Screening of Cyclic Peptide Phage Libraries Identifies Ligands That Bind Streptavidin with High Affinities

Lutz B. Giebel,*,‡ Robert T. Cass,‡ Daniel L. Milligan,‡ Dennis C. Young,‡ Rafael Arze,§ and Charles R. Johnson§

Department of Cytokine Biology and Department of Combinatorial Chemistry, Arris Pharmaceutical Corporation, 385 Oyster Point Boulevard, Suite 3, South San Francisco, California 94080

Received February 16, 1995; Revised Manuscript Received June 19, 1995[⊗]

ABSTRACT: The screening of combinatorial peptide libraries has emerged as an important tool in the discovery of novel substrates or ligands for enzyme and receptor targets. For example, screening linear peptide libraries using streptavidin as a model receptor system has previously identified many low-affinity peptide ligands, all of which contain the common motif His-Pro-Gln (HPQ). We reasoned that constraining the conformational freedom of linear peptides by cyclization in a library would yield peptide ligands of increased affinity. Three different cyclic peptide libraries were constructed in an M13 phage display system as N-terminal pIII protein fusions. The random peptide sequences were flanked by two cysteine residues, which allows efficient disulfide bond formation and cyclization during phage assembly. These cyclic peptide libraries were screened with streptavidin as the model receptor system. Many sequences, all of which contained the motif His-Pro-Gln (HPQ), were discovered, and in the preceding paper, the structures of complexes of streptavidin-bound cyclic and linear peptides are described (Katz, 1995). Analysis of binding kinetics and affinities demonstrated that the conformationally constrained cyclic peptides bound streptavidin with affinities up to 3 orders of magnitude higher than linear peptides identified in previous library screens. These results demonstrate the potential of screening conformationally constrained peptide libraries for high-affinity novel receptor ligands or enzyme substrates.

The screening of combinatorial peptide libraries has emerged as a promising tool in drug discovery. Traditionally, the pharmaceutical industry has screened large numbers of organic compounds to identify novel drug leads. Combinatorial peptide libraries have vastly expanded the number of compounds available for screening processes, as they typically contain more than 10^7 different peptides. Peptide libraries have been generated in biological systems (Cwirla et al., 1990; Devlin et al.; 1990, Scott & Smith, 1990; Smith, 1985) or by combinatorial chemistry (Geysen et al., 1986; Houghton et al., 1991; Lam et al.; 1991) and have been successfully screened against many biological targets including antibodies and cell surface or intracellular receptors [for recent reviews, see Clackson and Wells (1994) and Hoess (1993)]. Such screens have also lead to the identification of novel enzyme substrates (Matthews & Wells, 1993). Most peptide libraries reported to date have been conformationally unconstrained and have consisted of linear peptides in solution or peptides tethered at one end to a protein or to a solid support (Gallop et al., 1994). Unconstrained linear peptides can assume millions of different conformations, only very few of which may be able to bind to a target receptor. Constraining the conformational freedom and thus decreasing the entropy of unbound peptides in a library may result in isolation of higher-affinity ligands for a receptor target. O'Neil et al. (1992) screened a conformationally constrained cyclic peptide library against the integrin receptor, GPIIb/IIIa, and identified a cyclic peptide containing the well-known motif Arg-Gly-Asp (RGD) that bound the receptor with an affinity 3 orders of magnitude higher than its linear counterparts.

To further examine the utility of screening conformationally constrained peptide libraries, we constructed three different cyclic, disulfide-containing peptide libraries displayed on phage and screened them for binding to streptavidin as a model receptor target. We then made a rigorous comparison of the binding affinities of a wide variety of linear and cyclic peptides, and in the preceding paper (Katz, 1995) describe the crystal structures of representative linear and cyclic peptides bound to streptavidin. Previously, Devlin et al. (1990) have screened a phage library displaying random 15-residue linear peptides, and Lam et al. have screened a synthetic library of random pentapeptides against streptavidin. In both cases, almost all peptides that bound streptavidin shared the common motif His-Pro-Gln (HPQ) with no detectable sequence conservation at either side of the motif. In subsequent studies, Kay et al. (1993) screened a 38 amino acid phage display library, and McLafferty et al. (1993) screened a disulfide-constrained peptide library for streptavidin binders. In both of these studies, HPQ-containing peptides were identified, and in the latter study, treatment of the peptide-expressing phage with 1 mM DTT was found to decrease binding to streptavidin. However, none of the studies described above included an analysis of the binding affinities of the peptides. Subsequently, Weber et al. (1992) determined the streptavidin-binding affinities of two linear HPQ peptides by titration calorimetry. The two linear heptapeptides FSHPQNT and HDHPQNL, derived from Devlin's screen, had low affinities for streptavidin with dissociation constants of 125 and 282 μM, respectively. We screened three phage libraries displaying cyclic peptides of

* To whom correspondence should be addressed. Telephone: 415-737-8575. Fax: 415-737-8590 or 415-737-3560.
‡ Department of Cytokine Biology.
§ Department of Combinatorial Chemistry.
[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1995.

0006-2960/95/0434-15430$09.00/0 © 1995 American Chemical Society

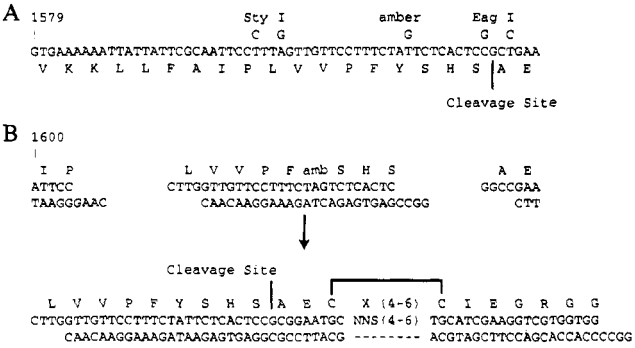

FIGURE 1: Construction of phage libraries. (A) Point mutations introduced at the 5′ end of gene III to create an amber stop codon and the two unique restriction sites, *Sty*I and *Eag*I. (B) Sequence of the 5′ end of gene III in the library construction vector, m13EAST, and the oligonucleotide cassettes used for construction of phage libraries.

different ring size. As observed in previous studies, all of the sequences identified in our screens contained the HPQ motif. Systematic analyses of binding of these cyclic peptides to streptavidin by surface plasmon resonance demonstrated that the binding affinities of cyclic HPQ peptides are 2−3 orders of magnitude higher than those of linear HPQ peptides. The structures of these peptides bound to streptavidin (Katz, 1995) provide clues to the mechanism by which peptide cyclization leads to such a remarkable increase in binding affinity.

MATERIALS AND METHODS

Bacterial Strains. Escherichia coli strains used were as follows: JS5, araD139, Δ(ara, leu)7697, Δ(lac)$_{c74}$, galU, galK, hsdR2(r_k^- m_k-), mcrA, mcrBC, rpsL (Str^r) thi, recA1/F'::Tn10 (tet^r) proAB, lacI^q, lacZ M15; K91Kan, thi/HfrC (Lyons & Zinder, 1972); LE392, el4$^-$, hsdR514, supE44, supF58, lacY, galK2, galT22, metB1, trpR55/F'kan^r.

Construction of Phage Peptide Libraries. Peptides were displayed on the surface of filamentous phage M13 as fusions to the N-terminus of minor coat protein, pIII, using the expression vector M13East. This vector is a derivative of sequencing vector M13mp19 (Yanisch-Perron et al., 1985) and was constructed using the following strategy: A β-lactamase gene was amplified by the polymerase chain reaction (PCR) from pUC19 using oligonucleotide primers 5′-GGGAAATGTGCGCGGAACCC-3′ and 5′-GTATATAT-GAGTAAACTTGG-3′ and cloned into the *Hin*cII site of the M13mp19 polylinker. Using PCR-mediated site-directed mutagenesis (Giebel & Spritz, 1990), the sequence CTTTAG at nucleotide position 1604 (within the pIII leader peptide sequence) of M13mp19 was converted to CCTTGG to create a unique *Sty*I restriction site, and the sequence CCGCTG at position 1631 was converted to CGGCCG to create a unique *Eag*I site (Figure 1A). None of these changes alter the amino acid sequence of pIII. The TAT tyrosine codon at amino acid position 15 of the pIII signal peptide was converted to a TAG amber codon. The final construct M13East was propagated in the *E. coli* host LE 392 (supF) which suppresses the amber codon and restores the original leader sequence. After insertion of an oligonucleotide to construct a library, the amber codon is converted back to the original tyrosine (Figure 1B). Thus, only phage with a productive oligonucleotide insert will grow in a nonsuppressing strain, completely eliminating any vector background.

Our three cyclic libraries displayed the following peptide sequences: AEC-X_6-C (cyclic octapeptide), AEC-X_5-C (cyclic heptapeptide), and AEC-X_4-C (cyclic hexapeptide). The first two amino acids of the mature pIII fusion protein, alanine and glutamic acid (A, E), are identical to the N-terminus of mature pIII protein and ensure efficient cleavage of the signal peptide. There is evidence that cleavage is insufficient if cysteine immediately follows the cleavage site (Shen et al., 1991; L. B. Giebel, unpublished results). Two glycine residues were inserted as a spacer between the cyclic peptide and pIII to prevent interference of the cyclic peptide with correct folding of pIII (Figure 1).

To construct a library, 30 μg of M13East DNA was digested with *Sty*I and *Eag*I and ligated overnight to the appropriate oligonucleotides encoding one of three different cyclic disulfide-linked peptide libraries.

For the cyclic hexa-, hepta-, and octapeptide libraries, these oligonucleotides were (Figure 1C) 5′-CTTGGTTGTTC-CTTTCTATTCTCACTCCGCGGAA-TGC-$(NNS)_{4-6}$-TGC-ATCGAAGGTCGTGGTGG-3′ where N = G, A, T, C and S = G, C. For each library, this oligonucleotide was hybridized to two complementary shorter oligonucleotides, 5′-GCATTCCGCGGAGTGAGAATAGAAAGGAA-CAAC-3′ and 5′-GGCCCCACCACGACCTTCGATGCA-3′ (all with phosphorylated 5′ ends) to produce gapped double-stranded inserts with *Sty*I and *Eag*I compatible ends. The resulting gapped-duplex molecules were ligated into the vector at 5-fold molar excess. The ligation mixtures were ethanol-precipitated and electroporated into *E. coli* strain JS5 as described (Scott & Smith, 1990). For each library, more than 10^8 independent recombinants (measured as ampicillin-resistant colonies) were generated. Libraries were amplified overnight in 2 L of LB media containing 50 μg/mL ampicillin, and PEG-precipitated phage particles were purified by CsCl gradient equilibrium centrifugation.

Screening of Libraries by Biopanning. We used streptavidin-coated paramagnetic particles (Promega, Madison, WI) in the biopanning procedure described previously (Parmley & Smith, 1988). For each library screen, 0.6 mL aliquots of the streptavidin-coated particles were incubated with 10^{12}−10^{13} pfu (plaque forming units) in 400 μL volumes of buffer A (150 mM NaCl, 50 mM Tris-HCl, pH 7.5, and 0.5% Tween 20) at 4 °C. First rounds of biopanning were incubated overnight, and subsequent rounds were incubated for at least 3.5 h. After incubation, particles were washed 8−10 times in buffer A and eluted in 400 μL of elution buffer (1 mg/mL BSA, 0.1 N HCl adjusted to pH 2.2 with glycine). After neutralization with 75 μL of 1 M Tris-HCl, pH 9.1, eluted phage were titered, amplified overnight in the presence of 50 μg/mL ampicillin, and rescreened. All screens were performed in duplicate. After three rounds of panning, individual clones were picked and sequenced, and the peptide sequences were deduced from the DNA sequence of the phage.

Peptide Synthesis. Peptides were assembled on a Millipore 9600 using the *N*-9-fluorenylmethoxycarbonyl/*N-tert*-butoxycarbonyl (Fmoc/*t*-Boc) strategy and acetamidomethyl (Acm) protection for cysteine. Couplings were performed with benzotriazol-1-yloxytris(Pyrrolidino)phosphonium hexafluorophosphate (PyBop)/hydroxybenzotriazole (Hobt) in dimethyl formamide (DMF). Rink (Novabiochem) or Peg-Pal (Millipore) resins were used as the solid support. In general, the peptides were cleaved from the support with

reagent R [90:5:3:2 trifluoroacetic acid (TFA)/thioanisole/ethanedithiol/anisole], precipitated with ether, and chromatographed on a reverse phase column (Vydac C-18; 0.05% TFA in a water/acetonitrile gradient]. The bis(Acm)-Cys peptides were cyclized in 8:1 acetic acid/water with iodine according to the method of Kamber (Kamber et al., 1980). Cyclization was judged to be complete by electrospray mass spectrometry (loss of Acm groups) and a negative Ellman's test. All peptides were acetylated at the N-terminus and synthesized as amides with a C-terminal linker, IEGRK. Thus, the structure of the peptides synthesized is AEC-X_{4-6}-CIEGRK. The side chain amino group of the C-terminal lysine was the sole amine in all of the peptides. This amino group served as the site for specific immobilization on the BIAcore surface.

Measurement of Peptide Binding to Streptavidin by Surface Plasmon Resonance (SPR). The binding of chemically synthesized peptides to streptavidin was measured by surface plasmon resonance (SPR) (Karlsson et al., 1991), using the BIAcore biosensor system (Pharmacia AB, Sweden). CM5 sensor chips and immobilization reagents were supplied by the manufacturer. Unless otherwise noted, the running buffer for all binding measurements by SPR was 10 mM Hepes, pH 7.4, 150 mM NaCl, and 3.4 mM EDTA (HBS) containing 0.005% Tween 20. Peptides were immobilized via the ϵ-amino group of a C-terminal lysine residue. Standard amine coupling, employing *N*-ethyl-*N*'-[3-(diethylamino)propyl]carbodiimide and *N*-hydroxysuccinimide, was performed as described by the manufacturer. For the kinetic measurements and competition assays described below, the surface loading of peptides was lowered empirically to approximately 16 RU (by adjusting the peptide concentration used during immobilization) in order to reduce artifacts due to mass transport and rebinding phenomena (as described in the BIAcore Manual). At this density of peptide on the sensor chips, binding involves only one of streptavidin's four binding sites, as shown by linear $\log(R/R_o)$ vs time plots. Peptide densities of 40 RU or greater are required for multimeric streptavidin binding, as evidenced by nonlinear $\log(R/R_o)$ vs time plots (data not shown). Standard errors of the data from the BIAcore experiments were always less than 8%.

Affinities of peptides for streptavidin were determined by injection of streptavidin at concentrations of 150−500 nM over an immobilized peptide surface and analysis of kinetic and equilibrium binding parameters. K_D values were calculated as the ratio of dissociation (k_d) and association (k_a) rate constants or by Scatchard analysis of equilibrium binding experiments (Felder et al., 1993; Payne et al., 1993). In order to reduce rebinding of streptavidin during measurement of dissociation rate constants, a competing, soluble peptide (identical to the immobilized peptide) was injected at a concentration of 500 μM immediately following the streptavidin injection.

Binding constants of lower-affinity peptides ($K_D > 1$ μM) could not be measured directly by SPR. Therefore, in order to compare the relative streptavidin affinities of the peptides identified by phage panning, we calculated IC_{50} values for each peptide in a BIAcore competition assay using an immobilized peptide of high affinity. In these experiments, streptavidin (167 nM) was preincubated with varying concentrations of a given peptide. This streptavidin−peptide mixture was then injected over a BIAcore surface coupled with an immobilized high-affinity peptide, AEC-HPQGPP-CIEGRK. IC_{50} values were calculated using the program Softmax (Molecular Devices, Mountain View, CA) as the concentration of the competing peptide required to decrease the maximum equilibrium binding by 50%.

Table 1: Important Features of the Cyclic Hexa-, Hepta-, and Octapeptide Libraries[a]

library	independent clones generated	possible AA permutations[b]	possible NT permutations[b]
C-X_4-C	3.8×10^8	1.6×10^5	1.0×10^6
C-X_5-C	2.4×10^8	3.2×10^6	3.3×10^7
C-X_6-C	6.1×10^8	6.4×10^7	1.1×10^9

[a] The genetic code used to construct these libraries was reduced to 32 possible codons described by NNS, where N = G, A, T, or C and S = G or C. The reduced code encodes all 20 natural amino acids plus 1 termination codon. [b] AA denotes amino acid or peptide permutations; NT indicates oligonucleotide permutations.

RESULTS

Three cyclic peptide libraries were constructed by phage display of peptides fused to the N-terminus of M13 gIII protein. The random amino acid sequences were flanked by cysteine residues which allow disulfide bond formation and cyclization of the peptides. Disulfide bond formation during phage assembly is very efficient, and proteins containing multiple correctly formed disulfide bonds can be expressed on the surface of phage (Hoess, 1993). The size and complexity of the cyclic hexa-, hepta-, and octapeptide libraries we constructed are summarized in Table 1. The use of NNS codons (N = G, A, T, or C, S = G or C) in the oligonucleotides encoding the libraries decreases the number of potential codons from 64 to 32. Thus, rare amino acids encoded by a single codon in our libraries will occur, on average, with a probability of 1/32 at each position. Despite this bias, we calculate that the cyclic hexa- and heptapeptide libraries contain a sufficient number of independent clones to allow representation of all possible peptide sequences. On the other hand, the cyclic octapeptide library of 6.1×10^8 independent clones probably does not contain all possible peptides, since it is encoded by a total of 1.1×10^9 nucleotide sequences. In addition, it should be noted that in each library some sequences will probably not be present due to inefficient infectivity of the phage or bacterial intolerance of the peptide−pIII fusion. As a quality control, we selected and sequenced at least 50 random clones from each library. All clones analyzed contained productive and correct inserts. Analyses of the nucleotide sequences of these random clones showed equal distribution of nucleotides at the randomized positions. While sequencing of 50 independent clones is not sufficient for a statistical analysis of amino acid distribution at the randomized positions, we did find each amino acid at each of the randomized positions with the exception of cysteine. We never observed cysteine at any position. The presence of an additional cysteine residue at the randomized positions would probably interfere with correct disulfide bond formation.

The three libraries were screened for streptavidin binding (Table 2). Analysis of sequences after the second round revealed that most but not all sequences contained this motif. After three rounds of biopanning, all clones analyzed contained the HPQ motif. In our screens, the HPQ motif was almost always adjacent to the N-terminal cysteine

Table 2: Streptavidin-Binding Peptide Sequences Isolated from Three Different Libraries

peptide	frequency	library
C-HPQGPP-C	56	C-X_6-C
C-HPQFPL-C	4	
C-HPQFTL-C	2	
C-HPQFNL-C	1	
C-HPQGDR-C	1	
C-HPQFSN-C	1	
C-HPQFRH-C	1	
C-HPQSGR-C	1	
C-HPQFP-C	28	C-X_5-C
C-HPQVP-C	4	
C-HPQFN-C	4	
C-HPQVA-C	3	
C-HPQFM-C	2	
C-HPQFR-C	2	
C-HPQNA-C	1	
C-HPQVS-C	1	
C-HPQFA-C	1	
C-HPQVR-C	1	
C-WHPQF-C	1	
C-HPQF-C	20	C-X_4-C

residue. This differs from the linear peptides found by others in which the positioning of the HPQ motif within the peptide sequence was random: it could be found at either end of the peptide or within the sequence. Apparently, the cyclic peptides are constrained in a manner that requires the HPQ motif to be adjacent to the N-terminal cysteine for optimal display of the amino acid side chains and binding to streptavidin. Furthermore, these results suggest that the peptides displayed on the phage do indeed contain the predicted disulfide bond since otherwise our screens should have yielded results more similar to previous linear peptide screens. The vast majority of peptides contain the motifs HPQG or HPQF. Again, the conformational constraints imposed by the disulfide bond probably do not allow the extensive sequence variation found in linear HPQ peptides.

For each of the three libraries, one HPQ sequence clearly predominates after three rounds of biopanning. C-HPQGPP-C is the most dominant sequence for the octapeptide library, C-HPQFP-C dominates the heptapeptide library screen, and C-HPQF-C is the only sequence found in the hexapeptide library screen. This suggests that the predominant sequence had the highest affinity for streptavidin and therefore successfully competed against the others during biopanning. Furthermore, 8 different DNA sequences encode the 18 clones with the peptide sequence HPQGPP that are present after 3 rounds of the octapeptide screen. This finding suggests that the best peptide sequences are being selected from the library, regardless of their encoding DNA sequences. The most frequently isolated clones in the octa- and heptapeptide libraries contain a proline adjacent to the C-terminal cysteine. Prolines are known to introduce bends or kinks into peptide sequences. The presence of this proline residue may provide a structural motif favorable for display of the HPQ side chains and binding to streptavidin.

The isolation of a single predominant sequence in each of the three cyclic libraries suggests that these sequences were selected because they bind with higher affinity to streptavidin than do other peptides in the library. In order to test this and to compare the binding of linear versus cyclic peptides to streptavidin, representative peptides were synthesized, and their affinities for streptavidin were determined by surface plasmon resonance.

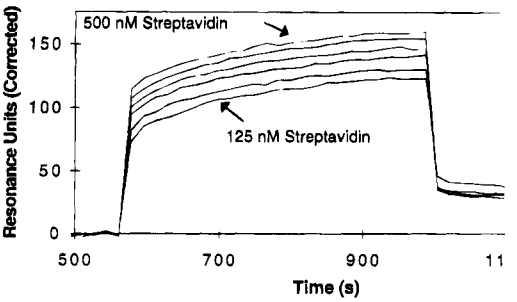

FIGURE 2: Analysis of streptavidin binding by a cyclic peptide by direct surface plasmon resonance measurements. The cyclic peptide AEC-HPQGPP-CIEGRK was immobilized to the BIAcore surface (16 RU), as described in the text. Streptavidin was injected at concentrations ranging from 125 to 500 nM, and the binding affinity was determined by Scatchard analysis of the equilibrium binding data or by kinetic analysis of the on- and off-rates. Free peptide (500 μM) was injected at the end of the streptavidin injection to prevent rebinding of the protein to the immobilized peptide. The surface was regenerated with 6 M GuHCl, pH 2.1, after each binding experiment.

For high-affinity interactions, we calculated the on- and off-rates to determine the K_D (see Materials and Methods). Such direct measurements were made for two cyclic peptides, AE-C-HPQGPP-C-IEGRK and AE-C-HPQF-C-IEGRK. The sensorgrams for the octapeptide AE-C-HPQGPP-C-IEGRK which dominated the cyclic octapeptide library screen are shown in Figure 2. Injection of increasing concentrations of streptavidin leads to corresponding increases in the binding signal observed. The on- and off-rates and equilibrium binding constants (K_D) deduced from these curves are $k_{on} = 8.7 \times 10^5$ M^{-1} s^{-1}, $k_{off} = 0.20$ s^{-1}, and $K_D = 230$ nM. This peptide showed the highest affinity of any of the peptides identified in our screens, an affinity that is 3 orders of magnitude higher than those of the linear peptides FSHPQNT and HDHPQNL, identified in previous screens (Devlin et al., 1990; Weber et al., 1992). Analysis of the cyclic peptide AE-C-HPQF-C-IEGRK yielded $k_{on} = 7.2 \times 10^5$ M^{-1} s^{-1}, $k_{off} = 0.47$ s^{-1}, and $K_D = 660$ nM.

We determined empirically that we could not accurately measure the on- and off-rates for peptides with K_D values higher than 1 μM and thus could not calculate the dissociation rate constants directly. In order to allow comparison of peptides with high and low affinities for streptavidin, IC_{50} values were calculated for all peptides in competition assays using the high-affinity peptide AE-C-HPQGPP-C-IEGRK immobilized on the BIAcore surface. A representative sample of the competition curves is shown in Figure 3, and the IC_{50} values are summarized in Table 3. The IC_{50} values range from 8.1 μM for AE-C-HPQGPP-C-IEGRK (the high-affinity peptide used for the immobilization) to 2.7 mM for the linear FSHPQNT peptide reported previously. The IC_{50} value for the cyclic octapeptide AE-C-HPQGPP-C-IEGRK is 70-fold lower than those of its linear counterparts, AE-S-HPQGPP-S-IEGRK and the linear (ACM-group protected) version of AE-C-HPQGPP-C-IEGRK. Clearly, the conformational constraints imposed by the disulfide bond are important for the high-affinity interaction of this peptide with streptavidin. The best binders of the hepta- and hexapeptide libraries have IC_{50} values of 32.6 and 16 μM, respectively. Comparison of the IC_{50} values of the best octa- and hexapeptides to that of the best heptapeptide suggests that

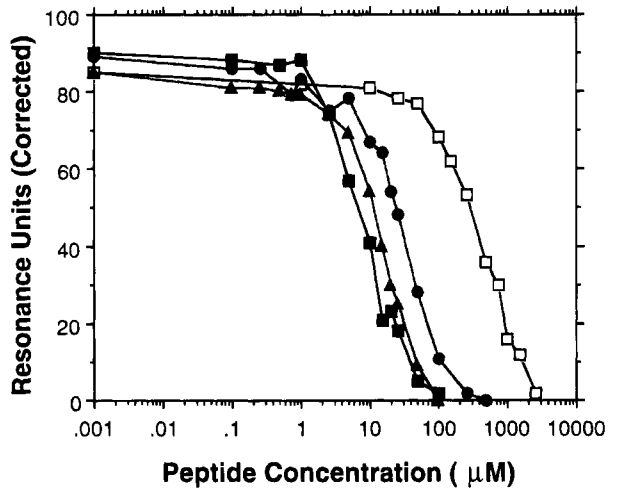

FIGURE 3: Analysis of peptide affinities by binding competition. The cyclic peptide AEC-HPQGPP-CIEGRK was immobilized to the BIAcore surface as described in the legend to Figure 2. Streptavidin (10 μg/mL, 167 nM) was injected alone or with increasing concentrations of each peptide, and IC_{50} values for inhibition of streptavidin binding to the immobilized peptide were calculated using curve-fitting routines in the program SoftMax. The peptide sequences shown below indicate the amino acids between the two cysteine residues; the flanking sequences in each peptide are the same as those of the immobilized peptide: cyclic HPQGPP (■), cyclic HPQF (▲), cyclic HPQFP (●), and linear HPQGPP (□). In each case, the equilibrium plasmon resonance values were corrected by subtracting the value obtained by injection of the same mixture over a blank BIAcore chip.

Table 3: Streptavidin-Binding Affinities of HPQ Peptides[a]

peptide sequence	IC_{50} (μM)	K_D (μM) apparent	K_D (μM)
AEC-HPQGPP-CIEGRK	8.1	0.230	0.230
AES-HPQGPP-SIEGRK	532.0	15.0	ND[c]
AEC[A]-HPQGPP-C[A]-IEGRK	574.0	16.3	ND
AEC-HPQFSN-CIEGRK	14.0	0.400	ND
AEC-HPQFP-CIEGRK	32.6	0.930	ND
AEC-HPQFN-CIEGRK	246.0	7.0	ND
AEC-HPQF-CIEGRK	16.0	0.47	0.660
FSHPQNT	2760.0	78.0	125.0[b]
HDHPQNL	ND	ND	282.0[b]

[a] K_D and IC_{50} values were determined by surface plasmon resonance. Apparent K_D values were derived from IC_{50} values using a calculated conversion factor described under Results. C^A denotes cysteine residues that still retain their ACM protecting group to prevent disulfide bond formation and cyclization of the peptide. [b] The K_D values for the linear peptides FSHPQNT and HDHPQNL were determined by Weber et al. (1992) using microcalorimetry. [c] ND indicates that affinity values have not been determined.

even-numbered rings may display the HPQ motif in a more favorable conformation than odd-numbered ones.

Further comparison of the IC_{50} values reveals that the most prevalent peptide from the octapeptide library screen (AE-C-HPQGPP-C-IERGK) has a 2-fold higher affinity for streptavidin than a peptide found only once in the screen (AE-C-HPQFSN-C-IEGRK). A similar comparison for the heptapeptide screen reveals that the most prevalent sequence (AE-C-HPQFP-C-IEGRK) has a 7-fold higher streptavidin affinity than a peptide found only once (AE-C-HPQFN-C-IEGRK). The biopanning procedure apparently selects for higher affinity sequences over those with lower affinity, despite multivalent display of peptides on the phage surface and the potential for binding avidity to the tetrameric streptavidin molecule.

To further address this point, we "spiked" 10^{12} PFU of the cyclic octapeptide library with 10^4 PFU of a phage encoding the cyclic C-HPQF-C sequence (affinity 660 nM). After 3 rounds of biopanning, we only recovered cyclic octapeptides, most of which (18 out of 20 clones) contained the sequence C-HPQGPP-C (affinity 230 nM). This result was surprising since there is only a 3-fold difference in affinity between the two peptides. Either the biopanning procedure can select for such subtle differences, or the cyclic hexapeptide-displaying phage are lost during biopanning due to significant disadvantages in phage propagation as compared to the octapeptide-displaying phage.

The observed IC_{50} values depend both on the affinity of the peptide for streptavidin and on the density of the immobilized peptide on the biosensor surface. In cases in which the comparison can be made, the observed IC_{50} is considerably higher than the K_D, due to rebinding of streptavidin to the surface. In order to allow comparison of affinities (K_D values) of peptides isolated in this study to those reported in the literature, we calculated a conversion factor (specific for each sensor chip with a given amount of peptide immobilized) to determine apparent K_D values from the observed IC_{50} values (Table 3). This was accomplished by comparing the IC_{50} and K_D values for the high-affinity peptide used for the immobilization, AE-C-HPQGPP-C-IEGRK. The ratio of the two values (IC_{50} = 8.1 μM, K_D = 0.23 μM, conversion factor = 35.2) was used to convert other IC_{50} values to apparent K_D values. In cases in which the comparison can be made, the affinity constants we calculated in this manner agree well with directly measured affinities. For example, direct measurement of the streptavidin affinity of the cyclic peptide AE-C-HPQF-C-IEGRK yielded K_D = 660 nM, while conversion of its IC_{50} value gave an apparent K_D of 470 nM. Analysis of these apparent K_D values indicates that all three of the cyclic peptide libraries we constructed and screened yielded streptavidin ligands with affinities at least 2 orders of magnitude higher than those of previously isolated linear peptides.

In the preceding paper, the crystal structures of the best cyclic hexa- and octapeptide hits bound to streptavidin are described (Katz, 1995). To improve the chances of obtaining crystals of these peptides bound to streptavidin, truncated versions of the peptides, which lack the flanking sequences, were used for crystallization. These peptides, C-HPQF-C and C-HPQGPP-C, have affinities for streptavidin that are similar to the longer versions described above. The affinity of the truncated cyclic hexapeptide, C-HPQF-C (K_D = 270 nM), is higher than that of its longer counterpart (K_D = 470 nM), while the affinity of the truncated octapeptide, C-HPQGPP-C (K_D = 670 nM), is somewhat lower than that of its longer version (K_D = 230 nM).

DISCUSSION

Several groups have previously screened linear peptide libraries using streptavidin (Devlin et al., 1990; Lam et al., 1991; Scott & Smith, 1990). In all cases, the sequences identified contained a common motif: His-Pro-Gln (HPQ), and binding of these peptides to streptavidin was competitive with biotin. In addition, two groups have identified cyclic peptides that bind to streptavidin, but the binding affinities of these peptides were not determined (Kay et al., 1993; McCafferty et al., 1990). Affinities or binding kinetics have

only been reported for two linear HPQ heptapeptides: the two peptides, FSHPQNT and HDHPQNL, bind streptavidin with low affinities of K_D = 125 μM and 282 μM (Weber et al., 1992). The cyclic peptides identified in our constrained peptide library screens are considerably better streptavidin ligands than the previously reported linear peptide sequences. The best streptavidin ligand we identified, the octapeptide AE-C-HPQGPP-C-IEGRK, has an affinity for streptavidin which is more than 2 orders of magnitude higher than those of the linear peptides reported in the literature. The lowest-affinity cyclic peptide we identified and tested, AE-C-HPQFN-C-IEGRK, had an apparent K_D of 7 μM, which is 18 times better than the best of the linear sequences.

Previous cocrystallization of the low-affinity linear peptide FSHPQNT (K_D = 125 μM) and streptavidin indicated that the peptide binds in a turn conformation with the histidine, proline, and glutamine side chains oriented inward at the biotin-binding site (Weber et al., 1992). In the preceding paper, we present a novel structure of the complex of streptavidin and this linear peptide, FSHPQNT, along with the structures of streptavidin bound to two cyclic peptides isolated in this study (Katz, 1995). The structures of these cyclic peptides in the biotin-binding pocket suggest that the increased affinity of our cyclic peptides results from presentation of the HPQ residues in a preordered fashion that is favorable for streptavidin binding. The structures of the peptides bound to streptavidin indicate that a specific H-bond exists between the Nδ1 of the unprotonated His residue and the main-chain amide NH of the adjacent glutamine residue. An additional β-turn-like H-bond exists between a backbone carbonyl and a backbone NH in the bound peptides. An increase in the probability of forming one or both of these H-bonds in unbound cyclic peptides compared to unbound linear peptides may contribute to the higher affinities observed for the cyclic peptides. Cyclization has been shown to increase the propensity for β-turn formation in peptides (Lee et al., 1995; Uma et al., 1993).

We have demonstrated that constraining the conformational freedom of peptides in a combinatorial library screen leads to the identification of peptide ligands with highly improved affinities for the model receptor streptavidin, compared to linear peptides found in previous screens. This concept should be applicable to any receptor target. In the field of medicinal chemistry, the incorporation of conformational constraints into flexible lead compounds is a proven powerful strategy to increase lead potency. It has been shown to be particularly useful in the field of peptidomimetic design (Al-Obeidi et al., 1989; Barker et al., 1992). On the other hand, introducing conformational constraints in a peptide library may eliminate many conformations capable of binding a given receptor target. This effect is likely to be offset by the vast number of peptides (several hundred million) present in our phage libraries. The sequence diversity of the cyclic HPQ peptides isolated in this study was substantially lower than those derived from linear peptide libraries, which is consistent with the notion that the conformational diversity in a cyclic peptide library is reduced. We conclude that screening structurally constrained peptide libraries can produce better ligands for a given receptor. Such libraries could be an important tool for drug discovery, in which initial leads (and their receptor-bound structures) will provide important starting points for the design of small-molecule peptidomimetics.

ACKNOWLEDGMENT

We thank Dr. George P. Smith for *E. coli* strain K91Kan and Dr. David Matthews for valuable discussions while this work was in progress.

REFERENCES

Al-Obeidi, F., De L. Castrucci, A. M., Hadley, M. E., & Hruby, V. J. (1989) *J. Med. Chem. 32*, 2555−2561.
Barker, P. L., Bullens, S., Bunting, S., Burdick, D. J., Chan, K. S., Deisher, T., et al. (1992) *J. Med. Chem. 35*, 2040−2048.
Clackson, T., & Wells, J. A. (1994) *Trends Biotechnol. 12*, 173−184.
Cwirla, S. E., Peters, E. A., Barret, R. W., & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6378−6382.
Devlin, J. J., Panganiban, L. C., & Devlin, P. E. (1990) *Science 249*, 404−406.
Felder, S., Zhou, M., Hu, P., Urena, J., Ullrich, A., Chaudhury, M., White, M. F., Shoelson, S. E., & Schlessinger, J. (1993) *Mol. Cell. Biol. 13*, 1449−1455.
Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, P. A., & Gordon, E. (1994) *J. Med. Chem. 37*, 1233−1251.
Geysen, H. M., Rodda, S. J., & Mason, T. J. (1986) *Mol. Immunol. 23*, 709−715.
Giebel, L. B., & Spritz, R. A. (1990) *Nucleic Acids Res. 18*, 4947.
Hoess, R. H. (1993) *Curr. Opin. Struct. Biol. 3*, 572−579.
Houghton, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T., & Cuervo, J. H. (1991) *Nature 354*, 84−86.
Kamber, B., Hartmann, A., Eisler, K., Riniker, B., Rink, H., Sieber, P., & Rittel, W., (1980) *Helv. Chim. Acta 63*, 899−915.
Karlsson, R., Michaelsson, A., & Mattson, L. (1991) *J. Immunol. Methods 145*, 229−246.
Katz, B. A. (1995) *Biochemistry* (preceding paper in this issue).
Kay, B. R., Adey, N. B., He, Y.-S., Manfredi, J. P., Mataragnon, A. H., & Fowlkes, D. M. (1993) *Gene 128*, 59−65.
Lam, K. S., Salmon, S. E., Hersch, E. M., Hruby, V. J., Kazmierski, W. M., & Knapp, R. J. (1991) *Nature 354*, 82−86.
Lee, M. S., Gardner, B., Kahn, M., & Nakanishi, H. (1995) *FEBS Lett. 359*, 113−118.
Lyons, L. B., & Zinder, N. D. (1972) *Virology 49*, 45−60.
Matthews, D. J., & Wells, J. A. (1993) *Science 260*, 1113−1117.
McCafferty, J., Griffiths, A. D., Winter, G., & Chiswell, D. J. (1990) *Nature 348*, 552−554.
McLafferty, M. A., Kent, R. B., Ladner, R. C., & Markland, W. (1993) *Gene 128*, 29−36.
O'Neil, K. T., Hoess, R. H., Jackson, S. A., Ramachandran, N. S., Mousa, S. A., & DeGrado, W. F. (1992) *Proteins: Struct., Funct., Genet. 14*, 509−515.
Parmley, S. F., & Smith, G. P. (1988) *Gene 73*, 305−318.
Payne, G., Shoelson, S. E., Gish, G. D., Pawson, T., & Walsh, C. T. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4902−4906.
Scott, J. K., & Smith, G. P. (1990) *Science 249*, 386−390.
Shen, L. M., Lee, J. L., Cheng, S., Jutte, M., Kuhn, A., & Dalbey, R. E. (1991) *Biochemistry 30*, 11775−11781.
Smith, G. P. (1985) *Science 228*, 1315−1317.
Uma, K., Kishore, R., & Balaram, P. (1993) *Biopolymers 33*, 865−871.
Weber, P. C., Pantoliano, M. W., & Thompson, L. D. (1992) *Biochemistry 31*, 9350−9354.
Yanisch-Perron, C., Viera, J., & Messing, J. (1985) *Gene 33*, 103−119.

BI950357T