The Crystal Structure of Iron-free Human Serum Transferrin Provides Insight into Inter-lobe Communication and Receptor Binding*\$\|\[\]

Jeremy Wally, Peter J. Halbrooks, Clemens Vonrhein, Mark A. Roudi, Stephen J. Everse, Anne B. Mason, and Susan K. Buchanan

From the NIDDK, National Institutes of Health, Bethesda, Maryland 20892, the Department of Biochemistry, University of Vermont, College of Medicine, Burlington, Vermont 05405, Global Phasing Ltd., Sheraton House, Castle Park, Cambridge CB3 0AX, United Kingdom, and the Department of Molecular Physiology and Biophysics, University of Vermont, College of Medicine, Burlington, Vermont 05405

Serum transferrin reversibly binds iron in each of two lobes and delivers it to cells by a receptor-mediated, pH-dependent process. The binding and release of iron result in a large conformational change in which two subdomains in each lobe close or open with a rigid twisting motion around a hinge. We report the structure of human serum transferrin (hTF) lacking iron (apo-hTF), which was independently determined by two methods: 1) the crystal structure of recombinant non-glycosylated apo-hTF was solved at 2.7-Å resolution using a multiple wavelength anomalous dispersion phasing strategy, by substituting the nine methionines in hTF with selenomethionine and 2) the structure of glycosylated apo-hTF (isolated from serum) was determined to a resolution of 2.7 Å by molecular replacement using the human apo-N-lobe and the rabbit holo-C1-subdomain as search models. These two crystal structures are essentially identical. They represent the first published model for full-length human transferrin and reveal that, in contrast to family members (human lactoferrin and hen ovotransferrin), both lobes are almost equally open: 59.4° and 49.5° rotations are required to open the N- and C-lobes, respectively (compared with closed pig TF). Availability of this structure is critical to a complete understanding of the metal binding properties of each lobe of hTF; the apo-hTF structure suggests that differences in the hinge regions of the N- and C-lobes may influence the rates of iron binding and release. In addition, we evaluate potential interactions between apo-hTF and the human transferrin receptor.

The transferrins are a family of bilobal iron-binding proteins that play the crucial role of binding ferric iron and keeping it in solution, thereby controlling the levels of this important metal in the body (1, 2). Human serum transferrin (hTF) is synthesized in the liver and secreted into the plasma; it acquires Fe(III) from the gut and delivers it to iron requiring cells by binding to specific transferrin receptors (TFR) on their surface. The entire hTF-TFR complex is taken up by receptor-mediated endocytosis culminating in iron release within the endosome (3). Essential to the re-utilization of hTF, iron-free hTF (apo-hTF) remains bound to the TFR at low pH. When the apo-hTF-TFR complex is returned to the cell surface, apo-hTF is released to acquire more iron.

Strong homologies exist, both between TF family members, and between the two lobes of any given TF (4, 5). Each N- and C-lobe is divided into two subdomains (designated N1 and N2, and C1 and C2) connected by a hinge that gives rise to a deep cleft containing the iron-binding ligands. Iron is coordinated by four highly conserved amino acid residues: an aspartic acid (the sole ligand from the N1- or C1-subdomain), a tyrosine in the hinge at the edge of the N2- or C2-subdomain, a second tyrosine within the N2- or C2-subdomain, and a histidine at the hinge bordering the N1- or C1-subdomain. In addition, the iron atom is bound by two oxygen atoms from the synergistic anion (carbonate), which is itself stabilized by a conserved arginine residue (6).

A key feature of iron binding and release by TF family members is the large conformational change involving not only opening of the two subdomains in each lobe but also a twist between the N1- and N2-, or C1- and C2-subdomains (7, 8). Although the N- and C-lobes of hTF share 56% sequence similarity, many studies show that the rate of iron release from the C-lobe is considerably slower than the rate of release from the N-lobe, particularly at the putative endosomal pH of ∼5.6 (9–15). At least some of the difference is attributed to the presence of a “dilysine trigger” in the N-lobe, which is replaced by a

* This work was supported in part by U.S. Public Health Services Grant R01 DK21739 to A. B. M. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
§ The atomic coordinates and structure factors (code 2HAV and 2HAU) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).
| This on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S4 and Table S1.
| These authors contributed equally to this work.
| Supported by the Intramural Research Program of NIDDK, National Institutes of Health (NIH).
| To whom correspondence should be addressed: Laboratory of Molecular Biology, NIDDK, NIH, 50 South Drive, Rm. 4503, Bethesda, MD 20892-8030. Tel.: 301-594-9222; Fax: 301-480-0597; E-mail: skbuchan@helix.nih.gov.

The abbreviations used are: hTF, human serum transferrin; hTF-NG, recombinant non-glycosylated human serum transferrin; hTF-Gly, commercially available glycosylated human serum transferrin; TF, transferrin; apo-TF, transferrin lacking iron; oTF, ovotransferrin; LTF, lactoferrin; TFR, transferrin receptor; cryo-EM, cryo-electron microscopy; SeMet, selenomethionine; NCS, non-crystallographic symmetry; r.m.s., root mean square.
The dilsyne trigger is composed of Lys206 in the N2-subdomain and Lys296 in the N1-subdomain, which reside on opposite sides of the iron binding cleft and are oriented with side chains extending toward one another allowing them to share a hydrogen bond in the iron-bound (closed) conformation (16, 17). In the C-lobe, a triad of residues (Lys534 in the C2-subdomain, Arg632 and Asp634 in the C1-subdomain) replaces the lysine pair (14, 18).

The release of iron from hTF depends upon a number of factors, including pH, a chelator (physiologically relevant chelators include citrate, pyrophosphate, and ATP), and ionic strength, as well as the specific TFR (19–21). Raymond et al. (22) suggest that a complete model must explain the differences in the rate of iron release from the two lobes, the observed variable chelator concentration dependence, and the effect of anions, as well as the presence or absence of cooperativity between the lobes.

A 7.5-Å cryoelectron microscopy (cryo-EM) model of diferric hTF bound to TFR was created by docking a human TF N-lobe structure and a rabbit TF C-lobe structure into the electron density map of the complex (because there is no full-length human TF structure available) (23). This model offers a preliminary view of the regions of hTF and TFR that interact; it is suggested that both the N1- and N2-subdomains of the hTF N-lobe contact the TFR, whereas only the C1-subdomain of the hTF C-lobe appears to be involved in the interaction. Interestingly, a translation of ~9 Å of the ferric N-lobe (relative to the ferric C-lobe) is required to dock the two lobes into the cryo-EM density. Of relevance, at pH 7.4, the TFR discriminates between diferric, the two monoferric species, and apo-hTF, although the basis of this discrimination has not been explained (24–26). Significantly, our studies with authentic monoferric hTF constructs established that each lobe contributes equally (and non-additively) to the binding energy of this interaction with the TFR (15). Clearly a structure of apo-hTF is required to determine whether a change in orientation of the two lobes could provide both a rationale for discrimination and further insight into the receptor interaction.

Here we report the structure of full-length apo-hTF that has been independently determined by two methods; both a nonglycosylated recombinant form of hTF (pH 6.5) and a glycosylated native form of hTF (pH 7.0) were solved to a resolution of 2.7 Å. These two structures, which are identical within the limits of the resolution, find both the N- and C-lobes in the open conformation. This work represents the first mammalian TF structure with an apo-C-lobe, the first published structure of full-length hTF, and the first report of a baby hamster kidney expression system to substitute the methionine residues in hTF with selenomethionine (SeMet). The apo-hTF structure allows comparisons to other relevant structures, including those for diferric pig (2.15 Å, 72% identical) and rabbit TF (2.6 Å, 79% identical) (27), and an unpublished model for an unrefined monoferric hTF with iron in the C-lobe (3.3 Å) (28).

**EXPERIMENTAL PROCEDURES**

**Production of hTF-NG**—To produce recombinant non-glycosylated hTF (hTF-NG) with SeMet substituted for methionine, baby hamster kidney cells transfected with the pNUT plasmid containing the sequence of the N-His-tagged hTF-NG were placed into four expanded surface roller bottles (13). Addition of culture media containing SeMet results in a significant deterioration of the cells within 24–48 h. To maximize the incorporation, medium containing SeMet was added when production of hTF was at a maximum as determined by a competitive immunoassay (29). Briefly, adherent baby hamster kidney cells were grown in Dulbecco’s modified Eagle’s medium/F-12 containing 10% fetal bovine serum. This medium was changed twice at 2-day intervals, followed by addition of Dulbecco’s modified Eagle’s medium/F-12 containing 1% Ultroser G and 1 mm butyric acid. After one or two changes in this medium, 200 ml of SeMet containing Dulbecco’s modified Eagle’s medium/F-12 (lacking normal methionine), with butyric acid and Ultroser G was added to each roller bottle. Following a 4-h wash-in period this media was discarded and replaced with 250 ml of the same medium for an additional 48 h of incubation. The recombinant hTF-NG was purified from the medium as described in detail (30). In two production runs between 8 and 16 mg of SeMet containing N-His hTF-NG was produced, of which approximately half was recovered. Electrospray mass spectrometry analysis indicated a mass consistent with incorporation of 8–9 SeMet residues (data not shown).

**Purification of Apo-hTF-Gly**—In independent experiments, lyophilized human serum TF lacking iron (apo-hTF-Gly) was obtained from Sigma-Aldrich and reconstituted in 50 mM Tris-HCl, pH 8.0, and 20 mM sodium carbonate at a protein concentration of 5 mg/ml. The protein was applied to a 10-ml Q-Sepharose High Performance column (GE HealthCare) equilibrated with 50 mM Tris-HCl, pH 8.0, and 20 mM sodium carbonate. The apo-hTF was eluted using a linear gradient from 0 to 150 mM NaCl over 3 column volumes. Peak fractions were pooled and dialyzed overnight into 20 mM Tris-HCl, pH 8.0, 20 mM sodium carbonate, and 200 mM sodium chloride.

**Crystallization**—Recombinant diferric hTF-NG (with or without a His tag) at a concentration of 15 mg/ml in 0.1 M ammonium bicarbonate, was mixed with an equal volume of reservoir solution composed of 0.3 M ammonium citrate (pH 6.5) and 16–18% PEG 3350 at 20 °C. The SeMet N-His hTF-NG required slightly lower levels of PEG 3350 and streak seeding using the non-SeMet labeled hTF-NG crystals. Clear crystals (0.2 mm × 0.3 mm × 0.4 mm) formed in 3–5 weeks. Crystals of the SeMet-labeled protein were essentially isomorphous with those of the native protein, showing similar cell dimensions and crystallizing in the orthorhombic space group P2₁2₁2₁ (Table 1). All crystals were cryoprotected by addition to the hanging drop of 0.5 μl of a solution of 25% PEG 3350/30% ethylene glycol.

Native apo-hTF-Gly was concentrated in a Centriprep 30 concentrator (Millipore) to 30 mg/ml and screened against commercially available 96-condition screens using a Mosquito crystallization robot (TTP Labtech) with a hanging drop format (drop size, 200 nl of protein plus 200-nl well solution). Conditions yielding the best crystals were then further refined using 24-well VDX plates (Hampton Research). The crystals used for native data collection were obtained from hanging drops with a well solution of 0.2 M ammonium citrate, pH 7.0, 20% PEG 3350, and 15% glycerol, incubated at 21 °C. Drop sizes varied
from 2 to 16 µl and consisted of equal parts of protein and well solution. Crystals grew in ~24 h and were flash frozen in propane cooled to ~170 °C.

Data Collection and Refinement—For the apo-hTF-NG crystals that contained SeMet, multiple wavelength anomalous dispersion data were collected at ~170 °C on beamlines X26C (peak and remote) and X25 (inflection) at Brookhaven National Laboratory. The data were reduced, scaled, and merged using DENZO/SCALEPACK (31) (Table 1). To find the selenium sites, multiple wavelength anomalous dispersion data were prepared with XPREP and analyzed with ShelxD (32). The data sets were combined, and refinement of the selenium sites was carried out using autoSHARP (33, 34). Profess (35) was used to find NCS, which revealed two molecules in the asymmetric unit. Following a round of density modification, the structure of apo-hTF N-lobe (Protein Data Bank (PDB) 1BP5) was used as a search model for a phased translation and rotation function using MOLREP (36). Subsequent model building of the C-lobe was done using O (37) in a stepwise manner by incorporating fragments of pig holo-TF (converted to the human sequence). Iterative rounds of density modification and phase recombination were performed with SOLOMON (38). Refinement was accomplished using both CNS (39) and REFMAC (35, 40).

Native diffraction data from an apo-hTF-Gly crystal were collected at the Advanced Photon Source on SER-CAT beamline 22ID at ~170 °C. The crystal belonged to the orthorhombic space group P2B1B2B1B2B1B1B, with two molecules in the asymmetric unit and a solvent content of 59.3%. The images were reduced, scaled, and merged using HKL2000 (31) (Table 1), and the structure was solved using the molecular replacement program Phaser (41). A search model consisting of the human apo-N-lobe structure (1BP5) and the rabbit holo-C1-subdomain (residues 342–424 and 579–676 from 1JNF) was used, leading to a single solution containing two copies of each component. After a single round of rigid body refinement using the program REFMAC (35, 40) the rabbit C2-subdomain was fit into the electron density. The C1- and C2-subdomains were then

| Table 1: Data collection, crystallographic refinement, and model statistics | Native 1 | Native 2 | Inflection | Peak | Remote |
| --- | --- | --- | --- | --- | --- |
| Protein | hTF-Gly | hTF-NG | hTF-NG | hTF-NG | hTF-NG |
| Cell parameters | | | | | |
| Space group | P2,2,2 | P2,2,2 | P2,2,2 | P2,2,2 | P2,2,2 |
| a, b, c (Å) | 88.32 | 88.99 | 88.49 | 88.28 | 88.47 |
| 103.26 | 102.16 | 103.35 | 103.39 | 103.55 |
| 200.36 | 197.04 | 199.17 | 198.94 | 198.38 |
| Data collection statistics | | | | | |
| Location | APS | Chess A1 | BNL | BNL | BNL |
| Wavelength (Å) | 1.0000 | 0.935, 1.100 | 0.9794 | 0.9790 | 0.9641 |
| Resolution range (Å) | 30–2.7 | 50–2.7 | 50–2.9 | 25–3.2 | 50–3.3 |
| Unique reflections | 49,871 | 47,968 | 37,498 | 31,124 | 28,343 |
| Completeness (%) | 96.7 (90.5) | 95.9 (83.6) | 99.9 (98.9) | 99.4 (93.8) | 100 (99.9) |
| Redundance | 6.6 (6.5) | 6.0 (4.5) | 6.0 (5.3) | 7.1 (6.5) | 7.3 (7.5) |
| Rmerge (%) | 11.6 | 11.6 | 11.6 | 11.6 | 11.6 |
| Detector | Mar300 | Q210 Q4 | Q315 | Q4 | Q4 |
| Integration software | HKL2000 | HKL v1.98.2 | HKL2000 | HKL2000 | HKL2000 |
| Molecular replacement | Search models | 1A8E (N-lobe) | 1H76 (C-lobe) | | |
| MAD phases | Figure of merit (FOM) | 100–2.9 Å | 0.41 | 0.38 | 0.80 |
| Acentrics | 1,352 | 4 | | |
| Centrics | 7 | 2 | | |
| Density-mod. mean FOM | 10.486 | 10.488 | 64 | 4 | 2 |
| Model statistics | Resolution (Å) | 15–2.7 | 15–2.7 | | |
| Rwork | 23.2 | 27.3 | 4 | 3 | 5 |
| Rfree | 29.3 | 32.3 | 5 | 5 | 5 |
| Refined model | No. amino acid residues | 1,352 | 1,352 | | |
| No. citrate molecules | 7 | 4 | | |
| No. glycerol molecules | 2 | 2 | | |
| No. atoms, non-hydrogen | Protein | 10.486 | 10.488 | | |
| Ligand | 115 | 64 | | |
| r.m.s. deviation, bonds (Å) | 0.009 | 0.009 | | |
| r.m.s. deviation, angles (°) | 1.57 | 1.33 | | |
| Ramachandran plot (%) | Most favored regions | 77.0 | 78.2 | | |
| Additionally allowed | 20.6 | 19.4 | | |
| Generously allowed | 1.9 | 1.9 | | |
| Disallowed regions | 0.5 | 0.5 | | |
| Average B factor (Å²) | 75.2 | 67.4 | | |
| Range of B factors (Å²) | 32–125 | 14–135 | | |

*a Rmerge = Σ|I(hk) − (1/n)Σ|I(hk)|, where |I(hk)| is the jth measurement of diffraction intensity of reflection h, and (1/n)Σ|I(hk)| is the average intensity of reflection h for all j measurements.

*b Rwork = Σ|Fobs| – |Fc|) / Σ|Fobs|.

*c Rfree is calculated using a test set of 5% of the reflection excluded from refinement.
RESULTS AND DISCUSSION

Comparison of (Recombinant) Apo-hTF-NG and (Native) Apo-hTF-Gly—Both protein preparations crystallized in the orthorhombic space group P2₁2₁2₁, having almost identical unit cell dimensions and two molecules per asymmetric unit (Table 1). The recombinant and native structures had an r.m.s. deviation of only 0.73 Å for 1352 equivalent CP positions, and therefore we are, for the most part, presenting and discussing them as a single structure (apo-hTF). As in all other serum TF structures, no electron density for residues 1–3 is visible. Even the His-tagged recombinant apo-hTF-NG, with 14 extra amino acids at the N terminus, showed no density in this region, implying that the amino terminus is very flexible. Additionally, no density is found in the vicinity of Asn₄₁₃ and Asn₆₁₁ in the native apo-hTF-Gly, indicating that the glycan moieties are unstructured, but visible in our model (although missing in the diferric pig TF structure (27)). Each of the lobes is further separated into two subdomains: the N1- (1–92 and 247–331) and C1- (339–425 and 573–679) subdomains are each composed of a single region of continuous polypeptide (Fig. 1 and Supplemental Fig. S1). Though kinetically distinct, the fold of the C-lobe is equivalent to the fold of other N- and C-lobes (Table 2).

Quality of Final Model—Data collection, refinement, and model statistics are summarized in Table 1. No breaks in the main-chain density were observed in apo-hTF and, as shown in Table 1, the geometry is good. The Ramachandran plot of the main-chain torsion angles shows that 97.6% of the residues lie in most favored or allowed regions and that only 6 of the 1352 residues of the two non-crystallographic symmetry-related molecules reside in the disallowed region. Four of these residues are Leu⁵⁹⁴ and Leu⁶³⁰ in each of the molecules, comprising the central residues in classic γ-turns (Leu-Leu-Phe) with phi and psi angles of ~77° and ~46°, respectively (44). This structural feature was first noted in the hTF N-lobe (45) and subsequently has been observed in each lobe of all mammalian and avian TF molecules and in the C-lobe of all fish TFs (5, 6). Interestingly, the Leu-Leu-Phe sequence is not conserved in any insect TF N-lobe and is only partially conserved in the insect C-lobe sequences (5, 6). However, because the function of insect TF remains unknown, the importance of this finding is unclear. Although the role of the γ-turn in mammalian TFs has not been definitively established, we believe that it is significant that it immediately precedes the dilysine trigger residue Lys⁴⁹⁶ in the N-lobe and triad residue Arg⁶³² in the C-lobe (45). Within this context, the γ-turn may help to stabilize the orientation of these two residues to provide better repulsion or a triggering action, aiding in opening of the cleft.

Overall Organization—As with all TF family members, the structure of hTF illustrates the bilobal nature of the molecule (Fig. 1) with an amino-terminal lobe (N-lobe, residues 1–331) and a carboxyl-terminal lobe (C-lobe, residues 339–679). The lobes are connected by a linker peptide (residues 332–338) that is unstructured, but visible in our model (although missing in the diferric pig TF structure (27)). Each of the lobes is further separated into two subdomains: the N1- (1–92 and 247–331) and C1- (339–425 and 573–679) subdomains are each composed of two discontinuous sections of the polypeptide chain, whereas the N2- (93–246) and C2- (426–572) subdomains are each composed of a single region of continuous polypeptide (Fig. 1 and Supplemental Fig. S1). Though kinetically distinct, the fold of the C-lobe is equivalent to the fold of other N- and C-lobes (Table 2).

When we superimpose our apo-hTF structure on the diferric pig TF structure, the N1- and C1-subdomains align closely (r.m.s. deviation = 1.18 Å) (Fig. 2 and Table 3). This finding suggests that the twist and opening of each lobe when iron is released are restricted to the N2- and C2-subdomains, which clearly have implications with regard to how the TFR discriminates between diferric, monoferric, and apo-conformations.

B Factors—The regions with the lowest B factors include core regions in the N1- and N2-subdomains as well as regions com-
Iron-free Human Transferrin Structure

TABLE 2
Comparison of apo-hTF secondary structure to holo-pig TF and holo-chicken oTF

| Nomenclature | hTF \(^a\) | pTF \(^b\) | oTF \(^c\) | hTF | pTF | oTF |
|--------------|----------|--------|--------|------|------|------|
| 1            | 13-25    | 12-29  | 13-28  | 349-362 | 353-366 | 351-365 |
| 2            | 45-53    | 43-53  | 41-53  | 374-383 | 378-387 | 376-388 |
| 3            | 64-71    | 62-70  | 60-68  | 393-401 | 396-406 | 395-405 |
| 3a           | 418-420  | 410-420 | 410-420 | 418-420 | 410-420 | 410-420 |
| 4            | 109-111  | 108-110 | 105-109 | 441-443 | 444-448 | 444-448 |
| 5            | 129-135  | 127-134 | 121-136 | 461-466 | 463-478 | 460-476 |
| 5a           | 137-138  | 135-137 |         | 468-470 |        |       |
| 6            | 146-153  | 145-152 | 147-156 | 476-478 | 481-486 | 480-483 |
| 6a           | 168-170  | 166-169 | 167-174 | 502-504 | 499-501 | 496-499 |
| 6b           | 175-177  |         |         |        |       |       |
| 7            | 187-196  | 191-201 | 189-201 | 516-526 | 525-536 | 523-534 |
| 7a           |          |         |         |        |       |       |
| 8            | 210-213  | 212-217 | 211-217 | 549-551 | 545-550 | 543-549 |
| 8a           | 217-220  | 223-225 | 220-223 | 556-557 | 565-567 | 565-567 |
| 8a1          |          |         |         |        |       |       |
| 8b           | 235-240  | 239-244 |         | 572-576 | 580-585 | 580-585 |
| 9            | 260-273  | 263-278 | 259-275 | 594-608 | 600-617 | 600-616 |
| 10           | 311-315  | 314-319 | 315-321 | 647-651 | 655-659 | 652-658 |
| 11           | 317-328  | 321-331 | 321-332 | 653-662 | 661-673 | 658-669 |
| 12           | 669-677  | 677-682 | 674-686 | 637-640 | 645-649 | 642-647 |

\(a\) The apo-hTF secondary structural prediction was determined using PROCHECK (84).
\(b\) From Ref. 27.
\(c\) From Ref. 7. The N2- and C2-subdomains in each lobe are highlighted in aqua.

This finding is consistent with an electrospray ionization mass spectrometry study indicating that the apo-N-lobe is significantly less flexible than the apo-C-lobe (43).

Hinge—The extent of the movement of the subdomains can be estimated by superimposing the N1- or C1-subdomains and determining the rotation and translation functions required to

prising residues 343–359, 368–407, 627–632, and 667–674 of the C1-subdomain (Supplemental Fig. S2). The regions with the highest B factors include a loop in the beginning of the N1-subdomain (residues 29–34), the hinge between the N- and C-lobes (residues 332–338), residues 415–420 and 611–618 in the C1-subdomain, and the majority of the C2-subdomain.
mediate between the hTF and LTF structures. Nevertheless, opening for the apo-oTF N-lobe (8); thus chicken oTF is inter-

cleft in the N-lobe opens 53.4°, whereas the C-lobe is closed

Table 3

| Pig       | Apo-hTF | N1-  | N2-  | C1-  | C1-  |
|-----------|---------|------|------|------|------|
| N1-       | 1.175 Å (371 aa) |      |      |      |      |
| N2-       | 1.006 Å (162 aa)  |      |      |      |      |
| C1-       | 1.308 Å (151 aa)  | 1.175 Å (371 aa) |      |      |
| C2-       | 1.097 Å (139 aa)  | 1.033 Å (171 aa) |      |      |

superimpose the N2- or C2-subdomains on the respective N2- or C2-subdomains from an open or closed molecule (46). By this analysis an average rotation of 59.4° has occurred in the apo-hTF N-lobe relative to the diferric pig structure, similar to the average 63.2° rotation (range of 62.9–64.8° for the four molecules in the asymmetric unit) reported for the isolated apo-hTF-N-lobe structure (46). In the case of apo-hTF C-lobe, a subdomain rotation of 51.8° has occurred relative to the closed monoferric C-lobe (28), and a 49.5° rotation is required relative to the diferric pig TF structure (27) (see supplemental Fig. S3).

In contrast to other family members, the N- and C-lobes of apo-hTF are open to almost the same degree. For example, in the human apo-lactoferrin (LTF) structure the iron-binding cleft in the N-lobe opens 53.4°, whereas the C-lobe is closed (similar to the diferric structure but lacking iron) (47). Two alternative crystal forms of LTF have been reported by Baker et al. (11) featuring C-lobes open by 18° or 27°. The chicken apo-ovotransferrin (oTF) C-lobe opens 35° compared with a 53° opening for the apo-oTF N-lobe (8); thus chicken oTF is intermediate between the hTF and LTF structures. Nevertheless, duck oTF and camel LTF have equivalent open N- and C-lobes, 51.6 and 49.9° and 57.8 and 57.3°, respectively, so the patterns are not consistent or predictable (48, 49).

Jeffrey et al. (46) delineated the hinge residues in the N-lobe of hTF as Thr93, Val246, and Pro247 due to the ~36° psi changes between the ferric and apo-forms of these residues. As expected from the close structural alignment of our N-lobe with the isolated apo-N-lobe, we observe similar changes in the same residues when compared with the iron-containing pig TF N-lobe. In contrast in the C-lobe the hinge appears larger, comprised of Ala522 and Gly525 preceding strand e, and Arg581, Ala582, and Pro583, which begin strand j (see supplemental Fig. S1 for helix and strand designations). On average these residues change ~36° in psi between apo-hTF and the diferric pig TF structure. Inspection of the secondary structure of the hinge regions of the N-lobe and C-lobe of apo-hTF reveals a striking difference. In the N-lobe, the hinge lies adjacent to an anti-parallel β-sheet formed from strands e and j. However, in the C-lobe, strand e is shorter (starting three residues from the hinge), and strand j is entirely absent (Fig. 3). We suggest that this difference in the hinge regions of the N- and C-lobes may contribute to differences in overall flexibility between the lobes (43) and could therefore play an important role in the differing rates of iron binding and release.

In our apo-structure, crystal packing could influence opening of the cleft in the N-lobe, because the N-lobe of one molecule in the asymmetric unit intercalates into the cleft of the N-lobe of the other molecule (Fig. 2). In fact the side chain of Gln20 in the A molecule lies within hydrogen bonding distance of Cys177 and Cys179 in the B molecule. Additionally, the side chain of Glu13 in molecule A interacts with the side chain of Arg124 in molecule B. This is of interest, because Arg124 is the critical anion binding residue in the N-lobe of hTF, whose important role in iron uptake and release has been extensively studied (6, 19, 50–52). Nevertheless, we conclude, that crystal packing is not a factor given the similarity of the degree of opening of the isolated hTF N-lobe (63.2°) to the current structure (59.4°).

Metal and Anion Binding Sites—Seven molecules of citrate were found bound to apo-hTF-Gly, and four citrate molecules were found in apo-hTF-NG (Table 1). The presence and location of citrate in the structure are notable due to the physiological relevance of citrate as both an anion and a chelator (53–55). Six of the seven citrates bound to the two NCS-related molecules are found within the open clefts of the N- and C-lobes (see supplemental Table S1). As shown in Supplemental Fig. S4, one
Iron-free Human Transferrin Structure

FIGURE 3. A superimposition of the N-lobe and C-lobe of apo-hTF illustrating the difference in secondary structure of the region surrounding the hinge residues between the two subdomains of each lobe. In the N-lobe (N1, blue; N2, red), the hinge is adjacent to an anti-parallel \( \beta \)-sheet formed from \( \beta \)-strands e and j. In the C-lobe (C1, green; C2, yellow) the hinge is located within an unstructured region, as \( \beta \)-strand e is shortened and \( \beta \)-strand j is entirely absent. The overlay used SPDBV (89).

citrate molecule (9207) located within the N-lobe cleft of apo-hTF-Gly molecule A is hydrogen-bonded to the side chains of dilylsine trigger residues Lys\(^{206} \) and Lys\(^{296} \) as well as coordinating the side chain of Arg\(^{124} \) (normally bound to the synergistic carbonate anion). Because it is thought that binding of an anion sets up the site for the iron to be captured this is an interesting finding. Given that the concentration of citrate within serum is 100 \( \mu \)M (56), it is plausible that citrate could be the anion that binds to the hTF and prepares the site to receive the iron. However, in the N-lobe of apo-hTF-Gly molecule B, the citrate (9206) that lies in a similar position does not coordinate the dilylsine trigger residues, contacting only the backbone nitrogens of residues 125–127.

In the C-lobes of both molecules in the asymmetric unit of apo-hTF-Gly, two citrate molecules are present. In molecule A, one citrate (9202), coordinates the side chains of iron-binding ligand His\(^{385} \) and triad member Arg\(^{632} \), as well as the backbone nitrogen of Thr\(^{457} \). The corresponding citrate (9204) in molecule B contacts Arg\(^{632} \) and Thr\(^{457} \). The second citrate (9203 in A and 9205 in B) in both molecules contacts the backbone nitrogens of the ligand Tyr\(^{517} \) as well as Thr\(^{518} \). The differences in the positions of the citrates in each molecule emphasize the flexible nature of the clefts within the two lobes.

The seventh citrate interacts with the backbone nitrogens of two His residues (349 and 350) at the beginning of a \( \alpha \) helix 1 (residues 349–362) of the C-lobe in molecule B with the same residues in a symmetry-related molecule, facilitating crystal packing. This feature of the structure is especially interesting because of the evidence that these two histidines, particularly His\(^{349} \), are important in interacting with the TFR to bring about the release of iron from the C-lobe at endosomal pH (57). Notably, only citrates equivalent to 9201, 9202, 9204, and 9207 are observed in apo-hTF-NG.

**TABLE 4**

**Surface area (Å\(^2\)) of TFs sequestered by the lobe/lobe interface**

|        | Apo | Monoferric* | Diferric |
|--------|-----|------------|---------|
| LTF    | 660 (human 1CB6) \(^a\) | 454 (human 1B0L) \(^a\) | 8800 (human 1B0L) |
|        | 567 (camel 1DTZ) | 253 (human 1LFG) | 5145 (human 1LFG) |
| oTF    | 496 (chicken 1AIU) | 430 (duck 1DOT) | 8997 (duck 1DOT) |
|        | 552 (human) | 379 (pig 1H76) | 524 (rabbit 1JNF) |
|        | 532 (human) | 524 (rabbit 1JNF) | 524 (rabbit 1JNF) |

\(^*\) Structure from Ref. 28.
\(^a\) Delineates PDB accession numbers.
\(^b\) From the structure of recombinant human LTF.

Triad—Lys\(^{534} \), Arg\(^{632} \), and Asp\(^{634} \) have been implicated in the pH-dependent conformational change in the C-lobe during iron release (14, 16), making their positions in the apo-C-lobe of great interest. Arg\(^{632} \) and Asp\(^{634} \) (in the C2-subdomain) are adjacent to one another on one side of the open C-lobe cleft. A salt bridge exists between the guanidinium group of Arg\(^{632} \) and the side chain carboxyl group of Asp\(^{634} \). Lys\(^{534} \) lies in the C1-subdomain \( \sim 10 \) Å across the cleft, a relationship highly reminiscent of that observed for the dilylsine pair in the N-lobe (46). Previously we demonstrated that Arg\(^{632} \) could serve as a kinetically significant anion binding site (18). We note that the guanidinium group of Arg\(^{632} \) is solvent-accessible in the apo-structure and is found bound to citrate, emphasizing both its reactivity and accessibility.

**Interactions between Lobes**—A cooperative effect of one lobe on the other lobe has been convincingly documented in a number of studies using a variety of techniques, which include NMR of hTF (58), absorption spectra of oTF (59), pH-dependent iron release of LTF (60), calorimetric studies of both hTF and oTF (61–63), electrospray ionization mass spectrometry studies of hTF (43), iron release from monoferric LTF (64), analysis by urea gels (22), and chemical relaxation studies (65–67). Significantly, a number of these studies specifically attribute cooperative effects between the two lobes of hTF to participation of helix 12 from the C-lobe (2, 68). However, this helix undergoes no discernable movement between the apo, monoferric, or diferric states of TF (using pig serum TF as a model for the diferric form). The absence of a significant change in the position of this helix between the apo, monoferric, or diferric states of TF (using pig serum TF as a model for the diferric form).

The absence of a significant change in the position of this helix between the conformations may simply reflect the absence of the TFR, which we believe may be critical in inducing such a change (see below) (69, 70).

Inspection of the interface between the lobes suggests a loose, predominantly hydrophobic interface in which only a small amount (550 Å\(^2\)) of the total accessible surface area (\( \sim 8800 \) Å\(^2\)) is buried (Table 4). This is similar to what has been observed in other TF structures (7, 8, 71–73). Although we assumed that, because the diferric structure is more compact than the apo-structure, it would sequester more surface area at the interface, the opposite trend prevails. Diferic TF family members tend to have smaller interfaces (\( \sim 421 \) Å) than the apo molecules (\( \sim 569 \) Å), although both are quite modest compared with the overall surface area of the molecule. In addition to the hydrophobic interface between the lobes, two residues from the N-lobe form salt bridges with residues in the C1-subdomain.
Asp240 at the end of /H9251-helix 8b in the N2-subdomain interacts with Arg678 adjacent to the C-terminal helix (helix 12) in the C1-subdomain, and Asp376 in the N2-subdomain to Arg308 in the C1-subdomain. The subdomains are colored as in Fig. 1, and the salt-bridging side chains are shown in stick representation. The inset shows the location of the inter-lobe contacts within the context of the complete apo-hTF structure.

**TABLE 5**

Residues in the hTF N-lobe that may interact with the human TFR compared to residues at equivalent positions in rabbit TF and pig TF

| Residues in hTF N-lobe | N1-subdomain | N2-subdomain |
|------------------------|--------------|--------------|
| Bind TFR               | HUMAN TF     | RABBIT TF    | PIG TF       |
|                        | Y          | G           | G           |
|                        | L          | L           | L           |
|                        | A          | T           | P           |
|                        | P          | P           | K           |
|                        |            |             |             |
| Do not Bind            | BOVINE TF   | CHICKEN oTF  | HUMAN LTF   |
|                        | G          | G           | G           |
|                        | L          | L           | L           |
|                        | K          | P           | P           |
|                        | P          | E           | E           |
|                        | S           | I           | E           |
|                        |           | G           | S           |
|                        |            |             |             |
| Bind TFR               | HUMAN TF     | RABBIT TF    | PIG TF       |
|                        | E          | E           | D           |
|                        | D          | N           | N           |
|                        | P          | P           | Q           |
|                        | Q          | P           | T           |
|                        |           |             |             |
| Do not Bind            | BOVINE TF   | CHICKEN oTF  | HUMAN LTF   |
|                        | D          | E           | R           |
|                        | N          | G           | Q           |
|                        | P          | S           | P           |
|                        | Q          | T           | R           |
|                        |           |             |             |

**FIGURE 4.** The interface between the two lobes of apo-hTF is shown to illustrate the inter-lobe contacts. Two pairs of salt-bridged side chains are present at the interface: Arg678 in the N1-subdomain to Asp376 in the C1-subdomain, and Asp240 in the N2-subdomain to Arg308 in the C1-subdomain. The subdomains are colored as in Fig. 1, and the salt-bridging side chains are shown in stick representation. The inset shows the location of the inter-lobe contacts within the context of the complete apo-hTF structure.

Iron-free Human Transferrin Structure
(residues 51–88) identified in the radiation footprinting studies (79) does not seem credible based on sequence comparisons. It is possible that the proximity to the iron binding ligands (Asp63 in the 51–88 sequence and Tyr95 in the 89–102 sequence) could make these regions particularly susceptible to radiation damage; binding to the TFR might provide a shielding effect and explain the susceptibility differences.

In general there is better agreement between the regions identified by various techniques in the C-lobe of hTF. A monoclonal antibody specific to a sequential epitope in the C-lobe of hTF blocks binding of hTF to the TFR on HeLa S3 cells (80). Interestingly, this same monoclonal antibody binds to diferric hTF with 2-fold higher affinity compared with apo-hTF. The epitope was mapped to the C1-subdomain of hTF (residues Lys365–Ile378). As shown in Table 6, 11 residues in the C1-subdomain that are in contact with the helical region of the TFR were identified in the cryo-EM model of hTF bound to TFR (23). These residues lie between His349 and Glu372, and include Glu367, Glu369 (Val in the human sequence), Ser370, and Glu372, all of which are part of the epitope recognized by the monoclonal antibody. Thus, the mapping study confirms the identification of this particular region of the C-lobe in binding to the TFR and because the affinity of the monoclonal antibody changes as a function of iron status, the results are consistent with a conformational change in this region of the C-lobe required for His349 to come into contact with the TFR. The only residue with a significantly different conformation in our apo-structure compared with the unrefined monoferric hTF is Arg352, whose side chain is oriented toward the iron-binding cleft, rather than away (28). This could be very important due to proximity of this residue to His349 (57). Radiation footprinting identified a region in the C2-subdomain (residues 457–470) that differed in susceptibility to radiation damage. Because the C2-subdomain makes no contact with the TFR in the model, it is difficult to ascribe the effect to TFR interaction.

Understanding the basis by which the TFR discriminates between various family members and between the four conformers of hTF is a longstanding pursuit in TF research. The cryo-EM model (23) was constructed with human TFR, human N-lobe TF, and rabbit C-lobe TF. As described above, placement of the C1-subdomain is convincingly supported by all of the available data. We therefore superimposed the C1-subdomains of apo- and/or iron-containing chicken oTF, human LTF, human TF, rabbit TF, and pig TF (Table 4). Two observations can be made: first, all of the sequences (oTF, LTF rabbit, and pig TF) have either insertions or deletions in the loop formed by human TF Cys495 and Cys506, which lie near the carboxyl terminus of the TFR. Second, the linker peptide, which connects the N- and C-lobes, closely approaches the TFR in the cryo-EM model. In LTF this region adopts a helical conformation, which could constrain the independent movement of the lobes relative to each other and potentially restrict access to a TF receptor. In mammalian TFs the linker region is not structured, but there is a disulfide bond between Cys339 and Cys596 (human numbering), which is missing in oTF and LTF. This feature would be predicted to constrain movement between the N- and C-lobes. In addition, although the unstructured linker of hTF is probably long enough to accommodate the 9-Å separation of the two lobes seen in the cryo-EM model, it would require disrupting two salt bridges formed between the N-lobe and C1-subdomain, as well as altering the regions of buried surface between the N- and C-lobes, seen in our apo-hTF structure. Perhaps the most logical explanation for the altered TF conformation observed in the cryo-EM structure is that there is a pH-induced conformational change in the TFR itself to accommodate TF. Support for this idea comes from the observation of conformational changes in the TFR when going from neutral to acidic pH (81). Additionally, changes in the relative orientations of the TFR subdomains in response to binding hemochromatosis protein (which competes with hTF for bind-
Iron-free Human Transferrin Structure

Acknowledgments—We thank Robert Sweet and the RapiData 2004 program for access to synchrotron radiation facilities. We also thank Lothar Esser (NCI, National Institutes of Health (NIH)), Gali Prag (NIDDK, NIH), and Sangho Lee (NIDDK, NIH) for critically reading the manuscript. Data were also collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. We want to thank Igor Kaltashov for providing the mass spectrometry data on the selenomethionine samples. The Mason and Buchanan laboratories shared equally in solving the structures and generating the manuscript.

REFERENCES

1. Harris, D. C., and Aisen, P. (1989) in Iron Carriers and Iron Proteins (Loehr, T. M., ed) pp. 241–351, VCH Publishers, Inc., New York.

2. Aisen, P., Leibman, A., and Zweier, J. (1978) J. Biol. Chem. 253, 1930–1937.

3. Klausner, R. D., Ashwell, G., van Renswoude, J., Harford, J. B., and Bridges, C. (1999) J. Mol. Biol. 281, 274–298.

4. Harris, D. C., and Aisen, P. (1989) in Comp. Biochem. Physiol. Physiol. B 140, 11–25.

5. Lamb, A. B., Perri, H., and Hallbrooks, P. J., and Mason, A. B. (2003) Comp. Biochem. Physiol. B 142, 129–141.

6. Kurokawa, H., Mikami, B., and Hirose, M. (1995) J. Mol. Biol. 254, 196–207.

7. Kurokawa, H., Dewan, J. C., Mikami, B., Sacchettini, J. C., and Hirose, M. (1999) J. Biol. Chem. 274, 28445–28452.

8. Princicat, J. V., and Zapolski, E. J. (1975) Nature 255, 87–88.

9. Baldwin, D. A., De Sousa, D. M. R., and Von Wandruszka, R. M. A. (1982) Biochim. Biophys. Acta 719, 140–146.

10. Baker, E. N., Baker, H. M., and Kidd, R. D. (2002) Biochem. Cell Biol. 80, 27–34.

11. J. Biol. Chem. 274, 28445–28452.

12. Kurokawa, H., Mikami, B., Hirose, M. (1995) J. Mol. Biol. 254, 196–207.

13. Mason, A. B., He, Q.-Y., Halbrooks, P. J., Everse, S. J., M Kumarov, D. R., Kaltashov, I. A., Smith, V. C., Hewitt, J., and MacGillivray, R. T. A. (2002) Biochemistry 41, 9448–9454.

14. Halbrooks, P. J., He, Q. Y., Briggs, S. K., Everse, S. J., Smith, V. C., MacGillivray, R. T. A., and Mason, A. B. (2003) Biochemistry 42, 3701–3707.

15. Mason, A. B., Halbrooks, P. J., James, N. G., Connolly, S. A., Larouche, J. R., Smith, V. C., MacGillivray, R. T. A., and Chasteen, N. D. (2005) Biochemistry 44, 8013–8021.

16. Dewan, J. C., Mikami, B., Hirose, M., and Sacchettini, J. C. (1993) Biochemistry 32, 11963–11968.

17. He, Q.-Y., Mason, A. B., Tam, B. M., MacGillivray, R. T. A., and Woodworth, R. C. (1999) Biochemistry 38, 9704–9711.

18. Halbrooks, P. J., Giannetti, A. M., Klein, J. S., Bjorkman, P. J., Larouche, J. R., Smith, V. C., MacGillivray, R. T. A., Everse, S. J., and Mason, A. B. (2005) Biochemistry 44, 15451–15460.

19. He, Q.-Y., and Mason, A. B. (2002) in Molecular and Cellular Iron Transport (Templeton, D. M., ed) pp. 95–123, Marcel Dekker, Inc., New York.

20. Baker, E. N. (1994) Adv. Inorg. Chem. 41, 389–463.

21. Zak, O., Tam, B., MacGillivray, R. T. A., and Aisen, P. (1997) Biochemistry 36, 11036–11043.

22. Hamilton, D. H., Turcot, I., Stintzi, A., and Raymond, K. N. (2004) J. Biol. Inorg. Chem. 9, 936–944.

23. Cheng, Y., Zak, O., Aisen, P., Harrison, S. C., and Witz, T. (2004) Cell 116, 565–576.

24. Young, S. P., Bomford, A., and Williams, R. (1984) Biochem. J. 219, 505–510.

25. Mason, A. B., He, Q.-Y., Tam, B. M., MacGillivray, R. T. A., and Woodworth, R. C. (1998) Biochem. J. 330, 35–40.

26. Evans, R. W., Crawley, J. B., Garratt, R. C., Grossman, J. G., Neu, M., Aitken, A., Patel, K. J., Melak, A., Wong, C., Singh, J., Bomford, A., and Hasnain, S. S. (1994) Biochemistry 33, 12512–12520.

27. Hall, D. R., Hadden, J. M., Leonard, G. A., Bailey, S., Neu, M., Winn, M., and Lindley, P. F. (2002) Acta Crystallogr. D. Biol. Crystallogr. 58, 70–80.

28. Zuccolla, H. J. (1993) The Crystal Structure of Monoferric Human Serum Transferrin, Georgia Institute of Technology, Atlanta, GA.

29. Mason, A. B., He, Q.-Y., Adams, T. E., Gometer, D. R., Kaltashov, I. A., Nguyen, V., and MacGillivray, R. T. A. (2001) Protein Exp. Purif. 23, 142–150.

30. Mason, A. B., Halbrooks, P. J., Larouche, J. R., Briggs, S. K., Moffett, M. L., Ramsey, J. E., Connolly, S. A., Smith, V. C., MacGillivray, R. T. A. (2004) Protein Exp. Purif. 36, 318–326.

31. Orwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326.

32. Schneider, T. R., and Sheldrick, G. M. (2002) Acta Crystallogr. D. Biol. Crystallogr. 58, 1772–1777.

33. La Fortelle, E., and Bricogne, G. (1997) Methods Enzymol. 276, 472–494.

34. Voirin, C., Blanc, E., Roversi, P., and Bricogne, G. (2006) in Crystallographic Methods (Doublie, S., ed) pp. 215–230, Humana Press, Totowa, NJ.

35. Collaborative Computational Project, No. 4. (1994) Acta Crystallogr. D. Biol. Crystallogr. 50, 760–763.

36. Vagin, A. A., and Tsvet, V. V. (2003) Acta Crystallogr. D. Biol. Crystallogr. 59, 580–587.

37. Jones, T. A. (1978) J. Appl. Crystallogr. 11, 268–272.

38. Abrahams, J. P., and Leslie, A. G. (1996) Acta Crystallogr. D. Biol. Crystallogr. 52, 30–42.

39. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. D. Biol. Crystallogr. 54, 905–921.

40. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Acta Crystallogr. D. Biol. Crystallogr. 57, 122–133.

41. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Acta Crystallogr. D. Biol. Crystallogr. 61, 458–464.
