Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Atomic resolution structure of biotin-free Tyr43Phe streptavidin: what is in the binding site?

Stefanie Freitag,[a,b]† Isolde Le Trong,[a]† Lisa A. Klumb,[b]‡ Patrick S. Stayton[b] and Ronald E. Stenkamp[a]*

[a]Department of Biological Structure and Biomolecular Structure Center, University of Washington, Box 357420, Seattle, Washington 98195-7420, USA, and [b]Department of Bioengineering, University of Washington, Box 352125, Seattle, Washington 98195-2125, USA

† S. Freitag and I. Le Trong should be considered co-first authors on this paper.
‡ Present address: Amgen, MS 8-1-C, One Amgen Center Drive, Thousand Oaks, CA 91320, USA.

Correspondence e-mail: stenkamp@u.washington.edu

Received 3 November 1998
Accepted 4 February 1999

PDB Reference: Y43F streptavidin mutant, 1swu.

The streptavidin–biotin system is an example of a high-affinity protein–ligand pair ($K_a \simeq 10^{13}$ mol^{-1}). The thermodynamic and structural properties have been extensively studied as a model system for protein–ligand interactions. Here, the X-ray crystal structure of a streptavidin mutant of a residue hydrogen bonding to biotin [Tyr43Phe (Y43F)] is reported at atomic resolution (1.14 Å). The biotin-free structure was refined with anisotropic displacement parameters (using the *SHELXL*97 program package). The high-resolution data also allowed interpretation of side-chain and residue disorder in 41 residues where alternate conformations were refined. The Y43F mutation is unambiguously observed in difference maps, although only a single O atom per monomer is altered. The atomic resolution enabled the identification of 2-methyl-2,4-pentanediol (MPD) molecules in the biotin-binding pocket for the first time. Electron density for MPD was observed in all four subunit binding sites of the tetrameric protein. This was not possible with data at lower resolution (1.8–2.3 Å) for wild-type streptavidin or mutants in the same crystal form using MPD in the crystallization. The impact of MPD binding on these studies is discussed.

1. Introduction

In recent years, an increasing number of atomic resolution X-ray protein structures have been reported (Dauter *et al.*, 1995, 1997; Merritt *et al.*, 1998). The advantage of high resolution in protein crystallography can be summarized by the following points.

(i) Disordered groups (*e.g.* side chains) can be detected more easily and refined in alternate conformations with occupancies adding up to 1.0.

(ii) The structural model for bound water molecules becomes more accurate.

(iii) The anisotropic refinement of thermal displacement parameters allows the interpretation of the direction of motion in molecular parts.

(iv) Overall, the refined model becomes more precise and the structure can be discussed in greater detail.

(v) H atoms are observed in difference electron-density maps and included in the refinements.

These results of high resolution are important for investigating biochemical mechanisms or protein–protein/protein–ligand interactions – especially because relatively small structural changes can play a vital role.

Streptavidin and its tight-binding ligand biotin ($K_a \simeq 10^{13}$ mol^{-1}) are widely used in biochemistry (Green, 1975; Bayer & Wilchek, 1990). Within this protein–ligand pair, three major binding motifs which are also observed in other systems

Table 1
Data collection and processing for the Y43F mutant.

Data set	Lower resolution	High resolution	All data
Detector distance (mm)	250	140	
Exposure	7 s	Dose 1500 counts	
Step size (°)	1.0	1.0	
Number of images	100	120	220
Resolution limit (Å)	50–1.8	50–1.14	50–1.14
Observed reflections	107083	520666	627749
Unique reflections	31852	148469	150170
Completeness (%)			
Overall	80	91	92
High-resolution shell	68 (1.90–1.83 Å)	81 (1.20–1.14 Å)	81 (1.20–1.14 Å)
R_{merge} (%)			
Overall	6.0	4.3	5.6
High-resolution shell	6.1 (1.86–1.80 Å)	14.3 (1.16–1.14 Å)	14.4 (1.16–1.14 Å)
$I/\sigma(I)$			
Overall	15.2	13.4	13.6
High-resolution shell	12.0 (1.90–1.83 Å)	6.1 (1.20–1.14 Å)	5.8 (1.20–1.14 Å)

Table 2
Reflection data statistics for refinement data set.

Resolution range (Å)	Unique reflections	Mean $I/\sigma(I)$	$I > 2\sigma(I)$ (%)	Completeness (%)
10.00–8.000	201	23.21	40.44	40.44
8.000–5.000	1402	23.96	93.96	94.03
5.000–3.000	6733	24.64	95.46	95.56
3.000–2.800	2041	25.34	98.55	98.69
2.800–2.600	2689	25.11	97.95	98.21
2.600–2.400	3667	24.66	97.56	98.26
2.400–2.200	5090	24.41	97.52	97.94
2.200–2.000	7293	23.81	97.09	97.67
2.000–1.800	10825	20.76	95.80	96.94
1.800–1.600	16735	16.43	93.20	96.00
1.600–1.400	27243	12.57	88.50	94.44
1.400–1.300	20103	9.41	81.42	92.31
1.300–1.200	26669	7.10	75.81	90.15
1.200–1.138	19287	5.76	64.30	81.19
All data	149978	13.62	84.10	92.13

play a major role: hydrophobic interactions, a disorder-to-order transformation of a binding loop and an elaborate hydrogen-bonding network. To examine the different ligand-binding motifs, extensive studies have been performed by a number of investigators: crystallographic studies of the wild-type protein (Hendrickson *et al.*, 1989; Weber *et al.*, 1989), calculations of binding properties (Miyamoto & Kollman, 1993*a*,*b*), variations of the binding ligand (for an overview, see Weber *et al.*, 1995; Katz, 1997*a*,*b*; Athappilly & Hendrickson, 1997; Voss & Skerra, 1997) and mutation of the protein (Sano & Cantor, 1995; Sano *et al.*, 1997; Chilkoti & Stayton, 1995; Chilkoti *et al.*, 1995).

Another approach to deciphering the binding energetics in the streptavidin–biotin system involves the generation of single-site mutants of the protein which are analyzed thermodynamically and kinetically as well as structurally. In our previous studies of streptavidin–biotin binding, we investigated X-ray crystal structures of wild-type and biotin binding-site mutants of the protein. All these structures were determined at resolutions of 1.8–2.3 Å (Freitag *et al.*, 1997, 1998; Chu *et al.*, 1998; Freitag, Chu *et al.*, 1999; Freitag, Le Trong *et al.*, 1999). This resolution range allowed us to refine the structures using *SHELXL*97 (Sheldrick & Schneider, 1997), giving the advantage of simultaneous atomic coordinate and isotropic *B*-factor refinement. With high-quality crystals of the streptavidin Tyr43Phe mutant (Y43F), we collected a data set at low temperature to a resolution of 1.14 Å using synchrotron radiation. The structural model of the Y43F variant was refined using *SHELXL*97 including all data in the resolution range 10–1.14 Å. The final *R* value for this model and data with $I > 2\sigma(I)$ is 0.121 with anisotropic refinement of temperature factors. This is in the range of 8–12% for *R* values of atomic resolution protein structures as described by Dauter *et al.* (1995). Streptavidin crystallizes with a tetramer in the asymmetric unit (4 × 13 kDa), making it one of the larger proteins refined at atomic resolution.

2. Materials and methods

The gene construction, protein expression and purification of the Y43F streptavidin mutant are described elsewhere (Klumb *et al.*, 1998). The protein was crystallized from solutions at 17 mg ml^{-1} using sitting-drop vapor-diffusion methods. Crystals grew in 50% MPD solutions at room temperature. The crystal chosen for data collection measured 0.2 × 0.4 × 1.0 mm. To assure rapid freezing, only half of this crystal was mounted in a loop with mother liquor at 113 K. The Y43F mutant crystallizes in space group $P2_1$ with unit-cell parameters a = 58.2, b = 84.9, c = 46.4 Å and β = 98.8°.

2.1. Data collection and processing

Data were collected at SSRL beamline 9-1 with a MAR Research imaging-plate scanner. The wavelength was 0.98 Å and the crystal diffracted to a resolution beyond 1.0 Å. Data collection to higher resolution than 1.14 Å was not possible owing to hardware and time restraints. Two data sets were collected at different crystal-to-detector distances and exposure times to avoid too many overloaded intensities at low resolution. First, we collected the high-resolution data to 1.14 Å at 140 mm distance (total time, 6 h). The lower resolution data to 1.8 Å were then collected on the same crystal at a distance of 250 mm (1.5 h). Because the current of the synchrotron is not constant, the high-resolution data were collected in dose mode instead of time mode. A crystal decay of about 30% in $I/\sigma(I)$ was observed after collection of the high-resolution data set and was corrected in data processing. Data were processed with *DENZO* and merged and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). 627749 observed reflections were merged to give 150170 unique reflections with an overall R_{merge} of 5.6% and a completeness of 92% to 1.14 Å (Table 1). Statistics for the data set used in the refinements are given in Table 2.

Table 3
Refinement progress.

Refinement step	Protein atoms (disordered)	H atoms	Water occupancy 1.0/0.5	MPD atoms	Number of parameters	Number of restraints	R/R_{free} $[I > 2\sigma(I)]$
Rigid-body tetramer	3354	—	—	—	7	13978	0.46/0.46
Rigid-body monomers	3354	—	—	—	25	13981	0.45/0.45
Coordinates	3354	—	—	—	10063	13961	0.38/0.40
Isotropic B	3354	—	—	—	13417	13953	0.27/0.30
Y43F mutation	3350	—	—	—	13401	13940	0.27/0.30
Loop, C-terminus	3478	—	—	—	13913	14451	0.26/0.29
Water	3478	—	414/0	—	15569	14468	0.20/0.24
MPD	3482	—	430/0	40	15809	14601	0.20/0.23
Disorder	3487 (89)	—	424/0	40	16161	15145	0.20/0.23
Solvent water correction	3487 (88)	—	435/0	40	16205	15159	0.20/0.22
Anisotropic protein	3487 (88)	—	426/0	32	34010	42891	0.15/0.18
Anisotropic MPD, water	3487 (88)	—	474/0	32	36764	46007	0.14/0.17
Hydrogen model	3487 (177)	3276	373/180	48	38388	48976	0.12/0.15
Final model	3487 (177)	3276	373/177	48	38361	48997	0.12/0.15

2.2. Structure solution and structural refinement

A wild-type streptavidin model in the same crystal form (Protein Data Bank code 1swa; Freitag *et al.*, 1997) without the binding loop (residues 45–51) or solvent molecules was used as a starting point for refinement. The refinement process utilized the program *SHELXL*97 (Sheldrick & Schneider, 1997), its auxiliary program *SHELXPRO* (map calculation, update of structure file for refinement, data and model analysis) and the graphics program *XTALVIEW* (McRee, 1992). The programs *WHATIF* (Vriend & Sander, 1993) and *PROCHECK* (Laskowski *et al.*, 1993) were also used for structure evaluation. The course of the refinements is shown in Fig. 1 and Table 3. Results of the refinement are summarized in Table 4.

The structure was refined against squares of the structure-factor amplitudes (F^2). All parameters were refined simultaneously. 10% of the data were used for the calculation of R_{free} (Brünger, 1992). Distance, planarity and chiral volume restraints were applied, as were anti-bumping restraints. The target values for 1–2 and 1–3 distances were based on the Engh & Huber (1991) study. A full-matrix least-squares rigid-body refinement was initially carried out for the whole tetramer and subsequently for all four separate subunits. The following conjugate-gradient least-squares refinement with

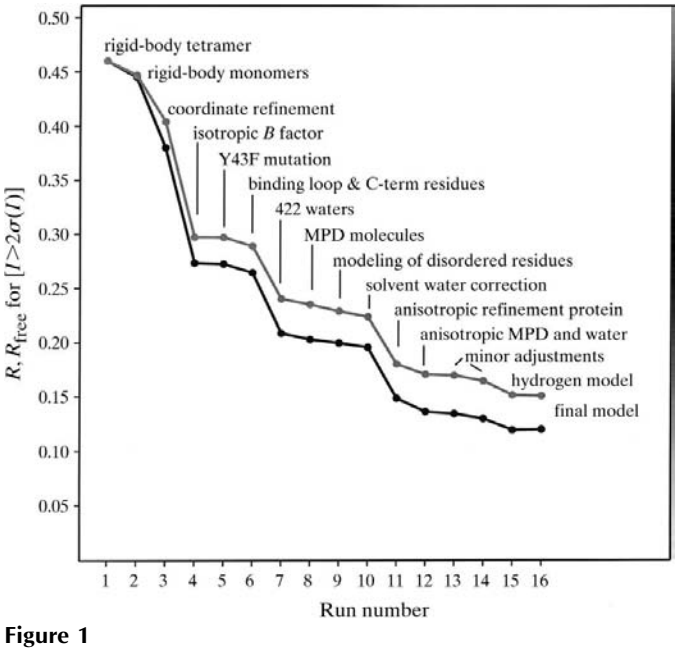

Figure 1
Course of refinement described by R (black) and R_{free} (gray) values for data with $I > 2\sigma(I)$. The circles represent the R values after the described refinement step. Only the major steps are represented for clarity.

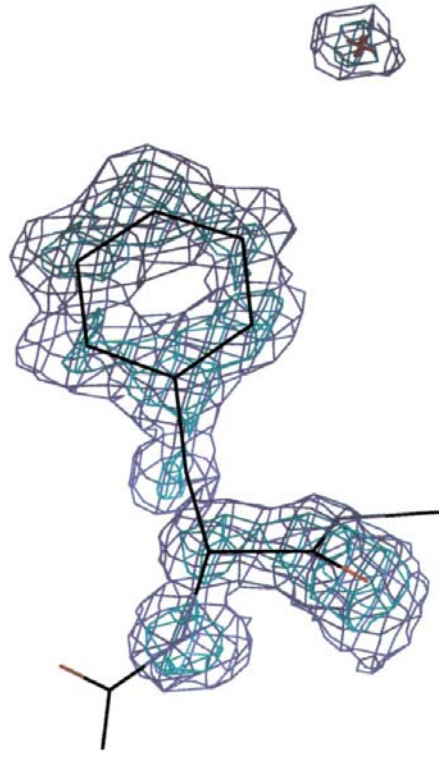

Figure 2
$|F_o| - |F_c|$ electron-density map for the mutation site Y43F when the residue is omitted from the refinement. The blue map is contoured at 3σ, the cyan map is contoured at 5σ. The additional density in the upper right was subsequently refined as a water O atom.

constant isotropic B value for all atoms resulted in a R value of 0.380 for data with $I > 2\sigma(I)$. (All following R values are for data fulfilling this criterion if not explicitly noted otherwise.) At this point in the refinement, R_{free} was 0.404. The biggest decrease in R values resulted from the introduction of isotropic B factors in the refinement ($R = 0.273$, $R_{free} = 0.297$). Similarity restraints were applied for isotropic and later anisotropic displacement parameters throughout the refinement. In the difference maps, negative density was observed on the side-chain O-atom positions of residue 43 when it was refined as tyrosine. Omit maps [σ_A-weighted (Read, 1986) $|F_o| - |F_c|$] showed positive density for phenylalanine residues at position 43 (Fig. 2). Residues of the biotin-binding loop (45–51) were partially observed in difference maps, as were additional C-terminal residues. The R value dropped to 0.264 ($R_{free} = 0.287$) on including these residues. 422 initial water positions were included using the automated water-searching program *SHELXWAT* (Sheldrick & Schneider, 1997). Water O atoms were rejected when B values increased above 63 Å ($U_{ij} \geq 0.8$ Å^2). They were refined with an occupancy of 0.5 if $U_{ij} \geq 0.5$ Å^2. After a visual check with *XTALVIEW* (McRee, 1992) and additional refinement cycles, the R value was 0.204 ($R_{free} = 0.237$). For bulk solvent, a correction using Babinet's principle (Moews & Kretsinger, 1975) was applied. Density for MPD molecules in the biotin-binding sites of the tetramer was clearly observed in a σ_A $|F_o| - |F_c|$ map and MPD molecules were included in the refinements (Fig. 3). While the B values for the MPD molecules in subunits 1 and 3 were relatively low, the molecules in subunits 2 and 4 showed higher B values. The B values were fixed at a reasonable value and the occupancies of the different MPD molecules were refined.

The quality of the maps allowed the refinement of alternate conformations of side chains for 37 residues and for alternate positions of complete residues in four cases. The refinement of anisotropic temperature factors for the protein atoms decreased R to 0.148 and R_{free} to 0.180. For anisotropic displacement parameters (ADP), rigid-bond restraints (Trueblood & Dunitz, 1983) were applied. The anisotropic refinement of water and MPD atoms resulted in an R value of 0.136 and an R_{free} of 0.170. In the case of water molecule O atoms, a soft linear restraint (effective standard deviation of 0.1) which holds the ADPs approximately isotropic was used. Minor adjustments (adding/omitting water O atoms, fitting of MPD positions and side-chain conformations and refinement of side-chain occupancies) and editing of the model based on σ_A-weighted $|F_o| - |F_c|$, $2|F_o| - |F_c|$ maps and the *SHELX* output further decreased R to 0.131 and R_{free} to 0.168. At this point in the refinement, densities for many of the H atoms in the hydrogen bonds between β-strands were observed in difference maps, as were some of the C$^\alpha$ H atoms (Fig. 4). H atoms are included in the refinements as a riding model ($R = 0.120$, $R_{free} = 0.151$).

Table 4
Structure refinement and final model for the Y43F mutant.

Resolution range (Å)	10–1.14
Number of unique reflections	149978
Number of parameters	38361
Number of restraints	48997
R [$2\sigma(I)$]†	0.121
R (all data)†	0.126
R_{free} [$2\sigma(I)$]‡	0.151
R_{free} (all data)‡	0.157
Goodness of fit/restrained	1.663/1.504
Number of atoms	
Protein	3487
Solvent (occupancy 1.0/0.5)	373/177
Heteroatoms	48
Average isotropic equivalent B value (Å^2)	
Protein, overall	15.0
Main chain (C, O, N, C$^\alpha$)	13.4
Side chain	16.7
Water	26.9
Heteroatoms	17.1

† For 100% of the data. ‡ For 10% of the data.

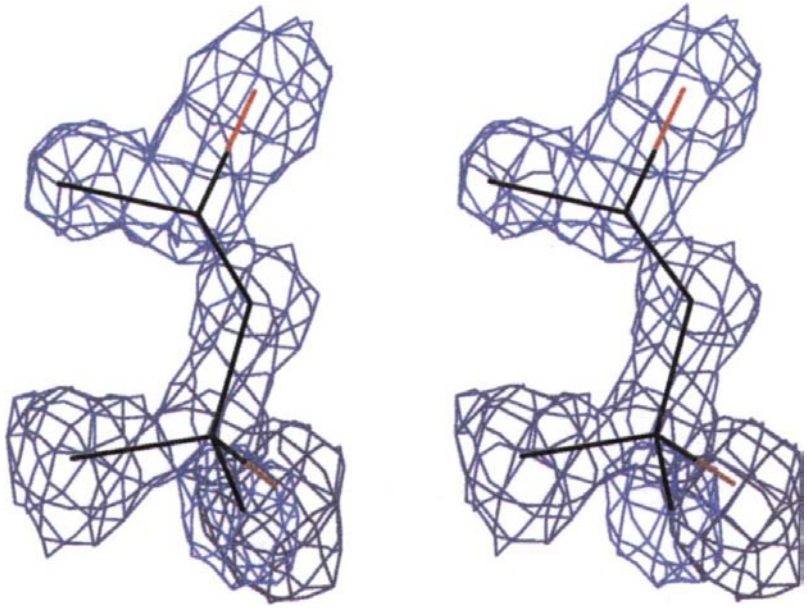

Figure 3
Stereoplot of an $|F_o| - |F_c|$ difference map showing the density for one molecule of MPD in the third subunit contoured at 3σ.

2.3. Final model

In the final streptavidin Y43F model, four subunits with residues 16–135 (1), 15–45 and 49–133 (2), 16–46 and 51–133 (3) and 16–44 and 48–133 (4) form the tetramer. The biotin binding-loop residues are in the open conformation in subunits 2 and 4 and in the closed conformation in subunit 1, as observed in wild-type streptavidin for the same crystal form (Freitag *et al.*, 1997). In subunits 2 and 4, electron density for this loop is not seen for all of the loop residues. In subunit 3, the electron density is even weaker. The loop is disordered in this subunit and not defined in our model. It is important to note that the binding loop does not fold over the MPD found in the

binding site in this structure. This is an important difference compared with biotin binding.

The final model of the Y43F streptavidin mutant shows excellent stereochemistry. In the Ramachandran plot (Fig. 5), 99% of the φ/ψ angles are in the core region typically observed for protein structures. It should be noted that the torsion angles of the protein chain were not restrained to any target values during the refinements. One outlier in the Ramachandran plot is Ser52 of subunit 1 (φ = 65.5°; ψ = −168.4°). This conformation is always found for this residue when the biotin-binding loop adopts the closed conformation. Another outlier is residue Glu101 in subunit 2 (φ = −115.2°; ψ = 70.4°). This residue is not well defined in electron-density maps, even at 1.14 Å resolution. This can be explained by the overall high flexiblity of this loop region. The φ/ψ angles of residue Asn23 of subunit 3 (φ = −104.8°; ψ = −169.2°) are very similar to those of the same residue in the other subunits and can be interpreted as systematic outliers of the main regions in the Ramachandran plot. This is also the case for the Asn81 torsion angles in all subunits. A non-systematic outlier with a greater ψ angle compared with its equivalents in the other three subunits is Thr66 in subunit 3 (φ = −117.8°; ψ = 59.9°). This residue is well defined in electron-density maps.

In the final model, six molecules of MPD were refined with occupancies varying from 0.3 to 1.0 (Fig. 6). 373 fully occupied and 177 half-occupied water O atoms were refined. The last refinement step includes the working (90%) as well as the $R_{\rm free}$ data (10%). The resulting R value is 0.121. Figs. 2, 3, 4, 6 and 7 are *XTALVIEW* plots (McRee, 1992); Figs. 1 and 5 are *SHELXPRO* (Sheldrick & Schneider, 1997) output; Figs. 8(*a*) and 8(*b*) are *PARVATI* (Merritt, 1999) output.

3. Results

In the high-resolution model of Y43F streptavidin, the mutation is well defined in omit maps of residue 43 (Fig. 2). There is clearly no density for a tyrosine O atom. The mutation did not alter the previously reported fold of the polypeptide chain of streptavidin (Weber *et al.*, 1989; Hendrickson *et al.*, 1989). Comparisons with three of our wild-type streptavidin structures (1swa, 1swb, 1swc; Freitag *et al.*, 1997) show only minor deviations in the Y43F mutant structure. The r.m.s.d. (root-mean-square distance) for the structures in the same crystal form is 0.24 Å (1swa) and 0.34 Å (1swb), superimposing 65 C^{α} atoms of the β-sheet regions (residues 19–23, 28–33, 38–42, 54–60, 71–80, 85–97, 103–112 and 123–131). Superimposing 1swc in the same way results in an r.m.s.d. of 0.36 Å. In some regions, the wild-type streptavidin model in the different crystal form (1swc) shows larger deviations from the mutant compared with the other two wild-type models. Deviating side-chain conformations in the high-resolution mutant structure are mainly observed in loop regions of the β barrel and include residues 22–26, 30–32, 35, 36, 39, 42–53, 66–69, 80–84, 97–103, 110, 113–118, 121 and 131–135. These residues show relatively high B values in the wild-type structures and have a tendency to be disordered. In some of these regions, disorder could be resolved and a second position for the side-chain atoms or even the whole residue was refined. Table 5 lists these residues, the disordered atoms and their refined occupancies. A comparison of the subunits shows some residues tend to be more disordered than others. A sign of the high mobility of the biotin-binding loop in streptavidin crystals is the lack of interpretable electron density for residues in this region. A nearly identical set of loop residues are missing in the atomic resolution Y43F mutant structure and in the lower resolution (1.9–2.0 Å) wild-type models in the same crystal form (Freitag *et al.*, 1997).

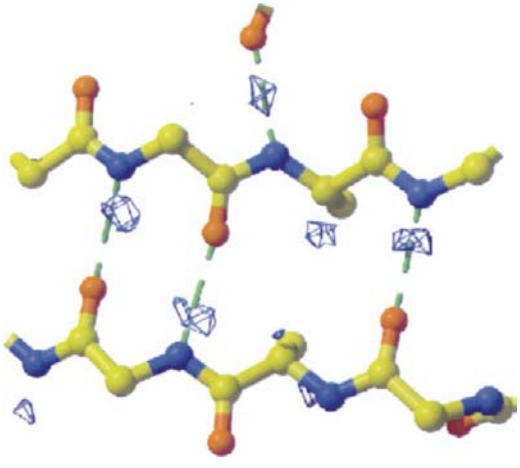

Figure 4
$|F_o| - |F_c|$ difference electron-density map at a 3σ level. The H-atom positions in some of the β-sheet hydrogen bonds can be observed, as well as some C^{α} H atoms.

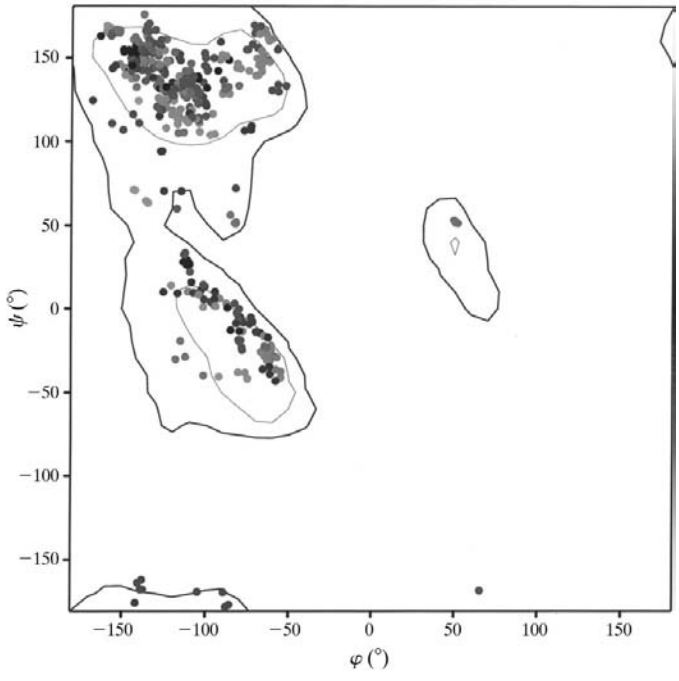

Figure 5
Ramachandran plot. φ/ψ angles for 393 standard residues excluding glycine residues are shown. 99.0% are in the core region and 82.4% are in the inner core region according to Kleywegt & Jones (1996).

Table 5
Discretely disordered residues and side chains in streptavidin mutant Y43F at 1.14 Å.

The atoms for which a second position was refined and the refined occupancies are given.

Residue	Subunit 1	Subunit 2	Subunit 3	Subunit 4
Tyr22	C^{β} to OH 0.5/0.5	—	—	—
Phe29	—	Residue† 0.5/0.5	—	—
Ile30	—	C^{β} to $C^{\delta1}$ 0.4/0.6	—	—
Thr32	—	—	—	C^{β} to $C^{\gamma2}$ 0.55/0.45
Thr40	—	C^{β} to $C^{\gamma2}$ 0.45/0.55	C^{β} to $C^{\gamma2}$ 0.65/0.35	—
Thr42	—	C^{β} to $C^{\gamma2}$ 0.45/0.55	—	—
Ser52	—	C^{β} to O^{γ} 0.5/0.5	—	C^{β} to O^{γ} 0.4/0.6
Leu56	—	C^{β} to $C^{\delta2}$ 0.55/0.45	C^{β} to $C^{\delta2}$ 0.5/0.5	C^{β} to $C^{\delta2}$ 0.45/0.55
Thr57	C^{β} to $C^{\gamma2}$ 0.5/0.5	C^{β} to $C^{\gamma2}$ 0.6/0.4	C^{β} to $C^{\gamma2}$ 0.5/0.5	—
Thr66	—	—	C^{β} to $C^{\gamma2}$ 0.6/0.4	—
Ser69	—	—	Residue† 0.65/0.35	—
Leu73	—	C^{β} to $C^{\delta2}$ 0.4/0.6	C^{β} to $C^{\delta2}$ 0.5/0.5	—
Lys80	—	N^{ζ} 0.4/0.6	—	—
Ser88	C^{β} to O^{γ} 0.45/0.55	C^{β} to O^{γ} 0.55/0.45	C^{β} to O^{γ} 0.5/0.5	C^{β} to O^{γ} 0.5/0.5
Glu101	C^{β} to $O^{\varepsilon2}$ 0.7/0.3	—	—	—
Gln107	C^{β} to $N^{\varepsilon2}$ 0.5/0.5	—	C^{β} to $N^{\varepsilon2}$ 0.75/0.25	C^{β} to $N^{\varepsilon2}$ 0.5/0.5
Leu110	—	C^{β} to $C^{\delta2}$ 0.45/0.55	C^{β} to $C^{\delta2}$ 0.6/0.4	C^{β} to $C^{\delta2}$ 0.45/0.55
Glu116	—	—	C^{β} to $O^{\varepsilon2}$ 0.5/0.5	—
Trp120	C^{γ} to $C^{\zeta3}$ 0.5/0.5	—	C^{β} to $C^{\zeta3}$ 0.5/0.5	—
Lys121	C^{β} to N^{ζ} 0.5/0.5	—	—	—
Asp128	—	Residue† 0.6/0.4	—	C^{β} to $O^{\delta2}$ 0.75/0.25
Thr129	—	—	C^{β} to $C^{\gamma2}$ 0.4/0.6	C^{β} to $C^{\gamma2}$ 0.6/0.4
Thr131	—	—	C^{β} to $C^{\gamma2}$ 0.4/0.6	—
Pro135	Residue† 0.55/0.45	—	—	—

† 'Residue' implies that all atoms (main-chain and side-chain) are refined in two positions.

The most striking result from the atomic resolution structure of Y43F is the identification of MPD molecules in the biotin-binding sites (Fig. 3). We have used MPD in many of our previous crystallizations and have never detected density for this compound in our 1.8–2.0 Å crystal structures. Other groups using MPD for streptavidin crystallization have also not reported MPD molecules in their electron-density maps. In the Y43F binding sites of subunit 1 and 3, two molecules of MPD were refined, while in subunits 2 and 4 only a single MPD was found. The atomic displacement parameters (ADPs) of these molecules vary over a large range, with two molecules (in subunits 1 and 3) having reasonable ADPs. The occupancies of the MPD molecules were refined fixing the *B* values to a reasonable value. Afterwards, the occupancies were set to the new values and the molecules were refined anisotropically. The refined occupancies are reported in the legend to Fig. 6. In all four subunits, MPD is observed in a position which overlaps with that of biotin when superimposed on the streptavidin–biotin complex (Fig. 7). The MPD molecules overlap the bicyclic part of biotin. In subunits 1 and 3, an additional molecule of MPD is observed overlapping with the valeric acid tail of biotin. Comparing our ligand-free wild-type and mutant structures with Y43F, there is positional overlapping between water O and MPD atoms in the binding site of the protein (Fig. 7).

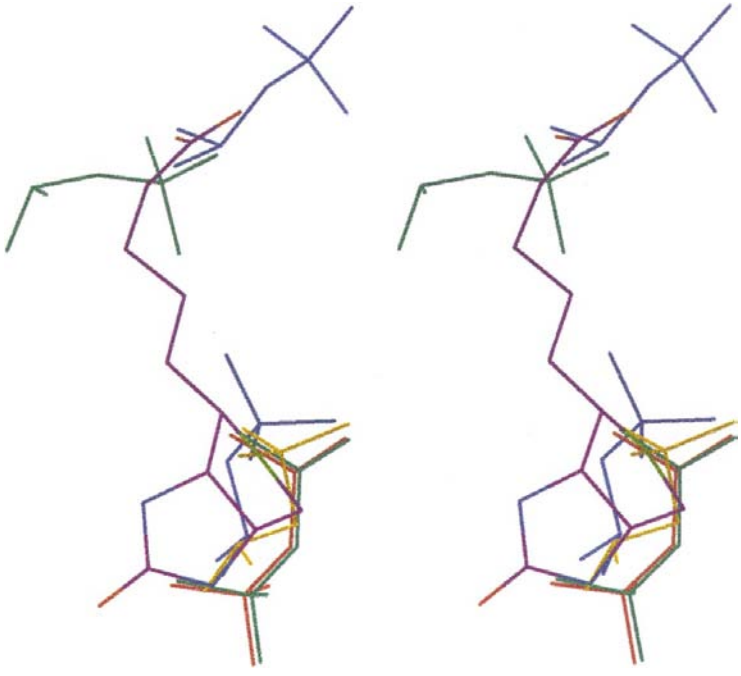

Figure 6
Stereo plot of a superposition of the MPD molecules in all four subunits in Y43F on one molecule of biotin in the streptavidin binding site. Subunit 1, blue (occupancy for lower molecule is 1.0, for the upper 0.65); subunit 2, red (occupancy 0.6); subunit 3, green (occupancies 0.6 and 0.3); subunit 4, yellow (occupancy 0.4); biotin from wild-type structure, purple.

The refinement of anisotropic displacement parameters in atomic resolution protein structures improves difference electron-density maps and allows interpretation of protein motion. The program *PARVATI* (Merritt, 1999) analyses the anisotropy of protein models. Anisotropy is here defined as the ratio of the minimum and maximum principal axes of the atomic

ellipsoids. The distributions of anisotropy by distance from the center of mass and by atom class are shown in Figs. 8(*a*) and 8(*b*), respectively, for the refined Y43F model. Fig. 8(*a*) displays a reasonable distribution of anisotropy for a globular protein. Fig. 8(*b*) shows that none of the protein atoms is perfectly isotropic (anisotropy = 1.00). Only seven of the protein atoms (0.2%) have a high anisotropy (<0.10). 278 atoms (6.5%) display anisotropy values smaller than 0.20.

In the biotin-binding site, Trp120 displays higher anisotropy owing to the higher mobility of this residue in the biotin-free state. This was also observed for the 1.8–2.0 Å wild-type streptavidin structures (Freitag *et al.*, 1997). Another residue which forms a hydrogen bond to one of the biotin carboxylate O atoms is Ser88. Its side chain was modeled in all four subunits in two positions, showing high mobility. For this residue, a comparison with the biotin-bound mutant structure at high resolution would be interesting. None of the other residues involved in biotin binding (tryptophan or hydrogen-bonding residues) show systematic motions in their anisotropic displacement parameters. Overall, there are no major structural changes upon mutation to Y43F streptavidin. Of course, the biotin-bound structure of this mutant could reveal

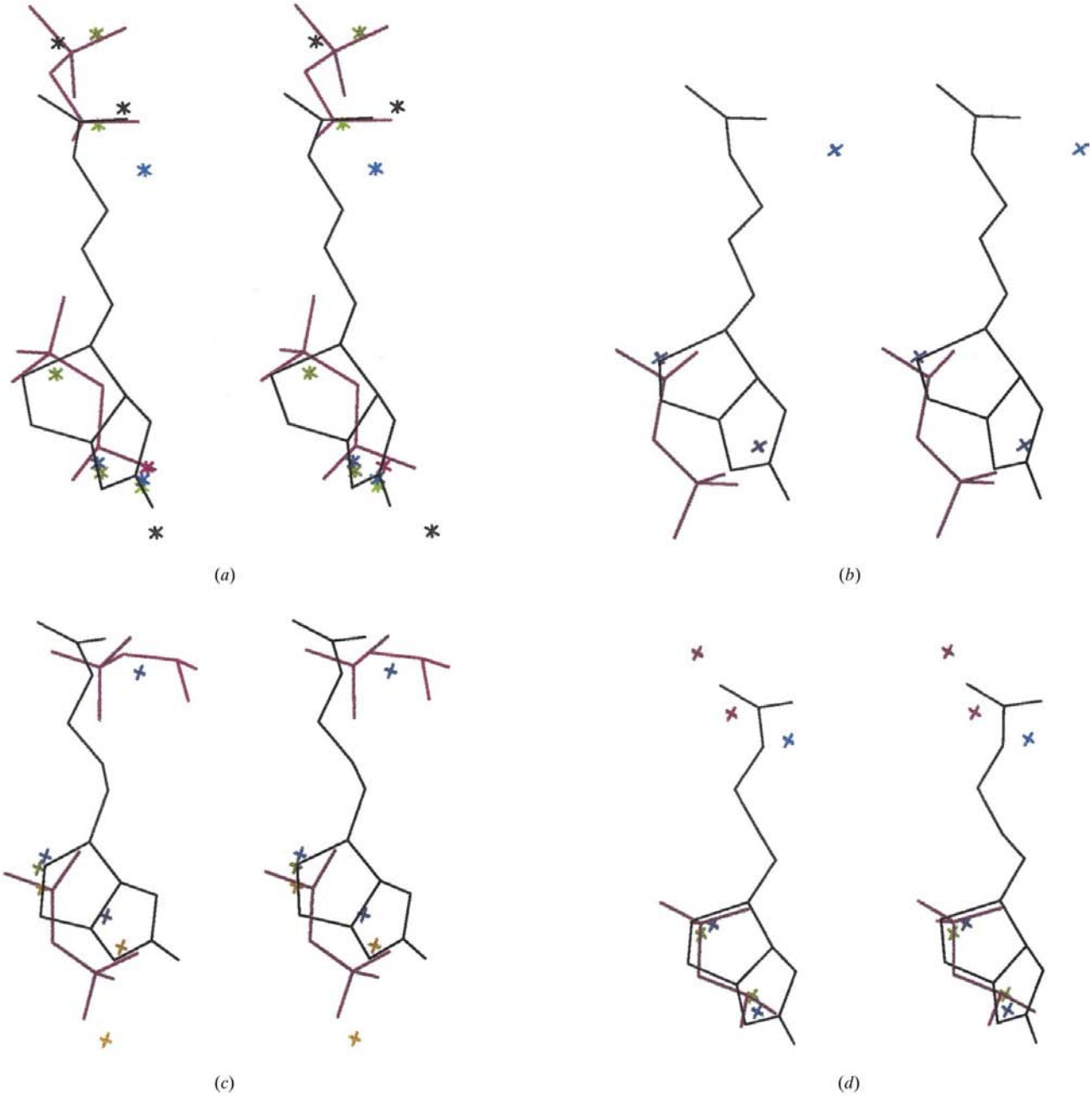

Figure 7
Superposition of the waters in the biotin-binding pocket of wild-type streptavidin 1swa (green), 1swb (blue), 1swc (brown) and biotin-bound 1swe (black) on the MPD molecules in the Y43F mutant (purple) for all four subunits (*a*)–(*d*).

more information. Studies of this complex and other biotin-free and biotin-bound streptavidin mutants are under way (Freitag *et al.*, in preparation).

4. Discussion

Biophysical results of studies including the Y43F mutant (Klumb *et al.*, 1998) will be discussed elsewhere, also taking into account the lower resolution (1.8 Å) biotin-bound structure of Y43F (Freitag *et al.*, in preparation).

The identification of MPD in the binding site of Y43F is only possible because of the high resolution of the data set and is very important for the interpretation of our previous results. In our previous work, we compared ligand-free and biotin-bound structures of streptavidin and its mutants and correlated the structures with thermodynamic and kinetic studies (Chilkoti *et al.*, 1995; Chilkoti & Stayton, 1995; Chu *et al.*, 1998; Klumb *et al.*, 1998). The possibility that the binding site in our biotin-unbound structures is not only occupied by water molecules but by another ligand could potentially affect structural interpretations if these MPD molecules alter the unbound structure. Eight of our published biotin-free streptavidin wild-type or mutant structures (three wild-type structures, two structures of the Trp79Phe mutant, mutant structures of Y43F, Asp128Ala and the circular permuted streptavidin CP51/46) were crystallized with MPD. Three mutant structures were crystallized under different conditions (Trp108Phe, 22–24% PEG 1000, 0.1 *M* Tris buffer pH 7.0; Trp120Ala, Trp120Phe, 12–20% PEG 4000, 0.1 *M* HEPES buffer pH 7.5) and the density in the binding site can only be identified as water molecules (Freitag *et al.*, 1997, 1998; Chu *et al.*, 1998; Freitag, Chu *et al.*, 1999). This observation leads to the question of how specific the streptavidin–MPD binding really is. The disorder and different occupancies of MPD in the four binding sites (Fig. 6) show that the binding of MPD is more randomized and not as specific as the binding of biotin. It can be compared with water binding. We believe it unlikely that the structural effects on the rest of the protein are significant when MPD is bound in the binding site. It is, however, important to note that at lower resolution, small molecules such as MPD can be mistaken for water molecules.

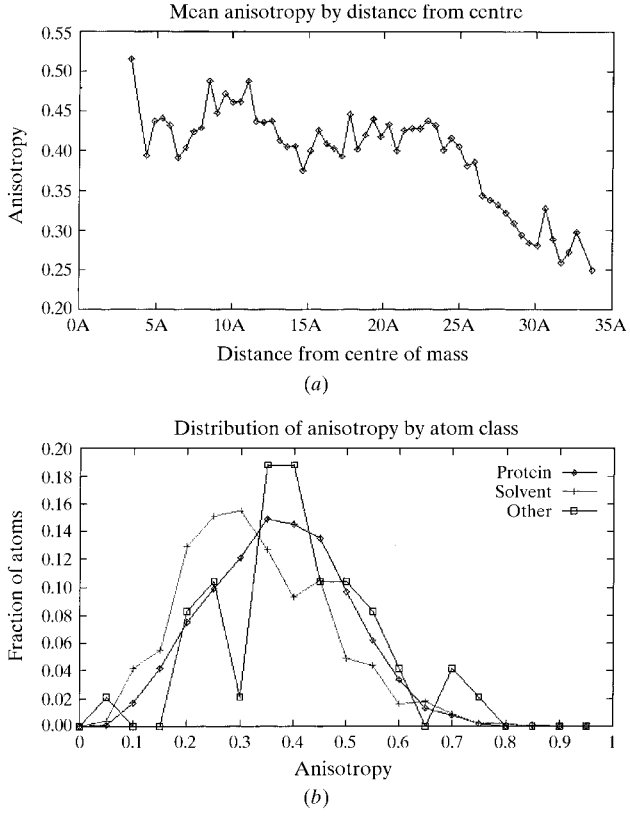

Figure 8
PARVATI (Merritt, 1999) plots describing the anisotropy of the atomic resolution model of the Y43F mutant. The anisotropy is the ratio of minimum and maximum principal axes of the atomic ellipsoids. A perfectly isotropic atom has an anisotropy of 1.0.

We thank Peter Kuhn, Craig Behnke, Richard Garzon and Brian Fox for their assistance with the data collection at SSRL. This work was supported by grant DK49655 from the National Institutes of Health. We thank the Murdock Foundation for financial support of the Biomolecular Structure Center.

References

Athappilly, F. K. & Hendrickson, W. A. (1997). *Protein Sci.* **6**, 1338–1342.
Bayer, E. A. & Wilchek, M. (1990). *Methods Enzymol.* **184**.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Chilkoti, A. & Stayton, P. S. (1995). *J. Am. Chem. Soc.* **117**, 10622–10628.
Chilkoti, A., Tan, P. H. & Stayton, P. S. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 1754–1758.
Chu, V., Freitag, S., Le Trong, I., Stenkamp, R. E. & Stayton, P. S. (1998). *Protein Sci.* **7**, 848–859.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). *Curr. Opin. Struct. Biol.* **5**, 784–790.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Freitag, S., Chu, V., Penzotti, J. E., Klumb, L.A., To, T., Le Trong, I., Lybrand, T. P., Stayton, P. S. & Stenkamp, R. E. (1999). In preparation.
Freitag, S., Le Trong, I., Chilkoti, A., Klumb, L. A., Stayton, P. S. & Stenkamp, R. E. (1998). *J. Mol. Biol.* **279**, 211–221.
Freitag, S., Le Trong, I., Klumb, L. A., Chu, V., Stayton, P. S. & Stenkamp, R. E. (1999). In preparation.
Freitag, S., Le Trong, I., Klumb, L. A., Stayton, P. S. & Stenkamp, R. E. (1997). *Protein Sci.* **6**, 1157–1166.
Green, N. M. (1975). *Adv. Protein Chem.* **29**, 85–143.
Hendrickson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A. & Phizackerley, R. P. (1989). *Proc. Natl. Acad. Sci. USA*, **86**, 2190–2194.
Katz, B. A. (1997*a*). *Annu. Rev. Biophys. Biomol. Struct.* **26**, 27–45.
Katz, B. A. (1997*b*). *J. Mol. Biol.* **274**, 776–800.
Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.
Klumb, L. A., Chu, V. & Stayton, P. S. (1998). *Biochemistry*, **37**, 7657–7663.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
McRee, D. E. (1992). *J. Mol. Graphics*, **10**, 44–46.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.

Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K. & Hol, W. G. J. (1998). *J. Mol. Biol.* **282**, 1043–1059.
Miyamoto, S. & Kollman, P. A. (1993*a*). *Proc. Natl Acad. Sci. USA*, **90**, 8402–8406.
Miyamoto, S. & Kollman, P. A. (1993*b*). *Proteins*, **16**, 226–245.
Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–228.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–325.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Sano, T. & Cantor, C. R. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 3180–3184.
Sano, T., Vajda, S., Smith, C. L. & Cantor, C. R. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 6153–6158.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Trueblood, K. N. & Dunitz, J. D. (1983). *Acta Cryst.* B**39**, 120–133.
Voss, S. & Skerra, A. (1997). *Protein Eng.* **10**, 975–982.
Vriend, G. & Sander, C. (1993). *J. Appl. Cryst.* **26**, 47–60.
Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J. & Salemme, F. R. (1989). *Science*, **243**,85–88.
Weber, P. C., Pantoliano, M. W. & Salemme, F. R. (1995). *Acta Cryst.* D**51**, 590–596.