Protein Science (1998), 7:848–859. Cambridge University Press. Printed in the USA.

Thermodynamic and structural consequences of flexible loop deletion by circular permutation in the streptavidin–biotin system

VANO CHU,[1,3] STEFANIE FREITAG,[1,2,3] ISOLDE LE TRONG,[2] RONALD E. STENKAMP,[2] AND PATRICK S. STAYTON[1]

[1]Department of Bioengineering, University of Washington, Box 357962, Seattle, Washington 98195-7962
[2]Department of Biological Structure and Biomolecular Structure Center, University of Washington, Box 357420, Seattle, Washington 98195-7420

(RECEIVED November 17, 1997; ACCEPTED January 12, 1998)

Abstract

A circularly permuted streptavidin (CP51/46) has been designed to remove the flexible polypeptide loop that undergoes an open to closed conformational change when biotin is bound. The original termini have been joined by a tetrapeptide linker, and four loop residues have been removed, resulting in the creation of new N- and C-termini. Isothermal titration calorimetric studies show that the association constant has been reduced approximately six orders of magnitude below that of wild-type streptavidin to 10^7 M^{-1}. The $\Delta H°$ of biotin association for CP51/46 is reduced by 11.1 kcal/mol. Crystal structures of CP51/46 and its biotin complex show no significant alterations in the binding site upon removal of the loop. A hydrogen bond between Ser45 and Ser52 found in the absence of biotin is broken in the closed conformation as the side-chain hydroxyl of Ser45 moves to hydrogen bond to a ureido nitrogen of biotin. This is true in both the wild-type and CP51/46 forms of the protein, and the hydrogen bonding interaction might thus help nucleate closure of the loop. The reduced entropic cost of binding biotin to CP51/46 is consistent with the removal of this loop and a reduction in entropic costs associated with loop closure and immobilization. The reduced enthalpic contribution to the free energy of binding is not readily explainable in terms of the molecular structure, as the binding contacts are nearly entirely conserved, and only small differences in solvent accessible surfaces are observed relative to wild-type streptavidin.

Keywords: circular permutation; ligand binding thermodynamics; loop deletion; molecular recognition; streptavidin; structure-based drug design

Flexible loops are protein structural elements often found near the binding sites or active sites of receptors and enzymes. With many flexible loops, ligand binding is accompanied by an open-to-closed (or disorder-to-order) conformation change in going from the unbound to the ligand-bound state (Wierenga et al., 1991; Tanaka et al., 1992; Noble et al., 1993; Falzone et al., 1994; Morton & Matthews, 1995). The loops presumably play an important role in gating ligand association and dissociation, but their energetic contributions to molecular recognition remain unclear. The free energy of binding is the result of balancing the entropic costs/benefits of ordering of loops and release of bound water with the enthalpic benefits of burying nonpolar surface area and establishing bonding contacts. It is expected that loop folding may lead to energetic signatures similar to those associated with protein folding (Murphy et al., 1993; Spolar & Record, 1994).

We are interested in defining the high-affinity reaction coordinate of streptavidin, where the conformational change of a flexible binding loop is a prominent feature accompanying biotin association (Hendrickson et al., 1989; Weber et al., 1989). High-affinity protein–ligand interactions play an important role in biology and often are the goal of structure-based drug design projects. The streptavidin–biotin affinity couple, with a K_a of approximately

Reprint requests to: Patrick S. Stayton, Department of Bioengineering, Box 357962, University of Washington, Seattle, Washington 98195; e-mail: stayton@bioeng.washington.edu. 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.

[3]V. Chu and S. Freitag should be considered co-first authors on this paper.

Abbreviations: CP51/46, circularly permuted streptavidin with Glu51 as N- and Ala46 as C-terminus, respectively, and GGGS connecting amino acids Ala13 and Ser139; RMSD, root-mean-square distance; MPD, 2-methylpentane-2,4-diol.

2.5×10^{13} M^{-1}, provides an interesting model system that may illuminate pathways for generating high-affinity, protein–small molecule interactions (Green, 1990).

In common with many other high-affinity protein–ligand systems, streptavidin utilizes three key molecular recognition mechanisms in its interaction with biotin: an extensive hydrogen bonding network, several direct aromatic side-chain contacts, and a flexible loop near the biotin binding site. We recently reported a crystallographic study of the flexible loop in core streptavidin (Freitag et al., 1997). The loop (residues 45–52) is in a closed conformation in the presence of biotin and in an open conformation in ligand-free streptavidin. In our monoclinic crystal forms, residues 49–52 are found in a 3_{10} helix, and the open conformation is stabilized by a hydrogen bonding interaction between residues 45 and 52. In a tetragonal crystal form, these residues are disordered in the open conformation (Weber et al., 1989). Ser45 terminates the β-strand leading into the loop, and the side-chain oxygen of this residue is hydrogen bonded to one of the ureido nitrogen atoms of biotin. Altogether, there are three hydrogen bonds to biotin mediated by residues near or in the loop. The hydroxyl of Tyr43 donates a hydrogen bond to the ureido-oxygen of biotin, and the backbone amide nitrogen of Asn49 is hydrogen bonded to the biotin carboxylate.

We describe here a thermodynamic and crystallographic study of loop deletion in streptavidin using a new approach based on circular permutation of the polypeptide (Fig. 1). Generation of circularly permuted proteins provides an experimental means of investigating the biophysical consequences of loop removal on ligand binding in ways not available using traditional deletion mutants. Circularly permuted proteins have been used previously to investigate the protein folding problem (Yang & Schachman, 1993; Graf & Schachman, 1996), and naturally occurring and synthetic circularly permuted proteins have been identified (Goldenberg & Creighton, 1983; Luger et al., 1989; Heinemann & Hahn, 1995; Lindqvist & Schneider, 1997). We have used circular permutation to remove the loop while leaving free N- and C-termini at the site of excision to provide insight into the energetic contribution of flexible loops to ligand association.

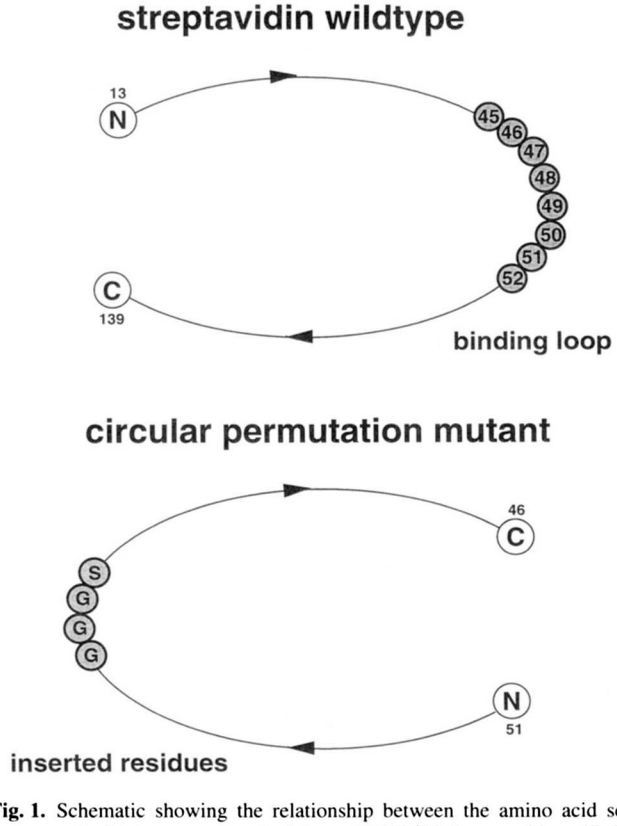

Fig. 1. Schematic showing the relationship between the amino acid sequences of wild-type core streptavidin and CP51/46. Loop residues 47–50 have been removed, and the old N- and C-termini have been joined by a four-residue linker.

Results

Thermodynamics of biotin binding to CP51/46

Isothermal titration calorimetry was used to characterize the thermodynamic consequences of deleting residues 47–50 of the flexible loop. The CP51/46 mutant has an unprocessed N-terminal methionine as demonstrated by high resolution electrospray mass spectrometry and N-terminal sequencing. The potential contributions of this residue must be considered, although it is disordered and not observed in electron density maps for either the biotin bound or unbound state.

Wild-type streptavidin displays a K_a that has been estimated to be 2.5×10^{13} M^{-1} (Green, 1990), and we have previously determined the standard binding enthalpy to be -24.9 kcal/mol at 25 °C (Chilkoti & Stayton, 1995). Using the estimated K_a and associated standard Gibb's free energy, the $T\Delta S°$ term for wild-type streptavidin would then be -6.6 kcal/mol at 25 °C. With the CP51/46 mutant, the K_a for biotin is reduced approximately six orders of magnitude to $2.28\ (\pm 0.44) \times 10^7$ M^{-1} ($\Delta G° = -10.0$ kcal/mol) at 25 °C (see Fig. 2A). The association of biotin is still enthalpically driven with a $\Delta H°$ of -13.8 (± 0.8) kcal/mol and the $T\Delta S°$ term is -3.8 (± 0.8) kcal/mol at 25 °C. The enthalpy values in both phosphate and Tris buffers are within experimental error of each other, suggesting that protonation effects are not significant in the mutant (Table 1).

The change in heat capacity was also measured for CP51/46 to provide further thermodynamic insight into the role of the loop (see Fig. 2B). The ΔC_p was significantly less negative at -95 cal/mol °C (standard deviation = 29 cal/mol °C) compared to the value of -345 cal/mol °C (standard deviation = 12 cal/mol °C) for wild-type streptavidin. This alteration is qualitatively consistent with the expected decrease in surface area buried in the CP51/46 bound state after loop deletion. Previous analysis of the avidin–biotin system by Spolar and Record (1994) suggested that ΔC_p for biotin association is dominated by the folding of the loop residues. The results of calculations relating ΔC_p and changes in the accessible surface area are presented in Table 1B. More detailed analysis of the structural thermodynamic calculations will be presented in a future publication.

Overall structure of CP51/46

The crystal structure of CP51/46 was determined at 2.0 Å resolution and refined to an R-value of 0.145. Results of the data collection and model refinement are summarized in Tables 2 and 3. Comparisons of the overall fold of the CP51/46 tetramer with other core-streptavidin structures indicate no major differences between them (Fig. 3). Least-squares fits of 4×65 β-sheet Cα atoms

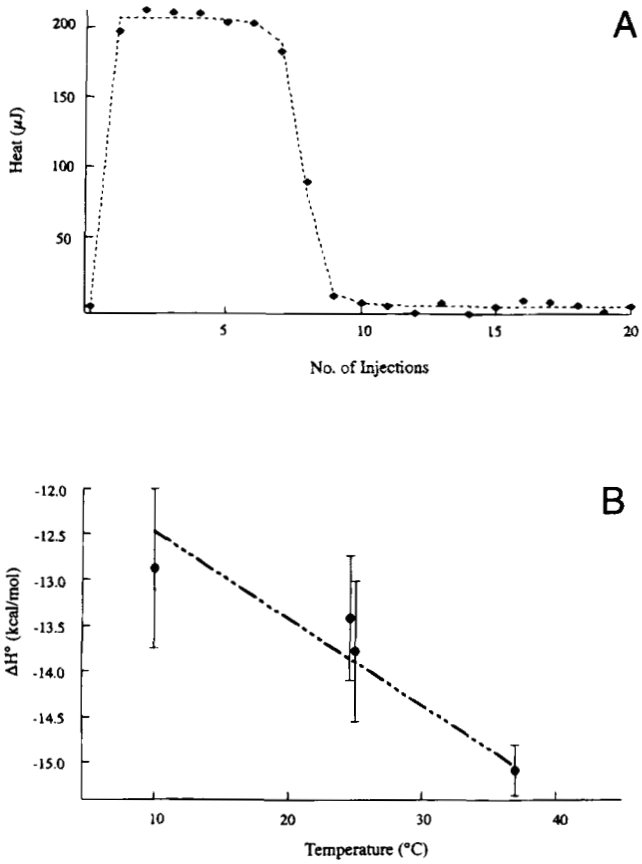

Fig. 2. A: Typical binding isotherm for CP51/46 at 25 °C. Heats for each injection are shown as solid symbols (filled diamond) and have been adjusted for heat of mixing. Parameterized fit is shown as a dotted line. **B:** Plot of ΔH versus T showing the linear fit for ΔC_p of CP51/46. Solid symbols (filled circle) represent average $\Delta H°$ for three experiments. Error bars indicate one standard deviation. The dotted line represents the linear fit of data using weights of 1/s for each point.

of CP51/46 (see Materials and methods) on the monoclinic wild-type structures (PDB entries 1SWA, 1SWB, 1SWC) result in RMS distances (RMSDs) between the Cα atoms of 0.3 Å in all cases. Fits of only one subunit at a time give RMSDs of 0.2 Å for the fitted subunit and values in the range of 0.2–1.0 Å for the other three subunits. Least-squares superpositions of the four individual CP51/46 subunits on each other show no significant differences (RMSDs = 0.2 Å, respectively), indicative of no systematic change in the β-barrel structure.

The crystal structure of the biotin complex of CP51/46 was determined and refined at 1.8 Å resolution. The final R-value is 0.181. Electron density for biotin was clearly defined for all atoms in the small molecule ligand (Fig. 4). The RMSD after superposing 4 × 65 Cα atoms of the CP51/46–biotin complex on the wild-type–biotin complex structure is 0.3 Å. Superposition of the wild-type tetrameric complex onto that of CP51/46, but based on superposition of only one subunit, gives RMSDs of 0.2–0.3 Å for the fitted subunit and 0.3–0.7 Å for the other three subunits.

Characterization of connecting residues

Although the overall structure of ligand-free CP51/46 is the same as that of wild-type streptavidin, there are structural changes at the two regions where the protein was specifically altered. A remarkable feature of the unbound CP51/46 structure is the observation (in two of the four subunits in the crystal) of the ordered, engineered polypeptide connecting the old N- and C-termini. This part of the structure (including 17 residues) is well ordered in subunits 2 and 3. Difference electron density in this region is shown in Figure 5A. Residues 13–15 and 134–139 have been disordered in other monoclinic wild-type structures (Freitag et al., 1997), but in the CP51/46 crystals, they can clearly be identified, along with the four inserted residues Gly140, Gly141, Gly142, and Ser143. The amino acid sequence of the connector between the old N- and C-termini is [. . .Val133-Lys-Pro-Ser-Ala-Ala-Ser-Gly140-Gly-Gly-Ser143-Ala13-Glu-Ala-Gly16. . .].

Table 1A. *Thermodynamic comparison of wild-type and CP51/46 streptavidin*[a]

Protein	K_a (M^{-1})	$\Delta G°$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$T\Delta S°$ (kcal/mol)	ΔC_p (cal/mol °C)
Wild-type	2.5×10^{13}	−18	−24.9 ± 0.4	−6.6 ± 0.4	−345 ± 12
CP51/46 (PB)	$(2.28 \pm 0.44) \times 10^7$	−10.0	−13.8 ± 0.8	−3.8 ± 0.8	−95 ± 29
CP51/46 (Tris)	$(2.47 \pm 1.87) \times 10^7$	−10.1	−13.4 ± 0.7	−3.3 ± 0.7	—

[a]Values reported are per mol of subunits.

Table 1B. *Comparison of observed and calculated values of* ΔC_p[a]

	Wild-type		CP51/46	
	Calculated	Observed	Calculated	Observed
Δ accessible surface (apolar) (Å^2)	−566	—	−446	—
Δ accessible surface (polar) (Å^2)	−146	—	−234	—
ΔC_p (cal/mol °C)	−216	−345	−139	−95

[a]Surface areas reported are for an average subunit in the tetramer.

Table 2. *X-ray data for streptavidin CP51/46 and its biotin complex*

Protein	CP51/46	CP51/46 + biotin
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters		
a (Å)	60.3	71.9
b (Å)	78.6	78.6
c (Å)	93.5	90.8
Number of tetramers (subunits)/ unit cell	4 (16)	4 (16)
Packing parameter V_M (Å^3/Da)	2.1	2.4
Resolution (Å)	2.0	1.8
Wavelength (Å)	1.54178	0.980
Measured reflections	93,577	417,564
Completeness		
Overall (%)	94.4	98.2
Outermost shell (%)	82.1	96.2
R_{merge}		
Overall	0.05	0.03
Outermost shell	0.32	0.17

Table 3. *Refinement statistics for streptavidin CP51/46 and its biotin complex*

Protein	CP51/46	CP51/46 + biotin
Resolution range (Å)	10–2.0	10–1.8
Unique reflections	28,548	42,736
Nonhydrogen atoms	3,576	3,525
Water molecules	214	335
R-factor[a]	0.145	0.181
Free *R*-factor[b]	0.229	0.231
Average *B*-factor[c] (Å^2)	29	27
Ramachandran quality[d]	0.92	0.90
RMSD between ideal and observed bond lengths (Å)	0.006	0.007
RMSD between ideal and observed bond angles (1–3 distances) (Å)	0.025	0.025

[a]For all data with $I > 2\sigma(I)$.
[b]For 10% of the data with $I > 2\sigma(I)$.
[c]For all atoms.
[d]Fraction of residues (except Gly and Pro) in "most favored regions" (Laskowski et al., 1993).

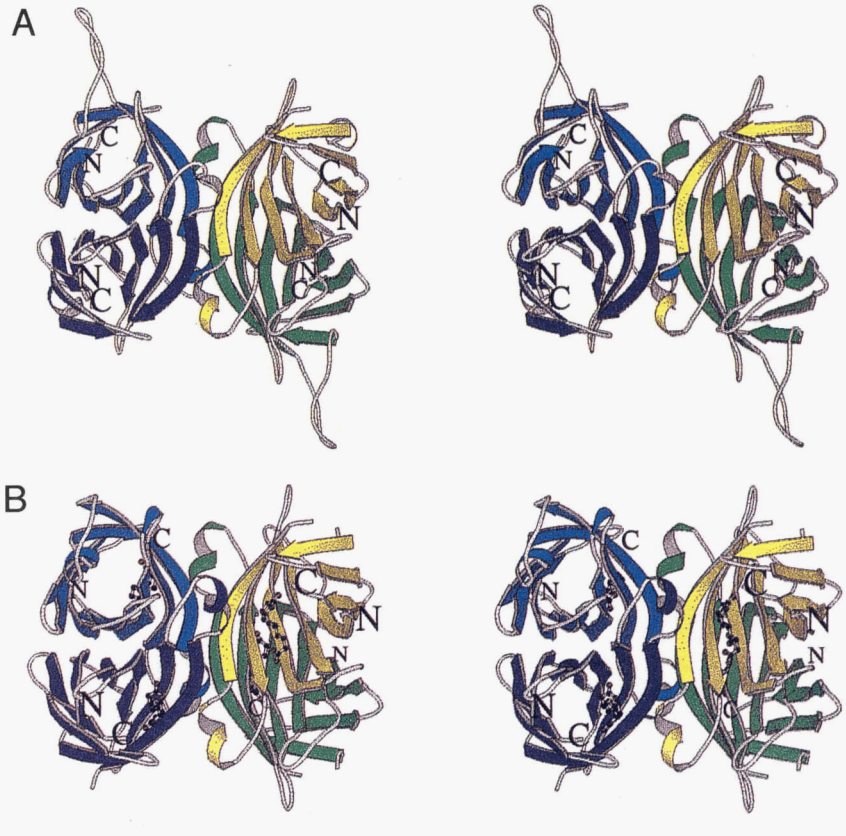

Fig. 3. A: MOLSCRIPT (Kraulis, 1991) stereoview of the tetrameric circularly permuted streptavidin CP51/46. In subunits 2 (green) and 3 (cyan), the engineered loops extend away from the globular structure and are stabilized by crystal packing interactions. **B:** MOLSCRIPT stereoview of the mutant–biotin complex. In subunit 3 (cyan), the new connecting loop adopts a different conformation more integrated in the β barrel structure and also stabilized by crystal packing.

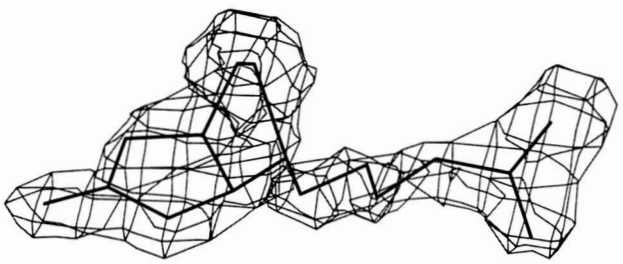

Fig. 4. Unbiased $|F_o| - |F_c|$ electron difference map contoured at 2.4σ in the region of biotin in one of the subunits in the mutant–biotin complex. Superposed is the refined biotin.

The two observed connectors (residues 133–16) form the regions of major packing interactions in this crystal form (Fig. 6). Two connectors from each tetramer form interactions with two neighboring tetramers, which also interact through one of their connecting regions with the first tetramer. The rigidity of the linking residues in subunit 2 and 3 compared to other loop regions in the protein can be explained by these packing interactions and is shown by their low temperature factors. The average *B*-values for the atoms of residues 133–18 in subunit 2 is 25 Å^2, in subunit 3, 24 Å^2 (average *B*-value for all atoms in the tetramer: 29 Å^2).

A Ramachandran plot for the CP51/46 structure including the engineered linkers shows no outliers, and the ϕ/ψ distribution in these regions is normal. The connecting regions in subunits 2 and 3 are very similar. Superposition of the structures between the Cα atoms of residues 133–18 yields an RMSD of 0.16 Å. The program PROMOTIF (Hutchinson & Thornton, 1996) finds no secondary structural elements other than a 3_{10}-helical portion formed by residues Ala137, Ala138, and Ser139. In the other subunits (1 and 4) only residues 143, 13, 14, and 15 are observed in the electron density beyond those seen in the wild-type structures. Residues 13–17 were identified by PROMOTIF as forming an α-helix in both subunits. Superpositions of subunits 2 and 3 (where the complete connectors are observed) on subunits 1 and 4 show that the rigid conformation for the linker would not fit in the crystal packing for the latter subunits. The connecting residues 134–142 are disordered in subunits 1 and 4. Because of the different orientation of the old C-terminal residues (Lys132 and Val133) in CP51/46, the adjacent Tyr22 side chain adopts a conformation with $\chi 1$ rotated by about 180° in all four subunits. Other side chain changes were not detected in the unbound form.

The biotin complex of CP51/46 crystallizes in a different crystal form from that of the uncomplexed mutant, and the connecting polypeptide is ordered for only subunit 3 (Fig. 5B). The conformation of the linker differs from that seen in the unbound structure, and includes a short segment of α-helix (residues 14–17). This can be seen on the upper left of Figure 3B and can be compared with the conformation of the connector in the unbound protein shown in Figure 3A.

The ordered linker in the crystal structure of the biotin complex contacts a biotin molecule bound in subunit 2 of a neighboring tetramer. A hydrogen bond (3.0 Å) is formed between the amide nitrogen atom of Ser139 (of a symmetry related molecule) and one of the oxygens (O1) of the carboxyl group on the aliphatic chain of biotin. There are also two interactions between Ser139 Oγ of the symmetry related molecule with both of the carboxyl oxygens

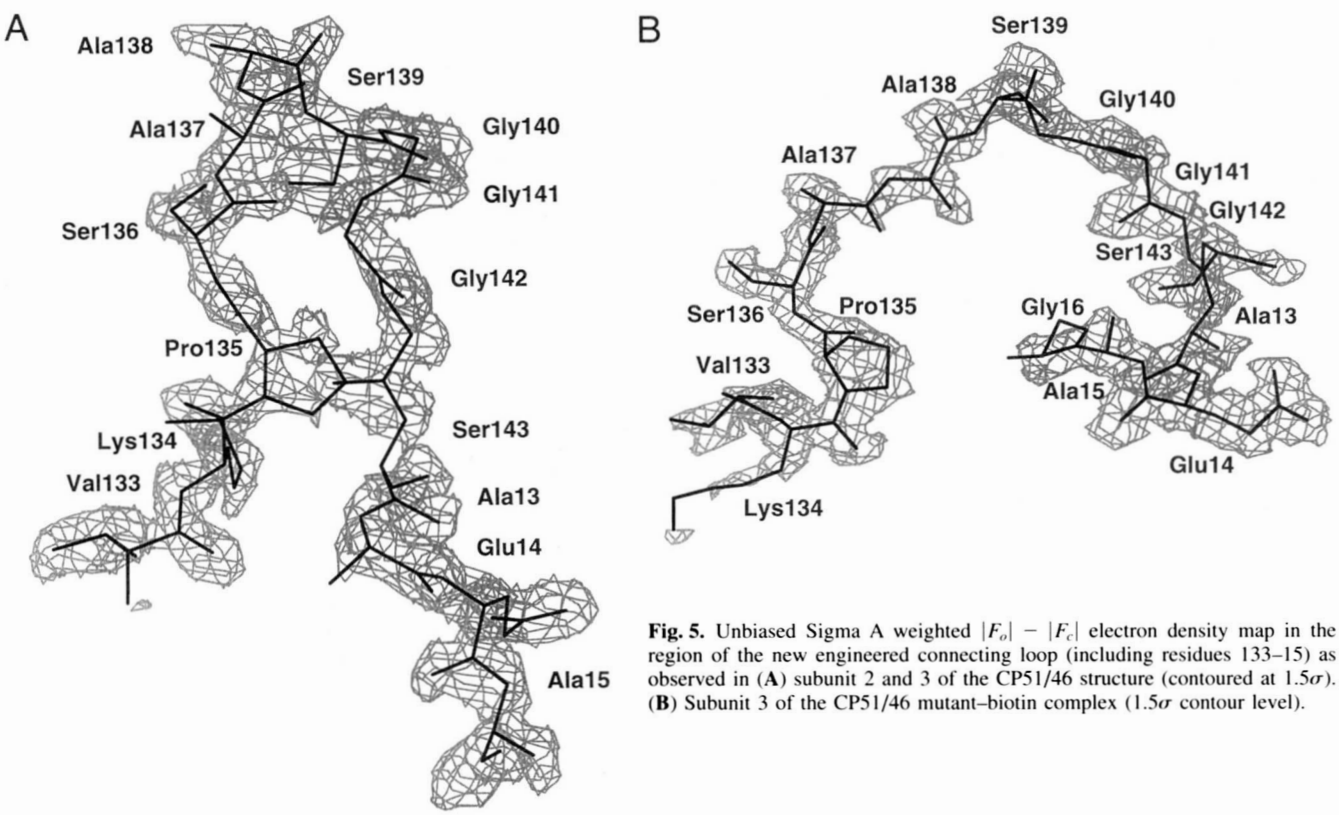

Fig. 5. Unbiased Sigma A weighted $|F_o| - |F_c|$ electron density map in the region of the new engineered connecting loop (including residues 133–15) as observed in (**A**) subunit 2 and 3 of the CP51/46 structure (contoured at 1.5σ). (**B**) Subunit 3 of the CP51/46 mutant–biotin complex (1.5σ contour level).

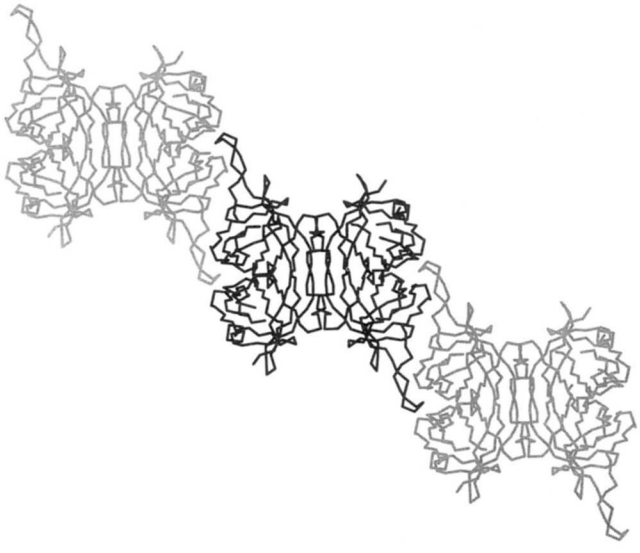

Fig. 6. Packing interactions between neighboring CP51/46 tetramers. The engineered loops of adjacent tetramers are highly involved in interactions with the next tetramer in the crystal.

(3.4 Å to O1 and 2.5 Å to O2). These are not sufficient to cause a conformational change in the biotin bound in this subunit, but the biotin temperature factors are lower in this subunit than in the others, consistent with the additional interactions between the ligand and the connecting loop.

Characterization of the binding site in ligand-free CP51/46 and the CP51/46–biotin complex

Introduction of new N- and C-termini near the biotin binding site could significantly change the protein structure and thus its binding affinity for biotin. Three residues at the new N- and C-termini are disordered in the unbound structure and unobserved in electron density maps. The N-terminal methionine is not seen in the electron density. Residue Glu51 is also mobile and invisible in electron density maps as is the C-terminal residue Ala46. The observed terminal residues are Ser52 and Ser45, which show distinctly higher *B*-values for main- and side-chain atoms than found in the rest of the structure. A comparison of CP51/46 in this region with unbound and biotin bound wild-type structures reveals more similarity to the unbound structure, where the binding loop adopts a more mobile open conformation than in the complex. The last β-sheet hydrogen bond in the wild-type unbound state is between atoms Ser45 N and Ser52 O. This interaction is also observed in CP51/46 where the distances in the four subunits range from 3.0–3.4 Å.

The termini in the four subunits in the biotin complex of CP51/46 align well with those residues in the wild-type structure. In subunits 1, 2, and 3, residue 46 was observed in the electron density maps. In subunit 4, an additional residue 51 was refined at the N-terminus, but the N-terminal methionine was never observed. There is a slight separation of the termini from each other that results in breakage of the hydrogen bond between Ser45 N and Ser52 O (see Fig. 7A). This was also observed in the structures of wild-type streptavidin (Freitag et al., 1997). Breakage of that hydrogen bond is accompanied by the formation of a hydrogen bond between Ser45 O and the ureido nitrogen atom of the bound biotin (Fig. 7B).

In all four subunits in the CP51/46–biotin complex, electron density for biotin was detected in the same orientation as in the wild-type biotin complex (Fig. 4). The only difference in the hydrogen bonding patterns for biotin between wild-type and CP51/46 involve the carboxyl O1 atom. In the wild-type complex, this oxygen atom is hydrogen bonded to the amide of Asn49, but the deletion of the binding loop removes this interaction. Crystal packing interactions replace the loop interactions in subunit 2 (as described above). All other first shell hydrogen bonds are very similar to those found in the wild-type complex (Freitag et al., 1997). The second shell of hydrogen bonds is disturbed by deletion of residue Val47, which interacts with Ser45 in the wild-type protein.

Figure 8A and B shows the biotin binding site of the mutant, including the tryptophan and hydrogen bonding residues. These components of the binding site are also very similar to those found in the wild-type complex. As in wild-type streptavidin, only minor alterations in the binding site are observed upon binding of biotin.

It might be anticipated that deletion of the loop that closes over the biotin site would result in significantly larger solvent accessible surfaces for the biotin ligands in CP51/46. In fact, the changes are not all that large. The average solvent accessible surface for biotin in the wild-type complex is 18.3 $Å^2$, indicative of the nearly complete burial of biotin when bound to streptavidin. (The solvent accessible surface for "free" biotin is 407.2 $Å^2$.) The only exposed biotin atoms are the carboxyl oxygen atoms. The accessible surface increases to 56.5 $Å^2$ for the CP51/46 complex. The oxygen atoms become more exposed as do portions of the aliphatic chain.

Discussion

Flexible loops near active sites or binding sites have been viewed as gates that are open to optimize substrate/ligand association and that close to prevent solvent access and to form bonding contacts that minimize dissociation. Most experimental work with flexible loops has been conducted in the context of enzymatic catalysis, where site-directed mutagenesis and crystallographic studies have probed their contributions to various states in the turnover cycle.

The detailed thermodynamic contributions of flexible loops to small molecule recognition have not been well studied, especially in terms of the free energy breakdown into enthalpic and entropic components. A theoretical consideration of flexible loop closure in the avidin–biotin system predicted that the energetic signature should resemble that of protein folding, because the process involves the partitioning of several loop residues from a solvent exposed state to a bonded, nonexposed state (Spolar & Record, 1994). We have used here a new circular permutation approach to delete the streptavidin flexible binding loop and subsequently conducted a study of the structural and thermodynamic consequences.

Comparison of the wild-type and circularly permuted streptavidin structures demonstrates that there are no significant quaternary changes introduced by the loop deletion. Previous crystallographic studies have shown that circular permutation does not perturb the overall folding of many proteins. For example, the structure of concanavalin A (Hardman & Ainsworth, 1972) is very similar to that of pea lectin (Einspahr et al., 1986) and favin (Reeke & Becker, 1986), although the amino acid sequences of these proteins are circularly permuted. Also, recent structural studies of five circularly permuted proteins [two variants of jellyroll proteins (Hahn et al., 1994), two mutants of α-spectrin SH3 domains (Viguera et al., 1996), and a circularly permuted β-lactamase (Pieper et al., 1997)] have shown that the overall folding of these molecules is

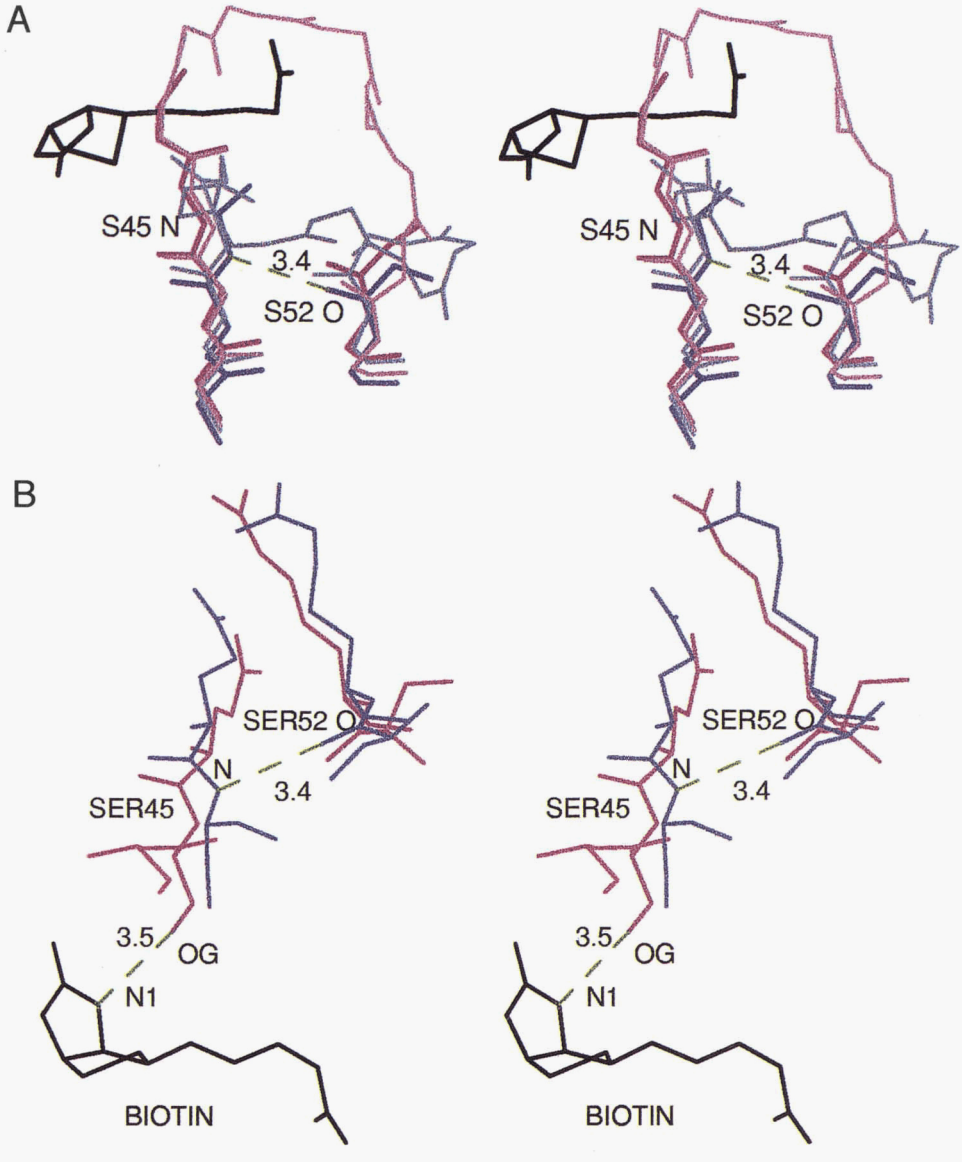

Fig. 7. A: Superposition of the new N- and C-terminal residues in one subunit of CP51/46 (blue) and the biotin complex (red). The binding loop as found in the wild-type structures is also shown in the closed (pink) and open (light blue) conformations. As in ligand-free streptavidin, the last β-sheet hydrogen bond is formed between Ser45 and Ser52. In the complex, this hydrogen bond is not formed. **B:** Instead, Ser45 Oγ hydrogen bonds to one ureido nitrogen atom of biotin, which is coupled with a shift of the C-terminal residue toward the binding site.

conserved when compared to the nonpermuted proteins. This is also true for streptavidin. In this case, the circular permutation approach leaves free ends at the base of the flexible loop, rather than the peptide connection left with standard deletion techniques. Similarly, the introduction of the residues joining the native N- and C-termini does not affect the structure of the tetramer. This engineered connector in both the unbound and complexed structures is observed only in subunits where the loop is involved in crystal packing interactions, and thus it is assumed that the connector is disordered in solution.

A detailed analysis of the CP51/46 biotin-binding residues demonstrates that the conformations of all of the direct hydrogen bonding and aromatic contacts are indistinguishable from wild-type in both the unbound and complexed states. A remarkable feature of the CP51/46 structural analysis relative to wild-type is the close reproduction of coupled hydrogen bonding alterations in the loop hinge points upon biotin association. Our previous crystallographic analysis of the flexible loop showed that the open conformation is stabilized by a β-sheet hydrogen bonding interaction between Ser45 of the β-strand leading into the loop and Ser52 of the β-strand leading out of the loop. This hydrogen bond is broken in the closed conformation as the side-chain hydroxyl of Ser45 moves into position to hydrogen bond to a ureido nitrogen of biotin. There is thus a direct structural link between the closure of the loop and biotin association. This structural alteration between Ser45 and Ser52 is preserved in the loopless CP51/46 mutant, and it appears that the formation of the Ser45 side-chain hydrogen bond to biotin may nucleate the conformational closure of the loop (Fig. 7).

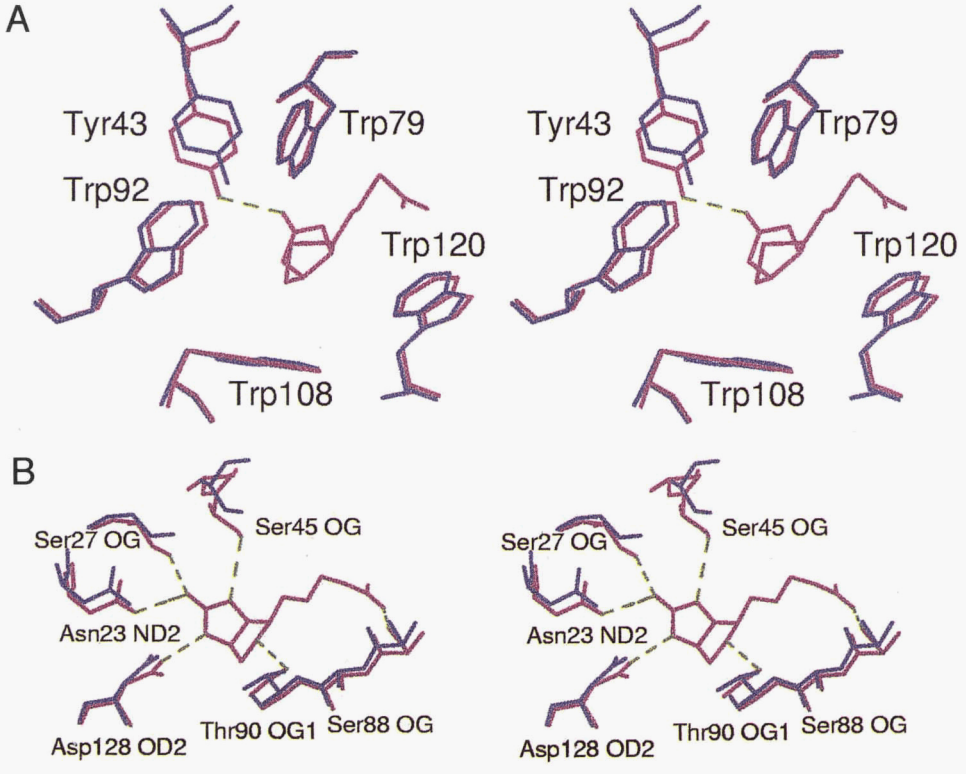

Fig. 8. Superposition of the streptavidin CP51/46 unbound structure (blue) and the CP51/46-biotin complex (red) in the region of the binding site (for subunit 1 in the tetramer). **A:** The tryptophan side chains that mediate hydrophobic interactions with biotin show only minor deviations from the unbound structure. **B:** Hydrogen bonding interactions with biotin.

Despite the close conservation of direct binding contacts to biotin, the CP51/46 mutant displays a dramatic decrease in affinity. This reduction in binding free energy is driven by a large decrease in the binding enthalpy, which is lowered from −24.9 to −13.8 kcal/mol at 25 °C. Qualitatively, it could be hypothesized that loop removal would result in a decrease in binding enthalpy and that the binding entropy would become less negative due to the reduction in configurational entropy costs associated with loop closure. If a K_a of 10^{13} M^{-1} for wild-type streptavidin is assumed, then the $T\Delta S$ term is less negative by about 3 kcal/mol at 25 °C. However, this calculated value is strongly dependent on the K_a assumption, as a K_a of 10^{15} M^{-1} would leave an unaltered $T\Delta S$ term. In addition, the entropy contributions related to the hydrophobic effect, e.g., potential release of bound water, must also be considered.

A comparison of the water molecules bound to the surface of the protein molecules in these crystal structures shows that the residues in the closed binding loop in wild-type streptavidin have not been replaced by bound water molecules in CP51/46. In one subunit of the CP51/46 complex with biotin (subunit 2), there is a bound water molecule close to the Asn49 nitrogen position in the wild-type streptavidin–biotin complex. This is a tightly bound water (2.7 Å from the biotin carboxylate oxygen atom) with a relatively low B-value of 25 $Å^2$. There is no water in this position in the other three subunits. In the unbound CP51/46 structure, there is no consistent pattern of water molecules bound in the biotin pocket. Each subunit has a slightly different set of water molecules in this region. This is also the pattern found in the wild-type structure. A more complete analysis of the bound water molecules will be discussed elsewhere.

To provide more insight into the thermodynamic alterations associated with the CP51/46 mutant, we conducted a structural thermodynamic analysis that utilized ΔC_p measurements and the high resolution unbound and complexed crystallographic structures. If we simply take the bound and unbound structures and perform the ASAcalc analysis of the accessible surface area (Freire et al., 1997), we calculate a ΔC_p of −216 cal/mol °C for the wild-type and −139 cal/mol °C for the CP51/46 mutant. Although these values are qualitatively consistent with the reduction in ΔC_p from −345 cal/mol °C for wild type to −95 cal/mol°C for CP51/46, there is poor quantitative agreement between the experimental and predicted numbers. However, we did not expect this algorithm to necessarily match the individual ΔC_p values because these treatments are based on an averaged parameterization over a protein folding data base. Significant alterations could result from the specific properties of biotin. However, the crystallographic analysis demonstrating the similarity in biotin-binding interactions and surface accessibility in wild-type and CP51/46 suggests that the biotin contribution should be largely the same for both proteins. Close examination of the wild-type and mutant structures reveals that there are many side chains not involved in binding of biotin that display significant alterations in accessible surface area between the unbound and bound states. This finding, which is due to relatively subtle alterations in the structures that vary from subunit to subunit and crystal form to crystal form, complicates the analysis and thus raises the concern that the isolated crystallo-

graphic structures may not be representative of the statistical conformational ensemble in solution. This might be sufficient to account for the differences between the experimental and the predicted $\Delta\Delta C_p$.

In summary, circular permutation represents a new approach to flexible loop deletion that leaves free ends instead of a constrained connection. High-resolution crystallographic analysis of the free and bound states demonstrates that biotin binding is accompanied by the same breaking of an ligand-free streptavidin Ser45–Ser52 β-sheet hydrogen bonding interaction and formation of Ser45 side-chain hydrogen bonding interaction to biotin that is found in the wild-type protein. This hydrogen bonding switch may thus represent the nucleation site for ligand-associated loop closure. Despite the preservation of all key binding interactions with biotin, the circularly permuted streptavidin mutant displays a large decrease in binding enthalpy and thus affinity. These results suggest that the flexible loop plays a crucial role in generating the exceptional biotin binding free energy of streptavidin.

Materials and methods

Unless otherwise noted, all oligonucleotides were obtained from Integrated DNA Technologies (Coralville, Iowa), plasmids and cells were from Novagen (Madison, Wisconsin), PCR reagents were obtained from Promega (Madison, Wisconsin), restriction enzymes and ligases were supplied by New England Biolabs (Beverly, Massachusetts), and chemical reagents were obtained from Sigma (St. Louis, Missouri).

Construction of the circularly permuted gene

The circularly permuted variant was constructed from a synthetic gene for core streptavidin (Chilkoti et al., 1995). A Gly-Gly-Gly-Ser linker was used to connect the original termini. A tandem gene (Horlick et al., 1992) of streptavidin was constructed first to serve as a "template" for the generation of desired circular permutations. The tandem streptavidin gene was constructed in two halves from the wild-type gene using PCR mutagenesis. Four primers were used pairwise to generate each half. The first half codes for the wild-type gene with a new linker sequence appended to the 3′ end. This linker sequence codes for the Gly-Gly-Gly-Ser segment, which bridges the original termini of streptavidin. The second half attaches the linker sequence to the 5′ end of the streptavidin gene. Both fragments were generated in separate PCR reactions and subcloned into pT7Blue plasmids. After cutting with the *NheI* restriction enzyme at a unique site in the linker region of both fragments, the two halves were ligated together to create the tandem streptavidin gene. The sequences of each set of fragments as well as that of the final assembly were checked for errors using dye-terminated DNA sequencing.

To create the CP51/46 mutant through PCR mutagenesis, two additional primers were designed and synthesized. The sense primer anneals at residue 51 in the first half of the tandem gene and adds an *NdeI* site to the beginning of the gene. The antisense primer anneals at residue 46 in the second half of the tandem gene and adds stop codons followed by a *HindIII* site. Thirty-five cycles of PCR mutagenesis (90 °C × 2 min; 50 °C × 2 min; 72 °C × 2 min) produced the circularly permuted gene that was ligated into Novagen (Madison, Wisconsin) pT7Blue plasmids and transformed into NovaBlue maintenance hosts. The gene was later subcloned into Novagen pET-21a plasmids in BL21(DE3) hosts for expression. The name CP51/46 reflects the relocation of the N-terminus to residue 51 of the original wild-type sequence and relocation of the C-terminus to residue 46. Again, DNA sequencing was used to confirm the integrity of the mutant gene.

Synthesis of CP51/46 in E. coli

BL21(DE3) cells containing the CP51/46 gene in pET-21a were cultured overnight at 37 °C in Luria-Bertani (LB) media. The cell pellet was washed and re-suspended in fresh LB before being used to inoculate 5 L of 2 × YT media supplemented with 100 μg/mL ampicillin. The culture was incubated at 37 °C with shaking until the A_{600} reached 1.0, when protein expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactoside (IPTG). Cells were cultured for an additional 3 h before harvesting by centrifugation.

Isolation and purification of CP51/46

Cell pellets were re-suspended in 50 mM Tris·HCl, 200 mM NaCl, 5 mM EDTA, 8% sucrose, 1% Triton X-100, and 1 mM phenylmethylsulfonyl fluoride (PMSF) at pH 8.0. Cells were lysed by sonication and centrifuged at 17,700 × *g* for 20 min. The insoluble fraction was sonicated and centrifuged twice more and pellets were then sonicated and centrifuged three times in the same buffer without Triton X-100. The remaining insoluble inclusion bodies containing the CP51/46 protein were dissolved in 6 M guanidine, 50 mM Tris·HCl at pH 7.5 to a concentration of no more than 10 mg/mL and allowed to equilibrate for several hours at 4 °C. Solubilized protein was then diluted dropwise with stirring at 4 °C in a 50× volume of 50 mM Tris·HCl, 100 mM NaCl, 5 mM EDTA, 0.1 mM PMSF at pH 7.5 and allowed to equilibrate overnight. The resulting solution was centrifuged to remove insoluble material and concentrated in a stirred Amicon (Beverly, Massachusetts) ultrafiltration cell.

CP51/46 was purified by affinity chromatography over Pierce (Rockford, Illinois) iminobiotin-agarose (Hofmann et al., 1980). Protein-containing fractions were pooled and exchanged into a storage buffer of 50 mM phosphate, 100 mM NaCl at pH 7.75.

Characterization of CP51/46

N-terminal sequencing was performed on an Applied Biosystems Model 477A Sequencer. SDS/PAGE analysis was done using precast Mini-Protean 10–20% gradient gels (Bio-Rad, Hercules, California). The concentration of CP51/46 was determined by absorption at 280 nm using an extinction coefficient (ϵ_{280}) of 34,000 M^{-1} cm^{-1} for the subunit (Sano & Cantor, 1990). Electrospray mass spectrometry was performed on a VG Quattro II Tandem Quadrupole Mass Spectrometer.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments on wild-type streptavidin were performed on a MicroCal Omega instrument. ITC experiments on CP51/46 were done using a Calorimetry Science Corporation 4200 Calorimeter (Provo, Utah). CP51/46 solutions of 30–40 μM concentration were titrated by the addition of 20 × 5 μL aliquots of 750 μM biotin dissolved in the same buffer as the protein. All ITC experiments were done in either phosphate

(50 mM sodium phosphate, 100 mM NaCl, pH 7.75) or Tris (50 mM Tris·HCl, 100 mM NaCl, pH 7.75) buffers. Biotin concentrations were determined gravimetrically.

Data were analyzed using the proprietary software supplied by Calorimetry Science Corporation with the instrument. Heats of dilution for each injection were subtracted from the reaction heats before data analysis. Nonlinear fitting of the data allowed the number of binding sites (n), association constant (K_a), and binding enthalpy (ΔH°) to be determined assuming noncooperative binding and one site per subunit.

Crystallization and diffraction data collection

A CP51/46 protein solution with a concentration of 30 mg/mL in water was used for crystallization experiments (hanging drop vapor diffusion method). The mutant crystallized in the form of rods from solutions containing 52% MPD (2-methyl-pentane-2,4-diol). A crystal with dimensions of 0.05 × 0.05 × 0.5 mm was mounted in a glass capillary, and diffraction data were collected on an R-AXIS II image plate detector system attached to a Rigaku RU-200 rotating anode (CuKα = 1.54178 Å) at 293 K. The crystal diffracted to 2.0 Å resolution, and data were collected to a completeness of 94% with a mean I/σ of 5.2. The orthorhombic unit cell dimensions are a = 60.3, b = 78.6, c = 93.5 Å, and the space group is $P2_12_12_1$ (Z = 4). One tetramer of the protein is found in the asymmetric unit. Data processing was carried out using DENZO (Otwinowski & Minor, 1994). The overall $R(I)_{merge}$ was 0.069. Table 2 gives an overview of the collected data.

Crystals of the biotin complex of CP51/46 were obtained by cocrystallization of biotin and the mutant streptavidin from hanging drop experiments. The protein solution was 12 mg/mL in CP51/46, 10 mM biotin. The reservoir solution was 52% MPD. The crystals were long plates with dimensions 0.1 × 0.3 × 0.7 mm. Diffraction data were collected at beamline 9-1 at the Stanford Synchrotron Radiation Laboratory at 100 K (λ = 0.98 Å). Earlier attempts to collect room temperature data on an R-AXIS II resulted in lower resolution data sets (2.3 and 2.6 Å). The shock frozen crystal diffracted to 1.8 Å resolution with a mean I/σ of 15.8. The overall completeness of the data set is 98%. The unit cell parameters are a = 71.9 Å, b = 78.6 Å, c = 90.8 Å. The orthorhombic space group is $P2_12_12_1$. One tetramer of the protein–biotin complex is in the asymmetric unit in this crystal form. The data processing and scaling were carried out using DENZO and SCALEPACK (Otwinowski & Minor, 1994). The overall $R(I)_{merge}$ was 0.037 (Table 2).

Structure solution and refinement

A tetrameric wild-type core-streptavidin model (residues 16–44 and 52–133, PDB entry 1SWA; Freitag et al., 1997) was employed as a search model in the molecular replacement structure solution for CP51/46 using X-PLOR (Brünger, 1992a). After application of the results of a cross-rotation function and a translation function, the R-value was 0.457. The twofold symmetry axes of the tetramer were not aligned with the crystallographic axes, confirming the existence of a tetramer in the asymmetric unit.

The resulting model was subjected to full matrix, least-squares rigid body refinement with the β-test version of SHELXL-97 (Sheldrick, 1997), giving R = 0.384 for data with $I > 2\sigma(I)$. Throughout the refinement, all data were included from 10 Å resolution to the highest limit (2.0 Å). Ten percent of the reflection data were held in a separate file and used for calculation of R_{free} (Brünger, 1992b). At this stage the value of R_{free} ($I > 2\sigma(I)$) was 0.422. Subsequent positional and B-factor refinement using conjugate gradient methods (Konnert & Hendrickson, 1980) as implemented in SHELXL lowered the R-value ($I > 2\sigma(I)$) to 0.265 and R_{free} ($I > 2\sigma(I)$) to 0.335. Electron density for the new engineered loop (residues 133–18) was observed in subunits 2 and 3 in $|F_o| - |F_c|$ maps after the first refinement steps. The modeling of these two loops and addition of 31 water positions in the refinement decreased the R-value to 0.213 and R_{free} to 0.300 (both for data with $I > 2\sigma(I)$). The final model contains residues 52–133 and 143–45 in subunits 1 and 4, and residues 52–45 in subunits 2 and 3, as well as 214 water molecules. The final R-value ($I > 2\sigma(I)$) is 0.145 (0.195 for all data) and R_{free} ($I > 2\sigma(I)$) is 0.229 (0.288 for all data).

Molecular replacement methods were again used to solve the structure of the biotin complex, starting with the same wild-type streptavidin model. The AMoRe program package (Navaza, 1994) was employed for the solution, and the correlation coefficient for the best solution from the rotation function was 0.298. The best solution after calculating the translation function for the eight best rotation solutions had a correlation coefficient of 0.523 and an R-value of 0.411. After rigid body refinement for the complete tetramer and for all four subunits separately, coordinate refinement with SHELXL-97 (Sheldrick, 1997) resulted in an R-value of 0.412 for data with $I > 2\sigma(I)$ and an R_{free} value of 0.442. All data in the range from 10–1.8 Å resolution were used throughout the refinement, as described above for the unbound structure. The R_{free} data set contained 10% of the data in the range. Biotin was clearly identified in the binding sites of all four subunits in the Sigma A weighted $|F_o| - |F_c|$ electron density maps (Read, 1986). Also, in subunit 3, the residues for the engineered loop were modeled in a conformation differing from that in the unbound structure. The final model for the biotin complex includes residues 52–132 and residues 16–46 in subunits 1 and 2; residues 52–46 in subunit 3 and residues 51–133 and 16–45 in subunit 4; as well as four biotin ligands and 335 water molecules. The final R-value is 0.181 for data with $I > 2\sigma(I)$ and 0.192 for all data. The final R_{free} values are 0.231 ($I > 2\sigma(I)$) and 0.245 (all data).

Both molecular models were refined against the squares of the structure factor amplitudes. All parameters, coordinates, and isotropic displacement parameters were refined together. Target values for 1,2- and 1,3-distance restraints were based on the study of Engh and Huber (1991). Planarity and chiral volume restraints were applied as were similarity restraints for the isotropic displacement parameters and anti-bumping restraints if nonbinding atoms came closer than a target distance. Diffuse solvent regions were modeled using Babinet's principle (Moews & Kretsinger, 1975). Anisotropic scaling of the observed structure factors (Parkin et al., 1995) as implemented in SHELXL was applied in the refinements. Hydrogen atoms were geometrically idealized and refined with a riding model in the last cycles.

XtalView (McRee, 1992) was used for graphical evaluation of the model during the refinement. Sigma A weighted $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ electron density maps (Read, 1986) were calculated with the interactive interface program SHELXPRO (Sheldrick, 1997). In addition, the programs PROCHECK (Laskowski et al., 1993) and WHATIF (Vriend & Sander, 1993) were employed to check the stereochemistry during the refinement process. Most of the refined water positions were found by SHELXWAT, an auxiliary program of SHELXL for automated water position searches. 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, and 123–131 in the β-sheet region. Figure 3 was produced with MOLSCRIPT (Kraulis, 1991). Figures 4, 5, 6, 7, and 8 are XtalView plots (McRee, 1992).

Accession numbers

The coordinates and structure factors for these structures have been deposited in the Brookhaven Protein Data Bank. The unbound CP51/46 structure has identifier 1SWF (release date April 23, 1998), while the biotin complex is 1SWG (release date July 12, 1998).

Surface area calculations

All calculations of accessible surface areas were performed using the program ASAcalc (Freire et al., 1997), which is based on the Lee and Richards algorithm (Lee & Richards, 1971). Calculations were done using a solvent radius of 1.4 Å and a slice width of 1.25 Å. Results are averaged over 12 rotations.

To allow meaningful comparison, files containing the structural coordinates for the ligand-free and bound forms of wild-type streptavidin (PDB entries 1SWC and 1SWE) and those of CP51/46 (1SWF and 1SWG) were modified as follows: all files were truncated to exclude residues 13–15, residues 133–139, and all water molecules; copies of the binding loop in the open conformation from one subunit of the ligand-free wild-type structure were added to subunits in which the loop residues were not visible; the surface area of free biotin was calculated by isolating the biotin coordinates from 1SWE. Areas associated with carbon and hydrogen atoms were treated as apolar surface area; those of oxygen and nitrogen were treated as polar. The changes in accessible surface area upon binding were calculated as the difference between the surface area of the bound form and the combined surface areas of the ligand-free form and the four biotin molecules.

Acknowledgments

This work was supported by grant DK49655 from the National Institute of Health and by a fellowship to V.C. from The Whitaker Foundation. We thank the Murdock Foundation for computational and X-ray equipment in the Biomolecular Structure Center, Wim Hol and David Hyre for assistance with the isothermal titration calorimetry, Lisa Klumb and Ashutosh Chilkoti for their calorimetry work on wild-type streptavidin, the support staff at SSRL for their assistance with data collection, and Ernesto Freire and the Biocalorimetry Center at Johns Hopkins University (a Biomedical Research Technology Resource Center sponsored by NIH Grant RR04328) for their program and assistance with the calculations of ΔC_p.

References

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.

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 tryptophan residues 79, 108 and 120. *Proc Natl Acad Sci USA 92*:1754–1758.

Einspahr H, Parks EH, Suguna K, Subramanian E, Suddath FL. 1986. The crystal structure of pea lectin at 3.0 Å resolution. *J Biol Chem 261*:16518–16527.

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

Falzone CJ, Wright PE, Benkovic SJ. 1994. Dynamics of a flexible loop in dihydrofolate reductase from *Escherichia coli* and its implication for catalysis. *Biochemistry 33*:439–442.

Freire E, Luque I, Townsend B. 1997. *ASAcalc: Calculation of accessible surface areas*, v. 1.0. Biocalorimetry Center: The Johns Hopkins University.

Freitag S, Le Trong I, Klumb L, Stayton PS, Stenkamp RE. 1997. Structural studies of the streptavidin binding loop. *Protein Sci 6*:1157–1166.

Goldenberg DP, Creighton TE. 1983. Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J Mol Biol 164*:407–413.

Graf R, Schachman HK. 1996. Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase. *Proc Natl Acad Sci USA 93*:11591–11596.

Green NM. 1990. Avidin and streptavidin. *Methods Enzymol 184*:51–67.

Hahn M, Piotukh K, Borriss R, Heinemann U. 1994. Native-like *in vivo* folding of a circularly permuted jellyroll protein shown by crystal structure analysis. *Proc Natl Acad Sci USA 91*:10417–10421.

Hardman KD, Ainsworth CF. 1972. Structure of concanavalin A at 2.4 Å resolution. *Biochemistry 11*:4910–4919.

Heinemann U, Hahn M. 1995. Circular permutation of polypeptide chains: Implications for protein folding and stability. *Prog Biophys Mol Biol 64*:122–143.

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

Hofmann K, Wood SW, Brinton CC, Montibeller JA, Finn FM. 1980. Iminobiotin affinity columns and their applications to retrieval of streptavidin. *Proc Natl Acad Sci USA 77*:4666–4668.

Horlick RA, George HJ, Cooke GM, Tritch RJ, Newton RC, Dwivedi A, Lischwe M, Salemme FR, Weber PC, Horuk R. 1992. Permuteins of interleukin 1β—A simplified approach for the construction of permuted proteins having new termini. *Protein Eng 5*:427–431.

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

Konnert JH, Hendrickson WA. 1980. A restrained parameter thermal factor refinement procedure. *Acta Crystallogr A36*:344–350.

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

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.

Lee B, Richards FM. 1971. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol 55*:379–400.

Lindqvist Y, Schneider G. 1997. Circular permutations of natural protein sequences: Structural evidence. *Curr Opin Struct Biol 7*:422–427.

Luger K, Hommel U, Herold M, Hofsteenge J, Kirschner K. 1989. Correct folding of circularly permuted variants of a $\beta\alpha$ barrel enzyme *in vivo*. *Science 243*:206–209.

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.

Morton A, Matthews BW. 1995. Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: Linkage of dynamics and structural plasticity. *Biochemistry 34*:8576–8588.

Murphy KP, Xie D, Garcia KC, Amzel LM, Freire E. 1993. Entropy in biological binding processes: Estimation of translational entropy loss. *Proteins 15*:113–120.

Navaza J. 1994. AMoRe: An automated package for molecular replacement. *Acta Crystallogr A50*:157–163.

Noble ME, Zeelen JP, Wierenga RK. 1993. Structures of the "open" and "closed" state of trypanosomal triosephosphate isomerase, as observed in a new crystal form: Implications for the reaction mechanism. *Proteins 16*:311–326.

Otwinowski Z, Minor W. 1994. *DENZO: A film processing program for macromolecular crystallography*. New Haven, Connecticut: Yale University.

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

Pieper U, Hayakawa K, Li Z, Herzberg O. 1997. Circularly permuted β-lactamase from *Staphylococcus aureus* PC1. *Biochemistry 36*:8767–8774.

Read RJ. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A42*:140–149.

Reeke GN Jr, Becker JW. 1986. Three-dimensional structure of favin: Saccharide binding-cyclic permutation in leguminous lectins. *Science 234*:1108–1111.

Sano T, Cantor CR. 1990. Expression of a cloned streptavidin gene in *Escherichia coli*. *Proc Natl Acad Sci USA 87*:142–146.

Sheldrick GM. 1997. *SHELXL, program for structure refinement*. Göttingen: University of Göttingen.

Spolar RS, Record MT. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science 263*:777–784.

Tanaka T, Kato H, Nishioka T, Oda J. 1992. Mutational and proteolytic studies on a flexible loop in glutathione synthetase from *Escherichia coli* B: The loop and arginine 233 are critical for the catalytic reaction. *Biochemistry 31*:2259–2265.

Viguera AR, Serrano L, Wilmanns M. 1996. Different folding transition states may result in the same native structure. *Nat Struct Biol 3*:874–880.

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

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

Wierenga RK, Noble ME, Postma JP, Groendijk H, Kalk KH, Hol WGJ, Opperdoes FR. 1991. The crystal structure of the "open" and the "closed" conformation of the flexible loop of trypanosomal triosephosphate isomerase. *Proteins 10*:33–49.

Yang Y, Schachman HK. 1993. Aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains. *Proc Natl Acad Sci USA 90*:11980–11984.