Peptide Modulators of

G Protein Signaling

Thesis by

William W. Ja

In partial fulfillment of the requirements

for the degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

2005

(Defended December 13, 2004)

Acknowledgments

I gratefully acknowledge my research advisor, Professor Richard W. Roberts, for his unparalleled encouragement and enthusiasm for our research. I was allowed great personal freedom to pursue opportunities of interest. Unfortunately, this often led to me working for months at a time without providing progress updates, then surprising Rich with odd data at group meetings or, better yet, names of new collaborators. I appreciate him putting up with me—and all of the practical jokes—over the years. I would also like to thank the Roberts lab members, both past and present, for their support and assistance. I am especially indebted to Shuwei Li, Terry Takahashi, and Christine Ueda for all of their help with experiments, critical scientific discussions, and numerous readings of my manuscripts and thesis chapters. I sincerely appreciate the mentorship provided by my committee members, Professors Doug Rees, Bob Grubbs, Ray Deshaies, and Mel Simon. I have often turned to my committee for advice and they have never failed to support me.

My time at Caltech has been one of the most enjoyable periods of my life thus far. I am grateful to the many friends and colleagues that have been so kind to me over the past 6.5 years. I have very fond memories of playing softball with our team, "Good Dog," the undefeated D league champions of 2003, and volleyball with the "Power Pups," spring 2004 champions. My thanks to everyone who played with us and made the games so enjoyable. I also thank Ernie for the countless "extra large" meals. I estimate that I've spent almost $5000 on food from The Taco Truck during my graduate career, a modest contribution to his daughter's Stanford education fund. The staff at Caltech have also been extremely accommodating over the years. My gratitude goes to Margot Hoyt,

Alison Ross, Dian Buchness, Steve Gould, Joe Drew and countless others that continue to play critical roles in running Caltech.

I would have never followed my current career path had it not been for several inspirational teachers in high school: Michael Ivanitsky, Robert Zafran, and I-Heng McComb. I am particularly indebted to Mrs. McComb for her brilliant and effective teaching of molecular biology. She is the reason that I made the switch away from physics and toward biological chemistry. That transition continues today as I prepare to join the Benzer laboratory to work in classical biology and genetics. The members of the Benzer lab have already taught me a great deal and I am especially grateful to Seymour Benzer, Ted Brummel, David Walker, and Viveca Sapin for their tutelage.

My time as an undergraduate at the University of California, Berkeley, was short but significant in shaping my career. Professors Richard Mathies and Alex Glazer were my first research mentors. Their wisdom and guidance continue to strongly affect my approach to science and I will always hope that my academic career reflects well upon them. I am also thankful for the friendship of Professor C. Bradley Moore, a scholar and gentleman, who provided much counsel during my years at Berkeley. Many graduate students and postdocs served as both friends and mentors: Melonie Hall, Judy Kim, Mike Tauber, Yuping Cai, Gary Wedemayer, and Indu Kheterpal. I continue to strive to be a patient mentor for younger scientists, as they were for me. My peers in the Mathies lab—Radbert Chin, Chris Balmaseda, Wendy Wong, Wendy Bao, Phung Tu, and Mary Lai—have become some of my closest and dearest friends. I thank them sincerely for their companionship. I also thank Jenny Chu and Alice Chen for their kindness and for helping me to recall fonder moments.

Finally, I would like to express my gratitude for the support and love from my family and relatives. My sister, Sylvia, and parents, Jason and Mary, have sacrificed much to grant me the freedom to pursue my dreams and interests. I grew up in a loving family and have never wanted for anything. Thanks, mom and dad. I have never worked a day in my life and, because of your encouragement, support, and love, I probably never will.

Abstract

The hundreds of transmembrane proteins that make up the superfamily of G protein-coupled receptors (GPCRs) mediate signaling from an enormous variety of extracellular stimuli—including odorants, pheromones, peptides, lipids, and neurotransmitters—to intracellular heterotrimeric G proteins. The identification of specific modulators of G protein signaling is highly relevant to drug discovery; approximately 50% of currently marketed drugs target a GPCR. Here, we use mRNA display to identify novel and potent peptide ligands for G protein targets. mRNA display is a robust technique that facilitates the isolation of peptides with specific activities (e.g., binding to a target of interest) from large libraries containing trillions of unique molecules. We first targeted the heterotrimeric G protein, $G_{i\alpha1}$, with peptide combinatorial libraries. Isolated peptides bind with high affinity to $G_{i\alpha1}$ and can potentially affect downstream signaling in a pathway-specific manner. A potent peptide core motif interacting with G_{α} subunits was identified and used to construct new mRNA display libraries for the isolation of class- and/or state-specific G_{α}-binding peptides. We have also identified a novel peptide (the RWR motif) that interacts with the *Drosophila* GPCR, Methuselah. These peptides are potent antagonists to Methuselah-mediated signaling and, as mutants of Methuselah are associated with longevity, may be useful in lifespan and aging studies of the fruit fly. Overall, these efforts demonstrate the successful use of mRNA display as an efficient and facile method for generating new solutions to molecular design problems.

Table of Contents

Thesis overview

G protein-coupled receptors (GPCRs) link a diverse array of extracellular signals—including peptides, hormones, odorants, and light—to an equally wide range of intracellular processes, allowing a cell to respond to external stimuli or communicate with other cells (*1-3*). There are approximately 750 human GPCRs, many of which are orphan receptors that respond to unknown ligands (*4*). The diversity of cell processes controlled by GPCRs and the accessibility of their extracellular domains (ligands do not have to cross the plasma membrane) have made them primary drug targets—approximately 50% of currently marketed drugs act on GPCRs (*5*). Intracellular G proteins mediate the signaling from activated GPCRs (*2, 3*). Although drugs directly targeting G proteins are not yet in clinical use, G proteins and their regulators have been increasingly regarded as potential pharmaceutical targets (*6-9*).

The work in this thesis covers our development of peptide ligands targeting G proteins and GPCRs. The long-term goals of our work are to produce new tools for probing G protein and GPCR structure and function, as well as to provide possible leads for future drug discovery and design. Our primary method is mRNA display, which allows us to isolate and identify specific peptide ligands from large libraries comprising over 10 trillion unique members (*10*). Chapter 1 is a brief review of published peptide selection experiments directly targeting G protein signaling pathways. In Chapter 2, we perform an mRNA display selection to isolate a novel peptide (R6A) that binds to the G protein, $G_{i\alpha1}$ (*11*). R6A and its derivatives are short, potent peptides that can modulate the active state of $G_{i\alpha1}$ and have the potential of specifically activating or deactivating particular G protein pathways. In Chapter 3, we extend our studies of the R6A peptide to other G_α family members and identify a core motif for the recognition of G_α subunits.

The core motif can be used as a starting point for new mRNA display libraries to isolate peptides with novel activities and/or specificity for particular G protein classes.

Selection techniques have had limited success against GPCRs because of the difficulty in expression, solubilization, and presentation of the receptor for recognition by peptide libraries. In Chapter 4, we demonstrate the successful isolation of peptides that bind to the GPCR, Methuselah (Mth), by targeting only the extracellular domain. Mth was previously determined to play a role in lifespan in the fruit fly, *Drosophila melanogaster* (*12*). Although the peptide ligands were identified by targeting only the ectodomain, subsequent studies demonstrated that they recognize the full-length receptor and antagonize Mth-mediated signaling.

Appendix A describes the epitope mapping of an anti-polyhistidine monoclonal antibody (mAb). The mRNA display library was originally intended to target $G_{i\alpha 1}$ immobilized on the mAb. However, only high affinity mAb-binding peptides were identified. These sequences revealed a different consensus than the cited epitope and demonstrated significantly higher affinity. To determine the minimal, functional epitope, a detailed procedure for the construction of unidirectional, nested deletion mRNA display libraries is described.

The work in this thesis was supported by grants from the NIH (RO160416) and the Beckman Foundation to R. W. R. W. W. J. was supported in part by a DOD National Defense Science and Engineering Graduate Fellowship and was a recipient of a Scholarship for Research in the Biology of Aging, sponsored by the Glenn Foundation for Medical Research and the American Federation for Aging Research. R. W. R. is an Alfred P. Sloan Foundation Research Fellow.

References

1. Strader, C. D., Fong, T. M., Tota, M. R., Underwood, D., and Dixon, R. A. F. (1994) Structure and function of G protein-coupled receptors, *Annu. Rev. Biochem. 63*, 101-132.
2. Gilman, A. G. (1987) G proteins: transducers of receptor-generated signals, *Annu. Rev. Biochem. 56*, 615-649.
3. Neves, S. R., Ram, P. T., and Iyengar, R. (2002) G protein pathways, *Science 296*, 1636-1639.
4. Vassilatis, D. K., Hohmann, J. G., Zeng, H., Li, F., Ranchalis, J. E., Mortrud, M. T., Brown, A., Rodriguez, S. S., Weller, J. R., Wright, A. C., Bergmann, J. E., and Gaitanaris, G. A. (2003) The G protein-coupled receptor repertoires of human and mouse, *Proc. Natl. Acad. Sci. U.S.A. 100*, 4903-4908.
5. Howard, A. D., McAllister, G., Feighner, S. D., Liu, Q., Nargund, R. P., Van der Ploeg, L. H., and Patchett, A. A. (2001) Orphan G-protein-coupled receptors and natural ligand discovery, *Trends Pharmacol. Sci. 22*, 132-140.
6. Freissmuth, M., Waldhoer, M., Bofill-Cardona, E., and Nanoff, C. (1999) G protein antagonists, *Trends Pharmacol. Sci. 20*, 237-245.
7. Höller, C., Freissmuth, M., and Nanoff, C. (1999) G proteins as drug targets, *Cell. Mol. Life Sci. 55*, 257-270.
8. Neubig, R. R., and Siderovski, D. P. (2002) Regulators of G-protein signalling as new central nervous system drug targets, *Nat. Rev. Drug Discov. 1*, 187-197.

9. Nürnberg, B., Tögel, W., Krause, G., Storm, R., Breitweg-Lehmann, E., and Schunack, W. (1999) Non-peptide G-protein activators as promising tools in cell biology and potential drug leads, *Eur. J. Med. Chem. 34*, 5-30.

10. 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.

11. 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.

12. Lin, Y. J., Seroude, L., and Benzer, S. (1998) Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*, *Science 282*, 943-946.

Chapter 1

G protein-directed ligand discovery with peptide combinatorial libraries

William W. Ja and Richard W. Roberts

Introduction

G protein-coupled receptors (GPCRs)[1] relay diverse extracellular signals to intracellular signal transduction pathways through heterotrimeric G proteins (*1, 2*). While drug discovery efforts have primarily focused on GPCRs, ligands for intracellular G proteins that directly modulate signaling have been increasingly regarded as potential drugs (*3-5*). Short peptides, both naturally occurring and synthetically derived from segments of GPCRs, G proteins, and effectors, have been used extensively to map critical interaction sites and antagonize or activate G proteins (*4-6*). While successful, most of these peptides are weak modulators of signaling, exhibiting their activities at μM to mM concentrations. Combinatorial methods have the potential of substantially increasing the potency of known ligands and identifying novel peptides with new functions from diverse, random libraries (*7, 8*). Here, we review several examples of *in vitro* selection applied to the isolation of peptide modulators of G protein signaling.

G protein signaling cycle

In the classical G protein signaling model, an inactive GPCR is coupled to a GDP-bound, $G_{\alpha\beta\gamma}$ heterotrimer (Figure 1). $G_{\beta\gamma}$ binds tightly to G_{α}-GDP, which enhances coupling of the inactive heterotrimer to specific GPCRs and acts as a guanine nucleotide dissociation inhibitor (GDI) by preventing GDP release (*9*). Activation by an extracellular agonist causes the GPCR to act as a guanine nucleotide exchange factor (GEF), exchanging GDP for GTP in the G_{α} subunit. GTP-binding to G_{α} induces $G_{\beta\gamma}$ release and subsequently both

[1] Abbreviations: GAP, GTPase-activating protein; GDI: guanine nucleotide dissociation inhibitor; GEF, guanine nucleotide exchange factor; GoLoco, $G_{\alpha i/o}$-Loco interaction; GPCR, G protein-coupled receptor; GPR, G protein regulatory; GTPγS, guanosine 5'-*O*-(3-thiotriphosphate); MBP, maltose-binding protein; RGS, regulator of G protein signaling; RT-PCR, reverse transcription-polymerase chain reaction.

G_α-GTP and $G_{\beta\gamma}$ can interact with downstream effectors. The intrinsic GTPase activity of G_α results in the eventual hydrolysis of GTP, leading to reformation of the inactive $G_{\alpha\beta\gamma}$ heterotrimer and re-coupling to the receptor. GTPase-activating proteins (GAPs) accelerate the hydrolysis of G_α-GTP, leading to shorter activation times and/or lower basal activities. This simple model of G protein signaling has grown increasingly complex because of (1) the numerous regulatory proteins that modulate or attenuate signaling by acting as GEFs, GDIs, or GAPs and/or by directly competing with receptor, G protein, or effector interactions (*10, 11*); (2) the immense diversity and crosstalk of signal transduction pathways controlled by heterotrimeric G protein activation (*2, 12*); and (3) the growing number of intracellular receptor partners discovered that activate signals through means other than classical G protein pathways (*13*).

In humans, there are 20 distinct, but highly homologous, G_α subunits that are divided into four classes based on their sequence and function: (1) $G_{i/o}$, (2) G_s, (3) $G_{q/11}$, and (4) $G_{12/13}$ (*2*). Despite their similarity, the G_α families can elicit different functions and have distinct and sometimes overlapping specificities for their binding partners (*2*). There are currently 6 known G_β and 11 G_γ subunits, making a large number of $G_{\beta\gamma}$ heterodimers possible. Each of the G_α and $G_{\beta\gamma}$ pairs can interact with a wide variety of effectors. While classical drugs targeting GPCRs usually antagonize natural agonist responses, direct G protein ligands can potentially modulate individual effector pathways, alter signals specifically from particular G protein classes or subclasses, and/or modify the kinetics of G protein signaling. Hence, there is a large degree of selectivity that can be

conferred by drugs that directly interact with G proteins or interfere with G protein signaling (*5*).

In vitro selection with combinatorial libraries

Naturally occurring peptides, as well as peptides derived from portions of GPCRs, G proteins, and effectors, have been used effectively to study the interactions between these proteins (*6*). These peptides are able to modulate G protein signaling in different ways (e.g., by antagonizing G protein interactions or directly stimulating GDP exchange in G_α subunits). Methods for the directed evolution of peptides can both optimize these ligands for higher affinity and activity or isolate novel sequences with desired properties from random libraries (*7, 8*). A typical selection experiment involves (1) construction of a DNA library, (2) expression to produce a peptide library where members are physically linked to their nucleic acid sequences, (3) affinity selection against an immobilized target to retain functional peptides, and (4) amplification of the recovered nucleic acid sequences to produce an enriched library (Figure 2). Typical selection libraries examine 10^8 to 10^9 unique molecules, whereas totally *in vitro* methods that do not require an *in vivo* transformation step can access even greater pool complexities ($>10^{13}$).

Receptor-G protein interface

While a complete structural characterization of GPCR-G protein coupling and activation has not yet been described, biochemical analyses have established that the receptor-G_α interface involves several regions on G_α, including the N- and C-termini, and the intracellular loops and C-terminus of the GPCR (*10, 14*). Synthetic peptides corresponding to the last 11 amino acids in the C-terminus of a number of G_α subunits

have been shown to block G protein-receptor coupling with low potency (μM IC_{50} values), as well as stabilize active forms of the GPCR, presumably by mimicking the conformational effects of heterotrimeric G proteins (*4-6, 10*). These C-terminal-derived peptides generally demonstrate receptor selectivity similar to the full-length G_α subunit.

To enhance the potency of a rhodopsin-binding peptide derived from the C-terminus of $G_{t\alpha}$ (transducin), a "doped" library was constructed using the "peptides-on-plasmids" approach (*15*). In this selection method, peptides are expressed as LacI fusions, which bind stably to *lacO* DNA sequences on the plasmid encoding the peptide. Peptide-LacI-plasmid complexes were affinity purified on activated rhodopsin and recovered plasmids encoding functional peptides were subsequently amplified (*16*). Selected peptides were significantly more potent than the wild-type sequence and the amino acid conservation highlighted several critical residues (Table I). Subsequent work demonstrated that the $G_{t\alpha}$ peptide analogs are able to modulate high and/or low affinity states of the A_1 adenosine receptor and reduce GPCR signaling responses in a receptor-selective fashion (*17*). These results suggest that selections targeting other GPCRs may be able to produce specific ligands, even though many receptor-G_α contacts are shared.

Interestingly, while amino acid conservation was not observed in the random region of the library (Table I), full-length, 15-residue peptides were significantly more potent than synthetic, C-terminal 11-mers derived from the selected sequences (*16*), suggesting that the structural context of the synthetic peptides is important for the high affinity interaction with rhodopsin. Indeed, recombinant N-terminal MBP-peptide fusions were several orders of magnitude more potent than their synthetic peptide counterparts. These

fusion proteins may display the selected peptides in a context similar to the LacI fusion used in the peptides-on-plasmids approach.

The receptor-G protein interface can also be disrupted using peptides derived from the intracellular loops and C-terminus of the GPCR, which presumably bind to G proteins and prevent coupling (*6, 14*). Thorough analyses of these peptides and optimization using combinatorial libraries have not yet been demonstrated. Previously, peptides derived from different intracellular regions of rhodopsin were shown to inhibit G protein coupling (*18*). These peptides demonstrated synergistic inhibition; the addition of multiple peptides dramatically decreased G protein coupling by binding to multiple contact sites on the G_α subunit. Hence, selection libraries based on protein scaffolds that present several receptor-derived loops, thereby mimicking the intracellular face of a GPCR, may be more effective for isolating more potent ligands.

G protein activators

Random peptide libraries have been an effective tool in the isolation of novel sequences with desired properties. Recently, $G_{i\alpha 1}$ was directly targeted in a phage display selection using a commercially available, 7-mer (X_7) peptide library (*19*). In phage display, peptide sequences are expressed on the surface of filamentous phages and selected against an immobilized target (*20*). Three classes of peptides with short consensus motifs were identified from the selection (Table I). Because the consensus sequences were short, database searches identified many (250 to 1000) proteins containing the motifs, only a few of which were implicated or known to be involved in signal transduction (*19*).

Curiously, the authors did not report any direct binding assays to assess peptide affinity or specificity for G_α subunits. However, two of the peptide classes demonstrated the ability to increase the rate of binding of GTP to G_α i, o, and s (*19*). These peptides bear little similarity to other known G protein activators that have cationic, amphipathic structures such as mastoparan (*21*). Whether the selected peptides act specifically as GEFs has yet to be determined, though in a reconstituted GPCR membrane assay, the active peptides were shown to increase the sensitivity of A_1 adenosine receptor agonist-binding to GTP, most likely due to an increase in the equilibrium level of G_α-GTP present in the reconstituted system (*19*).

G_α-$G_{\beta\gamma}$ interface

Activation of effectors by either G_α-GTP or $G_{\beta\gamma}$ is effectively blocked by formation of the GDP-bound heterotrimer, $G_{\alpha\beta\gamma}$. Hence, individual effectors most likely share overlapping binding sites at the G_α-$G_{\beta\gamma}$ interface. Extensive mapping of key residues for effector binding on $G_{\beta\gamma}$, for example, has shown that various signaling partners for $G_{\beta\gamma}$ rely on different subsets of residues for interaction (*22*). Hence, by targeting different sites on or adjacent to the G_α-$G_{\beta\gamma}$ interface, individual pathways might be affected.

Phage display peptides against $G_{\beta\gamma}$

Recently, phage display was used to identify peptides that bind to $G_{\beta\gamma}$ (*23*). A variety of libraries were used, both linear and constrained with disulfide bridges (Table I). Approximately 250 copies of peptide were displayed per phage, permitting the recovery of peptides with even very low affinity due to avidity effects (though higher affinity

peptides may be more difficult to isolate due to the narrower dynamic range of binding). The authors cleverly modified $G_{\beta\gamma}$ with an amine-specific biotinylation reagent in the presence of G_{α}, thereby "protecting" the G_{α}-$G_{\beta\gamma}$ interface from modification. After G_{α} was removed by affinity chromatography, biotinylated $G_{\beta\gamma}$ was immobilized on streptavidin and used as the selection target.

The selected peptides were grouped into four families, one of which had significant homology to peptides derived from phospholipase C-β (PLC-β) and to a short motif in phosducin that binds to G_{β} subunits (*23*). Peptides from all four families appeared to bind to a single site on $G_{\beta\gamma}$ based on competition experiments, suggesting a "hot spot" for binding interaction (*24, 25*). One synthesized peptide was shown to actively promote G_{α} dissociation from $G_{\beta\gamma}$, presumably through a non-competitive, allosteric effect (*26, 27*). Intriguingly, the peptide inhibited activation of PLC-β by $G_{\beta\gamma}$, but not $G_{\beta\gamma}$-mediated inhibition of voltage-gated calcium channels or adenylyl cyclase (*23*). Subsequent studies were also able to demonstrate an *in vivo* response to the application of the peptides, which presumably resulted from the disruption of heterotrimers and activation of downstream MAP kinase pathways in the absence of receptor activation (*26*).

The more recent description of an N-terminal, single-site biotinylation tag on G_{β} (*26*) suggests that homogeneously oriented, immobilized $G_{\beta\gamma}$ could be used in the future as a selection target. This may provide access to additional protein interaction sites that were blocked by biotinylation, due to protection of only the G_{α}-binding surface. Various sets of effectors might also be useful as competitors during selection experiments to identify rare peptides with highly specific functions.

mRNA display with the GoLoco/GPR motif

mRNA display is a completely *in vitro* method for selection where individual peptides are covalently coupled to the 3'-end of their encoding mRNA, resulting in stable RNA-peptide fusions (*28*). Pools of fusions are selected for binding against an immobilized target and recovered sequences are amplified by RT-PCR. The G protein regulatory (GPR) or GoLoco motif binds selectively to $G_{i/o\alpha}$ subunits and acts as a GDI (*29, 30*). mRNA display libraries, based on the C-terminal half of a GPR consensus sequence (*31*), were selected against $G_{i\alpha1}$ specifically biotinylated at the N- or C-terminus (*32*). A strongly conserved motif was identified and the dominant peptide after selection (named R6A) demonstrated high affinity (60 nM) and GDI activity for $G_{i\alpha1}$ (Table I).

The R6A peptide was subsequently minimized to a 9-residue sequence that retained high affinity and GDI activity and also competed with $G_{\beta\gamma}$ for binding to $G_{i\alpha1}$ (*32*). This 9-mer sequence retained only two residues from the original GPR consensus motif and, based on subsequent analysis, most likely exerts its effects though a different mechanism than the GPR consensus peptide. Recent results have demonstrated that the minimal peptide is able to bind to different G_α subunits representing all four G protein families.[2] Hence, this peptide acts as a core motif for G protein binding and most likely interacts with a conserved region in all G_α subunits. By starting with doped libraries based on this consensus sequence, peptides could be selected against various G_α subunits to produce peptides with class- and/or subclass-specificity.

[2] Ja and Roberts, manuscript in preparation.

G_α specificity using adapter peptides

The high sequence and structural similarity between the various G_α subunits makes it difficult to isolate small ligands that can distinguish between G protein classes. When comparing G_α subunits, it is evident that the helical domain represents the best target for developing class-specific molecules because of the high variability between all four G protein classes (Figure 3, left). This has been shown for RGS9, a GAP which differentiates between $G_{t\alpha}$ and $G_{i\alpha1}$ by recognizing subclass-specific residues in the helical domain (*33*). The crystal structure of a $G_{i\alpha1}$:RGS14-GoLoco complex revealed how a short peptide could selectively bind to a G_α subunit (*34*). The poorly conserved region C-terminal to the GoLoco motif makes numerous contacts with residues in the helical domain of $G_{i\alpha1}$ that differ in $G_{o\alpha}$, thereby imparting increased affinity and subclass specificity (Figure 3, right). RGS14 specificity has recently been extended to $G_{i\alpha1}$ over $G_{i\alpha2}$, which is remarkable due to the high protein sequence identity (88%) between these two isoforms (*35*).

The GoLoco peptide essentially acts as an efficient payload delivery system for directly affecting G protein interactions. While the GoLoco consensus sequence (the "payload") interacts with regions that interfere with nucleotide exchange and $G_{\beta\gamma}$-binding, the C-terminal region acts as an adapter peptide that delivers the required functional groups to a specific G_α target. Indeed, when replacing the RGS14-GoLoco peptide C-terminus with a sequence derived from Pcp2 (a GoLoco protein that acts on $G_{o\alpha}$ rather than $G_{i\alpha}$), the specificity of the RGS14-GoLoco-Pcp2 chimera is switched (*34*). Hence, it may be possible to design class-specific G_α ligands using various adapter

peptides to deliver small molecules or functional peptide motifs that modulate signaling activity.

While the G_{α} helical domain is an attractive target for designing specific ligands, several selective peptides have been characterized that likely interact primarily with the Ras-like domain. A GoLoco/GPR consensus peptide that binds to $G_{i\alpha}$ without the presence of the extended, non-conserved C-terminal region retains a strong preference for $G_{i\alpha}$ over $G_{o\alpha}$ subunits (*31, 36*). Assuming that the consensus peptide binds similarly to the strongly related RGS14-GoLoco peptide used in the $G_{i\alpha1}$ complex crystal structure (*34*), specificity for $G_{i\alpha1}$ over $G_{o\alpha}$ most likely results from different conformations of the G_{α} binding surface rather than the identity of specific residue contacts (Figure 3, right). While $G_{i\alpha1}$ has been extensively characterized by crystallography, structures of other $G_{i/o\alpha}$ isoforms are not yet available. These structures may reveal subtle conformational differences of interaction sites that establish subclass specificity between these strongly related proteins.

Several peptide activators of G_{α} subunits have also been studied. Mastoparan and its analogs demonstrate varying specificities for the $G_{i/o\alpha}$ and $G_{s\alpha}$ families (*21, 37, 38*). Competition binding studies suggest that mastoparan interacts with the C-terminus of $G_{i\alpha}$ (*39*). A 14-residue peptide derived from the IGF-II receptor preferentially activates $G_{i\alpha2}$ over $G_{i\alpha1}$ and $G_{i\alpha3}$, though the binding site is unknown (*40*). From these examples it is evident that class-specific peptide modulators of G protein signaling targeting the Ras-like domain can be developed, though the molecular design and mechanism of achieving this specificity is much less clear.

Future directions

The increasingly complex model for G protein signaling drives the need for new tools for probing G protein structure and function. Selection techniques have already enabled the discovery of novel peptide ligands with unique properties. Targeting of different G protein states (nucleotide-free (*41*), GDP, GDP-AlF_4^- (*42, 43*), or GTPγS) may facilitate the isolation of various G protein modulators that act as GEFs, GDIs, or GAPs. Similar effects may be achieved by targeting natural G protein regulators (e.g., RGS proteins (*44, 45*)). Assaying the effect *in vivo* of potential signal modulators will be crucial toward their use as drugs or drug leads. The utility of direct peptide modulators of G protein signaling is illustrated in the targeted expression of a C-terminal $G_{q\alpha}$ peptide that inhibited G_q-signaling in a murine model of cardiac pressure overload, thereby protecting the mice against subsequent myocardial hypertrophy (*46*). Hence, if peptide ligands can overcome the plasma membrane barrier and avoid proteolysis, they may indeed be useful as drugs *in vivo*.

Of the techniques used for peptide selection against G protein targets, mRNA display will be a significant tool for the rapid isolation of potent ligands. mRNA display has significant advantages over other peptide selection techniques, including access to higher complexity and monovalent display of library members, resulting in the identification of high affinity sequences (*47*). Access to extremely large libraries, comprising $>10^{12}$ molecules, most likely led to the successful isolation of high affinity $G_{i\alpha1}$-binding peptides that contain a critical mutation in the peptide constant region (*32*). The recent incorporation of unnatural amino acids into mRNA display libraries using sense (*48, 49*) and nonsense (*50*) suppression schemes provides further molecular diversity to explore.

Additionally, mRNA display libraries of peptide-drug conjugates (*51*) may be useful in the selection of molecules consisting of nucleotide analogs or other G protein-interacting ligands covalently coupled to peptides optimized for selectivity.

New discoveries of GPCR and G protein activation through non-traditional means continue to add complexity to the classical G protein signaling model (*13*). A number of diverse proteins (e.g., arrestins, GPCR kinases, and small GTP-binding proteins) have been found that associate with activated GPCRs and may represent additional targets for selection. Inhibition of G proteins may attenuate these alternate modes of signaling and demonstrate whether targeting G proteins for pharmaceutical purposes will be viable. Numerous molecules that interact with proteins involved in G protein signaling, including peptides derived from receptors, effectors, and G proteins, as well as natural peptides and, increasingly, designed small molecules, represent a rich source of potential starting points for selection libraries (*3-6, 52*).

Acknowledgments

We gratefully acknowledge Terry T. Takahashi for helpful discussions and comments on the manuscript.

References

1. Gilman, A. G. (1987) G proteins: transducers of receptor-generated signals, *Annu. Rev. Biochem. 56*, 615-649.
2. Neves, S. R., Ram, P. T., and Iyengar, R. (2002) G protein pathways, *Science 296*, 1636-1639.
3. Nürnberg, B., Tögel, W., Krause, G., Storm, R., Breitweg-Lehmann, E., and Schunack, W. (1999) Non-peptide G-protein activators as promising tools in cell biology and potential drug leads, *Eur. J. Med. Chem. 34*, 5-30.
4. Freissmuth, M., Waldhoer, M., Bofill-Cardona, E., and Nanoff, C. (1999) G protein antagonists, *Trends Pharmacol. Sci. 20*, 237-245.
5. Höller, C., Freissmuth, M., and Nanoff, C. (1999) G proteins as drug targets, *Cell. Mol. Life Sci. 55*, 257-270.
6. Taylor, J. M., and Neubig, R. R. (1994) Peptides as probes for G protein signal transduction, *Cell. Signal. 6*, 841-849.
7. Dower, W. J., and Mattheakis, L. C. (2002) *In vitro* selection as a powerful tool for the applied evolution of proteins and peptides, *Curr. Opin. Chem. Biol. 6*, 390-398.
8. Lin, H., and Cornish, V. W. (2002) Screening and selection methods for large-scale analysis of protein function, *Angew. Chem. Int. Edit. 41*, 4402-4425.
9. Higashijima, T., Ferguson, K. M., Sternweis, P. C., Smigel, M. D., and Gilman, A. G. (1987) Effects of Mg^{2+} and the βγ-subunit complex on the interactions of guanine nucleotides with G proteins, *J. Biol. Chem. 262*, 762-766.

10. Cabrera-Vera, T. M., Vanhauwe, J., Thomas, T. O., Medkova, M., Preininger, A., Mazzoni, M. R., and Hamm, H. E. (2003) Insights into G protein structure, function, and regulation, *Endocr. Rev. 24*, 765-781.

11. Lanier, S. M. (2004) AGS proteins, GPR motifs and the signals processed by heterotrimeric G proteins, *Biol. Cell 96*, 369-372.

12. Landry, Y., and Gies, J. P. (2002) Heterotrimeric G proteins control diverse pathways of transmembrane signaling, a base for drug discovery, *Mini Rev. Med. Chem. 2*, 361-372.

13. Hall, R. A., Premont, R. T., and Lefkowitz, R. J. (1999) Heptahelical receptor signaling: beyond the G protein paradigm, *J. Cell Biol. 145*, 927-932.

14. Wong, S. K.-F. (2003) G protein selectivity is regulated by multiple intracellular regions of GPCRs, *Neurosignals 12*, 1-12.

15. Cull, M. G., Miller, J. F., and Schatz, P. J. (1992) Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869.

16. Martin, E. L., Rens-Domiano, S., Schatz, P. J., and Hamm, H. E. (1996) Potent peptide analogues of a G protein receptor-binding region obtained with a combinatorial library, *J. Biol. Chem. 271*, 361-366.

17. Gilchrist, A., Mazzoni, M. R., Dineen, B., Dice, A., Linden, J., Proctor, W. R., Lupica, C. R., Dunwiddie, T. V., and Hamm, H. E. (1998) Antagonists of the receptor-G protein interface block G_i-coupled signal transduction, *J. Biol. Chem. 273*, 14912-14919.

18. König, B., Arendt, A., McDowell, J. H., Kahlert, M., Hargrave, P. A., and Hofmann, K. P. (1989) Three cytoplasmic loops of rhodopsin interact with transducin, *Proc. Natl. Acad. Sci. U.S.A. 86*, 6878-6882.

19. Hessling, J., Lohse, M. J., and Klotz, K.-N. (2003) Peptide G protein agonists from a phage display library, *Biochem. Pharmacol. 65*, 961-967.

20. Scott, J. K., and Smith, G. P. (1990) Searching for peptide ligands with an epitope library, *Science 249*, 386-90.

21. Higashijima, T., Uzu, S., Nakajima, T., and Ross, E. M. (1988) Mastoparan, a peptide toxin from wasp venom, mimics receptors by activating GTP-binding regulatory proteins (G proteins), *J. Biol. Chem. 263*, 6491-6494.

22. Ford, C. E., Skiba, N. P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L. R., Rosal, R., Weng, G., Yang, C. S., Iyengar, R., Miller, R. J., Jan, L. Y., Lefkowitz, R. J., and Hamm, H. E. (1998) Molecular basis for interactions of G protein βγ subunits with effectors, *Science 280*, 1271-1274.

23. Scott, J. K., Huang, S. F., Gangadhar, B. P., Samoriski, G. M., Clapp, P., Gross, R. A., Taussig, R., and Smrcka, A. V. (2001) Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein βγ subunits is used for recognition of a subclass of effectors, *EMBO J. 20*, 767-776.

24. Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface, *Science 267*, 383-386.

25. DeLano, W. L., Ultsch, M. H., de Vos, A. M., and Wells, J. A. (2000) Convergent solutions to binding at a protein-protein interface, *Science 287*, 1279-1283.

26. Goubaeva, F., Ghosh, M., Malik, S., Yang, J., Hinkle, P. M., Griendling, K. K., Neubig, R. R., and Smrcka, A. V. (2003) Stimulation of cellular signaling and G protein subunit dissociation by G protein βγ subunit-binding peptides, *J. Biol. Chem. 278*, 19634-19641.

27. Ghosh, M., Peterson, Y. K., Lanier, S. M., and Smrcka, A. V. (2003) Receptor- and nucleotide exchange-independent mechanisms for promoting G protein subunit dissociation, *J. Biol. Chem. 278*, 34747-34750.

28. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12297-12302.

29. Kimple, R. J., Willard, F. S., and Siderovski, D. P. (2002) The GoLoco motif: heralding a new tango between G protein signaling and cell division, *Mol. Interv. 2*, 88-100.

30. Willard, F. S., Kimple, R. J., and Siderovski, D. P. (2004) Return of the GDI: the GoLoco motif in cell division, *Annu. Rev. Biochem. 73*, 925-51.

31. Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., III, Graber, S. G., and Lanier, S. M. (2000) Stabilization of the GDP-bound conformation of Giα by a peptide derived from the G-protein regulatory motif of AGS3, *J. Biol. Chem. 275*, 33193-33196.

32. 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.

33. Skiba, N. P., Yang, C.-S., Huang, T., Bae, H., and Hamm, H. E. (1999) The α-helical domain of $G\alpha_t$ determines specific interaction with regulator of G protein signaling 9, *J. Biol. Chem. 274*, 8770-8778.

34. 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 Gα subunits, *Nature 416*, 878-881.

35. Mittal, V., and Linder, M. E. (2004) The RGS14 GoLoco domain discriminates among $G\alpha_i$ isoforms, *J. Biol. Chem. 279*, 46772-46778.

36. Peterson, Y. K., Hazard, S., III, Graber, S. G., and Lanier, S. M. (2002) Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins, *J. Biol. Chem. 277*, 6767-6770.

37. Higashijima, T., Burnier, J., and Ross, E. M. (1990) Regulation of G_i and G_o by mastoparan, related amphiphilic peptides, and hydrophobic amines. Mechanism and structural determinants of activity, *J. Biol. Chem. 265*, 14176-14186.

38. Sukumar, M., Ross, E. M., and Higashijima, T. (1997) A G_s-selective analog of the receptor-mimetic peptide mastoparan binds to $G_s\alpha$ in a kinked helical conformation, *Biochemistry 36*, 3632-3639.

39. Weingarten, R., Ransnäs, L., Mueller, H., Sklar, L. A., and Bokoch, G. M. (1990) Mastoparan interacts with the carboxyl terminus of the α subunit of G_i, *J. Biol. Chem. 265*, 11044-11049.

40. Okamoto, T., Katada, T., Murayama, Y., Ui, M., Ogata, E., and Nishimoto, I. (1990) A simple structure encodes G protein-activating function of the IGF-II/mannose 6-phosphate receptor, *Cell 62*, 709-717.

41. Ferguson, K. M., Higashijima, T., Smigel, M. D., and Gilman, A. G. (1986) The influence of bound GDP on the kinetics of guanine nucleotide binding to G proteins, *J. Biol. Chem. 261*, 7393-7399.

42. Tesmer, J. J. G., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997) Structure of RGS4 bound to AlF_4^--activated $G_{i\alpha1}$: Stabilization of the transition state for GTP hydrolysis, *Cell 89*, 251-261.

43. Berman, D. M., Kozasa, T., and Gilman, A. G. (1996) The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis, *J. Biol. Chem. 271*, 27209-27212.

44. Neubig, R. R., and Siderovski, D. P. (2002) Regulators of G-protein signalling as new central nervous system drug targets, *Nat. Rev. Drug. Discov. 1*, 187-197.

45. Zhong, H., and Neubig, R. R. (2001) Regulator of G protein signaling proteins: novel multifunctional drug targets, *J. Pharmacol. Exp. Ther. 297*, 837-845.

46. Akhter, S. A., Luttrell, L. M., Rockman, H. A., Iaccarino, G., Lefkowitz, R. J., and Koch, W. J. (1998) Targeting the receptor-G_q interface to inhibit *in vivo* pressure overload myocardial hypertrophy, *Science 280*, 574-577.

47. 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.

48. Frankel, A., and Roberts, R. W. (2003) *In vitro* selection for sense codon suppression, *RNA 9*, 780-786.

49. Frankel, A., Millward, S. W., and Roberts, R. W. (2003) Encodamers: unnatural peptide oligomers encoded in RNA, *Chem. Biol. 10*, 1043-1050.

50. Li, S., Millward, S., and Roberts, R. (2002) *In vitro* selection of mRNA display libraries containing an unnatural amino acid, *J. Am. Chem. Soc. 124*, 9972-9973.

51. Li, S., and Roberts, R. W. (2003) A novel strategy for *in vitro* selection of peptide-drug conjugates, *Chem. Biol. 10*, 233-239.

52. Chahdi, A., Daeffler, L., Gies, J. P., and Landry, Y. (1998) Drugs interacting with G protein α subunits: selectivity and perspectives, *Fundam. Clin. Pharmacol. 12*, 121-132.

53. 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.

Tables

Table I. Peptide selections against G protein-related targets.

Library	Diversity	Target	Result[a]	Activity	Ref
X_4-IKENLKDCGLF[b]	2×10^9	Light-activated rhodopsin	X_4-hXXXLKDCGLF	IC_{50}[c] 140 nM	(*16*)
X_7	$20^7 = 10^9$	$G_{i\alpha1}$[d]	(i) aPXXaHP (ii) QXPXSXP (iii) LPaXXXH	EC_{50}[e] (i) 16 μM (ii) >1000 μM (iii) 17 μM	(*19*)
X_m[f] XCX_nCX $X_5CX_3CX_4$ $X_4CX_pCX_4$ X_8CX_8 XCX_{15} $X_{15}CX$ $XCCX_3CX_5C_4$GIEGRG	10^8–10^9 (each library)	$G_{\beta1\gamma2}$	(i) KAXXLLG (ii) KaXXaaG (iii) CEKRXGXXXC (iv) CX_5C	IC_{50}[g] (i) ~5 μM	(*23*)
MSQSKRLDDQR-X_6	$20^6 = 6 \times 10^7$ [h]	$G_{i\alpha1}$-GDP	MSQTKRLDDQLYWWEYL[i]	K_D[j] 60 nM	(*32*)

[a] Amino acid types: h = hydrophobic; a = aromatic or aliphatic. Multiple sequences represent consensus classes.

[b] Each residue in the constant region was mutated at a 50% rate.

[c] IC_{50} of competition with $G_{t\alpha}$ for binding to light-activated rhodopsin (Meta II). Activity is for the most potent, full-length, synthetic peptide. MBP fusion proteins were several orders of magnitude more potent (*16*).

[d] Selection buffer was apparently not supplemented with nucleotide. Hence, the G_α nucleotide state is unclear, though it probably consisted of a mix between GDP-bound and nucleotide-free subunits.

[e] EC_{50} of rate enhancement of GTPγS binding to $G_{i\alpha1}$.

[f] Subscripts m = 6, 15, or 30; n = 4, 6, 8, 10, or 12; and p = 4, 5, or 6.

[g] IC_{50} of $G_{\beta\gamma}$-mediated phopholipase C activation. Peptides also disrupt $G_{\alpha\beta\gamma}$ heterotrimer formation at similar concentrations (*26*).

[h] Selected peptides encoded critical mutations in the constant region. The presence of these mutations implies that the initial diversity of the library was actually higher than indicated. The total number of molecules in the initial mRNA display pool was approximately 10^{12}. Hence, at least 10^4 copies of each unique (random region) peptide were present. This over-representation, coupled with a finite error-rate during PCR amplification, is most likely what permitted access to extremely rare sequences derived from mutations in the constant region.

[i] Underlined region represents the minimal active peptide (K_D = 200 nM to $G_{i\alpha1}$).

[j] K_D for binding to $G_{i\alpha1}$-GDP. Peptides also exhibited GDI activity and competed with $G_{\beta\gamma}$ for binding.

Figures

Figure 1. Classical G protein signaling. An intracellular, GDP-bound $G_{\alpha\beta\gamma}$ heterotrimer is coupled to a membrane-spanning GPCR (R). $G_{\beta\gamma}$ acts as a GDI for G_{α}-GDP, inhibiting nucleotide exchange and maintaining the inactive state. Extracellular agonists cause the GPCR to act as a GEF, catalyzing the exchange of GDP for cytosolic GTP in the G_{α} subunit. G_{α}-GTP and $G_{\beta\gamma}$ subsequently dissociate and are free to signal downstream effectors (E1 and E2). Hydrolysis of G_{α}-GTP to the GDP-bound state, a reaction that is catalyzed by GAPs, results in reassociation with $G_{\beta\gamma}$ and re-coupling to the receptor. Potential modulators of G protein signaling can interfere with protein-protein interactions (e.g., receptor coupling of G proteins, $G_{\alpha\beta\gamma}$ heterotrimer formation, or effector-G protein) and/or act as GDIs, GEFs, or GAPs.

Figure 2. General strategy for the selection of functional peptides. Starting from a DNA construct encoding a peptide library (top left), a selection pool is generated using various methods that localize each peptide with its encoding nucleic acid sequence. Examples of selection methods described in this review are peptides-on-plasmids (*15*), phage display (*20*), and mRNA display (*28*). After the library is affinity-selected against an immobilized target, functional peptides are "amplified" from the recovered nucleic acid sequences (e.g., by PCR). These peptides can be identified by DNA sequencing of individual clones and/or used as the library for the next round of selection. Each round of selection generates a new library that is enriched for functional members, eventually resulting in a pool that is dominated by active peptides.

Figure 3. (left) Amino acid conservation between a representative G_α subunit from each family [human G_α i1, q, s (short-form), and 12], overlaid on a surface representation of $G_{i\alpha1}$. Gaps in the protein sequence alignment generally appear in surface loops (not shown). Amino acid differences between the G_α subunits are colored on a scale from gray (helical domain) or white (Ras-like domain), for highly conserved or identical residues, to red, for highly variable residues. GDP and Mg^{2+} are colored cyan and magenta, respectively. (right) Structure of $G_{i\alpha1}$-GDP in complex with the RGS14-GoLoco peptide (*34*). The GoLoco consensus domain (blue) and C-terminal region (yellow) make extensive contacts through the Ras-like and helical domains of $G_{i\alpha1}$. Non-identical amino acids between $G_{i\alpha1}$ and $G_{o\alpha}$ are colored in pink or red for conserved or non-conserved differences, respectively. Specific contact residues in the helical domain that differ between $G_{i\alpha1}$ and $G_{o\alpha}$ have been described previously (*34*). Residues in the Ras-like domain that may be important to the specific binding of a GoLoco/GPR consensus peptide (*31, 36*) that lacks the C-terminal region are marked. Protein alignments were performed using ClustalW (*53*) from human cDNA sequences obtained from the UMR cDNA Resource Center (http://www.cdna.org). Both structure images were made from Protein Data Bank file 1KJY (*34*) using PyMOL (http://www.pymol.org).

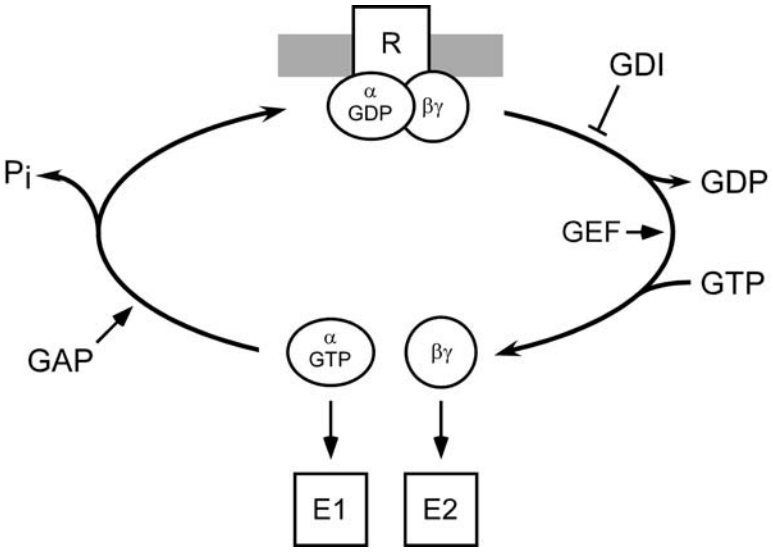


Figure 1

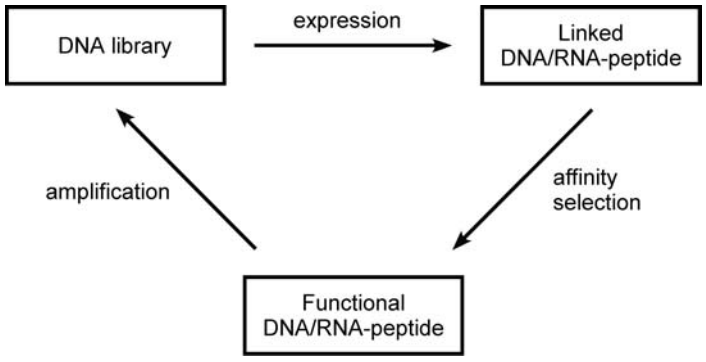

DNA library
expression
Linked
DNA/RNA-peptide
amplification
affinity
selection
Functional
DNA/RNA-peptide

Figure 2

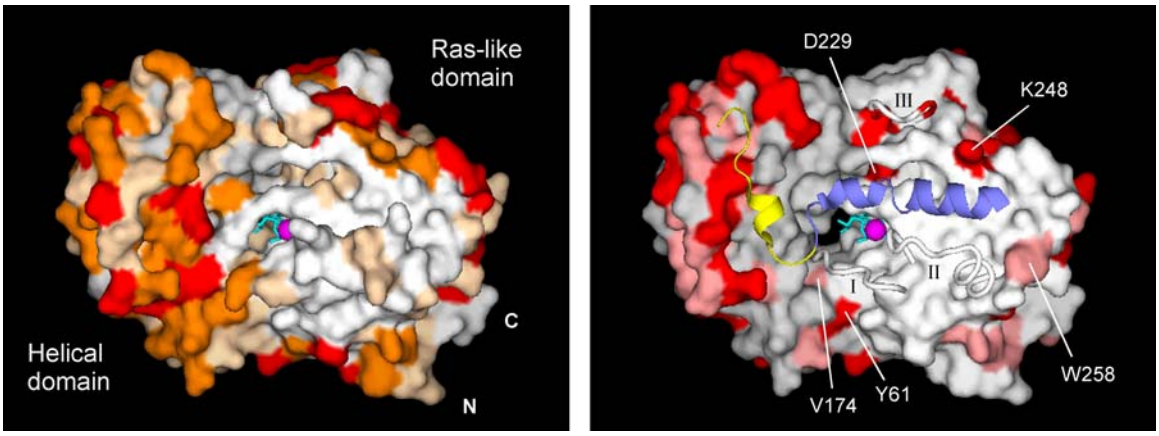

Ras-like
domain
Helical
domain
C
N
D229
III
K248
II
I
V174
Y61
W258

Figure 3

Chapter 2

State-specific peptide guanine nucleotide dissociation inhibitors for $G_{i\alpha 1}$

William W. Ja and Richard W. Roberts

Reproduced in part with permission from: 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.

The G protein regulatory (GPR) motif is a ~20-residue conserved domain that acts as a guanine nucleotide dissociation inhibitor (GDI) for $G_{i/o\alpha}$ subunits. Here, we describe the isolation of peptides derived from a GPR consensus sequence using mRNA display selection libraries. Biotinylated $G_{i\alpha1}$, modified at either the N- or C-terminus, serves as a high affinity binding target for mRNA displayed GPR peptides. *In vitro* selection using mRNA display libraries based on the C-terminus of the GPR motif revealed novel peptide sequences with conserved residues. Surprisingly, selected peptides contain mutations to a highly conserved Arg in the GPR motif, previously shown to be crucial for binding and inhibition activities. The dominant peptide from the selection, R6A, and a minimal 9-mer peptide, R6A-1, do not contain Arg residues yet retain high affinity (K_D = 60 nM and 200 nM, respectively) and specificity for the GDP-bound state of $G_{i\alpha1}$, as measured by surface plasmon resonance (SPR). The selected peptides also maintain GDI activity for $G_{i\alpha1}$, inhibiting both the exchange of GDP in GTPγS binding assays and the AlF_4^--stimulated enhancement of intrinsic tryptophan fluorescence. The kinetics of GDI activity however, are different for the selected peptides and demonstrate biphasic kinetics, suggesting a complex mechanism for inhibition. Like the GPR motif, the R6A and R6A-1 peptides compete with $G_{\beta\gamma}$ subunits for binding to $G_{i\alpha1}$, suggesting their use as activators of $G_{\beta\gamma}$-signaling.

Introduction

Intracellular heterotrimeric guanine nucleotide-binding proteins (G proteins) mediate signaling from cell-surface receptors (GPCRs)[1] to a wide variety of effectors (*1, 2*). In the inactive state, $G_{\beta\gamma}$ heterodimers bind tightly to GDP-bound G_α subunits, enhancing coupling to specific GPCRs and exhibiting guanine nucleotide dissociation inhibitor (GDI) activity by preventing GDP release from G_α (*3*). Activation by extracellular agonists causes the GPCR to act as a guanine nucleotide exchange factor (GEF), exchanging GDP with GTP in G_α and initiating signal transduction through G_α-GTP and/or $G_{\beta\gamma}$ subunits. The inherent guanosine triphosphatase (GTPase) activity of G_α returns the protein to the GDP-bound state, resulting in reassociation of $G_{\beta\gamma}$ and termination of signaling. Numerous other regulators of heterotrimeric G proteins acting as GDIs, GEFs, or GAPs (GTPase-activating proteins which accelerate the GTPase activity of G_α subunits and the termination of signaling) add further complexity to the intricate network of intracellular signaling pathways and the kinetics of G protein signaling (*4*).

Direct modulators of G protein signaling would be useful as molecular tools in studies on the involvement of particular G proteins in specific biochemical pathways, supplementing or replacing traditional genetic techniques. Potent molecules with marked specificity for individual G proteins would potentially act as leads for the development of

[1] Abbreviations: AGS3, activator of G protein signaling 3; Fmoc, Fluorenylmethoxycarbonyl; GAP, GTPase-activating protein; GDI: guanine nucleotide dissociation inhibitor; GEF, guanine nucleotide exchange factor; GoLoco, $G_{\alpha i/o}$-Loco interaction; GPCR, G protein-coupled receptor; GPR, G protein regulatory; GTPγS, guanosine 5'-*O*-(3-thiotriphosphate); HPLC, high performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time-of- flight; MBP, maltose-binding protein; RT-PCR, reverse transcription-polymerase chain reaction; SPR, surface plasmon resonance.

G protein-directed drugs. Drug discovery targeting G proteins has had limited success due to the broad spectrum of signaling events mediated at the G protein level, as well as the high sequence and structural similarities between G protein classes (*5, 6*). The ability to quickly assay combinatorial libraries for molecules with desired properties provides the potential to alleviate these difficulties (*7, 8*).

A selection experiment is an iterative process where a large pool of molecules (e.g., composed of nucleic acids, polypeptides, or synthesized compounds) is sieved for functionality (e.g., binding to a protein target) and active library members are retained. Techniques for peptide or protein selections generally involve the physical association or localization of a polypeptide with its encoding nucleic acid sequence, which allows for the identification of isolated peptides by DNA sequencing. *In vitro* selection has previously been used to recover high affinity peptides that bind to rhodopsin and compete with $G_{t\alpha}$ subunits for receptor-coupling (*9*). More recently, phage display selections produced several classes of peptides that appear to bind to the same site on $G_{\beta\gamma}$ subunits (*10*). Binding of these peptides to $G_{\beta 1\gamma 2}$ was subsequently shown to accelerate dissociation from $G_{i\alpha 1}$, most likely by inducing a conformational change in $G_{\beta\gamma}$ (*11*).

mRNA display is an *in vitro* peptide selection technique that gives access to high complexity libraries ($>10^{13}$ unique peptide sequences) in a robust format (*12, 13*). In mRNA display, an RNA library, produced by *in vitro* transcription from dsDNA template, is covalently linked to its encoded polypeptide via a 3'-puromycin moiety (Figure 1A). These libraries can be composed of random peptides or mutants of specific sequences, based on the DNA template construction. Pools of RNA-peptide fusions are selected against an immobilized target. Recovered, functional protein sequences are

amplified by RT-PCR to produce an enriched dsDNA pool suitable for the next round of selection.

The G protein regulatory (GPR) or GoLoco motif binds selectively to $G_{i/o\alpha}$ subunits and acts as a GDI, stabilizing the GDP-bound state (*14-18*). Single and multiple copies of the ~20-residue conserved GPR motif are found in a variety of signal-regulating proteins (*19*). Proteins encoding the GPR motif, as well as a synthetic, GPR consensus peptide, compete with $G_{\beta\gamma}$ for binding to G_{α} subunits (*14, 15*), thereby activating $G_{\beta\gamma}$-dependent pathways in the absence of nucleotide exchange (*21*). The high affinity and potency of the GPR motif makes it an ideal scaffold for peptide selection. Here, mRNA display with a GPR-derived library was used to select for novel peptides with high affinity for $G_{i\alpha1}$. The dominant, selected peptide (R6A) was minimized to a 9-residue sequence (R6A-1) that shares identity with only 2 amino acids from the core GPR motif yet retains sub-micromolar affinity for $G_{i\alpha1}$. The selected peptides retain GDI activity although the kinetics of inhibition differ significantly from that of the GPR consensus. R6A and R6A-1 also maintain the ability to compete with $G_{\beta\gamma}$ subunits for binding to $G_{i\alpha1}$.

Experimental Procedures

Materials

The *E. coli* strains, BL21 and BL21(DE3), were from Novagen, Inc. (Madison, WI). Restriction enzymes, T4 DNA ligase, and vector pTXB1 were from New England Biolabs, Inc. (Beverly, MA). The G protein expression vector, NpT7-5-H6-TEV-Giα1, was generously provided by Prof. Roger K. Sunahara (University of Michigan). The

cDNA clone of human $G_{i\alpha 3}$ was obtained from the Guthrie cDNA Resource Center (http://www.cdna.org). The *in vivo* biotinylation vector, pDW363, was kindly supplied by Dr. David S. Waugh (National Cancer Institute, Frederick, MD). L-[^{35}S]-methionine (1175 Ci/mmol) was purchased from PerkinElmer Life Sciences, Inc. (Boston, MA). The polyclonal antiserum BN1, which recognizes the N-termini of $G_{\beta 1}$ and $G_{\beta 2}$, was kindly provided by Prof. Melvin I. Simon. Other reagents were purchased from Sigma-Aldrich Corp. (St. Louis, MO) or VWR International, Inc. (West Chester, PA) unless otherwise stated. DEPC-treated ddH_2O was used for all RNA work. DNA oligos (including the modified oligo, pF30P) were synthesized at the Biopolymer Synthesis and Analysis Facility at the California Institute of Technology. DNA sequencing of generated ORFs on all expression vectors and selected peptide clones was performed at the California Institute of Technology DNA Sequencing Core Facility.

G_α subunit cloning and expression

Recombinant rat His_6-TEV-$G_{i\alpha 1}$ (N-terminal His_6 tag followed by a TEV protease cut site) was expressed and purified essentially as described (*22*). Briefly, *E. coli* BL21(DE3) cells harboring NpT7-5-H6-TEV-Giα1 were grown in 1 L of enriched media (2% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, 0.2% (v/v) glycerol, and 50 mM KH_2PO_4, pH 7.2, supplemented with 50 μg/mL ampicillin) to an OD_{600} of 0.5, induced with 0.1 mM IPTG, and collected by centrifugation after ~6 h of expression at 30 °C. Cells were lysed by French press and purified on Ni-NTA Superflow (Qiagen, Inc., Valencia, CA) using a Pharmacia FPLC system (Amersham Biosciences Corp., Piscataway, NJ). Pure protein fractions were combined and concentrated into HED

buffer (50 mM HEPES-KOH, pH 7.5, 1 mM EDTA, and 2 mM DTT) using a Centriprep YM-30 (Millipore Corp., Billerica, MA). The protocol yielded >95% pure protein at ~20 mg/L culture, and the protein was generally used without removal of the epitope tag.

The ORF of $G_{i\alpha1}$ was PCR-amplified from NpT7-5-H6-TEV-Giα1 with the primers 29.2 (5'-CCA TTC TCG AGC ATG GGC TGC ACA CTG AG) and 35.2 (5'-TCT TGG GAT CCT TAG AAG AGA CCA CAG TCT TTT AG) and ligated into vector pDW363 (*23*) using the XhoI and BamHI restriction sites to produce pDW363-Giα1. This vector encodes $G_{i\alpha1}$ with an N-terminal peptide tag that is biotinylated *in vivo* (Nb-$G_{i\alpha1}$). A 25 mL LB/ampicillin culture (supplemented with 50 μM D-biotin) of *E. coli* BL21 cells harboring pDW363-Giα1 was induced with 1 mM IPTG (at OD_{600} = 0.6), grown at 30 °C for 6 hours, and pelleted by centrifugation. Cell pellets were rinsed gently with ddH_2O, snap frozen in dry ice/ethanol, and stored at −80 °C until needed. Cells were thawed, lysed with B-PER (Pierce Biotechnology, Inc., Rockford, IL), and cleared as per the manufacturer's instructions. Cleared lysate was applied to a 2 mL monomeric avidin-agarose column (Pierce), washed with 8 × 2 mL of 1× PBS/0.1% (v/v) Triton X-100, and eluted with 7 × 2 mL of 1× PBS/2 mM D-biotin. The column could be regenerated with 0.1 M glycine, pH 2.8, and reused with negligible loss in binding capacity. Fractions containing Nb-$G_{i\alpha1}$ were combined and concentrated in a Centriprep YM-30 into HGD buffer (50 mM HEPES-KOH, pH 7.5, 10% (v/v) glycerol, and 1 mM DTT) for storage at −80 °C. The 25 mL culture yielded approximately 1 mg of >95% pure Nb-$G_{i\alpha1}$ (~40 mg/L culture).

Nb-$G_{i\alpha3}$ was expressed and purified using the same protocol as for Nb-$G_{i\alpha1}$. The coding region for human $G_{i\alpha3}$ was PCR-amplified from a cDNA clone using primers 30.4 (5'-CCA TTC TCG AGC ATG GGC TGC ACG TTG AGC) and 39.1 (5'-TCT TGG GAT CCT TAA TAA AGT CCA CAT TCC TTT AAG TTG) and ligated into pDW363 using the XhoI and BamHI restriction sites. Approximately 150 μg of biotinylated $G_{i\alpha3}$ was obtained from 50 mL of culture (3 mg/L culture), which was sufficient for our experiments. The lower yield of $G_{i\alpha3}$ compared with that of $G_{i\alpha1}$, despite the high sequence similarity, is consistent with previously published work (*22*).

To produce C-terminally biotinylated protein, $G_{i\alpha1}$ was expressed as an intein fusion (*24*). The $G_{i\alpha1}$-intein fusion protein was purified via a chitin binding domain within the intein which, in the presence of thiols, undergoes specific self-cleavage, releasing $G_{i\alpha1}$ from the chitin-bound intein. By using a biotinylated cysteine derivative, cleavage from the intein and biotinylation of $G_{i\alpha1}$ occur in a single step (*25, 26*). The ORF of rat $G_{i\alpha1}$ was PCR-amplified with primers 33.1 (5'- TTG GTG CCC GCA ACA TAT GGG CTG CAC ACT GAG) and 40.1 (5'- GGT GGT TGC TCT TCC GCA GAA GAG ACC ACA GTC TTT TAG G) and sequentially digested with SapI followed by FauI. Because the coding region of $G_{i\alpha1}$ contains an internal SapI site, aliquots were taken from the initial SapI digest over the course of a 4 min digestion (at 37 °C) and quenched immediately. The aliquots were pooled and desalted (QIAquick PCR purification, Qiagen) followed by a complete FauI digest and agarose gel purification to remove fragments that were cut at the internal SapI site. The FauI/SapI digested DNA was inserted into pTXB1 at the NdeI/SapI restriction sites to create a new ORF encoding a $G_{i\alpha1}$-intein fusion. A 300 mL

culture of *E. coli* BL21(DE3) harboring pTXB1-Giα1 was induced at an OD_{600} of 0.6 with 0.5mM IPTG, grown at 30 °C for 4 h, and collected by centrifugation. Cell pellets were snap frozen in dry ice/ethanol and stored at −80 °C until needed. Cells were resuspended in lysis buffer (20 mM HEPES-KOH, pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.1% Triton X-100) and lysed by French press. After clearing the cell debris by centrifugation (30 min at 12000 × g), 5 mL of chitin beads (New England Biolabs) was added to the supernatant and rotated at 4 °C for 2 h. The beads were collected in a gravity column and washed with 100 mL of column buffer (20 mM HEPES-KOH, pH 7.5, 500 mM NaCl, 0.1% Triton X-100). To cleave $G_{i\alpha1}$ from the intein and biotinylate the C-terminus, the beads were agitated at 4 °C for ~90 h in 5 mL of column buffer containing 1 mM TCEP (Molecular Biosciences, Inc., Boulder, CO) and 0.9 mM N,N'-D-biotinyl-2,2'-(ethylenedioxy)bis(ethylamine)-L-cysteine (Supporting Information). Sodium 2-mercaptoethanesulfonate was supplemented into the mixture at 20 h and 40 h (10 and 30 mM final, respectively). Cb-$G_{i\alpha1}$ was collected with several fractions of column buffer and concentrated using a Centriprep YM-30 into storage buffer (50 mM HEPES-KOH, pH 7.5, 1 mM DTT, 50 μM GDP, 0.1% Triton X-100, 10% glycerol). Approximately 80% of the protein (>90% purity) is biotinylated (determined by binding to streptavidin-agarose), with a yield of ~10 mg/L culture. Higher concentrations of the cysteine derivative result in nearly complete coupling without the need of supplementing 2-mercaptoethanesulfonate, which increases intein cleavage but reduces the percentage of coupled protein (data not shown).

Protein concentrations were determined by UV absorbance at 205 nm (*27*) or 280 nm using a calculated extinction coefficient (http://paris.chem.yale.edu/extinct.html). Values obtained from either method generally agreed within 5%.

mRNA display template preparation

A DNA template encoding the GPR consensus peptide was constructed from oligos GPR-top (5'-GGG ACA ATT ACT ATT TAC AAT TAC AAT GAC CAT GGG CGA GGA GGA CTT CTT TGA TCT GTT GGC CAA G) and GPR-bot (5'-GCC AGC CAG GTC CAC CCG TTG ATC GTC CAT CCG TTT GGA CTG AGA CTT GGC CAA CAG ATC AAA GAA G). These two oligos were PCR amplified together with primers 47T7FP (5'-GGA TTC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT AC) and mycRP (5'-AGC GCA AGA GTT ACG CAG CTG). The X23 library was constructed by step-wise PCR first with oligos GPR-top and 88.2 (5'-AGC GCA AGA GTT ACG CAG CTG GCC AGC CAG GTC AGA DNN TTG ATC GTC CAT CCG TTT GGA CTG AGA CTT GGC CAA CAG ATC AAA GAA G; N = A, C, G, or T; D = A, G, or T) and subsequently with the primers, 47T7FP and mycRP. The C-GPR extension library was generated by PCR amplification of the template C-GPR-X6 (5'-AGC GCA AGA GTT ACG CAG CTG SNN SNN SNN SNN SNN SNN CCG TTG ATC GTC CAG CCG TTT GGA CTG AGA CAT TGT AAT TGT AAA TAG TAA TTG TCC C; S = C or G) with primers 47T7FP and mycRP. The purified (QIAquick PCR purification) dsDNA constructs contained a T7 promoter, an untranslated region, and an ORF containing a 3' constant sequence encoding the peptide QLRNSCA.

In vitro transcription reactions (80 mM HEPES-KOH, pH 7.5, 2 mM spermidine, 40 mM DTT, 25 mM $MgCl_2$, 4 mM each of ATP, CTP, GTP, and UTP, and ~10 μg/mL DNA template) were treated with RNA*secure* (Ambion, Inc., Austin, TX) prior to initiating the reaction with T7 RNA polymerase (*28*). Transcription reactions were incubated at 37 °C for ~4 h, quenched with 0.1 volume 0.5 M EDTA, phenol-extracted using phase lock gel (Brinkmann Instruments, Inc., Westbury, NY), and desalted by isopropanol precipitation. Full-length mRNA was purified by denaturing urea-PAGE, collected from excised gel pieces by passive diffusion in water, and desalted by ethanol precipitation.

The puromycin-DNA linker, pF30P (5'-$dA_{21}[S9]_2$dAdCdC-P; S = spacer phosphoramidite 9; P = CPG-puromycin; 5'-phosphorylated using phosphorylation reagent II; Glen Research Corp., Sterling, VA) was ligated to mRNA templates using a splint oligo (5'-TTT TTT TTT TTN AGC GCA AGA GT). RNA (10 μM final), splint, and pF30P (1/1.1/0.5, respectively) were hybridized by heating at 95 °C for ~3 min, adding T4 DNA ligase buffer (1× final), and cooling on ice for 10 min. SUPERase·In (1 U/μL, Ambion) and T4 DNA ligase (1.6 U/pmol mRNA) were added and the reaction was incubated at room temperature for >2 h. Ligated mRNA-F30P was gel purified and desalted as described above.

RNA and RNA-F30P concentrations were estimated by their absorbance at 260 nm using the equation: c (pmol/μL) = $A_{260}/(10 \times S)$ where S is the length of the template in kilobases.

mRNA display

Purified mRNA-F30P templates were translated in rabbit reticulocyte lysate (Red Nova lysate, Novagen) with ^{35}S-methionine labeling under optimized conditions (100 mM KOAc, 0.5 mM MgOAc, 1 U/μL SUPERase·In, and 0.5 μM mRNA-F30P) and supplemented with unlabeled L-methionine (0.5 mM final). Following the 1 h incubation at 30 °C, additional KOAc and $MgCl_2$ were added to 585 mM and 50 mM (final), respectively. The reactions were then incubated on ice for 15 min to facilitate RNA-peptide fusion formation (*29*). Reactions were used directly or stored at −80 °C until needed. RNA-peptide fusions were purified by dilution into a 100-fold excess of 1× isolation buffer (50 mM HEPES-KOH, pH 7.5, 1 M NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol, 0.05% (v/v) Tween 20) and ~100 μL (dry volume) of pre-washed oligo dT-cellulose (New England Biolabs). After rotating at 4 °C for 1 h, the oligo dT-cellulose was washed thoroughly with 0.4× isolation buffer in a 0.45 μm centrifuge tube filter (Costar Spin-X, Corning, Inc., Corning, NY). RNA-peptide fusions were eluted with pre-warmed (50 °C) dT-elution buffer (10 mM tris-HCl, pH 7.5, 1 mM β-mercaptoethanol). Fusions were isopropanol precipitated with linear acrylamide (Ambion) as a carrier and subsequently reverse transcribed (Superscript II, Invitrogen Corp., Carlsbad, CA) with the oligo, mycRP.

The affinity matrix for selection was prepared by rotating Nb- and/or Cb-$G_{i\alpha1}$ (~10 μg each) with ~20 μL streptavidin-agarose (Immobilized NeutrAvidin on Agarose, Pierce) in buffer A (20 mM HEPES-KOH, pH 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.05% Tween 20) at 4 °C for >1 h. The slurry was supplemented with 1 mM D-biotin

(~0.1 mM final) and rotated for an additional 10 min to block biotin-binding sites. After washing thoroughly with buffer A2 (buffer A supplemented with 2 μM GDP, 1 mM β-mercaptoethanol, 0.2% (w/v) BSA, and 1 μg/mL yeast tRNA (Roche Diagnostics Corp., Indianapolis, IN)), reverse transcribed fusions were rotated with the affinity matrix in 1 mL of buffer A2 at 4 °C for 1 h. The matrix was then washed with 4 × 1 mL buffer A2 followed by 2 × 1 mL buffer A. Bound fusions were eluted with 2 × 0.1 mL 0.15% (w/v) SDS through a 0.45 μm centrifuge tube filter. After removal of the SDS using SDS-OUT (Pierce), cDNA was ethanol precipitated with linear acrylamide (Ambion). PCR amplification of the cDNA with primers 47T7FP and mycRP produced the dsDNA template for the next round of selection. DNA templates could also be directly cloned (TOPO TA cloning for sequencing kit, Invitrogen) for subsequent DNA sequencing.

For the C-GPR X6 extension library selection, RNA-F30P templates encoding R6A were removed by subtractive hybridization as described previously using the anti-R6A oligo, 25.2 (5'-CAA GTA CTC CCA CCA GTA CAG AAA-biotin) prior to the 7th and 8th rounds of selection (*30*).

Binding assays using RNA-peptide fusions on immobilized protein targets were performed similarly, except that translation reactions were prepared without supplementing with unlabeled L-methionine and washes were often performed using spin filters (0.45 μm, Costar Spin-X). Fusions used for binding assays were also often RNase-treated (RNase, DNase-free, Roche) prior to use.

Peptide preparation

Peptides were synthesized with amidated C-termini on a 432A Synergy peptide synthesizer (Applied Biosystems, Foster City, CA) using standard Fmoc chemistry. Following synthesis, peptides were deprotected and cleaved from the resin by agitation in TFA/1,2-ethanediol/thioanisole (90/5/5) for 2 hours at room temperature. Peptides were precipitated with methyl-tert butyl ether and pelleted by centrifugation. Crude peptides were dissolved in ddH_2O (hydrophobic peptides were dissolved in DMSO prior to being diluted in ddH_2O) and purified by reversed-phase HPLC (C18, 250 × 10 mm, Grace Vydac, Hesperia, CA) to >95% purity on an aqueous acetonitrile/0.1% TFA gradient. Peptide masses were confirmed by MALDI-TOF mass spectrometry. Peptide concentrations were determined by absorbance at 280 nm using a calculated extinction coefficient (http://paris.chem.yale.edu/extinct.html).

The L19 GPR and R6A peptides were also expressed as fusions to maltose-binding protein (MBP) using the *in vivo* biotinylation system. GPR or R6A dsDNA was PCR amplified with universal primer 29.4 (5'-TGA AGT CTG GAG TAT TTA CAA TTA CAA TG) and the specific primer 26.1 (5'-AAT CAT ACT AGT ACC GCC GGC CAG GT, for GPR) or 31.1 (5'-AAT CAT ACT AGT ACC GCC CAA GTA CTC CCA C, for R6A). After a BpmI/SpeI digest the dsDNA was co-ligated with synthesized, complementary linker oligos (5'-TCG AGC TCT GGA GGC ATC GAG GGT CGC AT and 5'-GCG ACC CTC GAT GCC TCC AGA GC) into pDW363A (Supporting Information) at the XhoI/SpeI sites to produce pDW363B-GPR and -R6A. These constructs encode the N-terminal biotinylation tag followed by a Factor Xa protease cut site, the inserted peptide, and a C-terminal MBP. L19 GPR was produced by site-

directed mutagenesis (QuikChange, Stratagene) of pDW363B-GPR. Expression and cell lysate preparation of MBP (using pDW363A), L19 GPR-MBP, and R6A-MBP were performed as described above. The cleared lysates were purified on Streptavidin Sepharose (High Performance, Amersham) and washed thoroughly with pDW buffer (50 mM HEPES-KOH, pH 7.5, 200 mM NaCl, 1 mM EDTA, and 0.1% Triton X-100). After washing once with Xa buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, and 1 mM $CaCl_2$), the protein was incubated on-column overnight with Factor Xa (20 U, Amersham) in Xa buffer at room temperature. Proteins were eluted with additional pDW buffer and the Factor Xa was removed with p-aminobenzamidine-agarose (Sigma). Purified proteins were desalted and concentrated in a Centriprep YM-30 into 1× PBS. A 50 mL culture yielded ~16 mg of >98% pure protein (~320 mg/L culture).

Binding analysis by surface plasmon resonance

Kinetic measurements were made at 25 °C on a Biacore 2000 instrument (Biacore, Inc., Piscataway, NJ) equipped with research-grade SA (streptavidin) sensor chips. Nb-$G_{i\alpha1}$ was immobilized to a surface density of ~1000 response units (RU). Modified HBS-EP (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20 (Tween 20), 8 mM $MgCl_2$, 30 μM GDP, and 0.05% (w/v) BSA) was used as the running buffer for all experiments. To collect kinetics data, a concentration series (25, 50, 2 × 100, 200, 400, and 800 nM) for each peptide was injected for 2 min at a flow rate of 100 μL/min. Sample injections were interspersed with a number of buffer blank injections for double referencing with a negative control surface without $G_{i\alpha1}$ used to monitor background binding (*31*). Dissociation was allowed to continue for ~6 min between injections which

allowed the signal to return to baseline, alleviating the need for injecting a regeneration solution. Raw data were processed with Scrubber and globally fit with CLAMP using a 1:1 bimolecular interaction model (*32*). K_D values were calculated (k_d/k_a) from the rates determined by CLAMP. For weaker affinity peptides, higher concentrations were used and the K_D values were determined from equilibrium binding responses using Scrubber. Results from repeated experiments produced similar results, with K_D values within 50% of those shown.

For the analysis of G protein binding states, L19 GPR- and R6A-MBP were immobilized by standard amine-coupling to separate flow cells of an NHS/EDC-activated CM5 sensor chip (Biacore) to a surface density of ~200 RU. Activated flow cells were subsequently blocked with ethanolamine. $G_{i\alpha1}$ (1 μM final) was incubated in HBS-EP+M (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20, 8 mM $MgCl_2$) supplemented with 25 μM GDP, 25 μM GDP with 25 μM $AlCl_3$ and 10 mM NaF, or 25 μM GTPγS for ~1 h at 30 °C. G protein solutions were then injected for 3 min at 35 μL/min across all flow cells and allowed to dissociate for 3 min between injections. BIAevaluation software version 3.2 (Biacore) was used to background subtract all traces with data from a negative control flow cell containing immobilized MBP.

Aluminum fluoride activation

Fluorescence measurements were made on a spectrofluorophotometer (RF-5301PC, Shimadzu Scientific Instruments, Columbia, MD) with excitation and emission wavelengths set at 292 nm and 333 nm, respectively (slit widths at 3 and 5 nm,

respectively). $G_{i\alpha1}$ (200 nM) was preincubated with and without 400 nM peptide in 2.5 mL of buffer A3 (buffer A supplemented with 100 μg/mL BSA, 1mM β-mercaptoethanol, and 5 μM GDP) at 25 °C for 15 min prior to starting the experiment. The temperature throughout the experiment was maintained at 25 °C using a circulating bath (RTE-101, Thermo NESLAB, Portsmouth, NH). Fluorescence was measured for 850 s with a data collection rate of 3 Hz. G proteins were activated by quickly adding 0.5 M NaF (2 mM final) and 10 mM $AlCl_3$ (30 μM final) at 150 and 200 s, respectively. Samples without $G_{i\alpha1}$ were used for baseline subtraction. Traces were smoothed by 5 point adjacent averaging using Origin 6.0 Professional (OriginLab Corp., Northampton, MA).

GTPγS binding

Solutions of $G_{i\alpha1}$ with varying concentrations of peptide were incubated in buffer B (20 mM HEPES-KOH, pH 7.5, 200 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$, 1 mM DTT, 0.005% Tween 20, and 100 μg/mL BSA) for ~20 min at room temperature. All measurements were made in black bottom 96 well plates (Nalge Nunc International, Rochester, NY). Reactions were initiated by diluting the $G_{i\alpha1}$ (100 nM final) samples into BODIPY FL GTPγS (0.8 μM final, Molecular Probes, Eugene, OR) in buffer B using a multichannel pipette, mixing by pipette, and scanning immediately in kinetics mode on a fluorescence plate reader (Flexstation, Molecular Devices, Sunnyvale, CA) for 6 h (45 s between reads, 15 reads/well) at ambient temperature (~25 °C). Excitation and emission wavelengths were set at 485 and 530 nm, respectively, and a 515 nm cutoff filter was used. PMT detection was set at high sensitivity. Data analysis and background

subtraction of reactions without protein were performed with Softmax Pro 4.3.1 (Molecular Devices). Fluorescence curves were fit to single ($A(1 - e^{-kt})$) or double ($A(1 - e^{-kt}) + A_2(1 - e^{-k_2t})$) exponential equations using Origin 6.0.

Immunoprecipitation

The interaction between $G_{i\alpha1}$ and $G_{\beta1\gamma2}$ subunits in the presence and absence of GPR-derived peptides was assayed using purified G protein subunits. Nb-$G_{i\alpha1}$ (40 ng) in 0.5 mL of IP buffer (25 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 5 mM EDTA, 10 mM $MgCl_2$, 1 mM DTT, 0.05% Tween 20, and 0.1% (w/v) BSA and 30 μM GDP or GTPγS) was supplemented with varying concentrations of the indicated peptide (0, 25, 250, and 2500 nM) and incubated at room temperature for 30 min. After addition of $G_{\beta1\gamma2}$ (50 ng, Calbiochem-Novabiochem Corp., La Jolla, CA) and incubation at 4 °C for several hours, NeutrAvidin-agarose (10 μL) was added and the samples were rotated overnight. The agarose was washed with 3 × 0.5 mL IP buffer in a 0.45 μm spin filter and resuspended in 2× SDS-loading buffer. Resuspended samples were incubated at 90 °C for 5 min prior to SDS-PAGE analysis. Proteins were electrotransferred to PVDF membranes (Amersham) and analyzed by Western blot using anti-G_β BN1 (1:5000) and anti-rabbit-peroxidase (1:8000, Roche) as the primary and secondary antibodies, respectively, and an ECL Plus kit for detection (Amersham).

Results

$G_{i\alpha 1}$ as a target for peptide selection

Specifically biotinylated $G_{i\alpha 1}$ subunits were expressed and purified to provide homogeneously presented targets for the peptide selection experiments. The recombinant proteins Nb-$G_{i\alpha 1}$ and Cb-$G_{i\alpha 1}$ contain a single N- or C-terminal biotin tag, respectively, and were produced by different techniques, as described in the Experimental Procedures (Figure 1B). Both Nb- and Cb-$G_{i\alpha 1}$ were protected from trypsin digest after loading with GTPγS (data not shown), demonstrating that the proteins were active for nucleotide exchange (*33, 34*). The biotinylated $G_{i\alpha 1}$ subunits were also tested for their ability to pull down the GPR consensus peptide, a sequence derived from the 4 GPR motif repeats of AGS3 (Figure 1C) (*20*). Radioactively labeled GPR RNA-peptide fusions were purified and assayed for binding against $G_{i\alpha 1}$ immobilized on streptavidin-agarose. Binding of the fusions was specific for Nb- and Cb-$G_{i\alpha 1}$ (80% and 30% binding, respectively) over the streptavidin-agarose matrix (0% binding). The binding of the GPR motif as an RNA-peptide fusion demonstrated the feasibility of performing further *in vitro* selection experiments using mRNA display of GPR-derived peptides. Because subsequent GPR-derived libraries would encode M19L and V24S mutations to facilitate library construction (L19 and S24 are "allowed" residues which are included in a number of the GPR motif repeats within AGS3), these mutants were also assayed for binding.[2] RNA-peptide fusions of M19L or V24S GPR demonstrated negligible differences in binding to

[2] Numbering of residues is based on the GPR consensus peptide (*20*), starting with Thr1 and ending with Gly28 (Figure 1C).

immobilized $G_{i\alpha1}$ compared to the fusions of the "wild-type" GPR consensus sequence (data not shown).

X23 control library

A control selection experiment using the GPR X23 library (Figure 1C) was performed against Nb- and Cb-$G_{i\alpha1}$ to evaluate the proteins as selection targets. R23 is a key amino acid in the GPR motif, as mutations to R23 greatly reduce or eliminate binding to $G_{i\alpha1}$ (*20, 35, 36*). Reverse-transcribed RNA-peptide fusions of the X23 library were allowed to bind to immobilized $G_{i\alpha1}$, non-binding fusions were removed with buffer washes, and viable peptide sequences were determined by PCR amplification of recovered cDNA and DNA sequencing of individual clones (Figure 1A). After 1 round of selection, 70% (4 of 6 sequences) and 80% (5 of 6) Arg at position 23 were recovered against the Nb- and Cb-$G_{i\alpha1}$ matrices, respectively, compared with 0% (0 of 6) for the original X23 pool.

In vitro *selection with C-GPR extension library*

Because R23, which marks the C-terminal residue of the conserved GPR motif, was determined to be crucial for G protein interaction, a C-terminal "extension" library was synthesized to establish whether amino acids just outside of the conserved region affect binding. The C-GPR X6 library (Figure 1C) also included an N-terminal truncation to reduce the binding affinity of the initial pool, allowing for higher enrichment of functional peptides. The initial pool of RNA-peptide fusions contained at least 10^{12} sequences, well encompassing the possible number of unique sequences in a random 6-mer library ($20^6 = 6.4 \times 10^7$ unique sequences). Six rounds of selection were performed

on a mixture of immobilized Nb- and Cb-$G_{i\alpha1}$ to reduce the effects of bias or steric hindrance with either terminus immobilized (Figure 2A). Detergent, bovine serum albumin (BSA), and salt were included in selection buffers to minimize recovery of non-specific binding peptides. DNA sequencing of the 6th round pool revealed a dominant peptide sequence, R6A (Figure 2B).

To recover other rare sequences that were active for binding, mRNA encoding R6A was removed by subtractive hybridization. After an additional two rounds of selection, each preceded by a subtractive hybridization step (Figure 2A), a variety of sequences with high similarity to R6A were discovered, revealing the conserved residues of the selected peptides (Figure 2B). Surprisingly, mutations were discovered in the constant region of R6A for all selected peptides, including the crucial R23. Despite the subtractive hybridization steps, sequences of R6A were still recovered after the 8th round, demonstrating the high selectivity for this peptide sequence.

A separate binding assay with RNA-peptide fusions from the 6th round of selection demonstrated the same preference for Nb-$G_{i\alpha1}$ (40% pull-down) over Cb-$G_{i\alpha1}$ (4%) as with the GPR consensus fusions, further indicating that GPR and GPR-derived peptides favor $G_{i\alpha1}$ immobilized via the N-terminus.

GPR-derived peptides favor the GDP-bound state of $G_{i\alpha1}$

To assay the nucleotide-dependence of the GPR-derived peptides for $G_{i\alpha1}$, binding interactions were observed in real-time using surface plasmon resonance (SPR). N-terminal L19 GPR or R6A peptide fusions with maltose-binding protein (MBP) were immobilized by random amine-coupling to biosensor surfaces. $G_{i\alpha1}$ subunits,

preincubated with either GDP (to maintain the inactive, GDP-bound state), GDP with AlF_4^- (to mimic the transition state of GTP hydrolysis), or GTPγS (a non-hydrolyzable GTP analog to mimic the active, GTP-bound state), were injected across these surfaces (Figure 3A). Both the L19 GPR- and R6A-MBP proteins favored the GDP-bound state of $G_{i\alpha1}$ although L19 GPR demonstrated detectable binding for the other states as well. No binding was detected in a control cell containing immobilized MBP.

Several GPR-derived peptides were also synthesized and purified for kinetic analysis by SPR. Nb-$G_{i\alpha1}$ was immobilized on streptavidin coated sensor chips and the binding of various concentrations of injected peptide was monitored (Figure 3B). The GDP-bound state of $G_{i\alpha1}$ was maintained by supplementing the running buffer with GDP. The running buffer also contained BSA, which was crucial for minimizing non-specific binding and obtaining high quality data. Kinetic parameters were derived from globally fitting the data with a 1:1 interaction model, resulting in dissociation constants (K_D) of 82 nM for L19 GPR and 60 nM for R6A (Table I) (*31*).

To determine a minimal binding peptide sequence, N-terminal truncations of R6A were also assayed by SPR. The shortest peptide tested, R6A-1, bound to $G_{i\alpha1}$ with a K_D of ~200 nM. Shorter peptides were not synthesized due to the hydrophobicity of the C-terminus of R6A. While the control C-GPR peptide did not bind to Nb-$G_{i\alpha1}$ at concentrations up to 20 μM, the mutant peptides R6A-R and L19 GPR R23L both demonstrated >100-fold weaker affinities (determined by fitting steady-state binding measurements) compared to their parent sequences (Table I). The full-length R6A library construct (with the C-terminal QLRNSCA tag) exhibited a similar affinity for $G_{i\alpha1}$

as R6A, indicating that the constant region did not bias the selection (data not shown). Using Cb-$G_{i\alpha1}$ as the immobilized ligand resulted in significantly lower affinities, confirming the preference for $G_{i\alpha1}$ immobilized via the N-terminus (data not shown). The K_D values determined for L19 GPR and R6A were verified by fluorescence titration experiments using C-terminal fluorescein-conjugated peptides (data not shown).

GPR and R6A act as guanine nucleotide dissociation inhibitors

GPR-derived peptides stabilize the GDP-bound state of $G_{i\alpha1}$ and inhibit the activation of $G_{i\alpha1}$ with aluminum fluoride (*37*). Binding of AlF_4^- causes an increase in intrinsic tryptophan fluorescence which can be measured in real-time by spectrofluorometry. While preincubation of $G_{i\alpha1}$ with the C-GPR control peptide had little effect, both the L19 GPR and R6A peptides significantly reduced aluminum fluoride activation, suggesting that R6A retains GDI activity (Figure 4A).

GDI activity of the peptides was also assayed by directly observing nucleotide exchange in $G_{i\alpha1}$. BODIPY FL GTPγS is a fluorescent, non-hydrolyzable analog of GTP that self-quenches in solution. Upon binding to a G_α subunit however, this analog exhibits an increase in fluorescence allowing real-time and high-throughput monitoring of GTP loading (*38*). The L19 GPR and R6A-1 peptides (1 μM), each preincubated with $G_{i\alpha1}$, reduced the initial rate of BODIPY FL GTPγS binding to ~20 and ~70%, respectively, of the initial rate for $G_{i\alpha1}$ without peptide. After 180 min however, both peptides demonstrated similar equilibrium inhibition activities, reducing the fluorescence to ~40% of the fluorescence of BODIPY FL GTPγS-bound $G_{i\alpha1}$ without peptide inhibitor

(Figure 4B). This disparity is caused by the biphasic kinetics of GTP-binding for $G_{i\alpha1}$ incubated with R6A-derived peptides.

The L19 GPR and $G_{i\alpha1}$ without peptide fluorescence curves fit well to single exponentials, and the GDI activity with L19 GPR was fairly well modeled by the simple scheme: G_α-GDP-GPR ↔ G_α-GDP ↔ G_α ↔ G_α-GTP (data not shown). The curves generated with higher concentrations (>50 nM) of R6A-derived peptides however, require a more complex inhibition model and were better described by double exponential equations which reveal a fast, "burst" phase and a ~10-fold slower second phase (Figure 4B). Both phases contribute significantly to the fluorescence amplitude (the slow phase represents 20 to 70% of the total amplitude depending on the inhibitor concentration). Appropriate blanks (with BODIPY FL GTPγS and peptide inhibitor but without $G_{i\alpha1}$) and controls with the R6A-R mutant peptide suggested that the effect was specific and not the result of background fluorescence or non-specific binding. The rate constants of the slow phase did not appear to correlate with peptide concentration, suggesting a parallel reaction pathway. Inhibition with R6A was similar to that of the minimal peptide, R6A-1 (see Supporting Information to view concentration series for all peptides).

IC_{50} values could be determined from the overall fluorescence at 180 min of BODIPY FL GTPγS-bound $G_{i\alpha1}$ with and without various concentrations of peptide inhibitor. L19 GPR and R6A-1 demonstrated comparable submicromolar IC_{50} values (~0.5 μM, Figure 4C) while the mutant peptides, L19 GPR R23L and R6A-R, demonstrated IC_{50} values consistent with their lower binding affinities (IC_{50} >10 μM, data not shown). IC_{50} values

determined by the peptide concentration dependence of the initial rate of BODIPY FL GTPγS binding were severely skewed for R6A-derived peptides due to the initial fast phase of binding (data not shown). Incubation with the C-GPR control peptide at concentrations up to 10 μM had no effect on either the initial rate or the steady-state fluorescence.

$G_{\beta\gamma}$ competition

Although GPR-derived peptides stabilize the inactive, GDP-bound state of $G_{i\alpha}$ subunits, previous studies demonstrated that the GPR motif competes with $G_{\beta\gamma}$ for binding to $G_{i\alpha}$-GDP, promoting subunit dissociation and $G_{\beta\gamma}$-specific signaling in the absence of nucleotide exchange (*14*). To examine this for the selected peptides, reconstituted $G_{i\alpha1\beta1\gamma2}$ was used in co-precipitation experiments. Control experiments first established that $G_{\beta1}$ subunits co-precipitated with $G_{i\alpha1}$ in the GDP state but not in the GTPγS-bound state (Figure 5A). To assay $G_{\beta\gamma}$ competition, increasing concentrations of peptide were incubated with the G protein prior to precipitation. Both the L19 GPR and R6A-1 peptides competed with $G_{\beta\gamma}$ for binding to $G_{i\alpha1}$ (Figure 5B). Results for the full-length R6A peptide were similar (data not shown).

Discussion

The GPR consensus peptide is the shortest, most potent peptide GDI known for the G_i family of G proteins (*20*). To demonstrate the feasibility of using *in vitro* selection to develop peptides with varying activities and specificities for various G protein alpha subunits, the GPR motif was used as a starting point for mRNA display selection

experiments, which requires immobilization of a target protein (*12*). Because G_α subunits putatively contain many regulatory/effector sites, random immobilization schemes (e.g., random amine-coupling or biotinylation of surface cysteine residues) that might restrict binding to favorable, "hot spots" for protein interaction (*39*) were avoided. Instead, specific biotinylation of the N- or C-terminus of $G_{i\alpha1}$ was accomplished using two different methods: *in vivo* biotinylation with *E. coli* biotin holoenzyme synthetase (*23*) and chemical ligation (*25, 26*). Both of these methods provided ample protein yields for the selection and subsequent assays. The *E. coli in vivo* biotinylation expression system was especially favorable as protein minipreps (5 mL) yielded sufficient material for hundreds of kinetics measurements by SPR.

In vitro selection with an extension library, where the conserved region of the GPR motif was extended by 6 random residues on the C-terminus, revealed a dominant peptide, R6A, as well as other highly similar sequences. Only the C-terminal half of the GPR motif was used in the library to allow for higher enrichment of viable peptides and to serve as an "anchor" for the selection, producing peptides that bound near the nucleotide-binding pocket. Surprisingly, selected peptides all contained mutations in the designed, conserved region, including the crucial R23. R6A and the L19 GPR peptide demonstrated comparable binding affinities for $G_{i\alpha1}$ based on SPR and fluorescence titration experiments although the association and dissociation rates were several fold faster for R6A.

N-terminal truncations of R6A bound nearly as well as the full-length peptide. The shortest peptide tested, R6A-1, is a 9-residue sequence that also retains both high affinity binding and GDI activity for $G_{i\alpha1}$. As R6A-1 preserves only 2 of the original residues

from the C-terminus of the GPR motif, this raises the possibility that the R23L mutation eliminated any "anchoring" effect that the constant region had for the nucleotide-binding pocket of $G_{i\alpha1}$ and allowed the library to localize to other regions. Several assays suggested that this was not the case. Both R6A and L19 GPR peptides favored binding to $G_{i\alpha1}$ immobilized by the N-terminus rather than by the C-terminus. This may result from steric hindrance as Cb-$G_{i\alpha1}$ was produced without the long peptide linker region that Nb-$G_{i\alpha1}$ includes. The peptides also competed with each other for binding to $G_{i\alpha1}$ based on SPR as well as radioactively-labeled pull-down experiments (data not shown). These results suggest that R6A and the GPR motif bind to the same or overlapping sites on $G_{i\alpha1}$, though this is not conclusive as binding to other regions (e.g., the flexible switch regions) could cause allosteric competition. The GDI activity of $G_{\beta1\gamma2}$ for example, stems from a rearrangement of switch regions I and II on $G_{i\alpha1}$, inducing new contacts with and tighter binding of GDP (*40*).

More surprising were observations that the minimal peptide, R6A-1, as well as its parent sequence, retained the ability to compete with $G_{\beta\gamma}$ subunits for binding to $G_{i\alpha1}$. The $G_{i\alpha1}$-GPR (GoLoco) crystal structure revealed direct contacts between the C-terminus of the GPR motif with the GDP-binding pocket and the N-terminus with switch II of $G_{i\alpha1}$, which is perturbed such that $G_{\beta\gamma}$ can no longer bind (*35*). R6A-1 is not long enough however, to fully span the same regions, implying that binding, GDI, and/or $G_{\beta\gamma}$ competition activities are produced by long-range effects. It is difficult to predict how the 9-residue R6A-1 could affect the switch II region as extensively as the GPR

consensus peptide, though perturbation of switch I from the nucleotide-binding site could lead to a restructuring of the switch regions and subsequent loss of $G_{\beta\gamma}$ binding.

Although the selected peptides are similar to the GPR consensus sequence in binding affinity and GDI activity for $G_{i\alpha1}$, aberrant inhibition kinetics were observed in the nucleotide exchange experiments using the BODIPY-labeled GTP analog. The inhibition by GPR was easily described by a direct competition model, however we were unable to determine a kinetics model describing the biphasic GTP binding curves from R6A-inhibited experiments. The double exponential fits suggest an alternate reaction pathway with a different reaction rate. Proposed models were unable to correlate the fast, initial phase of GTP-binding with the binding kinetics of the R6A peptide for $G_{i\alpha1}$-GDP determined by the SPR experiments. These peptide signal modulators may be useful in systems where it is desirable to attenuate the overall G protein activation, without significantly perturbing the initial kinetics.

Several studies have demonstrated the importance of neighboring residues outside of the conserved region of the GPR motif. Replacement of the non-conserved residues C-terminal to R23 of the GPR consensus sequence with a short peptide linker greatly reduces binding affinity for $G_{i\alpha1}$ (data not shown) demonstrating that flanking residues can strongly modulate the binding affinity. With the GPR (GoLoco) motif of RGS14, non-conserved C-terminal residues convey specificity for $G_{i\alpha}$ over $G_{o\alpha}$ subunits, winding through the helical domain and contacting $G_{i\alpha}$-specific residues (*35*). More recently, a comprehensive study of the 4 GPR motif repeats of activator of G protein signaling 3 (AGS3) confirmed that residues outside of the conserved GPR motifs strongly potentiate

binding and GDI activity for $G_{i\alpha1}$ (*41*). Studies with R6A and other peptides isolated from the selection may reveal additional specificities and activities for other G_α subunits.

The arginine finger has been a common theme in guanine nucleotide-binding proteins and GTPase activity (*35, 42-45*). In $G_{i\alpha1}$ for example, R178 within switch I stabilizes the γ-phosphate leaving group and is crucial for GTPase activity (*43*). The $G_{i\alpha1}$-GPR structure revealed extensive contacts with the nucleotide-binding pocket of $G_{i\alpha1}$ and the conserved tripeptide, Asp-Gln-Arg (Arg equivalent to R23 on the GPR consensus peptide), from the GPR motif. The Asp and Gln residues are positioned away from the GDP-binding site allowing the Arg residue to insert into the pocket and form hydrogen bonds with the α- and β-phosphates and their bridging oxygen (*35*). Mutation of Arg on the GPR motif has been shown to substantially diminish or eliminate GDI activity and binding affinity for $G_{i\alpha}$ (*20, 35, 36*). By our SPR experiments, the R23L mutation on the GPR consensus peptide resulted in a ~170-fold lower binding affinity ($\Delta\Delta G° = 3.0$ kcal/mol). It is unclear how the selected peptides bind and stabilize the GDP-bound state of $G_{i\alpha1}$ without an Arg residue, and whether the remaining conserved residues form the same contacts as in the GPR motif. However, the Arg to Leu mutation isolated by selection is crucial for binding and activity, as demonstrated by studies on the R6A-R peptide ($\Delta\Delta G° = 2.6$ kcal/mol between R6A-1 and R6A-R). Structural analysis of the $G_{i\alpha1}$-R6A complex will provide more insight into the mechanism of inhibition for the selected peptides.

We have demonstrated the use of mRNA display for the *in vitro* selection of peptides with high affinity for $G_{i\alpha1}$. By fine-tuning the selection methodology, we may be able to

further modulate peptide GDI or $G_{\beta\gamma}$ competition activities, or adjust the kinetics of G protein activation. The minimal 9-mer peptide, R6A-1, can serve as a short scaffold for the selection of new peptide sequences with affinity and specificity for other G_{α} targets. The recent development of mRNA display libraries of peptide-drug conjugates may facilitate the selection of molecules consisting of GDP or GTP analogs covalently coupled to peptides optimized for G_{α} selectivity (*46*). Selections on G proteins in various nucleotide-bound states may produce other peptide regulators that act as GDIs, GEFs, or GAPs. Small peptide modulators of G protein signaling will be useful for probing G protein function as well as serve as starting points for G protein-specific drug design (5, 6).

Acknowledgments

We thank Prof. Roger K. Sunahara (University of Michigan) for the G protein expression vector, Dr. David S. Waugh (National Cancer Institute at Frederick) for the *in vivo* biotinylation vector, Prof. Melvin I. Simon and Valeria Mancino for the polyclonal G_{β} antiserum and use of the Flexstation fluorescence plate reader, Prof. Pamela J. Bjorkman and Anthony M. Giannetti for time and technical assistance on their Biacore 2000 instrument, and Prof. Douglas C. Rees and Terry Takahashi for helpful discussions on the kinetics models. Prof. David G. Myszka (University of Utah) generously provided the analysis software, Scrubber and CLAMP. We greatly appreciate the preparative and technical expertise on protein purification provided by Dr. Shuwei Li (University of Texas Southwestern Medical Center) and Christopher T. Balmaseda. We are also grateful to Dr. Judy E. Kim for comments on the manuscript.

This work was supported by grants from the NIH (RO160416) and the Beckman Foundation to R. W. R. W. W. J. was supported in part by a DOD National Defense Science and Engineering Graduate Fellowship. R. W. R. is an Alfred P. Sloan Foundation Research Fellow.

References

1. Gilman, A. G. (1987) G proteins: Transducers of receptor-generated signals, *Annu. Rev. Biochem. 56*, 615-649.
2. Neves, S. R., Ram, P. T., and Iyengar, R. (2002) G protein pathways, *Science 296*, 1636-1639.
3. Higashijima, T., Ferguson, K. M., Sternweis, P. C., Smigel, M. D., and Gilman, A. G. (1987) Effects of Mg^{2+} and the βγ-subunit complex on the interactions of guanine nucleotides with G proteins, *J. Biol. Chem. 262*, 762-766.
4. Vetter, I. R., and Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions, *Science 294*, 1299-1304.
5. Höller, C., Freissmuth, M., and Nanoff, C. (1999) G proteins as drug targets, *Cell. Mol. Life Sci. 55*, 257-270.
6. Nürnberg, B., Tögel, W., Krause, G., Storm, R., Breitweg-Lehmann, E., and Schunack, W. (1999) Non-peptide G-protein activators as promising tools in cell biology and potential drug leads, *Eur. J. Med. Chem. 34*, 5-30.
7. Dower, W. J., and Mattheakis, L. C. (2002) *In vitro* selection as a powerful tool for the applied evolution of proteins and peptides, *Curr. Opin. Chem. Biol. 6*, 390-398.
8. Lin, H., and Cornish, V. W. (2002) Screening and selection methods for large-scale analysis of protein function, *Angew. Chem., Int. Ed. Engl. 41*, 4402-4425.
9. Martin, E. L., Rens-Domiano, S., Schatz, P. J., and Hamm, H. E. (1996) Potent peptide analogues of a G protein receptor-binding region obtained with a combinatorial library, *J. Biol. Chem. 271*, 361-366.

10. Scott, J. K., Huang, S. F., Gangadhar, B. P., Samoriski, G. M., Clapp, P., Gross, R. A., Taussig, R., and Smrcka, A. V. (2001) Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein βγ subunits is used for recognition of a subclass of effectors, *EMBO J. 20*, 767-776.

11. Ghosh, M., Peterson, Y. K., Lanier, S. M., and Smrcka, A. V. (2003) Receptor- and nucleotide exchange-independent mechanisms for promoting G protein subunit dissociation, *J. Biol. Chem. 278*, 34747-34750.

12. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12297-12302.

13. 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.

14. Bernard, M. L., Peterson, Y. K., Chung, P., Jourdan, J., and Lanier, S. M. (2001) Selective interaction of AGS3 with G-proteins and the influence of AGS3 on the activation state of G-proteins, *J. Biol. Chem. 276*, 1585-1593.

15. Natochin, M., Gasimov, K. G., and Artemyev, N. O. (2001) Inhibition of GDP/GTP exchange on Gα subunits by proteins containing G-protein regulatory motifs, *Biochemistry 40*, 5322-5328.

16. De Vries, L., Fischer, T., Tronchère, H., Brothers, G. M., Strockbine, B., Siderovski, D. P., and Farquhar, M. G. (2000) Activator of G protein signaling 3 is a guanine dissociation inhibitor for $G\alpha_i$ subunits, *Proc. Natl. Acad. Sci. U.S.A. 97*, 14364-14369.

17. Kimple, R. J., De Vries, L., Tronchère, H., Behe, C. I., Morris, R. A., Farquhar, M. G., and Siderovski, D. P. (2001) RGS12 and RGS14 GoLoco motifs are $G\alpha_i$

interaction sites with guanine nucleotide dissociation inhibitor activity, *J. Biol. Chem. 276*, 29275-29281.

18. Natochin, M., Lester, B., Peterson, Y. K., Bernard, M. L., Lanier, S. M., and Artemyev, N. O. (2000) AGS3 inhibits GDP dissociation from Gα subunits of the G_i family and rhodopsin-dependent activation of transducin, *J. Biol. Chem. 275*, 40981-40985.

19. Kimple, R. J., Willard, F. S., and Siderovski, D. P. (2002) The GoLoco motif: Heralding a new tango between G protein signaling and cell division, *Mol. Interv. 2*, 88-100.

20. Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., Graber, S. G., and Lanier, S. M. (2000) Stabilization of the GDP-bound conformation of Giα by a peptide derived from the G-protein regulatory motif of AGS3, *J. Biol. Chem. 275*, 33193-33196.

21. Takesono, A., Cismowski, M. J., Ribas, C., Bernard, M., Chung, P., Hazard, S., Duzic, E., and Lanier, S. M. (1999) Receptor-independent activators of heterotrimeric G-protein signaling pathways, *J. Biol. Chem. 274*, 33202-33205.

22. Lee, E., Linder, M. E., and Gilman, A. G. (1994) Expression of G-protein α subunits in *Escherichia coli*, *Methods Enzymol. 237*, 146-164.

23. Tsao, K.-L., DeBarbieri, B., Michel, H., and Waugh, D. S. (1996) A versatile plasmid expression vector for the production of biotinylated proteins by site-specific, enzymatic modification in *Escherichia coli*, *Gene 169*, 59-64.

24. Chong, S. R., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus,

H., and Xu, M.-Q. (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, *Gene 192*, 271-281.

25. Lesaicherre, M.-L., Lue, R. Y. P., Chen, G. Y. J., Zhu, Q., and Yao, S. Q. (2002) Intein-mediated biotinylation of proteins and its application in a protein microarray, *J. Am. Chem. Soc. 124*, 8768-8769.

26. Tolbert, T. J., and Wong, C.-H. (2000) Intein-mediated synthesis of proteins containing carbohydrates and other molecular probes, *J. Am. Chem. Soc. 122*, 5421-5428.

27. Scopes, R. K. (1974) Measurement of protein by spectrophotometry at 205 nm, *Anal. Biochem. 59*, 277-282.

28. Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods Enzymol. 180*, 51-62.

29. Liu, R., Barrick, J. E., Szostak, J. W., and Roberts, R. W. (2000) Optimized synthesis of RNA-protein fusions for *in vitro* protein selection, *Methods Enzymol. 318*, 268-293.

30. Barrick, J. E., Takahashi, T. T., Ren, J. S., 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.

31. Myszka, D. G. (2000) Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE, *Methods Enzymol. 323*, 325-340.

32. Myszka, D. G., and Morton, T. A. (1998) CLAMP: A biosensor kinetic data analysis program, *Trends Biochem. Sci. 23*, 149-150.

33. Fung, B. K.-K., and Nash, C. R. (1983) Characterization of transducin from bovine retinal rod outer segments. II. Evidence for distinct binding sites and conformational changes revealed by limited proteolysis with trypsin, *J. Biol. Chem. 258*, 10503-10510.

34. Hurley, J. B., Simon, M. I., Teplow, D. B., Robishaw, J. D., and Gilman, A. G. (1984) Homologies between signal transducing G proteins and ras gene products, *Science 226*, 860-862.

35. 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 Gα subunits, *Nature 416*, 878-881.

36. Peterson, Y. K., Hazard, S., Graber, S. G., and Lanier, S. M. (2002) Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins, *J. Biol. Chem. 277*, 6767-6770.

37. Higashijima, T., Ferguson, K. M., Sternweis, P. C., Ross, E. M., Smigel, M. D., and Gilman, A. G. (1987) The effect of activating ligands on the intrinsic fluorescence of guanine nucleotide-binding regulatory proteins, *J. Biol. Chem. 262*, 752-756.

38. McEwen, D. P., Gee, K. R., Kang, H. C., and Neubig, R. R. (2001) Fluorescent BODIPY-GTP analogs: Real-time measurement of nucleotide binding to G proteins, *Anal. Biochem. 291*, 109-117.

39. Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface, *Science 267*, 383-386.

40. Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., and Sprang, S. R. (1995) The structure of the G protein heterotrimer $G_{i\alpha1}\beta_1\gamma_2$, *Cell 83*, 1047-1058.

41. Adhikari, A., and Sprang, S. R. (2003) Thermodynamic characterization of the binding of activator of G protein signaling 3 (AGS3) and peptides derived from AGS3 with $G\alpha_{i1}$, *J. Biol. Chem. 278*, 51825-51832.

42. Bourne, H. R. (1997) G proteins. The arginine finger strikes again, *Nature 389*, 673-674.

43. Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994) Structures of active conformations of $G_{i\alpha1}$ and the mechanism of GTP hydrolysis, *Science 265*, 1405-1412.

44. Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmüller, L., Lautwein, A., Schmitz, F., and Wittinghofer, A. (1997) The Ras-RasGAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants, *Science 277*, 333-338.

45. Tesmer, J. J. G., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997) Structure of RGS4 bound to AlF_4^--activated $G_{i\alpha1}$: Stabilization of the transition state for GTP hydrolysis, *Cell 89*, 251-261.

46. Li, S. W., and Roberts, R. W. (2003) A novel strategy for *in vitro* selection of peptide-drug conjugates, *Chem. Biol. 10*, 233-239.

47. Schatz, P. J. (1993) Use of peptide libraries to map the substrate specificity of a peptide modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia Coli*, *Biotechnology 11*, 1138-1143.

Tables

Table I. Kinetic parameters for binding of various peptides with $G_{i\alpha 1}$-GDP, determined by surface plasmon resonance.[a]

Peptide		k_a	k_d	K_D	χ^2
		$M^{-1}s^{-1}$ ($\times 10^5$)	*s^{-1} ($\times 10^{-2}$)*	*nM*	
L19 GPR	TMGEEDFFDLLAKSQSKRLDDQRVDLAGYK	5.03 (1)	4.139 (7)	82	0.61
L19 GPR R23L	TMGEEDFFDLLAKSQSKRLDDQLVDLAGYK			14000	
R6A	MSQTKRLDDQLYWWEYL	15.51 (6)	9.28 (3)	60	0.76
R6A-1	DQLYWWEYL			200	
R6A-R	DQRYWWEYL			15000	
C-GPR	MSQSKRLDDQRVDLAGYK			NB	

[a] K_D values were calculated (k_d/k_a) from kinetic parameters when available. Other K_D values were obtained by fitting steady-state binding responses. Number in parentheses represents the error in the last digit from fittings. The C-GPR control peptide was non-binding (NB) at concentrations up to 20 μM.

Figures

Figure 1. (A) *In vitro* selection scheme using mRNA display. DNA containing a T7 promoter, an untranslated region, and an ORF is transcribed, ligated to a puromycin-DNA linker, and translated to produce a pool of RNA-peptide fusions. Purified fusions are reverse transcribed prior to selection on an immobilized target ($G_{i\alpha1}$). PCR amplification of the retained cDNA produces the dsDNA template for the next round of selection. (B) Biotinylated $G_{i\alpha1}$ protein constructs. Nb-$G_{i\alpha1}$ is expressed with an N-terminal peptide biotinylation tag (bio-tag, underlined) (*23, 47*). A specific lysine (bold) in the bio-tag is biotinylated *in vivo* by biotin holoenzyme synthetase. Cb-$G_{i\alpha1}$ is expressed as a fusion protein with a C-terminal intein. Cleavage and biotinylation of the C-terminus of $G_{i\alpha1}$ occur concurrently with the addition of a biotinylated cysteine derivative. (C) Peptides used for mRNA display. A C-terminal constant peptide sequence (QLRNSCA, not shown) results from the required priming site used in PCR amplification of the original DNA templates. **X** represents a random amino acid. Residues from the GPR motif consensus are underlined.

Figure 2. (A) Selection of the C-GPR X6 extension library against $G_{i\alpha1}$. ^{35}S-methionine-labeled RNA-peptide fusions from each round of selection and the original pool were assayed for binding to immobilized $G_{i\alpha1}$ (black) or to the matrix alone (gray). Subtractive hybridization (sub hyb) was performed prior to the 7th and 8th rounds of selection to remove the dominant sequence, R6A. (B) Sequences of selected peptides. A dash indicates the same residue as the wild-type (C-GPR X6 library). Sequences with internal deletions (spaces) have been aligned by their conserved residues (bold). R6A

(boxed) was the dominant peptide from the 6th round of selection which also reemerged after round 8 despite the subtractive hybridization step. The C-terminal constant region, which was frame-shifted in sequences with deletions, is not shown.

Figure 3. Binding interactions studied by surface plasmon resonance. (A) L19 GPR and R6A specifically recognize the GDP-bound state of $G_{i\alpha1}$. L19 GPR- and R6A-MBP fusion proteins (top and bottom, respectively) were immobilized by amine-coupling in separate flow cells to a surface density of ~200 response units (RU). State-specificity of the GPR-derived peptides was determined by injection (105 μL at 0 s, 35 μL/min flow rate) of preformed $G_{i\alpha1}$-GDP, $G_{i\alpha1}$-GDP-AlF_4^-, or $G_{i\alpha1}$-GTPγS (at 1 μM $G_{i\alpha1}$). (B) Kinetics of peptide interaction with $G_{i\alpha1}$-GDP. A peptide concentration series of L19 GPR (top) and R6A (bottom) was injected (200 μL at 0 s, 100 μL/min flow rate) across ~1000 RU of immobilized Nb-$G_{i\alpha1}$, maintained in the GDP-bound state. The global kinetic fits (black) are overlaid on the original sensorgrams (gray). The derived kinetic parameters are shown in Table I. Sensorgrams have been double referenced from response curves generated by an appropriate negative control flow cell and averaged buffer blank injections.

Figure 4. Guanine nucleotide dissociation inhibitor activity. (A) GPR-derived peptides stabilize the GDP-bound state of $G_{i\alpha1}$. Tryptophan fluorescence, which is enhanced upon activation by AlF_4^-, was measured on $G_{i\alpha1}$ (200 nM) preincubated with and without 400 nM peptide (L19 GPR, R6A, or C-GPR negative control). NaF and $AlCl_3$ were added at 150 and 200 s, respectively. Average fluorescence of the first 150 s was set to zero, and

all response curves were background subtracted with a buffer or peptide blank sample. (B) GPR-derived peptides inhibit binding of a fluorescent GTPγS analog. Binding of BODIPY FL GTPγS to $G_{i\alpha1}$ causes an enhancement of fluorescence which is measured in real-time. $G_{i\alpha1}$ (100 nM final) is preincubated with and without the indicated peptide (1 μM final) prior to dilution into buffer containing BODIPY FL GTPγS (0.8 μM final). After mixing, the measurements are quickly initiated in a fluorescence plate reader, allowing up to 96 samples to be assayed simultaneously. While the GPR and $G_{i\alpha1}$ without peptide inhibitor curves can be fit with single exponentials (gray), the R6A fluorescence curve appears biphasic, requiring a double exponential (gray) to fit appropriately (dotted line shows the single exponential fit). Fluorescence curves have been background subtracted with data generated from samples lacking $G_{i\alpha1}$. (C) Peptide concentration dependence of BODIPY FL GTPγS binding. Data for L19 GPR (■) and R6A-1 (▲) are expressed as a fraction of fluorescence (± s.d.) observed in the absence of peptide inhibitor at 180 min.

Figure 5. (A) $G_{\beta1\gamma2}$ subunits co-precipitate with $G_{i\alpha1}$-GDP. Nb-$G_{i\alpha1}$ reconstituted *in vitro* with $G_{\beta1\gamma2}$ subunits was precipitated with streptavidin-agarose. The equivalent of ~33 ng of $G_{\beta1\gamma2}$ was run in each lane, and membrane transfers were probed with a G_{β} antiserum. Preincubation of the G proteins with GTPγS prevented association and co-precipitation of G_{β} subunits. The −IP lane is a pull-down without Nb-$G_{i\alpha1}$. Approximately 60% of input $G_{\beta1}$ was precipitated in a 1:1 molar mix of Nb-$G_{i\alpha1}$ and $G_{\beta1\gamma2}$. (B) L19 GPR and R6A-1 peptides compete with $G_{\beta1\gamma2}$ for binding to $G_{i\alpha1}$. Reconstituted $G_{i\alpha1\beta1\gamma2}$ was preincubated

with increasing concentrations of the indicated peptide prior to precipitation and probing as in (A). The C-GPR control peptide did not compete for binding. Full-length R6A acted comparably to the minimal peptide (data not shown).

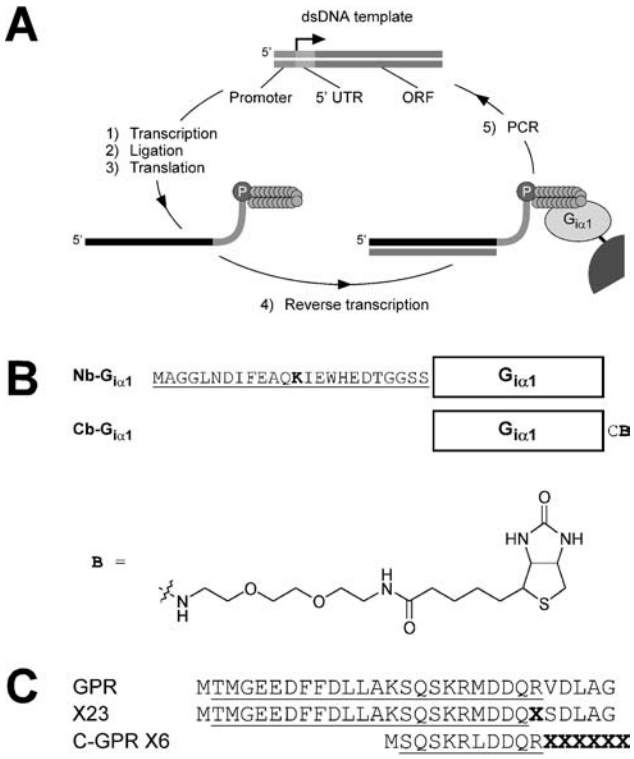

A
dsDNA template
5'
Promoter
5' UTR
ORF
1) Transcription
2) Ligation
3) Translation
4) Reverse transcription
5) PCR
P
$G_{i\alpha1}$
B
Nb-$G_{i\alpha1}$ MAGGLNDIFEAQKIEWHEDTGGSS
Cb-$G_{i\alpha1}$
$G_{i\alpha1}$
CB
B =
C
GPR MTMGEEDFFDLLAKSQSKRMDDQRVDLAG
X23 MTMGEEDFFDLLAKSQSKRMDDQXSDLAG
C-GPR X6 MSQSKRLDDQRXXXXXX

Figure 1

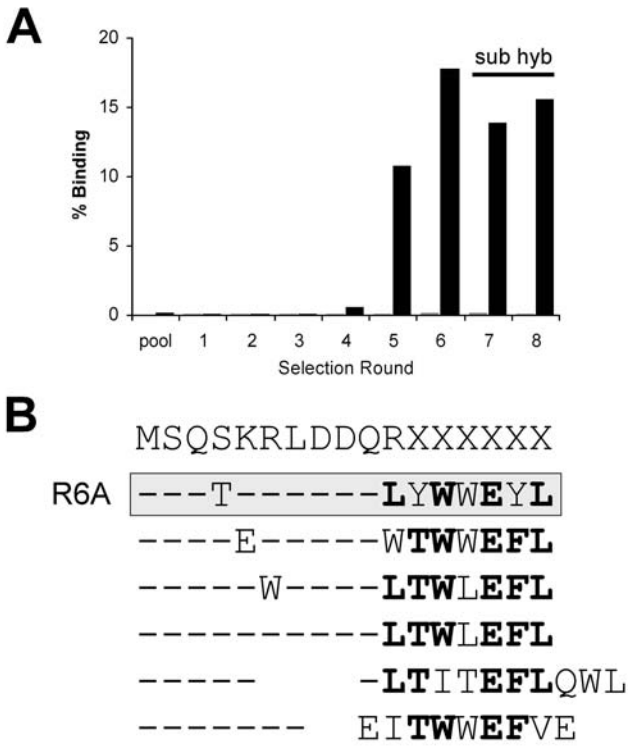

A
20
15
10
5
0
% Binding
sub hyb
pool
1
2
3
4
5
6
7
8
Selection Round
B
MSQSKRLDDQRXXXXXX
R6A
---T------LYWWEYL
----E-----WTWWEFL
-----W----LTWLEFL
----------LTWLEFL
----- -LTITEFLQWL
------- EITWWEFVE

Figure 2

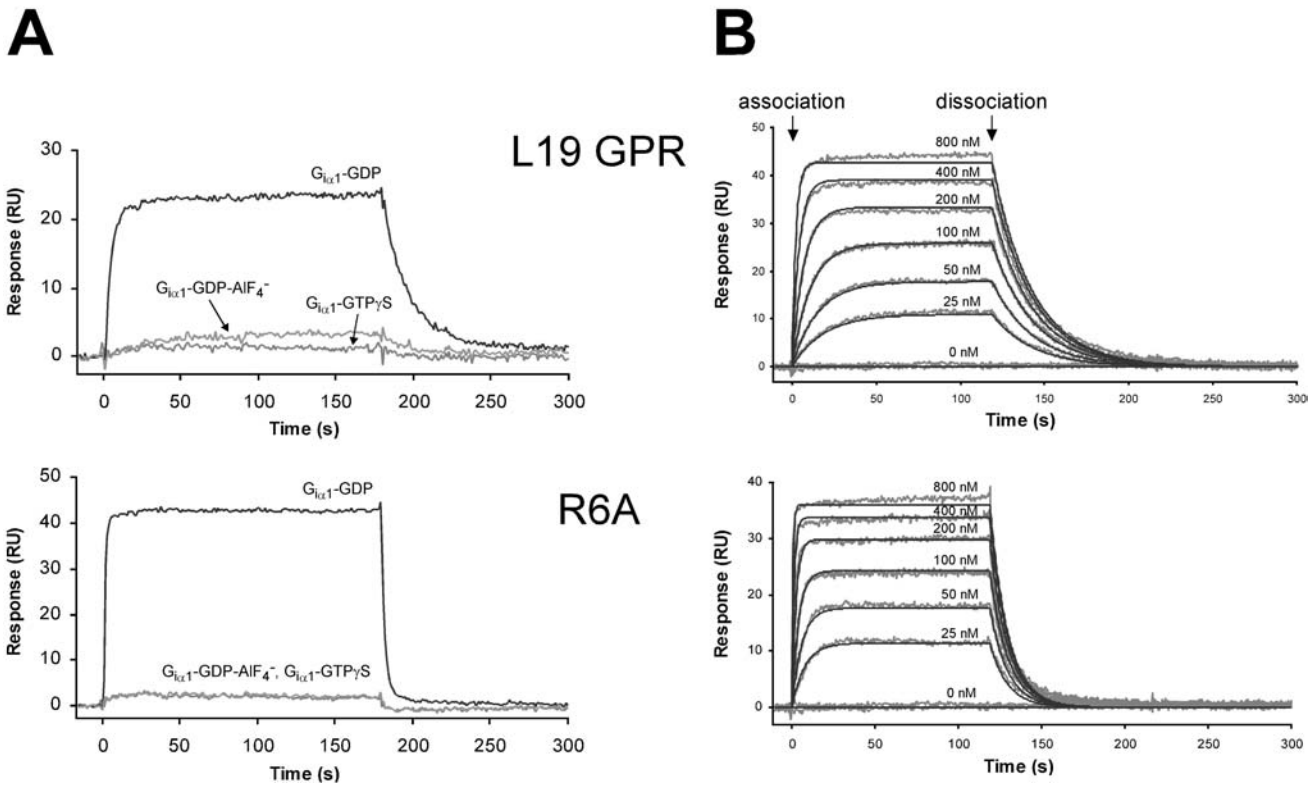

A
B
L19 GPR
R6A
Response (RU)
Time (s)
$G_{i\alpha1}$-GDP
$G_{i\alpha1}$-GDP-AlF_4^-
$G_{i\alpha1}$-GTPγS
$G_{i\alpha1}$-GDP-AlF_4^-, $G_{i\alpha1}$-GTPγS
association
dissociation
800 nM
400 nM
200 nM
100 nM
50 nM
25 nM
0 nM

Figure 3

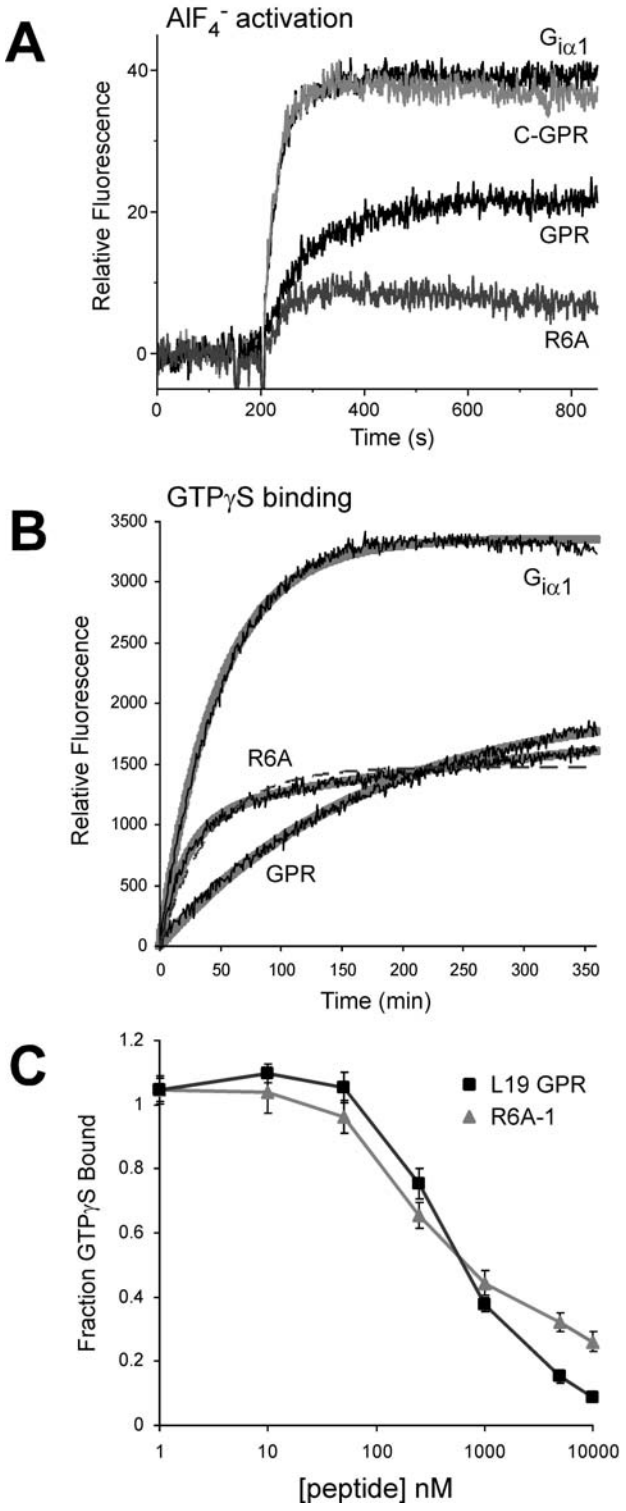

A
AlF4- activation
Gαi1
C-GPR
GPR
R6A
Relative Fluorescence
Time (s)
B
GTPγS binding
Gαi1
R6A
GPR
Relative Fluorescence
Time (min)
C
L19 GPR
R6A-1
Fraction GTPγS Bound
[peptide] nM

Figure 4

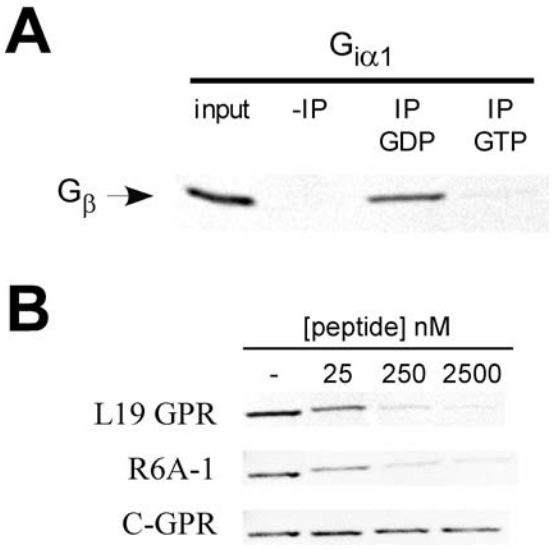

A
$G_{i\alpha1}$
input
-IP
IP GDP
IP GTP
G_β
B
[peptide] nM
-
25
250
2500
L19 GPR
R6A-1
C-GPR

Figure 5

Supporting Information

Synthesis of N,N'-D-biotinyl-2,2'-(ethylenedioxy)bis(ethylamine)-L-Cysteine

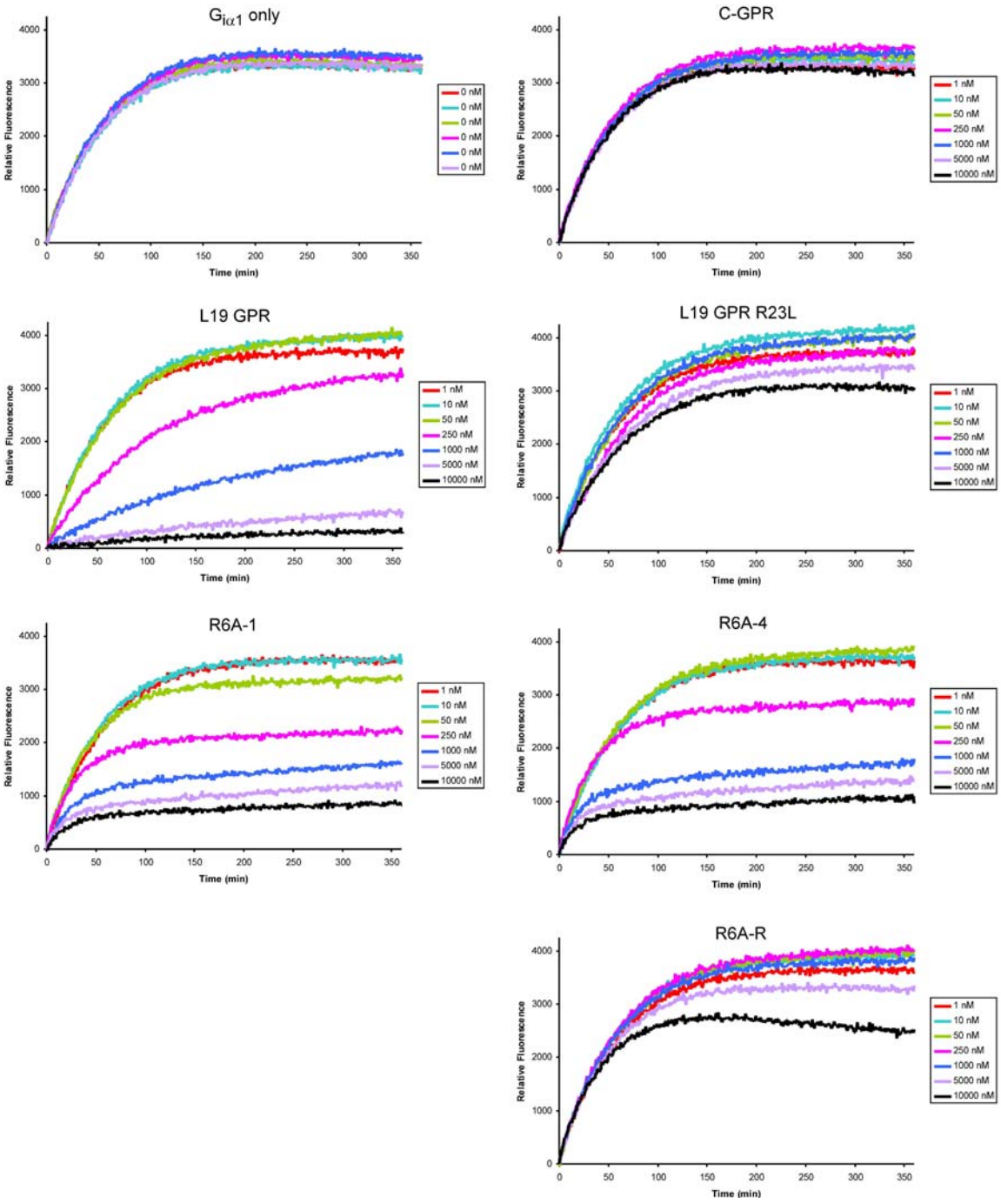
2-Cl-TrT-Cys(Mmt)-OBt resin (100 mg, 0.06 mmol capacity, Calbiochem-Novabiochem Corp., La Jolla, CA) was swelled in 3 mL of DMF at room temperature for 1 h followed by washing on a vacuum manifold with DMF and DCM. The resin was then rotated for 2.5 h at room temperature with 2,2'-(ethylenedioxy)bis(ethylamine) (500 μL, 3.42 mmol) in 3 mL of DCM. After washing as before, the resin was incubated in a solution containing D-biotin (60 mg, 0.25 mmol, dissolved in 1 mL DMSO), Pybop (130 mg, 0.25 mmol), HOBt (35 mg, 0.23 mmol), and DIPEA (90 μL , 0.52 mmol) in 2 mL of DMF. After rotating at room temperature for 9 h, the resin was washed with DMF and DCM and dried on vacuum. The resin could be stored at -20 °C until needed. For deprotection and cleavage, the resin was rotated with 5 mL of TFA/DCM/TIS (2/96/2) for 1.5 h. The cleaved biotinyl-Cys was collected by gravity filtration along with 2 additional collections using DCM. The compound was dried in vacuo, collected with MeOH, and dried again. The pellet was extracted 6 × 1 mL ether and dried in vacuo. The compound was used without further purification. ESI (MH^+) 478.2 Da (expected 478.2 Da).

Construction of pDW363A

The coding region for MBP from pDW363 was excised at the XhoI/BamHI restriction sites and purified by agarose gel electrophoresis (QIAquick gel extraction, Qiagen). The Factor Xa protease cut site was rearranged by PCR amplification of the MBP dsDNA first with primers 35.3 (5'-GGA CTA GTA AAA TCG AAG AAG GTA AAC TGG TAA TC) and 35.4 (5'-CCA TTG GAT CCT TAA TTA GTC TGC GCG TCT TTC AG)

and subsequently with primers 75.1 (5'-GAG CAC TCG AGC TCT GGA GGC ATC GAG GGT CGC ATG GGT GGC ACT AGT AAA ATC GAA GAA GGT AAA CTG GTA ATC) and 29.3 (5'-CCA TTG GAT CCT TAA TTA GTC TGC GCG TC) using Herculase DNA polymerase (Stratagene, La Jolla, CA). The MBP gene was then ligated back into pDW363 at the XhoI/BamHI sites to produce the vector, pDW363A.

Supplemental Figure 1. Effect of various peptides on GTPγS binding. Fluorescence enhancement of BODIPY FL GTPγS binding to $G_{i\alpha1}$ was observed in the presence of various peptides at the indicated concentrations, as described in the Experimental Procedures. The peptide sequences are given in Table I of the manuscript, except for R6A-4 (SQTKRLDDQLYWWEYL). "$G_{i\alpha1}$ only" demonstrates the repeatability of the fluorescence enhancement without peptide inhibitor in 6 separate wells of a 96-well plate experiment. The differing kinetics of inhibition between the L19 GPR consensus peptide and the selected peptides (R6A-1 and R6A-4) is easily seen. The L19 GPR R23L and R6A-R mutant peptides exhibit significantly reduced GDI activity.



Supplemental Figure 1

Chapter 3

A peptide core motif for binding heterotrimeric G protein α subunits

William W. Ja and Richard W. Roberts

Recently, *in vitro* selection using mRNA display was used to identify a novel peptide sequence that binds with high affinity to $G_{i\alpha1}$. The peptide was minimized to a 9-residue sequence (R6A-1) that retains high affinity and specificity for the GDP-bound state of $G_{i\alpha1}$ and acts as a guanine nucleotide dissociation inhibitor (GDI). Binding assays with immobilized R6A-1 reveal that the peptide interacts with G_{α} subunits representing all four G protein classes [i1–3, oA, q, s(s), 12, and 15], in contrast with the consensus G protein regulatory (GPR) sequence, a 28-mer peptide GDI derived from the GoLoco/GPR motif, which binds only to $G_{i\alpha1-3}$ in this assay. Binding to R6A-1 by G_{α} subunits completely excludes association with $G_{\beta\gamma}$. These findings suggest that the R6A-1 core motif might be suitable as a starting point for the identification of peptides exhibiting novel activities and/or specificity for particular G protein subclasses. A new mRNA display library based on the R6A-1 sequence has been constructed and used to select for peptides that bind $G_{i\alpha1}$, confirming that the 9-mer core is the minimal consensus. Negligible conservation is seen in residues flanking the core motif, suggesting that they play a minimal role in binding. However, these flanking regions may confer unique properties to the core peptide and the selected peptides are currently being characterized by their binding specificities to other G proteins.

Introduction

Heterotrimeric guanine nucleotide-binding proteins (G proteins), composed of α, β, and γ subunits, mediate signaling from cell-surface receptors (GPCRs)[1] to a wide variety of effectors (*1, 2*). In the inactive state, intracellular $G_{\alpha\beta\gamma}$ heterotrimers are coupled to the membrane-spanning GPCR. Activation of the receptor results in GDP exchange with GTP in the G_α subunit, dissociation of $G_{\beta\gamma}$ heterodimers from G_α, and subsequent signal transduction through G_α-GTP and/or $G_{\beta\gamma}$. The inherent guanosine triphosphatase (GTPase) activity of G_α, which is accelerated by various GTPase-activating proteins (GAPs), returns the protein to the GDP-bound state, resulting in reassociation with $G_{\beta\gamma}$ and termination of signaling.

Approximately 50% of currently marketed drugs target GPCRs (*3, 4*). Drug discovery targeting G proteins directly has traditionally been difficult due to (1) the broad spectrum of signaling events mediated at the G protein level, (2) the requirement that drugs must cross the cell membrane to reach intracellular G proteins, and (3) the high sequence and structural similarities between G protein classes (*5, 6*). Nevertheless, a number of diseases have been attributed to aberrant G protein activity (*7, 8*) and direct G protein ligands will provide new approaches and selectivities for drug treatment (*5, 6*).

Selection methodologies can facilitate the isolation of rare molecules with unique functions, such as specificity for particular G protein classes, from large libraries (*9, 10*). We recently demonstrated that mRNA display, a selection technique where peptides are covalently attached to their encoded RNA, could be used to isolate $G_{i\alpha1}$-binding

[1] Abbreviations: GAP, GTPase-activating protein; GDI: guanine nucleotide dissociation inhibitor; GoLoco, $G_{\alpha i/o}$-Loco interaction; GPCR, G protein-coupled receptor; GPR, G protein regulatory; MALDI-TOF, matrix-assisted laser desorption/ionization time-of- flight; MBP, maltose-binding protein.

sequences (*11*). The dominant peptide from the selection, as well as a minimized, active 9-mer sequence (R6A-1), acts as a guanine nucleotide dissociation inhibitor (GDI) and competes with $G_{\beta\gamma}$ for binding to $G_{i\alpha1}$.

To examine the specificity of R6A-derived sequences, we assayed binding of various *in vitro* translated G_α subunits to immobilized peptides. Surprisingly, the R6A-1 core motif binds strongly to all tested G_α subunits. Binding of R6A-1 is generally specific for the GDP-bound state of each G_α subunit and appears to exclude heterotrimer formation with $G_{\beta\gamma}$. A new mRNA display library based on the core motif was synthesized and used to select for peptides that bind $G_{i\alpha1}$ in either the GDP or the GDP-AlF_4^- state. Functional sequences were isolated quickly (within three rounds of selection), demonstrating the utility of the library for identifying G protein-binding sequences. We are currently characterizing several of the newly isolated peptides and using the core motif library to target other G protein subclasses.

Experimental Procedures

Materials

Human cDNA clones encoding various G proteins were obtained from the UMR cDNA Resource Center (http://www.cdna.org) in the pcDNA3.1+ vector (Invitrogen Corp., Carlsbad, CA). The G_α subunits used were i1, i2, i3, oA, q, s (short-form), 12, and 15. All *in vitro* translated G_β and G_γ subunits refer to $G_{\beta1}$ and N-terminal hemagglutinin (HA) tagged $G_{\gamma2}$, respectively. Reagents were obtained from Sigma or VWR, unless otherwise noted. DNA oligos were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA) except for the 115.1 library template which was synthesized at the W.M. Keck

Facility (Yale University, New Haven, CT). DNA sequencing of selected clones was performed by Laragen, Inc. (Los Angeles, CA) from purified plasmids.

Peptide/protein preparation

The C-terminal amidated peptides GPR-biotin (TMGEE DFFDL LAKSQ SKRLD DQRVD LAGQL RNSYA **K**, **K** = biocytin), L19 GPR (TMGEE DFFDL LAKSQ SKRLD DQRVD LAGYK), R6A-1-biotin (DQLYW WEYLQ LRNSY A**K**), R6A-1 (DQLYW WEYL), and R6A-4 (SQTKR LDDQL YWWEY L) were synthesized and purified as described previously (*11*). R6A-4 lacks an N-terminal methionine that the originally studied "full-length" R6A peptide contained. Peptide masses were confirmed by MALDI-TOF mass spectrometry and concentrations were determined using a calculated extinction coefficient (http://paris.chem.yale.edu/extinct.html) for absorbance at 280 nm. Biotinylated peptides were immobilized using streptavidin-agarose (Immobilized NeutrAvidin on Agarose, Pierce). Approximately 500–800 pmol of biotinylated peptide were used per 10 μL of agarose.

Full-length R6A (MSQTK RLDDQ LYWWE YL) was expressed as a fusion to maltose-binding protein (MBP) using an *in vivo* biotinylation system (*12*). Cloning, expression, and purification were performed as described previously (*11*). R6A-MBP or MBP was immobilized by random amine coupling on CNBr-sepharose 4B (Amersham Biosciences, Piscataway, NJ) as per the manufacturer's instructions at a concentration of approximately 1 mg/mL of the hydrated matrix.

N-terminal biotinylated $G_{i\alpha1}$ (Nb-$G_{i\alpha1}$) and $G_{i\alpha3}$ (Nb-$G_{i\alpha3}$) were expressed and purified as described previously (*11*). Nb-$G_{i\alpha2}$ was constructed and expressed similarly.

In vitro *translation*

All G protein subunits were translated separately in coupled transcription/translation reactions using the TNT reticulocyte lysate system (Promega, Madison, WI). Typically, 0.3–1.0 μg of plasmid DNA and 25 μCi of L-[^{35}S]-methionine (MP Biomedicals, Irvine, CA) were used in a 25 μL reaction. Translation efficiency of G_α subunits was quantitated by TCA precipitation of a 2 μL aliquot of each reaction, as per the manufacturer's instructions. G_γ reactions were supplemented with 10 μM mevalonic acid lactone to ensure complete polyisoprenylation (*13*). To make $G_{\beta\gamma}$ heterodimers, independently translated subunits were mixed together (3:1 by volume, G_β:G_γ) and incubated at 37 °C for 30 min. To reconstitute $G_{\alpha\beta\gamma}$ heterotrimers, equal volumes of G_α and preformed $G_{\beta\gamma}$ were mixed and incubated at 37 °C for an additional 30 min. For the heterotrimer immunoprecipitation assays, G_β was translated without radioactive labeling due to possible interference in the resolution of G_α subunits by SDS-PAGE. These unlabeled reactions were supplemented with L-methionine (40 μM final)

G_α interaction assay

G_α translation reactions were desalted and exchanged using MicroSpin G-25 columns (Amersham) into buffer [50 mM HEPES-KOH at pH 7.5, 6 mM $MgCl_2$, 75 mM sucrose, 1 mM EDTA, 1 μM GDP, and 0.05% (v/v) Tween 20 (Bio-Rad Laboratories, Hercules, CA)]. Equivalent aliquots (2 to 6 μL) of the desalted G_α subunits were used for the *in vitro* binding assays. G_α was added to 0.6 mL of binding buffer [25 mM HEPES-KOH at pH 7.5, 5 mM $MgCl_2$, 1 mM EDTA, 150 mM NaCl, 0.05% Tween 20, 0.05% (w/v) BSA,

1 mM β-mercaptoethanol, and 10 μM GDP] containing ~10 μL matrix with or without immobilized target. After rotating at 4 °C for 1 h, samples were briefly centrifuged and the supernatant was removed. The matrix was transferred to a 0.45 μm cellulose acetate spin filter (CoStar Spin-X, Corning, Inc., Corning, NY) and washed with 3 × 0.6 mL of binding buffer at 4 °C (1500 × g, ~40 s). The washed matrix was then removed from the spin filter for scintillation counting or analysis by SDS-PAGE. Relative binding is reported (+ standard deviation, when available) based on the bound cpm divided by the input protein counts, as determined from the TCA precipitation. Assays with aluminum fluoride were performed identically, except that the binding buffer was supplemented with 50 mM NaF and 25 μM $AlCl_3$.

$G_{\alpha\beta\gamma}$ heterotrimer immunoprecipitation

Equivalent aliquots (10 μL) of reconstituted $G_{\alpha\beta\gamma}$ heterotrimer were added to 0.6 mL of binding buffer containing ~10 μL matrix or 1 μL anti-HA mAb (Sigma, clone HA-7). After rotating at 4 °C for 1 h, ~10 μL of protein G-sepharose 4B Fast Flow was added to the mAb-containing samples. After an additional 30 min of rotating at 4 °C, immobilization matrices were washed in 0.45 μm spin filters (3 × 0.6 mL of binding buffer) as described above. A 4th wash was performed in batch, after transferring the matrices to new tubes, to prevent contamination from the spin filter membrane. The samples were resuspended in 2× SDS-loading buffer, incubated at 90 °C for 5 min, and analyzed by tricine SDS-PAGE. Gels were fixed, dried *in vacuo*, and imaged by autoradiography (Storm Phosphorimager, Amersham).

$G_{\beta\gamma}$ competition assay

Approximately 10 μL of the target matrix (~1 mg Nb-G_α i1, i2, or i3 per mL of NeutrAvidin-agarose) was incubated in 0.6 mL of binding buffer with and without various concentrations of added peptide (R6A-1, R6A, or L19 GPR) for 5 min at room temperature. Equivalent aliquots (~5 μL) of ^{35}S-methionine-labeled $G_{\beta\gamma}$ heterodimers were then added and the samples were rotated at 4 °C for 1 h. Samples were washed in spin filters as described for the G_α interaction assays (3 × 0.6 mL washes) and the amount of bound, radiolabeled protein was determined by scintillation counting of the matrices. For IC_{50} determinations, binding data were scaled relative to the bound counts in the absence of peptide competitor.

115.1 library mRNA display selection

The doped R6A-1 library was constructed by PCR amplification of oligo 115.1 [5'- AGC AGA CAG ACT AGT GTA ACC GCC $(SNN)_6$ (S13) (641) (542) (521) (521) (641) (S13) (543) (642) $(SNN)_6$ CAT TGT AAT TGT AAA TAG TAA TTG TCC C; 1 = 7:1:1:1, 2 = 1:7:1:1, 3 = 1:1:7:1, 4 = 1:1:1:7, A:C:G:T; 5 = 9:1, 6 = 1:9, C:G; N = A, C, G, or T; S = C or G (ratios have been adjusted for synthesis incorporation rates)] with primers 47T7FP (5'- GGA TTC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT AC) and 22.9 (5'-AGC AGA CAG ACT AGT GTA ACC G). PCR (40 × 200 μL reactions) was performed with 0.1 μM 115.1 template, 1 μM primers, and 200 μM each dNTP (cycling parameters: 97 °C 2 min, 52 °C 2 min, 72 °C 4 min, followed by 4 cycles of 97 °C 2 min, 58 °C 2 min, 72 °C 4 min and a 5 min at °72 C chase cycle). Amplified DNA was phenol-extracted with Phase Lock Gel (Brinkmann Instruments) and

desalted by isopropanol precipitation. *In vitro* transcription, ligation of the mRNA to the puromycin linker (pF30P), and purification of the RNA-F30P template were performed as described previously (*11*), except that the splint oligo 23.8 (5'-TTT TTT TTT TTN AGC AGA CAG AC) was used for the ligation reaction.

RNA-peptide fusions were prepared from rabbit reticulocyte lysate, purified on oligo-dT cellulose, reverse-transcribed, and selected against immobilized Nb-$G_{i\alpha1}$ as described previously (*11*) using a modified selection buffer [25 mM HEPES-KOH at pH 7.5, 150 mM NaCl, 0.05% Tween 20, 1 mM β-mercaptoethanol, 10 μM GDP, 20 μM EDTA, 5 mM $MgCl_2$, 0.05% BSA, and 1 μg/mL (w/v) yeast tRNA]. For selections against Nb-$G_{i\alpha1}$ in the GDP-AlF_4^- state, the selection buffer was supplemented with 10 mM NaF and 25 μM $AlCl_3$. For the selection against Nb-$G_{i\alpha1}$-GDP, stringency was increased in the 4th round by performing the binding at 37 °C and in the 5th and 6th rounds by allowing the target matrix to incubate in selection buffer containing free, non-biotinylated $G_{i\alpha1}$. Selected fusions were PCR amplified for use as the template in the subsequent round and for cloning and DNA sequencing.

Purified RNA-peptide fusions of individual clones were assayed for binding to Nb-$G_{i\alpha1}$ in selection buffer without yeast tRNA. ^{35}S-methionine-labeled fusions were RNase-treated (RNase, DNase-free, Roche) prior to addition to 1 mL of buffer containing ~10 μL of Nb-$G_{i\alpha1}$ (~10 μg) on NeutrAvidin-agarose. After binding at 4 °C for 1 h, 3 × 0.6 mL buffer washes were performed using spin filters (0.45 μm, Costar Spin-X) and the agarose was scintillation counted.

Results

R6A-1 is a core motif for G_α-binding

To investigate the specificity of the R6A-1 minimal peptide, a pull-down assay was developed using radiolabeled, *in vitro* translated G protein subunits (Figure 1A). Cell-free coupled transcription/translation offered a rapid way of screening multiple G proteins (*14-16*) and cDNA clones for human G proteins were readily available. R6A-1 and L19 GPR peptides were synthesized with a C-terminal biotin-containing linker peptide derived from the constant region used in the original selection (*11*). The full-length R6A peptide was also expressed as an N-terminal fusion to MBP, which was subsequently immobilized by random amine coupling. ^{35}S-methionine-labeled $G_{i\alpha1}$ was first tested against immobilized L19 GPR and full-length R6A, demonstrating specific pull-down of full-length $G_{i\alpha1}$, as well as a slightly lower molecular weight band that corresponds to an alternate translation initiation site (Figure 1B).

Previous results demonstrated that the consensus GPR peptide had high affinity for $G_{i\alpha}$ and weaker affinity for $G_{o\alpha}$ (*17, 18*). In our assay, the L19 GPR peptide exhibited binding only to $G_{i\alpha1-3}$ (Figure 2A). The R6A-1 minimal peptide exhibited strong binding for all heterotrimeric G_α subunits tested (Figure 2B). The full-length R6A sequence, however, demonstrated significantly weaker binding to a number of G proteins, especially to G_α o, s, and 15 (Figure 2C). It is not clear whether the differences in affinity to the various G proteins are due to the N-terminal flanking region of the full-length R6A sequence, the altered immobilization scheme (random amine coupling versus C-terminal biotinylation on the R6A-1 peptide), or steric effects from the comparatively

large MBP fusion. However, assuming that the various G proteins are structurally homologous and that the R6A peptide binds to each G_α subunit in the same manner, the differences in relative binding would seem to be a direct result of the R6A flanking residues.

Core motif remains specific for the GDP state

To confirm the interaction of the R6A-derived peptides to various G proteins and establish the nucleotide state specificity, the effect of aluminum fluoride on binding was determined. Previously, it was shown that R6A was highly specific for the GDP state of $G_{i\alpha1}$ and did not bind to either $G_{i\alpha1}$-GDP-AlF_4^- or $G_{i\alpha1}$-GTPγS (*11*). *In vitro* translated, radiolabeled G proteins were assayed for binding to immobilized R6A-1 and R6A-MBP in the presence and absence of aluminum fluoride. The minimal peptide was specific for the GDP state for all G proteins except for G_α 12 and 15, where the effect of aluminum fluoride was negligible (Figure 3A). Full-length R6A-MBP demonstrated strong specificity for the GDP state for all G proteins (Figure 3B), including $G_{12\alpha}$, where aluminum fluoride reduced binding to the background levels seen previously (Figure 2C). The minimal interaction seen with $G_{s\alpha}$ to R6A-MBP was also confirmed, based on the reduced binding in the presence of aluminum fluoride.

R6A competes with $G_{\beta\gamma}$ for binding to G_α subunits

Previously, GPR- and R6A-derived peptides had been shown to compete with $G_{\beta\gamma}$ heterodimers for binding to $G_{i\alpha1}$ (*11, 19-22*). To determine whether R6A would exclude $G_{\beta\gamma}$-binding for other G_α subunits, *in vitro* translated $G_{\beta1}$ and HA-tagged $G_{\gamma2}$ (Figure 4A)

were reconstituted with various G_α subunits and pulled down by immobilized L19 GPR, R6A-MBP, or an anti-HA monoclonal antibody. The tested G_α subunits (i1–3 and q) all appeared to couple to $G_{\beta\gamma}$ heterodimers (Figure 4B). Co-precipitation of G_γ subunits was not seen when heterotrimers were pulled down by L19 GPR or R6A-MBP (Figure 4B), clearly indicating that binding to these motifs excludes $G_{\beta\gamma}$ interaction (Figure 4C). Results for R6A-MBP with reconstituted $G_{12\alpha\beta1\gamma2}$ heterotrimers were similar (data not shown). While the GPR and R6A peptides recognize N-terminal truncations of $G_{i\alpha1-3}$, coupling to $G_{\beta\gamma}$ appears to require the full-length protein (Figure 4B, HA immunoprecipitation).

Peptide competition with $G_{\beta\gamma}$ heterodimers was also demonstrated in a reverse experiment with immobilized G_α subunits. N-terminal biotinylated $G_{i\alpha1-3}$ were immobilized on streptavidin and used to pull-down radiolabeled $G_{\beta\gamma}$. Immobilized G proteins were active for binding $G_{\beta\gamma}$, specifically in the GDP state (Figure 5A). Pull-down assays with single G_β or G_γ subunits on Nb-$G_{i\alpha1}$ resulted in ~10% and ~1% background binding, respectively, compared with reconstituted $G_{\beta\gamma}$-binding (data not shown). It is unclear whether this reflects non-specific binding to the matrix or the presence of free, unlabeled G proteins in the reticulocyte lysate, which would allow formation of intact heterodimers.

Competition with $G_{\beta\gamma}$ was measured by preincubation of immobilized $G_{i\alpha}$ with various concentrations of peptide. Surprisingly, $G_{\beta\gamma}$ competition could not be measured in this assay for the L19 GPR consensus peptide with $G_{i\alpha1}$ (Figure 5B). R6A-1 and full-length R6A-4 peptides demonstrated IC_{50} values of 3.0 and 1.0 μM, respectively, for

$G_{i\alpha1}$. L19 GPR and R6A-4 demonstrated similar IC_{50} values for immobilized $G_{i\alpha3}$ (Figure 5C). The differences in peptide competition for G_α i1 and i3 may result from variations in heterodimer coupling to or peptide recognition of the two G protein subclasses. The negative control peptide C-GPR had no effect on $G_{\beta\gamma}$-coupling for either G_α subunit at concentrations up to 10 μM (data not shown).

mRNA display with a doped R6A-1 library

While the R6A-1 9-mer demonstrated ubiquitous binding to the various G_α subunits, additional flanking residues may confer unique specificities and/or activities to the core peptide. A new mRNA display peptide library was designed and synthesized based on the R6A-1 core motif. The 115.1 template, after PCR amplification, contained a T7 promoter for transcription, an untranslated region (5' UTR), a start codon, the library X_6-DQLYWWEYL-X_6 where the core residues were doped to give approximately 50% wild-type at each R6A-1 residue, and a 3' constant sequence. Sequencing of randomly chosen clones from the initial pool revealed a reasonable distribution for wild-type residues in the core motif, in agreement with theoretical calculations (data not shown).

To demonstrate the utility of the 115.1 core motif library for identifying G protein ligands, *in vitro* selection was performed against immobilized $G_{i\alpha1}$ in the GDP and GDP-aluminum fluoride states. Addition of aluminum fluoride produces a stable $G_{i\alpha1}$-GDP-AlF_4^- complex which mimics the transition state for GTP hydrolysis (*23, 24*). Based on quantitation of the purified, ^{35}S-labeled RNA-peptide fusions and the estimated overall L-methionine concentration, the complexity of the starting library for each selection target was ~2×10^{13}. Selection was performed on immobilized Nb-$G_{i\alpha1}$ due to the previous

finding that the R6A- and GPR-derived peptides bound preferentially to Nb-$G_{i\alpha1}$ over the C-terminal biotinylated Cb-$G_{i\alpha1}$ (*11*). Six rounds of selection were performed against each target, with significant binding observed by the third rounds (Figure 6A and D). Sequences isolated from the selection are listed in Table I.

Alignment of the sequences from the $G_{i\alpha1}$-GDP selection confirmed the original R6A consensus (*11*), although Phe was preferred in the penultimate residue rather than Tyr (Figure 6B). Flanking residues in the random hexamer regions appeared to play a minimal role with no obvious sequence conservation. Positions 17 and 18, however, did seem to favor a Glu-Leu pair. There also seemed to be a preference for positively charged side chains in the N-terminal region, with significantly fewer Lys and Arg residues in the C-terminus (Figure 6C). A reduction in Ala, Ile, Val, and Glu and an increase in Lys, Pro, and Cys were observed when comparing the amino acid usage of the random domains between the 3rd and 6th round sequences. The increase in Cys may improve peptide affinity for the higher stringency selection rounds due to peptide cyclization, oligomerization, or disulfide-bridging with available surface Cys on $G_{i\alpha1}$ during the binding and wash steps. How the peptide properties changed with the shift in usage of other amino acids is unclear.

Selected peptides from the $G_{i\alpha1}$-GDP-AlF_4^- selection (Table II) demonstrated a slightly different consensus than sequences from the GDP state selection (Figure 6E and F). These differences may be critical residues in nucleotide state-specific recognition. Binding assays of several individual sequences, however, reveal only a marginal shift in propensities toward the aluminum fluoride state, with most peptides still favoring the GDP- over the GDP-AlF_4^--bound state. Nevertheless, peptides selected against GDP

(from the C-GPR X6 (*11*) or core motif libraries) clearly demonstrate different binding tendencies than the sequences isolated from the $G_{i\alpha1}$-GDP-AlF_4^- selection (Figure 7).

Discussion

Previously, *in vitro* selection with an mRNA display library was used to isolate novel peptide sequences that act as GDIs for $G_{i\alpha1}$ (*11*). The minimal 9-mer peptide, R6A-1, retained high affinity and competed with $G_{\beta\gamma}$ for binding to $G_{i\alpha1}$. Here, we have further characterized the R6A-derived peptides and have determined that the core motif binds to a variety of G_α subunits representing all four G protein classes. Binding appeared to be specific for heterotrimeric G proteins as there was negligible interaction with the small G protein, H-Ras. Impressively, R6A remained competitive with $G_{\beta\gamma}$ heterodimers for binding to G_α i1–3, q, and 12 (others were untested).

Full-length R6A and the R6A-1 core motif exhibited differences in state specificity and in relative binding to the various G proteins. These findings suggest that flanking residues may play a strong role in modulating the properties of the 9-mer core peptide. The use of flanking residues to gain G protein subclass specificity was recently demonstrated for a GoLoco or GPR peptide derived from RGS14. From the crystal structure of the $G_{i\alpha1}$:GoLoco peptide complex, it was determined that residues C-terminal to the GoLoco consensus, a region that is poorly conserved among GoLoco/GPR-containing proteins, interacted extensively with residues that differed between $G_{i\alpha1}$ and $G_{o\alpha}$ subunits, thereby controlling the specificity of GoLoco-G_α interactions (*25*).

The core motif library was designed to encode the R6A-1 peptide flanked by random hexamers. Nucleotide incorporation was controlled such that approximately 50%

conservation was seen for each "wild-type" residue in the core. This library will be useful for the selection of peptides that are specific for various G protein subclasses or nucleotide-bound states. As a demonstration of the utility of this library, *in vitro* selection was performed against immobilized $G_{i\alpha1}$ in two unique, nucleotide-bound states. $G_{i\alpha1}$-binding peptides were enriched remarkably quickly, as significant binding of the pool was seen in the 3rd selection rounds (Figure 6A and D). Assuming a maximum enrichment of 10- to 1000-fold per round, an estimated 10^7 to 10^{11} unique, $G_{i\alpha1}$-binding peptide sequences were present in the 3rd round pools. Many of these sequences may have unique functions, such as specificity for $G_{i\alpha1}$ over other G_α subunits, which have not yet been identified. We are currently assaying individual clones using a G_α-binding screen to obtain a general gauge of each peptide's properties.

We have previously used a naïve, random 27-mer library to target the $G_{i\alpha1}$-GDP-AlF_4^- state.[2] After eight rounds of selection, enrichment for functional peptides was not seen. The successful selection of peptides against $G_{i\alpha1}$-GDP-AlF_4^- clearly demonstrates the utility of the core motif library in rapidly generating G protein ligands. The $G_{i\alpha1}$-GDP-AlF_4^- complex mimics the transition state for GTP hydrolysis (*23, 24*). RGS proteins, which bind strongly to $G_{i\alpha1}$-GDP-AlF_4^-, act as GAPs, potently catalyzing GTP hydrolysis (*23, 24, 26, 27*). Whether the selected peptides can accelerate GTP hydrolysis for $G_{i\alpha1}$ is under investigation.

We have demonstrated that the core motif library, based on the R6A-1 peptide, is useful for the rapid isolation of G protein-binding peptides. The structural determination

[2] Ja and Roberts, unpublished results.

of an R6A-derived peptide:$G_{i\alpha 1}$ complex will greatly facilitate the molecular design of other specific, potent modulators of G protein signaling. The Ras family of small G proteins represents another rich source of bona fide drug targets (*28-30*). Whether the core motif library can be used against this G protein superfamily, which shares significant structural homology to the nucleotide-binding (Ras-like) domain of heterotrimeric G_α subunits, is unknown. Future directions include using the library to target other heterotrimeric G protein classes and nucleotide-bound states (e.g., G_α-GTP or nucleotide-free). The use of free, non-immobilized G protein competitors will enable the direct selection of peptides that are specific for particular G_α subclasses.

Acknowledgments

We are grateful for suggestions on G protein translation provided by Dr. Bradley M. Denker (Harvard Institutes of Medicine) and Prof. Carl Schmidt (University of Delaware). Dr. David S. Waugh (National Cancer Institute at Frederick) generously provided the original pDW363 vector. We also thank Dr. Yuri K. Peterson (Duke University) for advice on characterizing the R6A-derived peptides and Terry T. Takahashi and Christine T. Ueda for comments on the manuscript. This work was supported by grants from the NIH (RO160416) and the Beckman Foundation to R. W. R. W. W. J. was supported in part by a DOD National Defense Science and Engineering Graduate Fellowship. R. W. R. is an Alfred P. Sloan Foundation Research Fellow.

References

1. Gilman, A. G. (1987) G proteins: transducers of receptor-generated signals, *Annu. Rev. Biochem. 56*, 615-649.
2. Neves, S. R., Ram, P. T., and Iyengar, R. (2002) G protein pathways, *Science 296*, 1636-1639.
3. Drews, J. (2000) Drug discovery: a historical perspective, *Science 287*, 1960-1964.
4. Howard, A. D., McAllister, G., Feighner, S. D., Liu, Q., Nargund, R. P., Van der Ploeg, L. H., and Patchett, A. A. (2001) Orphan G-protein-coupled receptors and natural ligand discovery, *Trends Pharmacol. Sci. 22*, 132-140.
5. Höller, C., Freissmuth, M., and Nanoff, C. (1999) G proteins as drug targets, *Cell. Mol. Life Sci. 55*, 257-270.
6. Nürnberg, B., Tögel, W., Krause, G., Storm, R., Breitweg-Lehmann, E., and Schunack, W. (1999) Non-peptide G-protein activators as promising tools in cell biology and potential drug leads, *Eur. J. Med. Chem. 34*, 5-30.
7. Radhika, V., and Dhanasekaran, N. (2001) Transforming G proteins, *Oncogene 20*, 1607-1614.
8. Spiegel, A. M., and Weinstein, L. S. (2004) Inherited diseases involving G proteins and G protein-coupled receptors, *Annu. Rev. Med. 55*, 27-39.
9. Dower, W. J., and Mattheakis, L. C. (2002) *In vitro* selection as a powerful tool for the applied evolution of proteins and peptides, *Curr. Opin. Chem. Biol. 6*, 390-398.

10. Lin, H., and Cornish, V. W. (2002) Screening and selection methods for large-scale analysis of protein function, *Angew. Chem. Int. Edit. 41*, 4402-4425.

11. 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.

12. Tsao, K.-L., DeBarbieri, B., Michel, H., and Waugh, D. S. (1996) A versatile plasmid expression vector for the production of biotinylated proteins by site-specific, enzymatic modification in *Escherichia coli*, *Gene 169*, 59-64.

13. Neer, E. J., Denker, B. M., Thomas, T. C., and Schmidt, C. J. (1994) Analysis of G-protein α and βγ subunits by *in vitro* translation, *Methods Enzymol. 237*, 226-239.

14. Garcia-Higuera, I., Thomas, T. C., Yi, F., and Neer, E. J. (1996) Intersubunit surfaces in G protein αβγ heterotrimers. Analysis by cross-linking and mutagenesis of βγ, *J. Biol. Chem. 271*, 528-535.

15. Mende, U., Schmidt, C. J., Yi, F., Spring, D. J., and Neer, E. J. (1995) The G protein γ subunit. Requirements for dimerization with β subunits, *J. Biol. Chem. 270*, 15892-15898.

16. Schmidt, C. J., and Neer, E. J. (1991) *In vitro* synthesis of G protein βγ dimers, *J. Biol. Chem. 266*, 4538-4544.

17. Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., III, Graber, S. G., and Lanier, S. M. (2000) Stabilization of the GDP-bound conformation of Giα by a peptide

derived from the G-protein regulatory motif of AGS3, *J. Biol. Chem. 275*, 33193-33196.

18. Peterson, Y. K., Hazard, S., III, Graber, S. G., and Lanier, S. M. (2002) Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins, *J. Biol. Chem. 277*, 6767-6770.

19. Bernard, M. L., Peterson, Y. K., Chung, P., Jourdan, J., and Lanier, S. M. (2001) Selective interaction of AGS3 with G-proteins and the influence of AGS3 on the activation state of G-proteins, *J. Biol. Chem. 276*, 1585-1593.

20. Natochin, M., Gasimov, K. G., and Artemyev, N. O. (2001) Inhibition of GDP/GTP exchange on Gα subunits by proteins containing G-protein regulatory motifs, *Biochemistry 40*, 5322-5328.

21. Natochin, M., Lester, B., Peterson, Y. K., Bernard, M. L., Lanier, S. M., and Artemyev, N. O. (2000) AGS3 inhibits GDP dissociation from Gα subunits of the G_i family and rhodopsin-dependent activation of transducin, *J. Biol. Chem. 275*, 40981-40985.

22. De Vries, L., Fischer, T., Tronchère, H., Brothers, G. M., Strockbine, B., Siderovski, D. P., and Farquhar, M. G. (2000) Activator of G protein signaling 3 is a guanine dissociation inhibitor for $G\alpha_i$ subunits, *Proc. Natl. Acad. Sci. U.S.A. 97*, 14364-14369.

23. Berman, D. M., Kozasa, T., and Gilman, A. G. (1996) The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis, *J. Biol. Chem. 271*, 27209-27212.

24. Tesmer, J. J. G., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997) Structure of RGS4 bound to AlF_4^--activated $G_{i\alpha1}$: Stabilization of the transition state for GTP hydrolysis, *Cell 89*, 251-261.

25. 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 Gα subunits, *Nature 416*, 878-881.

26. Berman, D. M., Wilkie, T. M., and Gilman, A. G. (1996) GAIP and RGS4 are GTPase-activating proteins for the G_i subfamily of G protein α subunits, *Cell 86*, 445-452.

27. Srinivasa, S. P., Watson, N., Overton, M. C., and Blumer, K. J. (1998) Mechanism of RGS4, a GTPase-activating protein for G protein α subunits, *J. Biol. Chem. 273*, 1529-1533.

28. Barbacid, M. (1987) *ras* genes, *Annu. Rev. Biochem. 56*, 779-827.

29. Bos, J. L. (1989) *ras* oncogenes in human cancer: a review, *Cancer Res. 49*, 4682-4689.

30. Malumbres, M., and Barbacid, M. (2003) *RAS* oncogenes: the first 30 years, *Nat. Rev. Cancer 3*, 459-465.

31. Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences, *Nucleic Acids Res. 18*, 6097-6100.

32. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator, *Genome Res. 14*, 1188-1190.

Tables

Table I. Sequences of peptides selected against $G_{i\alpha 1}$-GDP.[a]

	Clone	Peptide sequence
Library	115.1	**XXXXXX *DQLYWWEYL* XXXXXX** GGYTSLSA
Round 3	R3-01	**MINTGD DELYWWQFL AELPVL** GGYTSLSA
	R3-02	**ASVHFT DKLHWWEFL EMSRDI** GGYTSLSA
	R3-03	**LEISGL DQVYWWEFL NELLSE** GGYTSLSA
	R3-04	**RLEMAS DKIYWWEYL AELASV** GGYTSLSA
Round 4	R4-01	**RDNMNR DELYWWEFL LEAVSE** GGYTSLSA
	R4-02	**ITIGAD DQLYWWEFL SDFHPQ** GGYTSLSA
	R4-03	**KEMWMD DQLYWWEFV LDTPLL** GGYTSLSA
	R4-04	**KRCNLT DELYWWEYL QSPHVA** GGYTSLSA
	R4-05	**NDWEST HRLYWWEFL EGMSTS DGYTSLSA**
	R4-07	**MMDSSN DQIYWWEFL DSWPLK** GGYTSLSA
	R4-08	**HTKLGN AKLSLEEFL LWLNDS** GGYTSLSA
Round 5	R5-01	**MHWHNT YQLSWWEFL DELDYN** GGYTSLSA
	R5-02	**DKENWH DQLYWWEFL ADYTNG** GGYTSLSA
	R5-03	**EESSLM DLMHWWEFL SELDCA** GGYTSLSA
	R5-04	**GSLNQW DRLYWWEFL ALCDSA** GGYTSLSA
	R5-05	**IESRLQ DLVYWWEAL LPTDSG** GGYTSLSA
	R5-06	**KGVSKR DQMTWWEFL SSPTGE** GGYTSLSA
	R5-07	**MLNCDN DKIYWWEYL REAPEA** GGYTSLSA
Round 6	R6-01	**KTNFWT AELNLCEFL CELDEL** GGYTSLSA
	R6-02	**HGLSMR DKLYWWEFL LDSTPN** GGYTSLSA
	R6-03	**TKCSLN DRVYWWEFL QCNSQK** CGYTSLSA
	R6-04	**TMNSLC DQLFWWEFL AQTSNL** DGYTSLSA
	R6-05	**KKPHER ESCCGRTGC RPCRSS AVTLVCL**
	R6-06	**LLTDLA AQLYWWEFL DMESGS** DGYTSLSA
	R6-07	**MENFWM DQLYWWEFI MELHDL** GGYTSLSA
	R6-08	**RTCNPD DLIYWWEYL SCPSCE** GGYTSLSA

[a] Sequences are in bold except for the C-terminal constant region. The 9-mer core (in italics) of the 115.1 core motif library was encoded by the DNA template to be conserved approximately 50% "wild-type" at each position. Clones R4-05, R6-03, R6-04, and R6-06 contained point mutations in the constant region, while clone R6-05 had a deletion resulting in a frame-shift. Except for the C-terminal constant sequence, residues are colored by amino acid type (red, positively charged; blue, negatively charged; black, polar; gray, non-polar). The methionine start codon is not shown.

Table II. Sequences of peptides selected against $G_{i\alpha 1}$-GDP-AlF_4^-.[a]

	Clone	Peptide sequence
Library	115.1	XXXXXX *DQLYWWEYL* XXXXXX GGYTSLSA
Round 6	R6-02	GSASDT DLMYWWEFL REPNRG GGYTSLSA
	R6-04	TKLRMT DNLGWGFLI LPSQF GGYTSLSA
	R6-05	DESDPE ELMYWWEFL SEDPSS GGYTSLSA
	R6-06	AHAKNL DLLTWWEFL SETNST GGYTSLSA
	R6-07	KLGNES DLLYWWEFL DQNEDD GGYTSLSA
	R6-09	KRHKLT DQLYWWEFL RDSYDD GGYTSLSA
	R6-11	EMRNQN ALLYWWEYL DELARS DGYTSLSA
	R6-12	MTSWLD DQLYWWEYL DECSRA GGYTSLSA
	R6-13	NMDRLN DLLYWWEFL EDEAPH GGYTSLSA
	R6-14	ITTMDD ELLYWWEYL DSLPQL GGYTSLSA
	R6-17	RKTHLS DLVYWWEFL AEDEDD GGYTSLSA
	R6-18	YWVDRY DERSGVCLG RQKNR GGYTSLSA
	R6-19	KLNFTN DELDWWESL MLALTT SGYTSLSA
	R6-20	YMDDND DLVYWWEFL LEPFPS GGYTSLSA
	R6-21	ALRLDV EPRNGWGFV LNPYNL GGYTSLSA
	R6-22	SDEYLD EKLYWWDFL SQMNDL GGYTSLSA
	R6-23	HKMMGS DLIYWWEFL DEINNE GGYTSLSA

[a] Sequences are shown as in Table I. Clones R6-04 and R6-18 had 3 bp deletions, while clones R6-11 and R6-19 contained point mutations in the constant region. All sequences were from the 6th round of selection.

Figures

Figure 1. Binding analysis with *in vitro* translated G proteins. (A) The indicated G_{α} subunits or H-Ras were directly translated from human cDNA vectors in a coupled transcription/translation reaction with ^{35}S-methionine labeling. A blank reaction (−) did not contain vector. The slightly lower molecular weight bands (seen clearly for $G_{i\alpha1}$ and $G_{i\alpha3}$) correspond to translation initiation at alternate methionine codons. (B) Pull-down of radiolabeled $G_{i\alpha1}$ on full-length R6A-MBP or L19 GPR peptide. R6A-MBP (+) was immobilized on cyanogen bromide-activated sepharose, while the negative control (−) contained MBP only. The L19 GPR peptide was immobilized on streptavidin-agarose, using the matrix without peptide as a negative control (−).

Figure 2. Binding of various *in vitro* translated G proteins to (A) L19 GPR, (B) R6A-1, and (C) R6A-MBP. Binding is shown relative to the estimated protein translation efficiency, as determined by TCA precipitation (see methods). The negative control matrices used in the assay were (A and B) streptavidin-agarose and (C) immobilized MBP-sepharose. $G_{12\alpha}$ consistently exhibited high non-specific binding which was especially noticeable on the MBP-sepharose in (C).

Figure 3. Binding of various *in vitro* translated G proteins to (A) R6A-1 and (B) R6A-MBP in the presence and absence of aluminum fluoride. R6A-MBP was highly specific to the GDP state while the minimal peptide, R6A-1, was GDP state specific only for G_{α} i1–3, oA, q, and s(s). Binding of $G_{12\alpha}$ to R6A-MBP in the presence of aluminum fluoride

was approximately equal to the non-specific binding seen previously (compare with Figure 2C).

Figure 4. R6A-MBP and L19 GPR compete with $G_{\beta\gamma}$ for binding to G_{α} subunits. (A) *In vitro* translated $G_{\beta 1}$ and HA-tagged $G_{\gamma 2}$ subunits. (B) Reconstituted $G_{\alpha\beta\gamma}$ heterotrimers (with G_{α} i1–3 or q) were pulled down with an anti-HA antibody, R6A-MBP, or L19 GPR. Only the G_{α} and $G_{\gamma 2}$ subunits were radiolabeled. Immunoprecipitation with anti-HA confirmed the presence of reconstituted heterotrimers in the reaction mix. G_{γ} was not co-precipitated when G_{α} subunits were pulled down by R6A-MBP or L19 GPR. Results were similar for $G_{12\alpha}$ (data not shown). (C) Binding of G_{α}-GDP to $G_{\beta\gamma}$ and R6A appear to be exclusive events.

Figure 5. R6A and GPR peptides compete with $G_{\beta\gamma}$ for binding to $G_{i\alpha}$ subunits. (A) Binding of radiolabeled $G_{\beta\gamma}$ to immobilized $G_{i\alpha 1-3}$ in the presence and absence of aluminum fluoride. The negative control (−) represents binding of $G_{\beta\gamma}$ to the matrix without immobilized G_{α}. IC_{50} values for peptide competition with $G_{\beta\gamma}$ heterodimers were determined for (B) $G_{i\alpha 1}$ and (C) $G_{i\alpha 3}$ by preincubating immobilized G_{α} with increasing concentrations of peptide prior to the binding assay. For the competition studies, binding limits have been scaled relative to the amount of $G_{\beta\gamma}$ bound in the absence of peptide (± aluminum fluoride for the lower and upper bounds, respectively). Sigmoidal fits were performed with Origin 6.0 Professional (OriginLab Corp., Northampton, MA) with the lower bound fixed at zero.

Figure 6. Selection of the 115.1 core motif library against $G_{i\alpha1}$ in the (A-C) GDP- and (D-F) GDP-AlF_4^--bound states. (A and D) Binding of RNA-peptide fusions from each round of selection. In rounds 4, 5, and 6 of the $G_{i\alpha1}$-GDP selection, stringency was increased by performing the binding at higher temperature or by adding free $G_{i\alpha1}$ as a competitor. (B and E) Sequence logo (*31*) representations of all sequences recovered from the selections (Tables I and II), generated using WebLogo (*32*) at http://weblogo.berkeley.edu. Only the X_6-R6A-1-X_6 region is shown. Residues are colored according to amino acid type: black, polar [CHNQSTWY]; gray, non-polar [AFGILMPV]; blue, negatively charged [DE]; and red, positively charged [KR]. (C and F) Percentage of amino acid types at each position using all sequences recovered from the selection. Color-coding is the same as in B and E.

Figure 7. Binding of individual peptide clones from the C-GPR X6 (*11*) and 115.1 core motif library selections to $G_{i\alpha1}$-GDP or $G_{i\alpha1}$-GDP-AlF_4^-. Purified, RNase-treated RNA-peptide fusions of each clone were assayed for binding against immobilized $G_{i\alpha1}$ in the GDP- or GDP-AlF_4^--bound states. Data are plotted as the fraction of input peptides bound to $G_{i\alpha1}$-GDP (Y-axis) versus $G_{i\alpha1}$-GDP-AlF_4^- (X-axis). While only one of the peptides from the core motif library selection against $G_{i\alpha1}$-GDP-AlF_4^- actually favored the GDP-AlF_4^--state (1.0% versus 0.6% binding for the aluminum fluoride and GDP states, respectively), the binding data from this selection clearly indicate a loss of nucleotide-state preference (toward the black line, which represents 1:1 binding to both states), compared with the other two selections that targeted $G_{i\alpha1}$-GDP. Non-specific

binding to the matrix was generally less than 0.1%. Binding data is tabulated in Supplemental Table I (see Supporting Information).

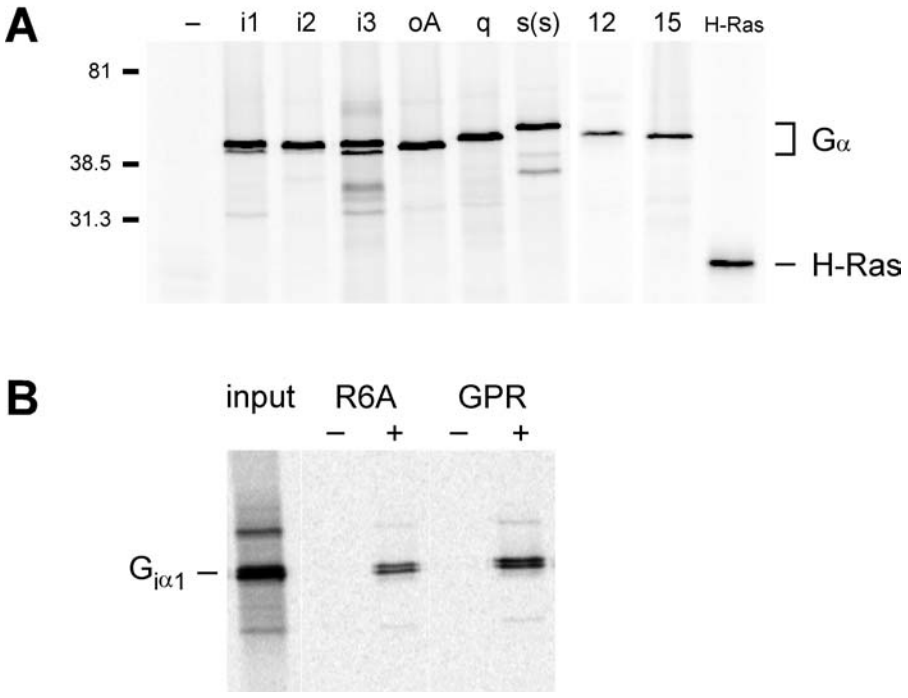

A
–
i1
i2
i3
oA
q
s(s)
12
15
H-Ras
81
38.5
31.3
Gα
H-Ras
B
input
R6A
GPR
–
+
–
+
Gi α1

Figure 1

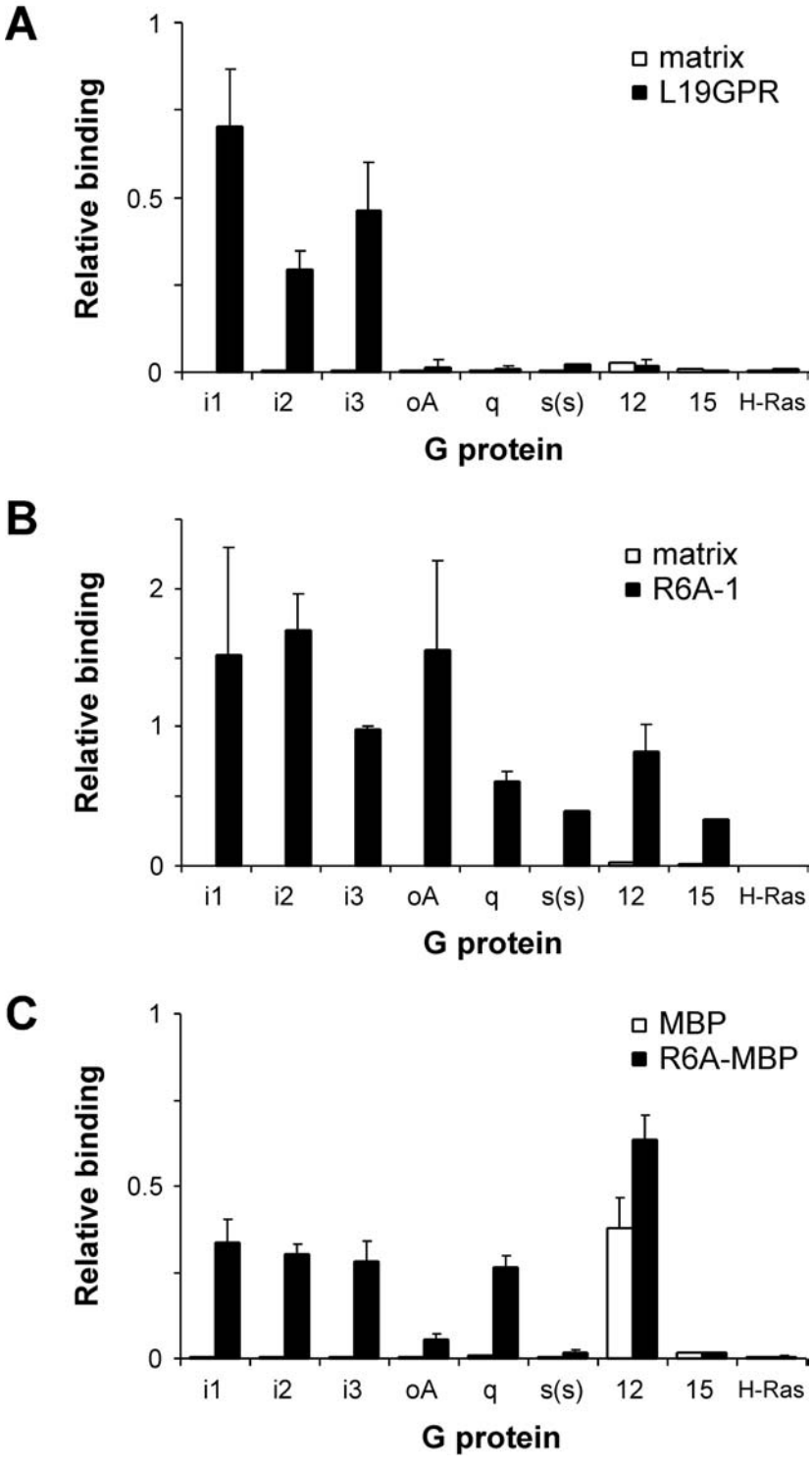

A
Relative binding
1
0.5
0
matrix
L19GPR
i1
i2
i3
oA
q
s(s)
12
15
H-Ras
G protein
B
Relative binding
2
1
0
matrix
R6A-1
i1
i2
i3
oA
q
s(s)
12
15
H-Ras
G protein
C
Relative binding
1
0.5
0
MBP
R6A-MBP
i1
i2
i3
oA
q
s(s)
12
15
H-Ras
G protein

Figure 2

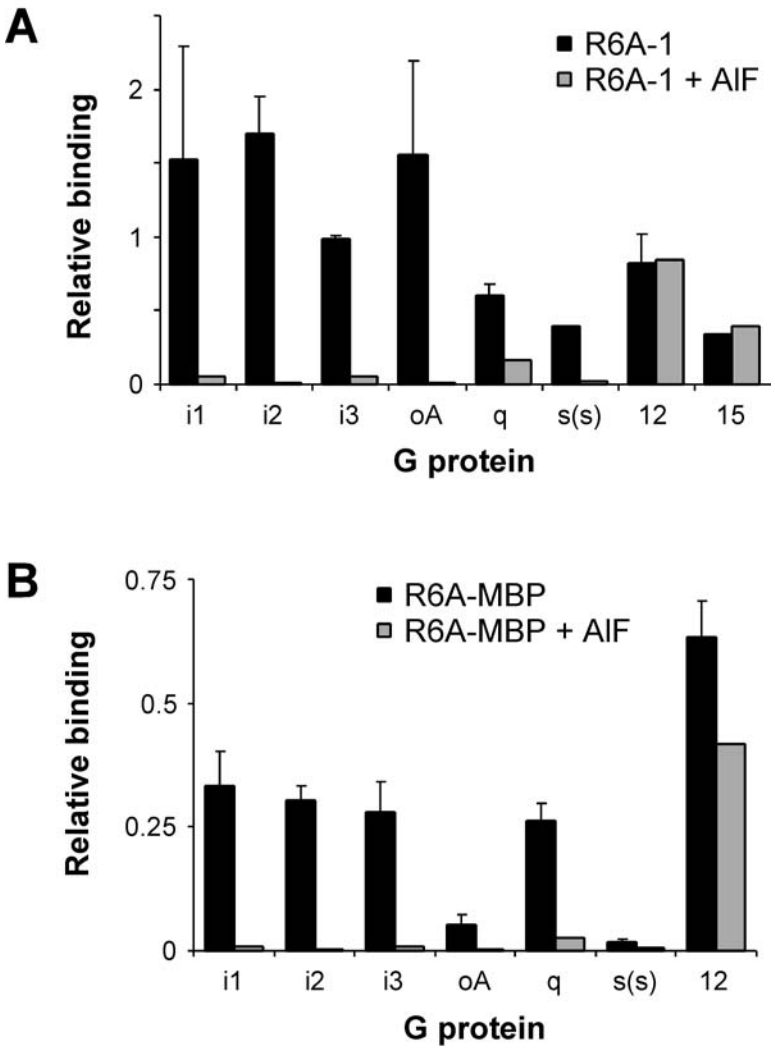

A
R6A-1
R6A-1 + AlF
Relative binding
2
1
0
i1
i2
i3
oA
q
s(s)
12
15
G protein
B
R6A-MBP
R6A-MBP + AlF
Relative binding
0.75
0.5
0.25
0
i1
i2
i3
oA
q
s(s)
12
G protein

Figure 3

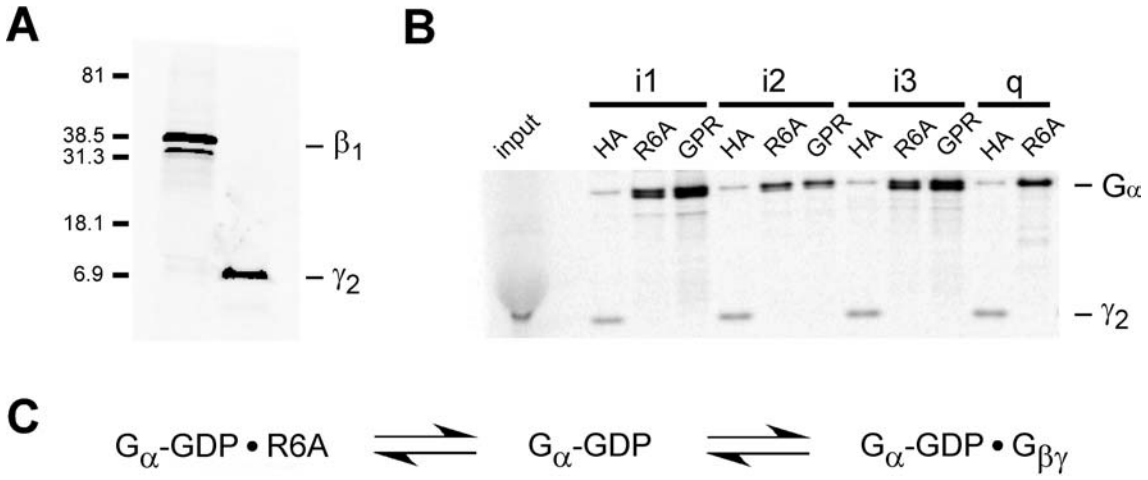

A
81
38.5
31.3
18.1
6.9
– β1
– γ2
B
input
i1
i2
i3
q
HA
R6A
GPR
– Gα
– γ2
C
Gα-GDP • R6A
Gα-GDP
Gα-GDP • Gβγ

Figure 4

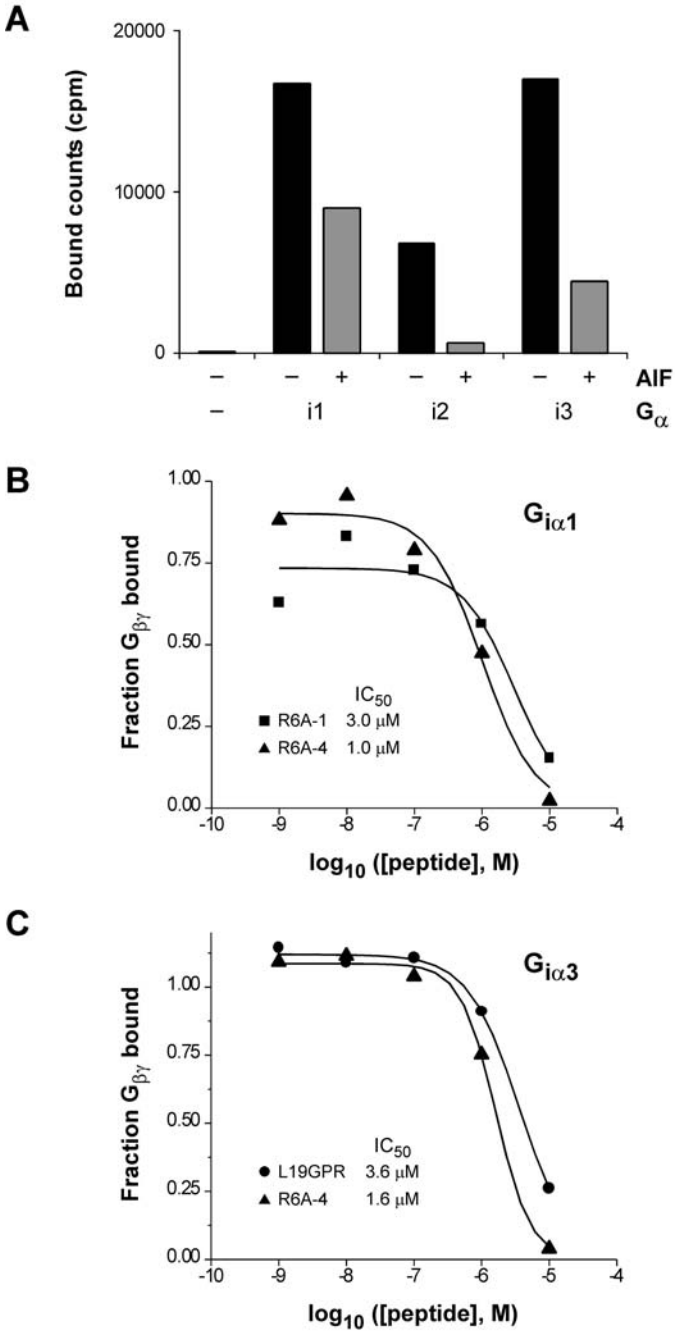

A
Bound counts (cpm)
20000
10000
0
– – + – + – + AlF
– i1 i2 i3 Gα
B
Fraction Gβγ bound
1.00
0.75
0.50
0.25
0.00
Giα1
IC50
R6A-1 3.0 μM
R6A-4 1.0 μM
-10 -9 -8 -7 -6 -5 -4
log10 ([peptide], M)
C
Fraction Gβγ bound
Giα3
IC50
L19GPR 3.6 μM
R6A-4 1.6 μM
log10 ([peptide], M)

Figure 5

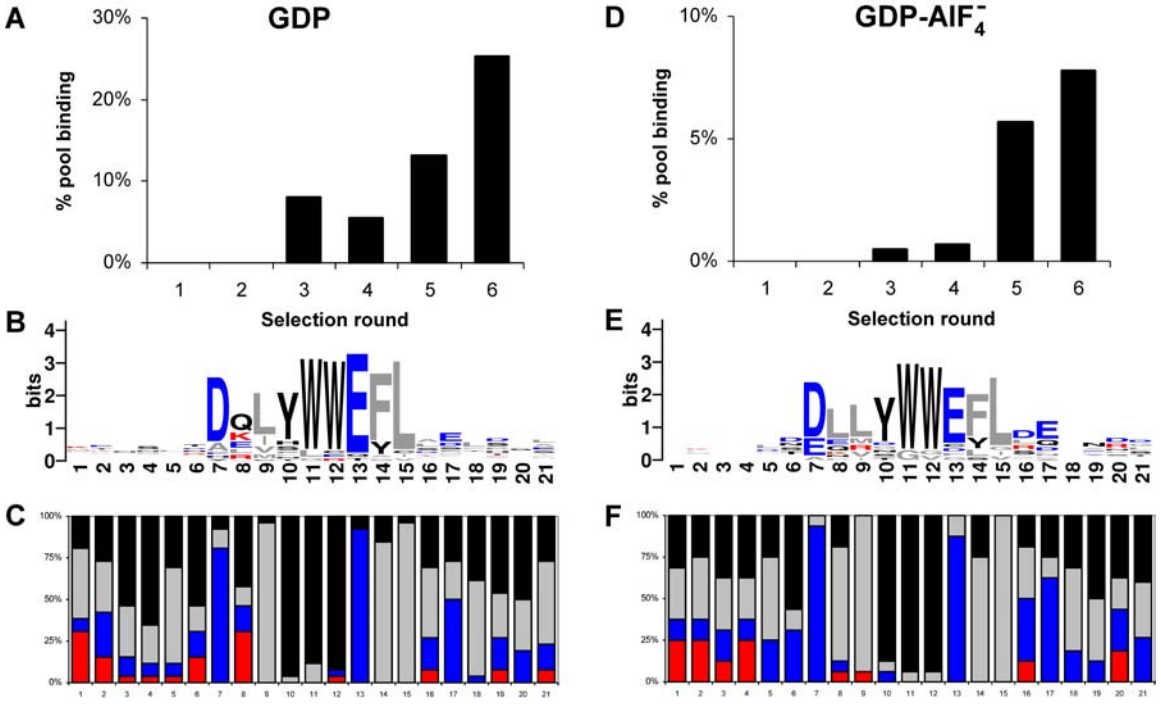

A
GDP
% pool binding
30%
20%
10%
0%
1 2 3 4 5 6
D
GDP-AlF4-
% pool binding
10%
5%
0%
1 2 3 4 5 6
B
Selection round
bits
E
Selection round
bits
C
F

Figure 6

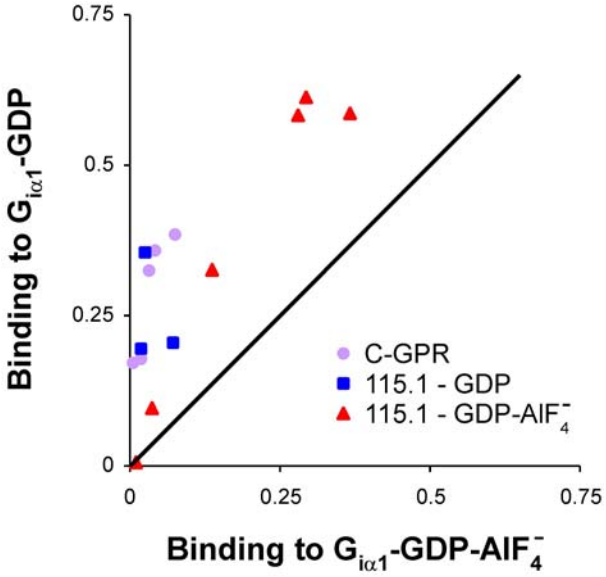

0.75
0.5
0.25
0
0
0.25
0.5
0.75
Binding to $G_{i\alpha1}$-GDP
Binding to $G_{i\alpha1}$-GDP-AlF_4^-
C-GPR
115.1 - GDP
115.1 - GDP-AlF_4^-

Figure 7

Supporting Information

Supplemental Table I. Tabulated binding data for individual peptide clones assayed in Figure 7.

Selection	Peptide [a]	% Binding	
		$G_{i\alpha 1}$-GDP	$G_{i\alpha 1}$-GDP-AlF_4^-
C-GPR X6 GDP	R6A	32.5%	3.4%
	CR7-A	17.0%	0.6%
	CR7-H	17.7%	2.0%
	CR8-B	38.3%	7.7%
	CR8-C	35.6%	4.2%
115.1 GDP	R4-01	20.2%	7.3%
	R4-04	35.3%	2.6%
	R4-05	19.4%	2.0%
115.1 GDP-AlF_4^-	R6-02	61.2%	29.4%
	R6-04	0.6%	1.0%
	R6-05	58.7%	36.6%
	R6-06	32.5%	13.8%
	R6-07	58.3%	28.1%
	R6-08	9.5%	3.7%

[a] Selected peptides from the C-GPR X6 library selection (*11*) are shown in in Supplemental Table II. Peptide sequences from the 115.1 core motif library selections are shown in Tables I and II.

Supplemental Table II. Sequences of peptides from the C-GPR X6 library selection (*11*).

Peptide	Sequence
R6A	MSQTKRLDDQLYWWEYL
CR7-A	MSQSKRLDDQLTWLEFL
CR7-H	MSQSKQLTITEFLQWL
CR8-B	MSQSERLDDQWTWWEFL
CR8-C	MSQSKRLEITWWEFVEQL

Chapter 4

Peptide ligands for Methuselah, a *Drosophila* G protein-coupled receptor associated with extended lifespan

William W. Ja, Anthony P. West, Jr., Silvia L. Delker, Pamela J. Bjorkman, Seymour Benzer, and Richard W. Roberts

Methuselah (Mth) is a G protein-coupled receptor (GPCR) associated with extended lifespan and stress resistance in *Drosophila melanogaster*. Eight rounds of *in vitro* selection have been performed using mRNA display of a random 27-mer peptide library against the extracellular domain of Mth. Isolated peptide sequences reveal the consensus, (R/P)xxWxxR (RWR motif). Synthesized RWR motif peptides (R8-01, -12, and -14) exhibit high affinity for the Mth ectodomain (K_D = 15 to 30 nM), as determined by surface plasmon resonance. The recently identified Mth agonist, Stunted, as well as a selected non-RWR motif peptide (R8-04), both competed with R8-01 for binding to the ectodomain, indicating that this site is a "hot spot" for interaction. A low-resolution crystal structure of the Mth:R8-01 complex suggests that the peptide binds at an interface between the ectodomain and extracellular loops. Despite being selected against the immobilized ectodomain, peptides tagged with rhodamine recognize and label cells expressing full-length Mth. RWR motif peptides act as antagonists to Stunted-induced Mth signaling in a cell-based GPCR calcium mobilization assay. Furthermore, characterization of mutant variants of the selected peptides has revealed a novel, potent agonist for Mth.

Introduction

G protein-coupled receptors (GPCRs)[1] mediate cell signaling from a diverse array of extracellular ligands (e.g., light, hormones, neuropeptides, odorants, and other small molecules) to intracellular signal transduction proteins (*1*). GPCRs are defined by a seven α-helical transmembrane domain, an extracellular N-terminus, which assists in ligand recognition, and an intracellular C-terminus, which is involved in G protein coupling and signal propagation. Because of their involvement in numerous cell processes, GPCRs are the target of approximately 50% of marketed drugs and new GPCR ligands continue to be pursued and developed (*2, 3*).

Commonly used, naïve approaches toward GPCR ligand identification involve high-throughput screening of a molecular library (10^2 to 10^5 unique members) in functional, cell-based assays (*3-5*). *In vitro* selection is an alternative approach for the rapid isolation of new ligands against biological targets (*6, 7*). Large, diverse libraries of unique molecules, composed of polypeptides, nucleic acids, or small molecules, are allowed to bind to a target of interest. Non-binding members are removed and functional molecules are recovered and identified. mRNA display, where each peptide in a library is covalently linked to its encoding RNA sequence, is a selection technique that allows access to very high library complexities ($>10^{13}$) in a robust format. An mRNA display library was recently used to isolate peptides that modulate G protein signaling by affecting intracellular, heterotrimeric G proteins (*8*).

---

[1] Abbreviations: CHO, Chinese hamster ovary; FACS, fluorescence-activated cell sorting; Fmoc, Fluorenylmethoxycarbonyl; GPCR, G protein-coupled receptor; HEK, human embryonic kidney; HPLC, high performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time-of- flight; Mth, Methuselah; SPR, surface plasmon resonance; TFA, trifluoroacetic acid.

Selection techniques have had limited success targeting GPCRs because of the difficulty in expression, solubilization, and presentation of the target (e.g., by immobilization on agarose). Several groups have performed successful selections with phage display libraries based on known ligands against cells expressing the GPCR of interest (*9-13*). Naïve phage display libraries, displaying either random peptides or antibody-based scaffolds, have also been selected against GPCRs expressed in cells (*14-17*), often producing ligands that act as antagonists with a range of affinities and/or IC_{50} values (nM to high μM). These results are tempered by several published failures where selected peptides were not receptor-specific, despite best efforts to target the expressed GPCR on whole cells (*18, 19*). Selections against live cells remain complicated by the low level of expression of the targeted GPCR in a large background of other cell surface proteins.

The GPCR, Methuselah (Mth), was previously determined to play a role in lifespan in the fruit fly, *Drosophila melanogaster* (*20*). The *mth* mutant exhibits a 35% increase in lifespan and enhanced resistance to various stresses including dietary paraquat, high temperature, and starvation. How Mth actually affects lifespan is unknown, although the *mth* mutant appears defective in synaptic transmission, suggesting a relation between the nervous system with stress resistance and longevity (*21*). As lower expression levels of Mth are associated with its lifespan and stress resistance phenotypes (*20, 21*), ligands for the GPCR, both agonists and antagonists, would potentially be useful as modulators of *Drosophila* aging and as tools for studying GPCR function. The structure of the large, 195-residue extracellular domain was previously determined by crystallography (*22*) and

proposed to contain the natural ligand binding site by analogy to other hormone GPCRs whose isolated ectodomains function in ligand binding (*23-26*).

Here, we have used the Mth ectodomain as a target for *in vitro* selection using mRNA display of a random peptide library. The selected peptides describe a consensus sequence, (R/P)xxWxxR, dubbed the RWR motif. Synthetic peptides bind with high affinity to the ectodomain and recognize full-length Mth expressed in cells. Cell-based GPCR signaling assays were performed using the recently identified Mth agonist, Stunted (*27*), distinguishing several of the selected peptides as strong antagonists. Competition assays suggest that the peptides bind to an interaction "hot spot" (*28*) on the ectodomain, shared by the Stunted binding site. This site appears to be located at an interface between the ectodomain and extracellular loops of Mth, based on a low-resolution crystal structure of a Mth:peptide complex. Additionally, studies of mutant variants of one of the RWR motif peptides has identified a potent Mth agonist that has little homology to Stunted.

Experimental Procedures

Materials

L-[^{35}S]-methionine (1175 Ci/mmol) was purchased from Perkin-Elmer Life Sciences, Inc. (Boston, MA). Other reagents were purchased from Sigma-Aldrich Corp. (St. Louis, MO) or VWR International, Inc. (Boston, MA), unless otherwise specified. Oligos were synthesized at the California Institute of Technology Biopolymer Synthesis and Analysis Facility, except for the 142.1 library template, which was synthesized at the W.M. Keck Facility (Yale University, New Haven, CT). HEK 293 cells expressing Mth-B (HEK-

Mth cells) were generously provided by Prof. Xin-Yun Huang (Cornell University Weill Medical College, New York, NY). DEPC-treated ddH_2O was used for all RNA work.

Expression and immobilization of Methuselah ectodomain

A hexahistidine-tagged Mth ectodomain construct (*22*) was appended with a C-terminal biotinylation tag (*29*) and subcloned into the baculovirus transfer vector pVL1392 (BD Biosciences Pharmingen, San Diego, CA). Mth was purified from supernatants of baculovirus-infected High 5 cells as described previously (*22*). Enzymatic biotinylation with biotin holoenzyme synthetase (BirA) resulted in specific biotinylation at a single lysine residue in the C-terminal tag. Desalting of the biotinylated Mth and immobilization on streptavidin-agarose (Immobilized NeutrAvidin on Agarose, Pierce Biotechnology, Inc., Rockford, IL) produced the target matrix for *in vitro* selection (~1 mg Mth/mL matrix).

mRNA display library preparation

PCR of the 142.1 template (5'-TTA AAT AGC GGA TGC ACG CAG ACC GCC ACT AGT $(SNN)_{27}$ CAT TGT AAT TGT AAA TAG TAA TTG TCC C; N = A, C, G, or T; S = C or G) with the primers 47T7FP (5'-GGA TTC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT AC) and 21.2 (5'-TTA AAT AGC GGA TGC ACG CAG) produced the initial DNA pool [0.1 μM initial template, 5 total cycles of PCR amplification using Herculase DNA polymerase (Stratagene, La Jolla, CA)]. This library encoded a T7 promoter for transcription, a 5'-UTR sequence, and an ORF for the peptide, M-X_{27}-TSGGLRASAI. After phenol extraction and ethanol precipitation of the DNA, an *in vitro* transcription reaction (80 mM HEPES-KOH at pH 7.5, 2 mM

spermidine, 40 mM DTT, 25 mM $MgCl_2$, 4 mM each of ATP, CTP, GTP, and UTP, and ~6 μg/mL dsDNA template) was treated with RNA*secure* (Ambion, Inc., Austin, TX). T7 RNA polymerase was added and the reaction was incubated at 37 °C for >2 h (*30*). Reactions were quenched with 0.1 volume of 0.5 M EDTA, followed by DNase I treatment (Epicentre, Madison, WI). The mRNA was phenol-extracted with Phase-Lock Gel (Brinkmann Instruments, Inc., Westbury, NY) and isopropanol precipitated prior to gel purification via urea-PAGE. RNA was eluted from gel slices by crushing the acrylamide gel pieces and incubating them in water at 4 °C overnight. Eluted mRNA was filtered and desalted by isopropanol precipitation. Resuspended mRNA was re-treated with DNase I (Invitrogen Corp., Carlsbad, CA), phenol-extracted, and ispropanol-precipitated to ensure complete removal of template DNA.

Cross-linking of a puromycin-psoralen linker to the mRNA was performed essentially as described (*31*). Briefly, mRNA was annealed to oligo 28A.1 (5'-[Ps]-UAG CGG AUG C-dA_{16}-$[S9]_2$-dCdC-[Pu]; where unlabeled bases are 2'-OMe RNA, Ps = psoralen C6, S9 = spacer phosphoramidite 9, and Pu = puromycin-CPG, Glen Research Corp., Sterling, VA) by incubating the mixture (~12.4 μM mRNA, 17.75 μM 28A.1, 12.5 mM HEPES-KOH at pH 7.5, and 250 mM NaCl) at 95 °C for 3 min and slow cooling to 4 °C over 10 min. The reaction was then irradiated under UV light (UVL-56 365 nm lamp, UVP, Inc., Upland, CA) in an open petri dish for ~45 min. RNA-28A.1 was purified by urea-PAGE as described above. Cross-linking reactions could be irradiated in sealed Eppendorf tubes for ~20 min with comparable yields (data not shown).

In vitro translation with the mRNA-28A.1 library was performed in Red Nova Lysate (Novagen) as per the manufacturer's instructions with optimized conditions (100 mM

KOAc, 0.5 mM MgOAc, 0.4 μM mRNA-28A.1, and ~25 μM overall L-methionine; 10 mL total reaction volume) and ^{35}S-methionine (0.5 mCi/mL final). Following the 30 °C incubation, KOAc and $MgCl_2$ were added to 585 mM and 50 mM (final), respectively, and the reaction was incubated on ice for 15 minutes to facilitate RNA-peptide fusion formation (*32*). RNA-peptide fusions were purified by diluting the reaction mixture ~150-fold into 1× isolation buffer [50 mM HEPES-KOH at pH 7.5, 1 M NaCl, 1 mM EDTA, and 0.05% Tween 20 (Bio-Rad Laboratories, Hercules, CA)] and incubating with oligo-dT cellulose (40 mg/mL of translation, New England Biolabs, Inc., Beverly, MA) at 4 °C for 1 h. Oligo-dT cellulose was pelleted by centrifugation (1500 × g for 1 min) and washed extensively with 0.4× isolation buffer in either glass wool-packed disposable columns (Poly-Prep, Bio-Rad) or in 0.45 μm cellulose acetate centrifuge tube filters (Costar Spin-X, Corning, Inc., Corning, NY). RNA-peptide fusions were eluted with a warmed solution of 1 mM β-mercaptoethanol (65 °C), desalted by isopropanol precipitation using linear acrylamide (Ambion) as a carrier, and reverse-transcribed (Superscript II, Invitrogen) with oligo 21.2. The reverse transcription reaction was supplemented with ~0.05% Tween 20 to ensure that the precipitated RNA-peptide fusions were solubilized. Reverse-transcribed fusions were desalted and exchanged into Mth buffer [50 mM HEPES-KOH at pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.1% (w/v) BSA, 1 μg/mL yeast tRNA (Roche Diagnostics Corp., Indianapolis, IN) and 0.05% Tween 20] using a NAP-25 column (Amersham Biosciences Corp., Piscataway, NJ) prior to the selection step.

In vitro *selection*

RNA:cDNA-peptide fusions were incubated with ~0.1 mL of Mth-NeutrAvidin-agarose at 4 °C for 1 h, then filtered and washed on a Poly-Prep column with 4 ×1 mL Mth buffer followed by 2 × 1 mL Mth buffer without BSA/tRNA. Bound fusions were eluted with 2 × 100 μL of 0.15% SDS at room temperature through a 0.45 μm spin filter. After removal of the SDS using SDS-OUT (Pierce) as per the manufacturer's instructions, fusions were isopropanol-precipitated (50 μg/mL linear acrylamide, 1/40 volume of 3 M NaOAc at pH 5.2, and 1 volume of isopropanol). The reduced salt used for isopropanol precipitation was necessary to prevent inhibition of subsequent PCR, due to the high salt introduced by the SDS-OUT reagent. Precipitated cDNA was PCR-amplified to produce the dsDNA pool for the next round of selection or for cloning and subsequent DNA sequencing (TOPO TA cloning for sequencing kit, Invitrogen).

Further rounds of selection were performed as described for the initial round except that *in vitro* translation reactions were smaller (~0.3 mL), less immobilized Mth was used for the selective step (~20 μL), washes were performed in batch, and in rounds 5 through 8, bound fusions were eluted by competition with non-biotinylated Mth (0.5 mg/mL) in Mth buffer without BSA/tRNA. Additionally, rounds 5 through 8 included a pre-clearing step where the precipitated RNA:cDNA-peptide fusions were passed through multiple columns containing NeutrAvidin-agarose and/or protein G-sepharose (Sigma) to remove peptides with high non-specific binding for the immobilization matrix. The flow-through was then used for selection as described above.

In vitro *binding assays of mRNA display libraries*

To assess the binding activity of mRNA display pools from each round of selection, ^{35}S-methionine-labeled RNA-peptide fusions were prepared and purified as described above. Fusions were treated with RNase (DNase-free, Roche) and mixed with ~10 μL of immobilized Mth in Mth buffer. After rotating at 4 °C for 1 h, the matrix was washed with 3 × 0.6 mL of Mth buffer in a 0.45 μm spin filter and bound 28A.1-puromycin-peptide fusions were quantitated by scintillation counting.

Individual clones were also tested in this format by PCR amplification of specific clones (e.g., R8-01 and R8-04) with oligos 47T7FP and 21.2. Competition studies were performed by introducing various concentrations of unlabeled peptides into the binding buffer during the initial binding step.

Peptide synthesis

Peptides were synthesized on a 432A Synergy peptide synthesizer (Applied Biosystems, Foster City, CA) using standard Fmoc chemistry. Synthesized peptides were cleaved and deprotected from the resin by agitation in TFA/1,2-ethanediol/thioanisole (90:5:5) for 2 h at room temperature. After desalting by precipitation in methyltertbutyl ether, crude peptides were purified by reversed-phase HPLC (C18, 250 × 10 mm, Grace Vydac, Hesperia, CA) to >95% purity on an aqueous acetonitrile/0.1% (v/v) TFA gradient. Peptide masses were confirmed by MALDI-TOF mass spectrometry and concentrations were determined by absorbance at 280 nm using a calculated extinction coefficient (http://paris.chem.yale.edu/extinct.html).

Several peptides were synthesized with an N-terminal spacer (ethylene glycol, Bachem California, Inc., Torrance, CA) for conjugation to various functional groups (e.g., biotin, fluorescein, or rhodamine). Prior to deprotection of the peptide and cleavage from the resin, the free N-terminal amine could be coupled to NHS-conjugated compounds (Pierce). Briefly, the peptide resin was incubated with agitation at room temperature for 2 to 4 hours in DMF containing 2- to 5-fold molar excess of the NHS-conjugate. Reactions were quenched with 0.5 M ethanolamine and the resin was washed thoroughly with DMF, DMSO, followed by methanol through a Poly-Prep column on a vacuum manifold. The resin was dried on vacuum before deprotecting and purifying the peptide as described above.

Peptide truncation and mutagenesis series were synthesized with C-terminal glycines by JPT Peptide Technologies GmbH (MicroScale Peptide Sets, Berlin, Germany). Crude, dried peptides were provided in 96-well plates at approximately 50 nmol of full-length peptide per well. Peptides were reconstituted in 20 μL of DMSO (~2.5 mM stock concentration) prior to their use in cell signaling assays.

Kinetics determination by surface plasmon resonance (SPR)

SPR measurements were performed at 25 °C on a BIAcore 2000 instrument (Biacore, Inc., Piscataway, NJ) equipped with research-grade SA (streptavidin) sensor chips. Biotinylated Mth was immobilized to a surface density of 450 to 700 response units (RU). HBS-EP [10 mM HEPES at pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005% polysorbate 20 (Tween 20), Biacore] was used as the running buffer for all experiments. To collect kinetics data, a concentration series of each peptide was injected for at least 60

s at a flow rate of >45 μL/min. Peptides were allowed to freely dissociate to background between injections (~5 min). Buffer blank injections were used for double referencing with a negative control surface (*33*). Raw data was processed with Scrubber and globally fit with CLAMP using a 1:1 bimolecular interaction model (*34*). K_D values were calculated (k_d/k_a) from the determined rate constants.

Spectrofluorometric analysis of Mth:peptide complexes

Measurements of intrinsic tryptophan and tyrosine fluorescence spectra were performed on a spectrofluorophotometer (RF-5301PC, Shimadzu Scientific Instruments, Columbia, MD) at excitation wavelengths of 280 or 295 nm (slit width set at 3 nm). Spectra were taken at room temperature from 260 to 450 nm at 0.2 nm intervals (medium speed, 3 or 5 nm slit widths for R8-12 or R8-04 peptides, respectively). Complexes were formed by mixing equimolar concentrations of peptide and Mth ectodomain in buffer (10 mM KH_2PO_4 at pH 7.2 and 150 mM NaCl) for 5 min before taking fluorescence readings in a stirred cell. For R8-04, 3 μM peptide and wild-type Mth ectodomain were used. The mutant ectodomain, Mth W120S, was used for complexes with R8-12 (both at 6.25 μM concentrations) because the peptide contains a single Trp residue. For the fluorescence titration study, 300 nM of R8-12 peptide was used and aliquots of Mth W120S were directly added to a stirring cell (excitation and emission wavelengths were set at 295 and 345 nm, with slit widths at 5 and 10 nm, respectively).

Crystallography – data collection, structure solution, and refinement

Purified His_6-tagged Mth ectodomain (15 mg/mL) was mixed with 1.5-fold molar excess R8-01 15-mer peptide in 10 mM Tris at pH 8.0. Crystals of the Mth:R8-01 complex were

grown by hanging drop, in which the protein mixture was mixed 1:1 with well solution [0.1 M HEPES-KOH at pH 7.5, 1.7 M ammonium sulfate, and 2% (w/v) PEG 400]. The resulting 2 μL drop was suspended over 0.8 mL of well solution. Crystals were grown for several months at room temperature and formed long rods with maximum dimensions of 200 μm × 50 μm × 50 μm.

Crystals were transferred to a cryoprotectant solution [(0.1 M HEPES-KOH at pH 7.5, 1.8 M ammonium sulfate, 2% PEG 400, and 20% (v/v) glycerol] prior to flash cooling. Data were collected at −150 °C at the Advanced Light Source beamline 9.2.2 with λ = 1.0781 Å. Data were processed and scaled with DENZO and SCALEPACK (*35*). The diffraction was strongly anisotropic, extending to ~2.5 Å along the 4-fold axis but only to ~3.8 Å perpendicular to the 4-fold with strong streaking in these directions. The crystals were initially indexed in space group $P4_2$, with a = 94.25 Å and c = 173.84 Å. R_{merge} from 20–3.50 Å (3.62–3.50 Å) was 15.5% (37.0%).

The structure was determined by molecular replacement using AMoRe (*36*) from version 5 of the CCP4 suite (*37*) with the 2.3 Å structure of the Mth ectodomain (PDB code 1FJR) as a search model. Examination of the native Patterson map revealed three very strong non-origin peaks implying translational non-crystallographic symmetry (NCS). A translational NCS vector was used during the translation function search. Two molecules were placed in the asymmetric unit, which explained two of the three peaks in the native Patterson. Addition of a third molecule, necessary to give the third native Patterson peak, always resulted in steric clashes. The data was eventually explained by a certain type of stacking disorder in the crystals, which was consistent with the observed symmetry of the Mth packing. This disorder is the likely cause of the strong anisotropy

of the data and of the streaking of the diffraction spots. We modeled the disorder with one Mth molecule with unit occupancy and two with half occupancy. Maps were calculated with solvent flattening, histogram matching, and NCS averaging using the program DM in the CCP4 suite (*37*). Anisotropy and bulk solvent corrections were applied and the model was refined with NCS constraints using grouped temperature (*B*) factors using the program CNS (*38*). The peptide has been left unmodeled. The current model has R_{cryst} = 37.5% and R_{free} = 40.8%.

Fluorescence labeling of cells expressing Methuselah

Chinese hamster ovary (CHO) cells were transiently transfected with a Mth-GFP fusion construct and grown in alpha-MEM with 5% fetal bovine serum (FBS). Cells were detached by trypsination and washed with P4F (1× PBS with 4% FBS). Cells were then incubated with 150 nM rhodamine-R8-12 peptide in P4F for 1 hour at 4 °C, washed extensively with P4F, and either mounted on glass slides for fluorescence microscopy or sorted by FACS (BD FACSCalibur System, BD Biosciences Immunocytometry Systems, San Jose, CA).

Cell-based GPCR signaling assay

HEK 293 cells expressing Mth-B were plated in clear-bottom 96-well plates (Corning) at ~25% confluency using Matrigel (BD Biosciences). Supernatant was removed by careful aspiration and each well was washed with 200 μL of buffer F [20 mM HEPES-KOH at pH 7.5, 0.1% BSA, and 2.5 mM probenicid (dissolved first in 1/100 volume of 1 M NaOH) in Hanks' balanced salt solution without phenol red (HBSS, Invitrogen)]. Fluo-4 (4 μM, dissolved first in 20 μL of 1:1 DMSO:10% Pluronic F-127, Molecular Probes,

Eugene, OR) in buffer F with 1% BSA was added to each well (100 μL) and the plate was incubated at room temperature for 45 min, followed by incubation at 37 °C for 15 min. Wells were washed with 3 × 200 μL of buffer F while the plate was on ice. For agonist assays, 100 μL of buffer F was added to each well and the plate was incubated at 37 °C for 15 min prior to starting the experiment. Fluorescence measurements and automated sample delivery were performed in Flexstation mode on a robotic plate reader (Flexstation, Molecular Devices, Sunnyvale, CA). Agonist peptides in reagent buffer (20 mM HEPES-KOH at pH 7.5 and 5% DMSO in HBSS) were added (50 μL) after an initial baseline reading (30 to 45 s). Continuous readings were made for ~5 min (7 reads/well, 2 s intervals, 494 nm excitation wavelength, 520 nm emission, and 515 nm cutoff filter). For antagonist assays, 80 μL of buffer F was added to each well prior to the 15 min 37 °C incubation. Potential antagonists were added (20 μL) after the baseline readings followed by addition of the agonist peptide (50 μL). Data analysis and background subtraction were performed with Softmax Pro 4.7.1 (Molecular Devices). Sigmoidal fits were calculated using Origin 6.0 Professional (OriginLab Corp., Northampton, MA).

Results

In vitro *selection of ligands for the Methuselah ectodomain*

The mature, N-terminal ectodomain of Mth is a stably folded, glycosylated protein of 195 residues (*22*). By analogy to other GPCRs with large, extracellular N-terminal domains that maintain recognition of their cognate ligands independently of their transmembrane cores (*23-26*), we targeted the Mth ectodomain for *in vitro* selection to isolate putative modulators of Mth signaling. A random, 27-mer peptide library was constructed for

selection using mRNA display. Based on quantitation of the purified, ^{35}S-labeled RNA-peptide fusions and the estimated L-methionine concentration, the complexity of the starting library was approximately 1.5×10^{13}. Because the hexahistidine tag is a weak epitope for peptide selections (see Appendix A), a specifically biotinylated construct of the Mth ectodomain was expressed and purified using a C-terminal peptide substrate sequence recognized by BirA (*29*). Four rounds of selection were initially performed and, although binding to the Mth ectodomain was above background levels, a high degree of non-specific binding to the matrix was observed (Figure 1A). Four subsequent rounds of selection were performed which included preclearing steps on matrix without target and specific elution with free, non-biotinylated Mth. The final 8th round pool exhibited negligible non-specific binding and high activity for Mth.

DNA sequencing of individual clones from the final selection round (Table I) revealed a highly conserved consensus motif, (R/P)xxWxxR, dubbed the RWR motif (Figure 1B). Interestingly, 6 out of 7 sequences in the 5th round pool (Supplemental Table I) encoded at least WxxR. Amino acid analysis of the selected peptides, however, revealed a shift toward more polar residues in the 8th round pool. This suggests that the multiple rounds of preclearing and specific elution succeeded in retaining the binding motif (and hence, affinity for the Mth ectodomain) while reducing non-specific interactions.

Analysis of the amino acid types at each position in the aligned peptides revealed a degree of covariation in the RWR motif depending on whether the first residue of the motif is Arg or Pro (Figure 1C and D). Arg is generally followed by two non-polar residues, while Pro is followed by charged or polar amino acids. This trend is most

evident in the 8th round sequences, but is also observed in RWR motif peptides in the previous rounds. Residues 6 and 7 in the WxxR consensus are generally non-polar, while residue 9, immediately C-terminal to the RWR motif, is almost always polar, especially favoring Ser or Thr. There also seems to be a weak preference for Arg an indiscriminant number of residues C-terminal to the RWR motif. Phe-Arg, Ala-Arg, and especially Leu-Arg were often observed 5 to 7 residues downstream of the RWR motif (Figure 1B and Supplemental Table I).

Binding kinetics determined by surface plasmon resonance

Several selected peptides were synthesized and purified for binding analysis by SPR (Table II). Peptides were designed to include the RWR motif and an additional number of flanking residues. *In vitro* binding studies with C-terminal truncations of an outlier peptide R8-04 (a recovered peptide that lacked an RWR motif), suggested that the full-length peptide was important for high affinity binding. Hence, R8-04 was synthesized through the first 3 residues encoded by the 3' constant region. Mth-binding peptides containing the RWR motif demonstrated high affinity (K_D <30 nM) while R8-04 exhibited a K_D of 330 nM (Table III). Binding was not observed at concentrations of up to 1 μM for the W7A and R10A mutants of R8-14, demonstrating the importance of these conserved residues for Mth-binding. Weak binding was observed with 1 μM of a scrambled R8-01 peptide (R8-01 SCR). Higher concentrations, however, were not tested by SPR.

Hot spot for Methuselah interaction

Sequential binding of peptides R8-01, R8-04, and R8-12 was also analyzed by SPR and strongly suggested that the selected peptides share the same binding site (data not shown). This was demonstrated more clearly in competition binding studies (Figure 2A). Synthetic, unlabeled peptide R8-01 competed for binding to immobilized Mth ectodomain with both radiolabeled, full-length R8-01 and R8-04. The weak affinity of R8-01 SCR was confirmed as it also competed for binding to Mth. Interestingly, a synthetic variant of the recently isolated peptide agonist for Mth, Stunted (*27*), also competed for binding, demonstrating that the natural ligand binding site is an interaction "hot spot" (*28*) and at least partially reconstituted by the Mth ectodomain.

The Mth ectodomain contains a single solvent-exposed Trp residue in a shallow groove that was proposed to be a potential binding site for the natural ligand (*22*). Mth W120S, where the Trp was mutated to a Ser, was constructed to assay the effect of this mutation on binding for the R8-12 peptide, which contains a single Trp at position 5. Mth W120S expresses similarly to the wild-type ectodomain and CD spectra of the two proteins were nearly identical, suggesting that the mutant folds stably (data not shown). Excitation at 295 nm, which is specific for tryptophan residues, revealed a significant fluorescence enhancement of Trp5 in the R8-12 peptide upon binding to Mth W120S (Figure 2B). The maximum fluorescence wavelength (λ_{max}) of 351 nm for R8-12 is indicative of denatured proteins, suggesting that the peptide is unstructured in solution. The blue-shift of λ_{max} to 343 nm, as well as the increase in fluorescence, suggests that R8-12 Trp5 is at least partially buried upon binding to Mth W120S and protected from solvent quenching. The fluorescence contributions from Tyr residues (λ_{max} = ~302 nm)

could be observed by excitation at 280 nm, which revealed a negligible change in signal upon peptide binding (data not shown). The K_D of R8-12 binding to Mth W120S was also determined by fluorescence titration and calculated to be 15 nM, which compares well with results from the SPR studies (Figure 2B, inset).

Fluorescence spectra were also measured for wild-type Mth ectodomain in complex with peptide R8-04, which does not contain a Trp residue. The fluorescence spectra were nearly identical, however, between Mth and the Mth:R8-04 complex (Figure 2C). The concentration of each component (3 μM) was ~10-fold greater than the K_D determined by SPR, suggesting that the lack of a fluorescence change was probably not due to inadequate complex formation. If the R8-04 peptide were interacting near Trp120, a much greater effect on Trp fluorescence would be expected upon binding. The resulting spectra, along with the R8-12 data, suggest that the Mth interdomain groove containing Trp120 is not the shared binding site for the selected peptides and the N-Stunted agonist.

To determine the precise location of the peptide binding site, the structure of the Mth ectodomain in complex with an RWR motif peptide was determined by crystallography. Although the resulting data was low-resolution (3.5–4 Å), electron density which putatively corresponds to the R8-01 15-mer peptide places the binding site near the C-terminus of the ectodomain. This suggests that the peptides bind at or near an interface between the Mth ectodomain and extracellular loops (Figure 3). Interestingly, the 2nd extracellular loop (EL2) between the 4th and 5th transmembrane helices contains a WxxR peptide sequence (a partial RWR motif) which may interact in the same region as the selected peptides.

Fluorescent peptide probes recognize the full-length GPCR

While the peptides bound with high affinity to the Mth ectodomain, it was important to determine whether they would still recognize the full-length GPCR. A full-length Mth-GFP fusion construct was expressed in a CHO cell line, stained with a rhodamine-tagged R8-12 peptide, and examined by fluorescence microscopy, revealing strong labeling of cells expressing the Mth-GFP fusion protein (Figure 4A). Staining was not observed on cells incubated with a rhodamine-labeled, scrambled version of R8-12 (data not shown). The CHO/Mth-GFP cells were also amenable to sorting by FACS, demonstrating a linear correlation between Mth-GFP expression and the degree of rhodamine staining (Figure 4B). These rhodamine-peptide probes were functional for labeling Mth-GFP over-expressed in *Drosophila* larvae and adults, but the use of wild-type and control flies did not result in any specific, punctate staining of endogenous Mth (data not shown). This may be due to a low level of endogenous Mth expression or cross-reactivity of our probes to Mth homologs.

Selected peptides act as antagonists for Methuselah signaling

Recently, Stunted, the endogenous agonist for Mth, was isolated from adult *Drosophila* extracts using a cell-based, fluorescence reporter system (*27*). HEK-Mth cells were loaded with a fluorescent calcium indicator and washed prior to the assay. Addition of a Mth agonist induced calcium flow into the cell and a subsequent increase in fluorescence. Synthetic 30-mer peptides corresponding to the N- and C-terminal halves of the ~60-mer Stunted were also tested, isolating the agonist activity to the N-terminus (N-Stunted). By pre-incubating cells with selected peptides prior to the addition of the N-Stunted agonist,

any antagonist activity from the peptides could be observed (Figure 5A). R8-12 and R8-14 were strong antagonists of N-Stunted-induced Mth signaling, with IC_{50} values of 70 and 170 nM, respectively (Figure 5B). R8-01 also acted as an antagonist, though an IC_{50} was not determined. The W7A and R10A mutants of R8-14 did not exhibit any antagonist activity. In all cases, signaling was not observed on non-transfected, control HEK 293 cells (data not shown).

Biochemical characterization of Stunted and a novel peptide agonist

Surprisingly, strong calcium signaling was observed upon addition of the R8-01 scrambled peptide to HEK-Mth cells (Figure 6A). The activity was specific to cells expressing Mth, in comparison with a non-transfected control cell line (data not shown), and suggested that R8-01 SCR is a specific Mth agonist. The R8-01 SCR peptide shares few residues in common with N-Stunted and appeared to be a more potent activator of Mth (EC_{50} of 2.5 μM, compared with 11 μM for N-Stunted, Figure 6B). The inhibition of R8-01 SCR-mediated signaling by the selected peptides, R8-12 and R8-14, provides further evidence that the scrambled peptide is indeed a specific agonist for Mth (Figure 6C).

A series of 15-mer peptides for N-Stunted and 12-mers for R8-01 SCR were synthesized to determine the minimal peptide that exhibited agonist activity. Crude peptides were tested directly in the cell-based assay for Mth activation. For N-Stunted, the region of activity was localized to the N-terminus (Figure 7A). As the first 3 peptides in the series were functional, this suggests that the minimal peptide sequence required for maximal activity is AWRAAGITYIQYS. The C-terminal peptides of N-Stunted also

exhibited some activity above background, although it is inconclusive whether a second local region of agonist activity exists, or if this is caused by something non-specific. For R8-01 SCR, the minimal peptide with high activity appears to be LQAPRRSVMRW.

Discussion

High affinity peptide ligands that act as antagonists to the recently isolated, endogenous Mth agonist, Stunted, were isolated by *in vitro* selection using a naïve, random 27-mer mRNA display library targeting the Mth ectodomain. Most of the unique sequences encoded a consensus, the RWR motif (R/P)xxWxxR. However, it appears that all the selected peptides, including R8-04 which was an "outlier" sequence lacking the RWR motif, bind to the same site on the Mth ectodomain, a putative "hot spot" for interaction.

Curiously, the final selection pool was dominated by the outlier peptide, R8-04, which also exhibited the weakest binding. The competitive elution with free, non-biotinylated Mth ectodomain in the last few rounds of selection most likely enriched the pool for the weakest binding peptides. Those peptides with the fastest off rates would be competed first. While we attempted to optimize the incubation time for the elution, our main goal was to eliminate the non-specific binding peptides. A selection against the Mth ectodomain with a "doped" peptide library, encoding the RWR motif, may be useful in isolating new sequences with even higher affinities. It is interesting that although the earlier 5^{th} round pool demonstrated moderate non-specific binding to various immobilization matrices, the clones had already encoded the RWR motif. More stringent selection conditions in the later rounds enriched RWR motif peptides with fewer hydrophobic residues and low non-specific binding.

The RWR motif is a very strong consensus, with other amino acids generally conserved. The C-terminal preference for Arg and especially certain Arg pairs were found a variable number of residues away from the RWR motif. Hence, the region immediately following the motif may not be highly structured. The placement of the RWR motif in the random region of the selected peptides appears random. Hence, the preference for a Leu-Arg sequence C-terminal to the RWR motif may not be very strong, as the constant region encoded a Leu-Arg–containing sequence.

In vitro studies suggested that R8-04 binds significantly better with the C-terminal constant region than without (data not shown). Because *in vitro* translation can start at alternate methionine codons, N-terminal truncations of R8-04 were also observed. Interestingly, N-terminal truncations were non-binding. Hence, the full-length, 38-mer R8-04 peptide (which includes the entire C-terminal constant region) may bind significantly better than the 31-mer used in our experiments and may do so in some folded structure that utilizes both termini.

The random scrambling of the R8-01 peptide was performed to ensure that the RWR motif was not retained. However, while the RxxWxxR motif was eliminated, the new peptide did encode a PxxSxxR sequence which may still have minimal affinity. Further study of the R8-01 scrambled agonist, as well as mutational studies of the other peptides, may reveal the critical residues in converting an antagonist (R8-01) to an agonist (R8-01 SCR). Not only was the R8-01 SCR a more potent agonist than Stunted, but the Stunted peptide also exhibited delayed activation of the receptor (Figure 6A). The isolation of a novel, potent agonist for Mth with little homology to Stunted suggests that other, natural endogenous Mth ligands may exist that have yet to be identified. The minimal, active

region of N-Stunted identified in our studies corresponds well to the most conserved region of Stunted homologs, the epsilon subunit of the eukaryotic F1-ATPase (*39*), further suggesting that N-Stunted is indeed one of the natural endogenous ligands for Mth.

The localization of the peptide binding site to a putative interface between the Mth ectodomain and extracellular loops suggests that the WxxR on EL2 (a partial RWR motif) may form similar interactions with the ectodomain as our peptides. An interaction between a synthetic EL2 peptide (at concentrations up to 150 μM) and the Mth ectodomain was not observed in *in vitro* competition studies (Table II and data not shown). This may not preclude an *in vivo* interaction, however, as the affinity between EL2 and the ectodomain may be extremely low, but compensated for in the full-length receptor where the EL2 sequence and the ectodomain are co-localized. Additionally, a high affinity, optimized interaction may be undesirable for natural signaling.

We have successfully isolated novel peptide ligands for a GPCR using mRNA display selection libraries targeting the extracellular domain of Mth. The application of these peptides toward the synthetic extension of *Drosophila* lifespan is currently being tested. Whether mRNA display can be generally used against full-length GPCR targets has yet to be determined. Targeting GPCRs expressed in cells remains an unattractive option due to the large background of cell surface proteins. The recent success in assembling paramagnetic proteoliposomes containing pure GPCRs is a favorable alternative for a selection target (*40*). The development of suitable expression and display platforms for transmembrane proteins will greatly facilitate the selection of peptide modulators of GPCR signaling.

Acknowledgments

We thank Dr. Anthony M. Giannetti for technical expertise on the Biacore, Prof. David G. Myszka (University of Utah) for the SPR analysis software, Scrubber and CLAMP, Prof. Melvin I. Simon for use of the Flexstation automated fluorescence plate reader, and Prof. Theodore Brummel (Sam Houston State University, TX) and Dr. David Walker for helpful discussions. We also thank Terry T. Takahashi for comments on the manuscript. We are greatly indebted to Dr. Svetlana Cvejic and Prof. Xin-Yun Huang (Cornell University Weill Medical College, NY) for providing the HEK-Mth cell lines and details on their protocols. We are especially grateful to Dr. Peter M. Snow (deceased 2004) for preparative expertise on protein purification.

This work was supported by grants from the NIH (RO160416) and the Beckman Foundation to R. W. R. W. W. J. was supported in part by a DOD National Defense Science and Engineering Graduate Fellowship and a Scholarship for Research in the Biology of Aging, sponsored by the Glenn Foundation for Medical Research and the American Federation for Aging Research. A. P. W., Jr., was supported by a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund. R. W. R. is an Alfred P. Sloan Foundation Research Fellow.

References

1. Strader, C. D., Fong, T. M., Tota, M. R., Underwood, D., and Dixon, R. A. F. (1994) Structure and function of G protein-coupled receptors, *Annu. Rev. Biochem. 63*, 101-132.

2. Drews, J. (2000) Drug discovery: a historical perspective, *Science 287*, 1960-1964.

3. Howard, A. D., McAllister, G., Feighner, S. D., Liu, Q., Nargund, R. P., Van der Ploeg, L. H. T., and Patchett, A. A. (2001) Orphan G-protein-coupled receptors and natural ligand discovery, *Trends Pharmacol. Sci. 22*, 132-140.

4. Milligan, G. (2002) Strategies to identify ligands for orphan G-protein-coupled receptors, *Biochem. Soc. Trans. 30*, 789-793.

5. Szekeres, P. G. (2002) Functional assays for identifying ligands at orphan G protein-coupled receptors, *Receptor. Channel. 8*, 297-308.

6. Dower, W. J., and Mattheakis, L. C. (2002) *In vitro* selection as a powerful tool for the applied evolution of proteins and peptides, *Curr. Opin. Chem. Biol. 6*, 390-398.

7. Lin, H., and Cornish, V. W. (2002) Screening and selection methods for large-scale analysis of protein function, *Angew. Chem. Int. Edit. 41*, 4402-4425.

8. 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.

9. Cain, S. A., Williams, D. M., Harris, V., and Monk, P. N. (2001) Selection of novel ligands from a whole-molecule randomly mutated C5a library, *Protein Eng. 14*, 189-193.

10. Heller, T., Hennecke, M., Baumann, U., Gessner, J. E., zu Vilsendorf, A. M., Baensch, M., Boulay, F., Kola, A., Klos, A., Bautsch, W., and Köhl, J. (1999) Selection of a C5a receptor antagonist from phage libraries attenuating the inflammatory response in immune complex disease and ischemia/reperfusion injury, *J. Immunol. 163*, 985-994.

11. Szardenings, M., Törnroth, S., Mutulis, F., Muceniece, R., Keinänen, K., Kuusinen, A., and Wikberg, J. E. (1997) Phage display selection on whole cells yields a peptide specific for melanocortin receptor 1, *J. Biol. Chem. 272*, 27943-27948.

12. Rousch, M., Lutgerink, J. T., Coote, J., de Bruïne, A., Arends, J.-W., and Hoogenboom, H. R. (1998) Somatostatin displayed on filamentous phage as a receptor-specific agonist, *Br. J. Pharmacol. 125*, 5-16.

13. Doorbar, J., and Winter, G. (1994) Isolation of a peptide antagonist to the thrombin receptor using phage display, *J. Mol. Biol. 244*, 361-369.

14. Goodson, R. J., Doyle, M. V., Kaufman, S. E., and Rosenberg, S. (1994) High-affinity urokinase receptor antagonists identified with bacteriophage peptide display, *Proc. Natl. Acad. Sci. U.S.A. 91*, 7129-7133.

15. Houimel, M., Loetscher, P., Baggiolini, M., and Mazzucchelli, L. (2001) Functional inhibition of CCR3-dependent responses by peptides derived from phage libraries, *Eur. J. Immunol. 31*, 3535-3545.

16. Moulard, M., Phogat, S. K., Shu, Y., Labrijn, A. F., Xiao, X., Binley, J. M., Zhang, M.-Y., Sidorov, I. A., Broder, C. C., Robinson, J., Parren, P. W. H. I., Burton, D. R., and Dimitrov, D. S. (2002) Broadly cross-reactive HIV-1-neutralizing human monoclonal Fab selected for binding to gp120-CD4-CCR5 complexes, *Proc. Natl. Acad. Sci. U.S.A. 99*, 6913-6818.

17. Osbourn, J. K., Earnshaw, J. C., Johnson, K. S., Parmentier, M., Timmermans, V., and McCafferty, J. (1998) Directed selection of MIP-1α neutralizing CCR5 antibodies from a phage display human antibody library, *Nat. Biotechnol. 16*, 778-781.

18. Hoogenboom, H. R., Lutgerink, J. T., Pelsers, M. M. A. L., Rousch, M. J. M. M., Coote, J., van Neer, N., de Bruïne, A., van Nieuwenhoven, F. A., Glatz, J. F. C., and Arends, J.-W. (1999) Selection-dominant and nonaccessible epitopes on cell-surface receptors revealed by cell-panning with a large phage antibody library, *Eur. J. Biochem. 260*, 774-784.

19. Sui, J., Bai, J., St Clair Tallarico, A., Xu, C., and Marasco, W. A. (2003) Identification of CD4 and transferrin receptor antibodies by CXCR4 antibody-guided Pathfinder selection, *Eur. J. Biochem. 270*, 4497-4506.

20. Lin, Y.-J., Seroude, L., and Benzer, S. (1998) Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*, *Science 282*, 943-946.

21. Song, W., Ranjan, R., Dawson-Scully, K., Bronk, P., Marin, L., Seroude, L., Lin, Y.-J., Nie, Z., Atwood, H. L., Benzer, S., and Zinsmaier, K. E. (2002) Presynaptic regulation of neurotransmission in *Drosophila* by the G protein-coupled receptor Methuselah, *Neuron 36*, 105-119.

22. West, A. P., Jr., Llamas, L. L., Snow, P. M., Benzer, S., and Bjorkman, P. J. (2001) Crystal structure of the ectodomain of Methuselah, a *Drosophila* G protein-coupled receptor associated with extended lifespan, *Proc. Natl. Acad. Sci. U.S.A. 98*, 3744-3749.

23. Cao, Y.-J., Gimpl, G., and Fahrenholz, F. (1995) The amino-terminal fragment of the adenylate cyclase activating polypeptide (PACAP) receptor functions as a high affinity PACAP binding domain, *Biochem. Biophys. Res. Commun. 212*, 673-680.

24. Chow, B. K. (1997) Functional antagonism of the human secretin receptor by a recombinant protein encoding the N-terminal ectodomain of the receptor, *Recept. Signal Transduct. 7*, 143-150.

25. Grauschopf, U., Lilie, H., Honold, K., Wozny, M., Reusch, D., Esswein, A., Schäfer, W., Rücknagel, K. P., and Rudolph, R. (2000) The N-terminal fragment of human parathyroid hormone receptor 1 constitutes a hormone binding domain and reveals a distinct disulfide pattern, *Biochemistry 39*, 8878-8887.

26. Wilmen, A., Göke, B., and Göke, R. (1996) The isolated N-terminal extracellular domain of the glucagon-like peptide-1 (GLP)-1 receptor has intrinsic binding activity, *FEBS Lett. 398*, 43-47.

27. Cvejic, S., Zhu, Z., Felice, S. J., Berman, Y., and Huang, X.-Y. (2004) The endogenous ligand Stunted of the GPCR Methuselah extends lifespan in *Drosophila*, *Nat. Cell Biol. 6*, 540-546.

28. Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface, *Science 267*, 383-386.

29. Schatz, P. J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia coli*, *Bio-Technol. 11*, 1138-1143.

30. Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods Enzymol. 180*, 51-62.

31. Kurz, M., Gu, K., and Lohse, P. A. (2000) Psoralen photo-crosslinked mRNA-puromycin conjugates: a novel template for the rapid and facile preparation of mRNA-protein fusions, *Nucleic Acids Res. 28*, e83.

32. Liu, R., Barrick, J. E., Szostak, J. W., and Roberts, R. W. (2000) Optimized synthesis of RNA-protein fusions for *in vitro* protein selection, *Methods Enzymol. 318*, 268-293.

33. Myszka, D. G. (2000) Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE, *Methods Enzymol. 323*, 325-340.

34. Myszka, D. G., and Morton, T. A. (1998) CLAMP (c): A biosensor kinetic data analysis program, *Trends in Biochemical Sciences 23*, 149-150.

35. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307-326.

36. Navaza, J. (1994) AMoRe: an automated package for molecular replacement, *Acta Crystallogr. A 50*, 157-163.

37. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D 50*, 760-763.

38. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R.

J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D 54*, 905-921.

39. Tu, Q., Yu, L., Zhang, P., Zhang, M., Zhang, H., Jiang, J., Chen, C., and Zhao, S. (2000) Cloning, characterization and mapping of the human *ATP5E* gene, identification of pseudogene *ATP5EP1*, and definition of the ATP5E motif, *Biochem. J. 347*, 17-21.

40. Mirzabekov, T., Kontos, H., Farzan, M., Marasco, W., and Sodroski, J. (2000) Paramagnetic proteoliposomes containing a pure, native, and oriented seven-transmembrane segment protein, CCR5, *Nat. Biotechnol. 18*, 649-654.

41. Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences, *Nucleic Acids Res. 18*, 6097-6100.

42. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator, *Genome Res. 14*, 1188-1190.

43. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor, *Science 289*, 739-745.

Tables

Table I. Peptide sequences from the 8th round of selection.[a]

Peptide	Sequence		Frequency
R8-08b	SSLSPPWPASWSPSRPSAPRAAPSTPT	*	2/20
R8-04	VRIGYTSKPGGMNPGNSYTMSIIRMLI		7/20
R8-07b	STAGSRARSTS**W**GT**R**SPWTWPTPARTG	*	
R8-01	NVSWGSF**P**SS**W**LQ**R**YYLAKRREADVTL		4/20
R8-07	LKY**P**DT**W**LA**R**SLSVFYLRKSARQGKSV		
R8-13	ELGQFQRLSL**P**YQ**W**YL**R**TISYVSLRTA		
R8-03	GDDMYRIREFLANY**R**PI**W**VM**R**SNLAQL		
R8-12	**R**LV**W**IV**R**SRHFGPRLRMALLGSDRKMW		
R8-14	AP**R**AV**W**IQ**R**AIQAMFRLASRQESKAFN		
R8-09b	**R**YV**W**YL**R**TKHRRSLRLRSACARGSSA	*	

[a] The frequency (out of 20) is shown for peptides that appeared more than once from DNA sequencing of individual clones. The methionine start codon and C-terminal constant region (TSGGLRASAI), which was frame-shifted or mutated in marked sequences (*), are not shown.

Table II. Synthesized peptide sequences.

Peptide	Sequence
R8-01	MNVSWGSFPSSWLQRYYLAKRR
R8-01 SCR	SWSLQAPRRSVMRWYGFYLNKS
R8-01 15-mer	FPSSWLQRYYLAKRR
R8-04	MVRIGYTSKPGGMNPGNSYTMSIIRMLITSG
R8-12	MRLVWIVRSRHFGPRLRMA
R8-12 SCR	MLRARRHGVWSPLRFRIMV
R8-14	MAPRAVWIQRAIQAMFRLA
R8-14 W7A [a]	MAPRAVAIQRAIQAMFRLAGY
R8-14 R10A	MAPRAVWIQAAIQAMFRLA
N-Stunted	MTAWRAAGITYIQYSNIAARILRESLKTGL
Mth EL2 [b]	DNIVENQDWNPRVGHEGH

[a] The R8-14 W7A mutant was synthesized with a C-terminal Gly-Tyr for quantitation by UV absorbance.

[b] The Mth EL2 sequence is derived from a section of the 2nd extracellular loop of Mth.

Table III. Kinetic parameters for binding of selected peptides with Mth, as determined by SPR.[a]

Peptide	$\mathbf{k_a}$	$\mathbf{k_d}$	$\mathbf{K_D}$[b]	$\boldsymbol{\chi^2}$
	$M^{-1}s^{-1}$ ($\times 10^5$)	s^{-1} ($\times 10^{-2}$)	*nM*	
R8-01	6.3	1.9	31	1.5
R8-01 15-mer	9.5	5.4	57	0.46
R8-01 SCR			WB	
R8-12	4.1	0.72	18	1.3
R8-14	7.0	1.2	18	1.7
R8-14 W7A			NB	
R8-14 R10A			NB	
R8-04[c]	6.1	19.9	326	0.46

[a] A sensorgram for one of the SPR experiments with R8-04 is shown in Supplemental Figure 1.
[b] K_D values were calculated (k_d/k_a) from kinetic parameters. Control peptides were non-binding (NB) or weakly binding (WB) at concentrations up to 1 μM.
[c] Peptide sequences are shown in Table II.

Figures

Figure 1. Selection of a 27-mer peptide library against the Mth ectodomain. (A) RNase-treated, ^{35}S-methionine-labeled mRNA displayed peptides from each round of selection were assayed for binding to immobilized Mth (black) or to matrix alone (white). Preclearing and competitive elutions were performed in the 5th through 8th rounds to eliminate non-specific binding peptides. (B) Sequence logo (*41*) representation of RWR motif peptides (13 sequences), indicating the conservation at each amino acid position. Selected peptides containing the RWR motif, (R/P)xxWxxR, were aligned manually and parsed (keeping 1 residue N-terminal and 10 residues C-terminal to the motif). The sequence logo was generated using WebLogo (*42*) at http://weblogo.berkeley.edu. Residues are colored according to amino acid type, as indicated. The same color-coding is used to illustrate the percentage of amino acid types at each position in (C) PxxWxxR (9 sequences) and (D) RxxWxxR peptides (4 sequences).

Figure 2. Analysis of peptide ligand binding. (A) *In vitro* translated, ^{35}S-methionine-labeled, full-length peptides (left, R8-01; right, R8-04) were assayed for binding to immobilized Mth in the presence or absence of various synthesized peptides. Concentrations of the peptide competitors were 30 μM for N-Stunted (N-St) and 10 μM for R8-01 and R8-01 SCR. Bound peptides are expressed in cpm (+ standard deviation, when available) and have been background subtracted with a negative control. (B) Intrinsic tryptophan fluorescence of Mth W120S ectodomain, R8-12 peptide, and their complex upon excitation at 295 nm. (B, inset) Fluorescence titration of 300 nM R8-12 peptide with Mth W120S at an emission wavelength of 345 nm. (C) Fluorescence

spectra of wild-type Mth ectodomain with and without R8-04 peptide upon excitation at 295 nm.

Figure 3. Structure of the Mth ectodomain in complex with the R8-01 15-mer peptide. (left) Ribbon diagram of the Mth ectodomain. Electron density shows the putative peptide binding site from an averaged 3.5 Å $F_O - F_C$ map contoured at 9 σ. (right) Scaled model depicting the full-length structure of Mth (adapted from (*22*)). The transmembrane region is represented by the structure of rhodopsin (*43*). The putative peptide binding site is shown, along with the sequence of the 2nd extracellular loop that contains a partial RWR motif.

Figure 4. Selected peptides recognize the full-length Mth receptor. (A) Fluorescence microscopy of CHO cells transiently transfected with a Mth-GFP fusion construct. Cells expressing Mth-GFP (left) were stained with 150 nM rhodamine-tagged R8-12 (right). (B) Unstained cells, sorted by FACS, exhibited a range of Mth-GFP expression (left). Staining with rhodamine-R8-12 (right) revealed a direct correlation between the level of Mth-GFP expression (X-axis) and the degree of staining (Y-axis). The slight tailing at high GFP levels (left panel) is due to uncompensated crosstalk between the GFP and rhodamine channels.

Figure 5. Selected peptides act as antagonists of Mth signaling. (A) The N-Stunted agonist peptide (20 μM final) was added to HEK-Mth cells pre-incubated with and without the indicated concentration of R8-12 peptide antagonist. Mth activation resulted in intracellular calcium mobilization and enhanced fluorescence. Fluorescence spectra

were divided by a baseline average, calculated from the region of data prior to the addition of N-Stunted. The dashed line indicates a control where only buffer (without agonist) was added. (B) Concentration dependence of the inhibition of Mth signaling by peptides R8-12 and R8-14. Signals were taken at a time point ~13 sec after the addition of 10 μM N-Stunted agonist and are expressed as a fraction of the fluorescence observed in the absence of antagonists.

Figure 6. The scrambled R8-01 peptide acts as a potent agonist for Mth signaling. (A) Comparison of agonist activity between R8-01 SCR and N-Stunted. Fluorescence spectra were measured as in Figure 5A. (B) Concentration dependence of signaling by Mth agonists. Higher concentrations of N-Stunted were not tested due to problems with peptide solubility. Hence, the max fluorescence determined from fitting the R8-01 SCR data was used as a fixed limit for the N-Stunted fit. (C) Concentration dependence of the inhibition of R8-01 SCR-induced Mth signaling by peptides R8-12 and R8-14. Data are reported as in Figure 5B except that the signals were taken at a time point ~7 sec after the addition of agonist.

Figure 7. Minimal peptides for Mth activation. (A) A series of 15-mer peptides derived from N-Stunted were tested for their ability to activate Mth in the cell-based calcium mobilization assay. Reported values (average of two experiments, + standard deviation) represent the maximum fluorescence achieved after the addition of peptide divided by the baseline average. The shaded region highlights the putative minimal peptide agonist. “Blank” is a negative control where buffer without peptide was added. (B) A series of

12-mer peptides were tested as in (A) for the R8-01 SCR agonist. All peptide sequences listed have an additional C-terminal glycine (not shown).

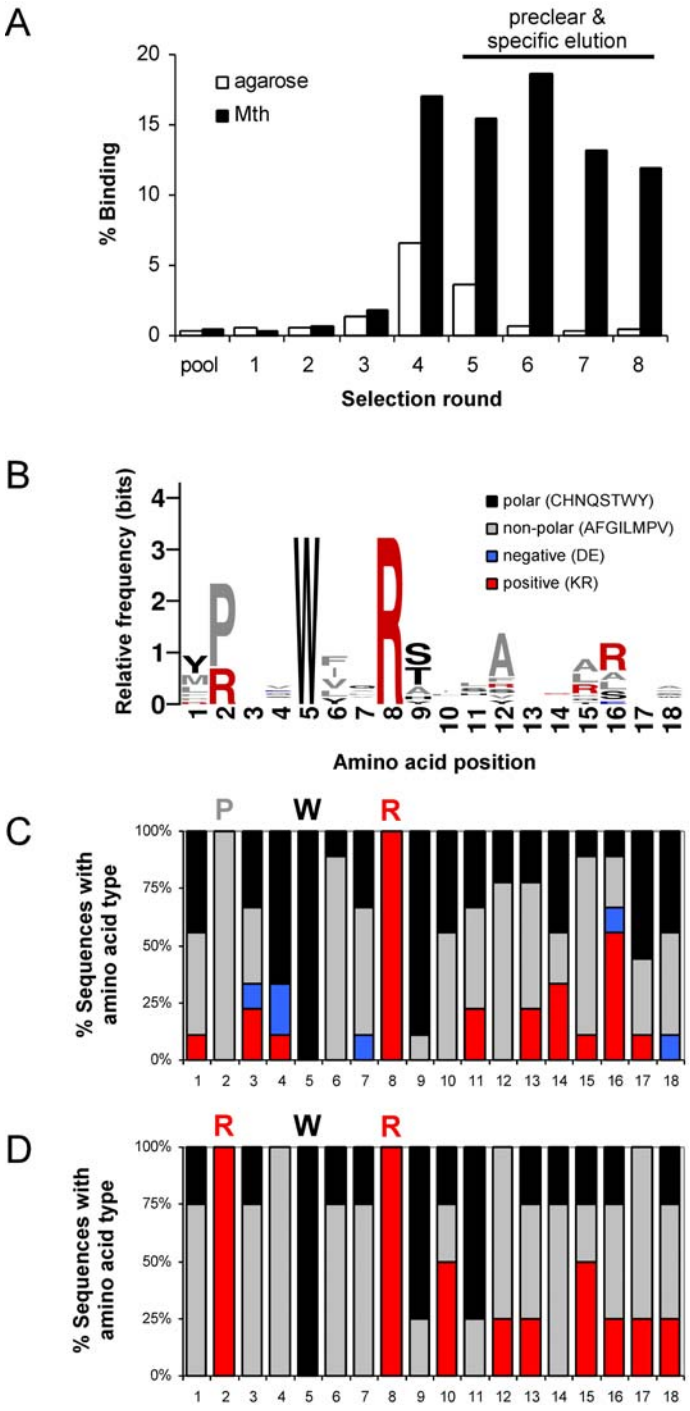

A
preclear &
specific elution
agarose
Mth
% Binding
Selection round
pool
B
Relative frequency (bits)
polar (CHNQSTWY)
non-polar (AFGILMPV)
negative (DE)
positive (KR)
Amino acid position
C
P
W
R
% Sequences with amino acid type
D
R
W
R
% Sequences with amino acid type

Figure 1

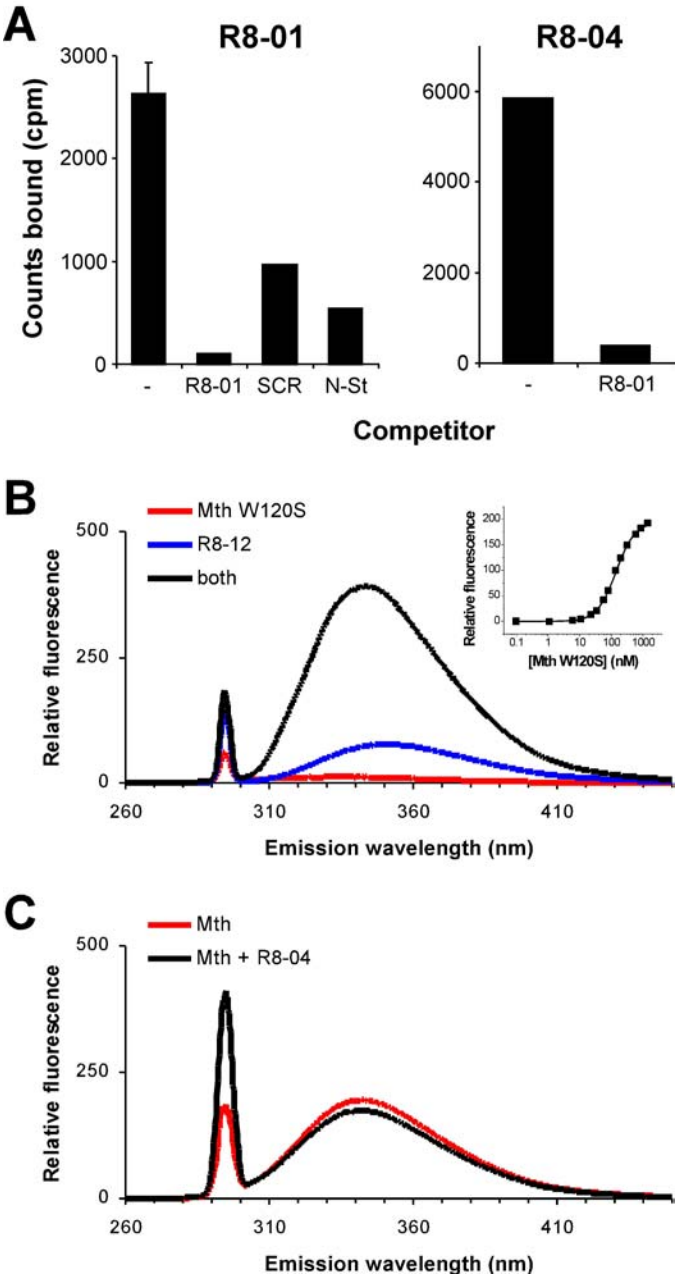

A
R8-01
R8-04
Counts bound (cpm)
3000
2000
1000
0
6000
4000
2000
0
-
R8-01
SCR
N-St
-
R8-01
Competitor
B
Mth W120S
R8-12
both
Relative fluorescence
500
250
0
260
310
360
410
Emission wavelength (nm)
200
150
100
50
0
0.1
1
10
100
1000
[Mth W120S] (nM)
C
Mth
Mth + R8-04
Relative fluorescence
500
250
0
260
310
360
410
Emission wavelength (nm)

Figure 2

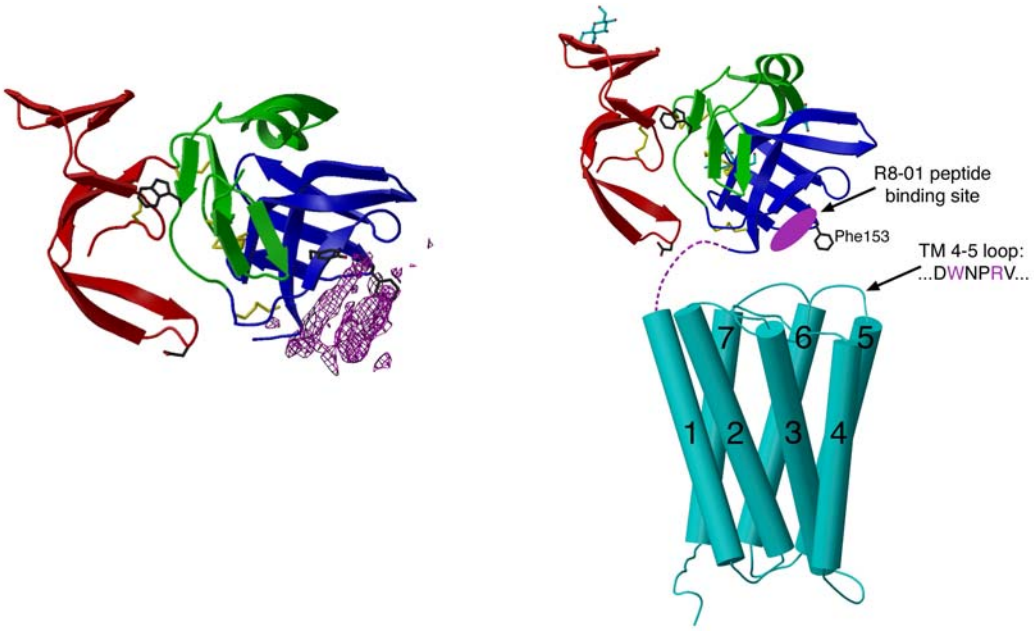


Figure 3

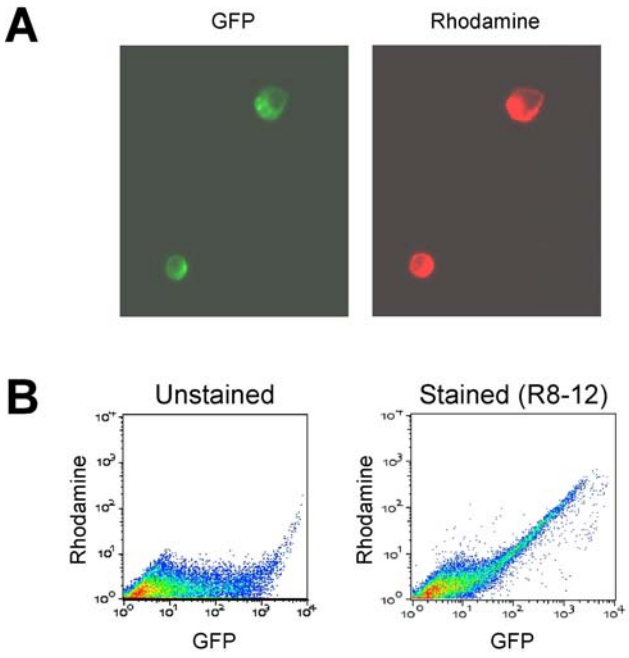

A
GFP
Rhodamine
B
Unstained
Stained (R8-12)
Rhodamine
GFP
Rhodamine
GFP

Figure 4

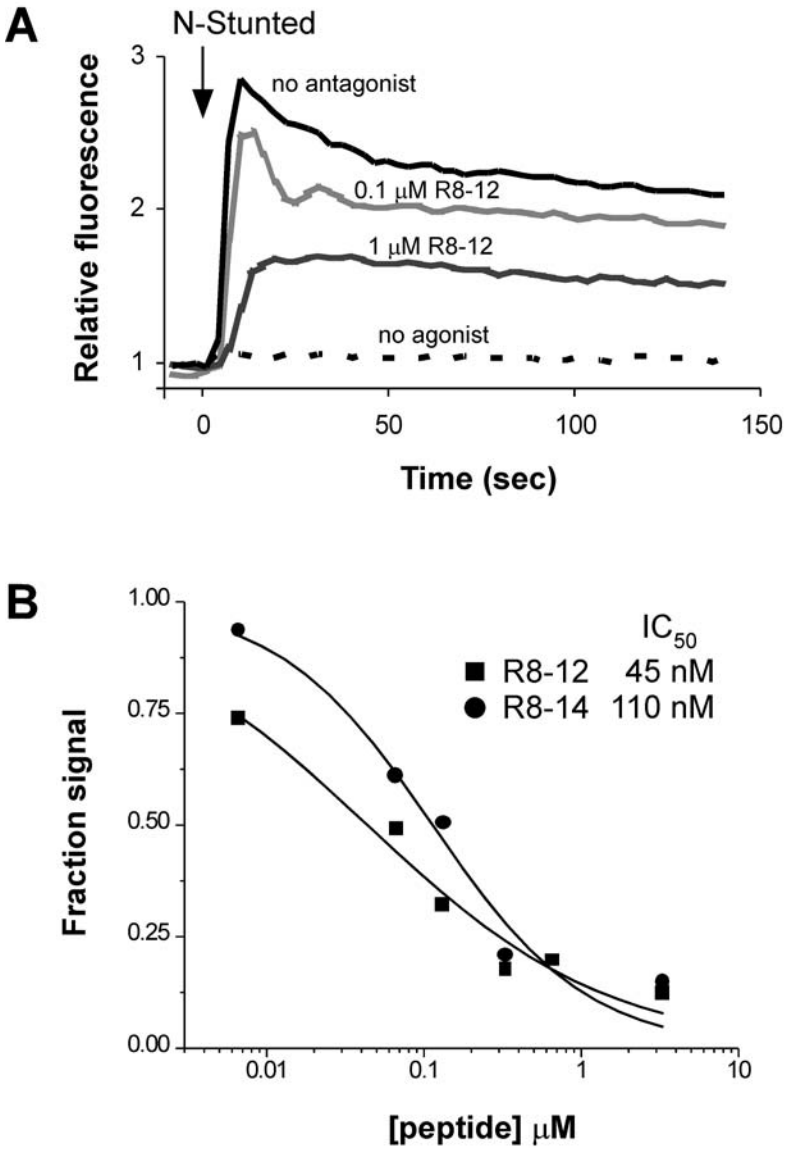

A
N-Stunted
no antagonist
0.1 μM R8-12
1 μM R8-12
no agonist
Relative fluorescence
Time (sec)
B
IC50
R8-12 45 nM
R8-14 110 nM
Fraction signal
[peptide] μM

Figure 5

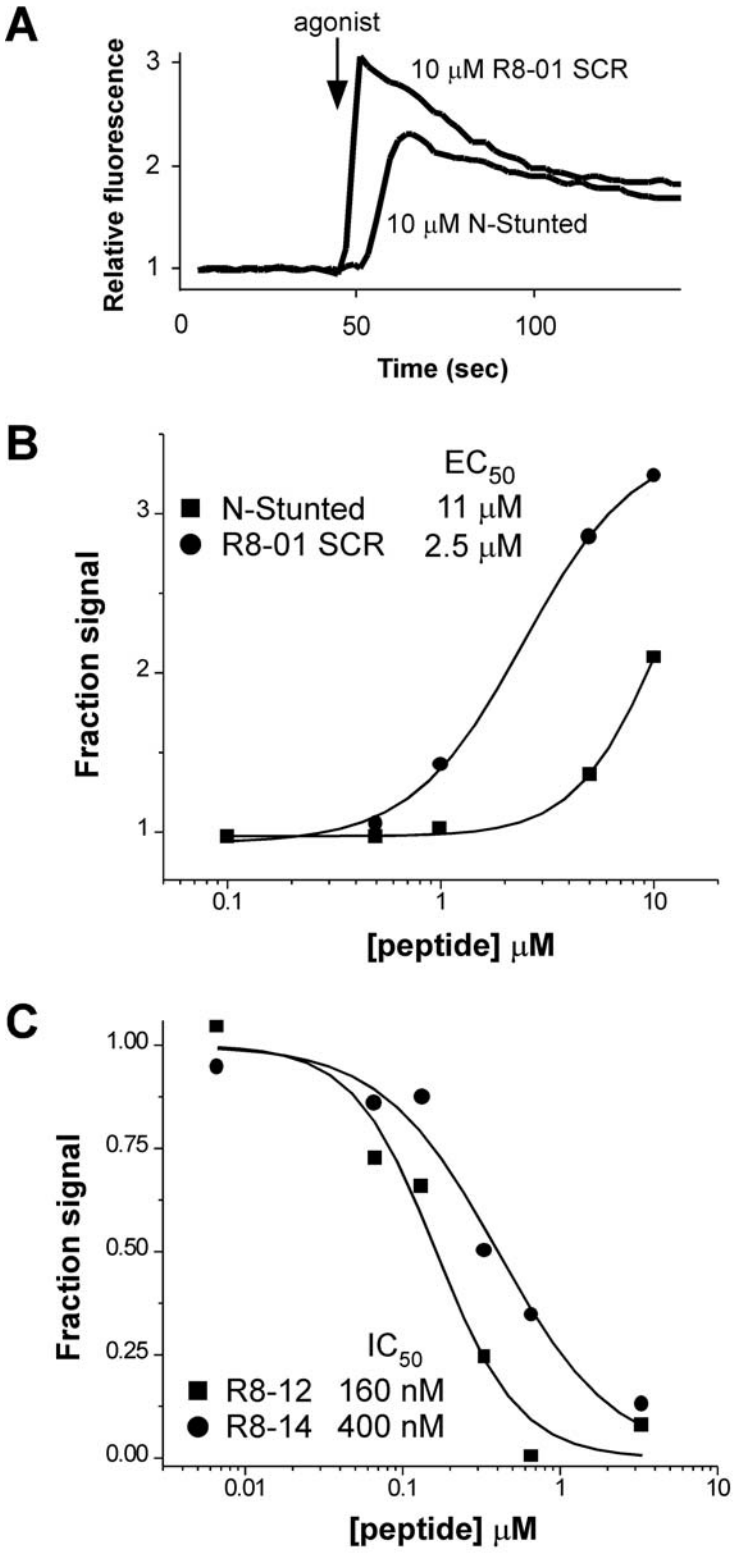

A
agonist
10 μM R8-01 SCR
10 μM N-Stunted
Relative fluorescence
3
2
1
0
50
100
Time (sec)
B
EC50
N-Stunted 11 μM
R8-01 SCR 2.5 μM
Fraction signal
3
2
1
0.1
1
10
[peptide] μM
C
Fraction signal
1.00
0.75
0.50
0.25
0.00
IC50
R8-12 160 nM
R8-14 400 nM
0.01
0.1
1
10
[peptide] μM

Figure 6

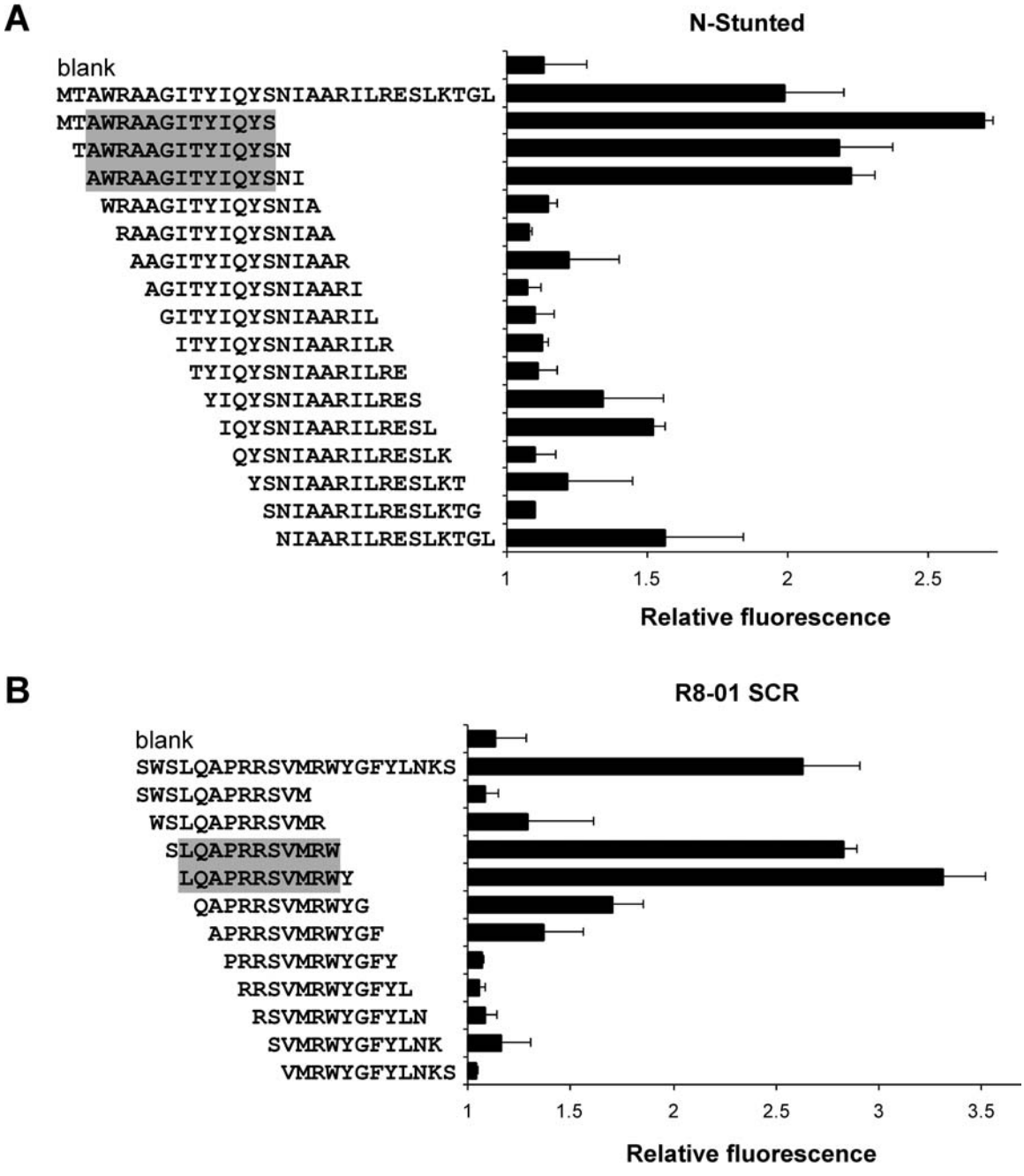

A
N-Stunted
blank
MTAWRAAGITYIQYSNIAARILRESLKTGL
MTAWRAAGITYIQYS
TAWRAAGITYIQYSN
AWRAAGITYIQYSNI
WRAAGITYIQYSNIA
RAAGITYIQYSNIAA
AAGITYIQYSNIAAR
AGITYIQYSNIAARI
GITYIQYSNIAARIL
ITYIQYSNIAARILR
TYIQYSNIAARILRE
YIQYSNIAARILRES
IQYSNIAARILRESL
QYSNIAARILRESLK
YSNIAARILRESLKT
SNIAARILRESLKTG
NIAARILRESLKTGL
1
1.5
2
2.5
Relative fluorescence
B
R8-01 SCR
blank
SWSLQAPRRSVMRWYGFYLNKS
SWSLQAPRRSVM
WSLQAPRRSVMR
SLQAPRRSVMRW
LQAPRRSVMRWY
QAPRRSVMRWYG
APRRSVMRWYGF
PRRSVMRWYGFY
RRSVMRWYGFYL
RSVMRWYGFYLN
SVMRWYGFYLNK
VMRWYGFYLNKS
1
1.5
2
2.5
3
3.5
Relative fluorescence

Figure 7

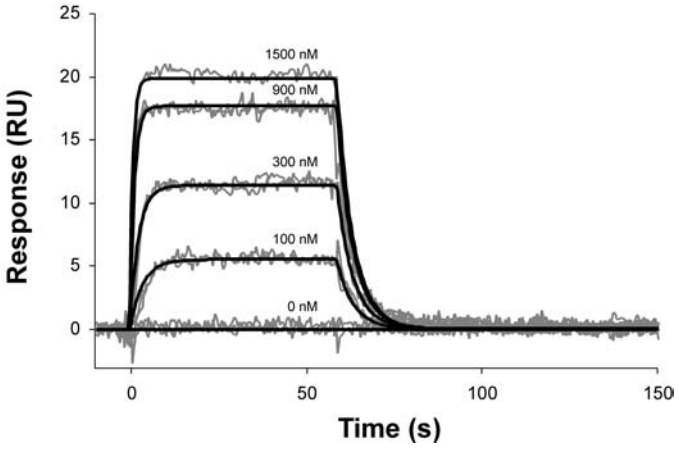
Supporting Information

Table I. Sequences of selected peptides.[a]

	Clone	Peptide sequence		Frequency
Round 5	R5-01	NQKFSPERFTVWWLRASSALLRVPGLR		2/14
	R5-02	IQLVNMPRVGTLRRANMNMSPWRARCR	*	
	R5-03	RYVWYLRTKHRRSLRLRSACARGSSA	*	
	R5-04	HLFSWRDYPWHWVYRSLLARAPRP		3/14
	R5-05	SSLSPPWPASWSPSRPSAPRAAPSTPT	*	
	R5-07	APRAVWIQRAIQAMFRLASRQESKAFN		
	R5-08	KWLVLGRPVQWFVRTLMAMHQAGGSMI		
	S5-01	SNPKMPSLWLVLLSLHTRNEFPNSVSV		2/14
	S5-04	PKKWIQRHIRALRARTWSYFFLSRTR		
	S5-06	LPLEWFERSSSAAAAASWGRPPRRSG	*	
Round 7	R7-01	IVSWGSFPSSWLQRYYLAKRREADVTL		
	R7-02	VRIGYTSKPGGMNPGNSYTMSIIRMLI		
	R7-03	SSLSPPWPASWSPSRPSAPRAAPSTPT	*	
	R7-05	ELGQFQRLGLPYQWYLRTISYVTFRTA		
	R7-06	VLYPREWFFRAWKSYNASNAGLKDTPR	*	
	R7-07	RSPWARQFPEWDRMRNHMNPI	*	
	R7-08	IYSAYPVSWVARTCAATRARSAGARSA	*	
Round 8	R8-01	NVSWGSFPSSWLQRYYLAKRREADVTL		4/20
	R8-03	GDDMYRIREFLANYRPIWVMRSNLAQL		
	R8-04	VRIGYTSKPGGMNPGNSYTMSIIRMLI		7/20
	R8-07	LKYPDTWLARSLSVFYLRKSARQGKSV		
	R8-12	RLVWIVRSRHFGPRLRMALLGSDRKMW		
	R8-13	ELGQFQRLSLPYQWYLRTISYVSLRTA		
	R8-14	APRAVWIQRAIQAMFRLASRQESKAFN		
	R8-07b	STAGSRARSTSWGTRSPWTWPTPARTG	*	
	R8-08b	SSLSPPWPASWSPSRPSAPRAAPSTPT	*	2/20
	R8-09b	RYVWYLRTKHRRSLRLRSACARGSSA	*	

[a] Only the random region of the peptide library is shown. For peptides that appeared more than once, the frequency out of the total number of clones sequenced for that round is shown. Marked sequences (*) contained deletions or mutations that modified the constant C-terminal peptide (not shown). The sequence for R7-07 assumes that translation started at an alternate methionine codon in the random region, as the originally encoded methionine was followed by several stop codons. Related peptides that appeared in more than one selection round are color-coded.

Supplemental Figure 1. Binding interaction of R8-04 peptide with immobilized Mth ectodomain, measured by SPR. The indicated concentrations of R8-04 peptide were injected (45 mL at 0 s, with a 45 mL/min flow rate) across ~500 RU of immobilized Mth-biotin. The global kinetic fits (black) are overlaid on the original sensorgrams (gray). The derived kinetic parameters are shown in Table III. Sensorgrams have been double-referenced from response curves generated by an appropriate negative control flow cell and averaged buffer blank injections. Three buffer blanks are shown (0 nM R8-04).



Supplemental Figure 1

Appendix A

Epitope mapping using mRNA display and a unidirectional nested deletion library

William W. Ja, Brett N. Olsen, and Richard W. Roberts

Epitope mapping of an anti-polyhistidine monoclonal antibody has been performed by *in vitro* selection using mRNA display with a random, unconstrained 27-residue peptide library. After 6 rounds of selection, peptides were identified that contain 2 to 5 consecutive, internal histidines and are biased for arginine residues, without any other identifiable consensus. The epitope was further refined by constructing a high complexity, unidirectional fragment-library from the final selection pool. Selection by mRNA display minimized the dominant peptide from the original selection to a 15-residue functional sequence. Other peptides recovered from the fragment-library selection reveal a separate consensus motif (ARRXA) C-terminal to the histidine-track. Kinetics measurements made by surface plasmon resonance, using purified Fab fragments to prevent avidity effects, demonstrate that the selected peptides bind with 10- to 75-fold higher affinities than a hexahistidine peptide. The highest affinity peptides (K_D = ~10 nM) encode both a short histidine-track and the ARRXA motif, suggesting that the motif and other flanking residues make important contacts adjacent to the core polyhistidine-binding site and can contribute >2.5 kcal/mol of binding free energy. Besides epitope mapping, the fragment-library construction methodology described here is applicable to the development of high complexity protein or cDNA expression libraries for the identification of protein-protein interaction domains.

Introduction

Epitope mapping, the identification of regions of an antigen recognized by an antibody, is an important subset of protein-protein interaction analysis that is relevant in a wide range of disciplines where antibodies are used as molecular reagents. Conventional methods for epitope mapping involve the synthesis or expression of numerous overlapping polypeptides followed by probing for antibody reactivity (1-5). Although these methods can achieve very fine-mapping (single amino acid resolution) of antibodies, they involve tedious, time-consuming, and often cost-intensive steps. These techniques also require *a priori* knowledge of one of the interacting partners (i.e., the antigen sequence).

Display technologies such as phage (6) and cell surface display on *E. coli* or yeast (*7, 8*) permit the assay of millions of polypeptides simultaneously for the identification of functional properties. In these systems, each display vehicle expresses multiple copies of a single polypeptide sequence on its surface. Active peptides are recovered by affinity selection (e.g. by biopanning or fluorescence-activated cell sorting) and identified by DNA sequencing of the library inserts. Random peptide libraries (9-11), antigen- or gene-fragment libraries (12-14), or a combination of both (*15, 16*) have previously been used for the epitope mapping of a wide variety of monoclonal antibodies (mAbs)[1] (reviewed in (17)). Generally, these libraries suffer from low starting complexities and do not always achieve fine-mapping of antibodies unless the epitope is short (~5 residues) and well-defined. Peptide selection in combination with immunoassay of overlapping synthetic peptides has been used to fully delineate the physicochemical

[1] Abbreviations: β-ME, β-mercaptoethanol; DEPC, diethyl pyrocarbonate; DROP, directional random oligonucleotide primed; IPTG, isopropyl-β-D-thiogalactopyranoside; Fab, fragment antigen-binding; mAb, monoclonal antibody; MBP, maltose-binding protein; RU, resonance units; SPR, surface plasmon resonance; UDG, uracil-DNA glycosylase; UTR, untranslated region.

requirements for functional epitopes and accessory factors that influence binding affinity (*16, 18, 19*).

More recently, entirely *in vitro* techniques for protein selection such as ribosome (20-22) and mRNA display (23) have emerged. In mRNA display, peptides are covalently attached to the 3'-end of their encoding mRNA via a tethered puromycin moiety. Pools of RNA-peptide fusions are selected for binding via their attached peptides and recovered fusions are RT-PCR-amplified for the next round of selection and/or cloned for DNA sequencing (Figure 1). The mRNA display system generates libraries that are robust (functional in a wide variety of conditions), encode high complexities ($>10^{13}$ unique sequences, compared with $\sim 10^{8}$-10^{9} for techniques requiring an *in vivo* transformation step), and lack avidity effects as only one peptide is displayed per mRNA sequence. By accessing larger libraries, extremely rare sequences (such as long, discontinuous epitopes or peptides with better functional properties) can be selected and amplified (24). Epitope-like consensus motifs that define the core determinants of binding for the trypsin active site and for the anti-c-Myc antibody, 9E10, have previously been identified using mRNA display with a random peptide library (25).

A further advancement of mRNA display technology is described here, where a unidirectional nested deletion library is constructed. A number of methods have been described for generating gene- or fragment-libraries from DNA, typically involving degenerate oligonucleotide priming (26-28), random fragmentation of DNA (29), or iterative removal of bases from either end of the gene (30-32), followed by ligation to a vector or PCR for subsequent amplification of the library. These techniques have been employed for a variety of purposes, including epitope mapping and the determination of

protein interaction domains (*12-16, 33*). Because of the random nature of library construction, the majority of sequences in these libraries are non-viable due to frame shifts and ligations in the anti-sense orientation. Techniques have been described to maintain gene orientation using a pair of degenerate primers with constant 5' sequences used sequentially in the amplification of cDNA (DROP synthesis, (26)) or mRNA (*33, 34*). However, these methods are technically challenging and may be prone to poor library coverage due to biased hybridization to target sequences (*35, 36*).

As mRNA display facilitates selection from peptide libraries larger than previously possible, improvements are needed for generating libraries with broad coverage while maintaining high sequence complexity. The method described here uses a partial DNase I digestion to fragment the DNA pool randomly. These fragments are then directionally amplified, maintaining the sense orientation, and used to generate an mRNA display library. We first developed a pool of active members by performing *in vitro* selection with a random peptide library against a His_6-tagged protein immobilized by an anti-polyhistidine mAb. Due to the weak affinity of the mAb for the cited His_6 epitope, we inadvertently selected for peptide sequences with high affinity for the antigen-binding region of the mAb. This pool of mAb-binding peptides was subsequently used as the template for a nested deletion library. A 35-residue "winning" peptide was minimized to a 15-mer sequence using the mRNA display fragment-library. Selected peptides were analyzed by surface plasmon resonance and demonstrated 10- to 75-fold higher affinities than the cited epitope. The fragment-library selection also revealed a new motif important for high affinity binding, demonstrating how sequence length may be an important factor in delineating an epitope. The nested deletion construction methods

should be highly applicable toward the isolation of minimal protein interaction domains from cDNA or protein expression libraries using mRNA display.

Experimental Procedures

General

Enzymes were purchased from New England Biolabs unless otherwise noted. Other reagents and solvents were obtained from Sigma-Aldrich or VWR International. All buffer components for RNA and RNA-peptide fusions were made with DEPC-treated ddH_2O. DNA oligos were synthesized at the Caltech Biopolymer Synthesis and Analysis Facility and were desalted by OPC purification with the exception of DNA template 130.2 which was synthesized at the W. M. Keck Foundation Biotechnology Resource Laboratory (http://keck.med.yale.edu) and purified by urea-PAGE. Oligo and peptide concentrations were determined by UV spectrophotometry using a calculated extinction coefficient (http://paris.chem.yale.edu/extinct.html). Fab and MBP fusion protein concentrations were determined by UV absorbance at 205 nm (37). The values obtained with this method were within 5% of those obtained using a calculated extinction coefficient at 280 nm.

mRNA display library construction

The anti-sense DNA template 130.2 (5'-AGC GCA AGA GTT ACG CAG CTG $(SNN)_{27}$ CAT TGT AAT TGT AAA TAG TAA TTG TCC C, S = C or G, N = A, C, G, or T) was PCR-amplified with primers 47T7FP (5'-GGA TTC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT AC) and mycRP (5'-AGC GCA AGA GTT ACG CAG CTG) to produce the initial template containing a T7 promoter, a 5'-untranslated

region (UTR), an ATG methionine start codon, 27 random amino acids each encoded by NNS, and a constant 3'-end that encoded the peptide, QLRNSCA. *In vitro* transcription, purification of the mRNA templates, and ligation of the puromycin linker oligo were performed essentially as described (38). Transcription reactions were pretreated with RNA*secure* (Ambion) to inhibit RNase activity and library DNA was removed by DNase I (Epicentre) digestion prior to purification of the mRNA pool. The ligation was performed with the puromycin-DNA linker, pF30P (5'-A_{21}[S9]$_3$ACC-P, S9 = spacer phosphoramidite 9, P = puromycin, 5'-phosphorylated with phosphorylation reagent II, Glen Research) and a splint oligo (5'-TTT TTT TTT TTN AGC GCA AGA GT) (38). Puromycin-conjugated templates (mRNA-F30P) were purified by urea-PAGE.

RNA-peptide fusion preparation and selection

Purified mRNA-F30P templates were translated in rabbit reticulocyte lysate (Red Nova lysate, Novagen) as per the manufacturer's instructions with optimized conditions (100 mM KOAc, 0.5 mM MgOAc, and 0.5 μM mRNA-F30P) and additional L-Met (0.5 mM final, 1 mL total reaction volume) or ^{35}S-Met labeling (150 μL reaction, New England Nuclear, now PerkinElmer Life Sciences). Following the incubation step at 30 °C, KOAc and $MgCl_2$ were added to 585 mM and 50 mM (final), respectively, and the reactions were incubated on ice for 15 min to facilitate RNA-peptide fusion formation. Radioactively labeled and non-labeled RNA-peptide fusions were pooled and subsequently purified with oligo dT-cellulose (New England Biolabs) as described (38). Purified fusions were concentrated (Microcon YM-30, Millipore) and reverse transcribed as per the manufacturer's instructions (Superscript II, Invitrogen) with the mycRP primer.

The matrix preparation and all selection steps were performed at 4 °C. The reverse-transcribed fusions, in 1 mL of selection buffer (50 mM HEPES-KOH, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM NaF, 30 μM $AlCl_3$, 0.05% Tween 20, 1 mM β-ME, and 5 μM GDP), were precleared by rotating with 20 μL of protein G-sepharose (4B Fast Flow, Sigma) for >1 h. The supernatant was transferred to the target matrix (80 μg of His_6-$G_{i\alpha 1}$ (39) immobilized by 40 μg of anti-polyhistidine mAb (H1029, Sigma) on 20 μL of protein G-sepharose) and rotated for 1 h. The matrix was washed with 3 × 1 mL selection buffer and the bound RNA-peptide fusions were eluted with 2 × 200 μL 4% acetic acid through a 0.45 μm spin filter (SpinX, Costar). Washes and an aliquot of the elution were scintillation counted (LS 6500, Beckman Coulter) to determine the amount of bound fusions.

The eluted fusions were either desalted by ultrafiltration (Microcon YM-30, Millipore) or frozen and dried by vacuum centrifugation. After resuspension in ddH_2O or 10 mM Tris-HCl, pH 8, samples were PCR-amplified for the next cycle of selection and/or for DNA sequencing (TOPO TA cloning, Invitrogen). Subsequent selection rounds were performed similarly, except that smaller translation reactions were used (300 μL non-labeled, 100 μL ^{35}S-Met labeled). Unblocked mAb (without the His_6-tagged protein) was used as the target in the 6th round of selection, when it was realized that the peptides were specific for the mAb.

RNA-peptide fusion binding assay

Aliquots of purified ^{35}S-Met labeled RNA-peptide fusions were treated with RNase (DNase-free, Roche) and added to ~15 μL of protein G-sepharose matrix (with or without

~10 μg of anti-polyhistidine mAb) in 1 mL of selection buffer. Mixtures were rotated at 4 °C for 1 h and washed with 3 × 1 mL selection buffer. The percent binding was determined by scintillation counting of the washes and the matrix.

Fragment-library preparation

To generate the fragment-library, first-strand cDNA from a selected library was synthesized with dUTP instead of dTTP nucleotides (Superscript II). After RNase H treatment (Roche) to remove mRNA, the cDNA was purified by spin-column (QIAquick, Qiagen) and randomly digested with DNase I (0.25 U DNase I (Invitrogen) added to 30 pmol cDNA (~1.2 μM final) in ice-cold 1× DNase I buffer (10 mM Tris-HCl, pH 7.4, 2.5 mM $MgCl_2$, and 0.1 mM $CaCl_2$)) at 15 °C for 10 min. DNase I was removed using DNase Removal Reagent (Ambion). A fill-in reaction (Sequenase v2.0, Amersham Biosciences) was performed according to the manufacturer's instructions with 125 pmol of myc6-N6-FP (5'-ATC TCT GAA GAG GAC CTG NNN NNN) and 200 μM of each dNTP (~0.6 μM cDNA final). First-strand cDNA was digested with uracil-DNA glycosylase and ssDNA >50 bases was extracted with QiaEX II (Qiagen) from a 4% agarose gel (40). A second fill-in reaction was performed with 3myc-N6-RP (5'-AAA TGC ACA AGA GTT GCC CTC GNN NNN N) as before. The dsDNA was subsequently agarose gel-purified by spin-column (QIAquick).

PCR using primers T7mycFP (5'-GGA TTC TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA ATG GAA CAG AAA CTG ATC TCT GAA GAG GAC CTG) and psn3mycRP (5'-AAA TGC ACA AGA GTT GCC CTC G) resulted in a smear of products ranging from 100 to 200 bp on an agarose gel. DNA

corresponding to 150 to 200 bp was extracted by spin column (QIAquick). Amplification of the dsDNA by PCR using the primers 47T7FP and psn3mycRP produced the initial library for selection. The selection was performed against the anti-polyhistidine mAb as before, except that the puromycin moiety was coupled to the mRNA by UV photo-crosslinking with oligo psn-mycF15P (5'-[Ps]-TGC ACA AGA GTT GA_{15}-$[S9]_2$-CC-P, Ps = Psoralen C6, Glen Research) as described previously (41). The selection buffer used for the fragment selection was 1× PBS, 1 mM β-ME, 1 mM EDTA, 0.05% Tween 20, 0.2% (w/v) BSA, and 1 μg/mL yeast tRNA (Roche). In rounds 2 and 3 of the selection, the matrix was more stringently washed by incubation in buffer containing poly-L-His (0.15 mg/mL) and His_6 peptide (60 μM, Covance Research Products) for ~40 min at 4 °C (42).

Direct binding assay of in vitro *translated peptides in lysate*

Individual clones (in pCR4-TOPO vector, Invitrogen) were PCR amplified with primers 47T7FP and mycRP, *in vitro* transcribed, urea-PAGE-purified, and *in vitro* translated (Red Nova Lysate) with ^{35}S-Met labeling as per the manufacturer's instructions. 4 μL of the translation reaction was added directly into an assay tube (600 μL fragment selection buffer, 10 μL protein G-sepharose, 5 μg anti-polyhistidine mAb). After rotating at 4 °C for 1 h, the sepharose was washed with 6 × 600 μL fragment selection buffer in a 0.45 μm spin filter (SpinX) and bound peptides were eluted with 2 × 20 μL 0.05% SDS. Half of the sample was analyzed via tricine SDS-PAGE along with 2 μL of the original translation reaction for comparison. After electrophoresis, gels were destained (40% methanol and 10% acetic acid) for 20 min, dried under vacuum, and imaged via

autoradiography (Storm Phosphorimager, Amersham Biosciences). Peptide band intensities were analyzed with ImageQuant software (Amersham Biosciences).

Peptide synthesis/protein purification

Peptides were synthesized on an ABI 432A Synergy peptide synthesizer (Applied Biosystems) using Fmoc chemistry. Peptides included the sequence GGY**K**-NH_2 at their C-terminus, where **K** is biotinyl-lysine (biocytin, BAchem) and -NH_2 represents C-terminal amidation. The tyrosine residue, used for quantitation by UV absorbance, was omitted from the synthesis for peptides that already contained a tryptophan and/or tyrosine. Crude peptides were deprotected in TFA:thioanisole:1,2-ethanediol (450:25:25 μL, 2 h at room temperature), precipitated with methyl tert-butyl ether, purified to >95% purity by reverse-phase HPLC on a semi-preparative C18 column (250 × 10mm, Vydac), and confirmed by MALDI-TOF mass spectroscopy.

Several peptide sequences were expressed in *E. coli* as *in vivo* biotinylated maltose-binding protein (MBP) fusions using a vector derived from pDW363 (43). The MBP gene from pDW363 was amplified by successive PCR (primers 35.3 5'-GGA CTA GTA AAA TCG AAG AAG GTA AAC TGG TAA TC and 35.4 5'-CCA TTG GAT CCT TAA TTA GTC TGC GCG TCT TTC AG, then primers 84.1 5'-GAG CAC TCG AGC GGT GCG AAT TCA AAC AAC ATC GAG GGG CGC GCC GGT GGC ACT AGT AAA ATC GAA GAA GGT AAA CTG GTA ATC and 29.3 5'-CCA TTG GAT CCT TAA TTA GTC TGC GCG TC). The PCR-amplified fragment and pDW363 were digested with XhoI/BamHI, purified, and ligated to produce the pDW363B vector.

DNA templates encoding peptides B and C were amplified by PCR using the universal forward primer 29.4 (5'-TGA AGT CTG GAG TAT TTA CAA TTA CAA TG) and a template-specific reverse primer that added a SpeI site. BpmI/SpeI digested dsDNA was co-ligated into XhoI/SpeI digested pDW363B with DNA linkers (XhoI linker 5'-TCG AGC TCT GGA GGC ATC GAG GGT CGC AT and BpmI linker 5'-GCG ACC CTC GAT GCC TCC AGA GC) to produce the expression vector. Inserts contained an N-terminal bio-tag, peptide B or C, and a C-terminal MBP fusion. The vectors produce a dicistronic mRNA which encode the bio-tag-peptide-MBP fusion and biotin holoenzyme synthetase (birA), an enzyme that attaches biotin to the bio-tag *in vivo*.

Protein expression with 30 mL cultures of *E. coli* BL21 cells was performed as described (43). Cells were lysed with B-PER (Pierce) and MBP fusions were purified on monomeric avidin-agarose (Pierce) as per the manufacturer's instructions. Purified proteins were concentrated and desalted into 1× PBS by ultrafiltration (Centriprep YM-10, Millipore).

Anti-polyhistidine mAb in ascites fluid was affinity purified on protein G-sepharose in 1× PBS/0.1% triton X-100, eluted with 0.1 M citric acid buffer, pH 3, and immediately neutralized with buffer. After concentration and buffer exchange (Centriprep YM-50) into papain buffer (20 mM phosphate, pH 7, 10 mM EDTA), Fab fragments were generated and purified using the ImmunoPure Fab Preparation Kit (Pierce) as per the manufacturer's instructions.

Surface plasmon resonance

SPR measurements were made at 25 °C on a Biacore 2000 (Biacore) equipped with either SA (streptavidin) sensor chips or research-grade CM5 sensor chips (Biacore) with amine-coupled streptavidin (ImmunoPure, Pierce). The CM5-streptavidin chips were prepared in-house by standard NHS/EDC amine coupling (Biacore) and achieved >1100 RU of immobilized streptavidin per flow cell. HBS-EP (20 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005% surfactant P20 (Tween 20)) was used as the running buffer for all experiments. Biotinylated ligands were diluted in HBS-EP to 1 nM and immobilized to individual flow cells (~10 RU for peptides and ~100 RU for proteins). Flow cell 1 was left as a streptavidin negative control in all sensor chips. To collect kinetics data, a concentration series of Fab in HBS-EP was injected for 2 min at 35 μL/min over all flow cells and dissociation was observed for 3 min. The Fab samples were injected in random order, interspersed with a number of buffer blank injections for double referencing (44). Flow cells were regenerated between Fab injections with a 0.5 min wash of 2.5 M NaCl at 100 μL/min. Raw data was processed with Scrubber and analyzed with CLAMP using a 1:1 bimolecular interaction model (45). K_D values were calculated (k_d/k_a) from the on and off rates determined by CLAMP. Standard free energies of binding were calculated from the K_D values ($\Delta G° = -RT \ln(C/K_D)$, R = 1.987 x 10^{-3} kcal mol^{-1} K^{-1}, T = 298.15 K, and C = 1 mol L^{-1}).

Results

Selection of a random peptide library against an anti-polyhistidine mAb

The peptide selection experiment, originally designed to target a His_6-tagged protein immobilized by an anti-polyhistidine mAb, utilized a random, unconstrained 27-mer peptide library. During PCR and transcription the complexity of the library was maintained by having at least 7×10^{13} sequences at the start of each reaction. The initial mRNA display pool contained at least 10^{12} unique peptide sequences, estimated from the initial mRNA and methionine concentrations in the translation reaction, out of a maximum complexity of 20^{27} peptides ($\sim 1.3 \times 10^{35}$).

Five rounds of selection were performed on the immobilized anti-polyhistidine mAb, pre-saturated with an N-terminal His_6-tagged protein (Figure 2A). Bound RNA-peptide fusions were eluted with acetic acid, which generally recovered >80% of the remaining ^{35}S counts. To determine the progress of the selection, a separate ^{35}S-Met labeled RNA-peptide fusion pool from the 5^{th} round was purified, RNase-treated, and assayed for binding (Figure 2B). This assay revealed specific binding of the peptide pool (now modified only at the C-terminus with puromycin and a short DNA linker) to the antibody rather than to the immobilization matrix (protein G-sepharose) or to the His_6-tagged protein. The reduced binding observed when a His_6 peptide was added as a competitor further evinced that the selected peptide sequences specifically targeted the antigen-binding region of the mAb. A 6^{th} round of selection, performed with unblocked mAb as the target, demonstrated that the enrichment for active peptides against the mAb was essentially complete (Figure 2A).

DNA sequencing of the final 6th round pool revealed a variety of peptides containing 2 to 5 consecutive His residues with no other apparent consensus except a bias for Arg. The His-track was seen in various positions in the random region of the library suggesting that the mAb had little preference for the epitope at either terminus or as an internal binding site. One sequence, peptide C, emerged as the dominant member of the selection (Table I). Further rounds of selection using His_6 peptide and/or poly-L-His as competitors in the selection buffer generally resulted in changes in the percentage of peptide C in the pool rather than the emergence of new, beneficial mutations or peptides defining a single consensus (data not shown). Peptide C remained the most prevalent sequence in all subsequent selection rounds, with a collective frequency of 20 out of 53 sequences (Table I).

Selection for a minimal binding epitope

To narrow down the epitope and isolate shorter, high-affinity peptide sequences, a nested deletion library was constructed from the peptide C-dominated library. This library is composed of fragments of DNA that encode shorter stretches of the parent peptides. By using the fragment-library in an mRNA display selection, minimal binding sequences can be identified. Initial attempts to generate nested deletions using random priming on cDNA resulted in nearly full-length sequences, possibly due to the strand-displacement abilities of the polymerases used (46).[2] This attribute was exploited in the final fragmentation scheme (Figure 3A). DNase I was used to generate random fragments from the cDNA of a functional library (any pool after the 6th round of selection). Various

[2] Unpublished observations and I. N. Hampson, personal communication.

dilutions of DNase I were used to find the optimal conditions for producing a range of ssDNA products from ~50 to 130 bases (data not shown). Successive random priming and fill-in reactions with a modified T7 polymerase (Sequenase v2.0) and primers containing 3'-random hexamers produced the initial DNA pool. PCR-amplified dsDNA was agarose gel-purified to retain fragments between 150 and 200 bp, corresponding to peptides approximately 10 to 30 amino acids long.

Because stop codons hinder RNA-peptide fusion formation, the 3'-constant sequence of the fragment-library was chosen such that TAA, TAG, and TGA codons did not exist in any frame. The 5'-constant region added a c-Myc epitope tag and provided a primer site for subsequent PCR amplification (for additional attachment of the T7 promoter and UTR sequence). This method resulted in a unidirectional fragmented pool; all transcribed RNA maintained the sense orientation. DNA sequencing of the initial pool demonstrated reasonable representation of the dominant sequence (peptide C) and confirmed the expected 1/3 fraction of in-frame sequences (Figure 3B). DNA alignments with peptide C derivatives typically contained several mismatches at the beginning and end of the fragment region, most likely due to imperfect annealing of the random hexamer primers.

The nested deletion library was used for selection against the anti-polyhistidine mAb (Figure 4). Poly-L-His and His_6 peptide were used as competitors in the 2^{nd} and 3^{rd} rounds. Although the binding of the 2^{nd} and 3^{rd} round pools was similar, more RNA-peptide fusions were retained after the stringent, competitive wash in the 3^{rd} round, suggesting that the washes were indeed enriching the pool for the highest affinity peptides. DNA sequencing of the final pool revealed three distinct classes of peptides

(Table II). Class 1 sequences were fragments corresponding to N- and C-terminal deletions of peptide C. A sequence alignment of the fragments identified RHDAGDHHHHHGVRQ (peptide Cmin) as a minimal functional sequence for peptide C.

The majority of fragments recovered after the selection came from parent sequences other than peptide C (Table II, Class 2). An alignment of peptides D and E (which collectively represented 40% of the final, 3rd round selection pool) revealed the consensus motif ARRHA. This exact motif was not seen in the original selection, although three peptide sequences contained ARRXA (X = R, G (peptide A), or K (peptide B)) two residues C-terminal to the His-track (Table I), as in peptide D. Additional N- and C-terminal deletions for peptides D and E were not observed. Hence, these sequences may already represent minimal high affinity binding epitopes. Alternatively, there may have been an insufficient number of clones sequenced to find other corresponding fragments. Other recovered sequences in this peptide class retained at least part of the ARRXA, suggesting that the first few residues of the consensus motif are more critical for high affinity.

Several additional peptides were discovered that encoded a weak consensus sequence non-related to the mAb-binding peptides (Table II, Class 3). Binding assays with a couple of these peptides revealed significantly weaker affinity for the mAb than a His_6-containing peptide control (data not shown). These peptides may bind to an alternate interaction site and were consequently enriched when high stringency, competitive washes were introduced for the last rounds of selection. Site-specific, competitive washes (e.g., with poly-L-histidine) would result in the enrichment of peptides with

higher affinity for the antigen-binding region, as well as for peptides with affinity for other sites.

Immunoprecipitation of selected peptides

Selected clones were qualitatively assessed for binding by immunoprecipitation with the anti-polyhistidine mAb (Figure 5A). ^{35}S-Met labeled peptides were assayed directly from the *in vitro* translation reactions. The selected peptides demonstrated significantly increased binding compared with a C-terminal His_6-tagged peptide control (Figure 5B). Non-specific binding was shown to be minimal with a c-Myc epitope control peptide. Correct translation of the fragment-selected peptides and the Myc control was confirmed by immunoprecipitation on the 9E10 anti-c-Myc mAb (data not shown).

Kinetics by surface plasmon resonance

Various peptides from the fragment selection were synthesized and purified for kinetics analysis by surface plasmon resonance (SPR). In an SPR experiment, one binding partner (ligand) is immobilized on the surface of a sensor chip while the other reactant (analyte) is in solution. Binding of the analyte is seen as a refractive index change on the sensor chip surface and is measured in real-time in resonance units (RU). Peptides were synthesized with a C-terminal biocytin residue for immobilization on streptavidin-coupled surfaces. Full-length peptides B and C were also assayed by expressing the peptides as fusion proteins with a C-terminal MBP and an N-terminal bio-tag, which is biotinylated *in vivo* by biotin holoenzyme synthetase (BirA). By purifying these proteins via monomeric avidin, they retained their biotin moieties and a homogeneous ligand surface could be produced on the SPR sensor chips.

Rebinding and bivalency effects of mAb interactions with immobilized antigens have previously been shown to offset kinetics measurements considerably, rendering both absolute and relative binding constants unreliable (47). To avoid these problems, Fab fragments were prepared from anti-polyhistidine mAb and used as the analyte. Using the peptides as the immobilized ligands and Fab as the analyte ensured fair comparisons between the kinetics measurements, avoiding bias in protein quantitation, since all Fab concentrations were prepared from a single stock solution. Kinetics parameters were determined using a 1:1 bimolecular interaction model (Table III).

The assayed peptides could be categorized by their dissociation rates from the Fab (Figure 6). The cited epitope, His_6, bound weakest to the Fab; the His_6 peptide and the His_6-tagged protein used in the original selection exhibited dissociation constants of 0.6 and 3 μM, respectively. Additional His residues (His_{10} peptide) increased the association rate 6-fold without changing the dissociation rate significantly. Peptides from the selection demonstrated dissociation constants less than 75 nM, approximately 10- to 75-fold better than the control His_6 sequence, with increased affinities as a result of faster association (up to 5-fold) and considerably slower (6- to 21-fold) dissociation rates (Table III). Class 2 peptides with the ARRXA motif demonstrated the highest affinities, with ~3-fold slower dissociation rates compared to sequences derived from peptide C (Figure 6C). While the flanking residues on the minimized peptide C contribute at least 1.6 kcal/mol to the binding free energy compared with the His_6 peptide, sequences with the ARRXA motif demonstrate 2.6 (peptide B) and 2.2 (peptide D) kcal/mol improvements. The contributions from these flanking residues is likely even greater, as

these calculations do not account for any loss of binding free energy from having shorter (<6) stretches of His residues in the core site.

Discussion

During an *in vitro* selection experiment against a target protein immobilized using an anti-polyhistidine antibody, mAb-binding peptides were inadvertently enriched. The weak affinity of the His_6-tagged fusion protein for the mAb and the existence of alternative peptide motifs that confer significantly higher affinity are the likely causes for the inability to enrich for peptides that bind the original target protein. A preclearing step that included the mAb may not have been totally effective in preventing the selection of antibody-specific peptides, as even the final selection round resulted in an incomplete, ~40% pull-down of the RNA-peptide fusions. Although the cited mAb epitope is hexahistidine, the recovered peptides surprisingly each contained a shorter (≤5) stretch of consecutive His residues and a bias for Arg.

To better characterize the mAb epitope and demonstrate the feasibility of gene fragment mRNA display, a nested deletion library was constructed from the final selection pool. A modified DROP-amplification of cDNA was performed to maintain as many viable library fragments as possible (26). Due to the difficulty in obtaining a broad size distribution of sequences with degenerate oligos, the protocol was modified to use DNase I for the random fragmentation of cDNA. DROP-synthesis using a highly processive DNA polymerase, capable of potent strand-displacement, yielded intact copies of the cDNA fragments while maintaining the sense strand (Figure 3A).

In vitro selection with the fragment-library resulted in the identification of a 15-mer functional sequence derived from the full-length 35-mer, peptide C. Because the initial fragment-library was produced from a pool dominated by peptide C, we expected to recover and identify numerous overlapping peptides that defined a minimal epitope for this sequence. Surprisingly, the majority of recovered sequences came from unknown parents. The enrichment of these peptides implies that these fragments were more highly favored after truncation. The flanking regions of the original peptides may have hindered access to the epitope by the mAb, suggesting that peptide length may be an important attribute in the fine-tuning of affinity and/or function. Alternatively, these particular sequences may have been negatively biased by the constant C-terminal peptide used in the original random peptide library. The 3-frame constant sequence used in the fragment-library construction increases the sensitivity of the selection when one of the translation frames causes negative bias. Additionally, a random distribution between the three translation frames would indicate that the constant region does not affect selectability. The 6 independent clones of peptide D, for example, had all 3 frames represented in the 3' constant region (Table II and data not shown).

Based on the selected peptide sequences, two major protein interaction motifs were identified: a core epitope consisting of at least three consecutive His residues and a 2nd interaction site encoded by the consensus motif, ARRXA. SPR experiments demonstrated a significant increase in the association rate of His_{10} compared with His_6, suggesting that additional His residues present a more accessible core interaction, rather than slow dissociation by enhancing rebinding from multivalency effects. Only additional contacts, made by the addition of interacting residues such as the ARRXA

motif, result in significantly slower dissociation rates. These flanking residues can contribute significantly to the binding free energy—at least 2.6 kcal/mol in the case of peptide B in comparison with His_6, which assumes the loss of 2 out of 6 histidines in the core has no effect. The two interaction cassettes we have identified here are likely juxtaposed sites from the fusion protein used as the original antigen, a proprietary sequence.[3]

Our results also highlight the importance of flanking residues outside of the two consensus motifs and their contribution to binding affinity with antibodies. Residues adjoining core amino acids in an epitope can substantially influence antibody binding, the effects of which can only be assessed through quantitative affinity measurements (*15, 19*). This is demonstrated in our experiments, where the rank order of binding in the immunoprecipitation assay did not entirely correspond with quantitative kinetics measurements. Epitope tags are often appended to proteins and used as molecular handles for detection, isolation, and analysis of protein-protein interactions. Their functionality in this context, however, is highly variable. Tandem repeats of tags (e.g., the popular c-Myc or FLAG epitopes) have been used to ensure robust affinity and recognition by antisera (*48, 49*). By identifying longer functional peptides with appropriate flanking residues, high affinity can be maintained with less variability depending on the linker region and the protein to which the epitope is attached.

The ability to access high complexity libraries is a great advantage for mRNA display over other selection systems. Library construction methods that involve PCR and DNA reassembly are better suited for the mRNA display format, thereby avoiding cloning steps

[3] Sigma-Aldrich Corp., technical specifications for unconjugated mouse anti-polyhistidine mAb.

that are required in techniques such as phage display. A comparative study on epitope mapping using random 6-mer and 15-mer peptide phage display libraries successfully identified consensus motifs for only 2 of the 4 mAbs examined (16). For one of the mapped mAbs, the random peptide selection succeeded only with the 6-mer library, identifying a short consensus motif that was not discovered with the 15-mer library, which the authors attributed to a statistical lack of representation. Previously, mRNA display with a random 27-mer library revealed epitope-like consensus motifs for the trypsin active site and the anti-c-Myc antibody, 9E10 (25). These experiments achieved relatively fine-mapping of the epitopes, uncovering the core residues as well as some of the allowed flanking amino acids. By utilizing high complexity, long peptide libraries, mRNA display selections can identify rare sequences of high affinity and determine linear or discontinuous epitopes. The full-length consensus peptide, H_m-X_2-ARRXA, for example, may not have been identified with more traditional X_6 or X_{10} phage display libraries.

One of the difficulties noticed in the fragment selection was the disproportionate number of peptides that did not contain an N-terminal deletion. Because of the 5'-UTR on the mRNA used to make the fragment-library, more fragments containing the first start codon (with varying lengths of UTR sequence) were probably present in the initial fragment pool. 5'-UTR and/or promoter sequences most likely do not hinder the fragment selection process, as ribosome scanning can initiate translation at the correct start codon, regardless of which frame was amplified. This was seen in several of the selected fragment sequences (Table II). This property increases the number of viable

(i.e., translatable) templates, but introduces some bias favoring intact N-terminal sequences.

Although not utilized in this experiment, the c-Myc tag introduced in the fragmentation library can be used to generate and purify a fragment-library enriched with in-frame sequences. Although the tag is at the N-terminus of the library, in general RNA-peptide fusions will form only when the ribosome can translate most of the sequence and reach the end of the mRNA (unpublished results). Hence, only sequences that lack stop codons (and therefore are most likely in-frame) will form fusions and be purified and amplified after a Myc-epitope pre-selection. Another improvement to the protocol includes using Exonuclease I to remove excess degenerate primers during DROP-synthesis, preventing the amplification of sequences without "inserts," as DNA size fractionation by agarose gel is not completely effective in removing these smaller fragments (data not shown).

Due to the higher efficiency of synthesizing the nested deletion library completely *in vitro*, the fragment-library construction described here maintains a higher number of unique sequences, in contrast to DNA libraries produced by enzymatic ligation and cloning, which are limited by *in vivo* transformation efficiencies. Additionally, the DROP-synthesis is unidirectional for all amplified sequences so that the sense orientation is maintained and only the minimal 2/3 of the fragments are non-viable due to frame shifts. This protocol produces a well-distributed library and is technically less challenging as the random oligonucleotide priming is used only to "copy" the cDNA fragments produced by DNase digestion, and need not be optimized for generating a fragment distribution. mRNA display with fragment-libraries combine the ease and

versatility of working with cDNA *in vitro* with the benefits of expression cloning. The method permits the minimization of functional domains, as well as the isolation of optimal binding contexts through the removal of negative-acting flanking regions. Although the technique may not be sufficiently processive for the fine-mapping of short peptide sequences, it should be highly applicable for constructing cDNA or tissue-specific expression-libraries and the subsequent determination of minimal binding domains and novel protein-protein interactions.

Acknowledgments

We thank Dr. David S. Waugh (National Cancer Institute at Frederick) for the pDW363 *in vivo* biotinylation vector, Prof. Pamela J. Bjorkman and Anthony M. Giannetti for time and support on the Biacore 2000, William Hunter (Biacore, Inc., Piscataway, NJ) for technical advice on SPR, Cindy I. Chen and Christopher T. Balmaseda for preparative and technical assistance on library construction and protein purification, and Prof. David G. Myszka (University of Utah) for generously providing the kinetics analysis software, Scrubber and CLAMP. We greatly appreciate Dr. Yuri Peterson (Duke University) for suggestions on the paper. We are indebted to Dr. Ian N. Hampson (St. Mary's Hospital, Manchester, UK) for valuable discussions and technical expertise on random priming and the synthesis of the fragment-library. This work was supported by grants from the NIH (RO160416) and Beckman Foundation to R. W. R. W. W. J. was supported in part by a DOD National Defense and Engineering Graduate Fellowship. R. W. R. is an Alfred P. Sloan Foundation Research Fellow.

References

1. Lenstra, J. A., Kusters, J. G., and van der Zeijst, B. A. (1990) Mapping of viral epitopes with prokaryotic expression products, *Arch. Virol. 110*, 1-24.
2. Frank, R. (1992) Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support, *Tetrahedron 48*, 9217-9232.
3. Frank, R., and Overwin, H. (1996) SPOT synthesis. Epitope analysis with arrays of synthetic peptides prepared on cellulose membranes, *Methods Mol. Biol 66*, 149-169.
4. Kramer, A., Reineke, U., Dong, L., Hoffmann, B., Hoffmüller, U., Winkler, D., Volkmer-Engert, R., and Schneider-Mergener, J. (1999) Spot synthesis: observations and optimizations, *J. Pept. Res. 54*, 319-327.
5. Reineke, U., Kramer, A., and Schneider-Mergener, J. (1999) Antigen sequence- and library-based mapping of linear and discontinuous protein-protein-interaction sites by spot synthesis, *Curr. Top. Microbiol. Immunol. 243*, 23-36.
6. Scott, J. K., and Smith, G. P. (1990) Searching for peptide ligands with an epitope library, *Science 249*, 386-390.
7. Boder, E. T., and Wittrup, K. D. (1997) Yeast surface display for screening combinatorial polypeptide libraries, *Nat. Biotechnol. 15*, 553-557.
8. Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L., and Curtiss, R., 3rd. (1997) Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines, *Nat. Biotechnol. 15*, 29-34.

9. Miceli, R. M., DeGraaf, M. E., and Fischer, H. D. (1994) Two-stage selection of sequences from a random phage display library delineates both core residues and permitted structural range within an epitope, *J. Immunol. Methods 167*, 279-287.

10. Parhami-Seren, B., Keel, T., and Reed, G. L. (1997) Sequences of antigenic epitopes of streptokinase identified *via* random peptide libraries displayed on phage, *J. Mol. Biol. 271*, 333-341.

11. Murthy, K. K., Shen, S.-H., and Banville, D. (1998) Epitope mapping of SHP-1 monoclonal antibodies using peptide phage display, *Biochem. Biophys. Res. Commun. 248*, 69-74.

12. Kuwabara, I., Maruyama, H., Kamisue, S., Shima, M., Yoshioka, A., and Maruyama, I. N. (1999) Mapping of the minimal domain encoding a conformational epitope by λ phage surface display: factor VIII inhibitor antibodies from haemophilia A patients, *J. Immunol. Methods 224*, 89-99.

13. Christmann, A., Wentzel, A., Meyer, C., Meyers, G., and Kolmar, H. (2001) Epitope mapping and affinity purification of monospecific antibodies by *Escherichia coli* cell surface display of gene-derived random peptide libraries, *J. Immunol. Methods 257*, 163-173.

14. Mullaney, B. P., Pallavicini, M. G., and Marks, J. D. (2001) Epitope mapping of neutralizing botulinum neurotoxin A antibodies by phage display, *Infect. Immun. 69*, 6511-6514.

15. Coley, A. M., Campanale, N. V., Casey, J. L., Hodder, A. N., Crewther, P. E., Anders, R. F., Tilley, L. M., and Foley, M. (2001) Rapid and precise epitope mapping of monoclonal antibodies against *Plasmodium falciparum* AMA1 by

combined phage display of fragments and random peptides, *Protein Eng. 14*, 691-698.

16. Fack, F., Hügle-Dörr, B., Song, D., Queitsch, I., Petersen, G., and Bautz, E. K. F. (1997) Epitope mapping by phage display: random versus gene-fragment libraries, *J. Immunol. Methods 206*, 43-52.

17. Irving, M. B., Pan, O., and Scott, J. K. (2001) Random-peptide libraries and antigen-fragment libraries for epitope mapping and the development of vaccines and diagnostics, *Curr. Opin. Chem. Biol. 5*, 314-324.

18. du Plessis, D. H., Wang, L.-F., Jordaan, F. A., and Eaton, B. T. (1994) Fine mapping of a continuous epitope on VP7 of bluetongue virus using overlapping synthetic peptides and a random epitope library, *Virology 198*, 346-349.

19. Choulier, L., Laune, D., Orfanoudakis, G., Wlad, H., Janson, J.-C., Granier, C., and Altschuh, D. (2001) Delineation of a linear epitope by multiple peptide synthesis and phage display, *J. Immunol. Methods 249*, 253-264.

20. Mattheakis, L. C., Bhatt, R. R., and Dower, W. J. (1994) An *in vitro* polysome display system for identifying ligands from very large peptide libraries, *Proc. Natl. Acad. Sci. U.S.A. 91*, 9022-9026.

21. Hanes, J., and Plückthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display, *Proc. Natl. Acad. Sci. U.S.A. 94*, 4937-4942.

22. He, M., and Taussig, M. J. (1997) Antibody-ribosome-mRNA (ARM) complexes as efficient selection particles for *in vitro* display and evolution of antibody combining sites, *Nucleic Acids Res. 25*, 5132-5134.

23. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12297-12302.

24. 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.

25. Baggio, R., Burgstaller, P., Hale, S. P., Putney, A. R., Lane, M., Lipovsek, D., Wright, M. C., Roberts, R. W., Liu, R., Szostak, J. W., and Wagner, R. W. (2002) Identification of epitope-like consensus motifs using mRNA display, *J. Mol. Recognit. 15*, 126-134.

26. Hampson, I. N., Hampson, L., and Dexter, T. M. (1996) Directional random oligonucleotide primed (DROP) global amplification of cDNA: its application to subtractive cDNA cloning, *Nucleic Acids Res. 24*, 4832-4835.

27. Santi, E., Capone, S., Mennuni, C., Lahm, A., Tramontano, A., Luzzago, A., and Nicosia, A. (2000) Bacteriophage lambda display of complex cDNA libraries: A new approach to functional genomics, *J. Mol. Biol. 296*, 497-508.

28. Whitcomb, J. M., Rashtchian, A., and Hughes, S. H. (1993) A new PCR based method for the generation of nested deletions, *Nucleic Acids Res. 21*, 4143-4146.

29. Gupta, S., Arora, K., Sampath, A., Khurana, S., Singh, S. S., Gupta, A., and Chaudhary, V. K. (1999) Simplified gene-fragment phage display system for epitope mapping, *Biotechniques 27*, 328-330, 332-334.

30. Henikoff, S. (1984) Unidirectional digestion with exonuclease-III creates targeted breakpoints for DNA sequencing, *Gene 28*, 351-359.

31. Milavetz, B. (1992) Preparation of nested deletions in single-strand DNA using oligonucleotides containing partially random base sequences, *Nucleic Acids Res. 20*, 3529-3530.

32. Pues, H., Holz, B., and Weinhold, E. (1997) Construction of a deletion library using a mixture of 5'-truncated primers for inverse PCR (IPCR), *Nucleic Acids Res. 25*, 1303-1304.

33. McPherson, M., Yang, Y., Hammond, P. W., and Kreider, B. L. (2002) Drug receptor identification from multiple tissues using cellular-derived mRNA display libraries, *Chem. Biol. 9*, 691-698.

34. Hammond, P. W., Alpin, J., Rise, C. E., Wright, M., and Kreider, B. L. (2001) *In vitro* selection and characterization of Bcl-X_L-binding proteins from a mix of tissue-specific mRNA display libraries, *J. Biol. Chem. 276*, 20898-20906.

35. Telenius, H., Carter, N. P., Bebb, C. E., Nordenskjöld, M., Ponder, B. A. J., and Tunnacliffe, A. (1992) Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer, *Genomics 13*, 718-725.

36. Zhang, J., and Byrne, C. D. (1999) Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR, *Biochem. J. 337*, 231-241.

37. Scopes, R. K. (1974) Measurement of protein by spectrophotometry at 205 nm, *Anal. Biochem. 59*, 277-282.

38. Liu, R., Barrick, J. E., Szostak, J. W., and Roberts, R. W. (2000) Optimized synthesis of RNA-protein fusions for *in vitro* protein selection, *Methods Enzymol. 318*, 268-293.

39. Lee, E., Linder, M. E., and Gilman, A. G. (1994) Expression of G-protein α subunits in *Escherichia coli*, *Methods Enzymol. 237*, 146-164.

40. Frohlich, M. W., and Parker, D. S. (2001) Running gels backwards to select DNA molecules larger than a minimum size, *Biotechniques 30*, 264-266.

41. Kurz, M., Gu, K., and Lohse, P. A. (2000) Psoralen photo-crosslinked mRNA-puromycin conjugates: a novel template for the rapid and facile preparation of mRNA-protein fusions, *Nucleic Acids Res. 28*, e83.

42. Boder, E. T., and Wittrup, K. D. (1998) Optimal screening of surface-displayed polypeptide libraries, *Biotechnol. Prog. 14*, 55-62.

43. Tsao, K.-L., DeBarbieri, B., Michel, H., and Waugh, D. S. (1996) A versatile plasmid expression vector for the production of biotinylated proteins by site-specific, enzymatic modification in *Escherichia coli*, *Gene 169*, 59-64.

44. Myszka, D. G. (2000) Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE, *Methods Enzymol. 323*, 325-340.

45. Myszka, D. G., and Morton, T. A. (1998) CLAMP: a biosensor kinetic data analysis program, *Trends Biochem. Sci. 23*, 149-150.

46. Hamilton, S. C., Farchaus, J. W., and Davis, M. C. (2001) DNA polymerases as engines for biotechnology, *Biotechniques 31*, 370-376, 378-380, 382-383.

47. Nieba, L., Krebber, A., and Plückthun, A. (1996) Competition BIAcore for measuring true affinities: large differences from values determined from binding kinetics, *Anal. Biochem. 234*, 155-165.

48. Hernan, R., Heuermann, K., and Brizzard, B. (2000) Multiple epitope tagging of expressed proteins for enhanced detection, *Biotechniques 28*, 789-793.

49. Nakajima, K., and Yaoita, Y. (1997) Construction of multiple-epitope tag sequence by PCR for sensitive Western blot analysis, *Nucleic Acids Res. 25*, 2231-2232.

Tables

Table I. Peptide sequences from anti-polyhistidine mAb selection using a random 27-mer library.[a]

YRTNHHYDVGRFAARGRRD	
NGRSSMNWRSQEITRYTSEHHYRMAFL	
PEQYDHHHLEARRRASSTRQVRARARR	
RAYTPHHHAEGRLVRLEPHPAPYKNRT	
YYVKNRLHHHRLARLVAAEHAHRLRVQ	
NKRNLSYPWSHHHQVARRTHMRAQHTM	
RPTKNFEAEVVRSTGPMHHHDTAKQRY	
DFLTYNKSMGGRPTNFRHHHSSVVQSQ	
DEPEVVGRVLGERPAGALADHHHMMKW	
EVLH*G*HHHVVARVRA*S*CTGPTRRASCA	(6/53)
HVYEKANNRLGHKHHHLAARRRSKSWN	
SNKGFSWRKKGMAVTPNRHLHHHMVAH	
TNHRHHHGVLERRQDILTGSLIEHKH	
ILKRLREQHRHHHAAAHHVRVRRRGRH	
NYTT*R*RAEWNRQDAHRHHHQEARRGAL	A (3/53)
SKKDNAVGLQELRLREGHRHHHDVMLT	
KKVRGHHRHHHQVALLDAAERGPGRMS	
GIHHHHAMAVLAELGMNPMGFALPDMW	
AGVHHHHDAARGGTRSRRSTPRSATRR	*
TMNWHHHHENGLRARMYDAGRR	
KVRRDVMRWHHHHRMARRKANR	B (4/53)
RVQDRLGHRAVQPVLHHHHQAARRRVR	
AALHHHHHDAGRASAMRRPGTPATSWR	
DGHPERHDAG*D*HHHHHGVRQWR*L*ISTG	C (20/53)

[a] Only the random domain is shown. Sequences contained between 2 and 5 consecutive histidines and were aligned at the C-terminal end of the His-track. A consensus was not observed except for a strong bias for Arg several residues C-terminal to the His-track. His and Arg residues are shown in bold. The frequency (out of 53) is shown for peptides that appeared more than once from DNA sequencing of individual clones. For these sequences, amino acids that differed between clones are in italics, with the most common residue at that position shown. Several sequences contained multiple deletions that shortened the random domain but left the C-terminal constant region intact and in-frame. The sequence marked with an asterisk contained a 2 bp insertion which resulted in a frame-shift of the C-terminal constant region (not shown). Peptides A, B, and C are named.

Table II. Peptide sequences from fragment-library selection.[a]

Class 1		
*	MDGHPERHDAGDHHHHHGVRQ	
	ERHDAGDHHHHHGVRQWRLIS	
	RHDAGDHHHHHGVRQWRLIS	
Class 2		
	*IT*NSPGRF**RHHH***V*L**ARR**H**A**LY*R*	D (6/20)
	MTSAGW*T*AM**H**YIS**ARR**H**A**M*R*SMKFAQ	E (2/20)
	NYTTQRAEWNRQDA**HRHHH**QE**ARR**GQ	A1
*	MKVRRDVMRW**HHHH**RM**ARR**K**A**NR	B
	D**HHHHH**GA**AR**PVFRRGLYQKRG	F
	D**HRHHH**GV**AR**VREQMARYV	
Class 3		
	VTM**FD**V**D**A**Y**F**G**LAVWSSGDLRAFQ	
	VTM**FD**V**D**A**Y**F**G**LAVW	(2/20)
*	M**FD**Y**D**AFY**G**YNGSAVGSPTLQHVRLQP	
*	MNF**D**E**Y**LRLLR	

[a] Only the fragment domain of the peptides is shown. Class 1 peptides are derived from peptide C (Table I) and the putative minimal epitope is underlined. Class 2 sequences contain portions of the ARRXA motif. Conserved residues are in bold. Sequences derived from parent peptides A and B, as well as new peptides D, E, and F, are labeled. The C-terminal RGQ in the sequence derived from peptide A is encoded by part of the 3'-constant region. Class 3 peptide sequences were aligned using CLUSTALW (http://npsa-pbil.ibcp.fr) with key residues determined automatically. Clone frequency (out of 20) is shown and differing residues are italicized as described in Table I. Peptide sequences translated from alternate start codons are marked (*).

Table III. Kinetic parameters for peptide interactions with Fab determined by surface plasmon resonance.[a]

	Peptide sequence	k_a	k_d	K_D	χ^2	$\Delta G°$
		$M^{-1} s^{-1}$ ($\times 10^4$)	*s^{-1} ($\times 10^{-2}$)*	*nM*		*kcal/mol*
	HHHHHH	9.9	5.78	580	0.72	−8.5
	HHHHHH-protein	7.3	23.82	3260	0.80	−7.5
	HHHHHHHHHH	62.4	6.56	105	1.19	−9.5
	MDGHPERHDAGDHHHHHGVRQ	11.8	0.85	72	1.28	−9.7
Cmin	RHDAGDHHHHHGVRQ	21.5	0.82	38	1.19	−10.1
C	MDGHPERHDAGDHHHHHGVRQWRLISTG-MBP	52.5	0.97	18.5	1.48	−10.5
B	MKVRRDVMRWHHHHRMARRKANR-MBP	40.9	0.31	7.6	1.28	−11.1
D	NSPGRFRHHHVLARRHALYR	17.4	0.27	15.5	0.77	−10.7

[a] SPR experiments monitored binding between immobilized peptides and purified Fab fragments. On and off rates were determined by global fit analysis on CLAMP using a 1:1 bimolecular interaction model (45). K_D values were calculated from k_d/k_a.

Figures

Figure 1. *In vitro* selection scheme using mRNA display. The starting dsDNA pool (top, center) which encodes the peptide library is transcribed *in vitro*. Purified mRNA is enzymatically ligated to a puromycin-DNA oligo prior to RNA-peptide fusion formation via *in vitro* translation. Purified RNA-peptide fusions are reverse transcribed and affinity selected onto the immobilized antibody target. Eluted cDNA is used as the template for PCR for the next cycle of selection.

Figure 2. Selection of peptides against the anti-polyhistidine mAb. (A) Percent binding from each round of selection was determined by scintillation counting of an aliquot of the ^{35}S-Met labeled RNA-peptide fusions before and after affinity selection on the immobilized antibody. (B) Binding assay of 5th round mRNA display library. Purified, RNase-treated ^{35}S-labeled fusions from the 5th round pool were assayed on protein G-sepharose matrix with and without immobilized anti-polyhistidine mAb. The addition of 10 mM His_6 peptide competitor resulted in reduced binding to the mAb, suggesting that the selected peptides interact specifically with the antigen-binding site.

Figure 3. Construction of a unidirectional nested deletion library. (A) cDNA library reverse transcribed with dUTP is partially digested with DNase I. A randomly-primed fill-in reaction is performed with degenerate DNA hexamers containing a constant 5' sequence, resulting in complete second-strand cDNA for each fragment. After UDG digestion to remove first-strand cDNA, the anti-sense strand is filled-in again by random priming. The constant region of the 2nd primer encodes a suitable peptide sequence in all 3 frames (lacking stop codons) and serves as the reverse primer site for subsequent PCR.

PCR of the resulting dsDNA produces the initial library suitable for *in vitro* selection. (B) Representation of the peptide C parent DNA sequence in the initial fragment-library. The 5'-UTR, peptide coding region, and 3'-constant region are in black, white, and gray, respectively. The bases spanned by each library member are shown. Sequences marked with an asterisk are in-frame with the 5'-constant region added during the generation of the library. The sequence spanning bases 4 through 89 is also viable assuming translation occurs at the first Met codon.

Figure 4. Selection of the peptide fragment-library on anti-polyhistidine mAb. The percentage of recovered fusions (black) was determined as in Figure 2. In rounds 2 and 3, the competitive washes (gray) removed a portion of the initially bound counts.

Figure 5. Binding of *in vitro* translated peptides to anti-polyhistidine mAb. (A) ^{35}S-labeled peptides were assayed for binding directly from the translation reaction. Myc is a peptide encoded by the constant regions of the fragment-library primers with only an arginine residue in between. The His_6 sequence encoded a 31-mer peptide with a C-terminal His_6 tag. Equivalent aliquots of the translation reactions (left lanes) were analyzed by tricine SDS-PAGE adjacently to immunoprecipitated peptides (right lanes). (B) Quantitation of peptide binding in *A*. Relative binding is shown as a fold-change versus the His_6 sequence. Peptide sequences are given in Tables I and II.

Figure 6. Representative sensorgrams from SPR experiments. Purified anti-polyhistidine Fab fragments at concentrations corresponding to ~0.5 K_D were injected over immobilized peptides or peptide-MBP fusions. Peptides fell into three categories

describing weak (A, His_6, His_{10}, and His_6-tagged blocking protein), intermediate (B, peptide C-derived sequences), and strong (C, sequences containing the ARRXA motif) binding for the Fab fragments. For comparison, sensorgrams were divided by the computed maximum signal.

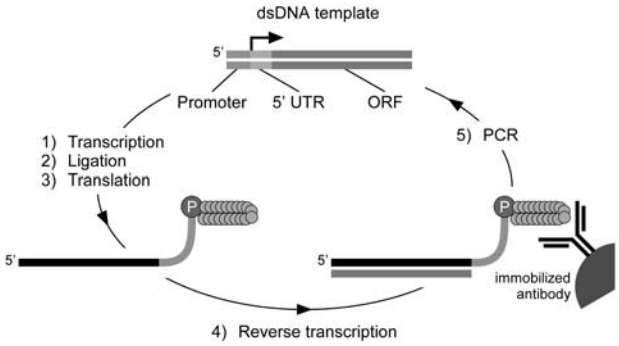

dsDNA template
5'
Promoter
5' UTR
ORF
1) Transcription
2) Ligation
3) Translation
5) PCR
P
5'
P
5'
immobilized
antibody
4) Reverse transcription

Figure 1

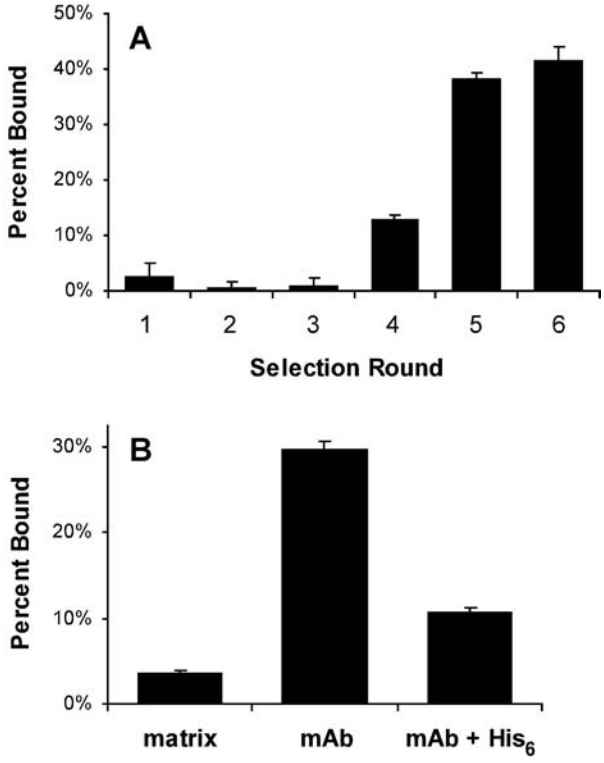

A
Percent Bound
50%
40%
30%
20%
10%
0%
1
2
3
4
5
6
Selection Round
B
Percent Bound
30%
20%
10%
0%
matrix
mAb
mAb + His6

Figure 2

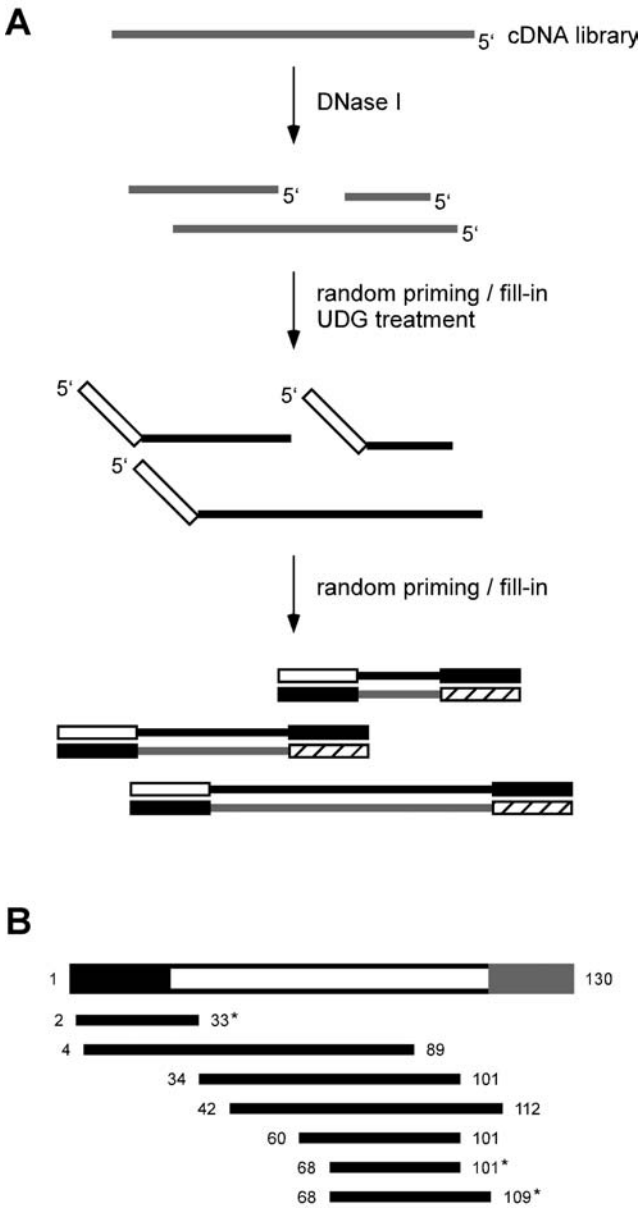

A
5' cDNA library
DNase I
5'
5'
5'
random priming / fill-in
UDG treatment
5'
5'
5'
random priming / fill-in
B
1
130
2
33*
4
89
34
101
42
112
60
101
68
101*
68
109*

Figure 3

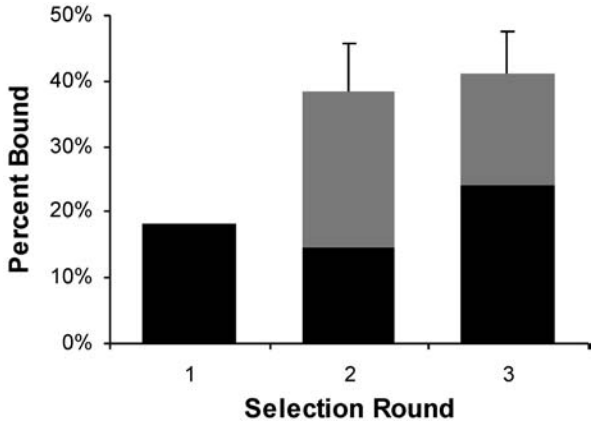


Figure 4

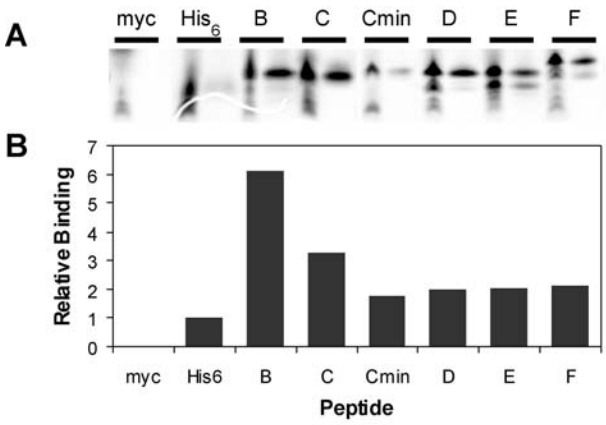

A
myc
His6
B
C
Cmin
D
E
F
B
Relative Binding
0
1
2
3
4
5
6
7
myc
His6
B
C
Cmin
D
E
F
Peptide

Figure 5

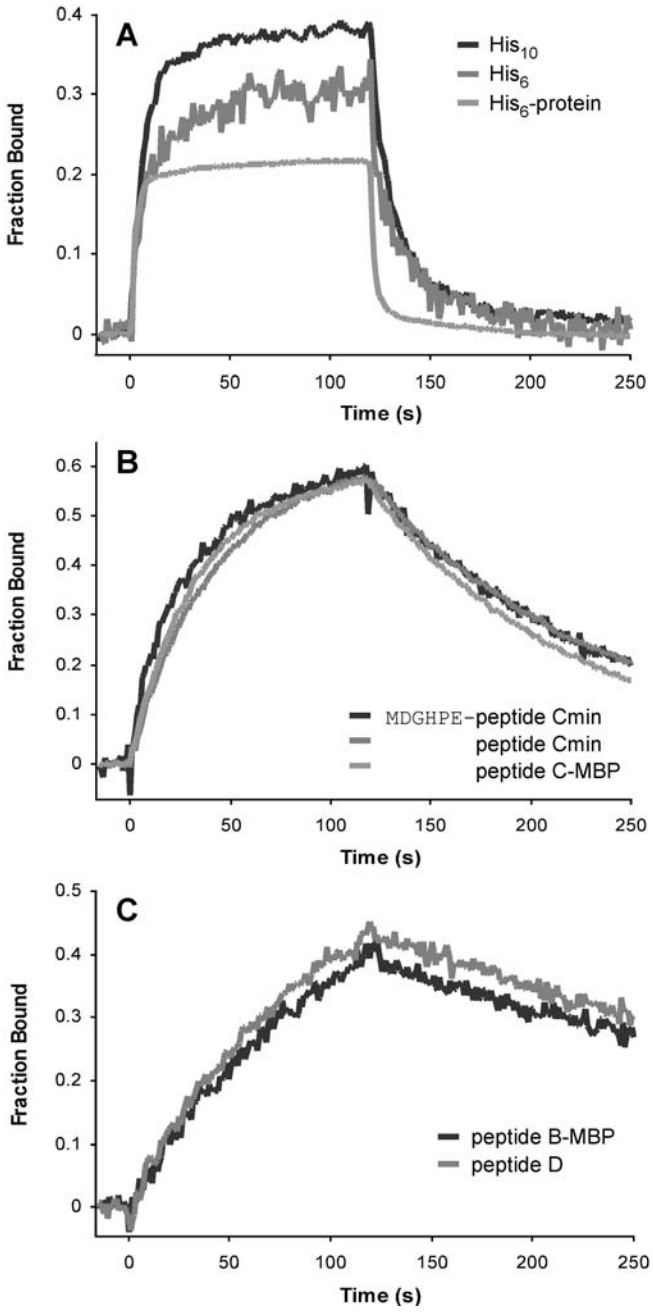

A
His10
His6
His6-protein
Fraction Bound
Time (s)
B
MDGHPE-peptide Cmin
peptide Cmin
peptide C-MBP
Fraction Bound
Time (s)
C
peptide B-MBP
peptide D
Fraction Bound
Time (s)

Figure 6