Chapter 5

Design of a Cyclic Peptide Ligand for Gαi1 Using a Cyclic, Unnatural mRNA Display Library

This chapter is adapted from the following reference (in preparation):

Millward, S.W., Fiacco, S., Austin, R.J., Roberts, R.W. (2007) Combinatorial Gα Ligand Design Using a Trillion-Member Peptide Macrocycle Library *ACS Chem. Biol.*

Abstract

There is a pressing need for new molecular tools to target protein surfaces with high affinity and specificity. Here, we describe "cyclic mRNA display" with a trillion-member covalent peptide macrocycle library. Using this library, we have designed a number of high-affinity, redox insensitive, cyclic peptides that target the signaling protein Gαi1. In addition to cyclization, our library construction took advantage of an expanded genetic code—utilizing nonsense suppression to insert N-methyl phenylalanine (NMF) as a 21st amino acid. The designed macrocycles exhibit several intriguing features. First, the core motifs in all selected variants share an identical context with respect to the macrocyclic scaffold, consistent with the idea that selection simultaneously optimized both the cyclization chemistry and the structural placement of the binding epitope. Detailed characterization of one molecule, cycGIBP (cyclic Gαi Binding Peptide), demonstrated substantially enhanced proteolytic stability relative to the parent linear molecule. Third, and perhaps most important, the cycGIBP peptide binds the target with very high affinity (K_i ~ 1 nM), similar to many of the best monoclonal antibodies and higher than the βγ heterodimer, an endogenous Giα1 ligand. Overall the work provides a general route to design novel, low-molecular-weight, high-affinity ligands that target protein surfaces.

Introduction

Although networks of protein-protein interactions control function inside cells, it is increasingly clear that many of these players cannot be targeted using traditional drug-like molecules (*1*). High-affinity, high-specificity ligands targeting protein surfaces are thus of considerable interest as tools for chemical genetics, as potential lead/surrogate compounds, and as new drugs. Nanotechnology could also greatly benefit from such molecules, particularly robust, inexpensive, low-molecular-weight ligands that could replace monoclonal antibodies (*2*). Designing these ligands still remains a significant challenge despite the tremendous advances in structural biology and computational chemistry.

Our laboratory has been working to design new peptide ligands that target heterotrimeric G-protein signaling. We have previously used mRNA display (*3*) to isolate new linear peptides that target Gαi1 (*4-6*). Gαi1 is a member of the Gα subunit family, and serves as a molecular router—connecting the cell-surface G-protein coupled receptors (GPCR) to downstream effector pathways [see (*7*) for a review of GPCR signaling]. Gα subunits function by collaboration with the Gβγ heterodimer and a transmembrane receptor. The Gαβγheterotrimer associates with the cytosolic portion of GPCRs, with GDP bound in the nucleotide pocket of Gα. Extracellular ligand binding to the GPCR results in intracellular exchange of bound GDP for GTP, dissociation of Gα from Gβγ, and activation or inhibition of downstream effectors. Hydrolysis of GTP on the Gα subunit results in reformation of the stable heterotrimer and termination of the signal.

GPCRs are clinically important and a large fraction of marketed drugs target these receptors (*8*). Additionally, a number of disease states are linked to aberrant Gα expression or Gα mutations, including McCune-Albright syndrome; pituitary, thyroid, and adrenal tumors (*9*); and ovarian stromal and adrenal cortex tumors (*10*). The selective modulation of specific Gα subclasses by drug-like compounds would facilitate the dissection of G-protein control circuits and provide starting points for the development of G-protein-targeted therapeutics.

Peptide ligands represent one route to target Gαi1 (Table 5.1). Examples include the GoLoco motif found in RGS domains (Kd = 82 nM) (*5, 11*), the KB-752 peptide isolated by phage display (16 amino acids; K_D = 3.9 μM) (*12*), and a number of peptides isolated by mRNA display (17 amino acids; K_D = 60 nM) (*5*). While these peptides are potent *in vitro* ligands and function upon microinjection inside cells (*6, 13*) they are unlikely to be functional as traditional drugs *in vivo* due to the proteolytic susceptibility and poor cell membrane permeability traditionally associated with linear peptides.

The peptidic products of non-ribosomal peptide synthetases are examples of peptides that can act as drugs (reviewed in (*14*)). Cyclosporin, a cyclic 11-residue peptide produced commercially by the fungus *Beauveria nivea* acts as a potent, orally bioavailable immunosuppressant (*15*). The enhanced proteolytic stability and cell permeability of cyclosporin are likely the result of both unnatural N-methyl amino acids (*16, 17*) and macrocyclization (*18, 19*) . Using this molecule as a guide, we reasoned that ribosomally synthesized peptide libraries could be improved by incorporating 1) cyclic structure and 2) N-methylated amino acids in the backbone. Previously, we demonstrated that the ribosome could be used to construct N-methylated peptides and that these

oligomers were highly resistant to protease degradation (*20*). Disulfide-based cyclization has been used extensively as a structural constraint in phage display libraries (*21, 22*), but is compromised by reduction in the intracellular environment (*23*). More recently, we designed and synthesized a trillion-member cyclic peptide library in mRNA display format using a bis-NHS cross-linking reagent to join the N terminus to an internal lysine residue (*24*).

Here, we describe "cyclic mRNA" display and employ it to design macrocyclic peptide ligand for Gαi1. These libraries were constructed with an expanded genetic code, using N-methyl phenylalanine (NMF) as the 21st amino acid. The resulting cyclic peptides, such as cycGIBP (cyclic Gαi Binding Peptide) bind Gαi1 with monoclonal antibody-like affinity (K_i = 1 nM), the highest reported affinity for a small peptide ligand to a Gα subunit. This work presents a potentially general route to design cyclic peptide ligands against a broad range of protein targets.

Peptide/ Protein	Sequence	MW (Da)	K_D/K_I (nM)	T (° C)	Ref
L19 GPR (GoLoco motif)	TMGEEDFFDLLAKSQSKRLDDQLVDLAGYK	3433.8	82	25	(5)
KB-752	SRVTWYDFLMEDTKSR	2034.3	3900	25	(12)
R6A	MSQTKRLDDQLYWWEYL	2275.6	60	25	(5)
AR6-05	DESDPEELMYWWEFLSEDPSS	2591.7	10	25	(6)
linGIBP	MITWYEFVAGTK	1444.7	29	4	
cycGIBP	*cyclo*[MITWYEFVAGTK]	1540.8	1	4	
Gβγ	Heterodimer	~43000	15	25	(33)

Table 5.1: Natural and selected peptides that bind to Gα subunits. The conserved core motif sequence is shown in red.

Results and Discussion

Selection for Binding to Gαi1

We began with a library of the form MX10K, where X represents a random position encoded by an NNG/C codon. We had previously shown that a library of this general form could be cyclized with good efficiency by linking the N terminus to the fixed lysine side-chain via the bis-NHS reagent disuccinimidyl glutarate (DSG) (*24*). This work also demonstrated that it was possible to construct a trillion member macrocyclic library containing an unnatural amino acid. In the present work, we combined these features to generate a trillion member library containing N-methyl phenylalanine as the 21st amino acid.

This library was translated in the presence of amber-suppressor NMF-tRNA^{CUA} and cyclized with DSG (Figure 5.1). The random region was constructed with tandem NNS repeats, resulting in UAG (amber) stop codons at a frequency of ~3% at each of the random positions. The estimated diversity of the round 0 pool was 1.67×10^{12} unique sequences with 3-fold oversampling. The library was panned against immobilized Gαi1 and the tightly binding sequences amplified to generate the DNA for the subsequent round. The results of the selection are shown in Figure 5.2. The binding of each pool to the immobilized target began to increase in the fourth round and continued to increase until round 7. At this point, the library was sequenced and the results shown in Figure 5.2B. As can be seen from the sequence alignment, glutamic acid and phenylalanine are 100% conserved at positions 6 and 7 respectively. Tryptophan is found at position 4 in seven of the eight sequences. Threonine or serine residues are found at position 3 while hydrophobic residues (Leu, Ile, or Val) are found at position 8. R7.3 and R7.4 are

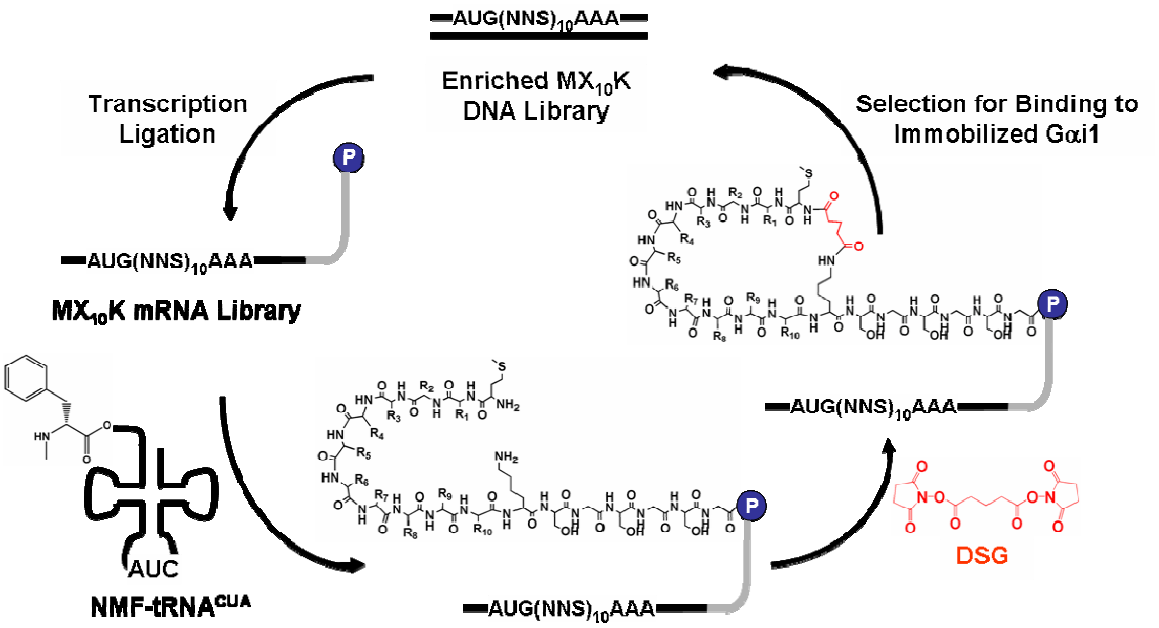


Figure 5.1: Schematic of the selection cycle described in this work. The MX10K library was transcribed from synthetic DNA and ligated to a short DNA linker (shown in grey). The resulting template was translated in rabbit reticulocyte lysate in the presence of THG73 amber suppressor tRNA chemically aminoacylated with N-methyl phenylalanine (NMF-tRNA^{CUA}). The resulting mRNA-peptide fusion library was purified by dT-cellulose and cyclized with disuccinimidyl glutarate (DSG). Following a second dT-cellulose purification, the resulting library was reverse transcribed and incubated with immobilized Gαi1. Members of the library which bound to the immobilized target were eluted with SDS and the corresponding sequences were amplified and enriched by PCR.

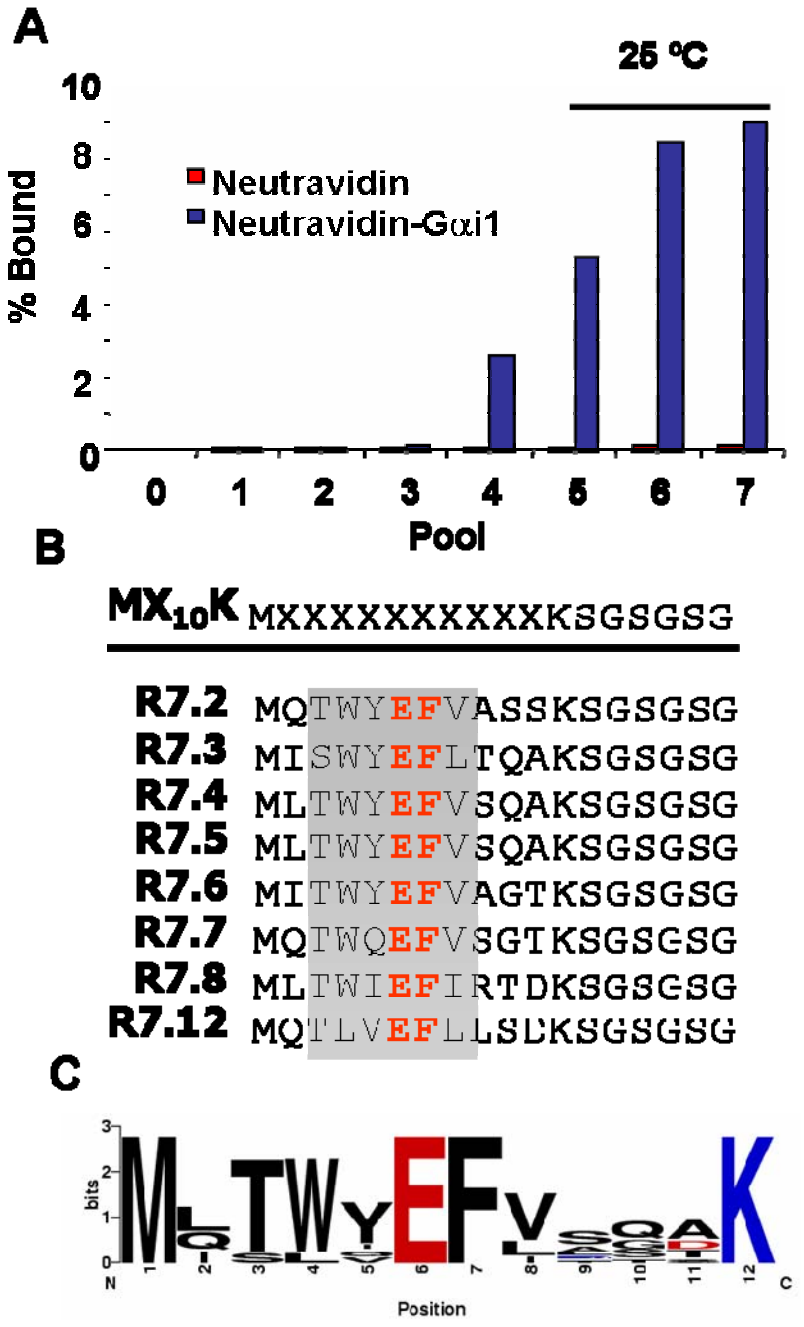


Figure 5.2: Selection against Gαi1 using the MX10K Library. (A) The progress of the selection was monitored by measuring the fraction of radiolabeled library bound to NeutrAvidin-immobilized Gαi1 (*blue bars*) or NeutrAvidin alone (*red bars*). All binding studies were carried out in 1X selection buffer at 4 °C (see materials and methods). Initial selection rounds were carried out at 4 °C, rounds 6 and 7 were performed at room temperature. (B) Peptide sequences of clones obtained from pool 7. Residues that are conserved in all the clones appear in red. The core motif is shaded in grey. (C) WebLogo (*25*) representation of Pool 7 consensus sequence.

identical and represent the only sequence to appear twice in the round 7 pool. Interestingly, the TWYEFV sequence appears at four times in pool 7 and represents a conserved core motif similar to that observed in previous selection experiments (Table 5.1 and references therein).

Positions 9-11 show little consensus, although at least one serine or threonine is found in this region in all of the round 7 clones. This pattern was also observed in clones from the previous round (data not shown). This may imply the presence of a hydrogen bond donor is required although its position within the sequence may not be strictly enforced. No lysine residues were observed within the randomized region providing initial evidence that cyclization proceeds via cross-linking of the primary amines on the N-terminus and the side chain of Lys12.

It is striking that the core motif remains in the same position within the random region in all of the clones under study. This may indicate that the location of the core motif within the macrocycle is critical for interaction with Gαi1 and that this context has been enforced by the selection process.

To rapidly screen the clones for Gαi1 binding activity, five pool 7 sequences were converted to radiolabeled mRNA-peptide fusions and the RNA removed to simplify the binding analysis (Figure 5.3). The fusions were cyclized with DSG or mock treated with DMF. In all of the clones, treatment with DSG resulted in dramatically higher binding efficiencies. The ratio of binding efficiency between DSG-treated fusions and DMF-treated fusions ranged from 5-fold (R7.3) to 52-fold (R7.2). The average binding efficiency of the fusions treated with DSG ranged from 2.3% (R7.12) to 8.4% (R7.6).

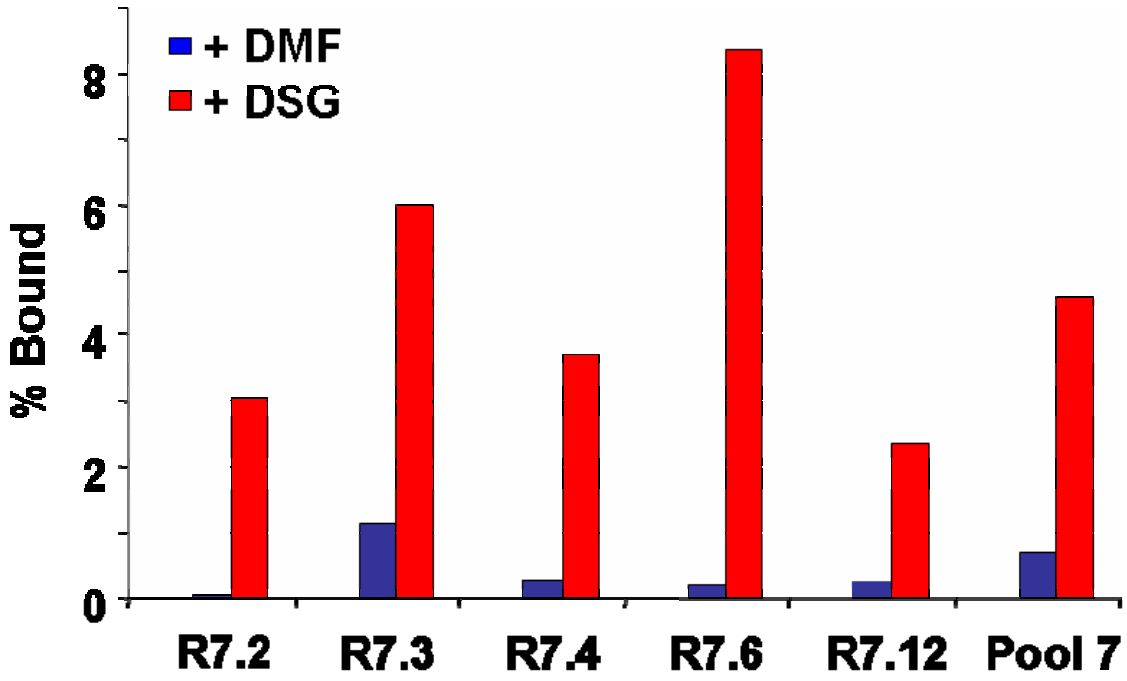


Figure 5.3: Binding of radiolabeled pool 7 peptide fusions to immobilized Gαi1. Following translation, the RNA was removed by treatment with RNase cocktail (Ambion). Following cyclization, fusions were treated with either DMF (*blue bars*) or DSG (*red bars*) for 1 hr and assayed for binding.

The average binding for the individual DSG-treated fusions was found to be 4.7%, in excellent agreement with the binding efficiency for the DSG-treated round 7 pool (4.8%).

DSG-treated R7.6 was the most active fusion with a binding efficiency of 8.4%, a 38-fold enhancement relative to the DMF-treated fusion. This clone was selected for further characterization. R7.12, the only clone to have a leucine instead of a tryptophan at position 4, showed the lowest binding efficiency. This residue has been observed in other Gαi1-binding peptides (Table 5.1 and references therein) which argues for its importance for high-affinity binding. R7.2, R7.4, and R7.6 contain the same core motif (TWYEFV) yet the differences in binding efficiency implicate the importance of the non-conserved residues (position 2 and positions 9-11). These residues may increase or decrease the efficiency of cyclization or may play a more direct role in the binding between the macrocycle and the Gαi1 target.

None of the clones from round 7 contained a TAG stop codon, indicating that NMF was selected against in the context of the cyclic peptide. Examination of the sequences in pool 7 might suggest that NMF could have been incorporated at position 7 in place of phenylalanine, or in the random region where almost no sequence conservation is observed. However, examination of the crystal structure of the closely related peptide KB-752 bound to Gαi1 (*12*) shows the backbone amide at this position is situated to make either of two intramolecular hydrogen bonds to maintain the α-helical character of the core motif. Incorporation of the N-methylated variant would oblate these interactions and presumably destabilize the binding conformation of the peptide.

Characterization of linGIBP-Bio and cycGIBP-Bio

The R7.6 peptide was synthesized as linGIBP-Bio (MITWYEFVAGTKGG-Bio) to allow for immobilization and subsequent pull-down experiments (Figure 5.4A). Cyclization of linGIBP-Bio to form cycGIBP-Bio was performed in 4:1 100 mM phosphate (pH=8):DMF; conditions similar to those used in the cyclization of the bulk MX10K library during the selection (Figure 5.4B). As can be seen in Figure 5.4C, the reaction proceeded smoothly in 1 hr at room temperature without noticeable formation of side products.

linGIBP-Bio and cycGIBP-Bio were immobilized and used to determine binding to radiolabeled Gα subunits (Figure 5.5A). For all subunits, the binding efficiency with cycGIBP-Bio was dramatically increased relative to the linear peptide. cycGIBP-Bio bound to members of the Gαi class with almost identical efficiencies. This is not surprising given that Gαi1, Gαi2, and Gαi3 share >85% sequence identity at the amino acid level (*26*). cycGIBP-Bio showed no binding to Gαq in contrast to R6A-1 which showed significant binding to this subunit in previous experiments (*4*). The cyclic peptide did show significant binding to Gαs, consistent with previous data obtained with R6A-1 (*4*). From these experiments we conclude that cycGIBP-Bio preferentially binds to subunits in the Gαi subclass with specificity comparable to Gαi-binding peptides previously selected with mRNA display. The linear and cyclic peptides were also found to specifically recognize the GDP-bound form of Gαi1 versus the GTPγS and GDP-AlF_3 forms indicating nucleotide-state specificity (Figure 5.5B).

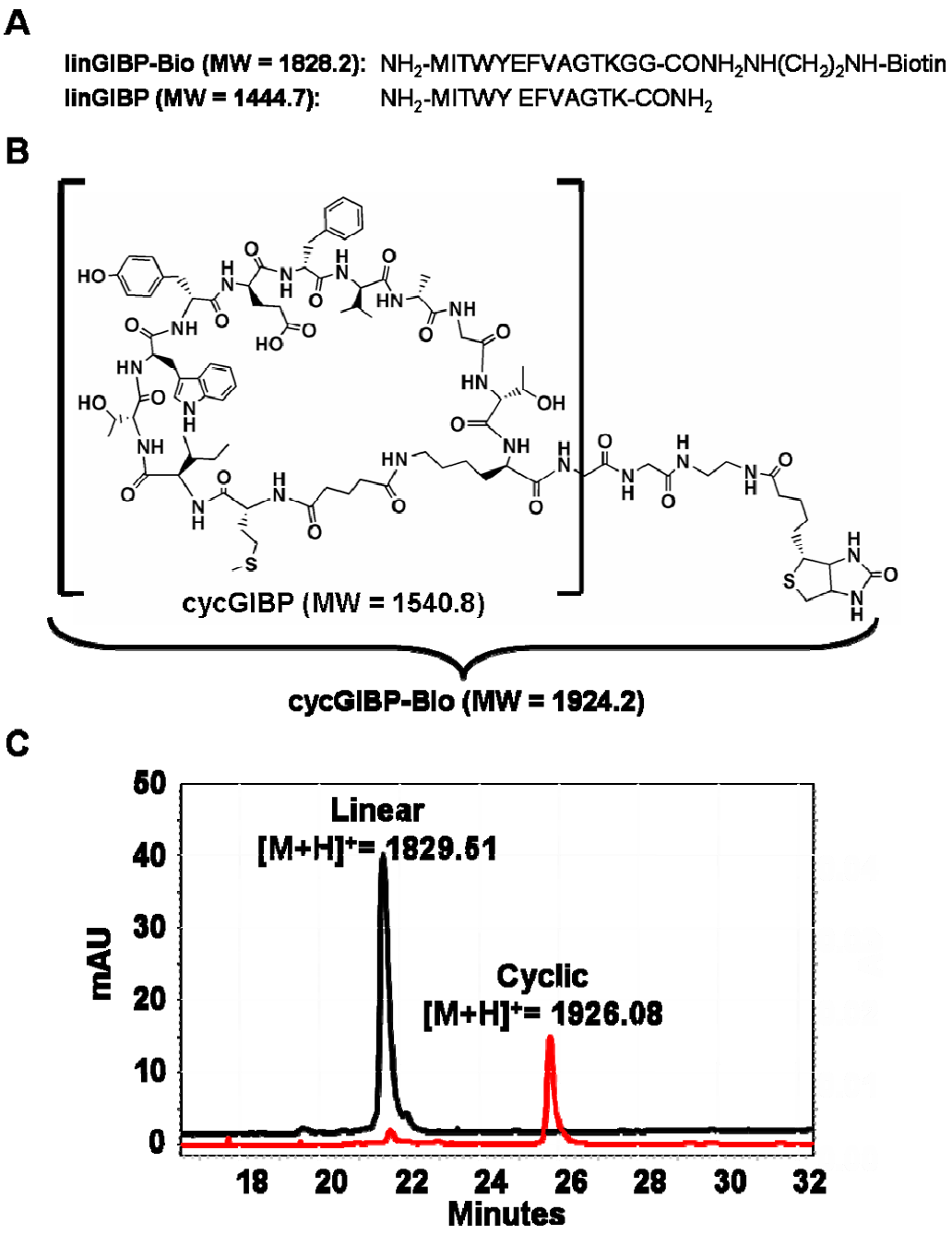


Figure 5.4: Design and synthesis of peptides derived from R7.6. (A) The peptide sequences of linGIBP and linGIBP-Bio. (B) The structures of cycGIBP and cycGIBP-Bio. (C) HPLC Traces of the DSG-mediated cyclization reaction of the linGIBP-Bio peptide. At t=0 (*black trace*) and t=60 min (*red trace*) a portion of the reaction was separated on a C-18 reverse phase column and monitored by absorbance at 280 nm. Peaks were collected, lyophilized, and analyzed by MALDI-TOF MS.

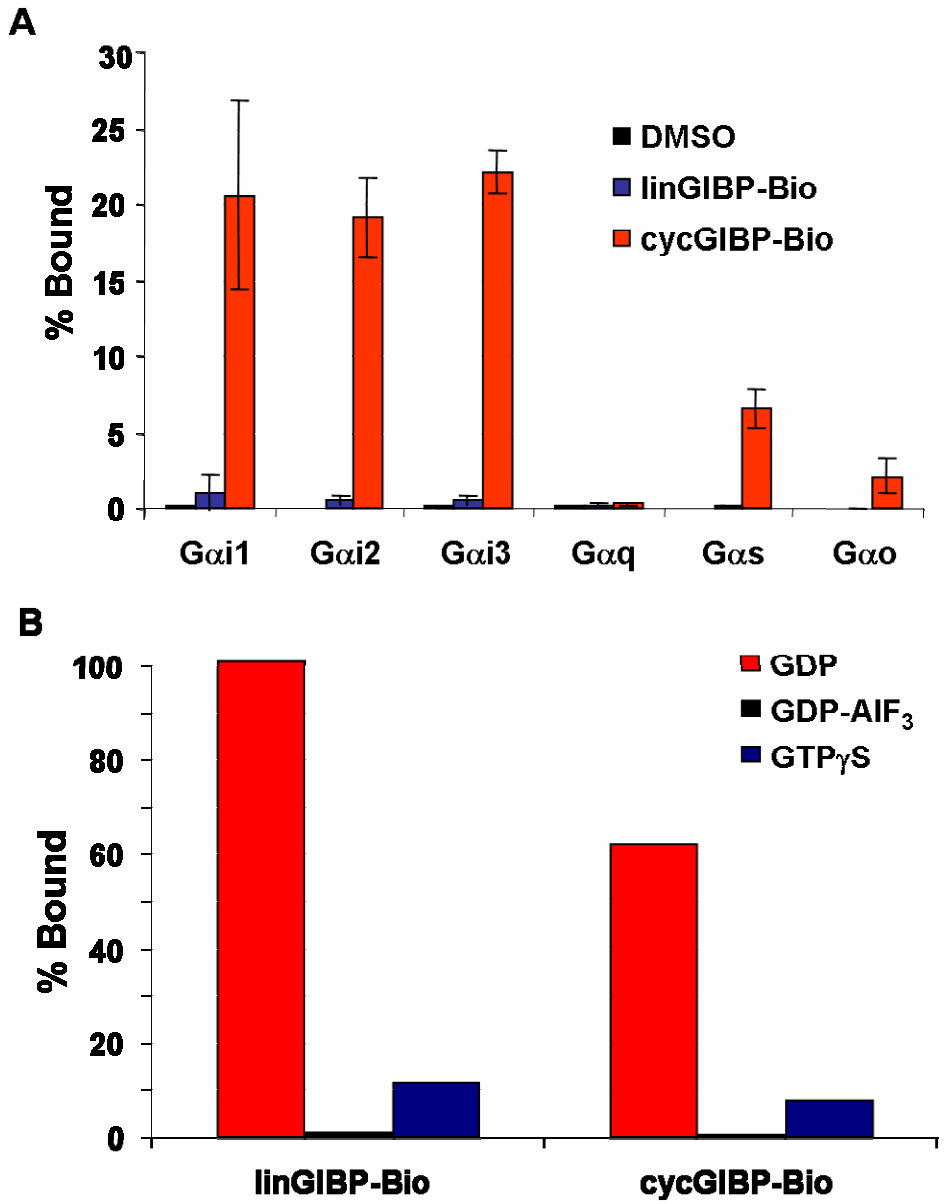


Figure 5.5: Binding of radiolabeled Gα subunits to immobilized R7.6 peptides. (A) NeutrAvidin-agarose was pre-treated with DMSO (*black bars*), linGIBP-Bio (*blue bars*), or cycGIBP-Bio (*red bars*) and incubated with various radiolabeled Gα subunits. The % bound values represent the mean value from three experiments and the error bars represent the standard deviation. (B) Nucleotide state specificity of GIBPs. linGIBP and cycGIBP were immobilized on NeutrAvidin and incubated with radiolabeled Gαi1 pre-bound with GDP, GDP-AlF_3, or GTPγS.

Synthesis and Characterization of cycGIBP

linGIBP, a derivative of linGIBP-Bio lacking the C-terminal biotin and glycine spacer was synthesized. Poor water solubility of linGIBP prompted cyclization to be carried out in DMF with DIEA as the organic base to generate cycGIBP (Figure 5.4B).

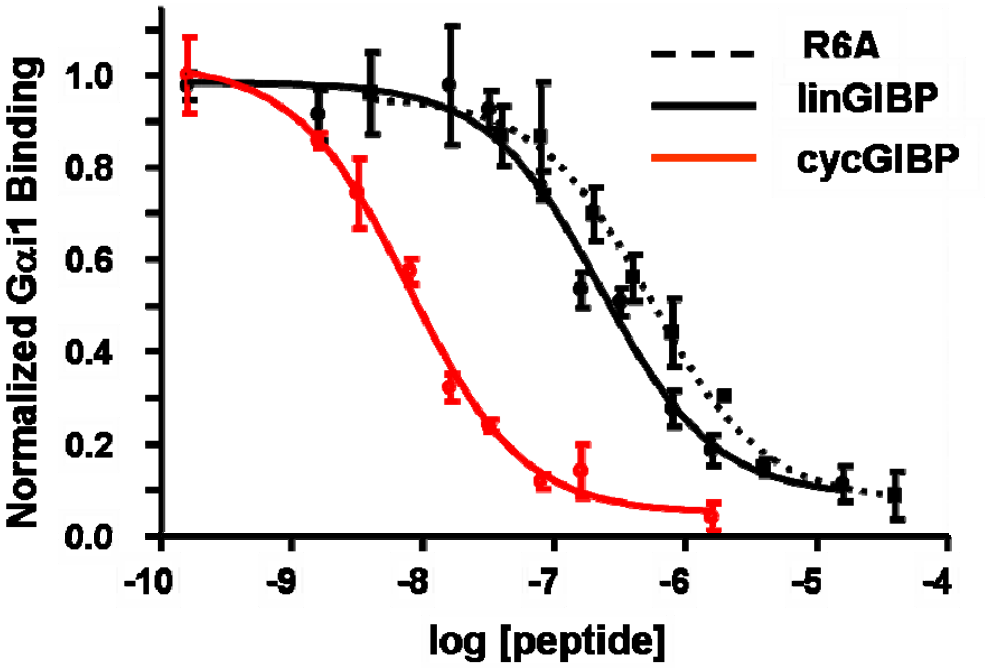
We next sought to determine if cycGIBP could compete with R6A for Gαi1 binding. N-terminally biotinylated R6A was immobilized on NeutrAvidin-agarose and varying concentrations of free peptide (linGIBP, cycGIBP, or free R6A) were added, along with radiolabeled Gαi1 (Figure 5.6). The results of the equilibrium binding experiment show that linGIBP and cycGIBP compete with the R6A for binding to Gαi1, thus supporting a common binding site. The K_i for linGIBP binding is approximately 2-fold lower than that of free R6A while the K_i for cycGIBP is 29-fold lower than its linear counterpart. It is striking that linGIBP has a higher affinity than R6A for Gαi1, given its smaller sequence length and molecular weight. linGIBP and R6A were selected under similar conditions and have homologous core motifs. It is possible that the core motif in linGIBP (TWYEFV) is more optimal for Gαi1 binding and that this core motif failed to appear in the previous selections which were carried out under significantly lower initial diversity.

To evaluate the energetic effect of cyclization, the K_i values can be used to calculate the change in the free energy of binding between the linear and cyclic forms of linGIBP. The $\Delta\Delta G^o$ is given by (*27*)

$$\Delta\Delta G^o = -RT[\ln(K_{i,\,cyc}/K_{i,\,lin})]$$

where R is the universal gas constant and T is the temperature in Kelvin. Substitution of the K_i values from Figure 5.6B gives a $\Delta\Delta G^o = 1.8$ kcal mol^{-1}. This value can also be

A



B

Peptide	K_i (nM)
linGIBP	29 +/- 5 nM
cycGIBP	1 +/- 0.1 nM
R6A	70 +/- 12 nM

Figure 5.6: (A) Inhibition of radiolabeled Gαi1 binding to immobilized R6A peptide. Radiolabeled Gαi1 was incubated with immobilized biotinylated R6A and varying concentrations of either R6A (---), linGIBP (—), or cycGIBP (—) for 2 hr at 4 ºC and the radiolabel bound was normalized to that which bound in the absence of ligand. The plotted values represent the mean of three experimental values and the error bars correspond to the standard deviation. (B) Experimentally determined Ki values for the three peptides shown above. The error is calculated based on the standard error of the mean (SEM) for the log EC_{50}.

interpreted as the conformational entropy lost upon binding of linGIBP to Gαi1. We can also calculate $\Delta\Delta G^{o}$ in terms of the number of independent rotatable bonds in the cycle to normalize this value for comparison to other cyclic molecules in the literature. For a cyclic molecule with N rotatable bonds, there are N-6 independent rotatable bonds (*28*). In the case of cycGIBP, there are 32 rotatable bonds and therefore 26 independent rotatable bonds. From this we can calculate $\Delta\Delta G^{o}$ = 0.07 kcal mol^{-1} per independent rotatable bond. In previous work with small cyclic peptide inhibitors of penicillopepsin, this value was found to vary between 0.4-1.4 kcal mol^{-1} per independent rotatable bond (*29, 30*). The lower value obtained for cycGIBP may indicate higher conformational flexibility in the cyclic variant relative to the conformational flexibility of small peptidomimetic inhibitors. The choice of cyclization chemistry for cycGIBP introduces 3 methylene carbons in the DSG linker and 4 methylene carbons in the lysine side chain, corresponding to 9 rotatable bonds relative to the backbone of the linear peptide. The use of a shorter cross-linker and judicious choice of lysine derivatives may be useful in reducing the conformational flexibility of future peptide macrocycles.

Finally, we determined the effect of cyclization on the proteolytic stability of GIBP. The linear and cyclic peptides were incubated with immobilized chymotrypsin at room temperature for 40 minutes and the amount of starting material was measured at various time points by RP-HPLC (Figure 5.7C). Cyclization increased the half-life from 3.7 min to 9.7 min. Previous work with small cyclic peptidomimetic inhibitors of HIV protease reported complete resistance to cathepsin D and pepsin A degradation over the course of 1 hr (*29*). However, for larger cyclic peptides it is not unreasonable to expect considerable proteolytic digestion within a 15 min time scale (*31*). In addition to

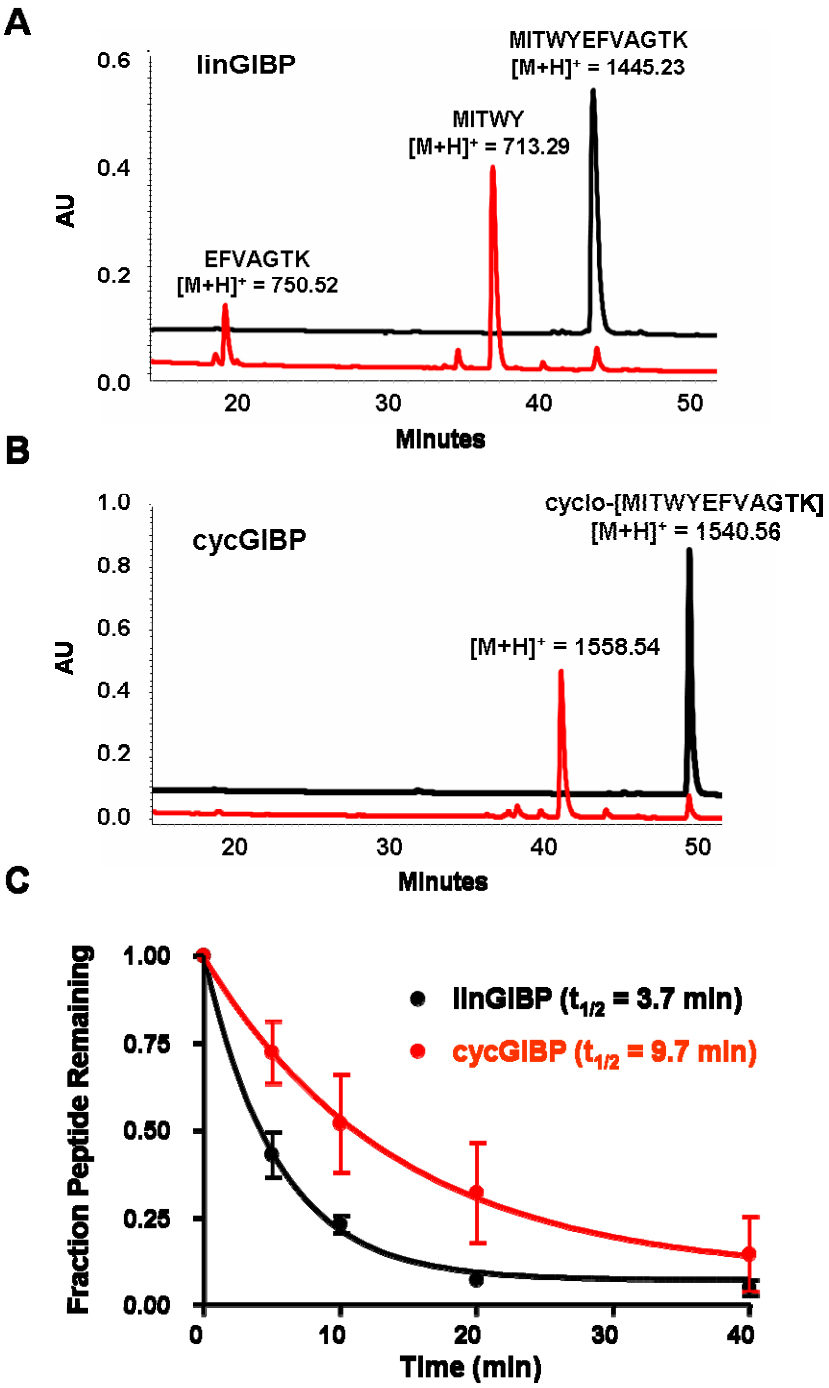


Figure 5.7: Proteolysis of linGIBP and cycGIBP. linGIBP (A) or cycGIBP (B) were incubated with chymotrypsin for 60 min. The major products at t=0 and t=60 min were collected and analyzed by MALDI-TOF MS. (C) Time course of chymotrypsin digestion. linGIBP or cycGIBP were incubated with immobilized α-chymotrypsin-argarose. The amount of starting material was quantitated at various times by C18 RP-HPLC. The plotted values represent the mean of three experimental values and the error bars represent the standard error of the mean.

demonstrating increased cyclic peptide stability, protease digestion of cycGIBP revealed a single product with $[M+H]^+$ =1558.5, indicating the addition of 18 Da (Figure 5.7B). This corresponds to the hydrolysis of a single bond within the macrocycle. Proteolysis of the linear peptide revealed two products with masses consistent with hydrolysis of the Tyr5-Glu6 bond (Figure 5.7A). This provides additional evidence of a cyclic structure for cycGIBP.

Conclusion

In this work we present the first selection of a post-translationally cyclized peptide from a biological display library. The selected cyclic peptide showed specificity for the GDP-bound state of the Gαi1 and moderate specificity for the Gαi class. Competition experiments showed cycGIBP bound to Gαi1 with extremely high affinity (K_i = 1 nM) and that cyclization resulted in a 29-fold improvement in binding affinity relative to the linear peptide. The cyclic peptide was also shown to have increased proteolytic stability relative to its linear counterpart. Taken together, cycGIBP represents an excellent lead compound for additional medicinal chemistry and *in vivo* studies. In the future, the use of cyclic, unnatural mRNA display libraries may enable the selection of high-affinity, high-specificity ligands with increased stability for biological applications.

Materials and Methods

The pdCpA dinucleotide was a provided by Neurion Pharmaceuticals (Pasadena, CA).

Synthesis of N-methyl phenylalanine-tRNA (NMF-tRNA^{CUA})

N-methyl, N-nitroveratrylcarbonyl phenylalanine cyanomethyl ester

The synthesis of N-methyl, N-nitroveratrylcarbonyl phenylalanine cyanomethyl ester was carried out according to the published protocol (*20*) with minor modifications. The final product was purified by silica gel chromatography in 1:1 EtOAc:Hexanes. Yield = 187.5 mg (74%). Analysis by low resolution ESI-MS: expected $[M+Na]^+$ = 479.15; observed $[M+Na]^+$ = 479.6.

N-methyl phenylalanine-tRNA

The synthesis of N-methyl, N-nitroveratrylcarbonyl phenylalanine-dCA was carried out according to previous protocol. Yield = 0.5 mg (5.5%). Analysis by low resolution ESI-MS: expected $[M-H]^-$ = 1035.2; observed $[M-H]^-$ = 1035.2.

Following ligation to THG-73 tRNA (*32*), deprotection of the NVOC group was effected by photolysis with a xenon lamp equipped with a 315 nm cutoff filter. Following deprotection, the N-methyl phenylalanine-tRNA was immediately added to the translation reaction.

Synthesis of MX10K Library

Antisense MX10K ssDNA template was synthesized at the Keck Oligonucleotide Synthesis Facility (Yale) and the sequence is shown below.

5'…GCCAGACCCCGATTTSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNCATTGTAATTGTAAATAGTAATTG…3'

N = A,T,C,G; S = G,C

The reagent bottle used for the "N" positions was made by mixing A:C:G:T in the ratio 3:3:2:2. The reagent bottle for "S" positions was made by mixing C:G in a 3:2 ratio.

MX10K library dsDNA was amplified by 5 cycles of PCR using the forward primer Gen-FP (5'-TAATACGACTCACTATAGGGACAATTACTATTTACAATTACA-3') and the reverse primer MK10K-RP (5'…ACCGCTGCCAGACCCCGATTT…3'). The Round 0 mRNA pool was generated by T7 runoff transcription and purified by Urea-PAGE. The purified mRNA was ligated to F30P (5'-$dA_{21}[C_9]_3$dAdCdC-P; C_9=triethylene glycol phosphate (Glen Research), P= puromycin (Glen Research)), a flexible DNA linker containing puromycin via. an oligonucleotide splint (MX10K-splint: 5'…TTTTTTTTTTTTTACCGCTGCCAGAC…3'). Following PAGE purification of the ligation reaction, the template was dissolved in water and quantitated by absorbance at 260 nm.

Translation and Cyclization: Pool 0

Translation of the Round 0 pool was aimed at generating an initial complexity of 1.67×10^{12} unique mRNA-peptide fusion sequences with 3-fold oversampling. Accordingly, 300 pmol of MX10K template was translated in rabbit reticulocyte lysate under standard conditions. The translation reaction was supplemented with 60 μg of NMF-$tRNA^{CUA}$ in 1 mM NaOAc (pH=4.5) and ^{35}S-methionine. After 1 hr of translation at 30 ºC, KOAc and MgOAc were added to a final concentration of 600 mM and 50 mM, respectively, and the reactions were placed at -20 ºC for 1 hr.

Translation mixtures were diluted 1:10 in dT Binding Buffer (10 mg/mL dT cellulose, 1M NaCl, 20 mM Tris, 1 mM EDTA, 0.2% Triton X-100, pH=8) and agitated for 1 hr at 4 ºC. The dT cellulose was filtered and washed with dT Wash Buffer (300

mM NaCl, 20 mM Tris, pH=8). The DNA-peptide conjugates were eluted with 10 mM NH_4OH and ethanol precipitated in the presence of linear acrylamide (Ambion).

The Round 0 pool was cyclized by adding 190 μL of dT-purified fusions in 50 mM phosphate buffer (pH=8) to 50 μL of DSG (1 mg/mL in DMF). The reaction was allowed to proceed for 1 hr and the fusions re-purified by dT-cellulose and ethanol precipitated.

Translation and Cyclization: Pools 1-7

60 pmol of library template was translated in the presence of 12 μg of NMF-$tRNA^{CUA}$ in 1 mM NaOAc (pH=4.5). Cyclization was performed by adding 40 μL dT-purified fusions in 50 mM phosphate buffer (pH=8) to 10 μL of DSG (1 mg/mL in DMF). The reaction was allowed to proceed for 1 hr and the fusions re-purified by dT-cellulose and ethanol precipitated.

Selection: Round 0

Following ethanol precipitation of the DSG-treated fusions, the pellet was dissolved in 100 μL dH_2O (0.005% Tween-20) and reverse transcribed with Superscript II RNase H^- under standard conditions. Following reverse transcription, the library was added to 5 mL of 1X Selection Buffer (25 mM HEPES-KOH (pH=7.5), 150 mM NaCl, 0.05% Tween-20, 1 mM EDTA, 5 mM $MgCl_2$, 1 mM β-mercaptoethanol, 10 μM GDP, 0.05% BSA, 1 μg/mL tRNA) containing 50 μL Gαi1-NeutrAvidin argarose (pre-blocked with biotin) (*5*). The binding reaction was carried out at 4 ºC for 1 hr. The reaction was filtered and the resin was washed four times with 1X Selection Buffer and twice with 1X Selection Buffer (-BSA, -tRNA) at 4 ºC. The library was eluted with 0.15%SDS at room

temperature and the SDS was removed from the sample using SDS-Out (Pierce). Following ethanol precipitation, the library was amplified by 15 cycles of PCR.

Selection: Rounds 1-7

Selection was carried out according to the protocol described above except that the binding reaction was carried out in 1 mL of Selection Buffer containing 6 μL Gαi1-NeutrAvidin argarose. PCR amplification of the eluted library members was carried until a 96 bp band was observed on a 4% argarose gel (11-14 cycles). The binding for each round was determined by adding 5 μL radiolabeled library to 2 μL Gαi1-NeutrAvidin argarose in 800 μL 1X Selection Buffer. After 1 hr at 4 ºC, the resin was washed as described above and analyzed by scintillation counting to determine the fraction of counts bound.

Pool 7 Analysis

The Round 7 pool was amplified by 12 cycles of PCR and purified using the MinElute PCR Purification Kit (Qiagen). The purified PCR product was subcloned into the PCR 4-TOPO vector (Invitrogen) followed by transformation into TOP10 competent cells (Invitrogen). Individual clones were sequenced using M13F and M13R primers (Laragen). Five clones (R7.2, R7.3, R7.4, R7.6, R7.12) were further amplified and used to generate templates for mRNA-peptide fusion formation as described above. Following translation in the presence of ^{35}S-methionine, fusions were treated with RNase Cocktail (Ambion) for 15 min at 37 ºC and purified by dT-cellulose and ethanol precipitation. 16 μL aliquots of fusion (in 50 mM phosphate buffer) were treated with either 4 μL DMF or 4 μL DSG (1 mg/mL) in DMF and allowed to react at room temperature for 1 hr. The

reactions were quenched with NaOH, neutralized with HCl, and added directly to 900 μL 1X Selection Buffer containing 3 μL of Gαi1-NeutrAvidin argarose (pre-blocked with biotin). After 1 hr at 4 ºC, the solid resin was washed twice in 1X Selection Buffer and once with 1X Selection Buffer (-BSA, -tRNA). The counts remaining on the Gαi1-NeutrAvidin-argarose were assayed by scintillation counting. The bulk library from Round 7 was treated as described above and included as a positive control.

Synthesis of linGIBP-Bio and cycGIBP-Bio

linGIBP-Bio (MITWYEFVAGTKGG-Biotin) was synthesized by manual solid-phase peptide synthesis with 250 mg of Biotin NovaTag resin (Novabiochem). Couplings were carried out with 5 eq of monomer in 2 mL coupling reagent (0.3 M HATU, 0.6 M HOAt in DMF) with 314 μL DIEA at room temperature for 10-20 min. Fmoc deprotection was carried out with 20% piperidine at room temperature for 10 min. Following deprotection, cleavage, filtration and ether extraction, the crude product was purified on a Vydac C-18 reverse phase column using gradient elution (0% B for 5 min, 10-50% B in 30 min. Solvent A: 0.1% TFA, Solvent B: CH_3CN (0.035% B)). Lyophilized solid was quantitated by absorbance at 280 nm (ε_{280}=6970 L mol^{-1} cm^{-1}). Yield = 1.3%. Analysis by MALDI-TOD MS: expected $[M+H]^+$ = 1827.9; observed $[M+H]^+$ = 1830.1.

The cyclization of linGIBP-Bio was carried out by adding 80 μL DSG (4 mM in DMF) to linGIBP-Bio (320 nmol in 232 μL100 mM phosphate buffer). The reaction was allowed to proceed for 3 hr at room temperature then quenched with 6 mL 0.1% TFA. R7.6-Bio was purified by C-18 RP-HPLC chromatography with gradient elution (0% B for 5 min, 20%-60% B in 40 min. Solvent A: 0.1% TFA, Solvent B: CH_3CN

(0.035% B)). Lyophilized solid was quantitated by absorbance at 280 nm (ε_{280}=6970 L mol^{-1} cm^{-1}). Yield = 6.1%. Analysis by MALDI-TOD MS: expected $[M+H]^+$ = 1923.9; observed $[M+H]^+$ = 1925.3.

Synthesis of linGIBP and cycGIBP

Synthesis and quantitation of linGIBP (MITWYEFVAGTK-$CONH_2$) was carried out as described above except that 500 mg of Rink Amide Resin (0.34 mmol capacity, Novabiochem) was used as the solid phase. Yield = 2.6%. Analysis by MALDI-TOF MS: expected $[M+H]^+$ = 1444.7; observed $[M+H]^+$ = 1446.0.

2.8 μmol linGIBP was dissolved in 2.1 mL DMF and 3.4 μL DIEA (19.6 μmol, 7 eq). To this was added 77 μL DSG (40 mM in DMF, 1.1 eq). The reaction was stirred at room temperature for 3hr and poured into 10 mL 4:1 dH_2O (0.1% TFA):CH_3CN (0.035% TFA). cycGIBP was purified by C18 RP-HPLC chromatography under the conditions described above. Yield = 32%. Analysis by MALDI-TOD MS: expected $[M+H]^+$ = 1540.7; observed $[M+H]^+$ = 1542.0.

Specificity Analysis of GIBP-Bio

TNT pulldown experiments were carried out as described in (*4*) with some modifications. The DNA clones for Gαi1, Gαi2, Gαi3, Gαq, Gαs(s), and GαoA were translated into ^{35}S-Met labeled proteins using a TNT coupled transcription/translation system (Promega). Translation reactions were desalted using G-25 spin columns (Amersham) and the translation yield was calculated by TCA precipitation. Equal amounts of each labeled subunit were added to 8 μL of NeutrAvidin-agarose containing 14 pmol of pre-bound linGIBP-Bio or cycGIBP-Bio in 600 μL 1X Selection Buffer (+BSA, -tRNA). Control NeutrAvidin-agarose was treated with DMSO alone. Binding

reactions proceeded for 1 hr at 4 ºC followed by filtration and washing with 1X Selection Buffer (+BSA, -tRNA). The matrix was analyzed by scintillation counting and the percent bound was determined by the matrix counts divided by the total counts as determined by TCA precipitation.

R6A-1 Competition/Equilibrium Binding

R6A (MSQTKRLDDQLYWWEYL) and Bio-R6A (Bio-MSQTKRLDDQLYWWEYL) was synthesized by previously described methods. This peptide was immobilized to NeutrAvidin beads (Promega) using the manufacturer's instructions.

Gαi1 was expressed using the TNT reticulocyte lysate system (T7 promoter, Promega) and desalted using MicroSpin G-25 columns (GE healthcare) into modified HBS-EP buffer [10 mM HEPES at pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20 (Tween 20), 8 mM MgCl2, 30 μM GDP, and 0.05% (w/v) BSA]. Specific Activity was measured by TCA precipitation. 400 pmol of immobilized R6A, and 50,000 cpm of Gαi1 were suspended in 1 mL of HBS-EP buffer. Varying quantities of free peptide (R6A-1, linGIBP, or cycGIBP) were added. The final DMSO concentration was adjusted to 0.5%. Samples were rotated for 2 hr at 4 ºC followed by brief centrifugation and supernatant removal. Samples were washed four times by resuspension in modified HBS-EP buffer at 4 ºC, centrifugation, and supernatant removal. After the final wash, the immobilized sample was transferred to scintillation vials and analyzed by scintillation counting. CPM values were normalized by dividing by the average value obtained in the absence of added peptide. Binding analysis was performed by GraphPad Prism 4.0.

Protease Resistance Experiment

linGIBP and cycGIBP (1.8 mM) in DMSO were added to 50 mM sodium phosphate buffer to a final concentration of 145 mM. Immobilized α-chymotrypsin argarose (Sigma) was added to a final concentration of 21.5 mU/mL. The reaction proceeded at room temperature for 40 min with agitation. 100 μL was removed from the reaction at 0, 5, 10, 20, and 40 minutes and diluted in 500 mL 0.1% TFA. The resulting samples were filtered and injected onto a C-18 reverse phase column and separated by gradient elution (0-40% B in 25 min. Solvent A: 0.1% TFA; Solvent B: CH_3CN (0.035% B)). The area under the starting material peak was quantitated using the 32 KaratGold Software package (Beckman). The plotted values represent the mean of three experimental value and the error bars represent the standard error of the mean. The graph was generated by fitting the data to a one phase exponential decay equation (GraphPad Prism).

References

(1) Hopkins, A. L., and Groom, C. R. (2002) The druggable genome. *Nat Rev Drug Discov 1*, 727-30.

(2) Cheng, M. M., Cuda, G., Bunimovich, Y. L., Gaspari, M., Heath, J. R., Hill, H. D., Mirkin, C. A., Nijdam, A. J., Terracciano, R., Thundat, T., and Ferrari, M. (2006) Nanotechnologies for biomolecular detection and medical diagnostics. *Curr Opin Chem Biol 10*, 11-9.

(3) 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*, 12297-302.

(4) Ja, W. W., Adhikari, A., Austin, R. J., Sprang, S. R., and Roberts, R. W. (2005) A peptide core motif for binding to heterotrimeric G protein alpha subunits. *J Biol Chem 280*, 32057-60.

(5) 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*, 9265-75.

(6) Ja, W. W., Wiser, O., Austin, R. J., Jan, L. Y., and Roberts, R. W. (2006) Turning G proteins on and off using peptide ligands. *ACS Chem Biol 1*, 570-4.

(7) McCudden, C. R., Hains, M. D., Kimple, R. J., Siderovski, D. P., and Willard, F. S. (2005) G-protein signaling: back to the future. *Cell Mol Life Sci 62*, 551-77.

(8) Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? *Nat Rev Drug Discov 5*, 993-6.

(9) Weinstein, L. S., Chen, M., Xie, T., and Liu, J. (2006) Genetic diseases associated with heterotrimeric G proteins. *Trends Pharmacol Sci 27*, 260-6.

(10) Radhika, V., and Dhanasekaran, N. (2001) Transforming G proteins. *Oncogene 20*, 1607-14.

(11) Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002) Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. *Nature 416*, 878-81.

(12) Johnston, C. A., Willard, F. S., Jezyk, M. R., Fredericks, Z., Bodor, E. T., Jones, M. B., Blaesius, R., Watts, V. J., Harden, T. K., Sondek, J., Ramer, J. K., and Siderovski, D. P. (2005) Structure of Galpha(i1) bound to a GDP-selective peptide provides insight into guanine nucleotide exchange. *Structure 13*, 1069-80.

(13) Oxford, G. S., and Webb, C. K. (2004) GoLoco motif peptides as probes of Galpha subunit specificity in coupling of G-protein-coupled receptors to ion channels. *Methods Enzymol 390*, 437-50.

(14) Finking, R., and Marahiel, M. A. (2004) Biosynthesis of nonribosomal peptides 1. *Annu Rev Microbiol 58*, 453-88.

(15) Lipinski, C. L., F, Dominy, BW; Feeney, PJ. (1997) Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Advanced Drug Delivery Reviews 23*, 3-25.

(16) Burton, P. S., Conradi, R. A., Ho, N. F., Hilgers, A. R., and Borchardt, R. T. (1996) How structural features influence the biomembrane permeability of peptides. *J Pharm Sci 85*, 1336-40.

(17) Haviv, F., Fitzpatrick, T. D., Swenson, R. E., Nichols, C. J., Mort, N. A., Bush, E. N., Diaz, G., Bammert, G., Nguyen, A., Rhutasel, N. S., et al. (1993) Effect of N-

methyl substitution of the peptide bonds in luteinizing hormone-releasing hormone agonists. *J Med Chem 36*, 363-9.

(18) Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P., and Lokey, R. S. (2006) Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J Am Chem Soc 128*, 2510-1.

(19) Veber, D. F., and Freidinger, R. M. (1985) The design of metabolically-stable peptide analogs. *Trends in Neurosciences 8*, 392.

(20) Frankel, A., Millward, S. W., and Roberts, R. W. (2003) Encodamers: unnatural peptide oligomers encoded in RNA. *Chem Biol 10*, 1043-50.

(21) Giebel, L. B., Cass, R. T., Milligan, D. L., Young, D. C., Arze, R., and Johnson, C. R. (1995) Screening of cyclic peptide phage libraries identifies ligands that bind streptavidin with high affinities. *Biochemistry 34*, 15430-5.

(22) O'Neil, K. T., Hoess, R. H., Jackson, S. A., Ramachandran, N. S., Mousa, S. A., and DeGrado, W. F. (1992) Identification of novel peptide antagonists for GPIIb/IIIa from a conformationally constrained phage peptide library. *Proteins 14*, 509-15.

(23) Gilbert, H. (1995) Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods in Enzymology 251*, 8-28.

(24) Millward, S. W., Takahashi, T. T., and Roberts, R. W. (2005) A general route for post-translational cyclization of mRNA display libraries. *J Am Chem Soc 127*, 14142-3.

(25) Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: A Sequence Logo Generator. *Genome Res. 14*, 1188-1190.

(26) Strathmann, M., and Simon, M. I. (1990) G protein diversity: a distinct class of alpha subunits is present in vertebrates and invertebrates. *Proc Natl Acad Sci U S A 87*, 9113-7.

(27) Khan, A. R., Parrish, J. C., Fraser, M. E., Smith, W. W., Bartlett, P. A., and James, M. N. (1998) Lowering the entropic barrier for binding conformationally flexible inhibitors to enzymes. *Biochemistry 37*, 16839-45.

(28) Guarnieri, F., Cui, W., and Wilson, S. R. (1991) ANNEAL-RING: A New Algorithm for Optimization of Rings Using Simulated Annealing. *J Chem Soc, Chem Commun*, 1542-1543.

(29) March, D. R., Abbenante, G., Bergman, D. A., Brinkworth, R. I., Wickramasinghe, W., Begun, J., Martin, J. L., and Fairlie, D. P. (1996) Substrate-Based Cyclic Peptidomimetics of Phe-Ile-Val That Inhibit HIV-1 Protease Using a Novel Enzyme-Binding Mode. *J Am Chem Soc 118*, 3375-3379.

(30) Smith, W. W., and Bartlett, P. A. (1998) Macrocyclic Inhibitors of Penicillopepsin. 3. Design, Synthesis, and Evaluation of an Inhibitor Bridged between P2 and P1'. *J Am Chem Soc 120*, 4622-4628.

(31) Botti, P., Pallin, T. D., and Tam, J. P. (1996) Cyclic Peptides from Linear Unprotected Peptide Precursors through Thiazolidine Formation. *J Am Chem Soc 118*, 10018-10024.

(32) Saks, M. E., Sampson, J. R., Nowak, M. W., Kearney, P. C., Du, F., Abelson, J. N., Lester, H. A., and Dougherty, D. A. (1996) An engineered Tetrahymena tRNAGln for in vivo incorporation of unnatural amino acids into proteins by nonsense suppression. *J Biol Chem 271*, 23169-75.

(33) Rebois, R. V., Schuck, P., Northup, J. K. (2002) Elucidating kinetic and thermodynamic constants for interaction of G protein subunits and receptors by surface plasmon resonance spectroscopy. *Methods Enzymol 344*, 15-42.