Structure-Based Design of Synthetic Azobenzene Ligands for Streptavidin

P. C. Weber,* M. W. Pantoliano, D. M. Simons,† and F. R. Salemme‡

Contribution from the Crystallography and Biophysical Chemistry Group, DuPont Merck Pharmaceutical Company, P.O. Box 80228, Wilmington, Delaware 19880-0228, Medical Products Department, E. I. DuPont de Nemours and Company, Glasgow Site 709, P.O. Box 6101, Newark, Delaware 19714-6101, and 3-Dimensional Pharmaceuticals, Inc., 3700 Market Street, Philadelphia, Pennsylvania 19104

Received September 23, 1993•

Abstract: A structure-based design strategy was used to elaborate modifications of the dye 2-[(4′-hydroxyphenyl)-azo]benzoate (HABA) in order to increase its affinity for the biotin-binding site of streptavidin. Analogs of HABA incorporating methyl, *tert*-butyl, and methoxy hydrophobic substituent groups, as well as the compound in which a naphthyl ring was substituted for the phenyl ring, bound to streptavidin with up to 165-fold greater affinity than unmodified HABA. Binding was most enhanced for HABA derivatives with methyl substituents at the 3′- and 5′-hydroxyphenyl ring positions, where the free energy of binding increased from −5.27 kcal/mol for HABA to −7.23 kcal/mol for 2-[(3′-methyl-4′-hydroxyphenyl)azo]benzoate (3′-methyl-HABA), and to −8.29 kcal/mol for 2-[(3′,5′-dimethyl-4′-hydroxyphenyl)azo]benzoate (3′,5′-dimethyl-HABA). Measurements of ligand binding energetics using isothermal titration calorimetry indicate that the increases in ligand affinity arise primarily from favorable entropy contributions to the free energy of ligand binding. High-resolution X-ray structural studies (3′-methyl-HABA, space group $I222$, a = 95.2 Å, b = 106.5 Å, c = 47.9 Å, crystallographic R-value = 0.17 to 1.8-Å resolution; 3′,5′-dimethyl-HABA, R = 0.18 to 1.65-Å resolution) suggest that both water displacement from the ligand binding site and retention of ligand flexibility in the bound state may contribute to overall favorable entropy changes on ligand binding. Comparative thermodynamic binding data, derived from differential scanning calorimetry measurements, together with high-resolution crystallographic structures, are also reported for streptavidin:ligand complexes of 2-[(3′-*tert*-butyl-4′-hydroxyphenyl)-azo]benzoate (K_a = 2.5 × 10^5 M^{-1}, R = 0.18 to 2.0-Å resolution), 2-[(3′,5′-dimethoxy-4′-hydroxyphenyl)azo]benzoate (K_a = 3.5 × 10^6 M^{-1}, R = 0.19 to 2.2-Å resolution), and 2-[(4′-hydroxynaphthyl)azo]benzoate (K_a = 2.8 × 10^6 M^{-1}, R = 0.18 to 1.8-Å resolution).

Introduction

Structure-based ligand design provides a powerful method for the discovery and optimization of drugs with high specificity and potency. The method is predicated upon the direct, three-dimensional visualization of a lead compound bound in the active site of a therapeutic target molecule such as an enzyme, receptor, or nucleic acid, from which modifications can be elaborated with enhanced binding properties. The strategy has been successfully used to discover potent inhibitors of a number of enzymes,[1] on the basis of both diversification of substrate analogs and computer searches of organic structural databases. Although these studies utilized a variety of computational approaches to assist compound design and estimate ligand binding energies, there is presently only an imperfect understanding of the relative importance of the diverse factors that contribute to the free energy of ligand binding to proteins.

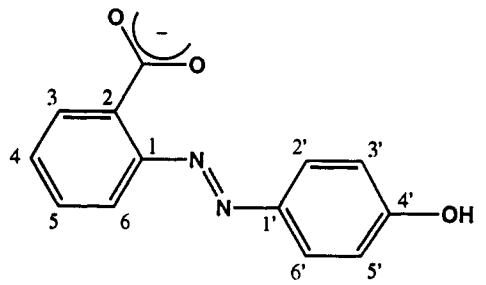


Figure 1. Structure and atomic numbering of 2-[(4′-hydroxyphenyl)-azo]benzoate (HABA).

The present work describes the high-resolution X-ray crystal structures and binding thermodynamics for a series of organic dye molecules elaborated through structure-based design methods to enhance their binding affinity to the biotin-binding site in streptavidin. Previous structural studies have shown that the dye, 2-[(4′-hydroxyphenyl)azo]benzoate (HABA, Figure 1), binds in the biotin-binding site of streptavidin and that several polar groups of the protein which interact with the biotin ureido group make structurally similar interactions with the HABA carboxylate group.[2] Nevertheless, in contrast to the binding of biotin which occurs with a large favorable enthalpy change ($\Delta H°$ = −32.0 kcal/mol),[2] thermodynamic measurements[2] of HABA binding

* Author to whom correspondence should be addressed at DuPont Merck Pharmaceutical Co.
† E. I. DuPont de Nemours and Co.
‡ 3-Dimensional Pharmaceuticals, Inc.
• Abstract published in *Advance ACS Abstracts*, February 1, 1994.

(1) (a) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M. *Science* **1990**, *249*, 527–533. (b) Appelt, K.; Bacquet, R. J.; Bartlett, C. A.; Booth, C. L. J.; Freer, S. T.; Fuhry, M. A. M.; Gehring, M. R.; Herrmann, S. M.; Howland, E. F.; Janson, C. A.; Jones, T. R.; Kan, C.; Kathardekar, V.; Lewis, K. K.; Marzoni, G. P.; Matthews, D. A.; Mohr, C.; Moomaw, E. W.; Morse, C. A.; Oatley, S. J.; Ogden, R. C.; Reddy, M. R.; Reich, S. H.; Schoettlin, W. S.; Smith, W. W.; Varney, M. D.; Villafranca, J. E.; Ward, R. W.; Webber, S.; Webber, S. E.; Welsh, K. M.; White, J. *J. Med. Chem.* **1991**, *34*, 1925–1934. (c) Ealick, S. E.; Babu, Y. S.; Bugg, C. E.; Erion, M. D.; Guida, W. C.; Montgomery, J. A.; Secrist, J. A., III. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11540–11544. (d) Shoichet, B. K.; Stroud, R. M.; Santi, D. V.; Kuntz, I. D.; Perry, K. M. *Science* **1993**, *259*, 1445–1450. (e) von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan, T. V.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. *Nature* **1993**, *363*, 418–423. (f) Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Springer, J. P.; Gautheron, P.; Grove, J.; Mallorga, P.; Viader, M.-P.; McKeever, B. M.; Navia, M. A. *J. Med. Chem.* **1989**, *32*, 2510–2513.

(2) Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R. *J. Am. Chem. Soc.* **1992**, *114*, 3197–3200.

0002-7863/94/1516-2717$04.50/0

Table 1. Thermodynamic Parameters for the Binding of Synthetic HABA Analogs to Streptavidin

ligand	T_m (°C)	$K_a^{T_m}$ (M^{-1})[a] (x 10^3)	K_a (M^{-1})[b] (x 10^3)	n	$\Delta G°_{bind}$ (kcal/mol)	$\Delta H°_{bind}$ (kcal/mol)	$T\Delta S°_{bind}$ (kcal/mol)
apostreptavidin	84.1						
2′,6′-dimethyl-HABA	88.1	16					
HABA[c]	90.2	31	7.3 (±0.5)	1.12 (±0.05)	−5.27	+1.70	+6.97
3′-*tert*-butyl-HABA	96.5	251					
3′-methyl-HABA	101.9	1590	200 (±30)	1.13 (±0.14)	−7.23	+1.28	+8.51
3′,5′-di-*tert*-butyl-HABA	103.2	2500					
naphthyl-HABA	103.4	2800					
3′,5′-dimethoxy-HABA	104.5	3450					
3′,5′-dimethyl-HABA	106.4	7700	1200 (±100)	1.00 (±0.03)	−8.29	+2.15	+10.44

[a] For a tight binding system of stoichiometry 1:1 in which there is one unfolding transition, one can estimate the binding constant at T_m from the following expression (ref 3):

$$K_a^{T_m} = \left[exp - \frac{\Delta H_u}{R}\left[\frac{1}{T_m} - \frac{1}{T_0}\right] + \frac{\Delta C_{pu}}{R}\left[\ln\left\{\frac{T_m}{T_0}\right\} + \frac{T_0}{T_m} - 1\right]\right]/[L]_{T_m}$$

where $K_a^{T_m}$ is the ligand association constant at T_m, T_m is the midpoint for the protein unfolding transition in presence of ligand (K), T_0 is the midpoint for the protein unfolding transition in absence of ligand (T_0 = 357.25 K for apostreptavidin), ΔH_u is the enthalpy of protein unfolding in the absence of ligand at T_0 (ΔH_u = 80.0 kcal/mol), ΔC_{pu} is the change in heat capacity upon protein unfolding in the absence of ligand (ΔC_{pu} = 0.95 kcal/mol K), $[L]_{T_m}$ is the free ligand concentration at T_m, and R is the gas constant. [b] Determined by isothermal titration calorimetry. Experiments were performed at 25 °C in unbuffered 0.1 M KCl adjusted to pH 6.9. The reported values are the mean for repeated experiments and errors are ±1 SD. The data were analyzed with the assumption of one set of sites with no cooperativity between subunits of the tetrameric streptavidin. [c] Data from ref 2.

show its energetics to be dominated by a favorable entropy change ($\Delta H°$ = +1.70 kcal/mol; $T\Delta S°$ = 6.97 kcal/mol), suggesting that hydrophobic effects are the predominant contributors to the HABA binding affinity. Here we describe the structural and thermodynamic properties of HABA derivatives whose hydrophobic character has been modified by addition of aliphatic and/or aromatic functional groups whose placement was guided by the crystal structure of the HABA:streptavidin complex.[2] Using this approach, derivatives were designed that preserved favorable properties of the HABA dye framework but introduced new functional groups at positions that enhanced binding affinity as much as 165-fold relative to HABA, the initial compound in the series.

Experimental Section

Ligand Synthesis. Substituted azobenzene HABA derivatives were synthesized by coupling diazotized 4-aminobenzoic acid to the appropriate phenol or naphthol derivatives (Aldrich Chemical, Milwaukee, WI). Reaction products were purified by multiple recrystallizations, and chemical composition and structure were confirmed by NMR spectrometry, and elemental and mass spectrometric analyses.

Ligand Binding Measurements. *Streptomyces avidinii* streptavidin was purchased from Calbiochem (La Jolla, CA) and used without further purification. Ligand affinity was initially screened by differential scanning calorimetric (DSC) measurements, where the observed unfolding transition of the protein:ligand complex can be used to estimate the ligand association constant at the complex melting temperature.[3] DSC measurements were performed with a Hart 4207 DSC heat conduction scanning microcalorimeter incorporating two matched pairs of removable ampoules and interfaced to an IBM personal computer. DSC scans ranged from 20 to 140 °C at a scan rate of 60 °C/h. Lyophilized protein was dissolved in 0.1 M *N,N′*-piperazinebis[2-ethanesulfonic acid] (PIPES), pH 7.0, in presence of a 2-fold molar excess of ligand per subunit of streptavidin. Final ampule protein concentrations were approximately 5×10^{-3} M as determined by amino acid composition analysis. Each experiment comprised four segments: (1) a first upward scan from 20 to 140 °C, (2) a first downward scan from 140 to 20 °C, (3) a second upward scan from 20 to 140 °C, and (4) a second downward scan to return to 20 °C. Since the thermal unfolding transitions of streptavidin are irreversible, as was reported for avidin,[4] the power input from the second upward scan was subtracted from that of the first upward scan to obtain the excess power input for the unfolding transitions. A sigmoidal curve was fit to the observed thermal transition data using a least squares computer fit of the pre- and post-transitional baselines to linear equations in T and α, where α is the fractional area under the transition curve at any particular T.[3] T_m values were measured as the midpoint in the unfolding transition for the various streptavidin:ligand complexes and are listed in Table 1. The calorimetric enthalpy of unfolding, ΔH_u, and change in heat capacity of unfolding, ΔC_{pu}, were measured as described previously.[5]

The binding energetics for two of the HABA derivatives, 3′-methyl- and 3′,5′-dimethyl-HABA, were further analyzed by isothermal titration calorimetry. Streptavidin solutions were titrated by addition of 18 × 15 μL aliquots of ligand solution at 26 °C in 5-min intervals using a MicroCal Omega titration calorimeter (Northampton, MA). The rate of the reaction of HABA derivatives with streptavidin was sufficiently fast under the conditions employed so that equilibrium was reestablished before the next ligand injection. The binding parameters, K_a, $\Delta H°$, and n (stoichiometry per subunit), were obtained through nonlinear least squares fit of the observed reaction heat for each titration step.[6] Lyophilized streptavidin was dissolved in unbuffered 100 mM KCl at a concentration between 0.16 and 0.75 mM (subunits) and dialyzed overnight against this same solution. All isothermal titration calorimetry experiments were conducted in 100 mM KCl in the absence of buffer to avoid heat effects due to ionization of buffer components. The pH was adjusted using small amounts of 1 M NaOH and 1 M HCl, and the protein was equilibrated with this solution by dialysis. The pH 6.9 was chosen for experiments to insure the full protonation of the dye phenolic hydroxyl (pK_a of 8.2–8.5 for HABA[7]), to minimize heat effects due to ligand ionization. An additional control experiment, diluting the ligand into 0.1 M KCl solution, was performed to correct for small contributions of the heat of dilution. Protein concentrations were determined by amino acid composition analyses. The solid ligands were dissolved in the dialyzate that resulted from the streptavidin dialysis to insure composition identity with respect to all solution components other than the reactants.

The small magnitude of the experimental signal (Figure 2) produced in the mixing experiments required the use of relatively high concentrations of streptavidin (0.16–0.75 mM subunits) to provide adequate signal-to-noise. The ligand concentration was 10 times higher than the streptavidin concentration to insure that at least 80% of the ligand per injection was bound to the protein for the first few injections. Moreover, these conditions also insured that at least 80% of the binding sites would be saturated after the last injection so that the binding reaction was followed over a wide range of fractional saturation. The thermodynamic binding parameters determined by isothermal titration calorimetry for the 3′-methyl- and 3′,5′-dimethyl-HABA derivatives are listed in Table 1.

Crystallographic Studies. To facilitate structural comparisons between different streptavidin:ligand complexes, X-ray diffraction experiments

(3) (a) Brandts, J. F.; Lin, L. *Biochemistry* **1990**, *29*, 6927–6940. (b) Schwarz, F. P. *Biochemistry* **1988**, *27*, 8429–8436.
(4) Donovan, J. W.; Ross, K. D. *Biochemistry* **1973**, *12*, 512–517.
(5) Pantoliano, M. W.; Whitlow, M.; Wood, J. F.; Dodd, S. W.; Hardman, K. D.; Rollence, M. L.; Bryan, P. N. *Biochemistry* **1989**, *28*, 7205–7213.
(6) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L. *Anal. Biochem.* **1989**, *179*, 131–137.
(7) (a) Baxter, J. H. *Arch. Biochim. Biophys. Acta* **1964**, *108*, 375–383. (b) Thomas, E. W.; Merlin, J. C. *Spectrochim. Acta* **1979**, *35A*, 1251–1255.

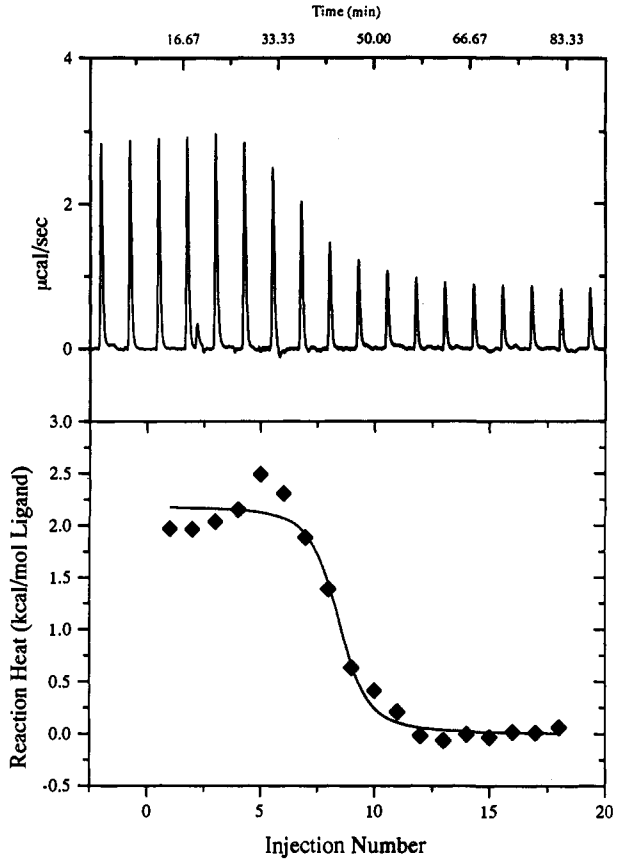


Figure 2. Characterization of 3′,5′-dimethyl-HABA binding to streptavidin by isothermal titration calorimetry. A 1.60 mM solution of 3′,5′-dimethyl-HABA was titrated into a 161 μM streptavidin (subunits) solution using an 18 × 15-μL injection schedule spaced at 5-min intervals (T = 25.1 °C). Each 15-μL injection was delivered over a 15-s duration. The streptavidin solution had been dialyzed against 2 × 1 L of 0.10 M KCl (adjusted to pH 6.9), and the 3′,5′-dimethyl-HABA solution was prepared gravimetrically by dissolving the recrystallized solid in the same dialyzate. (Top panel) Raw titration data for streptavidin showing the 18 individual changes in reaction heat, $Q_{(i)}$, after each injection of 3′,5′-dimethyl-HABA. The rate of reaction was sufficiently fast under these conditions that the baseline and equilibrium were reestablished before the next injection. (Lower panel) Resultant binding isotherm for 3′,5′-dimethyl-HABA titration of streptavidin after integration of the area under each injection peak and subtraction of the blank. The solid line represents a nonlinear least squares fit of the reaction heat for each injection ($\Delta Q_{(i)}$) with the assumption of a single binding site consisting of the three fitting parameters, n (stoichiometry per subunit), K_a, and $\Delta H°_{bind}$. In addition, independent noninteracting binding sites on the streptavidin tetramer were assumed for these fits. This experiment was carried out three times, and a summary of the binding parameters is described in Table 1.

were carried out with an orthorhombic crystal form of apostreptavidin[8] previously used for structure determinations of several streptavidin:ligand complexes.[2,9] In this crystal form, the biotin-binding sites of streptavidin are accessible to solvent channels so that ligands can be diffused into the crystal where complex formation can occur. Orthorhombic crystals (space group $I222$, dimer per asymmetric unit) of apostreptavidin grown from 35–45% saturated ammonium sulfate solutions at 30 °C were soaked in saturating concentrations of the ligand to prepare complex crystals. After soaking for at least 24 h, the clear apostreptavidin crystals assumed the yellow to red colors of the dye ligands due to protein:ligand complex formation.

X-ray diffraction data for complex crystals were collected using a Siemens imaging proportional counter and reduced to integrated intensities using XENGEN data reduction software.[10] Crystals were stable in the X-ray beam, and one crystal was used to collect a complete data set for each complex. Table 2 lists the crystallographic data parameters for streptavidin:ligand complex crystals.

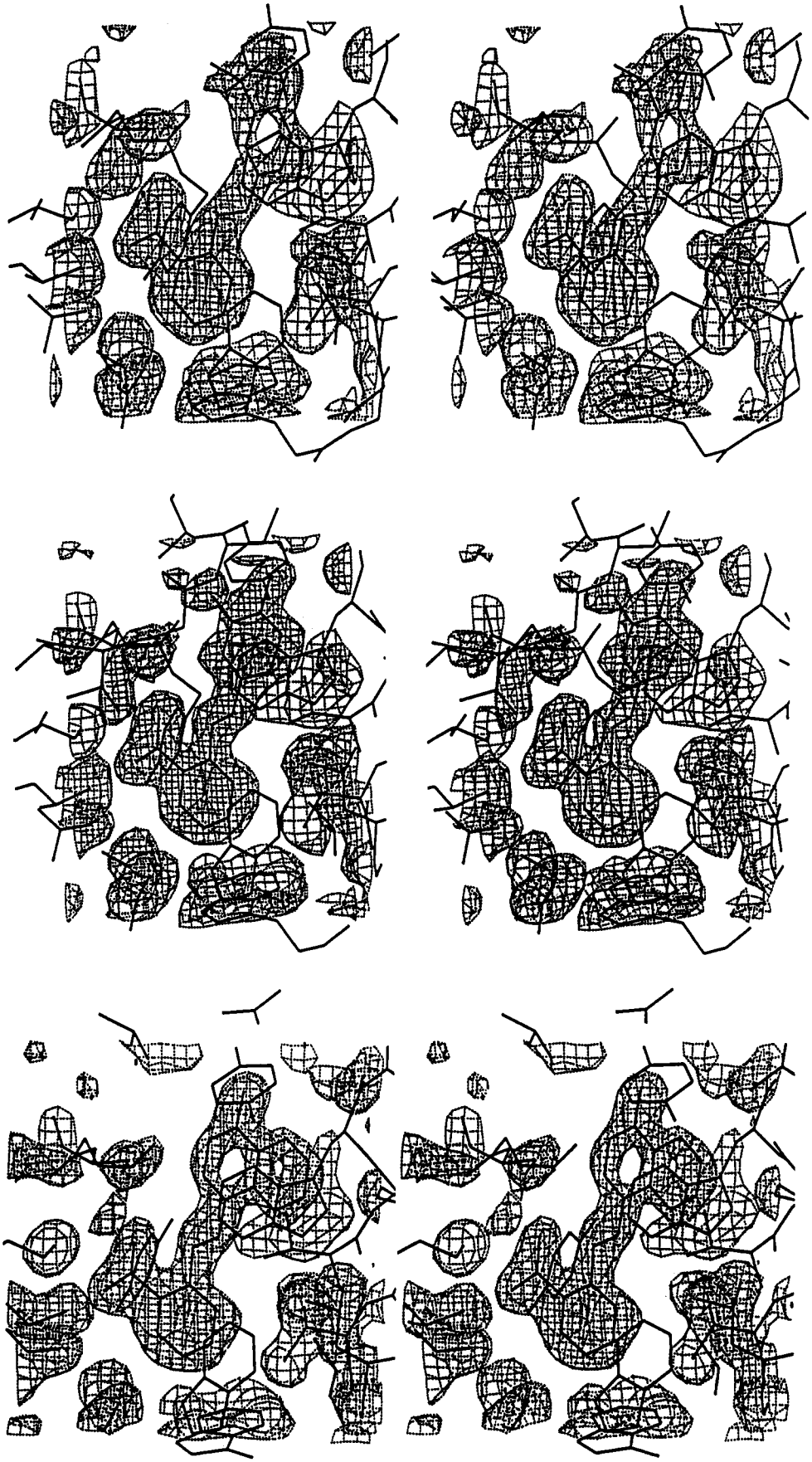
The structures of streptavidin:ligand complexes were refined at high resolution using restrained least squares methods,[11] alternating with manual rebuilding cycles into $(F_o - F_c)\alpha_{calc}$ and $(2F_o - F_c)\alpha_{calc}$ electron density maps displayed with the graphics program FRODO.[12] Chemically identical monomers in the crystallographic asymmetric unit were refined independently, starting with ligand-free models. Solvent molecules were introduced as they emerged as peaks greater than 3σ from the $(F_o - F_c)\alpha_{calc}$ electron density maps during successive refinement cycles. Ligand refinement parameters were derived from the small molecule crystal structure of 4-[(4′-hydroxyphenyl)azo]benzoic acid.[13] For the azobenzene ligands, aromatic ring geometry was constrained during the refinement, although planarity constraints were omitted for the diazo linkage (N–N bond, Figure 1) between the aromatic rings. Most of the protein structure corresponding to "core" streptavidin (residues 13–138[8]) is well-defined in the electron density maps, with the exception of residues 66–68 and carboxy terminal residues 134–138, which are not included in the protein model. The orientation of the various ligands was unambiguously defined in the electron density maps (Figure 3). Crystallographic refinement and model geometry statistics are listed in Table 2. Coordinates have been deposited in the Brookhaven Protein Data Bank.[14]

Results and Discussion

Previous thermodynamic and crystallographic studies of HABA binding to streptavidin[2] were used to guide the design of HABA analogs. Since it was desirable to retain spectroscopic properties of HABA in the modified ligands, which have potential utility as indicators in biotechnological applications,[7,15] the ligand modification strategy was focused on enhancing hydrophobic contributions to ligand binding through the introduction of aliphatic substituents on dye aromatic rings, which are expected to have minimal effects on dye electronic properties. The relatively tight packing of protein side chains about the benzoic acid portion of HABA (Figure 4) suggested that addition of substituents to this ring would be sterically restricted, thus focusing attention on the hydroxyphenyl ring, which is partially exposed to solvent in the HABA:streptavidin crystal structure. Two of four possible substitution sites on the hydroxyphenyl ring are buried in the biotin-binding site, consistent with the decreased binding affinity of the 2′,6′-dimethyl-HABA analog (Table 1). Subsequent HABA analogs were substituted at either or both the 3′- and 5′-carbons of the hydroxyphenyl ring (Figure 1) which, in the crystal structure of the HABA:streptavidin complex,[2] are respectively adjacent a small hydrophobic pocket in the protein or partially exposed to solvent.

Changes in ligand binding affinity for modified HABA derivatives were initially estimated from differential scanning calorimetry (DSC) measurements. Measurements of the melting temperatures of the streptavidin:ligand complexes as determined by DSC ranked the relative affinities of the HABA analogs in the order HABA < 3′-*tert*-butyl-HABA < 3′-methyl-HABA < 3′,5′-di-*tert*-butyl-HABA < naphthyl-HABA < 3′,5′-dimethoxy-HABA < 3′,5′-dimethyl-HABA. By incorporating values for additional protein physical parameters, variations in melting temperature of protein:ligand complexes can be related to ligand binding affinity.[3,4] Estimates of K_a values at the respective complex melting temperatures derived using this treatment are given in Table 1 for the ligands listed above.

(8) Pahler, A.; Hendrickson, W. A.; Kolks, M. A. G.; Argarana, C. E.; Cantor, C. R. *J. Biol. Chem.* **1987**, *262*, 13933–13937.
(9) Weber, P. C.; Pantoliano, M. W.; Thompson, L. D. *Biochemistry* **1992**, *31*, 9350–9354.
(10) Howard, A. J.; Gilliland, G. L.; Finzel, B. C.; Poulos, T. L.; Ohlendorf, D. H.; Salemme, F. R. *J. Appl. Crystallogr.* **1987**, *20*, 383–387.
(11) Hendrickson, W. A.; Konnert, J. H. *Biomolecular Structure, Function, Conformation and Evolution;* R. Srinivasan, Ed.; Oxford University Press: Oxford, U.K., 1980; pp 43–57.
(12) Jones, T. A. *J. Appl. Crystallogr.* **1978**, *11*, 268–272.
(13) Harlow, R. L.; Simons, D. M.; Weber, P. C. *Acta Crystallogr.* **1992**, *C48*, 48–50.
(14) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rogers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535–542.
(15) Green, N. M. *Adv. Protein Chem.* **1975**, *29*, 85–133.



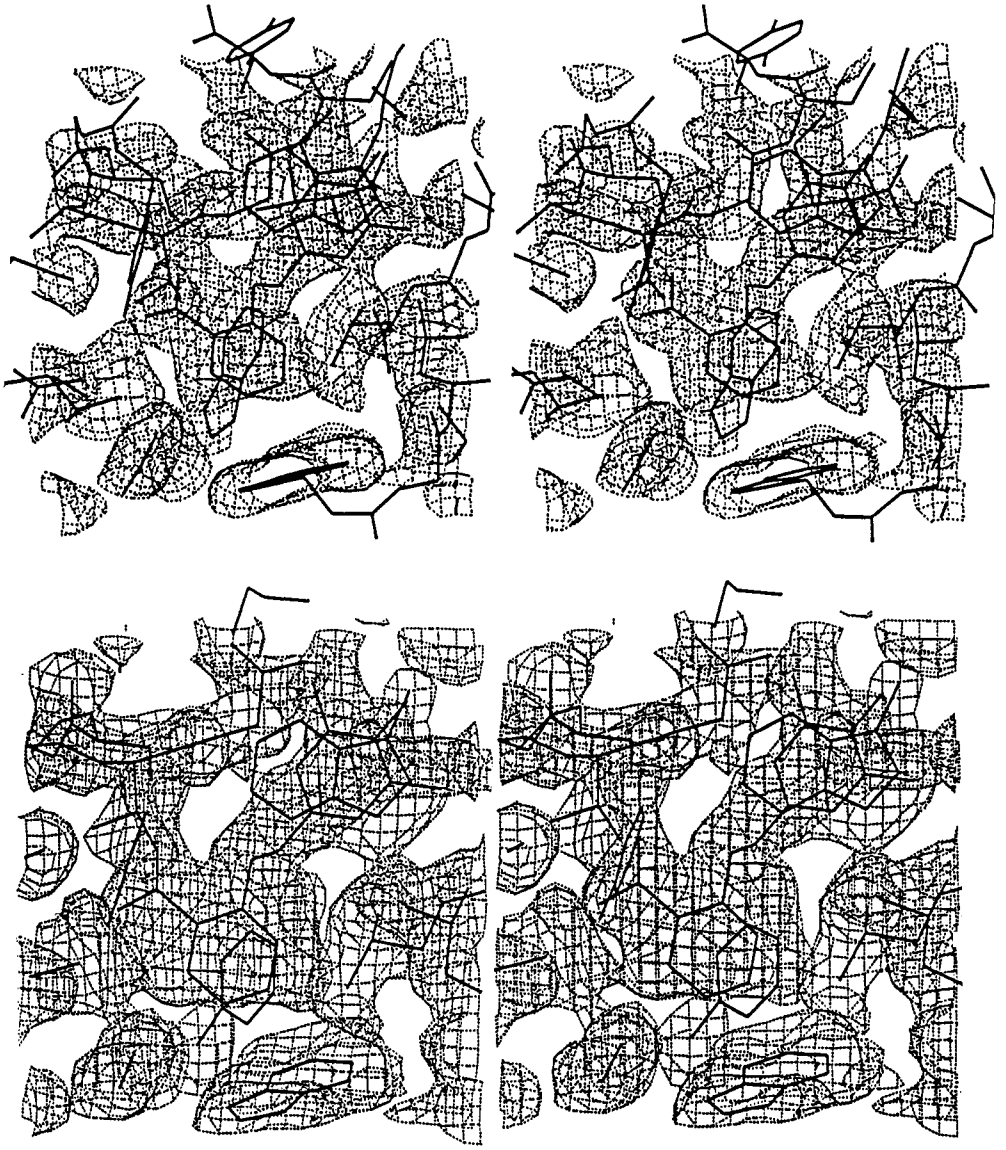


Figure 3. Electron density maps of streptavidin:ligand complexes. Stereoscopic views of $(2F_o - F_c)\alpha_{calc}$ electron density maps for streptavidin complexed with 3′-*tert*-butyl-HABA (2.0-Å resolution, left page top), 3′-methyl-HABA (1.8-Å resolution, left page center), naphthyl-HABA (1.8-Å resolution, left page bottom), 3′,5′-dimethoxy-HABA (2.0-Å resolution, right page top), and 3′,5′-dimethyl-HABA (1.7-Å resolution, right page bottom). The electron density in all five panels is presented for ligands in order of their increasing affinity for streptavidin. It is interesting that electron density quality does not correlate with compound affinity, i.e. the electron density envelope for 3′,5′-dimethyl-HABA, the most tightly bound ligand presented here, is less well-defined than that for 3′-methyl-HABA, a ligand that binds less tightly to streptavidin.

The energetics of binding 3′-methyl and 3′,5′-dimethyl analogs of HABA to streptavidin were studied in more detail using isothermal titrating calorimetry, which provides discrete estimates of enthalpy and entropy components of binding free energy. For both ligands, positive heats of reaction were observed on complex formation (Table 1, Figure 2). This behavior parallels unmodified HABA,[2] where enthalpy changes also oppose binding and the free energy of ligand binding is driven by favorable reaction entropy. Such entropy effects can produce substantial increases in ligand affinity (Table 1). In the best cases, addition of a single methyl group to HABA increased its affinity for streptavidin about 28-fold, and addition of two methyl groups increased the affinity over 165-fold. To better understand the structural details of the binding interactions, the crystal structures of several HABA analog:streptavidin complexes were determined and refined at high resolution.

Crystallographic studies showed qualitatively similar binding features for all HABA analogs bound to streptavidin. The dyes are oriented with the benzoic acid portion at the bottom of the biotin-binding pocket, with one edge of the hydroxyphenyl ring buried in the protein interior and the other partially exposed to the solvent (Figure 4). A key ionic interaction between ligand and protein is made by one of the dye carboxylate oxygens, which is situated in a pocket formed by three hydrogen-bond donating groups, Tyr43, Asn23, and Ser27, whose side chains are oriented to optimally stabilize a tetrahedral oxyanion.[2,16] The other HABA carboxylate oxygen is hydrogen bonded to Ser45. The aromatic rings of the dye occupy the hydrophobic interior of the biotin-binding site. The benzoic acid-containing ring is packed in a hydrophobic pocket formed by side chains of Thr90, and tryptophan residues 92 and 108, and the side chain of Trp120 from a diad-related subunit of the tetramer. This aromatic ring also packs perpendicular to the Trp108 indole side chain to form an energetically favorable edge-to-face packing interaction.[17] The hydroxyphenyl ring is sandwiched between the aromatic ring of Trp79 at the base of the pocket and loop residues 48–51. Residues of this loop are disordered in the apostreptavidin structure but become ordered in the biotin:streptavidin complex owing in part

(16) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. *Science* **1989**, *243*, 85–88.

(17) (a) Gould, R. O.; Gray, A. M.; Taylor, P.; Walkinshaw, M. D. *J. Am. Chem. Soc.* **1985**, *107*, 5921–5929. (b) Singh, J.; Thornton, J. M. *FEBS Lett.* **1985**, *191*, 1–6.

Table 2. Summary of Data Collection and Crystallographic Refinement Statistics for Steptavidin:Ligand Complexes

	HABA[a]	3′-methyl-HABA	3′,5′-dimethyl-HABA	naphthyl-HABA	3′-*tert*-butyl-HABA	3′,5′-dimethoxy-HABA
unit cell parameters						
a (Å)	95.2	95.2	95.6	95.2	94.2	94.5
b (Å)	106.5	106.5	106.0	106.5	104.8	104.7
c (Å)	47.9	47.9	47.9	47.9	47.0	47.4
number of observations	81 351	137 848	156 325	103 092	57 038	62 676
(unique reflections)	(19 376)	(22 702)	(28 764)	(20 894)	(14 692)	(14 498)
R_{sym}[b]	7.7	6.4	6.4	7.9	7.1	5.9
resolution range (Å)	10–1.78	10–1.8	10–1.65	10–1.8	10–2.0	10–2.2
number of reflections	17 650 (1σ)	18 744 (3σ)	24 078 (1σ)	15 599 (3σ)	12 231 (1σ)	10 975 (1σ)
crystallographic *R*-factor	0.181	0.168	0.175	0.177	0.180	0.190
number of atoms						
protein	1765	1751	1788	1767	1732	1750
solvent	211	237	220	222	234	232
ligand	36	38	40	44	44	44
rms deviation in						
bond distance (Å)	0.014	0.019	0.024	0.017	0.013	0.011
angle distance (Å)	0.026	0.029	0.035	0.029	0.025	0.022
planar group distance (Å)	0.013	0.011	0.015	0.014	0.008	0.006
chiral volume (Å^3)	0.194	0.257	0.289	0.225	0.174	0.145

[a] From ref 2. [b] $R_{sym} = [\sum_{hkl}\sum_{i=1}^{N}|\langle I^{hkl}\rangle - \langle I_i^{hkl}\rangle|]/\sum_{hkl}\sum_{i=1}^{N} I_i^{hkl}$

to hydrogen-bond formation between one of the biotin carboxylate oxygens and the backbone NH of Asn49.[16] In the structures of the HABA series of compounds, the loop is only partially ordered owing to packing interactions between ligands and loop residues.

As illustrated in Figure 4, the polar interactions formed with the dye carboxylate group in the biotin-binding site, together with local packing interactions, result in near superposition of the benzoate rings of HABA and all HABA derivatives. In contrast, the modified hydroxyphenyl groups show some diversity in binding.

The differences in binding geometry of the substituted HABA analogs are illustrated in Figure 4 and are summarized in Table 3. For reference, it is noted that HABA binding involves the displacement of five water molecules spatially and/or temporally localized in the unligated binding site of the apostreptavidin structure.[2] In addition to the five localized water molecules released, HABA binding immobilizes a water that forms a bridging hydrogen bond between the phenyl hydroxyl group and the hydroxyl group of Tyr54. The binding of 3′-methyl-HABA retains virtually all of these features, reflecting the presence of a small hydrophobic pocket adjacent to the hydroxyphenyl ring that is sterically complementary to the methyl substituent. Replacement of the methyl with the larger *tert*-butyl group forces the hydroxyphenyl ring to translate out of the pocket, breaking the water-mediated hydrogen bonding interaction with Tyr54. The binding of naphthyl-HABA retains essentially all of the features of HABA binding, but the larger fused naphthyl substituent displaces a sixth immobilized water molecule. Interestingly, both 3′,5′-dimethoxy-HABA and 3′,5′-dimethyl-HABA show large ring displacements and loss of the Tyr54-to-water hydrogen bond but displace six water molecules. Comparison of these data (Table 3) with the binding measurements of Table 1 illustrates the complexity of interpreting the structural origins of binding affinity. For example, the ability of a substituted ligand to displace six water molecules from the binding appears to be more effective in creating a tight binding ligand than preserving the ligand-to-Tyr54 hydrogen bond, perhaps because formation of the latter actually involves immobilizing an additional water molecule.

The ability to unambiguously define solvent molecules which are displaced by ligand binding is a result of the combination of the unusual diffracting power of the streptavidin:ligand complex crystals coupled with extensive crystallographic refinement. Because refinement at resolutions with Bragg spacings less than 2.0 Å typically involves variations of isotropic *B*-values, as well as atomic positions, high-resolution refinement can additionally reveal features of static or dynamic molecular disorder.

In the present context, a striking feature of the electron density maps of the streptavidin:ligand complexes is the apparent difference in mobility between side chains in the biotin-binding site, which typically show atomic resolution, and those of the ligand, which manifest varying extents of disorder. The extent of spatial or temporal disorder in the crystal structure is a parameter which is optimized during crystallographic refinement and expressed as an isotropic *B*-value, where $B = 8\pi^2/3(\langle x^2\rangle)$ and $\langle x^2\rangle$ is the mean squared displacement of an atom about its average position in the crystal structure. Because there are several experimental parameters attached to individual crystalline specimens that potentially contribute to *B*-values,[18] Table 3 tabulates average ligand *B*-values relative to the six tryptophan rings in streptavidin in each refined structure. Although crystallographers may be accustomed to equating tight binding to rigidly fixed ligands, some of the tightest binding streptavidin ligands, such as 3′,5′-dimethoxy-HABA and 3′,5′-dimethyl-HABA (Table 1), are characterized by significantly larger *B*-values than surrounding aromatic groups of the binding site. This suggests an unusual degree of flexibility for these ligands even when bound to the protein, including, in the case of the symmetrically substituted HABA analogs, the possibility of rotation of the hydroxyphenyl ring about the nitrogen-to-aromatic carbon bond (N–C1′, Figure 1).

Conclusions

Several points emerge from this study which are relevant to both a fundamental understanding of the thermodynamics of protein:ligand interactions and practical applications such as drug design.

The small positive enthalpy changes associated with binding HABA derivatives suggest that specific polar interactions formed with groups in the biotin-binding site reflect a compromise with compensating solvent interactions. These interactions consequently appear to structurally direct, but do not substantially stabilize, complex formation. This reflects a situation where the observed structural interactions, while not giving rise to a large negative enthalpy, are nevertheless those which contribute to the minimization of the total binding free energy.

It has been quantitatively shown that entropy effects resulting from the introduction of hydrophobic substituents that pack tightly in a protein binding site, and/or displace "immobilized" water molecules, can substantially enhance ligand binding affinity. In

(18) (a) Finzel, B. C.; Salemme, F. R. *Nature* **1985**, *315*, 686–688. (b) Hartmann, H.; Parak, F.; Steigemann, W.; Petsko, G. A.; Ponzi, D. R.; Frauenfelder, H. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4967–4971.

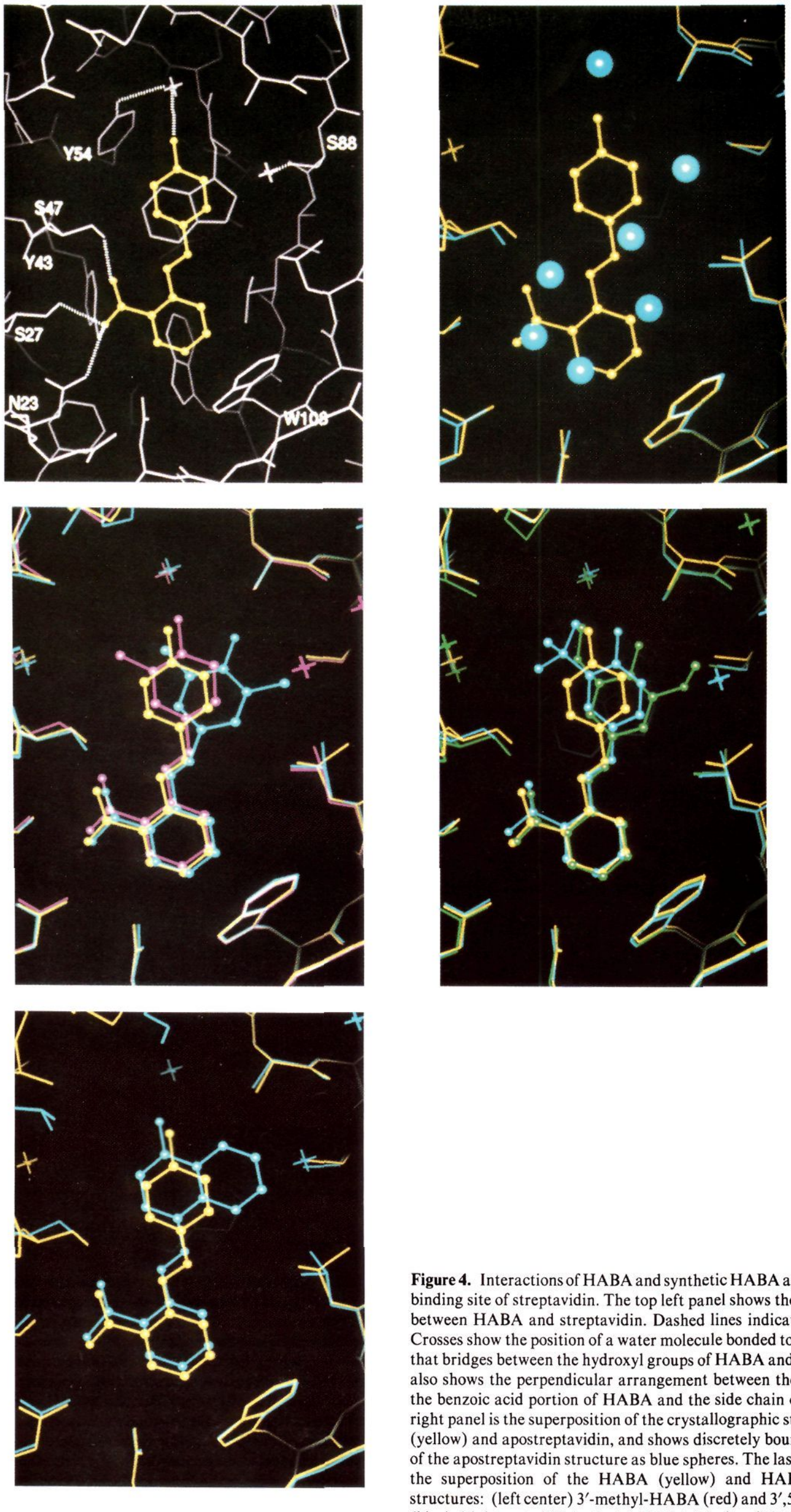


Figure 4. Interactions of HABA and synthetic HABA analogs in the biotin-binding site of streptavidin. The top left panel shows the interactions made between HABA and streptavidin. Dashed lines indicate hydrogen bonds. Crosses show the position of a water molecule bonded to Ser88 and another that bridges between the hydroxyl groups of HABA and Tyr54. This figure also shows the perpendicular arrangement between the aromatic rings of the benzoic acid portion of HABA and the side chain of Trp108. The top right panel is the superposition of the crystallographic structures of HABA (yellow) and apostreptavidin, and shows discretely bound water molecules of the apostreptavidin structure as blue spheres. The last three panels show the superposition of the HABA (yellow) and HABA analog crystal structures: (left center) 3′-methyl-HABA (red) and 3′,5′-dimethyl-HABA (blue), (right center) 3′-*tert*-butyl-HABA (blue) and 3′,5′-dimethoxy-HABA (green), and (bottom left) naphthyl-HABA (blue).

Table 3. Comparative Structural Properties of HABA Analogs Binding to Streptavidin

ligand	ring superposition[a] (Å)	ligand–H_2O–Y54 H-bond formation	"ordered" H_2O molecules displaced	resolution (Å)[b]	average ligand atomic *B*-value (Å²)	ligand:protein *B*-value difference (Å²)[c]	ligand:protein *B*-value ratio[d]
HABA		yes	5	1.78	25.2	16.4	2.9
3′-*tert*-butyl-HABA	no (0.7 Å)	no	5	2.00	22.0	12.1	2.2
3′-methyl-HABA	yes (0.4 Å)	yes	5	1.80	18.9	9.8	2.1
naphthyl-HABA	yes (0.3 Å)	yes	6	1.80	22.5	10.4	1.9
3′,5′-dimethoxy-HABA	no (0.9 Å)	no	6	2.00	25.4	18.9	3.9
3′,5′-dimethyl-HABA	no (1.2 Å)	no	6	1.65	33.8	19.9	2.4

[a] Displacement of the center of the analog hydroxyphenyl ring relative to HABA. [b] Other factors being equal, crystallographic refinement at increasingly high resolutions tends to increase the dynamic range of refined atomic *B*-values. [c] The *B*-value averaged over all ligand atoms, minus the *B*-value averaged over aromatic atoms of six tryptophan side chains per streptavidin monomer (averaged over two monomers per crystal asymmetric unit). [d] The ratio of the *B*-value averaged over all ligand atoms to the *B*-value averaged over all aromatic atoms of tryptophan side chains.

the case of 3′-methyl-HABA, the single methyl substituent enhanced binding free energy by −1.98 kcal/mol and increased ligand binding affinity by a factor of 28 relative to unmodified HABA. For 3′,5′-dimethyl-HABA, binding free energy was enhanced by −3.02 kcal/mol, producing a greater than 165-fold increase in binding affinity relative to that of HABA. In the first case, the apparent increase in binding free energy per additional methyl substituent is 2-fold higher than the 0.8–1.0 kcal/mol intrinsic binding energy for methyl groups found by Andrews et al.[19] in their analysis of over 200 drugs and enzyme inhibitors.

Crystallographic evidence suggests that water release from the protein binding site is a predominant contributor to the favorable entropy change, particularly as this appears to be the only feature of streptavidin:HABA complex structures that is shared among the tighter binding ligands such as naphthyl-HABA, 3′,5′-dimethoxy-HABA, and 3′,5′-dimethyl-HABA (Tables 1 and 3). Other details of the binding interaction, such as the precise packing orientation of the hydroxyphenyl ring or ability of the hydroxyphenyl hydroxyl group to form a water-mediated hydrogen bond to a protein tyrosine residue, appear energetically less important or are compensated by ligand solvation or other unidentified effects.

In contrast to the biotin:streptavidin interaction, where ligand binding is accompanied by both overall ligand immobilization (as manifested by the similarity in biotin *B*-values and those of surrounding protein atoms[2]) and immobilization of a loop which appears disordered in apostreptavidin, several of the HABA analogs, including those which bind most tightly, retain substantial ligand mobility within the binding pocket, and the biotin-binding loop of streptavidin remains somewhat disordered even when the HABA series of ligands are bound. In this context, it is notable that resonance Raman studies of HABA and 3′,5′-dimethyl-HABA indicate that ligand binding to avidin, a close homolog of streptavidin,[15] is accompanied by nearly complete conversion to the hydrazone form.[7] Hydrazone formation could potentially introduce additional flexibility into HABA derivatives when bound, which might facilitate rotation of symmetrical groups like the 3′,5′-dimethoxy- and 3′,5′-dimethylhydroxyphenyl groups about the N–C1′ bond (Figure 1). These groups manifest the highest relative *B*-values in the refined crystal structures, and it is interesting to speculate that this mobility could help compensate for losses in bound ligand translational entropy and contribute to the high binding affinities for these ligands.

The structure-based design strategy to enhance ligand affinity for HABA relies on the incorporation of both thermodynamic and crystallographic information. These types of data have also been determined for a streptavidin-binding peptide whose sequence was obtained by random screening methods and whose free energy of binding to streptavidin is similar to that of HABA.[9] Unlike the interactions of HABA and streptavidin, the thermodynamic binding parameters show that the peptide-binding reaction is favored enthalpically and disfavored entropically, and suggest that attempts to decrease ligand flexibility in solution might lead to increases in its binding affinity. While the thermodynamic data for the HABA- and peptide-binding reactions indicate alternate synthetic strategies for increasing ligand affinity, the value of integration of thermodynamic data into a structure-based ligand design process is clearly demonstrated.

Finally, as noted in previous structure-based design studies,[1,20] relatively conservative changes in ligand structure can produce significant and unanticipated alterations in protein:ligand interactions. In the present case, this behavior is exemplified by the very different pattern of protein interactions made by such closely related molecules as 3′-methyl-HABA and 3′,5′-dimethyl-HABA (Figure 4 and Table 3) and reiterates the potential difficulties of interpreting binding data using quantitative structure activity relationship (QSAR) methods[21] which implicitly assume that similar ligands preserve common sets of protein interactions.

Acknowledgment. We thank Drs. B. C. Finzel and S. T. Freer for assistance in setting geometric restraints in PROFFT, and also E. Asel, D. Heitmeyer, L. Janvier, and G. Koukedis for technical assistance.

(19) Andrews, P. R.; Craik, D. J.; Martin, J. L. *J. Med. Chem.* **1984**, *27*, 1648–1657.

(20) Aguilar, C. F.; Thomas, P. J.; Moss, D. S.; Mills, A.; Palmer, R. A. *Biochim. Biophys. Acta* **1991**, *1118*, 6–20.

(21) Martin, Y. C. *Quantitative Drug Design: A Critical Introduction*; Dekker: New York, 1978.