon, G., Ellington, A., and Patel, D.J. (1999). RNA architecture dictates the conformations of a bound peptide. Chem Biol *6*, 657-669.
33. Weiss, M.A., and Narayana, N. (1998). RNA recognition by arginine-rich peptide motifs. Biopolymers *48*, 167-180.
34. Patel, D.J. (1999). Adaptive recognition in RNA complexes with peptides and protein modules. Curr Opin Struct Biol *9*, 74-87.
35. Legault, P., Li, J., Mogridge, J., Kay, L.E., and Greenblatt, J. (1998). NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell *93*, 289-299.

36. Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A., and Patel, D.J. (1998). Solution structure of P22 transcriptional antitermination N peptide-boxB RNA complex. Nat Struct Biol *5*, 203-212.
37. Cilley, C.D., and Williamson, J.R. (2003). Structural mimicry in the phage phi21 N peptide-boxB RNA complex. RNA *9*, 663-676.
38. Faber, C., Scharpf, M., Becker, T., Sticht, H., and Rosch, P. (2001). The structure of the coliphage HK022 Nun protein-lambda-phage boxB RNA complex. Implications for the mechanism of transcription termination. J Biol Chem *276*, 32064-32070.
39. Weiss, M.A. (1998). RNA-mediated signaling in transcription. Nat Struct Biol *5*, 329-333.
40. Barrick, J.E., and Roberts, R.W. (2003). Achieving specificity in selected and wild-type N peptide-RNA complexes: the importance of discrimination against noncognate RNA targets. Biochemistry *42*, 12998-13007.
41. Cann, J.R. (1989). Phenomenological theory of gel electrophoresis of protein-nucleic acid complexes. J Biol Chem *264*, 17032-17040.
42. Carey, J. (1991). Gel retardation. Methods Enzymol *208*, 103-117.
43. Cilley, C.D., and Williamson, J.R. (1997). Analysis of bacteriophage N protein and peptide binding to boxB RNA using polyacrylamide gel coelectrophoresis (PACE). RNA *3*, 57-67.
44. Lacourciere, K.A., Stivers, J.T., and Marino, J.P. (2000). Mechanism of neomycin and Rev peptide binding to the Rev responsive element of HIV-1 as determined by fluorescence and NMR spectroscopy. Biochemistry *39*, 5630-5641.
45. Xia, T., Becker, H.-C., Wan, C., Frankel, A., Roberts, R.W., and Zewail, A.H. (2003). The RNA-protein complex: Direct probing of the interfacial recognition dynamics and its correlation with biological functions. Proc Natl Acad Sci U S A *100*, 8119-8125.
46. Xia, T., Wan, C., Roberts, R.W., and Zewail, A.H. (2005). RNA-protein recognition: single-residue ultrafast dynamical control of structural specificity and function. Proc Natl Acad Sci U S A *102*, 13013-13018.
47. Schreiber, G., and Fersht, A.R. (1995). Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J Mol Biol *248*, 478-486.
48. Albeck, S., Unger, R., and Schreiber, G. (2000). Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges. J Mol Biol *298*, 503-520.
49. Yang, J., Swaminathan, C.P., Huang, Y., Guan, R., Cho, S., Kieke, M.C., Kranz, D.M., Mariuzza, R.A., and Sundberg, E.J. (2003). Dissecting cooperative and additive binding energetics in the affinity maturation pathway of a protein-protein interface. J Biol Chem *278*, 50412-50421.
50. Kranz, J.K., and Hall, K.B. (1999). RNA recognition by the human U1A protein is mediated by a network of local cooperative interactions that create the optimal binding surface. J Mol Biol *285*, 215-231.
51. Jenkins, J.L., and Shapiro, R. (2003). Identification of small-molecule inhibitors of human angiogenin and characterization of their binding interactions guided by computational docking. Biochemistry *42*, 6674-6687.

52. Ye, X., Gorin, A., Ellington, A.D., and Patel, D.J. (1996). Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. Nat Struct Biol *3*, 1026-1033.
53. Cilley, C.D., and Williamson, J.R. (2003). Structural mimicry in the phage phi21 N peptide-boxB RNA complex. RNA *9*, 663-676.
54. Takahashi, T.T. (2005). In vitro selecion of RNA binding peptides. Dissertation thesis, California Institute of Technology, Pasadena.
55. Ueda, C.T. (2006). Targeting human telomerase RNA via biochemical and in vitro selection methods, California Institute of Technology, Pasadena, CA.
56. Ma, B., Wolfson, H.J., and Nussinov, R. (2001). Protein functional epitopes: hot spots, dynamics and combinatorial libraries. Curr Opin Struct Biol *11*, 364-369.
57. Neves, S.R., Ram, P.T., and Iyengar, R. (2002). G protein pathways. Science *296*, 1636-1639.
58. Downes, G.B., and Gautam, N. (1999). The G protein subunit gene families. Genomics *62*, 544-552.
59. Tesmer, J.J., Sunahara, R.K., Gilman, A.G., and Sprang, S.R. (1997). Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS. Science *278*, 1907-1916.
60. Slep, K.C., Kercher, M.A., He, W., Cowan, C.W., Wensel, T.G., and Sigler, P.B. (2001). Structural determinants for regulation of phosphodiesterase by a G protein at 2.0 A. Nature *409*, 1071-1077.
61. Chen, Z., Singer, W.D., Sternweis, P.C., and Sprang, S.R. (2005). Structure of the p115RhoGEF rgRGS domain-Galpha13/i1 chimera complex suggests convergent evolution of a GTPase activator. Nat Struct Mol Biol *12*, 191-197.
62. Tesmer, V.M., Kawano, T., Shankaranarayanan, A., Kozasa, T., and Tesmer, J.J. (2005). Snapshot of activated G proteins at the membrane: the Galphaq-GRK2-Gbetagamma complex. Science *310*, 1686-1690.
63. Ja, W.W., and Roberts, R.W. (2004). In vitro selection of state-specific peptide modulators of G protein signaling using mRNA display. Biochemistry *43*, 9265-9275.
64. Johnston, C.A., Willard, F.S., Jezyk, M.R., Fredericks, Z., Bodor, E.T., Jones, M.B., Blaesius, R., Watts, V.J., Harden, T.K., Sondek, J., Ramer, J.K., and Siderovski, D.P. (2005). Structure of Galpha(i1) bound to a GDP-selective peptide provides insight into guanine nucleotide exchange. Structure *13*, 1069-1080.
65. Johnston, C.A., Lobanova, E.S., Shavkunov, A.S., Low, J., Ramer, J.K., Blaesius, R., Fredericks, Z., Willard, F.S., Kuhlman, B., Arshavsky, V.Y., and Siderovski, D.P. (2006). Minimal Determinants for Binding Activated Galpha from the Structure of a Galpha(i1)-Peptide Dimer. Biochemistry *45*, 11390-11400.
66. Ja, W.W., Wiser, O., Austin, R.J., Jan, L.Y., and Roberts, R.W. (2006). Turning G Proteins On and Off Using Peptide Ligands. ACS Chemical Biology *1*, 570-574.
67. Millward, S.W., Fiacco, S., Austin, R.J., and Roberts, R.W. (2007). Combinatorial Gα Ligand Design Using a Trillion Member Peptide Macrocycle Library. ACS Chemical Biology (*submitted)*.
68. Adhikari, A., and Sprang, S.R. (2007). High Resolution Crystal Structure of the R6A-1-Gαi1 complex. (R.W. Roberts, ed.).

69. Dessauer, C.W., Tesmer, J.J., Sprang, S.R., and Gilman, A.G. (1998). Identification of a Gialpha binding site on type V adenylyl cyclase. J Biol Chem *273*, 25831-25839.
70. Kimple, R.J., Kimple, M.E., Betts, L., Sondek, J., and Siderovski, D.P. (2002). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.
71. Mukai, H., Munekata, E., and Higashijima, T. (1992). G protein antagonists. A novel hydrophobic peptide competes with receptor for G protein binding. J Biol Chem *267*, 16237-16243.
72. Dyson, H.J., and Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol *6*, 197-208.
73. Austin, R.J., Xia, T., Ren, J., Takahashi, T.T., and Roberts, R.W. (2002). Designed Arginine-Rich RNA-Binding Peptides with Picomolar Affinity. J. Am. Chem. Soc. *124*, 10966-10967.
74. Austin, R.J., Xia, T., Ren, J., Takahashi, T.T., and Roberts, R.W. (2003). Differential modes of recognition in N peptide-boxB complexes. Biochemistry *42*, 14957-14967.
75. Erlanson, D.A., Braisted, A.C., Raphael, D.R., Randal, M., Stroud, R.M., Gordon, E.M., and Wells, J.A. (2000). Site-directed ligand discovery. Proc Natl Acad Sci U S A *97*, 9367-9372.
76. Arkin, M.R., Randal, M., DeLano, W.L., Hyde, J., Luong, T.N., Oslob, J.D., Raphael, D.R., Taylor, L., Wang, J., McDowell, R.S., Wells, J.A., and Braisted, A.C. (2003). Binding of small molecules to an adaptive protein-protein interface. Proc Natl Acad Sci U S A *100*, 1603-1608.
77. Jones, S., Daley, D.T., Luscombe, N.M., Berman, H.M., and Thornton, J.M. (2001). Protein-RNA interactions: a structural analysis. Nucleic Acids Res *29*, 943-954.
78. Bogan, A.A., and Thorn, K.S. (1998). Anatomy of hot spots in protein interfaces. J Mol Biol *280*, 1-9.
79. Lo Conte, L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. J Mol Biol *285*, 2177-2198.
80. Ja, W.W., West, A.P., Delker, S.L., Bjorkman, P.J., Benzer, S., and Roberts, R.W. (2007). In vitro selection of peptides that extend Drosophila lifespan by antagonizing Methuselah. Nature Chemical Biology (*submitted*).
81. Olson, C.A., and Roberts, R.W. (2007). In vitro selection of a discriminate Fab protein against a phosphoserine peptide target.
82. Fellouse, F.A., Wiesmann, C., and Sidhu, S.S. (2004). Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. Proc Natl Acad Sci U S A *101*, 12467-12472.
83. Fellouse, F.A., Barthelemy, P.A., Kelley, R.F., and Sidhu, S.S. (2006). Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. J Mol Biol *357*, 100-114.
84. Yanez, J., Arguello, M., Osuna, J., Soberon, X., and Gaytan, P. (2004). Combinatorial codon-based amino acid substitutions. Nucleic Acids Res *32*, e158.
85. Takahashi, T.T., Austin, R.J., and Roberts, R.W. (2003). mRNA display: ligand discovery, interaction analysis and beyond. Trends Biochem Sci *28*, 159-165.

86. Cox, J.C., Hayhurst, A., Hesselberth, J., Bayer, T.S., Georgiou, G., and Ellington, A.D. (2002). Automated selection of aptamers against protein targets translated in vitro: from gene to aptamer. Nucleic Acids Res *30*, e108.
87. Kolonin, M.G., Bover, L., Sun, J., Zurita, A.J., Do, K.A., Lahdenranta, J., Cardo-Vila, M., Giordano, R.J., Jaalouk, D.E., Ozawa, M.G., Moya, C.A., Souza, G.R., Staquicini, F.I., Kunyiasu, A., Scudiero, D.A., Holbeck, S.L., Sausville, E.A., Arap, W., and Pasqualini, R. (2006). Ligand-directed surface profiling of human cancer cells with combinatorial peptide libraries. Cancer Res *66*, 34-40.
88. Battiste, J.L., Mao, H., Rao, N.S., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D., and Williamson, J.R. (1996). Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. Science *273*, 1547-1551.
89. Harada, K., Martin, S.S., and Frankel, A.D. (1996). Selection of RNA-binding peptides in vivo. Nature *380*, 175-179.
90. Scharpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S., and Rosch, P. (2000). Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. Eur J Biochem *267*, 2397-2408.
91. Stuart, A.C., Gottesman, M.E., and Palmer, A.G., 3rd (2003). The N-terminus is unstructured, but not dynamically disordered, in the complex between HK022 Nun protein and lambda-phage BoxB RNA hairpin. FEBS Lett *553*, 95-98.
92. Kawakami, J., Sugimoto, N., Tokitoh, H., and Tanabe, Y. (2006). A novel stable RNA pentaloop that interacts specifically with a motif peptide of lambda-N protein. Nucleosides Nucleotides Nucleic Acids *25*, 397-416.
93. Furusawa, H., Murakawa, A., Fukusho, S., and Okahata, Y. (2003). In vitro selection of N-peptide-binding RNA on a quartz-crystal microbalance to study a sequence-specific interaction between the peptide and loop RNA. Chembiochem *4*, 217-220.
94. Matsugami, A., Kobayashi, S., Ouhashi, K., Uesugi, S., Yamamoto, R., Taira, K., Nishikawa, S., Kumar, P.K., and Katahira, M. (2003). Structural basis of the highly efficient trapping of the HIV Tat protein by an RNA aptamer. Structure *11*, 533-545.
95. Willard, F.S., and Siderovski, D.P. (2006). The R6A-1 peptide binds to switch II of Galphai1 but is not a GDP-dissociation inhibitor. Biochem Biophys Res Commun *339*, 1107-1112.
96. Johnston, C.A., Ramer, J.K., Blaesius, R., Fredericks, Z., Watts, V.J., and Siderovski, D.P. (2005). A bifunctional Galphai/Galphas modulatory peptide that attenuates adenylyl cyclase activity. FEBS Lett *579*, 5746-5750.
97. Johnston, C.A., and Siderovski, D.P. (2007). Structural basis for nucleotide exchange on G alpha i subunits and receptor coupling specificity. Proc Natl Acad Sci U S A *104*, 2001-2006.
98. Ja, W.W., Adhikari, A., Austin, R.J., Sprang, S.R., and Roberts, R.W. (2005). A peptide core motif for binding to heterotrimeric G protein alpha subunits. J Biol Chem *280*, 32057-32060.
99. Sunahara, R.K., Tesmer, J.J., Gilman, A.G., and Sprang, S.R. (1997). Crystal structure of the adenylyl cyclase activator Gsalpha. Science *278*, 1943-1947.

100. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res *31*, 3497-3500.

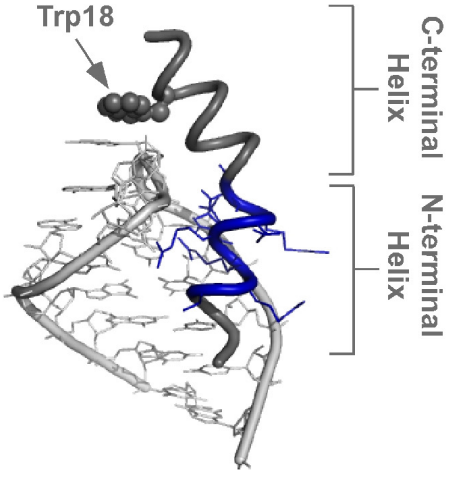
Figures

Figure 1.1 An RNA-binding arginine-rich motif (ARM). (A) ARM peptides are listed along with their amino-acid sequences and pertinent structural references [19, 20, 24-31, 35-38, 40, 88-91]. Natural ARM sequences are listed in black, with experimentally evolved sequences listed in gray. Cognate-RNA and evolved RNA-aptamer (Apt) targets are listed where applicable. Apt (References on the RNA-aptamers) HIV-1-REV [22, 32, 52]; λ N [22, 92, 93]; HIV-1 Tat [94]; HTLV-1 Rex [26]; P22 N, BIV Tat, BMV Gag, CCMV Gag [22]. **(B)** Structural model of the bacteriophage λ N peptide-boxB complex, viewed from the major groove. The N peptide is represented as a ribbon structure with the ARM-consensus motif shown in blue and amino- and carboxy-terminal helices indicated. λ N peptide residue Trp18, shown in ball and stick representation, makes shape-specific π–π stacking contacts with the λboxB loop). The structural image was made from Protein Data Bank file 1QFQ [90], using Pymol software (http://www.pymol.org). **(C)** Specificity-determining elements of λ N peptide. Gln4 and Arg8 residues confer binding specificity to the wild-type λ N peptide (Chapter 2). Shape-specific binding of the λ N peptide Trp18 residue is critical for biological function, but not binding specificity [21]. λ N peptide mutants: λ N(R15) and λ N(E14R15) bind λboxB with increased specificity [20], adopting a shortened α-helical fold that extends four residues to either side of the λ N peptide ARM-consensus motif. Both λ N(R15) and λ N(E14R15) peptides exhibit structural disorder 7 residues from the ARM-consensus motif (Chapter 4).

A

Sequence	peptide	Ref.	RNA-target
TRQARRNRRRRWRWRQR	HIV-1 Rev(34-50)	[88]	HIV-1 RRE, Apt
SRSSRRNRRRRRRR	RSNH-1	[89]	HIV-1 RRE
NHRRRRQRRRRR	RSNH-4	[89]	HIV-1 RRE
RCRRRRGSRRSGASRRRRR	RSG-1	[89]	HIV-1 RRE
RRGSRPSGAERRRRRAAAA	RSG-1.2	[24-25][89]	HIV-1 RRE, Apt
DRRQRRDRQRRRRRAAAA	K1	[19]	HIV-1 RRE
MDAQTRRRERRAEKQAQWKAAN	λ N(1-22)	[35][90]	λboxB, Apt
MDAQTRRRERRAALRNEAKWVV	11.10	[20][40]	λboxB
MDAQTRRRERRAMERATLPQVL	11.36	[20][40]	λboxB
MDAQTRRRERRANMRMYRSLVI	12.50	[20][40]	λboxB
GNAKTRRHERRRKLAIER	P22 N(14-30)	[36]	P22boxB, Apt
TAKTRYKARRAELIAERR	ϕ21 N(12-39)	[37]	ϕ21boxB, Apt
LTSRDRRRIARWEKRIAYALKN	HK022 Nun(22-44)	[38][91]	λboxB
MPKTRRRPRRSQRKRP	HTLV-I Rex (1-16)	[26]	HTLV RxRE, Apt
TRRQRTRRARRNR	HTLV-II Rex(4-16)		
RKKRRQRRR	HIV Tat(49-57)	[27-28]	HIV-1 TAR
RPRGTRGKGRRIRR	BIV Tat(68-81)	[29-30]	BIV TAR, Apt
KKRGTRGKGRKIHY	JDV Tat(68-81)	[31]	JDV TAR
KMTRAQRRAAARRNRWTAR	BMV Gag(7-25)		Apt
KLTRAQRRAAARKNKRNTR	CCMV Gag(7-25)		Apt
TRALRRQLAER	Spuma Gag(168-178)		
TRRNKRNRIQEQLNRK	Yeast PRP6(129-144)		
SQMTRQARRLYV	Human U2AF(142-153)		
RRRRNRTRRNRRRVR	FHV coat(35-49)		
—— $^{T}_{R}$RxxRR ——	ARM-Consensus Motif		

B



C

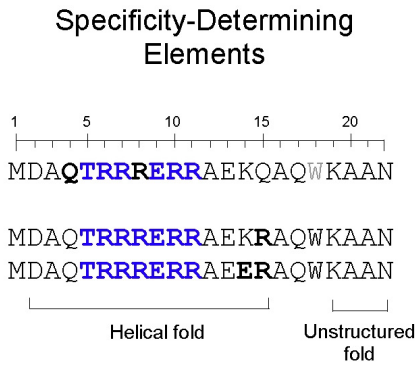
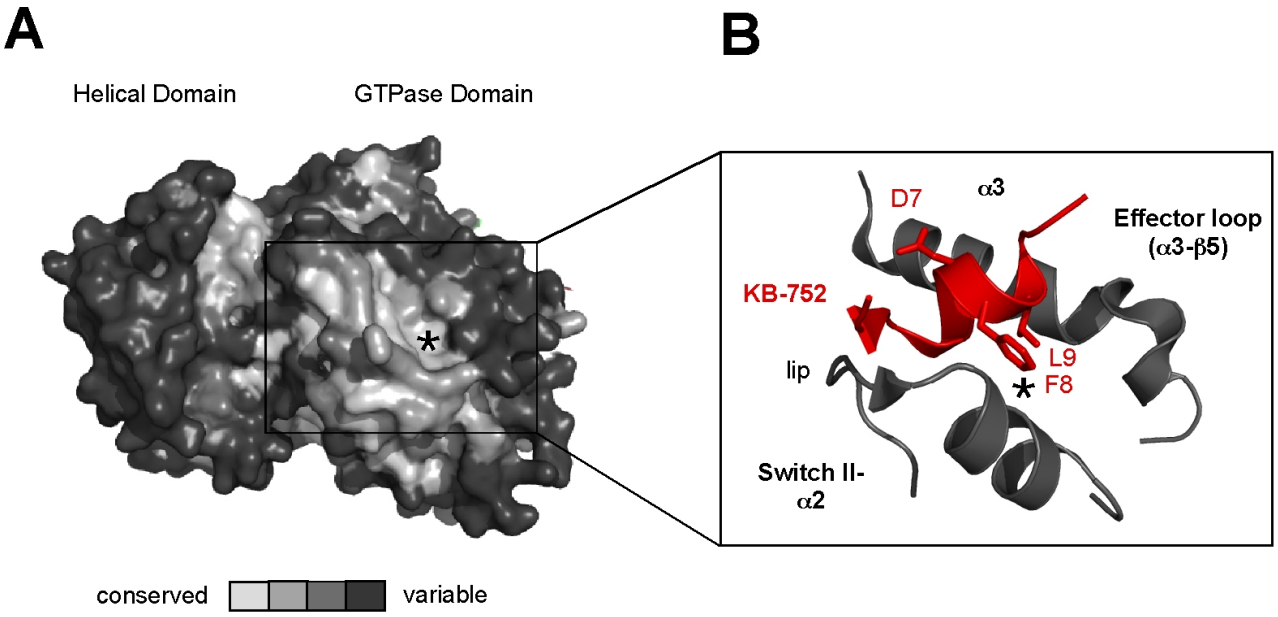


Figure 1.2 Gα-binding peptides. (A) Gα-binding peptides are listed along with their sequences and pertinent reference information. The class-binding specificity of ligands is indicated at right. Gray squares represent positive peptide binding whereas white squares indicate that the peptide does not bind a particular class of Gα. If binding data has not been measured for a Gα class, no square is indicated. An aromatic motif present in a subset of Gα-binding sequences (red) interacts non-specifically with the SII/α3 effector-binding site of Gα. References: [63-68, 70, 95-98], the specificities of AR6-05 and AR6-04 are from unpublished work by William Ja. **(B)** Molecular surface representation of Gα protein sequence homology superimposed on the Gαs(s)-GTPγS crystal structure [99]. A sequence alignment of Gα proteins (i1, i2, i3, oA, q, 11, 15, s(s), Olf, and 12) was performed by ClustalW [100], generating a list of identical (near white), conserved (light gray), similar (gray), and variable (slate) Gα residues, which were grafted onto the Gαs(s)-GTPγS crystal structure. The asterisk denotes an invariant hydrophobic binding pocket within the SII/α3 cleft [62]. **(C)** Ribbon diagram of KB-752 (red) binding within the SII/α3 cleft of Gαi1 (slate). Structural image was made from Protein Data Bank file 1Y3A [64]. Labels for the KB-752 consensus motif residues $D_7F_8L_9$ are shown in red, with structural elements of Gαi1 labeled in slate. Models were generated by Pymol software (http://www.pymol.org).



C

Sequence	Peptde	Ref.	Class Specificity (i/o, q/11, s, 12/13)
MTMGEEDFFDLLAKSQSKRMDDQRVDLAGYK	GPR	[70]	
TKLRMTDNLGWGFLILPSQF	AR6-04	[66]	
SSRGYYHGIWVGEEGRLSR	KB-1753	[65][96]	
MSQTKRLDDQ**LYWWEYL**	R6A	[63]	
DQ**LYWWEYL**	R6A-1	[68][95][98]	
SR**VTWYDLF**MEDTKSR	KB-752	[64][97]	
DESDPEEL**MYWWEFL**SEDPSS	AR6-05	[66]	
M**ITWYEFV**AGTK	Cyc GIBP	[67]	
MAMSDRNKR**LTVWEFL**ALPSST	GSP	[Chapter 5]	
MAMSDQNKR**MTVREFL**ALPSSL	mGSP-1	[Chapter 5]	
MYTSDHNKL**LTVREFL**ALPSST	mGSP-2	[Chapter 5]	

1 5 10

—— $h_Y^T W_Y^W$EFL —— Aromatic-Consensus Motif