Ligand-Induced Conformational Change in Transferrins: Crystal Structure of the Open Form of the N-Terminal Half-Molecule of Human Transferrin[†,‡]

Philip D. Jeffrey,[§,||] Maria C. Bewley,[§,||] Ross T. A. MacGillivray,[⊥] Anne B. Mason,[#] Robert C. Woodworth,[#] and Edward N. Baker*[,§,||]

Department of Biochemistry, Massey University, Palmerston North, New Zealand, Department of Biochemistry, University of British Columbia, Vancouver, Canada V6T 1Z3, and Department of Biochemistry, University of Vermont, Burlington, Vermont 05405

Received May 21, 1998; Revised Manuscript Received July 20, 1998

ABSTRACT: Serum transferrin binds ferric ions in the bloodstream and transports them to cells, where they are released in a process involving receptor-mediated endocytosis. Iron release is believed to be pH dependent and is coupled with a large conformational change. To help define the steps in iron release, we have determined the three-dimensional structure of the iron-free (apo) form of the recombinant N-lobe half-molecule of human serum transferrin ($ApoTf_N$) by X-ray crystallography. Two crystal forms were obtained, form 1 with four molecules in the asymmetric unit and form 2 with two molecules in the asymmetric unit. The structures of both forms were determined by molecular replacement and were refined at 2.2 and 3.2 Å resolution, respectively. Final *R*-factors were 0.203 (free $R = 0.292$) for form 1 and 0.217 (free $R = 0.312$) for form 2. All six copies of the $ApoTf_N$ structure are essentially identical. Comparison with the holo form ($FeTf_N$) shows that a large rigid-body domain movement of 63° has occurred in $ApoTf_N$, to give an open binding cleft. The extent of domain opening is the same as in the N-lobe of human lactoferrin, showing that it depends on internal constraints that are conserved in both proteins, and that it is unaffected by the presence or absence of the C-lobe. Although the conformational change is primarily a rigid-body motion, several local adjustments occur. In particular, two iron ligands, Asp 63 and His 249, change conformation to form salt bridges, with Lys 296 and Glu 83, respectively, in the binding cleft of the apo protein. Both salt bridges would have to break for iron coordination to occur. Most importantly, the structure, determined at a pH (5.3) that is close to the pH of physiological iron release, indicates that protonation of His 249 is a key step in iron release.

Iron is essential for all species but poses severe problems of toxicity and solubility. Under physiological conditions, the Fe(III) oxidation state is favored, but free Fe^{3+} ions are rapidly precipitated as insoluble ferric hydroxides and also have the capacity to catalyze the formation of damaging free-radical species. The proteins of the transferrin family thus play a crucial biological role of sequestering and solubilizing iron, thereby controlling the levels of free iron in body fluids (*1*−*3*). Serum transferrin, in particular, has the role of binding ferric ions in the bloodstream and transporting this bound iron to cells where it is released by a process of receptor-mediated endocytosis (*4*). Iron release occurs at the reduced intracellular pH (∼5.5), apparently with the active participation of the receptor (*5*), and the iron-free apotransferrin is returned to circulation without degradation.

The requirements for transferrin function are a high affinity for iron, as Fe(III), coupled with a mechanism for release of this tightly bound iron. Crystal structures of human lactoferrin (*6*−*8*), rabbit transferrin (*9*), and chicken ovotransferrin (*10*) have revealed the canonical transferrin fold and the nature of the two high-affinity ($K \approx 10^{20}$) iron-binding sites. In each protein, the polypeptide chain is folded into two lobes, representing its N- and C-terminal halves. Both lobes are of similar structure, comprising two α/β domains with a deep cleft between them; in each lobe a single Fe^{3+} ion is bound in this cleft to the side chains of two tyrosine residues, one aspartate, and one histidine, together with a CO_3^{2-} ion which bridges between the metal ion and an anion-binding site on one domain (*3*).

The sequence of steps leading to iron release is not well understood, although kinetic studies point to several key protonation events (*11*), and biophysical studies (*12, 13*) show that iron release is associated with a major conformational change which appears to be broadly similar in all

[†] Supported by grants from the U.S. Public Health Service (DK-21739 to R.C.W. and HD-20859 to E.N.B.) and the Health Research Council of New Zealand (to E.N.B.). E.N.B. also acknowledges research support as an International Research Scholar of the Howard Hughes Medical Institute, and R.T.A.M. acknowledges support from a Killam Senior Research Fellowship and U.S.P.H.S. Grant DK-35533 during his stay at Massey University.

[‡] Atomic coordinates for both structures have been deposited with the Protein Data Bank, Brookhaven National Laboratory, Upton, NY, with accession codes 1bp5 (form 1) and 1btj (form 2).

* Corresponding author. Phone (64) (9) 373-7599. Fax: (64) (9) 373-7619. E-mail: Ted.Baker@auckland.ac.nz.

[§] Massey University.

[||] Current addresses: P.D.J., Memorial Sloan Kettering Cancer Center, New York, NY 11021; M.C.B., Biology Department, Brookhaven National Laboratory, Upton, NY 11973; E.N.B., School of Biological Sciences and Department of Chemistry, University of Auckland, Private Bag 92019, Auckland, New Zealand.

[⊥] University of British Columbia.

[#] University of Vermont.

S0006-2960(98)01206-9 CCC: $15.00
Published on Web 09/15/1998

transferrins. The recently determined crystal structure of the iron-bound N-lobe of human transferrin has identified protonation and dissociation of the anion as a likely first step in iron release (*14*), consistent with the kinetic studies (*11*). A trigger mechanism, involving protonation of a pair of lysine residues from opposing domains of the N-lobe, has also been suggested from the crystal structure of the N-lobe half-molecule of ovotransferrin (*15*), but its importance relative to other protonation events is not clear.

To date, the detailed nature of the conformational change in transferrins has been defined only in the crystal structure of apolactoferrin (*16*), where a domain rotation of 54° gives rise to an open N-lobe binding cleft; this depends on a hinge movement in two antiparallel polypeptide strands that pass behind the iron site, which allows one domain to move as a rigid body relative to the other (*17*). There are, however, distinct structural and functional differences between lactoferrin and transferrin, notably the fact that transferrin releases iron from its N-lobe at a distinctly higher pH than lactoferrin (5.5, compared with 3.0). Amino acid substitutions, including the presence of an additional disulfide bond between Cys 137 and Cys 331 in human transferrin may also influence the conformational change; this joins two regions that move independently in lactoferrin (*16, 17*).

Here, we describe the structure of the unliganded, iron-free (apo) form of the recombinant N-terminal half-molecule of human transferrin ($ApoTf_N$,[1] residues 1−337). The structure has been determined at high resolution by X-ray crystallography at a pH (5.3) at which the protein would be iron-free in vivo. In concert with that of the iron-laden holo form (*14*), this structure therefore provides a snapshot of the two states of transferrin at physiologically relevant conditions of pH.

MATERIALS AND METHODS

The N-terminal half molecule of human transferrin (residues 1−337 of the native protein) was expressed in baby hamster kidney (BHK) cells and purified from the medium in the Fe(III) form, as previously described (*18*). To obtain the iron-free protein for crystallization trials, the purified $FeTf_N$ was dialyzed against 0.5 M sodium citrate buffer, pH 4.9, containing 1 mM EDTA and 1 mM nitrilotriacetate, as described by He et al. (*19*). For crystallization, the apoprotein was concentrated to 20−40 mg/mL in 20 mM sodium bicarbonate, 50 mM KCl.

Crystallization. Conditions for crystal growth were identified using a systematic screening procedure based on orthogonal arrays (*20*); ethanol and 2-methyl-2,4-pentanediol (MPD) were used as precipitants, over a pH range 4.9−9.1 and a concentration range 10−60% (v/v).

Two crystal forms were obtained under essentially identical conditions. The crystals were grown at 4 °C in sitting drops, on microbridges. A 1.5 μL drop of protein solution (20 mg/mL $apoTf_N$ in 20 mM $NaHCO_3$ and 50 mM KCl) was mixed with 1.5 μL of reservoir solution (50 mM potassium acetate, pH 5.3, 35% (v/v) MPD). Colorless crystals appeared over a period of 2−4 weeks. Crystal form 1 grew as cubic prisms, which were monoclinic, spacegroup $P2_1$, with unit cell dimensions $a = 83.5$, $b = 77.3$, $c = 107.5$ Å, $\beta = 98.3°$, and diffracted to 2.0 Å resolution. Assuming four molecules of 37.8 kDa/asymmetric unit, the solvent content is 46% ($V_m = 2.27$ Å^3/Da). Crystal form 2 grew as trapezoidal prisms, which were also monoclinic, spacegroup $P2_1$, with cell dimensions $a = 63.0$, $b = 77.3$, $c = 72.6$ Å, $\beta = 104.3°$, and diffracted to 3.0 Å resolution. Assuming two molecules per asymmetric unit, the solvent content is also 46%. The two crystal forms are closely related, with the same space group and the same *b*-axis.

Table 1: Data Collection Statistics

	form 1	form 2
resolution range (Å)	37.0−2.1	40.0−3.2
no. of crystals	1	1
temp	113 K	113 K
no. of unique reflections	73 066	11 269
redundancy[a]	3.2 (2.1)	3.3 (3.3)
completeness[a] (%)	79.9 (61.1)	99.7 (99.9)
merging *R*-factor[a] (%)	4.6 (27.1)	11.1 (25.7)
mean I/σ[a]	9.2 (4.0)	5.9 (2.6)

[a] Corresponding values for the outermost resolution shell given in parentheses.

Data Collection. Both crystal forms could be flash-frozen after increasing the concentration of MPD in the mother liquor to 45%. Intensity data for each were collected at 113 K, with an R-Axis IIC image plate detector on a Rigaku RU200 rotating anode generator. Both data sets were processed with DENZO (*21*). Data for form 1 were scaled with ROTAVATA and AGROVATA from the CCP4 program suite (*22*) and data for form 2 were scaled with SCALEPACK (*21*). Form 1 data were processed to 2.1 Å resolution, with an overall redundancy of 3.2 and a completeness of 80% (61% complete in the outermost, 2.1−2.3 Å, shell). For form 2, data were processed to 3.2 Å resolution, with an overall redundancy of 3.3 and completeness of 100% (100% complete in the outermost, 3.2−3.5 Å, shell). Full data collection statistics are in Table 1.

Structure Solution and Refinement. Form 1. The structure was determined by molecular replacement using the program AMORE (*23*). Search models corresponding to the N1 and N2 domains of the iron-bound $FeTf_N$ structure (*14*) were used to find the corresponding domains of the $apoTf_N$ structure. There was a clear solution corresponding to four N1 domains and four N2 domains, which were also consistent with the presence of a pseudocentering in the *h*0*l* zone of the diffraction pattern (the space group is pseudo-$B2_1$ at low resolution). The molecular replacement solution had an *R*-factor of 41.1% and a correlation coefficient of 70.1% for data in the range 10−4 Å.

The structure was refined using XPLOR (*24*). Tight NCS restraints were applied initially but were relaxed later in the refinement as the four molecules proved to be clearly nonidentical in some respects. Individual isotropic temperature factors were also refined later in the refinement. Initial difference maps showed clear density for the side chains of the residues involved in metal and anion binding in the holo protein (which were omitted from the original map calculation). Density was also visible for additional residues at the C-terminus (not modeled in the holo protein), that made a

[1] Abbreviations: Tf, serum transferrin; Lf, lactoferrin; $ApoTf_N$, apo form of recombinant N-lobe of human transferrin; $FeTf_N$, ferric form of recombinant N-lobe of human transferrin; ApoLf, apo form of human lactoferrin; BHK cells, baby hamster kidney cells; EDTA, ethylenediaminetetraacetic acid; rms, root-mean-square.

Table 2: Refinement Details

	form 1	form 2
resolution range (Å)	6.0−2.2	6.0−3.2
no. of reflections	58663	9438
R-factor (%)	20.3	21.7
free *R* (%)[a]	29.2	31.2
rms deviations from standard geometry		
bond lengths (Å)	0.013	0.013
bond angles (deg)	1.7	1.9
no. of protein atoms	10 258	5156
no. of water molecules	555	0
avg *B*-value		
main-chain atoms	39.3	15.4
side-chain atoms	41.6	16.0
water molecules	41.5	

[a] For 10% of reflections.

fortuitous packing interaction with a symmetry-related molecule. The final model comprised 10 258 protein atoms (from residues 4−331 of molecules A and B and 4−337 of molecules C and D) and 555 solvent molecules all modeled as water. The final crystallographic *R*-factor was 20.3% for all data in the resolution range 6−2.2 Å, and the free *R* (*25*) was 29.2%. Refinement statistics are shown in Table 2.

Form 2. This structure was determined independently by molecular replacement, also using AMORE (*23*), with the same search models as for form 1. A single unambiguous solution, for two N1 domains and two N2 domains, was found, with a correlation coefficient of 0.57 and an *R*-factor of 37.4% for data in the resolution range 10−4 Å. Ten percent of the data (930 reflections) was omitted from the refinement with XPLOR (*24*) to use in the free *R* calculation. A variety of refinement protocols was tested. Due to the limited resolution of the data, strict NCS constraints were applied until the last cycle of refinement when very tight NCS restraints were applied. Grouped isotropic temperature factor refinement (1/residue) resulted in a larger reduction of the free *R*-factor than simply refining an overall *B*-value and was deemed an appropriate protocol. The final model, comprising 5156 protein atoms (residues 5−336 of both molecules) and no water molecules, gives an *R*-factor of 21.7% (R_{free} 31.2%) for all data in the resolution range 6−3.2 Å. Refinement statistics are shown in Table 2.

Miscellaneous. Model quality was assessed with PROCHECK (*26*). Hydrogen bonding was analyzed with the program HBPLUS (*27*) using the geometrical criteria of Baker and Hubbard (*28*). Coordinates have been deposited in the Protein Data Bank (Brookhaven National Laboratory) with the accession codes 1bp5 (form 1) and 1btj (form 2).

RESULTS AND DISCUSSION

Structure Determination and Model Quality. Structures were determined for two different crystal forms. The high-resolution (2.2 Å) form 1 structure has four molecules in the asymmetric unit (referred to here as molecules A−D) and the lower resolution (3.2 Å) form 2 structure has two molecules (E and F). Both were determined independently, by molecular replacement, using the two domains N1 and N2 of the holo protein (*14*) as search models, and these two crystal structures therefore give six independent views of the $ApoTf_N$ structure. In both analyses NCS averaging gave maps of significantly higher quality than the data resolution might suggest; some small but genuine differences between the molecules in the form 1 crystals, however, led to relaxation of the NCS restraints late in the refinement.

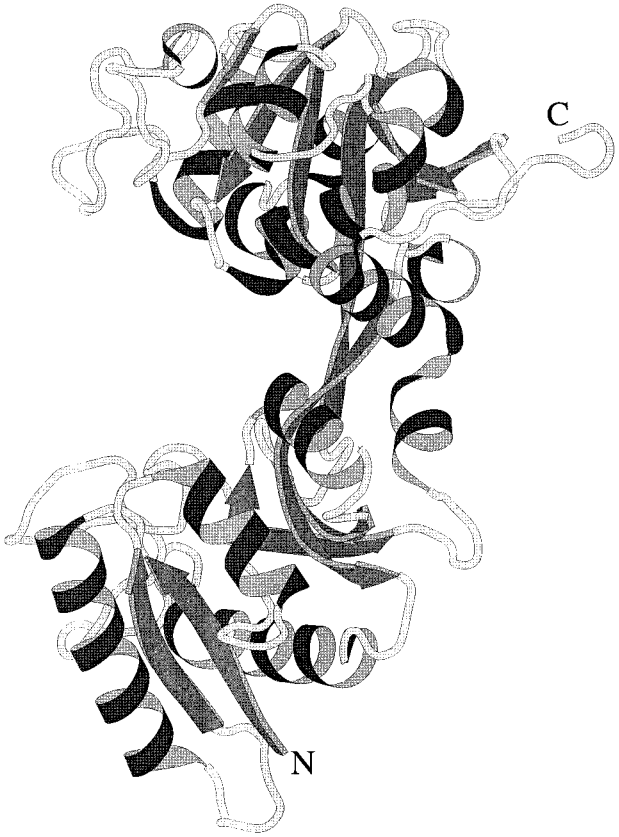

FIGURE 1: Ribbon diagram showing the structural organization of $ApoTf_N$, with the N1 domain below and the N2 domain above. The N- and C-termini are indicated. This figure and Figures 2, 4, and 5 drawn with MOLSCRIPT (*41*).

Details of the refinement and the two final models are summarized in Table 2. Residues 1−3 are not visible in any of the molecules. For crystal form 1, two of the molecules (A and B) do not have any clearly interpretable density beyond Cys 331, which provides the last link with the rest of the N-lobe, through the disulfide bridge 137−331. The other two molecules (C and D) have been modeled as far as the C-terminal residue, Asp 337. The conformation may be better ordered for these because of the crystal packing; the C-terminus is extended and projects into the open cleft of another molecule. For crystal form 2, both molecules have interpretable density for residues 4−336. In all of the molecules, residues Pro 74, Pro 142, and Pro 145 are in the cis conformation. Ramachandran plots show that only one residue in each molecule has ϕ, ψ angles that are normally disallowed. There is no doubt that this residue, Leu 294, has been correctly modeled; however, its electron density is very clear, and it is the central residue in a classic γ-turn with the expected ϕ, ψ angles of around (70, −60) (*28, 29*). This γ-turn is conserved in all other transferrin structures determined to date (*14, 30*).

Overall Conformation. All six molecules have the same open conformation that is characteristic of the unliganded (iron-free) form of transferrins; this is shown for molecule C in Figure 1. The polypeptide chain is folded into two domains, N1 (residues 1−93 + 247−315) and N2 (94−246), which are defined by the location of the hinge between them

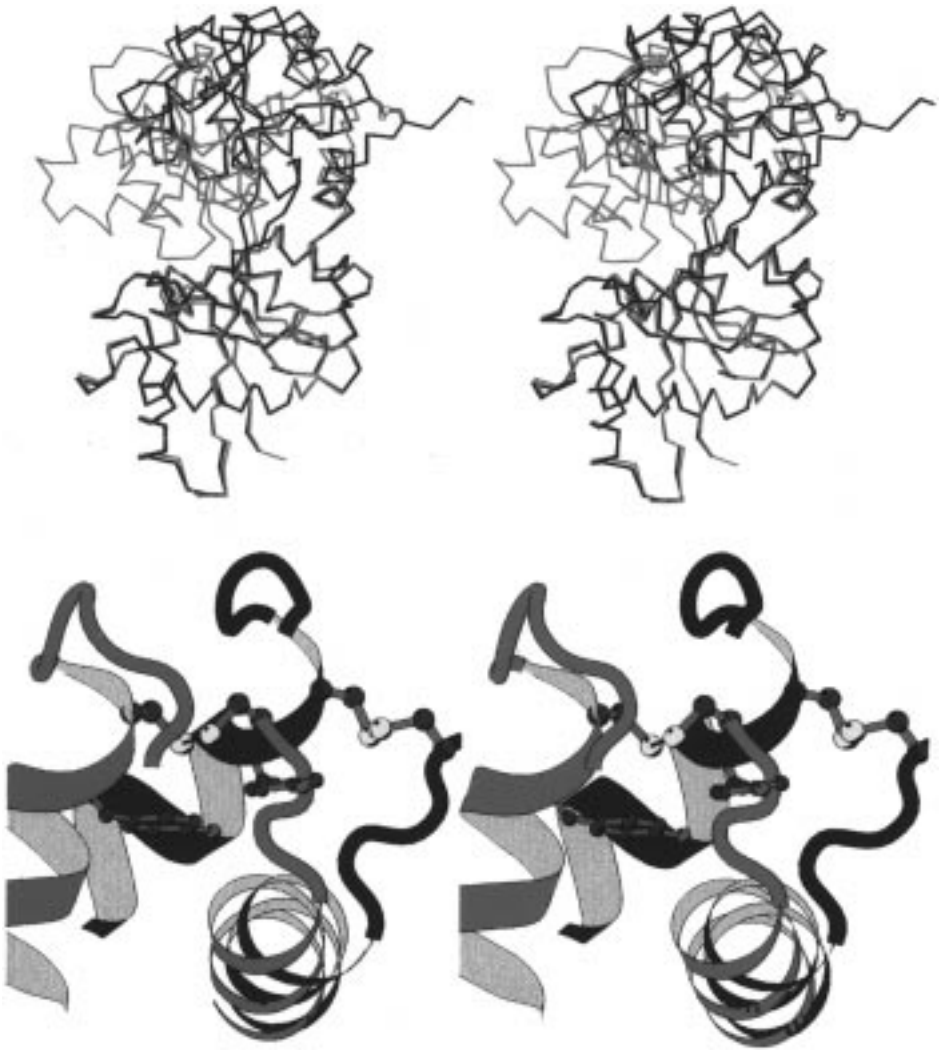

FIGURE 2: Two stereoviews of the conformational change in the N-lobe half-molecule of human transferrin. The holo form, $FeTf_N$, is shown in red, and the apo form, $ApoTf_N$, is shown in blue. In the top panel, the two structures have been superimposed using their N1 domains (residues 1−93 + 247−315); helix α11 (residues 316−327) also remains with the N1 domain, but the N2 domain is rotated 63° in the apo form, relative to its position in the holo form. The disulfide bond Cys 137−Cys 331, which links the N2 domain to the C-terminal region, following helix 11 (see text), is shown in yellow. In the bottom panel, a close-up of changes at the C-terminus is shown. Helix α5 from the N2 domain (left) pivots on helix α11 (lower right) and Tyr 136 and the disulfide Cys 137−Cys 331 push against the C-terminal polypeptide, seen on the right, displacing it.

(see below). Each domain has an α/β fold, in which the secondary structural elements are the same as described for the holo form, FeTfN (*14*). The molecule also contains seven disulfide bonds, two in the N1 domain (9−48 and 19−39), four in the N2 domain (118−194, 158−174, 161−179, 171−177, and 227−241), and one (137−331) that bridges between the N2 domain and the C-terminal region of the polypeptide.

Nature of the Conformational Change. The conformational change that differentiates $ApoTf_N$ from $FeTf_N$ involves a rigid body rotation of the N2 domain relative to the N1 domain (Figure 2, top), the extent of which is essentially the same in all six independent copies of the structure. The movement can be quantified by superimposing the N1 domain of $ApoTf_N$ on to the N1 domain of $FeTf_N$ and then determining what additional transformation is required to bring the two N2 domains into correspondence. By this analysis, a domain rotation of 63.2° has occurred in molecule C of $ApoTf_N$, relative to $FeTf_N$; in the other molecules, the domain rotation is very similar, ranging 62.9−64.8°. When the various $ApoTf_N$ molecules are superimposed, the rms differences in atomic positions between them (for the Cα atoms of residues 5−327) are 0.3−0.6 Å; these values are only marginally more than the rms differences when the individual domains are superimposed.

The rigid body nature of the conformational change is emphasized by superimposing the individual domains of the apo protein on to those of the holo protein. When this is done for molecule C of $ApoTf_N$, relative to $FeTf_N$, the rms difference in atomic positions in the N1 domain is 0.53 Å (for the Cα atoms of residues 5−93 + 249−315) and in the N2 domain is 0.41 Å (for the Cα atoms of residues 94−135 + 146−246). Overall, the polypeptide conformation within each of the domains remains essentially unchanged by the movement, except at two positions: the peptide 11−12 in the N1 domain is reoriented, and a rather mobile loop in the N2 domain, 136−145, also changes its conformation. Some of the side chains in the binding cleft are also reoriented. These changes are described below.

The C-terminal helix α11, residues 316−327, plays an important role in the conformational change, just as it does

in lactoferrin (*16*). Although it lies across the surface of the N2 domain, it does not move with it; instead, it remains tied to the N1 domain (Figure 2, top). Thus, if the N1 domain *plus* helix α11 (residues 5−93 + 249−327) is superimposed on to the corresponding residues in the holo protein, the rms difference is 0.56 Å, hardly any greater than for the main body of the N1 domain (0.53 Å). Helix α11 associates with the rest of the N1 domain primarily by burying a hydrophobic surface; the side chains of Tyr 319, Val 320, and Ile 323 from helix α11 pack against those of Tyr 68, Lys 312 (hydrophobic part), and Leu 315 of the N1 domain. A salt bridge, Asp 69···Arg 327, which remains unchanged in both apo and holo forms, also helps to tie this helix to the rest of the N1 domain; an analogous salt bridge (Glu 66···Arg 332) is present in lactoferrin (*30*). Helix α11 may be thought of as a pivot on which the N2 domain moves; helix α5 of the N2 domain lies across helix α11 at an angle of about 90° and during the conformational change the entire N2 domain moves as helix α5 pivots on helix α11 (Figure 2, bottom).

Beyond the C-terminus of helix α11, the polypeptide is displaced by the movement of the N2 domain, as can be seen in Figure 2, bottom. This arises both from the covalent connection provided by the Cys 137−Cys 331 disulfide bond, the conformation of which does not change, and from van der Waals contacts involving Tyr 136 and Cys 137 with the C-terminal polypeptide. Since the C-terminal polypeptide would lead into the interlobe connecting peptide in full-length transferrin, this has possible implications for signaling between the lobes (see later discussion).

What limits the extent of the domain movement? The most important factor in determining the nature and extent of the conformational change appears to be the exposure and burial of two alternative interfaces. In the open form, a large interface is exposed, comprising primarily residues in the binding cleft; these residues are identified in Table 3. This opening of the cleft exposes an extra ∼700 Å^2 of solvent-accessible surface. However, the opening of the cleft and exposure of this large interface is coupled with the burial of a second interface comprising primarily residues from helix α5, helix α11, and the polypeptide 243−245. A total of ∼450 Å^2 of surface area from this second interface becomes buried in the apo form. The movement can be visualized as a seesaw motion between two alternative close-packed interfaces, following the analysis of Gerstein et al. for lactoferrin (*17*). The extent of closure is defined by the first interface, and the extent of opening is defined by the second interface.

The Hinge. The molecular hinge in the transferrin N-lobe is located in the two antiparallel β-strands, strand βE (93−103) and strand βJ (243−254), that pass behind the iron-binding site. The hinge is quite localized, at Thr 93 in strand βE and between Val 246−Pro 247−Ser 248 in strand βJ. There is only one major change in the main-chain torsion angles in the hinge region of strand βE; this is at ψ(93), which changes by 48°, from 33° in $FeTf_N$ to −15° in $ApoTf_N$. In strand βJ, the torsion angle changes are smaller and more evenly distributed over several residues, but the two main torsion angle changes are in ψ(246) and ψ(247), which couple together to produce most of the conformational change; these change by 32 and 29°, respectively, taking the average for the four crystal form 1 molecules.

Table 3: Changes in Solvent Exposure Resulting from Domain Movement

residue	location	change in exposure[a]	residue	location	change in exposure[a]
		(a) Large Interface			
Val 11	N1 domain	13	Tyr 95	N2 domain	15
Ser 12	N1 domain	12	Gly 123	N2 domain	29
Glu 13	N1 domain	22	Arg 124	N2 domain	62
His 14	N1 domain	15	Ser 125	N2 domain	18
Asp 63	N1 domain	27	Asn 129	N2 domain	9
Leu 66	N1 domain	19	Cys 179	N2 domain	49
Asp 69	N1 domain	11	Ser 180	N2 domain	15
Tyr 85	N1 domain	25	Thr 181	N2 domain	56
His 249	N1 domain	−11	Leu 182	N2 domain	50
Lys 291	N1 domain	61	Tyr 188	N2 domain	29
Asp 292	N1 domain	11	Ser 189	N2 domain	5
Lys 296	N1 domain	17	Lys 206	N2 domain	21
Ser 298	N1 domain	25	Glu 212	N2 domain	52
Ile 323	helix α11	9			410[b]
Leu 326	helix α11	35			
		291[b]			
		(b) Small Interface			
Glu 318	Helix α11	−138	Ile 130	N2 domain	−20
Tyr 319	Helix α11	−23	Gly 133	N2 domain	−28
Ala 322	Helix α11	−48	Leu 134	N2 domain	−52
Asn 325	Helix α11	−27	Tyr 136	N2 domain	−16
Glu 328	Helix α11	−46	Cys 137	N2 domain	−17
		−282[b]	Leu 243	N2 domain	−15
			Ala 244	N2 domain	−22
					−170[b]

[a] Solvent exposure calculated by the method of Lee and Richards (*35*). Figures are calculated in angstroms squared. A positive value means greater solvent exposure in $ApoTf_N$. [b] Total change in exposure.

Just beyond the hinge, in strand βJ, Ser 248 changes its conformation radically from β-type in $FeTf_N$ (−51, 147) to α-type in $ApoTf_N$ (−66, −60), and the configuration of His 249 is changed from (−68, 151) in $FeTf_N$ to (−139, 121) in $ApoTf_N$. This change does not seem to be associated with the hinge movement, however, even though it is adjacent. Rather, it is linked to changes in side chains that are part of the transferrin iron-binding site, Glu 83, Tyr 85, His 249, and Lys 296, as described below.

Conformational Adjustments within the Domains. As noted above, the two domains move essentially as rigid bodies. There are, however, several structural adjustments within the domains that seem to be a consequence of the change between closed and open states.

In the N1 domain, the peptide between Val 11 and Ser 12 becomes reoriented during the change from closed to open form. In the closed form, these residues, which join strand βA to helix α1, pack against residues Ser 180 and Thr 181 in the opposing N2 domain. The interactions between these two regions involve van der Waals contacts and hydrogen bonds through two buried, highly ordered water molecules (both with *B*-values of ∼14 Å^2). In the change to the open form, the peptide 11−12 reorients and the Ser 12 side chain moves so that the hydroxyl group takes the place of one of the waters, forming a classic helix N-cap (*31, 32*); Ser 12 OG is hydrogen bonded to the NH of residue 15 at the N-terminus of helix α1 (12-OG···HN-15 = 2.93 Å). The rearrangement does not lead to increased interactions by the peptide, but the configuration of Ser 12 changes from the rarely found ε configuration (*33*) with (φ, ψ) values (69, −173) in the closed $FeTf_N$ structure to the much more common and presumably more favorable β configuration

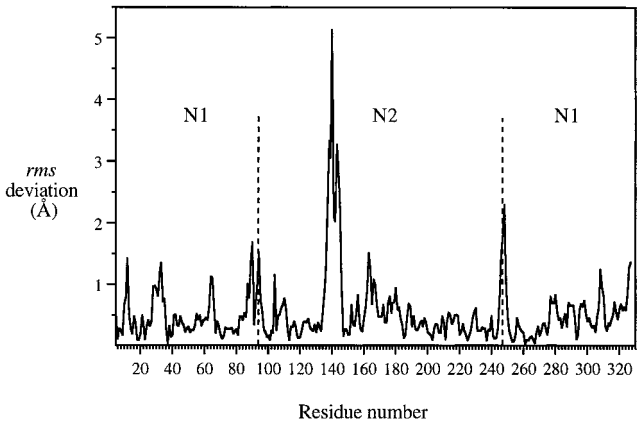

FIGURE 3: Plot showing the differences in Cα positions between $ApoTf_N$ and $FeTf_N$, as a function of residue number. In this plot the differences shown are those obtained when the individual domains of the two proteins are superimposed separately (N1 on N1, N2 on N2). The locations of the two hinges, which mark the boundaries between the N1 and N2 domains, are indicated with broken lines.

(−159, 162) in the open $ApoTf_N$ structure. The movement could thus be seen as releasing constraints inherent in the closed state.

A second rearrangement in the N1 domain involves the N-terminus of helix α3 (residues 63−72), which carries the iron-binding ligand Asp 63. In the iron-bound form, the N-terminus is distorted from regular α-helical geometry, with 63 O and 64 O making 3_{10}-type (1−4) hydrogen bonds to the NH groups of residues 66 and 67, respectively, and 65 O not hydrogen bonded at all. In the $ApoTf_N$ structure, these change to regular (1−5) α-helical hydrogen bonds. This could again be seen as relaxation (in the apo protein) from a more constrained holo structure imposed by iron binding and domain closure.

The other major polypeptide movement is in the N2 domain where the loop 135−145 is rearranged. The rearrangement is largest at Cys 137, which forms the disulfide bond with Cys 331, just beyond the C-terminus of helix α11 (Figure 2). Presumably, it is triggered, therefore, by the movement of the N2 domain; as helix α5 pivots on helix α11, the loop 135−145 pushes against the polypeptide that follows helix α11. Loop rearrangement then leads to enhanced van der Waals interaction involving Tyr 136 and Cys 137 in the $ApoTf_N$ structure. Although this loop is rather mobile in the $FeTf_N$ structure (B values 50−80 Å^2), it is less so in the apo structure (B values 30−40 Å^2), reflecting these increased van der Waals contacts.

A residue-by-residue comparison of Cα positions in the two domains, between $FeTf_N$ and $ApoTf_N$ (Figure 3), shows that there are very few other positions where differences occur; these are generally in mobile loops and do not seem to have any structural or functional implications.

Side Chains in the Binding Cleft. The location of the hinge results in a division of the iron-binding ligands, such that Tyr 95 and Tyr 188 remain with the N2 domain in the $ApoTf_N$ structure and Asp 63 and His 249 remain with N1. Tyr 95 and Tyr 188 project into the binding cleft, with their orientations hardly changed from the iron-bound form. The OH groups are 3.9 Å apart (cf. 3.3 Å in $FeTf_N$) and are close to the anion-binding site formed by Arg 124 and the N-terminus of helix α5 (residues 124−136). These two tyrosines appear poised to form the initial iron-binding locus, in association with the CO_3^{2-} ion when it binds, as is seen in an N2-domain fragment of ovotransferrin (*34*) and as proposed in schemes for iron binding by transferrins (*3*, *34*). The side chain of Arg 124 has an extended conformation, projecting into the binding cleft adjacent to the helix α5 N-terminus. The anion-binding site is thus largely preformed, but the Arg 124 side chain must fold down over the CO_3^{2-} ion when it binds. The structure of the holo form (*14*) has shown that conformational flexibility in this arginine also appears to play an important part in iron release (*14*). In four of the molecules of $ApoTf_N$, the C-terminus of a neighboring molecule projects into the binding cleft with the side-chain carboxylate of Asp 337 occupying the anion-binding site; this is clearly a crystal-packing artifact but emphasizes the high affinity for anions of this site.

The Asp 63 and His 249 ligands lie against the surface of the N1 domain, with the carboxylate oxygen of Asp 63 approximately 9 Å away from Tyr 95 OH (11 Å from Tyr 188 OH) and the imidazole ring of His 249 about 12 Å away from Tyr 95 OH (and 14 Å from Tyr 188 OH). In the apo protein, both of these side chains change conformation from their position in the iron-bound protein to form salt bridges that are not present in the latter (Figure 4).

In $FeTf_N$, Asp 63 binds to iron through OD1 and hydrogen bonds with Gly 65 NH (at the N-terminus of helix α3) and with Ser 125 NH (on domain N2); on the loss of iron and opening of the domains it swings away to form a salt bridge (3.06 Å) with the amino group of Lys 296. This movement is also associated with the rearrangement at the N-terminus of helix α3 in which the hydrogen bond geometry becomes more favorable (see above).

His 249 undergoes a similar movement. In $FeTf_N$, it binds to iron through NE2 and to a water molecule through ND1; this water molecule is in turn hydrogen bonded to the carboxylate of Glu 83. In the apo protein, the His side chain moves to displace the water molecule and form a direct salt bridge His 249 ND1···OE2 Glu 83 (2.61 Å). The Glu 83 side chain does not move at all, but the movement of the His side chain results from both a side-chain reorientation and a change in the main-chain configurations of Ser 248 and His 249 adjacent to the hinge.

A significant factor in the adjustments on the N1 domain appears to be the role of Lys 296. This is one of the two lysine residues that makes up the proposed "dilysine trigger" (*15*); in the iron-bound protein, it is hydrogen bonded to Lys 206 from the N2 domain and it is proposed that protonation converts this to a repulsive interaction which destabilizes the closed state. It also interacts with Tyr 85 and Tyr 188. In the apo protein, Lys 296 undergoes a small reorientation which places its amino group between the carboxylate of Asp 63 and the carbonyl oxygen of Ser 248. It thus plays a part both directly in positioning the Asp ligand and indirectly in orienting the His ligand (since the movement of His 249 depends in part on the reorientation of the Ser 248−His 249 peptide).

Comparison with Lactoferrin. Comparisons of the $ApoTf_N$ structure with the open N-lobe of apolactoferrin reveal a remarkable similarity in the nature and extent of the conformational change from their iron-bound forms, but at the same time, some intriguing differences involving functionally important residues.

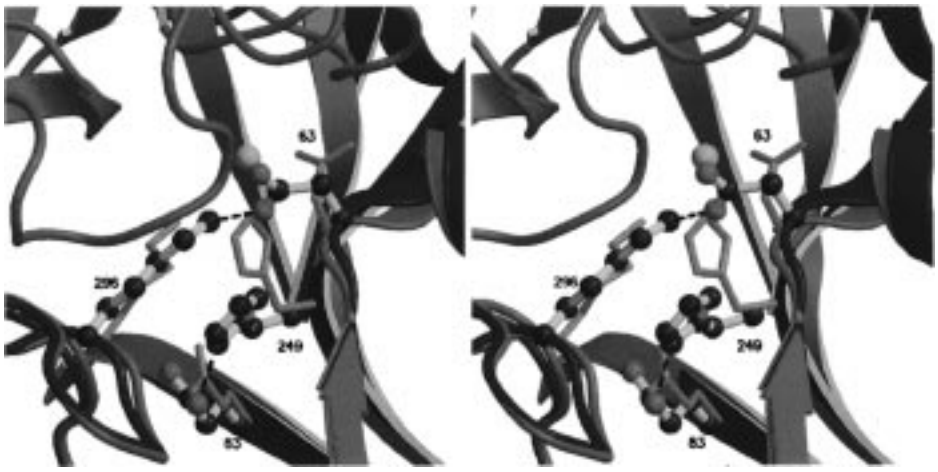

FIGURE 4: Stereo diagram showing the changed interactions involving Asp 63 and His 249. Here, the holo form, $FeTf_N$, is shown with its polypeptide backbone in red and side chains in green. The apo form, $ApoTf_N$, is shown with its polypeptide backbone in blue, and side chains in ball-and-stick representation, with pink bonds. In $FeTf_N$, Asp 63 and His 249 are bound to the iron atom (orange sphere), whereas in $ApoTf_N$ Asp 63 swings down to form a salt bridge with Lys 296, and His 249 swings down to form a salt bridge with Glu 83. Figure produced with MOLSCRIPT (*41*) and RASTER3D (*42*).

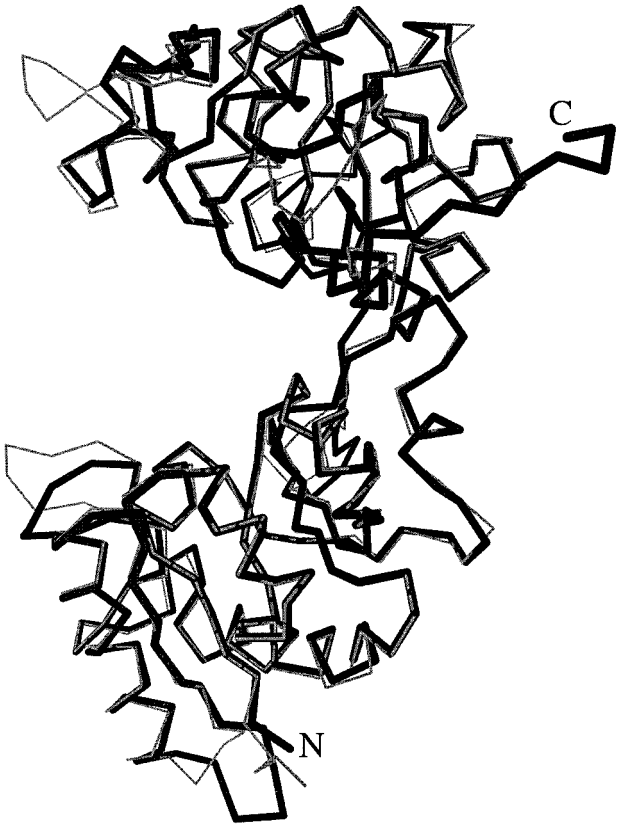

FIGURE 5: Cα plot showing the superposition of the open N-lobes of human transferrin (dark, thick bonds) and human lactoferrin (light, thin bonds), from the crystal structures of $ApoTf_N$ (this work) and ApoLf (*16*, *30*).

The larger domain rotation in the transferrin N-lobe, ~63° compared with 54° in lactoferrin (*16*), primarily reflects a difference in the closed, iron-bound forms rather than the open forms. The two open structures are, in fact, strikingly similar (Figure 5). Superposition of the entire N-lobes of the two apo proteins gives an rms difference of only 0.73 Å (for 304 corresponding C_α atoms), hardly any greater than for the individual N1 and N2 domains (0.69 and 0.62 Å, respectively). Domain-by-domain superpositions show that the difference in domain orientations between $ApoTf_N$ and the ApoLf N-lobe corresponds to a rotational difference of only 2.8°. Thus, the difference in the domain movement arises because transferrin has a slightly more open (2.8°) apo form, coupled with a more closed (5.5°) holo form. We stress, however, that the differences are extremely minor given the overall scale of the conformational change.

The similarity in the overall conformational changes between the transferrin and lactoferrin N-lobes results from the conserved nature of the two interfaces that are alternately buried and exposed during the domain movement (*17*). The residues in the small interface (Table 3) are highly conserved (10 out 15 residues identical between Tf and Lf), leading to very similar open forms. The large interfaces are less conserved, and some variation in the loops on either side of the binding cleft leads to the slightly greater closure of the iron-bound form of transferrin compared with lactoferrin. The conformational change also depends on the helix α11, across which the N2 domain pivots; this helix is tied to the N1 domain by a number of conserved interactions including a salt bridge that is present in both transferrin and lactoferrin. It is remarkable that helix α11 seems so firmly tied to the N1 domain considering the way it projects out like a rod, but superposition confirms that it is identically oriented in Lf and Tf.

In contrast, there are notable differences in several functionally important residues. The movements of Asp 63 and His 249 to form salt bridges in $ApoTf_N$ do not occur in ApoLf where the corresponding residues remain essentially unmoved from their positions in the holo form (*30*). Likewise, the flip of the Ser 248−His 249 peptide, which appears to be linked with the movement of His 249, also does not occur in lactoferrin. We believe that these differences may give some important insights into transferrin function (see below).

Two other points deserve comment. The disulfide bond 137−331 in transferrin, which has no counterpart in lactoferrin, does not appear to alter the nature of the domain movement. This is because the magnitude of domain opening depends on the residues belonging to helix α11 (316−327) and to the small interface. On the other hand,

Cys 331 follows four residues later at the start of the connecting peptide. In both proteins, the loop that follows helix α5, i.e., the loop 136−145 in transferrin or 137−145 in lactoferrin, is somewhat mobile. In transferrin, it is better defined in the open form than in the closed form, whereas in lactoferrin, the situation is reversed. In both proteins, however, a mobile loop in this position may help to cushion the N2 domain as it moves against the helix α11 and the interlobe connecting peptide.

Functional Implications. An unexpected feature of the structure of $ApoTf_N$ is the highly conserved nature of the structural transition between closed and open forms. Even though the comparison is between the isolated transferrin N-lobe and the N-lobe of the complete lactoferrin molecule, and the two proteins share no more than 60% sequence identity, the open forms are virtually identical and the nature and extent of the conformational change are the same. This implies that the conformational change reflects mechanisms inherent in the N-lobe structure that are probably conserved between different members of the transferrin family and that are not significantly perturbed by whether the N-lobe is connected to the C-lobe or not. This gives confidence that binding studies carried out with N-lobe half-molecules are likely to be broadly relevant to the equivalent studies carried out on the intact (bilobal) molecules.

The close similarity of the conformational changes between the transferrin and lactoferrin N-lobes correlates with the similar location of the hinge, at a well-defined position in the two antiparallel β-strands, βE and βJ, behind the iron sites. The location of the hinge is critical for function because it splits the two ligands provided by these strands so that one (His 249) remains with the N1 domain and the other (Tyr 95) with the N2 domain.

The movement of the N2 domain displaces the C-terminus of the polypeptide, beyond the end of helix α11 (see Figure 2). We believe that this may point to a mechanism by which domain opening or closing in the N-lobe of transferrin can be signaled to the C-lobe. The displacement results from van der Waals interactions and from the presence of the Cys 137−Cys 331 disulfide bond and implies that the interlobe connecting peptide, to which Cys 331 belongs in full-length transferrin, is likely to be displaced by movement of the N2 domain. This, in turn, could affect the positioning of the C-lobe relative to the N-lobe and so provide a structural explanation for the interlobe communication detected by thermodynamic studies (*37*). A very similar communication of domain movements is seen in lactoferrin, where the movement of the N2 domain affects the orientation of the interlobe connecting helix and the relative orientations of the two lobes (*30*).

The differences in the orientations of the two iron-binding residues, Asp 63 and His 249, in $ApoTf_N$, compared with ApoLf, give important clues about the effects of pH on transferrin structure. The residues they interact with in Tf (Glu 83 and Lys 296) are conserved in Lf, as are most of the neighboring residues (including Ser 248 and Tyr 85, which hydrogen bonds with Glu 83); the polypeptide conformations are also virtually identical. We suggest that the differences result from the different pH of the two crystal structure analyses. The $ApoTf_N$ crystals were obtained at pH 5.3, whereas the ApoLf crystals grew at pH 8.0 (*36*). At the lower pH, His 249 should be protonated and this could account for the formation of the His 249···Glu 83 salt bridge and the change in the 248−249 peptide orientation; neither is seen in ApoLf. Likewise, the lower pH should favor the formation of the Asp 63···Lys 296 salt bridge, also not seen in apo-lactoferrin. Thus, the $ApoTf_N$ and ApoLf structures may represent low pH and high pH forms common to both Tf and Lf and the $ApoTf_N$ structure is the one adopted when iron is released at low pH in vivo.

The existence of these two salt bridges, involving Asp 63 and His 249 in $ApoTf_N$, has important implications for both iron binding and iron release. The initial site of iron binding is assumed to involve the two Tyr residues and carbonate ion on the N2 domain (*3*, *34*). Completion of the iron coordination, to Asp 63 and His 249, requires *both* domain movement *and* the breakage of these two salt bridges. Moreover, the Asp 63···Lys 296 salt bridge covers His 249, blocking it from iron coordination until the salt bridge is broken (Figure 4). Thus, we favor a binding mechanism in which either Asp 63 binds before His 249 or both bind together.

The crystal structures of $ApoTf_N$ (this work) and $FeTf_N$ (*14*) between them give direct evidence for two key protonation events in the mechanism of iron release by transferrins. Kinetic studies (*11*) suggest that iron release begins with protonation of the carbonate ion and subsequent decarbonation of the binding site, followed by a rate-determining step in which one of the protein ligands for iron is protonated. The first step is apparent in the crystal structure of $FeTf_N$, in which two alternative positions for the anion are seen, one of which is consistent with protonation of the anion and partial removal from iron coordination (*14*). The second step is identified by the present $ApoTf_N$ structure; the formation of the salt bridge between His 249 and Glu 83 indicates that the protein ligand that is protonated as the pH is reduced is His 249. Loss of the His ligand from iron coordination will not necessarily cause the immediate breakup of the iron-binding site but will severely weaken it. Evidence for this statement comes from the His253Met mutant of human lactoferrin where Met is not coordinated, the iron is 5-coordinate, and iron binding is destabilized (*38*).

Both crystallographic and kinetic studies thus converge on a mechanism of iron release in which (i) the carbonate is protonated and becomes detached from iron, and (ii) the histidine is protonated and dissociates from the iron to form a salt bridge with Glu 83. The possible role of the "dilysine trigger" in iron release may now be reevaluated. It seems unlikely that domain opening could be triggered by protonation of one of the lysine residues (Lys 206 and Lys 296 in human Tf) so long as the iron coordination remains intact to hold the domains together. This is because the structure of bovine lactoferrin (*39*) shows that alternative positions for a pair of lysine side chains in the holo form can exist. On the other hand, it seems certain that these residues do play some part. We suggest that the proposed "dilysine trigger" comes into play at a later stage, when iron binding is sufficiently weakened by loss of the anion and His ligands. This may also be the site at which secondary, or nonsynergistic, anions (*40*) bind to the N-lobe of human Tf and perturb its stability of iron binding (*2*). In both $FeTf_N$ and $ApoTf_N$, the side chains of Lys 206 and His 207 are close together and exposed to solvent, and binding of anions to this pair of residues could thus disrupt the Lys 206···Lys 296 interaction.

ACKNOWLEDGMENT

We thank Heather Baker and Steven Shewry for generous help and advice with crystallization and Clyde Smith for help with preparation of figures.

REFERENCES

1. Brock, J. H. (1985) in *Metalloproteins* (Harrison, P. M., Ed.) Part 2, pp 183−262, Macmillan, London.
2. Harris, D. C., and Aisen, P. (1989) in *Iron Carriers and Iron Proteins* (Loehr, T. M., Ed.) pp 241−351, VCH Publishers, New York.
3. Baker, E. N. (1993) *Perspect. Bioinorg. Chem. 2*, 161−205.
4. Klausner, R. D., Ashwell, J. V., Van Renswoude, J. B., Harford, J., and Bridges, K. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 2263−2267.
5. Bali, P. K., and Aisen, P. (1991) *Biochemistry 30*, 9947−9953.
6. Anderson, B. F., Baker, H. M., Dodson, E. J., Norris, G. E., Rumball, S. V., Waters, J. M., and Baker, E. N. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1769−1773.
7. Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W., and Baker, E. N. (1989) *J. Mol. Biol. 209*, 711−734.
8. Haridas, M., Anderson, B. F., and Baker, E. N. (1995) *Acta Crystallogr., Sect. D 51*, 629−646.
9. Bailey, S., Evans, R. W., Garratt, R. C., Gorinksy, B., Hasnain, S., Horsburgh, C., Jhoti, H., Lindley, P. F., Mydin, A., Sarra, R., and Watson, J. L. (1988) *Biochemistry 27*, 5804−5812.
10. Kurokawa, H., Mikami, B., and Hirose, M. (1995) *J. Mol. Biol. 254*, 196−207.
11. El Hage Chahine, J.-M., and Pakdaman, R. (1995) *Eur. J. Biochem. 230*, 1102−1110.
12. Kilar, F., and Simon, I (1985) *Biophys. J. 48*, 799−801.
13. Grossmann, J. G., Neu, M., Pantos, E., Schwab, F. J., Evans, R. W., Townes-Andrews, E., Lindley, P. F., Appel, H., Thies, W.-G., and Hasnain, S. S. (1992) *J. Mol. Biol. 255*, 811−819.
14. MacGillivray, R. T. A., Moore, S. A., Chen, J., Anderson, B. F., Baker, H. M., Luo, Y., Bewley, M. C., Smith, C. A., Murphy, M. E. P., Wang, Y., Mason, A. B., Woodworth, R. C., Brayer, G. D., and Baker, E. N. (1998) *Biochemistry 37*, 7919−7928.
15. Dewan, J. C., Mikami, B., Hirose, M., and Sacchettini, J. C. (1993) *Biochemistry 32*, 11963−11968.
16. Anderson, B. F., Baker, H. M., Norris, G. E., Rumball, S. V., and Baker, E. N. (1990) *Nature (London) 344*, 787−790.
17. Gerstein, M., Anderson, B. F., Norris, G. E., Baker, E. N., Lesk, A. M., and Chothia, C. (1993) *J. Mol. Biol. 234*, 357−372.
18. Mason, A. B., Funk, W. D., MacGillivray, R. T. A., and Woodworth, R. C. (1991) *Protein Expression Purif. 2*, 214−220.
19. He, Q.-Y., Mason, A. B., Woodworth, R. C., Tam, B. M., Wadsworth, T., and MacGillivray, R. T. A. (1997) *Biochemistry 36*, 5522−5528.
20. Kingston, R. L., Baker, H. M., and Baker, E. N. (1994) *Acta Crystallogr., Sect. D 50*, 429−440.
21. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
22. Collaborative Computational Project No.4 (1994) *Acta Crystallogr., Sect. D 50*, 760−763.
23. Navaza, J., and Saludjian, P. (1997) *Methods Enzymol. 276*, 581−594.
24. Brunger, A. T. (1992) X-PLOR, Version 3.1. Yale University Press, New Haven, CT.
25. Brunger, A. T. (1992) *Nature (London) 335*, 472−474.
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
27. McDonald, I., and Thornton, J. (1994) *J. Mol. Biol*, *238*, 777−793.
28. Baker, E. N., and Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.
29. Matthews, B. W. (1972) *Macromolecules 5*, 818−819.
30. Jameson, G. B., Anderson, B. F., Norris, G. E., Thomas, D. H., and Baker, E. N. (1998) *Acta Crystallogr., Sect D* (In press).
31. Presta, L. G., and Rose, G. D. (1988) *Science 240*, 1632−1641.
32. Richardson, J. S., and Richardson, D. C. (1988) *Science 240*, 1648−1652.
33. Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) *J. Mol. Biol. 206*, 759−777.
34. Lindley, P. F., Bajaj, M., Evans, R. W., Garratt, R. C., Hasnain, S. S., Jhoti, H., Kuser, P., Neu, M., Patel, K., Sarra, R., Strange, R., and Walton, A. (1993) *Acta Crystallogr., Sect. D 49*, 292−304.
35. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
36. Norris, G. E., Baker, H. M., and Baker, E. N. (1989) *J. Mol. Biol. 209*, 329−331.
37. Lin, L.-N., Mason, A. B., Woodworth, R. C., and Brandts, J. F. (1994). *Biochemistry 33*, 1881−1888.
38. Nicholson, H., Anderson, B. F., Bland, T., Shewry, S. C., Tweedie, J. W., and Baker, E. N. (1997) *Biochemistry 36*, 341−346.
39. Moore, S. A., Anderson, B. F., Groom, C. R., Haridas, M., and Baker, E. N. (1997) *J. Mol. Biol. 274*, 222−236.
40. Folajtar, D. A., and Chasteen, N. D. (1982) *J. Am. Chem. Soc. 104*, 5775−5780.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
42. Merritt, E. A. and Murphy, M. E. P. (1994) *Acta Crystallogr., Sect. D 50*, 869−873.

BI9812064