Structure-Based Design Tools: Structural and Thermodynamic Comparison with Biotin of a Small Molecule That Binds to Streptavidin with Micromolar Affinity

Bradley A. Katz,* Beishan Liu, and Robert Cass

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

Received April 25, 1996[⊗]

Abstract: A small streptavidin-binding ligand was identified which incorporates some of the binding interactions seen in the crystal structures of streptavidin−biotin, and of streptavidin complexed with cyclo-Ac-[CHPQFC]-NH_2, a cyclic peptide ligand with a K_d of 2.3×10^{-7} M, discovered by phage display. The crystal structure of streptavidin−glycoluril is described and compared with the crystal structures of streptavidin−biotin and of streptavidin−cyclo-Ac-[CHPQFC]-NH_2. The K_d of glycoluril for streptavidin was determined by plasmon resonance measurements to be 2.5×10^{-6} M. The differences in the affinities of biotin and glycoluril for streptavidin were related to the differences in the crystal structures of the complexes and to differences in the determined solubilities of the ligands. Streptavidin-bound glycoluril has structural characteristics of, and makes interactions common to, both bound biotin and the bound cyclic peptide ligand. Binding of glycoluril and biotin is mediated by short, medium strength hydrogen bonds involving the ureido oxygen common to the two ligands. Binding of glycoluril is further mediated by a short hydrogen bond involving its unique ureido oxygen. The structural and physicochemical factors responsible for the weaker binding of glycoluril compared with biotin are discussed.

Introduction

A common challenge faced in drug design is the conversion of peptide ligands into small organic ligands of high affinity and selectivity. Drug development often relies on the structures or sequences of peptide ligands as initial leads in the subsequent design and evaluation of peptidomimetic compounds, since peptides and proteins are the normal substrates or ligands for macromolecules involved in ubiquitous biological processes. Peptides may also be lead compounds discovered by phage display, a technology which has quickly become widespread as an effective method for discovery of peptide ligands for diverse macromolecular targets (reviewed in ref 1−10). Phage display offers the opportunity to produce cyclic peptide libraries the conformational constraints of which imposed by disulfide bonds often contain molecules of significantly increased ligand affinity compared with their linear peptide counterparts.[11−19]

Using streptavidin as a model system, we recently probed the structural relationship between the natural, organic ligand biotin and linear and cyclic peptide ligands containing the HPQ sequence[20] discovered by phage display.[11] The remarkable affinity of streptavidin, a tetrameric protein secreted from *Streptomyces avidinii*, for the vitamin biotin, $K_d \sim 10^{-14}$ M,[21] forms the basis for its prevalent use in biotechnological, bioanalytical, diagnostic, and therapeutic applications.[21−25] Streptavidin has also been used as a model system for discovery of peptide ligands by phage display,[11,26−29] as a paradigm for probing the structural basis of high-affinity protein−ligand interactions,[20,30−33] and for developing and applying structure-based ligand design strategies.[34−37]

* Phone: (415) 829-1010. Fax: (415) 829-1001. E-mail: bak@arris.com.
⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1996.

(1) O'Neil, K.; Hoess, R. H. *Curr. Opin. Struct. Biol.* **1995**, *5*, 443−449.
(2) Daniels, D. A.; Dion, A.; Lane, D. P. *Expert Opin. Ther. Pat.* **1995**, *5*, 901−912.
(3) Burton, D. R. *Immunotechnology* **1995**, *1*, 87−94.
(4) Cortese, R.; Monaci, P.; Nicosia, A.; Luzzago, A.; Felici, F.; Galfré, G.; Pessi, A.; Tramontano, A.; Sollazzo, M. *Curr. Opin. Biotechnol.* **1995**, *6*, 73−80.
(5) Cortese, R.; Felici, F.; Galfré, G.; Luzzago, A.; Monaci, P.; Nicosia, A. *TIBTECH* **1994**, *12*, 262−267.
(6) Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R. *Annu. Rev. Immunol.* **1994**, *12*, 433−455.
(7) Winter, J. *Drug Dev. Res.* **1994**, *33*, 71−89.
(8) Clackson, T.; Wells, J. A. *Trends Biotechnol.* **1994**, *12*, 173−184.
(9) Barbas III, C. F. *Curr. Opin. Biotechnol.* **1993**, *4*, 526−530.
(10) Wells, J. A.; Lowman, H. B. *Curr. Opin. Struct. Biol.* **1992**, *2*, 597−604.
(11) Giebel, L. B.; Cass, R.; Milligan, D. L.; Young, D.; Arze, R.; Johnson, C. *Biochemistry* **1995**, *34*, 15430−15435.
(12) Wright, R. M.; Gram, H.; Vattay, A.; Byrne, S.; Lake, P.; Dottavio, D. *Biotechnology* **1995**, *13*, 165−169.
(13) Koivunen, E.; Wang, B.; Ruoslahti, E. *Biotechnology* **1995**, *13*, 265−270.
(14) Pasqualini, R.; Koivunen, E.; Ruoslahti, E. *J. Cell. Biol.* **1995**, *130*, 1189−1196.
(15) Hoess, R. H.; Mack, A. J.; Walton, H.; Reilly, T. M. *J. Immunol.* **1994**, *153*, 724−729.
(16) Koivunen, E.; Gay, D. A.; Ruoslahti, E. *J. Biol. Chem.* **1993**, *268*, 20205−20210.
(17) Luzzago, A.; Felici, F.; Tramontono, A.; Pessi, A.; Cortese, R. *Gene* **1993**, *128*, 51−57.
(18) O'Neil, K. T.; Hoess, R. H.; Jackson, S. A.; Ramachandran, N. S.; Mousa, S. A.; DeGrado, W. F. *Proteins* **1992**, *14*, 509−515.
(19) McLafferty, M. A.; Kent, R. B.; Ladner, R. C.; Markland, W. *Gene* **1992**, *128*, 29−36.
(20) Katz, B. A. *Biochemistry* **1995**, *34*, 15421−15429.
(21) Green, N. M. *Methods Enzymol.* **1990**, *184*, 51−67.
(22) Chilkoti, A.; Schwartz, B. L.; Smith, R. D.; Long, C. J.; Stayton, P. S. *Biotechnology* **1995**, *13*, 1198−1204.
(23) Wilchek, M.; Bayer, E. A. *Methods Enzymol.* **1990**, *184*, 5−13.
(24) Wilchek, M.; Bayer, E. A. *Methods Enzymol.* **1990**, *184*, 14−45.
(25) Sano, T.; Pandori, M. W.; Chen, X.; Smith, C. L.; Cantor, C. R. *J. Biol. Chem.* **1995**, *270*, 28204−28209.
(26) Saggio, I.; Laufer, R. *Biochem. J.* **1993**, *293*, 613−616.
(27) Kay, B. K.; Adey, N. B.; He, Y.-S.; Manfredi, J. P.; Mataragnon, A. H.; Fowlkes, D. M. *Gene* **1993**, *128*, 59−65.
(28) Roberts, D.; Guegler, K.; Winter, J. *Gene* **1993**, *128*, 67−69.
(29) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. *Science* **1990**, *249*, 404−406.
(30) Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R. *J. Am. Chem. Soc.* **1992**, *114*, 3197−3200.
(31) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. *Science* **1989**, *243*, 85−88.

S0002-7863(96)01357-1 CCC: $12.00

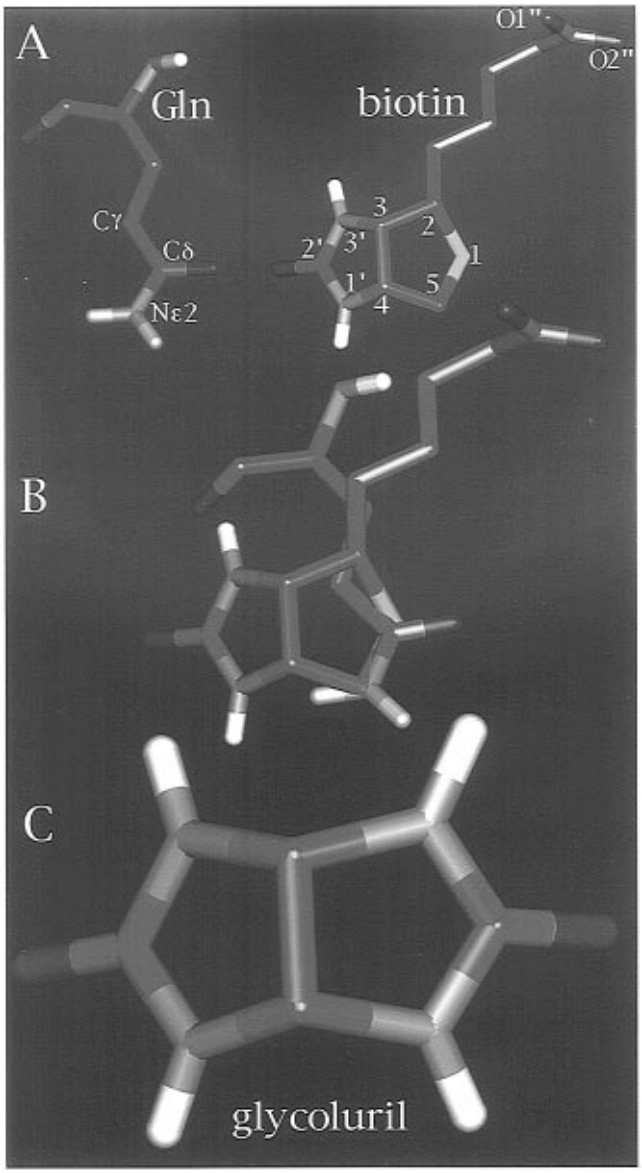

Figure 1. (a) Structures of Gln of streptavidin-bound cyclo-Ac-[CHPQFC]-NH_2 and of streptavidin-bound biotin with ligand labeling scheme. (b) Relationship of Gln and biotin after superposition of streptavidin−cyclo-Ac-[CHPQFC]-NH_2 onto streptavidin−biotin.

From the crystal structure of streptavidin−biotin, determined here at pH 7.3 and 1.55 Å resolution, and from the crystal structure at pH 7.5 and 2.00 Å resolution of streptavidin complexed with a cyclic peptide, cyclo-Ac-[CHPQFC]-NH_2,[20] discovered by phage display,[11] we identified a small organic ligand having structural features common to both biotin and the cyclic peptide ligand. In streptavidin−cyclo-Ac-[CHPQFC]-NH_2, three contiguous side chain atoms of the Gln residue, Cγ, Cδ, and Nϵ2, superimpose well on three contiguous atoms in the bicyclic ring of streptavidin-bound biotin, the C2, S1, and C5 atoms (Figure 1).

A search of the Fine Chemical Database identified glycoluril as a molecule which incorporates these aspects of the design. In glycoluril the group corresponding to the C2 carbon atom of biotin is also an NH group. Thus glycoluril, a C2-symmetric molecule with the same basic bicyclic framework as biotin but with different groups at the 1, 2, and 5 ring positions, was identified as a commercially available compound that was a potential streptavidin-binding ligand incorporating structural features of both bound biotin and bound cyclo-Ac-[CHPQFC]-NH_2. Glycoluril was shown by plasmon resonance measurements to bind to streptavidin with a K_d of 2.5×10^{-6} M. The structure of streptavidin−glycoluril was determined at pH 7.3 and 1.85 Å resolution and compared with the structures of streptavidin−biotin and streptavidin−cyclo-Ac-[CHPQFC]-NH_2, from which it was designed.

Dissection of the structural and physicochemical factors contributing to the unusually high affinity of streptavidin for biotin has been the focus of many studies. Hydrophobic and van der Waals contributions mediated primarily by tryptophan residues have been predicted[38] and demonstrated by site-directed mutagenesis studies[39,40] to be dominant determinants of binding. From free energy perturbation experiments, Miyamoto and Kollman predicted that van der Waals interactions contribute 18.2 kcal/mol of the binding energy compared with 3.0 kcal/mol for hydrogen-bonding interactions.[38] From microcalorimetry, Weber et al. determined that binding of biotin to streptavidin is driven predominantly by enthalpy changes.[30] Glycoluril can be considered as an analog of biotin in which the valerate substituent is removed. Thus the contribution of the valerate group of biotin to its binding was probed by comparing the affinities, solubilities, and streptavidin-bound structures of biotin and glycoluril.

Experimental Section

Design of Glycoluril-like Ligand and Identification of Glycoluril in the Fine Chemical Database. The structure of streptavidin−cyclo-Ac-[CHPQFC]-NH_2 was superimposed on the structure of streptavidin−biotin with a matrix determined from corresponding pairs of non-hydrogen atoms involving 56 well ordered ($B < 20$ Å^2) residues (29−32, 38−47, 54−61, 71−79, 88−96, 104−112, and 124−130) spread throughout the molecule. The superposition revealed that in streptavidin−cyclo-Ac-[CHPQFC]-NH_2 three contiguous side chain atoms of the Gln residue, Cγ, Cδ, and Nϵ2, superimpose well on three contiguous atoms of streptavidin-bound biotin, the C2, S, and C5 carbon atoms (Figure 1). With the program Insight (licensed from Molecular Simulations, Burlington, MA 01803), we built a model of a biotin-like ligand bound to streptavidin that incorporates substituents from both biotin as well as from the Gln side chain of cyclo-Ac-[CHPQFC]-NH_2. Using the crystal structure of streptavidin−biotin, we replaced the sulfur of biotin with a carbonyl group and the C5 carbon atom of biotin with an NH group. A search of the Fine Chemical Database for such a molecule as well as chemically similar molecules, using the program Navigator,[41] identified glycoluril as the compound the chemical structure of which is the most similar in this database with respect to the bicyclic ring of this biotin-like ligand. In glycoluril the S, C1, and C5 atoms of biotin are converted to C═O, NH, and NH groups, respectively. Visualization of the model of streptavidin−glycoluril with Insight following energy minimization to convergence with Discover (licensed from Molecular Simulations, Burlington, MA 01803) suggested that glycoluril would fit into the biotin binding site and make many of the interactions that biotin makes.

Determination of Ligand Solubilities. Biotin (hexahydro-2-oxo-1*H*-thieno[3,4-*d*]imidazole-4-pentanoic acid) was purchased from Sigma

(32) Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, Y. S.; Phizackerley, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2190−2194.
(33) Katz, B. A. *J. Am. Chem. Soc.* **1996**, *118*, 2535−2536.
(34) Katz, B. A.; Johnson, C.; Cass, R. T. *J. Am. Chem. Soc.* **1995**, *117*, 8541−8547.
(35) Katz, B. A.; Stroud, R. M.; Collins, N.; Liu, B.; Arze, R. *Chem. Biol.* **1995**, *2*, 591−600.
(36) Katz, B. A.; Cass, R. T.; Liu, B.; Arze, R.; Collins, N. *J. Biol. Chem.* **1995**, *270*, 31210−31218.
(37) Weber, P. C.; Pantoliano, M. W.; Simons, D. M.; Salemme, F. R. *J. Am. Chem. Soc.* **1994**, *116*, 2717−2724.
(38) Miyamoto, S.; Kollman, P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402−8406.
(39) Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1754−1758.
(40) Sano, T.; Cantor, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3180−3184.
(41) Chapman, D.; Harris, N.; Park, J.; Critchlow, R. E., Jr. *J. Mol. Graphics* **1995**, *13*, 242−249.

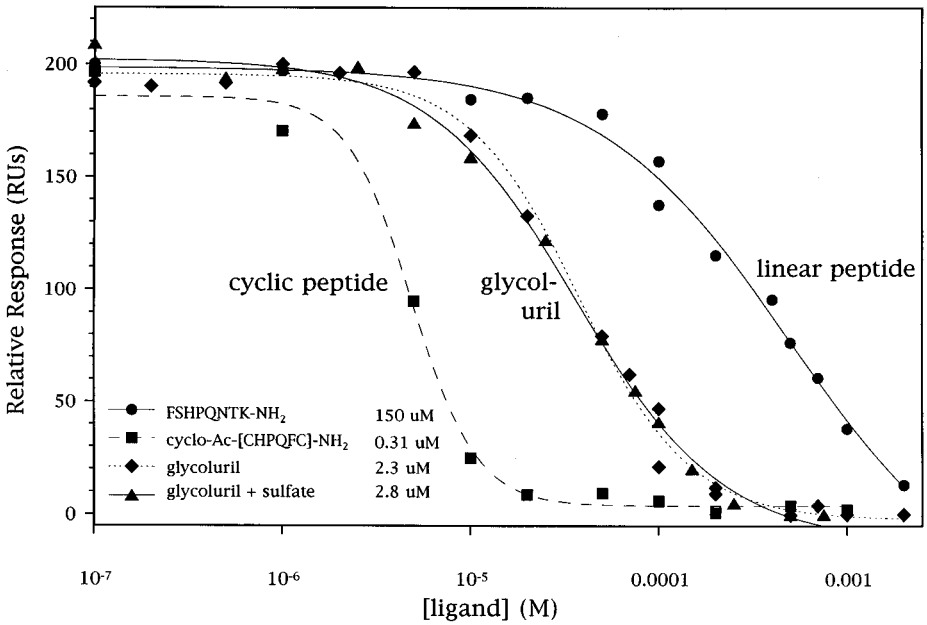

Figure 2. Inhibition curves for the competition of streptavidin between cyclo-[5-*S*-valeramide-HPQGPPC]K-NH_2 immobilized on the sensor chip and cyclo-[5-*S*-valeramide-HPQGPPC]K-NH_2, linear FSHPQNTK-NH_2, or glycoluril (in the absence or presence of 500 mM ammonium sulfate) in the solution flowing over the sensor chip. The IC_{50}s for the ligands were converted to K_d's by applying the factor relating the IC_{50} determined for cyclo-[5-*S*-valeramide-HPQGPPC]K-NH_2 in solution to the actual K_d determined directly in a separate BIAcore experiment by described methods.[11] The IC_{50}'s were determined by curve fitting the data to the function: $RU = (A - D)/(1 + (IC_{50}/conc)B) + D$, where RU is the BIAcore response, conc is the concentration of the free peptide ligand, and *A*, *B*, IC_{50}, and *D* are the parameters derived from the curve fit.

Table 1. Thermodynamic Properties of Glycoluril and Biotin

	glycoluril	biotin
solubility (pH 7.3) (mM)	6.37 (0.18)	150 (5)
$\Delta\Delta G_{solvation}$ (kcal/mol)[a]	−1.83 (0.05)	0
K_d (M)	$2.5\ (0.3) \times 10^{-6}$	4.0×10^{-14} [b]
$\Delta\Delta G_{binding}$(kcal/mol)[c]	10.4 (0.1)	0

[a] $\Delta G_{solvation}$(glycoluril) − $\Delta G_{solvation}$(biotin). [b] pH 6.9, 100 mM KCl.[30] [c] $\Delta G_{binding}$(glycoluril) − $\Delta G_{binding}$(biotin). Only the uncertainity in the K_d for glycoluril, based on a total of four independent determinations, was used to calculate the uncertainty in $\Delta\Delta G_{binding}$ (the uncertainty of which is thus underestimated because the uncertainty in the K_d for biotin is not known).

and glycoluril from Aldrich. Weighed quantities of biotin and glycoluril were each added separately to measured volumes of buffer followed by constant agitation for 6−18 h on a rotor. On the basis of whether or not total dissolution was attained, the quantities of ligands per volume of buffer were accordingly modified in subsequent iterations. Saturation was signaled by visualization of small amounts of undissolved material with a microscope after the agitation process. Solubility determinations were repeated twice to obtain averages and standard deviations given in Table 1. The buffer was 150 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.3. For biotin the final pH was adjusted with NaOH (150 mM) to attain the target pH. To attain a similar ionic strength for the glycoluril solubility determination, 150 mM NaCl was added. The added NaCl had an insignificant effect (<1.4%) on the solubility of glycoluril.

Determination of the Affinity of Glycoluril for Streptavidin. The affinity for glycoluril in the presence and absence of ammonium sulfate was determined by surface plasmon resonance measurements by methods previously described.[11,34,36] The BIAcore 2000 system, sensor chip, and coupling reagents *N*-hydroxysuccinimide, *N*-ethyl-*N*′-(3-diethylaminopropyl)carbodiimide, and ethanolamine hydrochloride were from Pharmacia Biotech, Inc, Uppsala, Sweden. In the first two K_d determination, done in the absence of sulfate, the running buffer used was 10 mM HEPES, 150 mM NaCl, 3.4 mM ethylenediaminetetraacetic acid, 0.005% P20 surfactant, pH 7.3. The second two determinations were done with the same running buffer, except containing 500 mM ammonium sulfate. The BIAcore 2000 system, sensor chip, and coupling reagents *N*-hydroxysuccinimide, *N*-ethyl-*N*′-(3-diethylaminopropyl)carbodiimide, and ethanolamine hydrochloride were from Pharmacia Biotech, Inc, Uppsala, Sweden. The affinity was determined in a competition assay involving a cyclic streptavidin-binding peptide, cyclo-[5-*S*-valeramide-HPQGPPC]K-NH_2,[34] immobilized on the sensor chip surface via the ϵ-amino group of the C-terminal lysine. The competition curves are shown in Figure 2.

Crystallography of Complexes of Streptavidin with Glycoluril and Biotin. Apostreptavidin was crystallized from ammonium sulfate solution, pH 4.0, in the diamond plate form, space group, *I*222, as described.[20,42] Streptavidin−biotin, pH 7.3, and streptavidin−glycoluril, pH 7.3, were prepared by soaking large diamond plate apostreptavidin crystals for several days in synthetic mother liquor solution (50% saturated ammonium sulfate solution, 50% 0.50 M Tris, pH 7.3) saturated with the ligands.

X-ray diffraction datasets were collected from single crystals on a Siemens IPC X-1000 multiwire area detector coupled to a Siemens three-circle goniometer (Table 2). X-rays were produced with Rigaku RU200 generator using a rotating copper 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 (Redmond, WA 98052). The following settings were used: crystal to detector distance = 8.5 cm, $2\theta = -27.5$ or $-32.0°$, oscillation widths in $\omega = -0.085°$, exposure time per frame = 300 s. Three sweeps of 90° in ω at φ values of 0.0, 90.0, 45.0° were collected.

Models of streptavidin with bound ligands were manually built and visualized by computer graphics using Quanta (licensed from Molecular Simulations, Burlington, MA 01803) from the structure of apostreptavidin[20] from which the sulfates and waters were removed from the biotin binding site. Models were based on initial electron density maps calculated with coefficients ($|F_o| - |F_c|$) or ($2|F_o| - |F_c|$) and phases (α_c) after refinement to convergence of structures where the ligand atoms were omitted from the refinement and map calculations.

The structures of streptavidin−ligand complexes were refined with X-PLOR[43] and with difference Fourier methods.[44] Bond lengths from the small molecule crystal structure of glycoluril[45] and biotin[46] were used in the refinements; other bond lengths and all force constants were

(42) Pähler, A.; Hendrickson, W. A.; Kolks, M. A. G.; Argaraña, C. E.; Cantor, C. R. *J. Biol. Chem.* **1987**, *262*, 13933−13937.

(43) Brünger, A. T. In *X-PLOR version 3.1: A system for X-ray crystallography and NMR*; Yale University Press: New Haven, CT 06511, 1992; pp 187−206.

(44) Chambers, J. L.; Stroud, R. M. *Acta Crystallogr.* **1979**, *B35*, 1861−1874.

(45) Li, N.; Maluendes, S.; Blessing, R. H.; Dupuis, M.; Moss, G. R.; DeTitta, G. T. *J. Am. Chem. Soc.* **1994**, *116*, 6494−6507.

Table 2. Crystallography of Streptavidin Complexes

	biotin	glycoluril
Parameters[a]		
no. atoms (including disorder)	2117	2070
no. waters (including sulfate(s))	227	193
no. discretely disordered groups[b]	7	7
no. side chains with refined occs[c]	22	21
Diffraction Statistics		
resolution (Å)	21−1.52	22−1.81
no. observations[d]	71142	56712
no. merged reflections	27764	17677
average redundancy	2.6	3.2
R_{merge} (%)[e]	5.0	7.4
Refinement Statistics		
refinement resolution	7.5−1.55	7.5−1.85
no. merged reflections	25217	15772
$\|F_o\|/\sigma$ cutoff	1.8	1.5
R_{cryst} (%) [f]	16.8	17.2
free R_{cryst} (%) [g]	19.9	21.2
overall completeness (%)	70.9	75.7
and at highest resolution (%)	27.6	34.3
highest resolution shell	1.62−1.55	1.93−1.85
RMS Deviations[h]		
bond lengths (Å)	0.018	0.018
bond angles (deg)	2.8	3.1
torsion angles (deg)	27.1	27.8

[a] Restrained, isotropic temperature factors were refined and bulk solvent contributions included for all structures. [b] Not including waters. [c] 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. [d] Data with $R_{sym} > 50\%$ were rejected along with data with values $>3.5\sigma$ from the mean for each bunch of symmetry equivalents. [e] $R_{merge} = \sum_h\sum_i|I(h)_i - \langle I(h)\rangle|/\sum_h\sum_i I(h)_i$, where $I(h)_i$ is the ith observation of the intensity of reflection h. [f] $R_{cryst} = \sum(||F_o| - |F_c||)/\sum|F_o|$ (for reflections from 7.5 Å to the highest resolution). [g] Cross validation R-factor using 10% of the data withheld from the refinement.[66] [h] Root mean square deviations from ideal bond lengths and bond angles.

from the X-PLOR force field.[43] In ($|F_o| - |F_c|$), α_c maps, positive and negative peaks the magnitudes of which were greater than 2.8σ were systematically identified and analyzed using a proprietary peak picking program. Water structure was determined and refined. New waters were identified on the basis of density peaks greater than 2.8σ in ($|F_o| - |F_c|$), α_c maps that coincided with density in ($2|F_o| - |F_c|$), α_c maps at stereochemically reasonable locations.[47] Waters were generally rejected if their temperature factors exceeded 60 Å^2 after refinement. Data collection and refinement statistics are given in Table 2. For comparison of structures, the streptavidin−ligand complexes were superimposed on one another with matrices determined from corresponding pairs of non-hydrogen main chain atoms involving 56 well-ordered ($B < 20$ Å^2) residues (29−32, 38−47, 54−61, 71−79, 88−96, 104−112, and 124−130) spread throughout the molecule. The structures will be deposited into the Brookhaven Data Bank.

Results

Affinity of Glycoluril Is Similar to Affinities of Other Designed Streptavidin-Binding Ligands. The K_d of glycoluril for streptavidin is 2.5×10^{-6} M at pH 7.3 (Figure 2). This affinity is within the range of other small molecule azobenzene streptavidin-binding molecules developed by structure-based design for which K_d's, determined by microcalorimetry, range from 8.3×10^{-7} M to 1.4×10^{-4} M.[37] The affinity for glycoluril at pH 7.3 is intermediate between the values obtained at pH 7.3 for linear ($K_d \sim 1.5 \times 10^{-4}$ M) and cyclic ($K_d \sim 3.1 \times 10^{-7}$ M) peptides discovered by phage display[11] (Figure 2). The K_d's of glycoluril, azobenzene ligands, and linear and cyclic peptide ligands at pH 7.3 are all significantly greater than that of biotin ($K_d = 4.0 \times 10^{-14}$ M, pH 6.9).[30]

Glycoluril Makes Interactions Common to Both Streptavidin−Biotin and Streptavidin−HPQ-Containing Peptides from Which It Was Designed. Superpositions of the ($2|F_o| - |F_c|$), α_c maps for streptavidin−glycoluril, pH 7.3, and streptavidin−biotin, pH 7.3, onto the respective refined structures are shown in Figure 3, parts a and b. Both ligands are defined by strong density; the average temperature factor for the glycoluril is 20 ± 3 Å^2 compared with 12 ± 3 Å^2 for biotin. The dihedral angles for the valerate group of bound biotin, defined by atoms starting from the bridgehead carbon and ending at the carboxylate oxygen, are $\chi1 = 159(3)°$, $\chi2 = 177(4)°$, $\chi3 = -173(4)°$, $\chi4 = -65(5)°$, $\chi5 = -38(4)°$. These dihedrals represent an optimal, low-energy conformation. The structure of streptavidin−biotin is the same as another high (1.55 Å) resolution structure determined in the same space group.[30]

The superposition of streptavidin−glycoluril onto streptavidin−biotin is shown in Figure 4a. The chemical structure of the left ureido ring of glycoluril is the same as that of biotin, and the streptavidin-bound structures of the left ureido rings of the two ligands are also the same. The C═O and NH groups of the left ureido ring of glycoluril make the same hydrogen bonds with streptavidin (O2′−Oη_{Tyr43}, O2′−Oγ_{Ser27}, O2′−N$\delta2_{Asn23}$, N1′−O$\delta2_{Asp128}$, and N3′−Oγ_{Ser45}) as the corresponding groups in biotin. Hydrogen-bond lengths are given in Table 3.

In streptavidin−glycoluril the hydrogen bond between the left ureido oxygen (O1) and O$\gamma1_{Thr90}$ replaces the corresponding hydrogen bond in streptavidin−biotin between the sulfur atom of biotin and O$\gamma1_{Thr90}$. In streptavidin−glycoluril the O1−O$\gamma1_{Thr90}$ and N5H−π_{Trp108} hydrogen bonds correspond to the O$\epsilon1_{Gln}$−O$\gamma1_{Thr90}$ and N$\epsilon2H_{Gln}$−π_{Trp108} hydrogen bonds, respectively, in streptavidin−cyclo-Ac-[CHPQFC]-NH_2. (Similar NH $\rightarrow \pi$ H-bonds have been described.[48] The structure of streptavidin−glycoluril is superimposed on the structure of streptavidin−cyclo-Ac-[CHPQFC]-NH_2,[20] in Figure 4b. Thus, as intended, streptavidin−glycoluril makes interactions common to both streptavidin−biotin as well as streptavidin−HPQ-containing peptide ligands.

Although Glycoluril Is Longer Than Biotin in One Direction, Streptavidin Binds Both Ligands Similarly. Replacement of the sulfur atom of biotin with a carbonyl group in glycoluril lengthens glycoluril in one direction; the O1−O2′ distance in glycoluril is 6.00 Å compared with the S1−O2′ distance of 5.01 Å in biotin. The additional length of the bicyclic ring of glycoluril in this direction due to the addition of the ═O substituent at position 1 is partially offset by the somewhat smaller ring dimensions in glycoluril due to the shorter N2−C1 and N5−C1 bond lengths (1.33 Å) compared with the corresponding C2−S1 and C5−S1 bond lengths (1.80 Å) in biotin. Also the van der Waals radius of the oxygen substituent at C1 in glycoluril is significantly smaller that of the sulfur in biotin. Finally, in the sulfur-containing ring of biotin all atoms are sp^3 hybridized and thus this ring is puckered, whereas the corresponding ring in glycoluril is planar. Thus streptavidin incorporates glycoluril and biotin in a similar binding mode and makes all the same or analogous hydrogen bonds to the bicyclic ring of glycoluril as to that of biotin (Figure 4a).

Binding of Glycoluril Is Mediated by a Sulfate Ion. An unexpected component of the streptavidin−glycoluril structure is a sulfate anion mediating the binding of glycoluril to streptavidin. The sulfate in streptavidin−glycoluril makes a hydrogen bond with the N1 group of glycoluril. The sulfate

(46) DeTitta, G. T.; Edmonds, J. W.; Stallings, W.; Donohue, J. *J. Am. Chem. Soc.* **1976**, *98*, 1920−1926.
(47) Finer-Moore, J. S.; Kossiakoff, A. A.; Hurley, J. H.; Earnest, T.; Stroud, R. M. *Proteins* **1992**, *12*, 203−222.
(48) Burley, S. K.; Petsko, G. A. *FEBS Lett.* **1986**, *203*, 139−143.

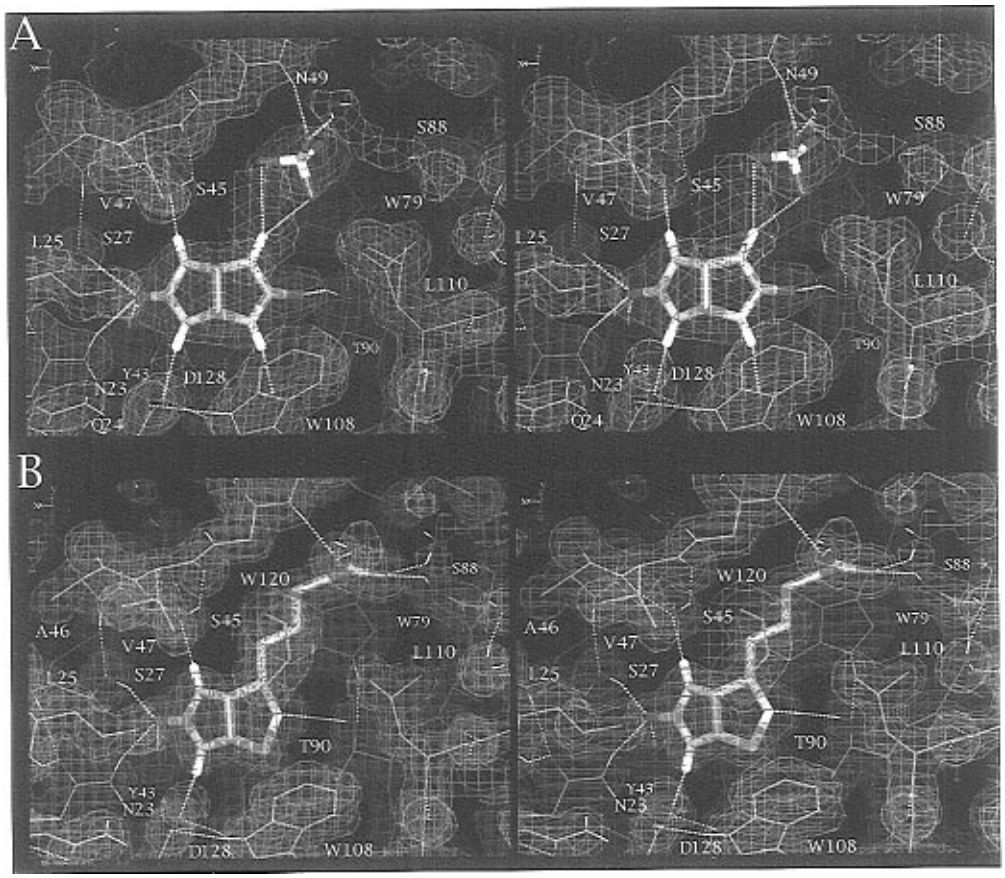

Figure 3. (a) $2(|F_o| - |F_c|)$ map for streptavidin−glycoluril, pH 7.3, superimposed on the refined structure with hydrogen bonds mediating ligand binding shown in yellow. (b) $2(|F_o| - |F_c|)$ map for streptavidin−biotin, pH 7.3, superimposed on the refined structure.

anion in streptavidin−glycoluril occupies a position similar to one of the two sulfates in the biotin binding site of streptavidin−sulfate, a form of apostreptavidin.[20] Both the sulfate in streptavidin−glycoluril and the corresponding sulfate in apostreptavidin make a hydrogen bond with NH_{Asn49}, replacing a corresponding hydrogen bond between NH_{Asn49} and the carboxyl group of biotin in streptavidin−biotin. Following crystallographic observation of the sulfate ion, the affinity was redetermined in 500 mM ammonium sulfate, but no significant difference in K_d was observed. Even though the concentration of the ammonium sulfate in the crystal soaking solution was high, 50% saturated, the temperature factors of the sulfate at the crystallographically independent binding sites 1 (60 ± 1 Å^2) and 2 (71 ± 1 Å^2) indicate much weaker, looser binding than for glycoluril, for which the temperature factors $= 20 \pm 1$ Å^2 at site 1 and 15 ± 2 Å^2 at site 2.

Discussion

Binding of Biotin and Glycoluril Involves Stabilization of a Tautomeric Form of the Ligands by Residues Surrounding an Oxyanion Hole. One determinant of the high affinity of biotin is the stabilization by streptavidin of a tautomeric form of biotin which makes better and more hydrogen bonds with the protein than are possible in water.[30] In the bound tautomer a negative charge is localized on the resulting sp^3 ureido oxygen occupying an oxyanion hole surrounded by the Tyr43, Ser27, and Asn23 side chains.[30] The resulting positive charge or partial positive charge on N1′ is stabilized by a hydrogen-bonded salt bridge with Asp128.[30] In glycoluril a similar tautomeric form is stabilized by the oxyanion residues, Tyr43, Asn23, and Ser27, and by Asp128.

Since both rings in the C2-symmetric glycoluril molecule are equivalent, the left ring is capable of a tautomeric form similar to that of the right ring described above. A partial positive charge on N2 can be stabilized by a hydrogen-bonded salt bridge with the sulfate ion. A partial positive charge on N5 can be stabilized by the N5H $\rightarrow \pi_{Trp108}$ interaction. A negative charge on O1 can be stabilized by the hydrogen bond between O1 and $O\gamma 1_{Thr90}$.

Binding of Biotin and Glycoluril Is Mediated by Medium Strength Hydrogen Bonds Involving the Ureido Oxygen(s). The majority of hydrogen bonds in proteins and protein−ligand complexes are classified as weak or conventional. For weak OH- - -O hydrogen bonds, d(O−O) > 2.80 Å, the hydrogen is attached to one of the oxygens by a covalent bond, and the interaction with the other oxygen is largely electrostatic.[49,50] When the pK_a's of the two oxygens are similar, OH- - -O hydrogen bonds are shorter than 2.6 Å and have covalent character because there is a low barrier to transferring the proton between the two oxygens.[49,50] Although low-barrier hydrogen bonds in transition states or intermediates of enzymes have been invoked as suppliers of stabilization energies of 10−20 kcal/mol,[50−52] critical for driving many enzyme reactions, there is controversy about whether or not such unusually strong hydrogen exist in solution or in enzyme active sites.[53] The O2′−$O\eta_{Tyr43}$ and O2′−$O\gamma_{Ser27}$ hydrogen-bond lengths in streptavidin−biotin and in streptavidin−glycoluril and the O1−$O\gamma 1_{Thr90}$ bond length in streptavidin−glycoluril (Table 2) are classified as medium-strength hydrogen bonds (2.65 Å < d(O−O) < 2.80 Å).[49] The observation of these short hydrogen bonds mediating

(49) Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. *J. Am. Chem. Soc.* **1994**, *116*, 909−915.
(50) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.
(51) Shan, S.; Loh, S.; Herschlag, D. *Science* **1996**, *272*, 97−101.
(52) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.
(53) Guthrie, J. P. *Chem. Biol.* **1996**, *3*, 163−170.

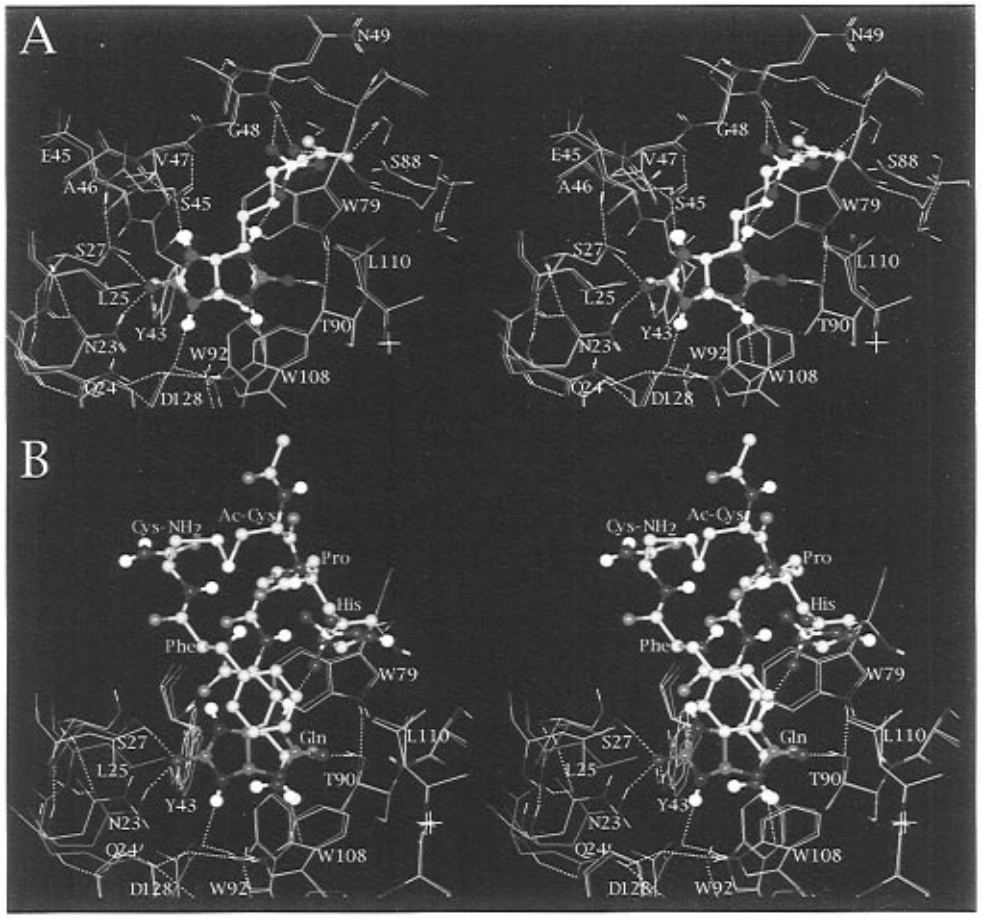

Figure 4. (a) Superposition of the structure of streptavidin−biotin onto that of streptavidin−glycoluril. In streptavidin−biotin carbon atoms are light blue, oxygens orange, nitrogens purple, and sulfurs yellow. In streptavidin−glycoluril carbons are green, oxygens red, nitrogens blue, and sulfurs yellow. Hydrogen bonds mediating binding of glycoluril are shown in yellow and the NH → π_{Trp108} H-bond in light cyan; other H-bonds are shown in white. In both structures Leu110 is disordered between two discrete conformations. (b) Superposition of the structure of streptavidin−glycoluril onto that of cyclo-Ac-[CHPQFC]-NH_2. For clarity, residues 46−52 of the loop that is in an ordered, closed conformation in streptavidin−glycoluril and that is in an open, mobile/disordered conformation in streptavidin−peptide complexes,[20] are not shown. In the cyclic peptide complex carbon atoms are light blue, oxygens orange, nitrogens purple, sulfurs yellow, and peptide ligand residues labeled in light blue font. Tyr43 is disordered between two conformations involving a rotation about the Cα−Cβ bond. In streptavidin−glycoluril carbons are green, oxygens red, nitrogens blue, and sulfurs yellow. Streptavidin residues are labeled in white font. Hydrogen bonds involving glycoluril are shown in yellow and the NH → π_{Trp108} H-bond in light cyan; other H-bonds are shown in white.

Table 3. Selected Hydrogen Bond Lengths in Complexes of Streptavidin with Glycoluril, Biotin, and cyclo-Ac-[CHPQFC]-NH_2

protein atoms	ligand atoms,[a] distances (Å)[b]		
	glycoluril	biotin	cyclo-Ac-[CHPQFC]-NH_2
Oη Tyr43	O2′, 2.69(08)	O2′, 2.72(06)	O Gln, 3.07(17)
Oγ Ser27	O2′, 2.77(15)	O2′, 2.69(07)	
Nδ2 Asn23	O2′, 2.92(11)	O2′, 3.02(05)	
Oδ2 Asp128	N1′, 2.98(05)	N1′, 2.94(04)	
Oγ Ser45	N3′, 2.93(08)	N3′, 3.05(01)	O Pro, 2.73(11)
Oδ1 Thr90	O1, 2.68(07)	S1, 3.41(02)	Oϵ1 Gln, 2.80(05)
π^c Trp108	N3, 3.48(13)		Nϵ2 Gln, 3.51(15)
SO_4^{2-}	N2, 3.10(07)		

[a] The bicyclic labeling scheme is shown in Figure 1. [b] For glycoluril, values and standard deviations are derived from the two crystallographically independent sites. For biotin they are derived from six distances involving both sites from three independently determined structures. For the cyclic peptide they are derived from eight distances involving both sites from four independent determinations at different pH's. [c] Closest distance to the aromatic ring system.

the binding of glycoluril and biotin suggests that incorporating into ligands the potential to form short, strong hydrogen bonds with protein targets may be an important principle in the engineering of specific, high-affinity ligands in drug design, to the extent that short "strong" hydrogen bonds can really provide significantly higher binding energies than "normal" hydrogen bonds.

Table 4. Comparison of Streptavidin−Glycoluril with Streptavidin−Biotin

	glycoluril: site 2	biotin	
		site 1	site 2
glycoluril site 1	0.140	0.130	0.157
glycoluril site 2		0.150	0.127
biotin site 1			0.107

[a] RMS deviations (in Å) are for well-determined residues (29−32, 38−47, 54−61, 71−79, 88−96, 104−112, and 124−130) after superposition. Site 1 is close to a crystallographically 2-fold related site.

Weaker Binding of Glycoluril Compared with Biotin Reflects Missing Valerate Group. The structures and solubilities of glycoluril and biotin and affinity for glycoluril were determined at pH 7.3, and the affinity for biotin at pH 6.9[30] (Table 1). There are no significant conformational differences between streptavidin−glycoluril and streptavidin−biotin (Figure 3a and Table 4). The ratio of the K_d for glycoluril, 2.5×10^{-6} M, to that of biotin, 4.0×10^{-14} M (Table 1), is 6.3×10^7, which corresponds to a change in the free energy of binding of glycoluril compared with that of biotin, $\Delta\Delta G_{binding}$, of 10.4 kcal/mol. Glycoluril is less soluble than biotin at pH 7.3 by a factor of 23.5, corresponding to a change in free energy of solvation, $\Delta\Delta G_{solvation}$, of −1.83 kcal/mol (Table 1). Since glycoluril is C2-symmetric it can bind to streptavidin in two indistinguishable

orientations, and thus has a favorable entropic component to its free energy of binding, $\Delta\Delta G_{entropy} = RT \ln(2) = 0.4$ kcal/mol.

Factors other than solubility differences or the C2 symmetry of glycoluril amount to a change in binding energy compared with that of biotin of $\Delta\Delta G_{residual} = \Delta\Delta G_{binding} - \Delta\Delta G_{solvation} - \Delta\Delta G_{entropy} = 10.4 + 1.8 - 0.4 = 11.8$ kcal/mol. The reduced affinity for glycoluril predominantly reflects the missing interactions involving the valerate moiety. There is no hydrogen bond in streptavidin−glycoluril corresponding to that between the carboxylate oxygen of biotin and the $O\gamma_{Ser88}$ in streptavidin−biotin, and the hydrophobic interactions involving the valerate methylene groups are also missing or replaced by fewer and weaker van der Waals interactions involving the sulfate oxygens. The hydrophobic interactions between the valerate methylene groups and the Trp79 side chain in streptavidin−biotin are large, estimated by analysis of the binding of 2-iminobiotin to the Ala79 streptavidin mutant to be 5.4 to 8.2 kcal/mol.[39] There are also van der Waals interactions between the valerate methylene groups and the Trp120 and Leu110 side chains and the main chain carbonyl of Val47 in streptavidin−biotin. The enthalpic contribution of the valerate group outweighs the entropic cost of ordering it. The latter entropy change involves ordering atoms associated with four bonds, rotation about each of which yields three low-energy conformations. The change is thus estimated as $4 \times RT \ln(3) = 2.5$ kcal/mol.

Binding of Glycoluril versus Biotin Entails Compensating Changes in Hydrogen Bonding and Hydrophobic Interactions Involving the Left Portions of the Bicyclic Rings. All hydrogen bonds involving the right ring in streptavidin−biotin have corresponding hydrogen bonds in streptavidin−glycoluril of comparable lengths. In the left ring, the shorter $O1-O\gamma1_{Thr90}$ hydrogen bond is a more favorable enthalpic component to the binding of glycoluril than is the corresponding weak $S1-O\gamma1_{Thr90}$ hydrogen bond to the binding of biotin. The stronger hydrogen bond in streptavidin−glycoluril may be offset locally in streptavidin−biotin by better hydrophobic interactions involving the sulfur of the $S1-O\gamma1_{Thr90}$ hydrogen bond. The $S1-C\zeta2_{Trp79}$ and $S1-C\zeta2_{Trp92}$ distances are shorter in streptavidin−biotin, 3.65 and 3.89 Å, respectively, than are the corresponding $C1-C\zeta2_{Trp79}$ and $C1-C\zeta2_{Trp92}$ distances in streptavidin−biotin, 4.19 and 4.34 Å, respectively. The $N5H \rightarrow \pi_{Trp108}$ hydrogen bond in streptavidin−glycoluril is replaced with hydrophobic interactions invloving the C5 atom in streptavidin−biotin.

Water and Ions Play Key Roles in Binding of Ligands to Streptavidin. Water plays a versatile and ubiquitous role in chemistry and biology, not only as an integral component of protein structure but also as a mediator of molecular recognition and a determinant of ligand affinity in physiological environments. In the complexes of streptavidin with cyclo-Ac-[CHPQFC]-NH_2, and other cyclic or linear peptides discovered by phage display, ordered waters mediate ligand binding.[20] The enthalpy of water-mediated protein−ligand hydrogen bonds is often substantial,[54] and the entropy change of displacing bound waters upon ligand complexation sometimes dominates ligand binding.[55] Entropy also contributes to the binding of glycoluril and biotin in this study, since they displace water from the binding site of apostreptavidin.[20]

Ions play pivotal roles in mediating many important protein−ligand recognition events.[56−62] The crystal structures of ligand complexes in streptavidin and other systems underscore the often critical and unpredicted contributions of both anions and cations as mediators of ligand binding in biology. In streptavidin−glycoluril, a sulfate-mediated hydrogen bond replaces the hydrogen bond between the valerate oxygen of biotin and NH_{Asn49} in streptavidin−biotin. Since the enthalpy contribution of solvent-mediated hydrogen bonds can be similar to those made directly by a ligand,[54] this replacement is not necessarily enthalpically unfavorable. However the entropy change accompanying the ordering of the sulfate is unfavorable. For water, the entropic cost of ordering one molecule reduces ligand affinity by a value estimated to be as little as 0.3 kcal/mol,[63] <0.6 kcal/mol,[54] or 1.4 kcal/mol,[64] or as much as 3.0 to 4.8 kcal/mol,[64,65] depending on the degree of ordering of the water. For sulfate with more atoms than water, the corresponding values are expected to be somewhat larger. Since the same K_d for binding of glycoluril was observed in the presence and absence of sulfate, water and/or other ions undoubtedly mediate binding of glycoluril in the absence of sulfate.

Conclusion

Phage display enables discovery of ligands to macromolecules involved in diverse biological processes. Identification of a small molecule streptavidin-binding ligand incorporating structural and hydrogen-bonding features common to both the nonpeptide biotin ligand and a peptide ligand discovered by phage display demonstrates the potential of structure-based design for engineering ligands targeted to proteins. The structure and affinity of glycoluril and comparisons with biotin also provide useful data to test and tune algorithms aimed at docking potential ligands into protein binding sites and at predicting ligand affinities. Crystallography combined with phage display or other combinatorial chemistry methodologies promises to expedite structure-based ligand and drug design.

JA961357Z

(54) Watson, K. A.; Mitchel, E. P.; Johnson, L. N.; Son, J. C.; Bichard, C. J. F.; Orchard, M. G.; Fleet, G. W. J.; Oikonomakos, N. G.; Leonidas, D. D.; Kontou, M.; Papageorgioui, A. *Biochemistry* **1994**, *33*, 5745−5758.

(55) Scapin, G.; Young, A. C. M.; Kromminga, A.; Veerkamp, J. H.; Gordon, J. I.; Sacchettini, J. C. *Mol. Cell. Biochem.* **1993**, *123*, 3−13.

(56) Willett, W. S.; Gillmor, S. A.; Perona, J. J.; Fletterick, R. J.; Craik, C. S. *Biochemistry* **1995**, *34*, 2172−2180.

(57) Smith, J. W. The Structural Basis of Integran-Ligand (RGD) Interactions. VI. *The Role of Divalent Cations in Ligand Binding in Integrans: Molecular and Biological responses to the Extracellular Matrix*; Cheresh, D. A., Mecham, R. P., Eds.; Academic Press: San Diego, 1994; pp 11−18.

(58) Husain, S. S. *J. Biol. Chem.* **1993**, *268*, 8574−8579.

(59) Christianson, D. W. *Adv. Prot. Chem.* **1991**, *42*, 280−355.

(60) Kelley, J. L.; McLean, E. W.; Crouch, R. C.; Averett, D. R.; Tuttle, J. V. *J. Med. Chem.* **1995**, *38*, 1005−1014.

(61) Matthews, D. J.; Wells, J. A. *Chem. Biol.* **1994**, *1*, 25−30.

(62) Somers, W.; Ultsch, M.; De Vos, A. M.; Kossiakoff, A. A. *Nature* **1994**, *372*, 478−481.

(63) Pearlman, D. A.; Connelly, P. R. *J. Mol. Biol.* **1995**, *248*, 696−717.

(64) Smith, G. M.; Alexander, R. S.; Christianson, D. W.; McKeever, B. M.; Ponticello, G. S.; Springer, J. P.; Randall, W. C.; Baldwin, J. J.; Habecker, C. N. *Protein Sci.* **1994**, *3*, 118−125.

(65) Rich, D. H. *J. Med. Chem.* **1985**, *28*, 263−273.

(66) Brünger, A. T. *Nature* **1992**, *355*, 472−474.