Binding to Protein Targets of Peptidic Leads Discovered by Phage Display: Crystal Structures of Streptavidin-Bound Linear and Cyclic Peptide Ligands Containing the HPQ Sequence†

Bradley A. Katz*

Arris Pharmaceutical Corporation, 385 Oyster Point Boulevard, Suite 3, South San Francisco, California 94080

Received February 16, 1995; Revised Manuscript Received August 2, 1995®

ABSTRACT: The streptavidin-bound crystal structures of two disulfide-bridged cyclic peptides (*cyclo*-Ac-[CHPQGPPC]-NH_2 and *cyclo*-Ac-[CHPQFC]-NH_2) and of a linear peptide (FSHPQNT) were determined, as well as the structure of apostreptavidin (streptavidin−sulfate). Both the linear and disulfide-bridged cyclic peptides studied share a common HPQ conformation and make common interactions with streptavidin, although significant differences in structures and interactions occur for flanking residues among the complexes. The conformation of the linear peptide in the crystal structure of streptavidin−FSHPQNT was found to differ from that in the same complex published [Weber, P. C., Pantoliano, M. W., & Thompson, L. D. (1992) *Biochemistry 31,* 9350−9354]. In the present investigation, the HPQNT portion of the ligand is well-defined with some density defining the Phe, whereas in the investigation of Weber et al. only the HPQ segment of the bound peptide could be interpreted. Both bound cyclic peptides adopt a β-turn involving an H-bond between the His main chain carbonyl and the main chain amide NH of the $i+3$ residue. In the streptavidin-bound *cyclo*-Ac-[CHPQFC]-NH_2 structure, there is an additional H-bond, indicative of α-helix, between the main chain His carbonyl and the main chain C-terminal Cys amide NH group. Binding interactions for both cyclic and linear peptides include direct H-bonds, H-bonds mediated by tightly bound water molecules, and hydrophobic interactions. The above structures and that of streptavidin−biotin in the literature are compared and discussed in the context of structure-based ligand design.

The discovery of drug leads by screening chemical or phage libraries has become an important mandate of many pharmaceutical companies. Technologies involving generation and screening of such libraries are often initially developed in model or validation systems for which reference data from other laboratories are available. One such system that has been utilized here and elsewhere in the discovery of novel, unnatural ligands is streptavidin (Giebel et al., 1995; Devlin et al., 1990; Lam et al., 1991). Streptavidin binds its natural ligand biotin with one of the highest affinities known, and the tight interaction between streptavidin and biotin has been successfully applied in several assay strategies (Savage et al., 1992).

Drug development often involves the design of organic, non-peptide ligands from the sequences and/or the bound structures of natural peptides or proteins that are the normal ligands or substrates for the macromolecular targets. The streptavidin system offers an opportunity for studying the reverse process of developing a peptide ligand from the structure of the protein-bound natural ligand, biotin, which is non-peptide. Probing the relationships between peptide and non-peptide ligands that bind in the biotin binding site of streptavidin might therefore prove useful in drug design in other systems. The discovery of unnatural peptide ligands that bind to streptavidin, and the structural relationship between the streptavidin-bound biotin ligand and a discovered linear peptide ligand, has been the subject of a publication (Weber et al., 1992). Many linear and cyclic peptides discovered here and elsewhere (in libraries comprised of natural amino acids) contain a conserved tripeptide sequence, HPQ.

One general principle of drug development is the potential of enhancing the affinity of a molecule, due to an entropy effect, by rigidifying or constraining it in a way that helps lock it into a conformation in which the binding motif is similar in the unbound state to that in the protein-bound state. For example, this strategy was successful in the design of high-affinity cyclic peptides (Jackson et al., 1994) or rigid non-peptide ligands (McDowell et al., 1994) that bind to the platelet glycoprotein IIb/IIIa. Such a principle has been studied here with many peptide sequences discovered by screening phage libraries for streptavidin binding: the measured affinities for streptavidin of constrained, disulfide-linked peptide ligands were found to be greater by several hundredfold than those of the unconstrained, linear counterparts (Giebel et al., 1995).

There are relatively few crystal structures of small cyclic peptides bound to proteins in the literature. Appraisal of the types of structures that cyclic peptides are likely to adopt when bound to proteins will be important in the eventual conversion of such leads, discovered by screening, into small tight binding organic molecules. Thus, the present study was directed at determining the crystal structures of the discovered disulfide-bridged cyclic peptides bound in the biotin binding site of the streptavidin model system.

† The structures for streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2, streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2, streptavidin−FSHPQNT, and streptavidin−sulfate have been submitted to the Brookhaven Protein Data Bank under Accession Numbers 1SLE, 1SLD, 1SLG, and 1SLF, respectively.

* Address correspondence to this author. Phone: (415) 737-1643. FAX: (415) 737-8590. e-mail: bak@arris.com.

® Abstract published in *Advance ACS Abstracts,* November 1, 1995.

0006-2960/95/0434-15421$09.00/0

The observation that the HPQ-containing cyclic peptides bind to streptavidin with much greater affinity (K_d approximately submicromolar) than their linear counterparts (K_d approximately submillimolar) directed a comparison of the structures and interactions made with streptavidin by both peptide classes, in an effort to probe the structural basis for the affinity differences. Similarly, a comparison of the structures and binding interactions in streptavidin−bound linear and cyclic peptides with those in streptavidin-biotin ($K_d \sim 10^{-15}$ M) was undertaken.

One issue that concerns structure-based drug or ligand design given a ligand-free structure is the degree of flexibility of the protein at the ligand binding site and the type of movement and reorganization in the protein, and associated water molecules that are likely to occur upon complexation. Thus, these liganded streptavidin structures were also compared with a ligand-free streptavidin structure to assess the nature, degree, and extent of reorganization incurred by the binding of the various streptavidin ligands.

EXPERIMENTAL PROCEDURES

Peptide Synthesis. Peptides were assembled on a Millipore 9600 synthesizer using N^{α}-Fmoc-protected[1] amino acids with acid-labile side chain protecting groups, and acetamidomethyl (Acm)-protected cysteines. Couplings were performed with PyBOP/HOBt in DMF. Rink (Novabiochem) or Peg-Pal (Millipore) resins were used as solid support. Peptides were cleaved from the solid support with reagent R (90:5:3:2 TFA/thioanisole/ethanedithiol/anisole), precipitated with ether, and chromatographed on a reverse phase column (Vydac C-18, 0.05% TFA/water/acetonitrile gradient) with a Gilson 715 HPLC system. The peptides with two Acm-protected cysteines were cyclized in 8:1 acetic acid/water with iodine according to the method of Kamber et al. (1980). Disulfide cyclization was shown to be complete by a negative Ellman test (Ellman, 1959). All peptides were characterized by electrospray mass spectra and amino acid analysis (data not shown). All peptides in this study that contain two cysteines have an intramolecular disulfide bridge between the cysteines.

Crystallization and Preparation of Streptavidin−Ligand Complexes. Apostreptavidin was purchased from Calbiochem and crystallized by vapor diffusion in 40 λ sitting drops under conditions described for crystals of space group *I222* (Pähler et al., 1987). Streptavidin−FSHPQNT was cocrystallized in the *I222* space group under similar conditions (pH 4.0) at an initial peptide concentration of 15.0 mg/mL. Crystals of streptavidin−FSHPQNT were soaked in synthetic mother liquor (50% SAS, 50% 0.1 M potassium acetate, and 15 mg/mL peptide) at pH 5.6 before data collection. Streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2 was cocrystallized at pH 4.0, at an initial peptide concentration of 15 mg/mL under conditions that normally produce *I222* diamond plates. However, these setups yielded square plates (space group $I4_122$, $a = 59.19$ Å, $c = 178.26$ Å). A crystal was transferred to synthetic mother liquor containing peptide at a concentration of 15.0 mg/mL, pH 7.5, before mounting and data collection. Streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 crystals were prepared by soaking large apostreptavidin crystals in synthetic mother liquor at pH 5.0 at a peptide concentrations of 10.0 mg/mL for several days. Before being mounted, crystals were transferred to fresh soaking solutions made with freshly dissolved peptide and left equilibrating for several hours. Soaking *I222* apostreptavidin crystals in peptide ligands tended to crack the crystals or induce strain lines, especially at higher pHs. One of several crystals of streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 soaked at pH 5.0 did not crack, and this crystal was used for X-ray crystallography.

Crystallographic Data Collection of Streptavidin−Ligand Complexes. X-ray diffraction data from single crystals of streptavidin−sulfate or streptavidin−peptide complexes were collected on a Siemens IPC area detector coupled to a Siemens three-circle goniometer mounted on a Rigaku rotating-anode target tube operating at 50 kV, 60 mA. Data were indexed and reduced to produce integrated intensities and structure factors with the programs Sadie and Saint supplied by Siemens. The following settings were used: crystal to detector distance = 8.4 or 9.0 cm; $2\theta = -22.5°$, $-25.0°$, or $-30.0°$; oscillation widths in $\omega = 0.10°$ (for streptavidin complexes with sulfate, FSHPQNT, and *cyclo*-Ac-[CHPQFC]-NH_2) or 0.15° for streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2. The exposure time per frame was 150−250 s, and a sufficient number of frames was collected to yield about 6-fold redundancy (Table 1). Typically 3−4 sweeps of 90−110° in ω at φ values of 0.0, 90.0, 45.0, or 180° were collected.

Structure Determinations of I222 Streptavidin−Ligand Complexes. The first complex determined was streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2. A model of streptavidin with no inhibitor or waters at or near the expected binding regions of the peptide was built from the structure of streptavidin−FSHPQNT (Weber et al., 1992), Brookhaven Protein Data Bank (Bernstein et al., 1977) Entry Code 1PTS. Initial electron density maps calculated with coefficients ($|F_o| - |F_c|$) or ($2|F_o| - |F_c|$) and phases α_c for the model eventually enabled determination of the structure of the bound cyclic peptide. Numerous attempts to fit the ligand density to a peptide structure in which the HPQ portion was in a position and conformation similar to those in the published structure of streptavidin−FSHPQNT (Weber et al., 1992) were all unsuccessful. Associated refinement experiments were also unsuccessful. In view of the unexpected differences in this region between the streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 structure determined here and the streptavidin−FSHPQNT structure described in the literature, the structure of streptavidin−FSHPQNT was redetermined in this study.

To compare the conformation of the streptavidin-bound FSHPQNT in the conformation determined in this investigation with the conformation described by Weber et al. (1992), the streptavidin-bound conformation of the HPQ segment of the peptide in the investigation of Weber et al. was built with the aid of Figure 1 and Figure 2a of the paper of Weber et al. The peptide H, P, and Q Cα coordinates (the only FSHPQNT peptide coordinates deposited in the Brookhaven Data Bank by Weber et al.) were also used in the construction to yield an initial structure that was then fit as best as possible

[1] Abbreviations: Fmoc, 9-fluorenylmethoxycarbonyl; Acm, acetamidomethyl; PyBOP, benzotriazol-1-yloxytris(pyrrolidino)phosphonium hexafluorophosphate; HOBt, *N*-hydroxybenzotriazole·H_2O; DMF, dimethylformamide; TFA, trifluoroacetic acid; SAS, saturated ammonium sulfate; HPLC, high-performance liquid chromatography; kV, kilovolt; mA, milliamp; F_o and F_c, observed and calculated structure factors; α_c, calculated phases; HIV, human immunodeficiency virus; rms, root-mean-square.

Table 1: Crystallography of Streptavidin-Bound FSHPQNT, *cyclo*-Ac-[CHPQGPPC]-NH_2, *cyclo*-Ac-[CHPQFC]-NH_2, and Streptavidin−Sulfate (Apostreptavidin)

	FSHPQNT, pH 5.6		Ac-[CHPQGPPC]-NH_2, pH 5.0	Ac-[CHPQFC]-NH_2, pH 7.5	sulfate, pH 5.6
	2.00 Å[a]	1.74 Å			
parameters[b]					
no. of atoms (including disorder)	2130	2130	2075	1042	2055
no. of waters (including disorder)	159	159	150	61	199
no. of discretely disordered groups[c]	8	8	0	5	8
no. of discretely disordered waters	4	4	0	0	8
no. of side chains with refined occs[d]	29	29	24	23	21
diffraction statistics					
resolution (Å)	50−2.00	50−1.74	50−1.90	50−1.92	50−1.74
no. of observations[e]	78608	94347	64474	52950	93159
no. of unique observations	16059	21587	16042	11347	22185
redundancy	4.9	4.4	4.0	4.7	4.2
R_{merge} (%)[f]	6.9	7.2	7.5	9.1	7.3
refinement statistics					
refinement resolution	10.0−2.00	7.5−1.76	7.5−2.00	7.5−2.00	7.5−1.76
no. of merged reflections	15417	20506	14948	9362	21115
$\|F_o\|/\sigma$ cutoff	1.5	1.0	1.0	1.7	1.0
R_{cryst}[g] (%)	17.8	18.7	19.1	19.4	17.6
free R_{cryst}[h] (%)	23.8	22.2	23.3	22.0	20.2
overall completeness (%)	91.6	84.0	84.7	85.5	87.3
and at highest resolution (%)	77.8	53.3	54.4	53.9	54.2
highest resolution shell	2.09−2.00	1.84−1.76	2.09−2.00	2.09−2.00	1.84−1.76
rms deviations[i]					
bond lengths (Å)	0.019	0.019	0.018	0.019	0.018
bond angles (deg)	3.5	3.4	3.2	3.2	3.1
torsion angles (deg)	27.7	27.7	27.9	28.3	27.8

[a] The statistics for the determination of Weber et al. (1992) are as follows: resolution, 2.0 Å; number of reflections, 15 592; percent completion, 92%; *R*-factor, 17.9%; deviation from ideality of bond lengths, 0.025 Å; scaling *R*-factor, 6.9%. [b] Restrained, isotropic temperature factors were refined and bulk solvent contributions included for all structures. [c] Not including waters. [d] Also includes ligand groups. Density for all side chain atoms or for terminal atoms in these groups was weak or absent, and temperature factors were high. Discretely disordered groups are not included in this category. Occupancies for poorly defined groups of atoms were refined. [e] Data with $R_{sym} > 50\%$ were rejected along with data with values $>3.5\sigma$ from the mean for each bunch of symmetry equivalents. [f] $R_{merge} = \sum_h\sum_i|I(h)_i - \langle I(h)\rangle|/\sum_h\sum_i I(h)_i$, where $I(h)_i$ is the *i*th observation of the intensity of reflection *h*. [g] $R_{cryst} = \sum(||F_o| - |F_c||)/\sum|F_o|$ (for reflections from 7.5 Å to the highest resolution). [h] Cross-validation *R*-factor using 10% of the data withheld from the refinement (Brünger, 1992a,b). The somewhat higher *R*-factors for streptavidin−peptide structures than for trypsin−inhibitor structures (Katz et al., 1995) of comparable or higher resolution using the same data collection software and hardware reflect weaker diffraction and a disorder in some of the loops. [i] Root-mean-square deviations from ideal bond lengths and bond angles.

to ($|F_o| - |F_c|$), α_c and ($2|F_o| - |F_c|$), α_c maps. The resulting streptavidin−HPQ segment was crystallographically refined to convergence.

Structure Determination of I4$_1$22 Streptavidin−Cyclo-Ac-[CHPQFC]-NH_2. This tetragonal crystal form of streptavidin has one subunit per asymmetric unit (Weber et al., 1989), and therefore the tetramer must be situated on the origin of the unit cell with the molecular 2-fold axes of the tetramer coincident with the crystallographic 2-fold axes. Of the three potential solutions (calculated by transformation of the *I222* structure), one yielded very good packing and an initial *R*-factor of 41%, whereas the packing of the others was impossible.

Crystallographic Refinement and Analysis of Streptavidin−Ligand Complexes. Models of streptavidin−sulfate or streptavidin−peptide complexes were built with Quanta[2] or Insight[2] and refined with X-PLOR (Brünger, 1992a,b) and with difference Fourier methods (Chambers & Stroud, 1977). In ($|F_o| - |F_c|$), α_c maps, positive and negative peaks whose magnitudes were greater than 2.8σ were systematically identified (using the in-house program Peak-pick) and analyzed. Water structure was determined and refined (Finer-Moore et al., 1992). Included waters were based on significant density in both ($|F_o| - |F_c|$), α_c and ($2|F_o| - |F_c|$), α_c maps in stereochemically reasonable locations, and were generally rejected if their refined temperature factors exceeded 60 Å^2. Deviations of bond lengths from ideality in the final structures are 0.018−0.019 Å (Table 1). For comparison with the determination of the 2.0 Å resolution structure of streptavidin−FSHPQNT by Weber et al. (1992), Table 1 also contains data collection and refinement statistics for the present determination of streptavidin−FSHPQNT when the data are limited to 2.0 Å resolution, and refinement conditions intended to correspond as closely as possible to those of Weber et al. (1992) are used. For comparison of streptavidin structures, all structures were superimposed with matrices determined from corresponding pairs of non-hydrogen atoms involving 48 well-ordered ($B < 20$ Å^2) residues spread throughout the molecule. The structures reported here have been submitted to the Brookhaven Data Bank (*cyclo*-Ac-[CHPQGPPC]-NH_2, 1SLE; *cyclo*-Ac-[CHPQFC]-NH_2, 1SLD; FSHPQNT, 1SLG; sulfate, 1SLF).

[2] Quanta is a molecular modeling program licensed from Molecular Simulations, 16 New England Executive Park, Burlington, MA 01803. Insight is a molecular modeling program licensed from Biosym Technologies, Inc., 10065 Barnes Canyon Rd., San Diego, CA 92121.

RESULTS

I222 Crystals of Streptavidin Complexes Contain Two Crystallographically Independent Ligand Sites. The *I222* streptavidin−peptide structures discussed here have two crystallographically independent subunits per asymmetric unit, each with a bound peptide whose structure was

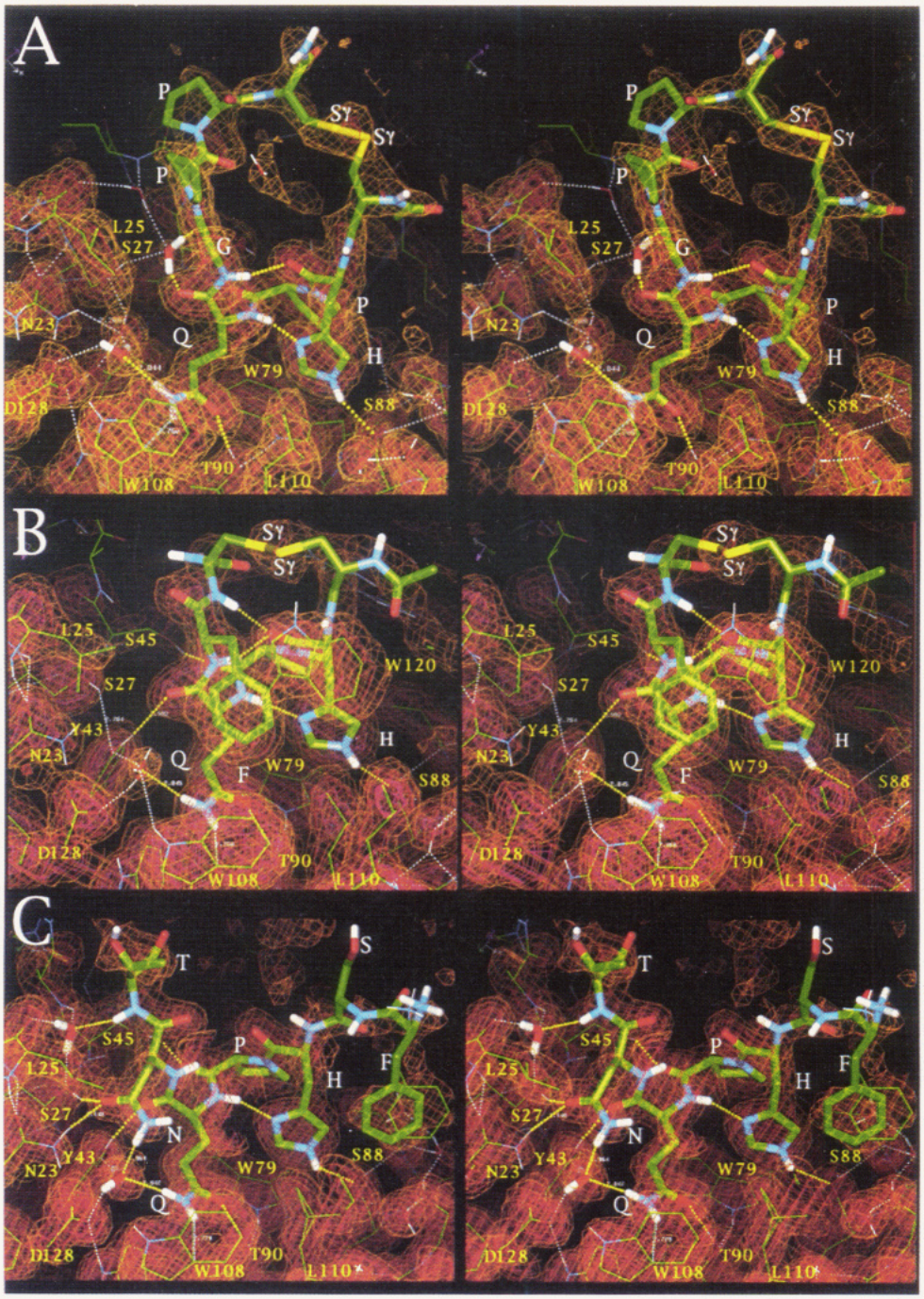

FIGURE 1: (A) $(2|F_o| - |F_c|)$, α_c map for streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 superimposed on the refined structure. The bound peptide is represented by "licorice" sticks. The density in the foreground in front of the peptide Gln residue corresponds to part of one surface of the Trp120 side chain of a neighboring subunit of the tetramer that forms one wall of the binding site. (B) $(2|F_o| - |F_c|)$, α_c map for streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2 superimposed on the refined structure. Streptavidin residues are labeled in white; peptide ligands in yellow. Intrapeptide H-bonds are in yellow. (C): $(2|F_o| - |F_c|)$, α_c map for streptavidin−FSHPQNT superimposed on the refined structure. The density in the foreground near the middle of the figure is due to Trp120 of a neighboring subunit of the tetramer, which for clarity is not shown.

independently determined, refined, and analyzed. For all these streptavidin−peptide complexes, the structure of a particular peptide at one site can be superimposed onto the corresponding structure at the other site with an rms deviation of less than 0.1 Å for the heavy atoms in the HPQ portion. One site (site A) is close to a 2-fold related site from an adjacent tetramer, and is therefore not as open to the solvent as the other site (site B). Density for the bound peptides is better defined at site A, and the temperature factors are lower. Thus, all *I222* structures and associated figures in the paper deal with ligands bound at site A.

Structure of Streptavidin−cyclo-Ac-[CHPQGPPC]-NH_2 Clearly Shows Conformation and Binding Interactions of the Cyclic Ligand. The $(2|F_o| - |F_c|)$, α_c map superimposed on the refined structure of streptavidin−*cyclo*-Ac-[CHPQG-PPC]-NH_2, pH 5.0 (Figure 1A), shows density for the entire cyclic peptide. The average temperature factor for the best defined peptide residues (HPQG) is 16.7 (6.5) $Å^2$. For the more mobile or disordered peptide residues on the surface that do not interact with the protein (Ac-Cys, Pro6, Pro7, and Cys-NH_2), the average temperature factor is 58 (11) $Å^2$. There are good van der Waals interactions between the

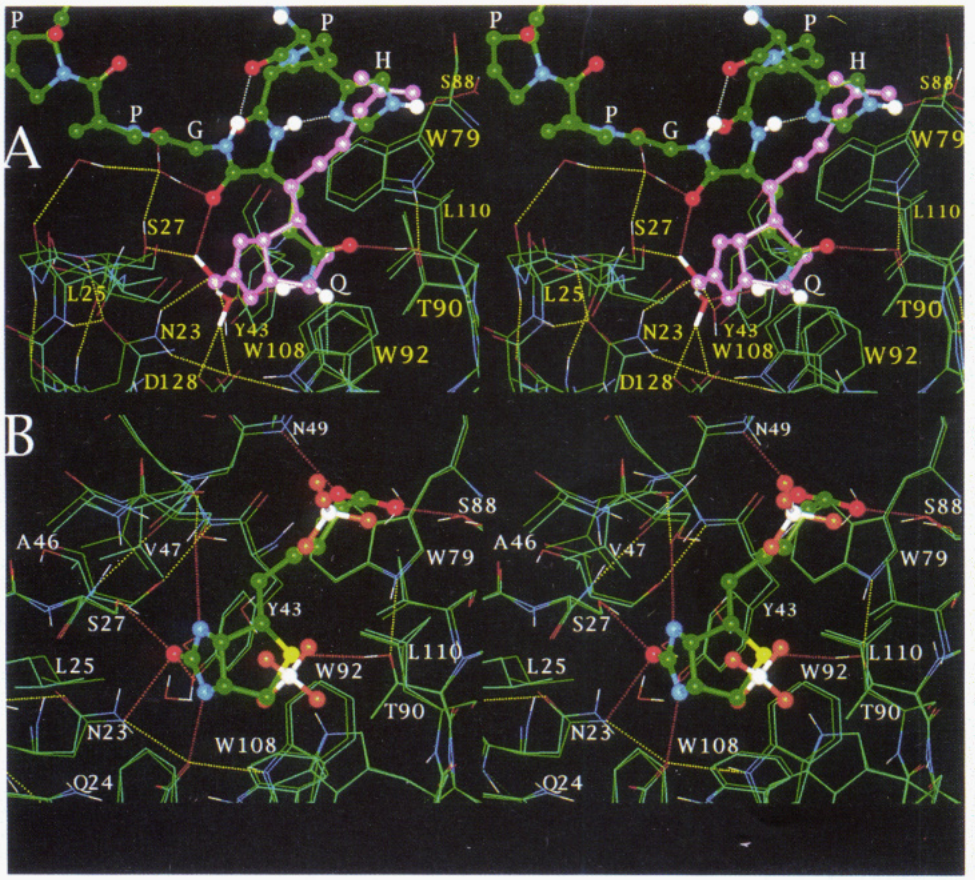

FIGURE 2: (A) Superposition of streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 onto streptavidin−biotin. The carbon atoms for the protein of the streptavidin−biotin structure are in cyan, and the biotin is represented by pink balls-and-sticks. The *cyclo*-Ac-[CHPQGPPC]-NH_2 ligand is represented by balls-and-sticks, and two of the associated waters, discussed in the text, are represented by sticks. For clarity, the flap residues, which are disordered in the streptavidin−peptide complexes, have been removed from both structures. Streptavidin residues are labeled in white, and cyclic peptide residues are labeled in yellow. Hydrogen bonds corresponding to the streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 are shown. Intrapeptide H-bonds are white, protein−peptide H-bonds are red, protein−protein H-bonds are yellow, and the peptide Gln → π interaction with Trp108 is cyan. (B) Superposition of streptavidin−sulfate onto streptavidin−biotin. The carbon atoms for streptavidin−sulfate are cyan, and the sulfate anions are represented by balls-and-sticks, with the sulfurs white. Biotin is represented by balls-and-sticks colored by atom type. Hydrogen bonds corresponding to the streptavidin−biotin structure are displayed.

peptide Gln Cα and Cβ atoms and the Trp79 side chain. The peptide His contacts Leu110 and Trp79, stacking approximately perpendicular to the latter side chain. The Cα of the Gly peptide packs against Trp120 of a neighboring subunit which is part of the tetramer and provides one side of the binding cavity. (For clarity, the Trp120 side chain is not shown in the electron density figure.)

There are Two Intrapeptide Hydrogen Bonds in cyclo-Ac-[CHPQGPPC]-NH_2, One Defining a Type I β-Turn. The H-bonds that mediate binding of the cyclic peptide to streptavidin are shown in yellow in Figure 1A. There is an intrapeptide H-bond between Nδ1 of His and the main chain Gln amide NH. Another H-bond between the Gly main chain amide NH and the His main chain carbonyl defines a type I β-turn within the cyclic peptide. There is an H-bonding interaction between one of the amide NH groups of the peptide Gln side chain and the Trp108 π ring system similar to those described by Burley and Petsko (1986). There are three highly ordered waters that mediate binding of the cyclic peptide to streptavidin. Hydrogen bond distances involving the bound peptide are listed in Table 2.

Figure 2A compares the structure of streptavidin−*cyclo*-Ac-[CHPQGPPC]-NH_2 with that of streptavidin−biotin (Brookhaven Data Base Entry 1STP; Weber et al., 1989). Several contiguously bonded atoms in biotin (labeled 1−4) are nearly superimposable on corresponding atoms (Cβ, Cγ, Cδ, Nϵ2) of the Gln side chain. In the streptavidin−cyclic peptide complex, one ordered water that mediates peptide binding is near the position of the cyclic urea oxygen, while another one is about 1 Å from the lower cyclic urea NH group. There are significant differences in the positions of

Table 2: H-Bonds Involving Streptavidin-Bound *cyclo*-Ac-[CHPQGPPC]-NH_2

(A) Intrapeptide

peptide	peptide	distance (Å)
NH_{Gly}	O_{His}	3.07
NH_{Gln}	Nδ1$_{His}$	2.98

(B) Peptide−Protein

peptide	water	streptavidin	distance (Å)
Nϵ2$_{His}$		Oγ_{Ser88}	3.17
Oϵ1Gln		Oγ1$_{Thr90}$	2.84
Nϵ2Gln	Wat600	Oδ1$_{Asp128}$	3.02
	Wat600	Oδ2$_{Asp128}$	3.15
	Wat600	Nδ2$_{Asn23}$	2.96
	Wat599		3.02
	Wat599	Oγ_{Ser27}	3.02
	Wat599	Oη_{Tyr43}	2.56
	Wat599−Wat600		2.82
O_{Gln}	Wat598		3.20
O_{Gly}	Wat598		2.60
	Wat598	Oγ_{Ser27}	2.93
	Wat598−Wat597		3.03

many side chains participating in the binding interactions in the two complexes.

In Streptavidin−cyclo-Ac-[CHPQFC]-NH_2 the Cyclic Peptide Contains an H-Bond That Defines a Short Segment of α-Helix, as Well as an H-Bond That Defines a Type I β-Turn. In Figure 1B, the $(2|F_o| - |F_c|)$, α_c map for streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2 is superimposed on the refined structure. The average temperature factor for the best-defined peptide residues (HPQF) is 24.1 (6.6) Å^2. For the more mobile or disordered peptide cysteines on the surface that do not interact with the protein, the average temperature factor is 58 (9) Å^2. The position and conformation of the HPQ part of this peptide are very similar to those in streptavidin-bound *cyclo*-Ac-[CHPQGPPC]-NH_2. As in the former complex, there is an intrapeptide H-bond between Nδ1 of His and the Gln main chain amide NH, and an intrapeptide β-turn H-bond between the main chain Phe amide NH and His carbonyl. In addition, there is yet another intrapeptide H-bond between the main chain C-terminal Cys amide NH and His carbonyl, defining a short stretch of distorted α-helix.

There are some protein−peptide interactions in this complex that differ from those in the other complexes. The peptide Phe side chain is sandwiched between the peptide Gln side chain and Trp120 of a neighboring subunit of the tetramer, and also interacts through hydrophobic interactions with the Leu25 and Trp108 side chains. There is a hydrogen bond from the protein Arg84 side chain to the peptide N-terminal Cys main chain carbonyl. The structure of the streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2 complex in the $I4_122$ space group, prepared by cocrystallization, is very similar to that prepared by soaking *I222* crystals (data not shown).

The HPQ Conformation in the Streptavidin-Bound Linear Peptide (FSHPQNT) Determined Here Is the Same as in the Bound Cyclic Peptides. Figure 1C shows the $(2|F_o| - |F_c|)$, α_c map superimposed on the refined structure of streptavidin−FSHPQNT, pH 5.6. Strong density for the HPQN segment is observed, with weaker density for the Phe and Thr, and very little density for the Ser. The average temperature factor for the HPQN segment is 22.0 (4.5) Å^2. For the more mobile or disordered peptide Phe, Ser, and Thr residues on the surface, the average temperature factor is 49 (5) Å^2. The structure of the streptavidin−FSHPQNT complex prepared and determined under the conditions described here shows that the position and conformation of the peptide HPQ segment are the same in the linear peptide as in the bound cyclic peptides (Figure 1A−C).

Like the streptavidin-bound cyclic peptides, the protein-bound linear peptide contains an H-bond between Nδ1 of the His and the Gln main chain amide NH. Unlike the protein-bound cyclic peptides, the streptavidin-bound linear peptide does not contain a β-turn H-bond. However the main chain dihedrals are still characteristic of this type of turn. The more open conformation may help accommodate the interactions made with the peptide Asn. There are two tetrahedrally coordinated waters that mediate binding of the peptide to the protein (Figure 1C). The tetrahedrally coordinated water molecule hydrogen-bonded to Asp128 is in a somewhat different position than in *cyclo*-Ac-[CHPQGPPC]-NH_2 complex, and some of the H-bonds involving it are different. The other water is in a new location.

Some of the van der Waals interactions that the linear peptide makes with streptavidin are similar to those made by the cyclic peptides. Different van der Waals interactions include contacts between the peptide Asn side chain carbonyl carbon and the Leu25 side chain.

The Structure and Conformation of Streptavidin-Bound FSHPQNT Determined Here Differ from Those Described in Another Study. The apparent difference between the structure of the HPQ segment of the streptavidin−cyclic peptide complexes of this study and the structure of the HPQ segment of the streptavidin-bound linear peptide described in the literature prompted a redetermination of the structure of the streptavidin-bound linear peptide at 1.76 Å resolution. The determination in this study was from a streptavidin−peptide cocrystal, whereas the previous determination by Weber et al. was from a streptavidin crystal soaked with the peptide. The quality of the data determined in both investigations at 2.0 Å resolution appears comparable (Table 1).

The structure of streptavidin-bound FSHPQNT in the conformation determined in this investigation (Figure 3A) is compared with that described by Weber et al. (1992) (Figure 3B). Note that since the only FSHPQNT peptide coordinates deposited by Weber in the Brookhaven Data Bank are the Cα coordinates for the H, P, and Q residues, the structure of the streptavidin-bound HPQ segment in the conformation in the investigation of Weber et al. was determined and refined with the data of this study (see Experimental Procedures). The structure presented here differs from the interpretation of Weber et al. in the following aspects. The peptide Asn residue of this structure occupies a region of space similar to that of the Gln in the structure of Weber et al., the peptide Gln of this structure occupies a region of space similar to the His in the structure of Weber et al., and the peptide His of this structure occupies a region of space where no residues have been placed in the structure of Weber et al. (Figure 3B). In this investigation, density clearly defines the Gln and Asn side chains and the linking peptide bond, whereas in the investigation of Weber et al., a proline was inserted between these side chain density features.

Figure 3A,B compares the fit to an unbiased $(|F_o| - |F_c|)$, α_c map of the FSHPQNT peptide in the conformation determined in this investigation with that described in the investigation of Weber et al. The $(|F_o| - |F_c|)$, α_c map was calculated after refinement to convergence of a starting structure in which the peptide was removed from site 1. (The peptide was also removed for the map calculation.) The peptide-free starting structure, for refinement and map calculation, was derived from the complex refined to convergence with the peptide either in the conformation described by Weber et al., or in the conformation determined here; no significant differences were observed between the $(|F_o| - |F_c|)$, α_c omit maps thus calculated.

The conformation determined in this investigation fits the density better. In the conformation described by Weber et al., part of the proline ring and the proline carbonyl are completely out of density, the histidine side chain is too short and not the right shape to optimally fit the density into which it is placed, and the glutamine side chain is too long. Furthermore, more residues of the peptide are accounted for, yielding a significantly lower *R*-factor (18.7%, free *R*-factor = 22.2%) in the structure with the peptide in the conformation determined here compared with 19.7% (free *R*-factor = 25.7%) in the conformation described by Weber et al.,

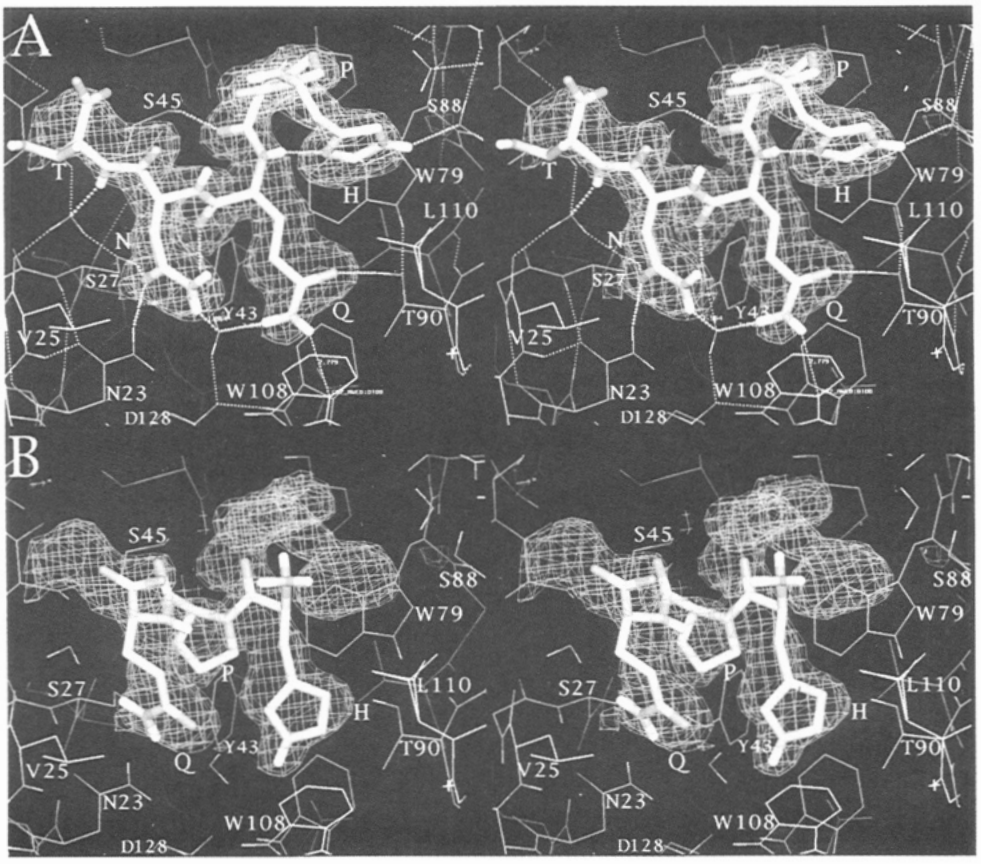

FIGURE 3: (A) ($|F_o| - |F_c|$), α_c omit map for streptavidin−FSHPQNT after refinement to convergence of the structure following removal of the ligand from site 1 superimposed on the refined structure of the complex of this investigation. The density is contoured at 2.5 and 3.0 σ. (B) Same omit map superimposed on the streptavidin−HPQ segment from the investigation of Weber et al. (refined with the data of this investigation, see Experimental Procedures). The ($|F_o| - |F_c|$) (nonomit) maps after refinement in each of the two conformations are also included in Figure 1A,B of the supporting information.

refined with the data of this investigation. The average temperature factor for the HPQ atoms of the peptide in the conformation of Weber is 40.2 (9.9) Å^2 compared with 21.2 (4.5) Å^2 for the same segment of the peptide in the conformation of this study. Finally, crystallographic refinement of several models of the streptavidin-bound peptide in the conformation of Weber et al. all led to significant distortion from planarity of the Pro−Gln peptide bond.

Apostreptavidin Crystallized from Ammonium Sulfate Contains Two Sulfate Anions in the Biotin Binding Site. The crystal structure of streptavidin−sulfate contains two sulfate anions in the biotin binding site, neither of which makes particularly good electrostatic interactions or many direct H-bonds with the protein. One sulfate is located near the position of the five-membered sulfur-containing ring of biotin in streptavidin−biotin, and the other sulfate is located near the position of the carboxyl group of biotin in streptavidin−biotin (Figure 2B).

The Flexible Flap Is in a "Closed" Conformation in Streptavidin−Biotin and Streptavidin−Sulfate, and in an "Open" Conformation in Streptavidin−Peptide Complexes. The flexible flap comprising residues 47−52 is completely ordered and well determined in streptavidin−sulfate, and is in the same position and conformation as in streptavidin−biotin (Figure 2B). The flap is in a "closed conformation" in these complexes. The ($2|F_o| - |F_c|$), α_c map superimposed on the refined structure of streptavidin−sulfate is provided in the supporting information (see paragraph at end of paper regarding available supporting information).

By contrast, this region in the crystal structure of apostreptavidin crystallized in the absence of sulfate is disordered (Weber et al., 1989). The flexible flaps in the streptavidin−peptide complexes are in a markedly different position and conformation than in streptavidin−sulfate or streptavidin−biotin. Although the density for the flap is weak in the streptavidin−peptide complexes, reflecting mobility and disorder, a crude determination of the coordinates for the backbone and some of the side chain atoms shows that the flap is in a much more open conformation (necessitated by binding of the relatively large peptide ligands).

DISCUSSION

The K_d's for streptavidin of *cyclo*-Ac-[CHPQGPPC]-NH_2 and *cyclo*-Ac-[CHPQFC]-NH_2, determined by plasmon resonance measurements using a Biacore instrument, are 670 and 270 nM, respectively (Giebel et al., 1995), whereas the K_d for FSHPQNT is 125 μM (Weber et al., 1992). Thus, the disulfide-bonded cyclic peptides bind with several hundred-fold greater affinity than their linear counterparts. This enhancement was suggested to reflect a lowered entropy of the unbound cyclic peptides compared with that of the linear ones (Giebel et al., 1995). The crystal structures are consistent with the idea that the effect of the disulfides on affinity is primarily entropic, as they show that the disulfides of the streptavidin-bound cyclic peptides are on the surface and do not interact with the protein (and thus do not directly contribute enthalpically to the binding energies).

In streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2, the Phe residue makes hydrophobic interactions not only with Leu25 and Trp108 of one subunit but also with Trp120 of a neighboring subunit. However, in streptavidin-bound *cyclo*-Ac-[CHPQGPPC]-NH_2, the Pro-Pro segment of the cyclic peptide does not make many close interactions with streptavidin. Thus, for the smaller cyclic peptide there are more favorable interactions with streptavidin per ligand residue than for the larger cyclic peptide. Although the increased number of favorable interactions per residue may explain why the smaller cyclic peptide has the greater affinity, the smaller cyclic peptide is also more conformationally restricted. Thus, both enthalpy and entropy effects may account for the difference in affinities. (Note that to estimate or rationalize changes in affinity one may also have to account for changes in solvation and desolvation of the free ligands.)

The bound linear FSHPQNT peptide is less compact than the bound cyclic peptides, and lacks the type I β-turn H-bond that the bound cyclic peptides contain. In the streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2 complex, there is an additional H-bond between the peptide main chain His carbonyl and the C-terminal main chain Cys amide NH, defining a small stretch of distorted α-helix. This bound peptide is the most compact. Because the cyclic ring size of *cyclo*-Ac-[CHPQFC]-NH_2 is relatively small, it might be anticipated that this peptide is ordered in solution. Indeed, 2D NMR solution spectra show that *cyclo*-Ac-[CHPQFC]-NH_2 has two roughly equally populated structured conformers (B. D. Sykes, private communication). The relatively ordered solution structure reflects a relatively low degree of conformational entropy in the unbound, solution state as suggested above.

In the *cyclo*-Ac-[CHPQGPPC]-NH_2 peptide, the role of the Pro-Pro stretch may be primarily to render the main chain conformation of the cyclic peptide compatible with disulfide formation. Similarly, in all HPQ peptides that bind streptavidin, the role of the Pro (in the HPQ portion) may be primarily to enable the formation of the type I β-turn. The role of the disulfides is to provide conformational constraints.

The interactions observed in the streptavidin−peptide complexes involving the HPQ portions are all very similar, involving common H-bonds and van der Waals interactions. Additional interactions involving the residues flanking the HPQ sequences, however, differ among the peptide complexes, as do the structures of these parts. In streptavidin−FSHPQNT, the Asn residue, which directly follows the HPQ sequence, makes three direct H-bonds to the protein, as well as one water-mediated H-bond. Thus, this Asn may contribute to the binding affinity of the linear peptide. Although in streptavidin−*cyclo*-Ac-[CHPQFC]-NH_2 the Phe residue also directly follows the HPQ sequence, it occupies a different region of space and makes totally different interactions (hydrophobic) than the Asn in FSHPQNT. Thus, the Phe residue contributes in a different way to the binding affinity of cyclo-Ac-[CHPQFC]-NH_2 than does the Asn in FSHPQNT.

In addition to the streptavidin binding peptides studied here, several additional HPQ-containing linear peptide sequences (HDHPQNL, SHPQGPPS) and disulfide-bonded cyclic peptides ([CHPQFSNC], [CHPQFPC], and [CHPQFNC]) were discovered in the study of Giebel et al. (1995). The crystal structures of this investigation suggest that the streptavidin binding interactions for the HPQ segment of all these sequences involve common H-bonds and van der Waals interactions described in the crystal structures of this study. The interactions made by the flanking residues are expected to differ from one complex to another.

In streptavidin−sulfate, one of the sulfates is near the position of the carboxyl group of biotin in streptavidin−biotin, and makes the same H-bond to the main chain amide of Asn49. Formation of this H-bond may help induce the flap into the same closed conformation in both streptavidin−biotin and streptavidin−sulfate. However, since these structures are in different space groups, differences in crystal packing forces may also affect the flap conformations. Indeed, there are two crystal forms of inhibitor-free HIV protease: one with an open flap and one with this flap closed. The latter closed-flap conformation is the same as observed in inhibitor complexes (Erickson, 1994). The access of new protein−ligand interactions from the flap residues when it closes the binding site of streptavidin is paralleled in HIV protease by a similar process which is one of the determinants of inhibitor affinity and activity in the latter system.

From the structure of the closed-flap conformation determined in the structure of streptavidin−biotin or streptavidin−sulfate, it might be possible to exploit additional interactions from flap residues in the design of a tight binding organic ligand different from biotin and smaller than the peptides in this study. The crystallographically determined sulfate and water structure observed in the biotin binding site of streptavidin−sulfate could be used to aid such ligand design. The streptavidin−ligand structures show that ordered water plays an integral role in the binding of peptides, biotin, and sulfate. Water is versatile in its reorganization in response to the binding of different ligands to macromolecules. The ability to predict how protein, ligand, water, and ions interact and coassemble to produce a defined bound structure of a defined ligand binding affinity remains a challenge in structure-based ligand and drug design.

Another challenge involves the structure-based design of a small organic tight binding molecule from that of a bound peptide. Peptides and proteins are the natural substrates for many pharmaceutically relevant macromolecular targets, and their bound structures are often determinable by crystallography or 2D NMR. The structures of the streptavidin−peptide and streptavidin−biotin complexes described here address the principles likely to be involved in such design efforts. Clearly there is no simple global relationship between the structure and binding interactions made by biotin and those made by any of the peptide ligands in this study. To rationally design biotin in one step based on the structure of a streptavidin−peptide structure with the present-day computational technologies would be difficult. But together, the structures of both streptavidin-bound biotin and peptides provide rationales for ligand design that combine features of both types of ligands.

The visualization of the β-turn motif involving the streptavidin binding residues of the cyclic peptides also allows the design of non-peptide β-turn mimetic streptavidin binding ligands. Since β-turns are key recognition elements in many important biological interactions (Rose et al., 1985), much effort has been focused on the design of small constrained β-turn mimetics [for reviews, see Ball and Alewood (1990) and Kahn (1993)] to provide potent and specific therapeutic agents. The design, assay, and determination of the streptavidin-bound structures of β-turn mimetics will provide valuable insight into the principles

involved in converting a peptide ligand into a small organic ligand with high affinity and specificity.

CONCLUSION

The screening of phage libraries followed by chemical synthesis of discovered sequences, assay, and structure determination of the protein-bound peptides in the streptavidin system has led to a good structural basis in which structure-based ligand design can be applied. The evaluation of designed compounds in this system should provide design experience useful in other pharmaceutically relevant systems. Thus, the streptavidin system offers an exciting opportunity to combine two powerful and potentially synergistic technologies that are maturing in drug discovery and development: combinatorial chemistry and structure-based design.

ACKNOWLEDGMENT

I thank Chuck Johnson, Nathan Collins, and Rafael Arze for peptide synthesis, Thutam Hopkins for crystallization of streptavidin–FSHPQNT, Robert Cass for BIAcore discussions, and Mike Ross, Robert Stroud, and Brian Sykes for discussion and review of the manuscript.

SUPPORTING INFORMATION AVAILABLE

($|F_o| - |F_c|$), α_c maps obtained after refinement to convergence of the streptavidin–FSHPQNT complex with the peptide in the conformation determined here (Figure 1A), and in the conformation published by Weber et al. (1992) (Figure 1B) are available. Figure 2 shows the ($2|F_o| - |F_c|$), α_c map for streptavidin–sulfate superimposed on the refined structure (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Ball, J. B., & Alewood, P. F. (1990) *J. Mol. Recognit. 3*, 55–64.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Brünger, A. T. (1992a) X-PLOR Manual, Version 3.1, Yale University, New Haven, CT.
Brünger, A. T. (1992b) *Nature 355*, 472–474.
Burley, S. K., & Petsko, G. A. (1986) *FEBS Lett. 203*, 139–143.
Chambers, J. L., & Stroud, R. M. (1979) *Acta Crystallogr. B33*, 1861.
Devlin, J. J., Panganiban, L. C., & Devlin, P. E. (1990) *Science 249*, 404.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Epp, O., Coleman, P., Fehlhammer, H., Bode, W., Schiffer, M., Huber, R., & Palm, W. (1974) *Eur. J. Biochem. 45*, 513–524.
Erickson, J. (1994) Speaker at the International Business Communications Third Annual Conference on Rational Drug Design, Dec 12–13, 1994, San Diego, CA.
Finer-Moore, J. S., Kossiakoff, A. A., Hurley, J. H., Earnest, T., & Stroud, R. M. (1992) Proteins: *Struct., Funct., Genet. 12*, 203–222.
Giebel, L. B., Cass, R., Milligan, D. L., Young, D., Arze, R., & Johnson, C. (1995) *Biochemistry* (following paper in this issue).
Jackson, S., DeGrado, W., Dwivedi, A., Parthasarathy, A., Higley, A., Krywko, J., Rockwell, A., Markwalder, J., Wells, G., Wexler, R., Mousa, S., & Harlow, R. (1994) *J. Am. Chem. Soc. 116*, 3220–3230.
Kahn, M. (Guest Ed., 1993) Tetrahedron 49, *Symp. 50*, 3433–3677.
Kamber, B., Hartmann, A., Eisler, K., Riniker, B., Rink, H., Sieber, P., & Rittel, W. (1980) *Helv. Chim. Acta 63*, 899–914.
Katz, B. A., & Kossiakoff, A. A. (1986) *J. Biol. Chem. 261*, 15480–15485.
Katz, B. A., Finer-Moore, J. S., Mortezaei, R., Rich D. H., & Stroud, R. M. (1995) *Biochemistry 34*, 8264–8280.
Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., & Knapp, R. J. (1991) *Nature 352*, 82.
McDowell, R. S., Blackburn, B. K., Gadek, T. R., McGee, L. R., Rawson, T., Reynolds, M. E., Robarge, K. D., Somers, T. C., Thorsett, E. D., Tischler, M., Webb, R. R., II, & Venuti, M. C. (1994) *J. Am. Chem. Soc. 116*, 5077–5083.
Pähler, A., Hendrickson, W. A., Kolks, M. A. G., Argaraña, C. E., & Cantor, C. R. (1987) *J. Biol. Chem. 262*, 13933–13937.
Rose, G. D., Gierasch, L. M., & Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1–109.
Savage, D., Mattson, G., Desai, S., Nielander, G., Morgensen, S., & Conklin, E. (1992) in *Avidin-Biotin Chemistry: A Handbook*, Pierce Chemical Co., Rockford, IL.
Wang, D., Bode, W., & Huber, R. (1985) *J. Mol. Biol. 185*, 595–624.
Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J., & Salemme, F. R. (1989) *Science 234*, 85–88.
Weber, P. C., Pantoliano, M. W., & Thompson, L. D. (1992) *Biochemistry 31*, 9350–9354.

BI9503561