"Dilysine Trigger" in Transferrins Probed by Mutagenesis of Lactoferrin: Crystal Structures of the R210G, R210E, and R210L Mutants of Human Lactoferrin[†,‡]

Neil A. Peterson,[§] Vickery L. Arcus,[||] Bryan F. Anderson,[§] John W. Tweedie,[§] Geoffrey B. Jameson,[^] and Edward N. Baker*[,||]

Institute of Molecular Biosciences and Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand, and School of Biological Sciences, University of Auckland, Auckland, New Zealand

Received June 28, 2002; Revised Manuscript Received September 17, 2002

ABSTRACT: The mammalian iron-binding proteins lactoferrin (Lf) and transferrin (Tf) bind iron very tightly, but reversibly. Despite homologous structures and essentially identical iron binding sites, Tf begins to release iron at pH 6.0, whereas Lf retains iron to pH ∼3.5. This difference in iron retention gives the two proteins different biological roles. Two lysine residues, Lys 206 and Lys 296, which form a hydrogen-bonded dilysine pair in human Tf, have been shown to strongly influence iron release from the N-lobe. The equivalent residues in human Lf are Arg 210 and Lys 301, and we have here mutated Arg 210 in the N-lobe half-molecule of human lactoferrin, Lf_N, to probe its role in iron release. The Lf_N mutants R210G, R210E, and R210L were expressed, purified, and crystallized, and their crystal structures were determined and refined at resolutions of 1.95 Å (R210G), 2.2 Å (R210E), and 2.0 Å (R210L). The overall structures are very similar to that of wild-type Lf_N, but with small differences in domain orientations. In each of the mutants, however, Lys 301 (equivalent to Lys 296 in Tf) changes its conformation to fill the space occupied by Arg 210 Nη2 in wild-type Lf_N, interacting with the two tyrosine ligands Tyr 92 and Tyr 192. By comparison with other Lf and Tf structures, we conclude that Lys 301 (or Lys 296 in Tf) only occupies this site when residue 210 (206 in Tf) is nonpositive (neutral as in R210G and R210L or negative as in R210E). Thus, Lys 206 in the Tf dilysine pair is identified as having a depressed pK_a. Three specific sites are variably occupied by polar groups in the Lf mutants and other Lf and Tf proteins, and when coupled with iron-release data, these give new insights into the factors that most influence iron retention at low pH.

The transferrins are a family of proteins that are involved in regulating the levels of free iron in the body fluids of animals. The principal members of this family are lactoferrin (Lf), present in mammalian secretions, serum transferrin (Tf), which transports iron in the bloodstream, and ovotransferrin (oTf), which is the avian transferrin and is also present in high concentrations in egg white. These proteins [reviewed in (*1*−*3*)] are monomeric glycoproteins of ca. 80 kDa, which are capable of binding two ferric ions together with two carbonate anions per molecule. The proteins are folded into two globular lobes of highly similar structure, termed the N- and C-lobes (*3*). Each lobe is further divided into two domains, the N1 and N2 domains in the N-lobe and the C1 and C2 domains in the C-lobe. The iron and carbonate binding sites are located within a deep cleft that is formed between the domains of each lobe. Four of the ligands to the iron are provided by amino acid side chains, two tyrosines from one domain (N2 or C2) and one aspartate and one histidine from the other (N1 or C1). The remaining two coordination sites are occupied by oxygens from the synergistically bound carbonate anion. The ferric ion is bound very tightly ($K_d \sim 10^{-20}$ mol^{-1} L at pH 7.4), but can be released without denaturation of the proteins. The release of iron from transferrins is accompanied by substantial conformational changes (*4*). Comparisons of the iron-bound and apo transferrin structures have shown that that the loss of iron produces a 50°−60° rigid-body domain rotation that opens the binding cleft (*5*−*7*). This movement is facilitated by a hinge-type motion within the two antiparallel β-strands that form the primary link between the two domains of each lobe (*5*).

Amino acid sequence alignment indicates that the transferrins are a highly conserved family of proteins. The levels of sequence identity are approximately 60−70% between the Lfs and 50−60% between the Lfs and the mammalian serum Tfs (*3*). The crystal structures of human Lf (*8*), bovine Lf (*9*), rabbit serum Tf (*10*), and hen oTf (*11*) show that the

[†] This work was supported by grants from the U. S. Public Health Service (HD20859), the Marsden Fund of New Zealand, and the Health Research Council of New Zealand. E.N.B. also acknowledges research support as an International Research Scholar of the Howard Hughes Medical Institute.

[‡] Atomic coordinates have been deposited with the Protein Data Bank, with accession codes 1h43 (R210E), 1h44 (R210L) and 1h45 (R210G).

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

[§] Institute of Molecular Biosciences, Massey University.

[||] University of Auckland.

[^] Institute of Fundamental Sciences, Massey University

[1] Abbreviations: Lf, lactoferrin; Tf, transferrin; oTf, ovotransferrin; Tf_N, recombinant N-lobe half-molecule of human transferrin, residues 1−337; Lf_N, recombinant N-lobe half-molecule of human lactoferrin, residues 1−334; R210G, R210E, R210L, and R210K, mutants of Lf_N in which Arg210 is changed to Gly, Glu, Leu, and Lys respectively; MPD, 2-methyl-2,4-pentanediol; λ_{max}, wavelength of maximum absorption in the visible spectrum; pH_{50}, pH at which 50% of iron is lost; rms, root-mean-square.

10.1021/bi020443a CCC: $22.00
Published on Web 11/08/2002

proteins share essentially the same fold. Their iron and carbonate binding sites are also essentially the same; the ligands are strictly conserved in all Lfs, Tfs, and oTfs, except in the evolutionarily more distant insect Tfs, and the coordination geometry remains identical within the error levels of the current structural analyses.

Despite the close similarity of their three-dimensional structures, there is significant variation in the binding properties of the transferrins (*3*). For example, human Lf is estimated to bind iron between 50 and 90 times more strongly than human serum Tf at pH 7.4 (*12*). Human Lf also has the capacity to retain iron under more acidic conditions than human serum Tf; human serum Tf releases iron over the pH range 6.0−4.5, whereas human Lf does not lose its iron until the pH is reduced to below 3.5 (*3*, *13*). More subtle differences have been observed between the two sites of individual transferrins. The C-lobe of human serum Tf binds iron approximately 20-fold more strongly than the N-lobe site (*14*) and has a slower rate of iron release (*15*). The two binding sites of human serum Tf also have different pH dependence of iron release, with the N-lobe losing iron from ca. pH 6.0 and the C-lobe site beginning to release iron at ca. pH 5.0 (*16*).

The structural origins of these differences must lie outside the primary iron coordination sphere, since the same ligands are found in both sites of each protein. Several factors may contribute to the differences (*3*), such as the asymmetry of the molecules due to the "head to tail" packing of the two lobes, the variation in the distribution of disulfide bonds, differences in the local environment of the iron ligands, and cooperativity between the N- and C-lobe binding sites. One important factor is an unusual interaction between two lysine residues (*17*), which appears to modulate the pH-dependent release of iron from the N-lobe of the serum transferrins. The crystal structures of the N-lobes of human serum Tf (*18*) and hen oTf (*19*) indicate that the side chains of the two lysines (206 and 296 in Tf, 210 and 301 in oTf) are hydrogen-bonded. This interaction has been described as a "dilysine trigger" for iron release (*19*), because protonation at this site would result in the formation of two adjacent positive charges and the loss of the hydrogen bond between the lysines. Since Lys 206 and Lys 296 are located on opposite domains of the N-lobe, a repulsive interaction between them would destabilize the closed form of the protein and could induce the loss of iron. In contrast, human Lf does not contain a dilysine trigger interaction. Here, the corresponding residues are an arginine and a lysine, but although Arg 210 occupies a similar position in human Lf to that of Lys 206 in human Tf, the side chain of Lys 301 in Lf adopts a different conformation from that of Lys 296 in Tf, to form a salt bridge with Glu 216 from the other domain. This localized difference may contribute toward the ability of human Lf to retain iron to lower pH than Tf.

The situation is further complicated by the fact that bovine, equine, and buffalo Lf each contain lysines at the sequence positions corresponding to Lys 206 and Lys 296 in human serum Tf, yet their iron-binding properties are typical of lactoferrins. The crystal structures of these Lfs (*9*, *20*, *21*) show that the dilysine pair does not form. Instead, the side chain of the second lysine (Lys 301, human Lf numbering) adopts the same conformation as in human Lf, forming a salt bridge with Glu 216. This is consistent with their ability to retain iron to low pH and suggests that the presence or absence of the dilysine interaction is a characteristic feature of the N-lobes of the serum Tfs and the Lfs, respectively.

There is evidence that iron release from the N-lobe of Lf is influenced by interactions from the C-lobe (*13*). However, the crystal structure of the R210K mutant of the human Lf N-lobe (*22*) has shown that in the half-molecule, as in the full-length bovine, buffalo, and equine Lfs, the presence of lysine residues at the appropriate sequence positions does not lead to the formation of a dilysine interaction, even though the interaction is sterically feasible. This confirms that the failure to form a dilysine interaction in Lfs similar to that in Tfs must result from differences that are inherent to the Lf and Tf N-lobes.

The underlying basis for this key structural and functional difference between lactoferrins and transferrins has remained unclear. To further investigate the factors that modulate the interactions in the dilysine trigger region, we have mutated Arg 210 to Gly, Glu, and Leu in the recombinant N-lobe of human Lf (Lf_N). The crystal structures of these mutants are presented and compared with those of wild-type Tf_N and Lf_N and the R210K mutant of Lf_N. In contrast to R210K, the R210G, R210E, and R210L structures reveal that the Lys 301 side chain rotates into a conformation that is similar to that of Lys 296 in human serum Tf. These results imply that the position of the Lys 301 side chain is governed by the charge of the side chain of residue 210. By extension, we conclude that Lys 206 in Tf (but not Lys/Arg 210 in Lf) has a depressed pK_a, allowing it to form the dilysine interaction at physiological pH.

MATERIALS AND METHODS

Preparation of Recombinant Proteins. The R210G, R210E, and R210L mutations were introduced into the human lactoferrin cDNA and cloned into the mammalian expression vector pNUT-Lf_N using the same protocols as for preparation of the R210K mutant (*22*). The pNUT-Lf_N plasmids containing these mutations were then transfected into baby hamster kidney cells, and the recombinant lactoferrins were purified from the tissue culture medium by ion-exchange chromatography, as described previously (*13*).

Ultraviolet−Visible Spectroscopy. Ultraviolet−visible absorption spectra were recorded over the range 250−700 nm, using a Hewlett-Packard HP8452A diode array spectrophotometer. The spectra were measured from solutions of recombinant transferrins with protein concentrations that ranged between 3.6 and 8.2 mg/mL.

Iron Release. The pH dependence of iron release for each mutant was measured by dialysing the protein (concentration 2 mg/mL) against a series of buffers over the pH range 8.0−2.5, as described previously (*13*, *22*). The protein was maintained for 2 days at each pH and the percent saturation with iron was estimated from the peak absorbance of the visible charge-transfer band. In wild-type Lf_N, this band had a maximum absorbance at 454 nm, but the mutants all varied slightly and λ_{max} was found to change with pH; therefore, the percent saturation was estimated from [absorbance at visible λ_{max} for each sample]/[absorbance at λ_{max} at pH 8.0], normalized to a standard protein concentration from the absorbance at 280 nm.

Crystallization. Crystallization experiments were undertaken with iron-saturated proteins, from which the carbohy-

Table 1: Data Processing Statistics

	R210G	R210E	R210L
resolution	1.95 Å	2.2 Å	2.0 Å
outer resolution shell	2.02−1.95 Å	2.28−2.2 Å	2.07−2.0 Å
space group	*C*2	*C*2	*C*2
unit cell	a =131.7 Å	a = 112.9 Å	a = 125.1 Å
	b = 58.9 Å	b = 57.9 Å	b = 57.1 Å
	c = 57.5 Å	c = 57.8 Å	c = 57.6 Å
	β = 114.4°	β = 101.4°	β =117.2°
R_{merge}[a]	5.3% (14.6%)	5.8% (30.0%)	7.7% (29.4%)
I/σ[a]	13.6 (4.2)	15.3 (4.4)	11.9 (3.5)
no. of unique reflns	27 857	18 711	23 271
redundancy	2.8	3.4	2.8
completeness[a]	94.2% (60.4%)	98.7% (97.0%)	94.4% (89.0%)

[a] Figures in parentheses are for the outer resolution shell.

drate chain had been removed by the enzymes endoglycosidase F and peptide *N*-glycosidase F from *Flavobacterium meningosepticum* (*23*). The deglycosylated proteins were prepared for crystallization by concentration to 28 mg/mL in 20 mM HEPES, 1 M NaCl, pH 8.0. Crystals suitable for X-ray diffraction studies were grown by a batch method, by the dropwise addition of the protein into 1 mL of water (*22*). The crystallizations were then left to slowly equilibrate at 4 °C. Under these conditions, large needlelike red crystals grew over a period of 6 weeks.

Data Collection and Processing. X-ray diffraction data were measured by oscillation photography, using a Rigaku R-axis IIC imaging plate system with graphite-monochromated Cu Kα radiation (1.5418 Å) from a Rigaku RU200 rotating-anode generator. Data for R210G and R210L were collected from flash-frozen crystals, maintained at 113 K by a stream of nitrogen gas from an Oxford Cryosystem. These crystals were prepared for data collection by immersion in a cryoprotectant solution (crystallization buffer made up in 30% MPD), mounted in a small nylon loop and then rapidly transferred into the cold nitrogen gas stream. Since a suitable cryoprotectant was not found for the R210E crystals, data were collected at ambient temperature. Profile-fitted intensities were obtained using DENZO (*24*) and were scaled and merged using SCALEPACK (*24*). Data processing statistics are summarized in Table 1.

Structure Determination. The initial models for each of the R210 mutant structures were prepared from the coordinates for the recombinant N-lobe of human Lf (*25*). The iron, carbonate, and all water molecules were removed and Arg 210 was mutated to a glycine (for R210G) or to an alanine (for R210E and R210L). Several N- and C-terminal residues were also omitted from the models as these regions have relatively high-temperature factors in other human Lf N-lobe structures, indicating some disorder of the polypeptide chain.

Both the R210E and R210L structures were solved by molecular replacement, using residues 5−312 of the search model described above. The rotation and translation functions were calculated by the program AMoRe (*26*), using 95% of the data in the resolution range 8−4 Å, and in each case gave a single, clear solution. For R210E, the *R*-factor was 28.3% and the correlation coefficient 74.3%; for R210L, the corresponding values were 35.9% and 57.4%. Molecular replacement was not required to solve the R210G structure, as the crystals of this mutant were isomorphous with those of wild-type Lf_N. Consequently, residues 6−319 of the search model were placed directly into the R210G cell and used as the starting model for refinement.

Refinement. The R210G, R210E, and R210L mutant structures were refined using the programs XPLOR (*27*) and CNS (*28*). A randomly chosen subset of reflections (5−10% of the total data) was selected from each dataset for the calculation of the free *R*-factor (*29*), as a monitor of refinement. The starting models were subjected to an initial rigid-body refinement, in which the molecule was first treated as one fragment and then as two rigid bodies that represent the N1 and N2 domains. Starting rounds of rigid-body refinement, using the data in the resolution range 8−4 Å, were followed by further rounds in which the maximum resolution was extended to the resolution limit.

Electron density maps were calculated from the rigid-body refined models, and the Fe^{3+} and carbonate ions were built into the difference electron density. A number of atoms in the dilysine trigger region (residues 301−303 and the side chains of residues 210 and 216) were omitted from the initial refinements to reduce model bias. These models were then subjected to molecular dynamics refinement by XPLOR or CNS in order to remove model bias from difference electron density maps. The atoms that had been omitted from the refinements were rebuilt into difference electron density maps, and the Glu 210 and Leu 210 side chains were built into the R210E and R210L models, respectively. The models were then improved by several rounds of molecular dynamics and restrained least-squares refinement, followed by model rebuilding using TURBO FRODO (*30*). Ordered solvent molecules, all treated as water, were added at locations where there was spherical $|F_o| - |F_c|$ density of greater than 3σ above the mean and there were potential hydrogen-bonding contacts with the neighboring structure. The iron−ligand bonds were initially restrained to conform to standard bond distances, but these restraints were removed from the later rounds of refinement.

The refined R210 mutant models show good agreement with the observed X-ray data, as shown by the values for *R* and R_{free} (Table 2). The final models comprise residues 2−322 for R210G, 4−312 for R210E, and 4−325 for R210L. In each case, interpretable electron density is lacking for a few residues at the N- and C-termini, and these residues are not included in the models. The remaining residues are well-defined and are represented by continuous electron density in $2|F_o| - |F_c|$ maps when contoured at the 1σ level. The structures have good overall stereochemistry, with each having root-mean-square (rms) deviations of around 0.005 Å from standard bond lengths and 1.3° from standard bond angles. Ramachandran plots of the main chain torsion angles show that in each model Leu 299 is the only residue that is outside the allowed regions, as defined by PROCHECK (*31*). However, this residue has well-defined electron density and is the central residue of a γ-turn that is conserved in all transferrin structures (*9*).

RESULTS

Iron Binding and Release. A characteristic feature of the transferrin family is the red-brown color that develops upon iron binding. This color is due to the appearance of a ligand-to-metal charge-transfer band at ca. 465 nm, which is associated with the excitation of an electron from a π-orbital

Table 2: Structure Refinement Statistics

	R210G	R210E	R210L
resolution limits	30−1.95 Å	32−2.2 Å	32−2.0 Å
no. of reflections	26 435 (1351)	16 352 (1215)	23 270 (2179)
R factor (R_{free})	22.2% (24.7%)	20.4% (23.6%)	20.4% (24.0%)
model atoms			
protein residues	2−322	4−313	4−325
ions	Fe^{3+}, CO_3^{2-}	Fe^{3+}, CO_3^{2-}	Fe^{3+}, CO_3^{2-}
solvent molecules	92	76	212
av temp factors			
protein main chain atoms	30.1 Å^2	39.4 Å^2	23.1 Å^2
protein side chain atoms	31.4 Å^2	42.1 Å^2	24.5 Å^2
ordered solvent molecules	28.6 Å^2	41.2 Å^2	31.1 Å^2
rms deviation from ideal geometry			
bond distances	0.006 Å	0.006 Å	0.005 Å
bond angles	1.2°	1.3°	1.3°
dihedral angles	23.5°	22.9°	23.4°

Table 3: Iron Binding and Release by Lf_N and the R210 Mutants

protein	λ_{max} (nm)[a]	pH_{50}[b]	protein	λ_{max} (nm)[a]	pH_{50}[b]
Lf_N	454	4.3	R210L	462	5.5
R210G	468	4.4	R210K	472	5.0
R210E	476	4.3			

[a] Wavelength of maximal visible absorption. [b] pH at which 50% of iron is released.

Table 4: Differences (in deg) in the Orientation of the Domains of Lf_N and the R210 Mutants[a]

	R210G	R210E	R210L
Lf_N	2.9	2.8	4.7
R210G		1.9	6.9
R210E			5.2

[a] The main chain atoms of the N1 domains of each pair of recombinant lactoferrins were initially superimposed. The values given are the rotation angles that were then required to superimpose the N2 domains of each pair of proteins.

of one of the tyrosine ligands into a π*-orbital of the bound iron (*32*). The R210 mutants all bind iron at neutral pH and have broadly similar charge-transfer bands. The λ_{max} values for the mutants are, however, consistently higher than that for Lf_N (Table 3) and are in fact closer to the value of 472 nm that has been measured for the iron-bound human Tf N-lobe half-molecule, Tf_N (*33*). This indicates that the mutation of Arg 210 changes the electronic character of the tyrosine ligands and/or the bound iron.

All transferrins characteristically lose iron as the pH is lowered. Although the profile of iron release as a function of pH is somewhat sensitive to the exact solution conditions used and there are potential problems arising from protein precipitation at low pH, it can provide a useful comparison when the proteins are treated identically. In the present case, both R210G and R210E show very similar pH profiles to wild-type Lf_N, losing 50% of iron at pH 4.3−4.4 (Table 3). The R210L mutant, on the other hand, loses iron at significantly higher pH, with a pH_{50} of approximately 5.5. When taken with the results for the R210K mutant (*22*), this means that the pH stability of iron binding decreases in the order $Lf_N \sim$ R210G $\sim$ R210E > R210K > R210L.

Protein Structure. Superpositions of R210G, R210E, and R210L on each other and on Lf_N shows that they all have the same overall polypeptide fold but that the relative domain orientations for each mutant are slightly different from those of Lf_N (see Table 4 and Figure 1). The difference between the Lf_N and R210G domains is primarily a small twist in the backbone strands, whereas the domains of R210E are slightly more closed over the interdomain cleft than the Lf_N domains. In the case of R210L there is a larger difference from Lf_N, which incorporates an increased closure of the domains in R210L, coupled with a significant lateral twist. These small changes in domain orientations must originate from the mutations of Arg 210, presumably through differences in interdomain contacts, since very similar crystallization conditions were used in each case. Small molecular differences may, however, become magnified by crystal packing effects as the molecules seek to maximize crystal contacts.

The iron-binding site is in a well-ordered region of the N-lobe of human Lf. The iron−ligand bonds were initially restrained to conform to standard bond distances, but these restraints were subsequently removed during the latter rounds of refinement. Only in the slightly lower resolution R210E mutant were any restraints retained; in this mutant, removal of the iron−carbonate bond restraints led to a conformation of the carbonate that had an unreasonably short Fe−O bond. With this exception, the iron−ligand bond lengths of the R210 mutants all refined to essentially the same values as those in the Lf_N structure, given the estimated error levels of their structure analyses (approximately 0.1 Å in bond lengths). Thus, the structure of the iron-binding site is apparently not significantly affected by the mutation of Arg 210.

Mutation Site. In each structure, the mutated residue is represented by clearly defined electron density (refer Figure 2). Structural changes accompanying the mutations are limited to (i) the nature of the mutated residue and the interactions it makes, (ii) the conformation of the neighboring residue Lys 301 and an associated β-turn 301−304, and (iii) local changes in the locations of ordered water molecules.

The structure of the R210G mutant reveals the strong preference for a positively charged side chain in the dilysine trigger region of the transferrins. Removal of the basic Arg 210 side chain causes the nearby Lys 301 side chain to undergo a conformational change (Figure 3a), such that it adopts a conformation similar to that of Lys 296 in human Tf_N. This is at the expense of the interdomain Lys 301···Glu 216 salt bridge, which is present in Lf_N but is absent in R210G. In this new conformation, the ϵ-amino group of Lys 301 fills the site occupied by Arg 210 Nη2 in

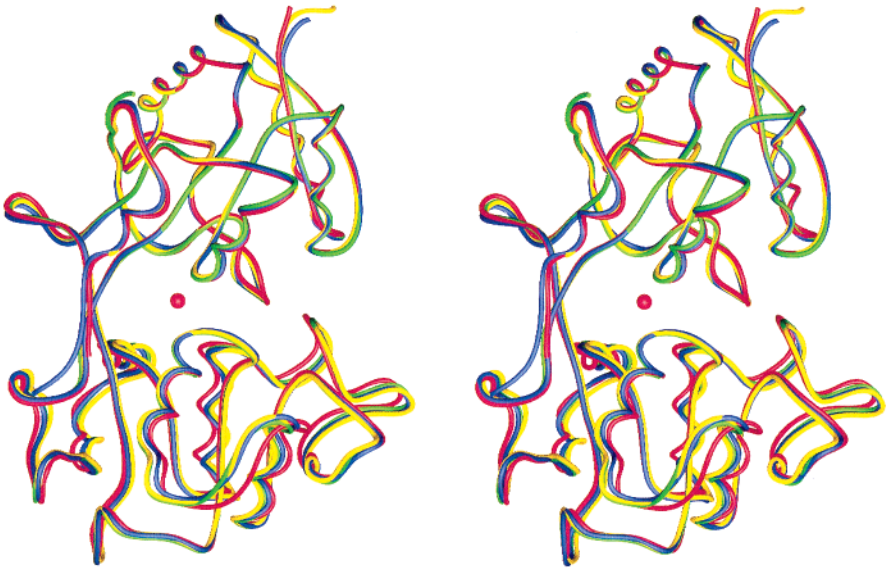

FIGURE 1: Stereoview of the polypeptide conformation of the Arg210 mutants. The R210G (yellow), R210L (red), and R210E (green) mutants are superimposed on to wild-type Lf_N (blue). The position of the bound iron atom is shown with a red sphere.

wild-type Lf_N and by Lys 296 Nζ in Tf_N and makes similar interactions. These interactions, which include a cation$-\pi$ interaction (3.4 Å) with the phenolate ring of one iron ligand, Tyr 92, and a hydrogen bond (2.8 Å) with the phenolate oxygen of the other tyrosine ligand, Tyr 192, are clearly important to the electronic structure of the iron-binding site. A second consequence of the R210G mutation is that a water molecule fills the space vacated by the Arg 210 Nη1 atom and hydrogen bonds to Lys 301 Nζ. Thus, the space occupied by the Arg 210 side chain in Lf_N is instead filled by water molecules and the Nζ atom of Lys 301 in the R210G structure.

Similar readjustments occur in both the R210L and R210E mutants (Figure 3a). In both structures, loss of the basic Arg 210 side chain causes the Lys 301 side chain to change conformation so that its ϵ-amino group occupies the site vacated by Arg 210 Nη2. Again, it makes the same interactions with the two tyrosine ligands, Tyr 92 and Tyr 192. There are also differences, however, that result from the different character of side chains substituted at residue 210. In R210E, the new position of Lys 301 Nζ allows it to make a salt bridge (2.7 Å) with the mutated Glu 210 carboxylate group, in addition to its interactions with Tyr 92 and Tyr 192. No such interaction is possible in the R210L mutant, where the hydrophobic leucine side chain is oriented away from Lys 301, into an adjacent solvent cavity. This solvent cavity allows the R210L mutation to be accommodated without significant disruption of the protein structure, but it is at the expense of disruption of the solvent structure. In fact, in neither R210L nor R210E is there a solvent molecule in the site vacated by the Arg 210 Nη1 atom, because in both cases such a solvent molecule would be too close to the mutated side chain.

In all three mutants a second conformational change occurs that may be linked to the conformational change of Lys 301 (Figure 3b). In all native diferric Lf structures [human Lf (*8*, *34*), bovine Lf (*9*), buffalo Lf (*21*) and equine Lf (*20*)], as well as in wild-type Lf_N (*25*) and its R210K mutant (*22*), Lys 301 forms a salt bridge with Glu 216 and the associated β-turn 301−304 has the relatively uncommon type II conformation. (Note that Lf_N was previously reported to have a type I conformation for this turn, but further refinement has shown that it has the same type II conformation as for native Lf; R. D. Kidd and E. N. Baker, unpublished results). In contrast, in the R210G, R210E, and R210L mutants, the 302−303 peptide bond is flipped so that the 301−304 turn adopts the more favored type I conformation. The same type I conformation occurs in all transferrin structures. We therefore identify a consistent pattern, whereby the two different conformations of Lys 301 (Lys 296 in Tf) are associated with two different conformations of the 301−304 β-turn (296−299 in Tf).

DISCUSSION

Many of the proposed biological functions of the transferrins are directly related to their capacity to bind iron. Each member of the Tf family has been optimized for the performance of particular functions, which has led to the divergence of their iron-binding properties. Probably the most crucial difference for transferrin function is that the Lfs have the capacity to retain iron under more acidic conditions than the serum Tfs. This allows Lf to provide protection against infection by binding free iron, even at low pH, whereas serum Tf transports iron to cells and then releases it in the acidic environment of the endosome (*35*).

The Dilysine Pair. This study has focused on two basic residues close to the N-lobe iron site that are proposed to play a key role in iron release from transferrins (*17*, *19*). In human transferrin these residues are Lys 206 and Lys 296. In the N-lobe of human serum transferrin the Nζ atoms of these two lysine residues are only 3.0 Å apart, implying that they are hydrogen bonded (*18*). Such an interaction further implies that the two side chains cannot both be positively charged and, hence, that at least one of them is perturbed by its protein environment such that its pK_a is below 5.75, the pH of the structure determination.

The interdomain hydrogen bond between Lys 206 and Lys 296 should stabilize the iron-bound form of human serum transferrin at neutral pH. However, under more acidic

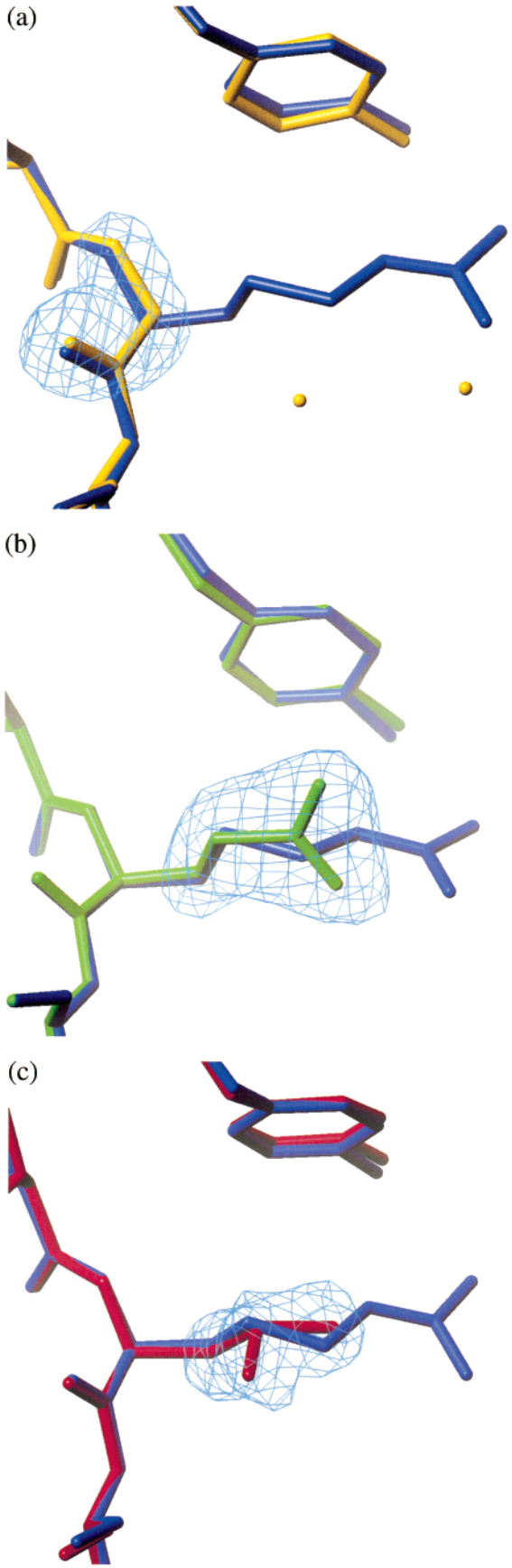

FIGURE 2: Superposition of the mutation site of (a) R210G, (b) R210E, and (c) R210L onto the wild-type Lf_N. The $|F_o| - |F_c|$ electron density is shown after simulated annealing refinement with the main chain atoms of Gly 210 in R210G, and the side chain atoms of residue 210 in R210E and R210L having been omitted from the models.

conditions, when both lysines are protonated, the hydrogen bond would be replaced by a repulsive interaction between the two positive charges. Since the two lysines are contributed by opposing domains, this should destabilize the closed form of human serum transferrin, leading to the description of these residues as providing a pH-dependent trigger for the release of iron (*19*). Mutation of either lysine results in loss of the dilysine interaction. Although there is no change to the rest of the transferrin N-lobe structure (*36*), iron release is altered drastically; iron is retained to significantly lower pH, and the rate of iron release is slowed on the order of 10^4-fold, depending on solution conditions (*36*−*38*). These studies confirm the importance of the dilysine interaction in modulating iron release from the N-lobe of transferrin at reduced pH.

The dilysine pair is a highly conserved feature of the N-lobe of the serum transferrins. The two lysines are present in all serum Tf sequences that are available to date, and the same hydrogen-bonded interaction has been found in the N-lobes of all serum Tfs or oTfs for which crystal structures are currently available. There are striking contrasts in the lactoferrins, however. In human Lf the equivalent residues, Arg 210 and Lys 301, are both basic, but the crystal structure shows that they do not interact. Arg 210 occupies a similar position in human Lf to that of Lys 206 in human Tf, but the Lys 301 side chain has a different conformation from that of Lys 296 in human Tf, such that the termini of the Arg 210 and Lys 301 side chains are 5.5 Å apart in diferric human Lf (*8*, *34*) and in Lf_N (*25*). Even more intriguingly, the amino acid sequences of several other Lfs, including bovine, equine, and buffalo Lf, have two lysine residues at the sequence positions corresponding to Lys 206 and Lys 296 in human Tf. However, the crystal structures of these Lfs (*9*, *20*, *21*) indicate that, instead of forming the dilysine interaction, Lys 301 in each case adopts the same conformation as in human Lf, forming a salt bridge with Glu 216. The lack of the dilysine trigger repulsion and the formation of this additional salt bridge thus appears to be a conserved feature of lactoferrins that contributes toward their ability to retain iron to much lower pH, compared with the serum transferrins and ovotransferrins.

Alternate Conformations in the R210 Mutant Structures. The structures of the R210G, R210E, and R210L mutants reported here, together with that of the R210K mutant (*22*), show that Lys 301 in lactoferrins is capable of adopting two alternative conformations (Figure 3a). In the R210K mutant, as in all native Lfs and in wild-type Lf_N, Lys 301 forms a salt bridge with Glu 216; we will refer to this as the A conformation. In the R210G, R210E, and R210L mutants, however, Lys 301 has the same conformation as Lys 296 in transferrins, which we will refer to as the B conformation. Both conformations are favorable for a protonated lysine residue; in the first, its ϵ-amino group forms the ion pair with Glu 216, and in the second, it interacts with the negative charge on the two tyrosine ligands, forming a hydrogen bond with the phenolate oxygen of Tyr 192 (2.8−3.0 Å) and a cation−π interaction (3.4 Å) with the aromatic ring of Tyr 92. With two potentially stable conformations for Lys 301, the structure adopted will be that which has the lowest energy in each protein.

Linked to the change in conformation of Lys 301 is that of the β-turn 301−304 (296−299 in Tfs) (Figure 3b). In Tfs,

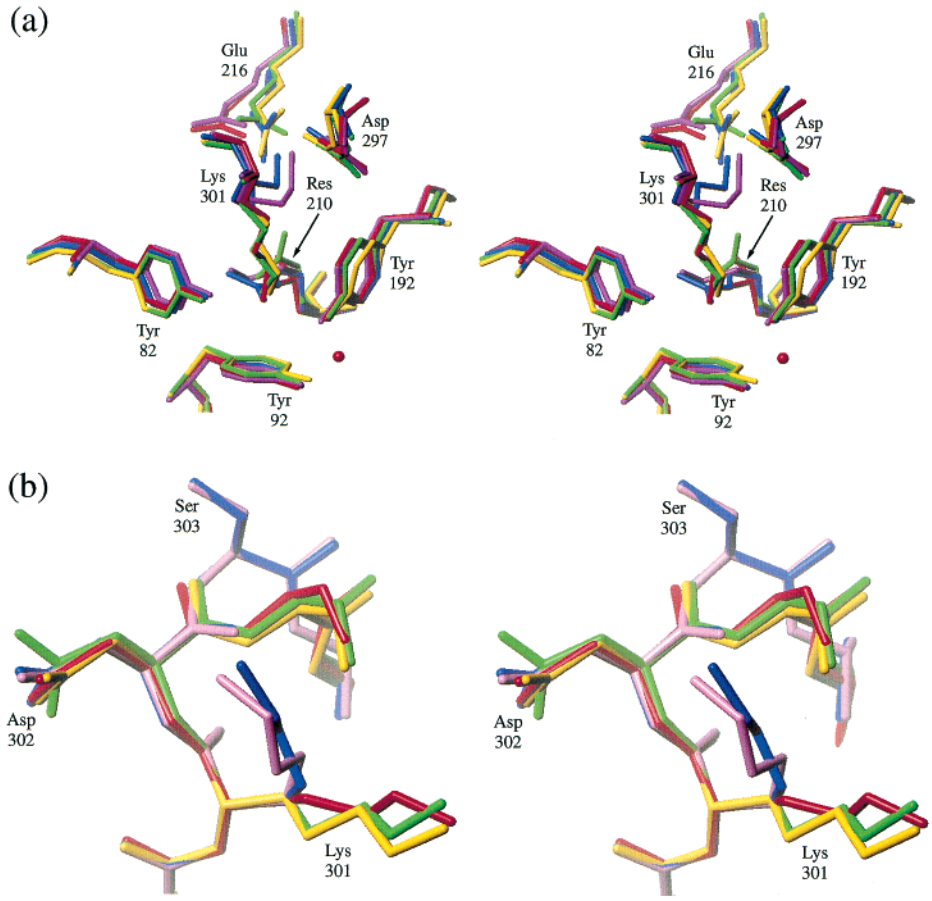

FIGURE 3: Stereoviews of the dilysine pair region in the lactoferrin N-lobe, for wild-type Lf_N and the R210 mutants. (a) Lys301 can be seen in its A conformation, interacting with Glu216 and Asp297, in wild-type Lf_N (blue) and R210K (purple), and in its B conformation, adjacent to residue 210, in the other mutants R210E (green), R210G (yellow), and R210L (red), where the side chain of residue 210 is no longer positively charged. The position of the iron atom in these mutants is shown with a red sphere. (b) The conformation of the β-turn 301−304 is seen to be correlated with the conformation of Lys301, with a type II conformation in Lf_N and R210K (302 C═O pointing toward Lys301 Nζ) and type I for R210E, R210G, and R210L, as in Tf_N (302 C═O pointing away).

and in the R210G, R210E, and R210L mutants of Lf, this turn has the energetically favorable type I configuration, and Lys 296 (Lys 301) has the B conformation, occupying the tyrosine interaction site. In Lfs, however, the turn has a type I configuration in the apo protein, but this changes to the energetically less favorable type II configuration in the closed, iron-bound form, and Lys 301 forms an ion pair with Glu 216. It appears that Lys 301 cannot adopt this alternative conformation A unless the 302−303 peptide is flipped and the turn is changed to type II. This is because the peptide flip directs the C═O group of residue 302 (297) toward the Lys 301 ϵ-amino group, forming a hydrogen bond; otherwise, the NH group of residue 303 would be directed toward the amino group, creating an unfavorable interaction.

We conclude from these structures that when the side chain of residue 210 is not positively charged (as in R210G, R210E, and R210L), the Lys 301 side chain has the B conformation, its Nζ atom filling the tyrosine interaction site that is occupied by Lys 296 Nζ in Tf and Arg 210 Nη2 in human Lf. Conversely, when a positively charged lysine residue is at position 210, which we assume to be the case in R210K and in bovine, buffalo, and equine Lfs, Lys 301 adopts the A conformation, rotated away to form a salt bridge with Glu 216. Concomitantly, a water molecule fills the tyrosine interaction site and hydrogen bonds to Lys 210 Nζ. The reasons for the conformational change are not steric but are to avoid charge repulsion with Lys 210. That is, the interactions in the dilysine trigger region are determined by the charge of residue 210.

These arguments strongly suggest, by analogy, that in serum transferrins it is Lys 206 that has a depressed pK_a and is uncharged. This would allow the protonated Lys 296 Nζ (the equivalent to Lys 301 Nζ in Lf) to fill the adjacent tyrosine interaction site (as in R210G, R210E, and R210L) and also form the dilysine pair hydrogen bond with Lys 206.

Environment of the Dilysine Trigger Residues. The depressed pK_a of Lys 206 in Tf must reflect local electrostatic factors that differ between Tfs and Lfs. The environments of Lys 206 and Lys 296 in Tfs and of Arg/Lys 210 and Lys 301 in Lfs are remarkably similar, in general, and the residues that interact with these basic residues, including Glu 216, are all conserved in both Lfs and Tfs. The only obvious difference is the presence in Tfs of a histidine residue, His 207, that lies only 4.5 Å from the ϵ-amino group of Lys 206, linked to it by a bridging water molecule. Among the vertebrate Tfs, only *Xenopus* Tf lacks a histidine at this position, but this protein also lacks the second lysine of the dilysine pair (Table 5). In most Lfs, in contrast, the residue equivalent to His 207 is acidic, Glu 211 in human Lf. The presence of a negatively charged side chain at residue 211 (human Lf numbering) would elevate the pK_a of a lysine side chain at residue 210, ensuring that it is protonated at

Table 5: Residues in the Dilysine Trigger Region of Transferrins and Lactoferrins

protein	residues				
human serum transferrin	Lys 206	Lys 296	His 207	Glu 212	Tyr 85
rabbit serum transferrin	Lys 206	Lys 296	His 207	Glu 212	Tyr 85
horse serum transferrin	Lys 208	Lys 298	His 209	Glu 214	Tyr 82
pig serum transferrin	Lys 210	Lys 300	His 211	Glu 216	Tyr 84
rat serum transferrin	Lys 206	Lys 296	His 207	Glu 212	Tyr 85
hen ovotransferrin	Lys 209	Lys 301	His 210	Glu 215	Tyr 82
Xenopus transferrin	Lys 210	Gln 300	Gln 211	Glu 216	Tyr 82
human lactoferrin	Arg 210	Lys 301	Glu 211	Glu 216	Tyr 82
mouse lactoferrin	Arg 210	Lys 301	Gly 211	Glu 216	Tyr 81
bovine lactoferrin	Lys 210	Lys 301	Glu 211	Glu 216	Tyr 82
pig lactoferrin	Lys 205	Arg 296	Glu 206	Glu 211	Tyr 78
buffalo lactoferrin	Lys 210	Lys 301	Glu 211	Glu 216	Tyr 82
horse lactoferrin	Lys 210	Lys 301	Asp 211	Glu 216	Tyr 82
goat lactoferrin	Lys 210	Lys 301	Glu 211	Glu 216	Tyr 83

physiological pH and that no dilysine interaction with Lys 301 can form. The only exception is in mouse Lf (Table 5), but in this case residue 210 is an arginine, whose steric bulk would prevent Lys 301 from occupying the tyrosine interaction site. Despite this pattern, which suggests a significant role for residue 207 (211) in controlling the pK_a of a lysine at position 206 (210), this cannot be the only contributor. This is indicated by studies of the H207E mutant of Tf_N (*39*); although this mutation does strengthen iron binding, the dilysine pair evidently remains intact.

Factors Influencing Iron Binding and Release. It is possible that a dilysine interaction could be present at elevated pH, in those Lfs with the appropriate pair of lysines, when one or both is deprotonated. This evidently has no influence on iron release, however. The current model for the acid-mediated release of iron involves a series of protonation events that progressively destabilize the iron coordination. The initial event is the protonation of the synergistic carbonate ion, to produce a bicarbonate-bound form with weakened iron binding (*18*). In serum Tfs, this is presumably followed by the protonation of the dilysine pair, which further destabilizes the closed form of the N-lobe and so contributes to the eventual breakdown of the iron−transferrin complex (*19, 36*−*38*). If the protonation of the dilysine pair were to occur at significantly higher pH than that of the carbonate, as might be the case in those Lfs with two lysines, we suggest that this simply results in the change of conformation of Lys 301, to occupy its alternative A site. The significant difference in Tfs is evidently that both the carbonate ion and the dilysine pair are protonated at rather similar pH, producing a synergistic impetus to iron release.

Comparisons of the Arg 210 mutants with each other and with other transferrin N-lobe proteins show that there are three sites adjacent to the iron site that are variably occupied in these proteins (Table 6). The most important appears to be site 1, which is occupied by Arg 210 Nη2 in Lf_N, by Lys 301 Nζ in R210G, R210E, and R210L, and by Lys 296 Nζ in Tf_N. A positively charged group at this site should depress the pK_a of one or both of the tyrosine ligands with which it interacts and thereby favor iron retention at low pH. The presence of a water molecule in place of a positively charged group at this site in R210K probably accounts for its weaker iron retention at low pH (pH_{50} 5.0 compared with 4.3 for Lf_N); by extension, bovine Lf and other lactoferrins with lysine at both residues 210 and 301 may share this property. Site 2 is occupied by Arg 210 Nη1 in Lf_N, where it has a cation−π interaction with Tyr 82. This site is only 2−3 Å from site 1, and it is the presence of lysine amino groups at both site 2 and site 1 (the dilysine pair) that significantly destabilizes iron retention in Tf_N at low pH. Otherwise, however, the group occupying site 2 does not seem to have much influence on iron retention, as shown by the fact that R210E (carboxylate group at site 2, forming an ion pair with Lys 301 at site 1) and R210G (water molecule at site 2) both have very similar iron retention to Lf_N. Similarly, the group present at site 3 does not appear to have significant impact on iron retention at low pH. In Lf_N and R210K, Lys 301 Nζ occupies this site, forming an interdomain ion pair with Glu 216, but R210G and R210E do not have such an interaction and have very similar pH_{50} values (Table 3).

Table 6: Atoms at the Three Polar Sites in the Dilysine Trigger Region of the Transferrins

protein	site 1[a]	site 2[b]	site 3[c]
human lactoferrin (Lf_N)	Arg 210 Nη2	Arg 210 Nη1	Lys 301 Nζ
bovine lactoferrin	water	Lys 210 Nζ	Lys 301 Nζ
human serum transferrin (Tf_N)	Lys 296 Nζ	Lys 206 Nζ	water
R210K	water	Lys 210 Nζ	Lys 301 Nζ
R210G	Lys 301 Nζ	water	water
R210E	Lys 301 Nζ		
R210L	Lys 301 Nζ		

[a] Hydrogen bond with Tyr 192 (Tyr 188 in Tf), cation−π interaction with Tyr 92 (Tyr 95 in Tf). [b] Cation-π interaction with Tyr 82 (Tyr 85 in Tf). [c] Hydrogen bond with Glu 216 (Glu 212 in Tf).

The R210L mutant deserves additional mention. In this case, the substitution of a hydrophobic group for Arg 210 destabilizes iron retention by more than 1 pH unit. This is attributed primarily to the disruption of the interdomain hydrogen-bonding network, since neither site 2 nor 3 is filled in this mutant; the altered domain orientations may also play a part, however.

Finally, although the metal−ligand bond lengths in the Arg 210 mutants cannot be distinguished crystallographically from those in Lf_N, the visible spectrum λ_{max} values (Table 3) show that there is a difference in the metal−tyrosine interaction. Significantly, R210K, R210G, R210E, and R210L, together with Tf_N, all have a red shift of λ_{max} compared with Lf_N. In fact, Lf_N is really the outlier, and we attribute this to the numerous and different interactions that are made with the tyrosine ligands by an arginine side chain at position 210. In Lf_N, Arg 210 Nη2 hydrogen bonds with Tyr 192 Oη, the guanidinium group forms a plane-to-plane stacking interaction with the aromatic ring of Tyr 92, and Arg 210 Nϵ makes a NH-π hydrogen bond with the aromatic ring of Tyr 192. These interactions differ significantly from the interactions made by lysine ϵ-amino groups when they occupy site 1, adjacent to the tyrosine ligands.

CONCLUSIONS

Our results clarify the roles of the two basic residues in the "dilysine trigger" regions of lactoferrins and transferrins and their different behaviors in these two subfamilies. The key finding is that, providing residue 210 in Lf is not positively charged, Lys 301 can occupy the site adjacent to the two tyrosine ligands that is occupied by Lys 296 in Tf. From this we conclude that the dilysine interaction in the N-lobe of Tfs arises because (a) the site adjacent to the two

tyrosine ligands (the tyrosine interaction site) is a very favorable one for the protonated ϵ-amino group of Lys 296, and (b) the pK_a of the ϵ-amino group of Lys 206 is depressed by its location between Lys 296 and His 207. Only below approximately pH 6 does Lys 206 become protonated, creating repulsive interactions with Lys 296 that must contribute to the more facile iron release shown by Tfs. In the N-lobe of Lfs, two effects prevent an equivalent interaction. Some Lfs (human and mouse) possess an arginine at position 210, and the greater bulk of the arginine side chain allows it to occupy the tyrosine interaction site, sterically preventing the approach of Lys 301. Other Lfs have two lysine residues, as in Tfs, but the pK_a of Lys 210 is raised by an adjacent acidic residue, such that it is positively charged at physiological pH. Charge repulsion then prevents Lys 301 from occupying the tyrosine interaction site, which is instead filled by water. In both cases, Lys 301 adopts an alternative conformation, coupled with the flip of the 302−303 peptide, and forms a salt bridge with Glu 216. In both cases also, the lack of a dilysine interaction allows Lfs to retain iron to significantly lower pH than Tfs. These results show how only a very small number of residue substitutions, at critical positions, can drastically change the physiological functions of otherwise extremely similar proteins.

REFERENCES

1. Brock, J. H. (1985) In *Metalloproteins* (Harrison, P., Ed.) pp 183−262, MacMillan, London.
2. Aisen, P., and Harris, D. C. (1989) In *Iron carriers and iron proteins* (Loehr, T., Ed.) pp 241−351, VCH Publishers, New York.
3. Baker, E. N. (1994) *Adv. Inorg. Chem. 41*, 389−463.
4. Grossmann, J. G., Neu, M., Pantos, E., Schwab, F. J., Evans, R. W., Towns-Andrews, E., Lindley, P. F., Appel, H., Thies, W.-G., and Hasnain, S. S. (1992) *J. Mol. Biol. 225*, 811−819.
5. Anderson, B. F., Baker, H. M., Norris, G. E., Rumball, S. V., and Baker, E. N. (1990) *Nature 344*, 784−787.
6. Jeffrey, P. D., Bewley, M. C., MacGillivray, R. T. A., Mason, A. B., Woodworth, R. C., and Baker, E. N. (1998) *Biochemistry 37*, 13978−13986.
7. Kurokawa, H., Dewan, J. C., Mikami, B., Sacchettini, J. C., and Hirose, M. (1999) *J. Biol. Chem. 274*, 28445−28452.
8. Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W., and Baker, E. N. (1989) *J. Mol. Biol. 209*, 711−734.
9. Moore, S. A., Anderson, B. F., Groom, G. R., Haridas, M., and Baker, E. N. (1997) *J. Mol. Biol. 274*, 222−236.
10. Bailey, S., Evans, R. W., Garratt, R. C., Gorinsky, B., Hasnain, S., Horsburgh, C., Jhoti, H., Lindley, P. F., Mydin, A., Sarra, R., and Watson, J. L. (1988) *Biochemistry 27*, 5804−5812.
11. Kurokawa, H., Mikami, B., and Hirose, M. (1995) *J. Mol. Biol. 254*, 196−207.
12. Harris, W. R. (1986) *Biochemistry 25*, 803−808.
13. Day, C. L., Stowell, K. M., Baker, E. N., and Tweedie, J. W. (1992) *J. Biol. Chem. 267*, 13857−13862.
14. Aisen, P., Liebman, A., and Zweier, J. (1978) *J. Biol. Chem. 253*, 1930−1937.
15. Kretchmar, S. A., and Raymond, K. N. (1986) *J. Am. Chem. Soc. 108*, 6212−6218.
16. Evans, R. W., and Williams, J. (1978) *Biochem. J. 173*, 543−552.
17. Baker, E. N., and Lindley, P. F. (1992) *J. Inorg. Biochem. 47*, 147−160.
18. MacGillivray, R. T. A., Moore, S. A., Chen, J., Anderson, B. F., Baker, H., Luo, Y., Bewley, M., Smith, C. A., Murphy, M. E., Wang, Y., Mason, A. B., Woodworth, R. C., Brayer, G., and Baker, E. (1998) *Biochemistry 37*, 7919−7928.
19. Dewan, J. C., Mikami, B., Hirose, M., and Sacchettini, J. C. (1993) *Biochemistry 32*, 11963−11968.
20. Sharma, A. K., Paramasivam, M., Srinivasan, A., Yadav, M. P., and Singh, T. P. (1998) *J. Mol. Biol. 289*, 303−317.
21. Karthikeyan, S., Paramasivam, M., Yadav, S., Srinivasan, A., and Singh, T. P. (1999) *Acta Crystallogr. D 55*, 1805−1813.
22. Peterson, N. A., Anderson, B. F., Jameson, G. B., Tweedie, J. W., and Baker, E. N. (2000) *Biochemistry 39*, 6625−6633.
23. Baker, H. M., Day, C. L., Norris, G. E., and Baker, E. N. (1994) *Acta Crystallogr. D 50*, 380−384.
24. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
25. Day, C. L., Anderson, B. F., Tweedie, J. W., and Baker, E. N. (1993) *J. Mol. Biol. 232*, 1084−1100.
26. Navaza, J. (1994) *Acta Crystallogr. A 50*, 157−163.
27. Brünger, A. T. (1992) *X-PLOR Version 3.1, A System for Crystallography and NMR*, Yale University Press, New Haven, CT.
28. 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 54*, 905−921.
29. Brunger, A. T. (1992) *Nature 355*, 472−474.
30. Cambillau, C., Roussel, A., Inisan, A.-G., and Knoops-Mouthuy, E. (1996) Bio-Graphics, AFMB−CNRS. Marseille. France.
31. Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
32. Patch, M. G., and Carrano, C. J. (1981) *Inorg. Chim. Acta 56*, L71−L73.
33. Funk, W. D., MacGillivray, R. T. A., Mason, A. B., Brown, S. A., and Woodworth, R. C. (1990) *Biochemistry 29*, 1654−1660.
34. Haridas, M., Anderson, B. F., and Baker, E. N. (1995) *Acta Crystallogr. D 51*, 629−646.
35. 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.
36. Nurizzo, D., Baker, H. M., He, Q.-Y., MacGillivray, R. T. A., Mason, A. B., Woodworth, R. C., and Baker, E. N. (2001) *Biochemistry 40*, 1616−1623.
37. Steinlein, L. M., Ligman, C. M., Kessler, S., and Ikeda, R. A. (1998) *Biochemistry 37*, 13696−13703.
38. He, Q.-Y., Mason, A. B., Tam, B. M., MacGillivray, R. T. A., and Woodworth, R. C. (1999) *Biochemistry 38*, 9704−9711.
39. Yang, A. H−W., MacGillivray, R. T. A., Chen, J., Luo, Y., Wang, Y., Brayer, G. D., Mason, A. B., Woodworth, R. C., and Murphy, M. E. P. (2000) *Protein. Sci. 9*, 49−52.

BI020443A