Protein Science (1997), *6*:1157–1166. Cambridge University Press. Printed in the USA.

Structural studies of the streptavidin binding loop

STEFANIE FREITAG,[1,2] ISOLDE LE TRONG,[1] LISA KLUMB,[2]
PATRICK S. STAYTON,[2] AND RONALD E. STENKAMP[1]

[1]Department of Biological Structure and Biomolecular Structure Center, University of Washington, Seattle, Washington 98195-7742
[2]Department of Bioengineering, University of Washington, Seattle, Washington 98195-7962

(RECEIVED February 25, 1997; ACCEPTED March 13, 1997)

Abstract

The streptavidin–biotin complex provides the basis for many important biotechnological applications and is an interesting model system for studying high-affinity protein–ligand interactions. We report here crystallographic studies elucidating the conformation of the flexible binding loop of streptavidin (residues 45 to 52) in the unbound and bound forms. The crystal structures of unbound streptavidin have been determined in two monoclinic crystal forms. The binding loop generally adopts an open conformation in the unbound species. In one subunit of one crystal form, the flexible loop adopts the closed conformation and an analysis of packing interactions suggests that protein–protein contacts stabilize the closed loop conformation. In the other crystal form all loops adopt an open conformation. Co-crystallization of streptavidin and biotin resulted in two additional, different crystal forms, with ligand bound in all four binding sites of the first crystal form and biotin bound in only two subunits in a second. The major change associated with binding of biotin is the closure of the surface loop incorporating residues 45 to 52. Residues 49 to 52 display a 3_{10} helical conformation in unbound subunits of our structures as opposed to the disordered loops observed in other structure determinations of streptavidin. In addition, the open conformation is stabilized by a β-sheet hydrogen bond between residues 45 and 52, which cannot occur in the closed conformation. The 3_{10} helix is observed in nearly all unbound subunits of both the co-crystallized and ligand-free structures. An analysis of the temperature factors of the binding loop regions suggests that the mobility of the closed loops in the complexed structures is lower than in the open loops of the ligand-free structures. The two biotin bound subunits in the tetramer found in the MONO-b1 crystal form are those that contribute Trp 120 across their respective binding pockets, suggesting a structural link between these binding sites in the tetramer. However, there are no obvious signatures of binding site communication observed upon ligand binding, such as quaternary structure changes or shifts in the region of Trp 120. These studies demonstrate that while crystallographic packing interactions can stabilize both the open and closed forms of the flexible loop, in their absence the loop is open in the unbound state and closed in the presence of biotin. If present in solution, the helical structure in the open loop conformation could moderate the entropic penalty associated with biotin binding by contributing an order-to-disorder component to the loop closure.

Keywords: biotin; flexible loop; molecular recognition; protein–ligand interactions; streptavidin

Streptavidin is a homotetrameric 159 residue protein isolated from *Streptomyces avidinii* (Chaiet & Wolf, 1964). Each monomer of the protein binds one molecule of the vitamin biotin non-covalently with an exceptionally high affinity ($K_a \sim 10^{13}\ M^{-1}$) (Green, 1975). This fact has been exploited to devise powerful and widely used tools for affinity chromatography, biochemical assays, and many other applications (Bayer & Wilchek, 1990). The dissection of the factors involved in streptavidin–biotin binding is of high interest, and several approaches have been undertaken to elucidate the structure–function relationships governing this high-affinity protein–ligand complex. Concerted structural and energetic studies represent a particularly important pathway to achieve this goal. The determination of the crystal structure of the core–streptavidin–biotin complex using Multiple Anomalous Dispersion methods (Hendrickson et al., 1989) and Multiple Isomorphous Replacement (Weber et al., 1989) provided the first key structural insight. (Core–streptavidin contains residues 13–139 (Hendrickson et al., 1989; this work) or residues 13–133 (Weber et al., 1989)). Several crystal forms of streptavidin have also been described (Pähler et al., 1987; Weber et al., 1987; Hemming et al., 1995). These different crystal forms are important because they allow the determination and comparison of molecular structures influenced by a variety of crystal packing interactions. Structural studies have also been carried

Reprint requests to: Ronald E. Stenkamp, Department of Biological Structure and Biomolecular Structure Center, University of Washington, Box 357420, Seattle, Washington 98195-7420; e-mail: stenkamp@u.washington.edu.

out in the related avidin–biotin system, where the binding motifs closely resemble those found in the streptavidin–biotin complex (Livnah et al., 1993; Pugliese et al., 1993). Here we report new information concerning conformational differences upon binding of biotin.

The streptavidin–biotin pair provides a remarkable example of a high-affinity protein–ligand system. The aim of our studies is to dissect the factors governing the binding using a combination of site-directed mutagenesis, kinetic and thermodynamic analyses, and X-ray crystallography. We have previously gained insight into the contribution of aromatic contacts to the biotin binding equilibrium and the activation barrier to dissociation through characterization of site-directed mutants at Trp 79, 108, and 120 (Chilkoti et al., 1995; Chilkoti & Stayton, 1995). Chemical modification studies have also been utilized to study aromatic contacts to biotin in the analogous avidin system (Morag et al., 1996). In addition to biotin, other studies have been aimed at characterizing the interactions of streptavidin with alternative ligands. Peptide libraries have been screened to identify peptides that specifically bind to streptavidin (Devlin et al., 1990; Lam et al., 1991; Giebel et al., 1995), and complexes of streptavidin and linear or cyclic peptides have been studied thermodynamically and structurally (Weber et al., 1992a, 1995; Katz, 1995, 1996; Katz et al., 1995a, 1995b; Schmidt et al., 1996). Energetic and structural studies have also been conducted with derivatives of the dye 2-(4′-hydroxyphenylazo)benzoic acid (HABA) (Weber et al., 1992b, 1994, 1995).

Streptavidin forms a homo-tetramer with D_2 symmetry (222) (Fig. 1A). Each monomer of streptavidin folds into an eight-stranded antiparallel β barrel (Fig. 1B). Biotin is bound in the open end of the twisted barrel, and one surface loop folds over the binding site when biotin is bound. Streptavidin displays several commonly observed molecular recognition motifs in the interaction with biotin. These include hydrophobic and van der Waals interactions and an extended hydrogen bonding network. The flexible binding loop consisting of residues 45 to 52 is disordered in the unbound state in the orthorhombic and tetragonal crystal forms

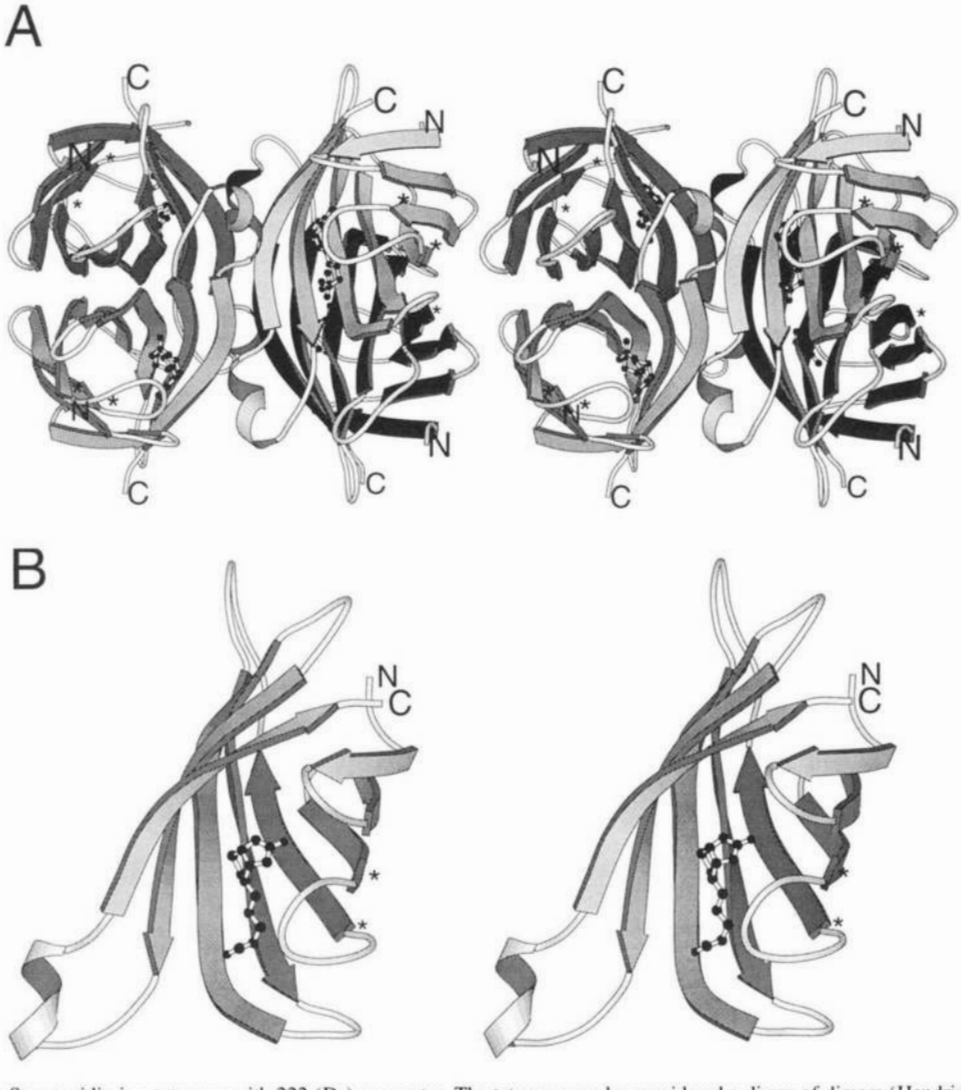


Fig. 1. A: Streptavidin is a tetramer with 222 (D_2) symmetry. The tetramer can be considered a dimer of dimers (Hendrickson et al., 1989; Weber et al., 1989) due to the extensive inter-subunit contacts shown by the pairs of subunits on the right (1 and 2) and left (3 and 4) sides of the figure. **B:** One monomer of streptavidin forms a β barrel with extended hairpin loops. Biotin is bound at the open barrel side and a surface loop (residues 45 to 52) folds over the ligand. Asterisks denote the termini of that loop.

(Weber et al., 1989, 1992b) and found in a closed conformation in the bound state of the co-crystallized tetragonal crystal form (Weber et al., 1989). Our initial studies of streptavidin in two different monoclinic crystal forms indicated that this loop can be found in an open or closed conformation, independent of the presence of biotin. Closer examination of these structures, and of a new co-crystallized complex, has provided new details about the loop in the open form and the effects of crystallographic packing interactions on the loop conformation.

Results

We collected diffraction data from five different monoclinic wild-type streptavidin crystals (Table 1). Crystals of recombinant core–streptavidin were grown at pH 4.5. Soaking of these crystals in a pH 7.5 buffer before data collection gave us the opportunity to compare the structure of streptavidin at two different pH values. We also obtained a new crystal form, MONO-4, by crystallizing the protein from a higher concentration MPD solution. Co-crystallization of the protein with 1.2 molar and 2.5 molar excess biotin, respectively, resulted in two different crystal forms, MONO-b1 and MONO-b2, containing streptavidin–biotin complexes. The five data sets are comparable with respect to the resolution, completeness, and quality of the data (Table 2). In all five cases (multiple crystal forms), the protein crystallized in the monoclinic space group $P2_1$ with one tetramer in the asymmetric unit. The structures were solved using molecular replacement methods implemented in X-PLOR (Brünger, 1992a). Initial refinement steps were carried out with the same program. In later stages the β-test version of SHELXL-96 (Sheldrick, 1996) was used to refine the models. Refinement results are presented in Tables 2 and 3.

The ligand-free streptavidin structures

Our attempts to grow suitable monoclinic crystals of streptavidin at physiological pH for X-ray diffraction experiments were not successful. The comparison of the structural models based on the data sets I (soaked at pH 7.5) and II (pH 4.5) shows no major differences in the protein structures caused by the difference in pH. Only minor deviations of the main-chain in loop regions and minor movements in several side chains—mainly in loop regions—are observed. These differences do not alter intra- or intermolecular interactions in this crystal form. A least-squares fit of the Cα atoms for the tetramer (4 × 65 β-sheet residues; see Materials and methods) in both structures gives a root-mean-square deviation (RMSD) of 0.2 Å.

The binding loops of the tetramer (residues 45 to 52) in the MONO-1 crystal form do not display the same conformation in all four subunits. The binding surface loop was reported to be disordered in other ligand-free streptavidin structures (Weber et al., 1989). In subunit 1 of the MONO-1 crystal form, this loop adopts a closed conformation that was previously reported only in streptavidin–ligand complexes (Weber et al., 1995) and a structure of streptavidin with two sulfate ions in the binding site (Katz, 1995). The surface loops in the other three subunits are partially disordered. Only residues 45 and 49 to 52 could be traced in electron density maps and have been refined in a more open conformation. Figure 2 shows a representation of the streptavidin tetramer (circles) for the unbound crystal forms, MONO-1 and MONO-4. Subunits 1 and 2, and 3 and 4, respectively, build the dimer pairs (Fig. 1A). Subunits 1 and 4, and 2 and 3, respectively, are connected via the Trp 120 residues that participate in the biotin binding site of the adjacent subunit. The missing sectors of the circles represent the biotin binding sites with the binding loops in the closed or open conformation. Disordered parts of loops are depicted as dotted lines. Crystal packing interactions cause the loop in the first subunit of structures I and II (MONO-1 crystal form; see Table 1) to adopt a closed conformation. A fit of that loop in the open conformation in subunit 1 would lead to steric hindrance in the crystal packing between side chains of two surface loops (residues 45 to 52 and 81 to 84, respectively) of subunit 1 and subunit 2 of the adjacent tetramer in the packing. Similar hindrances by packing interactions are found for a modeled closed loop in subunit 2. Crystal packing interactions do not allow the loop to close for subunit 2. The surface binding loops in subunits 3 and 4 are not involved in any packing interactions.

Table 1. *Crystallization conditions for ligand-free and biotin-bound core-streptavidin*

Structure	Crystal form[a]	Crystallization conditions[b]	Morphology	Unit cell
I	MONO-1	48% MPD 5.5 h soaking in 0.1 M Hepes pH 7.5	Rods	a = 58.9, b = 88.2, c = 47.4 Å β = 98.7°
II	MONO-1	47% MPD pH 4.5	Rods	a = 58.5, b = 87.1, c = 47.0 Å β = 98.9°
III	MONO-4	60% MPD soaking in 20 mM acetate pH 4.5	Blocks	a = 47.8, b = 65.9, c = 82.1 Å β = 97.4°
IV	MONO-b1	50% MPD pH 4.5 1.2 M excess of biotin	Rods	a = 57.8, b = 85.2, c = 46.6 Å β = 100.1°
V	MONO-b2	50% MPD pH 4.5 2.5 M excess of biotin	Rectangular plates	a = 52.8, b = 100.1, c = 52.1 Å β = 112.6°

[a]The MONO-1 crystal form was observed before by Pähler et al. (1987) and described as crystal form B1. (Crystal forms MONO-2 and MONO-3, not included in this paper, were described as B2 and B3 in the same reference.) MONO-b1 shows the same unit cell parameters as MONO-1. MONO-4 and MONO-b2 are new crystal forms. All crystal forms contain one tetramer in the asymmetric unit. The space group is $P2_1$ (two tetramers in the unit cell).

[b]Crystals for structures IV and V were co-crystallized with biotin.

Table 2. *Summary of the data collection and refinement for ligand-free and biotin-bound streptavidin*

Structure	R_{merge}[a]	$\langle I/\sigma(I)\rangle$ total	Space group	Resolution (Å)[b] [max.]	Completeness (%)	R1/R1$_{free}$[c,d]	No. of refined water molecules
I[e]	0.052	10.9	$P2_1$	1.9 [1.9]	78	0.155/0.231	209
II[f]	0.049	11.7	$P2_1$	2.0 [1.9]	79	0.142/0.246	211
III[f]	0.053	14.8	$P2_1$	1.8 [1.8]	79	0.157/0.227	226
IV[e]	0.042	14.1	$P2_1$	2.0 [1.9]	83	0.172/0.276	234
V[e]	0.066	9.25	$P2_1$	2.1 [2.1]	83	0.144/0.237	302

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|F^2_{hkl,i} - \overline{F^2_{hkl}}|/\Sigma_{hkl}\Sigma_i|F^2_{hkl,i}|$.
[b] The resolution refers to the limit where $\langle I/\sigma(I)\rangle$ was greater than 2 and data set completeness >50% in the last resolution shell (0.1 Å). In addition, the maximum observed resolution is reported.
[c] R1 for reflections with $I > 2\sigma(I)$. The final R1 value was calculated for all data in this range including the reference data.
[d] $R1 = \Sigma_{hkl}||F_o| - |F_c||/\Sigma_{hkl}|F_o|$.
[e] Data processed with Xengen (Howard et al., 1987).
[f] Data processed with Raxis software (Higashi, 1990).

The binding loops in all four subunits of structure III are found in the open conformation. In subunits 1 and 3 only residues 48 to 52 and 45, 46, 51, 52, respectively, are defined in difference electron density maps. The other residues appear to be disordered. An open conformation for this surface loop was also described for various streptavidin–peptide complexes (Katz, 1995, 1996; Katz et al., 1995a, 1995b; Schmidt et al., 1996). The open loops described by Katz are not identical to those in I, II, and III, but they are similar in their tendency to adopt a conformation that is opened away from the binding site. The structures of streptavidin in the MONO-1 and MONO-4 crystal forms are very similar. Least-squares fits of 4 × 65 β-sheet residues (Cα) of structures I and II, respectively, on structure III give RMSDs of 0.3 and 0.3 Å, respectively. The biggest differences are in the binding loop regions and C-termini.

Table 3. *Refined residues and average atomic displacement parameters in ligand-free and biotin-bound streptavidin*

Structure	Refined residues in subunit*‡ Average B-value for main chain atoms in italics, Å^2				Binding loop residues (45-52) refined in subunit (closed conformation is shaded)				Biotin refined in subunits			
	1	2	3	4	1	2	3	4	1	2	3	4
I	16-135	16-45	16-45	16-45	45-52	45,	45,	45,	-	-	-	-
		49-133	49-133	49-133		49-52	49-52	49-52				
	25.2	*23.6*	*22.6*	*26.3*	*44.8*	*37.6*	*45.8*	*54.5*				
II	16-135	16-45	16-45	16-45	45-52	45,	45,	45,	-	-	-	-
		49-133	49-133	49-133		49-52	49-52	49-52				
	28.8	*28.1*	*27.0*	*30.6*	*47.3*	*44.3*	*50.7*	*57.9*				
III	16-44	16-132	16-46	16-132	48-52	45-52	45,46	45-52	-	-	-	-
	48-132		51-132				51,52					
	28.9	*24.5*	*32.2*	*27.7*	*37.7*	*33.6*	*64.2*	*48.6*				
IV	16-132	16-44	16-45	16-133	45-52	49-52	45,	45-52	+	-	-	+
		49-133	49-133				49-52	(0.7)				(0.7)
	27.3	*30.1*	*31.3*	*31.9*	*28.2*	*32.8*	*71.7*	*30.2*	*25.8*			*22.2*
V	16-132	16-133	16-133	16-133	45-52	45-52	45-52	45-52	+	+	+	+
	18.6	*19.3*	*18.8*	*19.3*	*18.6*	*14.4*	*16.3*	*18.5*	*10.4*	*11.4*	*17.2*	*12.6*

( ) Occupancy if ≠ 1.0.
*Note: There are four subunits (one tetramer) in the asymmetric unit in each of the crystal forms.
‡Missing residue numbers imply disordered and unmodeled portions of the protein.

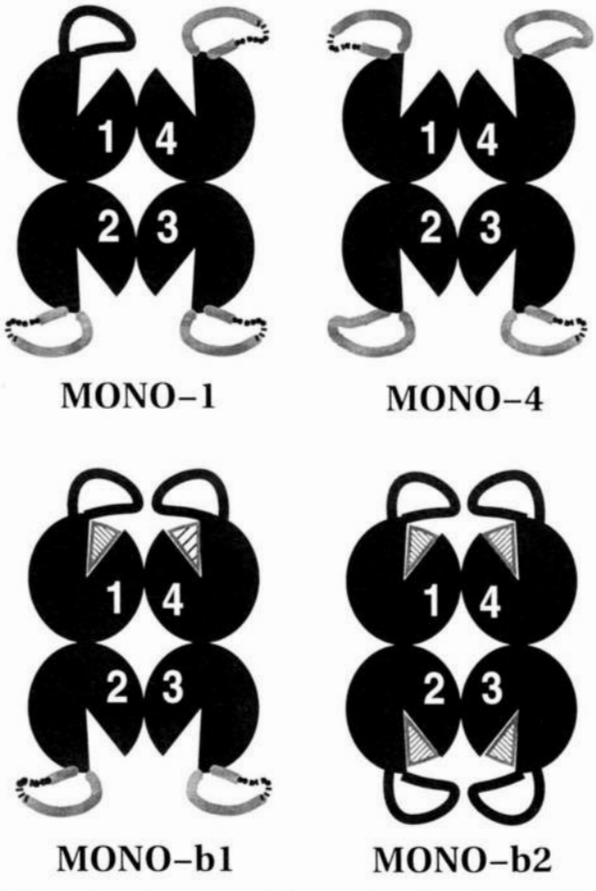


Fig. 2. The four observed crystal forms of ligand-free (top) and biotin-bound (bottom) streptavidin are depicted schematically to illustrate the behavior of the surface loop (residues 45 to 52) relative to the biotin binding site. The circles represent the streptavidin tetramer subunits with the binding sites (missing sectors). Subunits 1 and 2 , and 3 and 4, respectively, build the dimer pairs. Subunit 1 and 4, and 2 and 3, respectively, donate Trp 120 to each others binding site. The curved lines over the binding sites trace the loop conformations, with dotted lines representing disordered conformations. Triangles in the binding sites symbolize biotin.

The biotin-bound streptavidin structures

Soaking of the crystals of structure II in biotin-containing solutions for up to 24 h does not result in formation of crystalline streptavidin–biotin complexes. An increase in biotin concentration leads to dissolution of the crystals, indicating a conformational change that affects protein–protein contacts in the crystal. Co-crystallization of streptavidin and biotin at different biotin concentrations resulted in two different crystal forms, MONO-b1 and MONO-b2. MONO-b1 is isomorphous to the unbound MONO-1 form. The merging R-values for an isotropic scaling of the data sets for structures I and II, respectively, on IV are 0.136 and 0.126. The loop orientation and biotin distribution in IV and V are shown in Figure 2. Whereas biotin is bound in all four subunits in V with full occupancy, biotin is found only in subunits 1 and 4 in IV. As reflected by high B-factors, the biotin molecule and the residues of the closed binding loop in subunit 4 are not fully occupied in IV. The occupancy was refined to a value of 0.7 for the referred residues. Loop residues 49 to 52 (subunit 2) and 45, 49 to 52 (subunit 3) adopt the open conformation.

The binding of biotin alters the conformation of the binding loop (residues 45 to 52) in both crystal forms MONO-b1 and MONO-b2. When biotin is in the binding site, this surface loop is always in the closed conformation. Figure 3 shows a superposition of the loop in the open (red) and closed (black) conformation. A minor shift of the N-terminal loop (residues 24 to 27) in the direction of the bound ligand was also observed. This is not surprising, because this part of the molecule makes several important interactions with biotin. Asn 23, Glu 24, Leu 25, Gly 26, and Ser 27 contribute to the hydrogen bonding network. Additionally, this part of the protein neighbors the binding surface loop (residues 45 to 52) and could be influenced by its closure. The biggest shift is that of Leu 25, where the Cα atom moves 0.7–1.5 Å in the direction of biotin upon ligand association. The B-values of the loop residues 24 to 27 are lower in the bound state, corresponding to their lower mobility. The hydrogen bonding patterns of biotin and streptavidin in the monoclinic structures (Table 4) are very similar to those found in the tetragonal crystal form (Weber et al., 1989).

The RMSDs of a least-squares fit of the fully bound structure onto the unbound structure are 0.4 Å (65 Cα atoms of β sheet residues (see Materials and methods) of structures I and II on V) and 0.5 Å (structure III on V), respectively. These values are only slightly higher than the deviations between the unbound structures (0.2 and 0.3 Å, respectively). Figure 4 shows the tryptophan and Tyr 43 side chains of the binding pocket in a superposition of ligand-free structure II and bound structure V. No significant hydrogen bonding alterations or changes in tryptophan side-chain conformations in the binding site are observed other than those just described. The major differences upon binding are found in the binding loop regions including the hydrogen bonding residues Ser 45 and Glu 49. Comparisons of average B-values for the Trp residues in the unbound and biotin-bound structures show that Trp 120 is the most mobile of the tryptophan residues in the binding site and becomes more rigid when biotin is associated to the protein (Table 5). Striking examples for this result are the B-values for Trp 120 in structure IV, where biotin is bound only in subunits 1 and 4. The B-values of Trp 120 in these subunits are 27.6 and 26.0 Å^2, respectively, compared to 40.7 and 42.3 Å^2 in the unbound subunits 2 and 3.

Weber et al. (1989) reported quaternary structure changes upon biotin binding. The angle between the barrel domains increased when biotin was bound in the tetragonal crystal form and a superposition of their bound and unbound structures (67 Cα atoms of the β sheets) gave RMSDs of 0.7 Å for the tetramer and 0.3 Å after fitting individual subunits. For the structures described here, the fit of an individual subunit of structure V on I, II, or III, gives RMSDs of about 0.3 Å. The deviations of the other three subunits in these fits (0.4–1.1 Å) are always higher than the RMSDs for the "fitted" subunit. However, this is also observed for fits of unbound structures on each other, where an individually fit subunit shows RMSDs of 0.2 Å and the other three subunits give values in the range of 0.2–0.8 Å. In addition, a visual check of the superimposed models reveals no systematic change in the quaternary structure that can be associated with biotin binding.

Discussion

Structural comparison of the flexible loop in the unbound and bound state

Previous structural studies of streptavidin in a tetragonal crystal form demonstrated that the loop consisting of residues 45 to 52 was disordered in the apo state and ordered in the biotin-bound

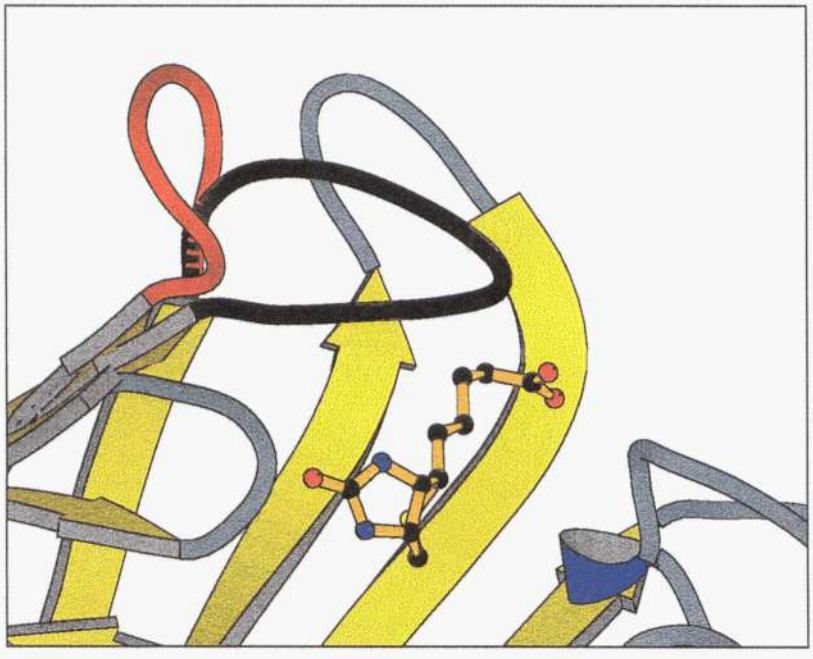
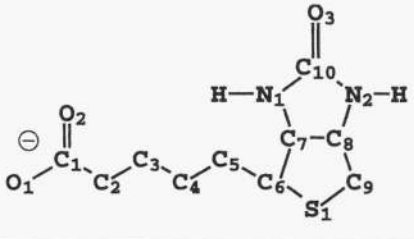


Fig. 3. Cα representation of a superposition of the binding loop region in subunit 2 of structure III (ligand free) on subunit 2 in structure V (biotin bound). This plot illustrates the relative open (red, unbound) and closed (black, biotin-bound) conformations of the binding loops.

state (Weber et al., 1989). Our initial crystallographic studies of the unbound and bound structures of streptavidin in different monoclinic crystal forms indicated that the open conformation of the loop was not always found in the absence of biotin. We have thus conducted a rigorous analysis of the loop conformation and the hydrogen bonding interactions in several different crystal forms and additionally characterized a new co-crystallized complex.

The loop is generally observed in the open conformation for the ligand-free structures, but packing interactions force the loop into the closed conformation in subunit 1 of crystal form MONO-1. The binding loop in streptavidin–biotin complexes always adopts the closed conformation. Figures 5 and 6 show the conformations of the open and closed binding loop in detail. The main-chain atoms of subunit 4 in the unbound structure III (Fig. 5) and subunit 2 in the fully bound structure V (Fig. 6) are depicted, as well as the main-chain hydrogen bonds in the loop region and loop–biotin interactions. The surface loop in the bound structure folds over the biotin (green) so that the backbone of residues 47, 48,

Table 4. *Hydrogen bonds of streptavidin and biotin in structure V (MONO-b2)*



Biotin atom	Streptavidin atom	Distance in Å to subunit			
		1	2	3	4
O1	Asn 49 N	2.8	3.1	2.5	2.7
O2	Ser 88 OG	2.8	2.7	3.1	3.0
S1	Thr 90 OG1	3.3	3.4	3.5	3.3
N1	Ser 45 OG	3.0	2.8	3.1	3.3
N2	Asp 128 OD2	2.8	2.8	2.7	2.7
O3	Asn 23 ND2	3.0	3.0	3.1	2.9
	Tyr 43 OH	2.8	2.7	2.6	2.9
	Ser 27 OG	2.6	2.7	2.5	2.3

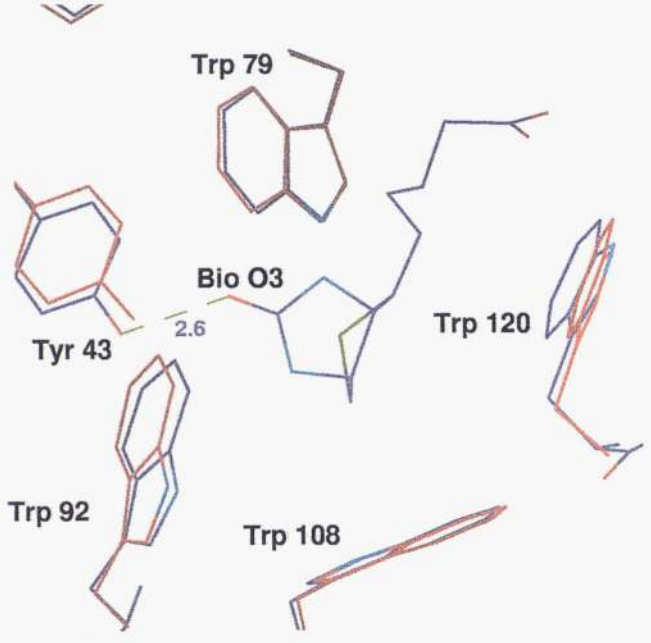


Fig. 4. Superposition of the tryptophan residues and Tyr 43 in the binding site of bound (blue, structure V, subunit 1) and unbound (red, structure II, subunit 1) streptavidin.

Table 5. *B-values of the Trp residues in unbound and biotin-bound streptavidin*

	Subunit			
	1	2	3	4
Structure I				
W79	18.8	18.7	17.5	20.4
W92	16.3	13.8	13.0	15.4
W108	16.6	18.4	17.4	17.7
W120	34.9	39.3	40.2	29.1
Structure II				
W79	22.5	18.8	24.1	21.2
W92	19.4	15.2	16.0	18.1
W108	21.1	20.4	18.9	20.6
W120	40.0	40.6	35.0	29.6
Structure III				
W79	22.1	21.3	26.8	25.3
W92	19.9	17.4	22.9	19.0
W108	19.8	19.8	23.0	18.8
W120	36.8	40.9	36.1	45.2
Structure IV				
W79	23.9	24.0	29.5	20.2
W92	21.9	21.0	17.5	23.0
W108	19.1	22.3	21.8	18.5
W120	27.6	40.7	42.3	26.0
Structure V				
W79	13.6	9.3	8.6	15.5
W92	12.1	10.8	9.9	11.0
W108	12.7	14.1	12.1	7.5
W120	16.0	14.9	14.9	14.2

and 49 runs parallel to the long axis of the bicyclic ring of biotin. Hydrogen bonding interactions to biotin (green) are formed by Tyr 43, Ser 45, and Glu 49. The main-chain atoms of Tyr 43 and Tyr 54 form two hydrogen bond interactions (blue) as in β-strands.

In the apo–streptavidin structure, considerable structural detail can be discerned in the open loop conformations of our monoclinic crystal forms. Figure 5 shows three additional main-chain hydrogen bonds for the surface loop in ligand-free streptavidin. The hydrogen bond between Ser 45 and Ser 52 extends the β-strands by one residue per strand. The geometry of the loop in the bound structure prevents this interaction. The hydrogen bonding pattern (Asn 49 O-Ser 52 N, Ala 50 O-Arg 53 N) and the main-chain conformational angles (ϕ, ψ) for residues 50 to 52 are characteristic for a 3_{10} helix. This pattern is not found in the closed conformation. The program PROMOTIF (Hutchinson & Thornton, 1996) describes residues 48 to 50 in the closed conformation as inverse γ turns. The ϕ angles vary in the range from -79.9 to $-87.8°$, the ψ angles from 68.8 to 79.9°. The other loop residues do not form regular secondary structure elements. For the open conformation, residues 50 to 52 are described as 3_{10} helix. The program PROCHECK (Laskowski et al., 1993) describes residues 49 and 53 as extensions of the 3_{10} helix. This helical part of the open loop is often observed in the unbound structures as in subunits 2, 3, and 4 in structures I and II as well as subunits 1, 2, and 4 in III and subunits 2 and 3 in V (Table 3).

The open and closed conformations are correlated with the atomic displacement parameters of the loop residues. A comparison of the B-values for the five structures (Table 3) shows that the B-values for V are much lower than for the others. The low B-values for V were observed with two refinement programs (X-PLOR and SHELXL-96), verifying that the difference in B-values is not a refinement artifact. The B-values of the loop residues in V (B-values from 14 to 19 Å^2) are in the range of the average main-chain B-values in the rigid β-sheet regions (19 Å^2). This corresponds to low mobility in these residues. A similar result was found for the two closed loops in structure IV (subunits 1 and 4) (B-values from 28 to 30 Å^2; main-chain B-values from 27 to 32 Å^2). Nearly all of the loops in the open conformations show distinctly higher B-values than the main-chain residues in the structures, indicating a higher mobility in the open loops. The exceptions to this are the binding loops in subunits 2 in structures III and IV. The B-values in the closed loop in subunit 1 of the unbound structures I (pH 4.5) and II (pH 7.5) are higher than those of the closed loops in the bound structures IV and V.

In summary, the loop residues have lower mobility when biotin is bound and the loop is closed. The loop is not totally disordered in the unbound state, but clearly shows increased mobility. The helical segment of the open loop conformation (residues 49 to 53) is observed in nearly every unbound subunit. If present in solution, the existence of a favored loop conformation with helical segments in the open structure would be expected to contribute different components to the equilibrium entropy and enthalpy of biotin binding when compared to a completely disordered loop. The ordering of the open loop conformation and the melting of the 3_{10} helix could lessen the conformational entropy loss associated with loop folding in the bound state. The loss of 3_{10} helix hydrogen bonding interactions could introduce an unfavorable enthalpic contribution, but the formation of several hydrogen bonding interactions to biotin will mitigate this term. We are currently characterizing site-directed mutants that probe the energetic contributions of this flexible surface loop to the binding of biotin.

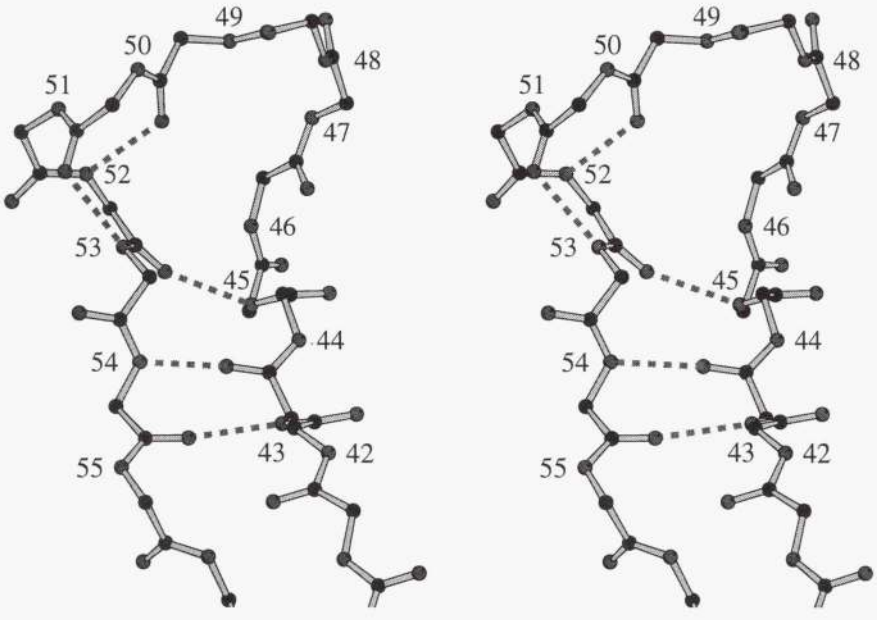


Fig. 5. Main-chain structure of the open loop in subunit 2 of structure III.

Implications for binding cooperativity from the co-crystal structures

Co-crystallization of streptavidin and biotin leads to two different crystal forms and structures. One is a fully bound structure V, that was obtained by co-crystallization with a 2.5 molar excess of biotin. The second structure (IV), obtained with 1.2 molar excess biotin, contains tetramers with biotin bound only in subunits 1 and 4. The restriction of biotin association to subunit 1 and partially to subunit 4 in this latter structure is not easily explained, given the high association constant for streptavidin and biotin. Virtually all of the binding pockets in the starting solution should be occupied by ligand molecules at both ratios of biotin. It seems unlikely that the concentration of the doubly occupied streptavidin–biotin complex in solution is high enough to support preferred crystal growth of this particular complex. One explanation for the sub-stoichiometric binding in this structure could be the low availability of ligand, as biotin crystals were observed in the crystallization setups. However, this would not account for the observation of a preferred distribution of biotin in structure IV. This structure is isomorphous with structures I and II. Crystal packing interactions promote the closed loop conformation in subunit 1, so this crystal environment is appropriate for a fully bound subunit. Likewise, the loop is open in subunit 2, requiring the presence of an unbound subunit at this position.

The loops in subunits 3 and 4 are open in the unbound structure, but there are no close protein neighbors in the lattice that require this conformation. Thus, the binding site could be occupied or unoccupied. We observe biotin bound to subunit 4, but not subunit 3.

In the streptavidin tetramer, adjacent binding sites are connected across the dimer–dimer interface through Trp 120, which contributes a key side-chain contact to biotin and which stabilizes the

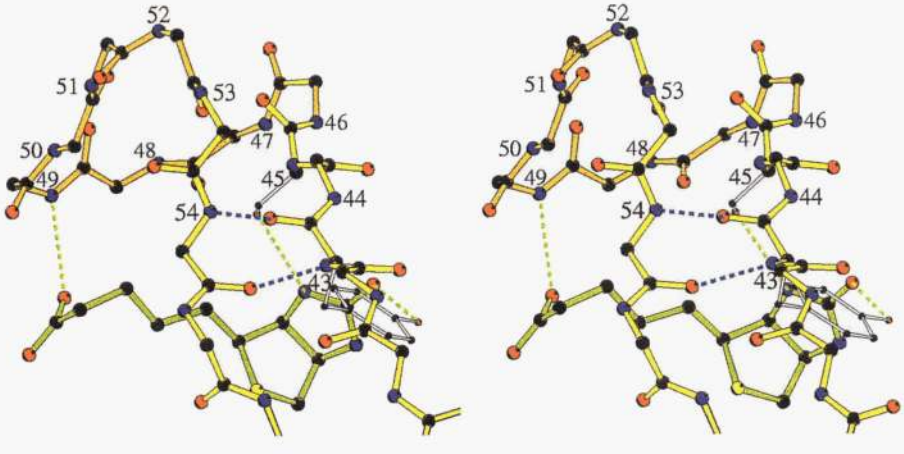


Fig. 6. Main-chain structure of the closed loop in subunit 2 of structure V.

tetramer in the bound state (Chilkoti & Stayton, 1995; Sano and Cantor, 1995) The asymmetry of binding site occupancy observed within structure IV raises the question of structural and energetic cooperativity, as subunits 1 and 4, and 2 and 3, respectively, represent the subunit pairs that contribute Trp 120 to adjacent binding pockets. Thus, the crystal packing interactions that destabilize subunit 2 could also alter the ability of subunit 3 to retain biotin when the crystals are formed if Trp 120 is involved in cooperative binding. Sano and Cantor have previously reported evidence for binding cooperativity using a gel filtration chromatography experiment (Sano & Cantor, 1990), but later studies by Kurzban and coworkers disputed these findings and suggest that the binding is not cooperative (Kurzban et al., 1991; Jones & Kurzban, 1995). The asymmetric binding of biotin in the co-crystal at low ratios of biotin to streptavidin suggests the possibility of communication between binding sites. However, there is no evidence of Trp 120 structural alterations between the bound and unbound states in the co-crystals (Fig. 4), and in general, there are no quaternary changes between the ligand-free and biotin-bound structures in these crystal forms. In summary, there is no readily apparent structural basis for the observed non-statistical distribution of biotin ligands in this crystal form, but small structural alterations with significant energetic consequences, of course, remain possible.

Materials and methods

The design of the recombinant core–streptavidin gene, its expression in a T7 expression system (pET-210, Novagen, Inc., Madison, Wisconsin), and the isolation, refolding, purification, and functional characterization of wild-type streptavidin have been reported previously (Chilkoti et al., 1995).

Crystallization

Solutions with protein concentrations of 25–30 mg/mL were used for hanging drop vapor diffusion crystallization experiments. The crystallization conditions are reported in Table 1. For the co-crystallization of core–streptavidin with biotin, 4 μL (structure IV) or 8 μL (structure V) of a saturated biotin solution in water and 18 μL of protein solution were incubated overnight at 4 °C. The solubility of biotin was low at pH 7 and the saturated (~10 mM) solution contained undissolved biotin crystals. Eighteen microliters of a 60% (structure IV) or a 100% MPD (2-methyl-pentane-2,4-diol) solution (structure V) were added before setting up the drops. The reservoir contained a 50% MPD solution. The crystals of core–streptavidin and the complex grew within three to six days to typical dimensions of 0.1 × 0.1 × 1.5 mm (I,II), 0.3 × 0.3 × 0.3 mm (III), 0.05 × 0.05 × 1.0 mm (IV), and 0.02 × 0.2 × 0.3 mm (V), respectively.

Diffraction data collection and processing

The crystals were mounted in capillaries and diffraction data were collected at room temperature. The diffraction data for structures I, IV, and V were collected on a Siemens-Nicolet-Xentronics area detector system (Huber goniostat, Rigaku RU-200 rotating anode X-ray source). The data processing was carried out using the program Xengen (Howard et al., 1987). Data sets for structures II and III were collected on an R-AXIS II image plate detector system (Rigaku RU-200 rotating anode X-ray generator). These data sets were processed with the program PROCESS (Higashi, 1990). Each data set was collected using one crystal, because only little crystal decay was observed. The data collection statistics for all data sets are presented in Table 2.

Structure solution and structural refinement

A streptavidin–peptide complex (Weber et al., 1992a) (Protein Data Bank identification 1PTS) was used for the structure solution of the MONO-1 and MONO-b1 structures (I, II, and IV) with molecular replacement techniques (Rossmann, 1972). Crystallization of this crystal form was previously reported (Pähler et al., 1987), but coordinates for the wild-type protein in this crystal form are not yet available. 1PTS contains a streptavidin dimer in the asymmetric unit so the structural model was expanded to a tetramer and all solvent, ligand, and binding loop atoms were omitted. The actual structure solution using the program X-PLOR (Brünger, 1992a) was carried out for an isomorphous streptavidin mutant, W79F, that will be described in subsequent publications. The rotation function yielded several possible rotational transformations that would align the search model with the tetramers in the new unit cell. These transformations were screened by Patterson correlation methods to eliminate false solutions. The best rotation solutions were applied to the starting coordinates, which were then used in a translation search to position the correctly oriented molecule in the unit cell. The resulting model was subjected to rigid-body refinement, first of the entire tetramer and then of each subunit. The R-value ($R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes) at this stage was 0.345, indicative of a good structure solution.

Structures III and V were also solved using molecular replacement methods, starting from a reduced model of an early MONO-1 refinement. This model included residues 16 to 133. The loop residues 45 to 52 and all solvent molecules were excluded. After calculation of a cross-rotation function and translation function, a rigid body refinement was carried out ($R = 0.351$ for structure III; $R = 0.393$ for V).

For the first steps of the refinement and the completion of the loop regions, the program X-PLOR was used. The refinement was continued with the β-test version of the program SHELXL-96 (Sheldrick, 1996) using the auxiliary programs SHELXWAT for automatically locating bound water molecules and SHELXPRO as an interactive interface program. For graphical evaluation, the program XtalView (McRee, 1992) was used. All RMSDs for least-squares fits were calculated with X-PLOR using residues 19–23, 28–33, 38–42, 54–60, 71–80, 85–97, 103–112, 123–131 in the β-sheet region. The stereochemistry was checked during the refinement process with $2F_o - F_c$ and $F_o - F_c$ maps and the programs PROCHECK (Laskowski et al., 1993) and WHATIF (Vriend & Sander, 1993). Most of the refined water positions were found by SHELXWAT and all were checked visually. Water oxygen atoms were rejected when B-values increased to 63 Å^2 or bigger ($U > 0.8$ Å^2).

SHELXL-96 refines against the squares of the structure factor amplitudes (F^2). This gives the advantage of incorporation of more experimental information and the chance of avoiding local minima. All data from 10 Å resolution to the highest limit were used in the refinement for all data sets except for data set V where an 8 Å resolution limit was applied. Conjugate gradient least-squares methods were used, and all parameters (coordinates and isotropic temperature factors) were refined together. Distance, planarity, and chiral volume restraints were applied, as well as anti-

bumping restraints. The target values for 1,2- and 1,3-distances were set by the program, based on the study of Engh & Huber (1991). Anti-bumping restraints are applied only if non-bonded atoms came closer to each other than a target distance. For the isotropic temperature factors, similarity restraints were applied. Although the asymmetric unit contains a streptavidin tetramer in all of the crystal structures, non-crystallographic symmetry (NCS) restraints were not used in the refinement because R_{free} tests for each structure showed that NCS restraints would not improve the model. Ten percent of the reflection data were held in a separate file and used throughout the X-PLOR and SHELXL-96 refinements as a reference set for calculations of R_{free} (Brünger, 1992b). Diffuse solvent regions were modeled using Babinet's principle (Moews & Kretsinger, 1975). Anisotropic scaling of the observed structure factors as suggested by Parkin et al. (1995) was applied in the refinements but not for structure I, because the R_{free} test showed no improvement of the structural model after this correction. The hydrogen atom positions were geometrically idealized and refined using a riding model. The occupancy of the loop residues 45 to 52 and the biotin molecule in subunit 4 of structure IV were refined with constant isotropic temperature factors and subsequently set to a value of 0.7. Refinement results are presented in Tables 2 and 3. The coordinates for the biotin-bound and ligand-free core–streptavidin structures have been deposited in the Brookhaven Protein Data Bank as 1swa, 1swb, 1swc, 1swd, and 1swe, respectively. Figures 1, 3, 5, and 6 were produced with MOLSCRIPT (Kraulis, 1991). Figure 4 is an XtalView plot (McRee, 1992).

Acknowledgments

This work was supported by grant DK49655 from the National Institutes of Health.

References

Bayer EA, Wilchek M. 1990. Avidin–biotin technology. *Methods Enzymol 184*. Forthcoming.

Brünger AT. 1992a. *X-PLOR, A system for crystallography and NMR, v. 3.1*. New Haven, Connecticut: Yale University Press.

Brünger AT. 1992b. Free R-value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*:472–475.

Chaiet L, Wolf FJ. 1964. The properties of streptavidin, a biotin binding protein produced by streptomyces. *Arch Biochem Biophys 106*:1–5.

Chilkoti A, Stayton PS. 1995. Molecular origins of the slow streptavidin–biotin dissociation kinetics. *J Am Chem Soc 117*:10622–10628.

Chilkoti A, Tan PH, Stayton PS. 1995. Site-directed mutagenesis studies of the high-affinity streptavidin–biotin complex: Contributions of the tryptophan residues 79, 108 and 120. *Proc Natl Acad Sci USA 92*:1754–1758.

Devlin JJ, Panganiban LC, Devlin DE. 1990. Random peptide libraries: A source of specific protein binding molecules. *Science 249*:404–406.

Engh RA, Huber R. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr A47*:392–400.

Giebel LB, Cass RT, Milligan DL, Young DC, Arze R, Johnson CR. 1995. Screening of cyclic peptide phage libraries identifies ligands that bind streptavidin with high affinities. *Biochemistry 34*:15430–15435.

Green NM. 1975. Avidin. *Adv Protein Chem 29*:85–143.

Hemming SA, Bochkarev A, Darst SA, Kornberg RD, Ala P, Yang DSC, Edwards AM. 1995. The mechanism of protein crystal growth from lipid layers. *J Mol Biol 246*:308–316.

Hendrickson WA, Pähler A, Smith JL, Satow Y, Merrit EA, Phizackerley RP. 1989. Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. *Proc Natl Acad Sci USA 86*:2190–2194.

Higashi T. 1990. *PROCESS: A program for indexing and processing R-AXIS II imaging plate data*. Tokyo: Rigaku Corp.

Howard AJ, Gilliland GL, Finzel BC, Poulos TL, Ohlendorf DH, Salemme FR. 1987. The use of an imaging proportional counter in macromolecular crystallography. *J Appl Crystallogr 20*:383–387.

Hutchinson EG, Thornton JM. 1996. PROMOTIF—A program to identify and analyze structural motifs in proteins. *Protein Sci 5*:212–220.

Jones ML, Kurzban GP. 1995. Noncooperativity of biotin binding to tetrameric streptavidin. *Biochemistry 34*:11750–11756.

Katz BA. 1995. 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. *Biochemistry 34*:15421–15429.

Katz BA, Stroud RM, Collins N, Liu B, Arze R. 1995a. Topochemistry for preparing ligands that dimerize receptors. *Chem Biol 2*:591–600.

Katz BA, Cass RT, Liu B, Arze R, Collins N. 1995b. Topochemical catalysis achieved by structure-based ligand design. *J Biol Chem 270*:31210–31218.

Katz BA. 1996. Preparation of a protein-dimerizing ligand by topochemistry and structure-based design. *J Am Chem Soc 118*:2535–2536.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Kurzban GP, Bayer EA, Wilchek M, Horowitz PM. 1991. The quaternary structure of streptavidin in urea. *J Biol Chem 266*:14470–14477.

Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmurski WM, Knapp RJ. 1991. A new type of synthetic peptide library for identifying ligand binding activity. *Nature 354*:82–84.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Livnah O, Bayer EA, Wilchek M, Sussman JL. 1993. Three-dimensional structures of avidin and the avidin–biotin complex. *Proc Natl Acad Sci USA 90*:5076–5080.

McRee DE. 1992. A visual protein crystallographic software system for X11/XView. *J Mol Graph 10*:44–46.

Moews PC, Kretsinger RH. 1975. Refinement of the structure of carp muscle calcium-binding parvalbumin by model building and difference fourier analysis. *J Mol Biol 91*:201–228.

Morag E, Bayer EA, Wilchek M. 1996. Reversibility of biotin-binding by selective modification of tyrosine in avidin. *Biochem J 316*:193–199.

Pähler A, Hendrickson WA, Gawinowicz Kolks MA, Argaraña CA, Cantor CR. 1987. Characterization and crystallization of core streptavidin. *J Biol Chem 262*:13933–13937.

Parkin S, Moezzi B, Hope H. 1995. XABS2: An empirical absorption correction program. *J Appl Crystallogr 28*:53–56.

Pugliese L, Coda A, Malcovati M, Bolognesi M. 1993. Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 Å resolution. *J Mol Biol 231*:698–710.

Rossmann MG. 1972. *The molecular replacement method*. New York: Gordon and Breach.

Sano T, Cantor CR. 1990. Cooperative biotin binding by streptavidin. *J Biol Chem 265*:3369–3373.

Schmidt TGN, Koepke J, Frank R, Skerra A. 1996. Molecular interactions between the strep-tag affinity peptide and its cognate target, streptavidin. *J Mol Biol 255*:753–766.

Sheldrick GM. 1996. *SHELXL-96, program for structure refinement, β-test version*. Göttingen: University of Göttingen.

Vriend G, Sander C. 1993. Quality control of protein models: Directional atomic contact analysis. *J Appl Crystallogr 26*:47–60.

Weber PC, Cox MJ, Salemme FR, Ohlendorf DH. 1987. Crystallographic data for streptomyces avidinii streptavidin. *J Biol Chem 262*:12728–12729.

Weber PC, Ohlendorf DH, Wendoloski JJ, Salemme FR. 1989. Structural origins of high-affinity biotin binding to streptavidin. *Science 243*:85–88.

Weber PC, Pantoliano MW, Thompson LD. 1992a. Crystal structure and ligand-binding studies of a screened peptide complexed with streptavidin. *Biochemistry 31*:9350–9354.

Weber PC, Wendoloski JJ, Pantoliano MW, Salemme FR. 1992b. Crystallographic and thermodynamic comparison of natural and synthetic ligands bound to streptavidin. *J Am Chem Soc 114*:3197–3200.

Weber PC, Pantoliano MW, Simons DM, Salemme FR. 1994. Structure-based design of synthetic azobenzene ligands for streptavidin. *J Am Chem Soc 116*:2717–2724.

Weber PC, Pantoliano MW, Salemme FR. 1995. Crystallographic and thermodynamic comparison of structurally diverse molecules binding to streptavidin. *Acta Crystallogr D51*:590–596.