# Tertiary templates for the design of diiron proteins

Christopher M Summa*†, Angela Lombardi‡, Mitchell Lewis* and William F DeGrado*

Diiron proteins represent a diverse class of structures involved in the binding and activation of oxygen. This review explores the simple structural features underlying the common metal-ion-binding and oxygen-binding properties of these proteins. The backbone geometries of their active sites are formed by four-helix bundles, which may be parameterized to within approximately 1 Å root mean square deviation. Such parametric models are excellent starting points for investigating how asymmetric deviations from an idealized geometry influence the functional properties of the metal ion centers. These idealized models also provide attractive frameworks for *de novo* protein design.

**Addresses**
*Department of Biochemistry and Biophysics, University of Pennsylvania, School of Medicine, Philadelphia, PA 19104-6059, USA
†e-mail: csumma@mail.med.upenn.edu
‡Department of Chemistry, University of Naples "Federico II", Via Mezzocannone, 4 I-80134 Napoli, Italy

**Current Opinion in Structural Biology** 1999, **9**:500–508

http://biomednet.com/elecref/0959440X00900500

 ISSN 0959-440X

**Abbreviations**

| | |
|---|---|
| **$\Delta^9$ ACP desaturase** | $\Delta^9$ stearoyl acyl carrier protein desaturase |
| **MMO** | methane mono-oxygenase |
| **PDB** | Protein Data Bank |
| **R2** | R2 subunit of ribonucleotide reductase |
| **rmsd** | root mean square deviation |

## Introduction

Dinuclear iron sites are found in a functionally diverse class of proteins involved in oxygen binding and activation [1–5]. Prototypes for this class of proteins include hemerythrin and myohemerythrin [6], which reversibly bind and transport oxygen. Ferritins and bacterioferritins catalyze the oxidation of Fe(II) to Fe(III) (ferroxidase activity) and sequester Fe(III) as an insoluble oxide within the core of a virus-like structure [7–10]. Diiron proteins also catalyze the hydroxylation of alkanes, alkenes and aromatic groups [1,2,4,5,11–14], and a diiron site in the R2 subunit of ribonucleotide reductase (R2) is responsible for the formation of an organic radical [15].

How the diiron proteins tune the properties of a single inorganic cofactor to obtain such a diversity of function is a question that is currently being addressed through spectroscopic and mechanistic studies [1,2,4,5,11,15]. There are a number of structural and mechanistic commonalities amongst this class of diiron proteins — their active sites are formed by the association of four antiparallel helices and their diiron sites bind dioxygen, initially in their diferrous (bis-$Fe^{2+}$) state. Oxygen is rapidly reduced to the peroxide, with the concomitant oxidation of the metal ions to the diferric (bis-$Fe^{3+}$) state. The metal-binding sites of the hemerythrins are relatively rigid and are rich in nitrogen-containing ligands, which specifically stabilize both the diferrous and the diferric/peroxo states. Thus, binding of oxygen to these centers is reversible, as required for oxygen transport. The active sites of hemerythrins are distinct from those of proteins involved in oxygen activation and will not be discussed further here.

The metal-ion-chelating sites of the other diiron proteins are more flexible, richer in carboxylates and tend to consume oxygen, rather than binding it reversibly. Oxygen binding leads to a diferric intermediate with a symmetrical bridging peroxo group [16–19], but the fate of this intermediate differs depending on the specific geometries of the individual active sites. In methane mono-oxygenase (MMO), two electrons are subsequently transferred from the diferric center to the bound peroxide, thus cleaving the peroxide O–O bond. This reductive cleavage of the O–O bond gives rise to a 'diamond core' diferryl (bis-$Fe^{4+}$) intermediate [20], which is stabilized by a bis-μ-oxo bridge. The resulting oxo-bridged center, a potent electrophile, is an essential intermediate in the subsequent hydroxylation reaction. Hydroxylation of the bound alkane leads to a diferric center, which is then primed for another catalytic cycle by reduction to the diferrous state. A similar mixed valency Fe(III)/Fe(IV) intermediate is formed in R2, which accepts a single electron from a neighboring tyrosine sidechain, leading to the active form of the protein with a tyrosine radical and a diferric iron center [1,2,4,5,11,15].

## Metal-ion-chelation geometries

As a diiron site cycles among its various functional oxidation states, the protein must accommodate large changes in the net charge of the metal ions, ranging from +4 to +8 in MMO. This stabilization is accomplished via shifts in the geometries of the liganding glutamic acid sidechains, as well as through changes in the number and protonation states of ligating oxo groups. In the reduced di-Fe(II) [or the structurally similar di-Mn(II)] state, diiron proteins tend to chelate each ion using four to five ligands arranged in a distorted trigonal bipyramid or octahedron, with open chelation sites available for binding oxygen (Figure 1a). Each of the four helices in the active site donates one carboxylate ligand to the diiron site. Two of the active-site helices contain a Glu-Xxx-Xxx-His motif; the δ-nitrogen of each histidine ligates to a single iron atom, while the glutamic acid sidechains often form bridging interactions with both metal ions. Two types of bridging interactions are possible (Figure 1) [1–4,21,22]. In a 1,3 bridge, each of the carboxylate oxygens forms a

**Table 1**

**Parameterization of the symmetric active-site helices of various diiron proteins*.**

| Protein name | α angle | β angle | γ angle | X disp | Y disp | Z disp | rmsd | Number of atoms superimposed |
|---|---|---|---|---|---|---|---|---|
| Bacterioferritin[a] | 162.76 | −14.03 | −4.13 | 4.54 | 5.64 | 7.47 | 1.04 | 48 |
| Rubrerythrin[b] | 161.35 | −11.13 | −3.74 | 4.14 | 6.00 | 7.29 | 0.89 | 48 |
| Ferritin [c] | 164.52 | −8.23 | −4.13 | 4.32 | 6.35 | 7.94 | 1.37 | 48 |
| R2[d] | 166.98 | −18.65 | −0.98 | 4.65 | 5.70 | 6.88 | 1.11 | 48 |
| MMO[e] | 170.67 | −20.04 | 2.56 | 4.12 | 5.41 | 6.53 | 0.96 | 36 |
| $\Delta^9$ ACP desaturase[f] | 182.64 | −15.55 | −2.17 | 4.44 | 5.73 | 9.12 | 1.29 | 36 |
| Consensus[g] | 167.51 | −13.31 | −1.77 | 4.44 | 5.93 | 7.47 | 1.83 | 264 |
| R2 Fe(III)[h] | 158.95 | −17.82 | −2.17 | 4.40 | 5.68 | 6.82 | 1.16 | 42 |
| R2 Fe(II)[i] | 140.47 | −18.66 | −1.38 | 4.47 | 5.70 | 7.41 | 1.11 | 42 |

*Parameters for generating 222-symmetrical diiron-binding sites. The Cα rmsd between a 222-symmetrical model and the crystal structure of a given protein was used as a metric. A genetic algorithm was used to optimize the parameters for the best fit of the model to the X-ray/NMR structure (lowest rmsd). The bundles were built starting from a single idealized helix aligned along the z axis. The angle α corresponds to the rotation of the helix about its major axis, the angle β corresponds to the angle between the helical major axis and the x–z plane and the angle γ corresponds to the angle between the helical major axis and the y–z plane. X disp, Y disp and Z disp are translations in x, y and z, respectively. Once the single helix has been moved into position, the four-helix bundle is generated by application of the 222-symmetry operators. (a) The PDB code for bacterioferritin is 1BCF [26], residues 14–25, 47–58, 90–101 and 123–134. (b) The PDB code for rubrerythrin is 1RYT [27], residues 16–27, 49–60, 90–98, 100–102 and 124–135. (c) The PDB code for ferritin is 2FHA [45], residues 23–34, 58–69,103–114 and 137–148. (d) The PDB code for R2 is 1RIB [46], residues 80–91, 111–122, 234–245, 200–204 and 206–212. (e) The PDB code for MMO is 1MTY [47], residues 110–121, 140–151 and 239–250. (f) The PDB code for $\Delta^9$ ACP desaturase is 1AFR [24], residues 139–150, 192–203 and 222–233. (g) The consensus parameters were obtained using a best fit of a single model to all of the above proteins. (h) The PDB code for R2 Fe(III) is 1R2F [38], residues A63–A74, A94–A105, A154–A158, A160–A166 and A188–A199. (i) The PDB code for R2 Fe(II) is 2R2F [38], residues A63–A74, A94–A105, A154–A158, A160–A166 and A188–A199.

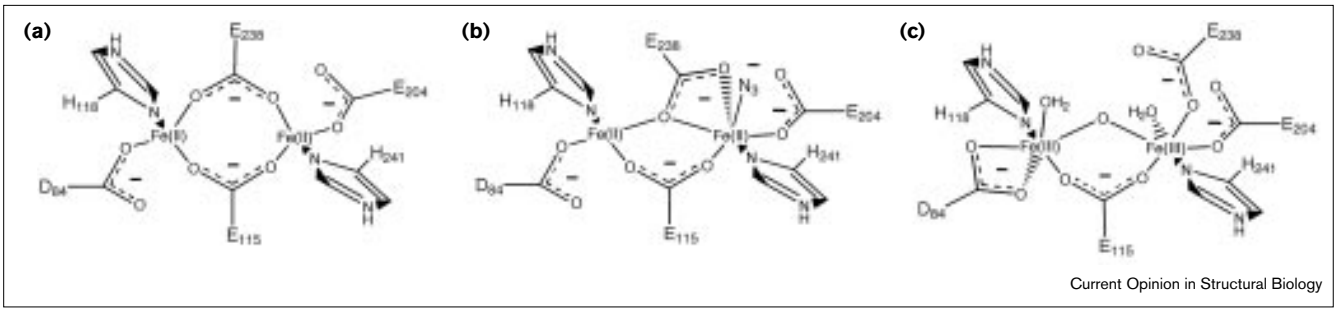
monovalent interaction with a different iron ion. In a 1,1 bridging interaction, a single carboxylate oxygen simultaneously interacts with both iron atoms, generally leading to a shorter Fe–Fe distance and enhanced antiferromagnetic coupling. The remaining two helices in the active site four-helix bundle each contribute a single, nonbridging glutamic acid carboxylate, which forms a monovalent or divalent interaction with a single iron atom. A water bridge is found in the active site of diferrous MMO [21], but not in diferrous R2 [23], diferrous $\Delta^9$ stearoyl acyl carrier protein desaturase ($\Delta^9$ ACP desaturase) [24], bis-Mn(II) R2 [25] or bis-Mn(II) bacterioferritin [26]. The metal-binding sites of these proteins often show approximate twofold symmetry (Figure 1a).

A number of diiron proteins have also been crystallographically characterized in the diferric state. In the Fe(III) state, the metal ions tend to bind additional hydroxide, oxo or carboxylate ligands, forming distorted octahedral complexes [1,2,4,5,11]. Often, these ligands bridge the two metal centers, leading to a shorter Fe–Fe distance. Figure 1 compares the diferrous and the diferric states of R2.

## A four-helix bundle forms the active site of diiron proteins

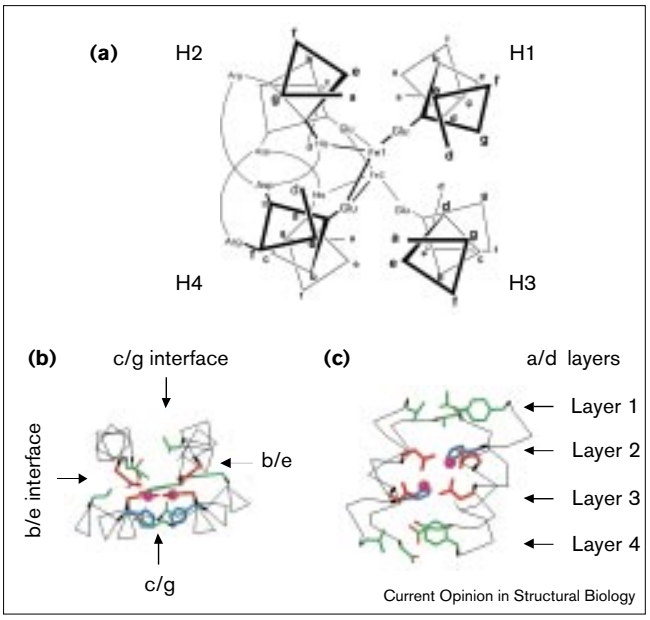
In order to identify common structural features in the Glu-Xxx-Xxx-His class of proteins, we have structurally superimposed the active sites of six functionally and

**Figure 1**




Schematic diagram of the diiron site in R2 in **(a)** the diferrous state, **(b)** the azide-modified diferrous state and **(c)** the diferric state (adapted from [23]). Note the carboxylate shift of E238 from 1,3 bridging in the diferrous state (a) to 1,1 bridging in the azide complex (b). Also, D84 shifts from being a monodentate ligand to being a bidentate ligand between the ferrous structures (a,b) and the ferric structure (c).

**Figure 2**



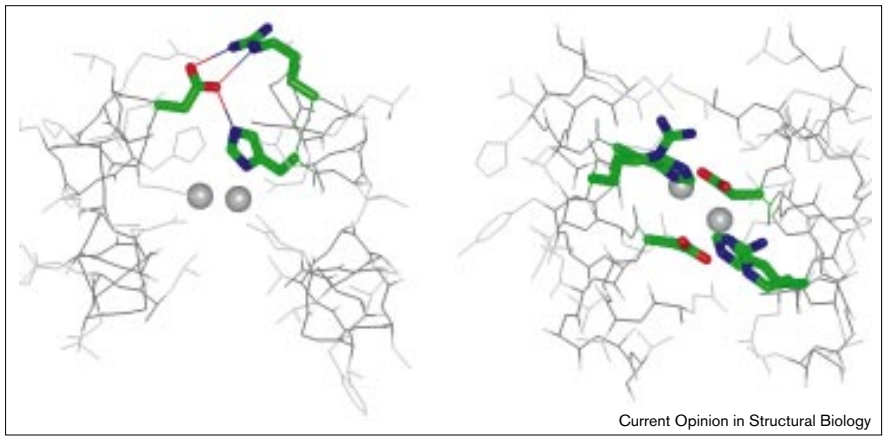

Architecture of a diiron-binding four-helix bundle. **(a)** Schematic diagram of the four-helix bundle defining the active site of a typical diiron protein. The helices are 12 residues in length and their positions are labeled 'a' through 'g' using the nomenclature of coiled coils. Helices are labeled H1 through H4, in order of their sequence in the protein. **(b,c)** Top and side view, respectively, of the four-helix bundle harboring the active site of bacterioferritin (PDB code 1BCF [26]). The backbone is drawn as lines connecting the Cα atoms. The liganding glutamic acid residues are colored red and the liganding histidine residues are colored blue. The other residues defining the diiron-binding site are colored by atom: carbon, green; oxygen, red; nitrogen, blue. The iron atoms are shown as magenta spheres. The interhelical distance at the c/g interface is longer than that seen at the b/e interface, producing a rectangular, rather than square, bundle when viewed from above. When viewed from the side, the 'a' and 'd' positions partition into layers that help to define the diiron-binding site.

structurally diverse members of this class (Table 1). These proteins include three small four-helix bundle ferroxidases — bacterioferritin [26], rubrerythrin [27] and H ferritin [28] — as well as three more complex proteins — R2 [22], MMO [21] and $\Delta^9$ ACP desaturase [24]. The diiron sites of these proteins are all contained within an approximately 18 Å segment of a four-helix bundle, whose helices will be designated H1, H2, H3 and H4 in order of their sequence. The helices are arranged in the order H1, H2, H4 and H3, when considered counterclockwise and viewed as in Figure 2a, b. The helices have a slight left-handed tilt, as in classical four-helix bundles [29] and antiparallel four-stranded coiled coils [30]. Coiled coils [31–34] show a structural seven-residue repeat and successive positions in this heptad are generally designated by the letters 'a' through 'g'. Four-helix bundles with straight or curved helices also approximately conform to a seven-residue repeat within the limited (18 Å) region considered here. Thus, we have adopted this heptad nomenclature for designating specific positions within the active sites of diiron proteins, although we do not assume that they adopt a classical coiled-coil motif.

The helices each contain a single chelating glutamic acid residue at position 'a' of the heptad motif that projects towards the center of the bundle (Table 2). Two exceptions to this rule involve a conservative replacement of the chelating glutamic acid with aspartic acid in H3 of R2 and with glutamine in H4 of ferritin. Histidine residues located four residues downstream at the 'd' position of H2 and H4 contribute the remaining sidechain ligands to the diiron site. Small sidechains frequently occur at the analogous 'd' positions of H1 and H3. These residues play a particularly important role in defining the accessibility of the active site to oxygen and other substrates. Thus, the residues at the central 'a' and 'd' positions of the active site four-helix bundle define the chelation and immediate environment of the diiron sites of these proteins.

The 'a' and 'd' sidechains of the diiron four-helix bundles segregate into distinct layers, each layer consisting of two

**Figure 3**




The Asp-Glu-Xxx-Arg-His motif. The left panel illustrates the hydrogen bonding between the aspartic acid sidechain of one helix and the histidine and arginine of a second helix. The right panel illustrates two interlocking Asp-Glu-Xxx-Arg-His motifs interacting in a quasi-symmetrical manner. The backbone is drawn as black lines and the interacting residues are shown in stick representation. Residues are colored by atomic identity, as in Figure 2b,c. The positions of the iron atoms are shown as gray spheres.

**Table 2**

**Amino acid sequences of the helices surrounding the active sites of diiron proteins*.**

| | | | Heptad position | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Protein | Helix number | Residues | d | e | f | g | a | b | c | d | e | f | g | a |
| $\Delta^9$ ACP desaturase | H3 | 192–203 | T | S | F | Q | **E** | R | A | T | F | I | S | H |
| MMO | H1 | 110–121 | L | E | V | G | **E** | Y | N | A | I | A | A | T |
| R2 | H1 | 80–91 | Q | T | L | L | **D** | S | I | Q | G | R | S | P |
| R2 | H3 | 200–212 | V | N | A | L | **E** | A | I | R F | Y | V | S | F |
| Bacterioferritin | H1 | 47–58 | L | L | G | N | **E** | L | V | A | I | N | Q | Y |
| Bacterioferritin | H3 | 90–101 | D | L | A | L | **E** | L | D | G | A | K | N | L |
| Ferritin | H1 | 23–34 | Q | I | N | L | **E** | L | Y | A | S | Y | V | Y |
| Ferritin | H3 | 103–114 | A | L | H | L | **E** | K | N | V | N | Q | S | L |
| Rubrerythrin | H1 | 16–27 | A | F | A | G | **E** | S | Q | A | R | N | R | Y |
| Rubrerythrin | H3 | 90–102 | G | E | H | H | **E** | Y | T | E | Y T | E | M | Y |
| $\Delta^9$ ACP desaturase | H2 | 139–150 | W | T | A | E | **E** | N | R | **H** | G | E | L | L |
| $\Delta^9$ ACP desaturase | H4 | 222–233 | I | A | A | D | **E** | K | R | **H** | E | T | A | Y |
| MMO | H2 | 140–151 | Q | V | L | D | **E** | I | R | **H** | T | H | Q | C |
| MMO | H4 | 239–250 | I | E | T | D | **E** | L | R | **H** | M | A | N | G |
| R2 | H4 | 111–122 | W | A | F | S | **E** | T | I | **H** | S | R | S | Y |
| R2 | H3 | 234–245 | I | A | R | D | **E** | A | L | **H** | L | T | G | T |
| Bacterioferritin | H2 | 47–58 | E | S | I | D | **E** | M | K | **H** | A | D | R | |
| Bacterioferritin | H4 | 90–101 | I | L | R | D | **E** | L | G | **H** | I | D | W | L |
| Ferritin | H2 | 58–69 | Q | S | H | E | **E** | R | E | **H** | A | E | K | L |
| Ferritin | H4 | 137–148 | Y | L | N | E | **Q** | V | K | **A** | I | K | E | L |
| Rubrerythrin | H2 | 49–60 | I | A | D | Q | **E** | R | E | **H** | A | K | R | |
| Rubrerythrin | H4 | 124–135 | I | A | V | A | **E** | E | F | **H** | E | K | R | F |

*Amino acid sequences of the structurally homologous helices of the diiron proteins considered in this review. Chelating residues are in bold. Most of H3 of MMO and H1 of $\Delta^9$ ACP desaturase deviate from α-helical geometry and are not included. Only a 12-residue segment surrounding the active site is provided. The helices are grouped into two classes. The first consists of H1 and H3, which contain a chelating glutamic acid at the central 'a' position. The following 'd' position controls the access of oxygen and other substrates into the diiron site. These residues are generally small, except in R2, which has a large hydrophobic Phe208 residue that helps to stabilize the Tyr122 radical [48,49]. The second class of helices includes H2 and H4, which show the highly conserved Glu-Xxx-Xxx-His motif. The residues preceding the chelating glutamic acid are Asp>Glu>Ser, Gln, which frequently hydrogen bond to the chelating histidine of a neighboring helix.

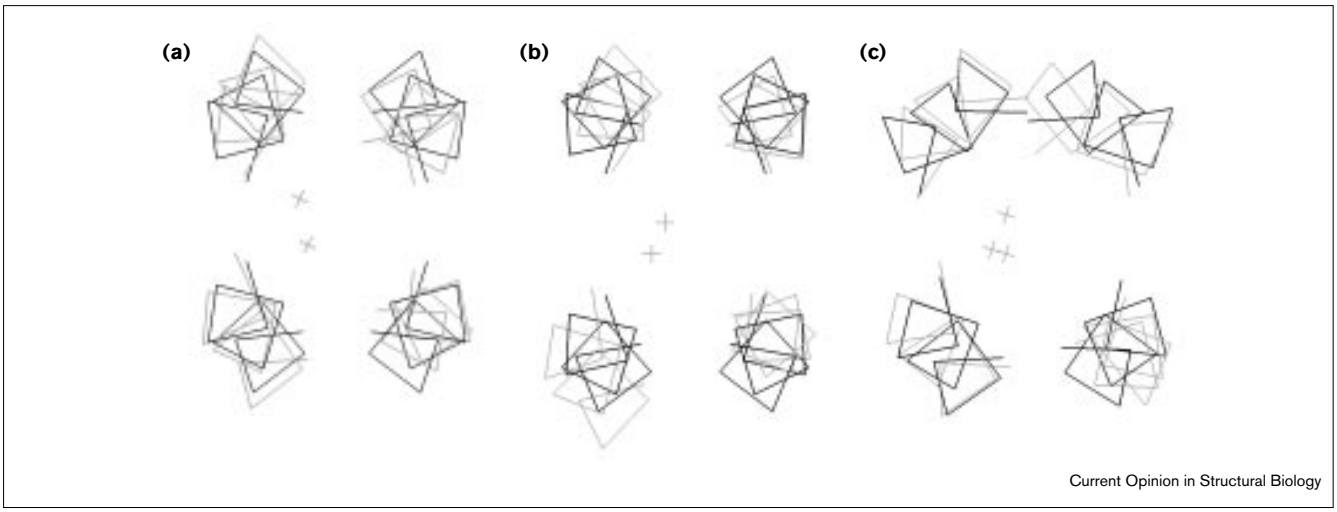
'a' residues and two 'd' residues (Figure 2c). The diiron site is located midway between the two central 'a/d' layers, labeled layers 2 and 3 in Figure 2c. The layers just above and below layers 2 and 3 define the top and the bottom of the active site; within a given functional family, these residues are often conserved. For example, in R2, a tyrosine residue in an 'a' position following the Glu-Xxx-Xxx-His motif forms the organic radical. Tyrosine residues also frequently occupy analogous 'a' positions in H1 and H2 of the ferritins and bacterioferritins. Interestingly, an organic radical has been detected in bacterioferritin using electron spin resonance (ESR) spectroscopy [35]. Also, the radical-forming tyrosine in H2 of R2 has a direct structural correlate in H4 of $\Delta^9$ ACP desaturase, although the role of radicals (if any) in this protein is presently unknown.

The sides of the active site are defined by residues at positions 'g' and 'c' on two faces of the four-helix bundle and 'b' and 'e' on the other two faces (Figure 2a,b). The residues along the b/e faces of the helical bundle tend to be tightly packed — often in an 'Ala-coil' interaction [36] — and appear to serve a structural role. By contrast, the residues along the c/g faces are particularly important for function. One c/g face contains a hydrogen-bonded network in MMO and $\Delta^9$ ACP desaturase, involving the sequence $Asp_g$-$Glu_a$-$Xxx_b$-$Arg_c$-$His_d$ in both H2 and H4 (Figure 3). In this motif, the $Asp_g$ of H2 accepts three hydrogen bonds from residues in H4 — two from $Arg_c$ and one from $His_d$. In a quasi-symmetrical manner, reciprocal interactions occur between the aspartic acid in H4 and the histidine and arginine in H2. R2 and bacterioferritin show a modification of this motif, in which an aspartic acid residue forms a hydrogen bond to the ligating histidine residue, without the accompanying arginine salt bridge. The g/c face on the opposite side of the active site contains residues from H1 and H3 that help to define the entry and binding of substrates. In most diiron proteins, the residues at one of the two 'g' positions of H1 and H3 serve to define a channel for the entry of the substrate and one or more sidechains at these positions tend to be small (glycine, alanine or serine).

## Parameterization of the backbone conformation

Weber and Salemme [29] have noted that the backbone atoms of four-helix bundle proteins may be reasonably described by a 222-symmetrical arrangement of equivalent helices. For convenience, we consider 222-symmetrical four-helix bundles in a Cartesian coordinate system, with

**Figure 4**




A superposition of the Cα traces of **(a)** bacterioferritin, **(b)** ferritin and **(c)** R2 with the 222-symmetrical bundles of best fit, obtained using the parameters described in Table 1. The traces from the X-ray/NMR structures are shown in gray, whereas the 222-symmetrical model trace is shown in black. The gray crosses illustrate the sites of the iron atoms. The bundles are oriented in the same manner as in Figure 2a.

the principal axis of the bundle coincident with the z axis and the other twofold symmetry axes directed along x and y. The six adjustable parameters defining the geometry of such a *222*-symmetrical bundle include the rotation of the monomer about its helical axis, its tilt and inclination relative to the z axis, and the displacement of its coordinates in x, y and z. For each of the active-site bundles, the values of these parameters were optimized using nonlinear least squares methods to evaluate the goodness of fit of the model four-helix bundles to the observed structures (Table 1).

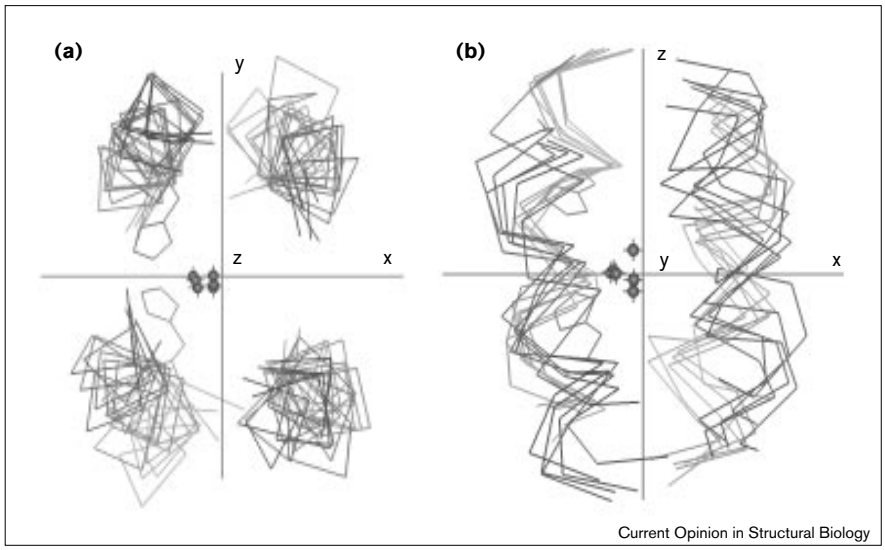
The backbone geometries of the active sites are remarkably well described by this simple *222*-symmetrical parameterization, with rmsds between each model and the appropriate experimental structure ranging from 0.9 to 1.4 Å. Also, the rmsd between a single consensus model and all six active-site structures was 1.8 Å. The models are closely related (Figure 4) and have a rectangular cross-section when viewed down the axis of the bundle. In each case, the structural b/c interface shows a shorter interhelical distance (4.1–4.7 Å, Table 1) than the functional g/c interface (5.4–6.4 Å). The helix is rotated such that the 'a' and 'd' positions are approximately equally accessible to the interior of the bundle, bringing the chelating glutamic acid and histidine residues from all four helices into close proximity. The helices have a left-handed tilt, relative to the z axis, as in classical four-helix bundle proteins. The symmetry axis relating any two antiparallel helices occurs between the active-site glutamic acid and histidine residues in the Glu-Xxx-Xxx-His motif, located approximately one residue beyond the chelating glutamic acid residue. Thus, chelating glutamic acid residues project from positions that are alternately approximately 1.5 Å above or below the x–y plane defined by the twofold symmetry axes of the models. The inclination of the helices, or γ angle, relative to the y–z plane, is very small and this parameter can be eliminated without significantly affecting the fit (changes in the rmsd between the models and the observed structure were less than 0.05 Å).

In conclusion, the active sites of these proteins comprise a relatively homogeneous set of helical bundles, which may be well described by a *222*-symmetrical model with five adjustable parameters. Antiparallel coiled coils were also considered, but showed a slightly poorer fit to the experimental structures. Also, less symmetrical models, such as dimers of asymmetrical dimers, failed to provide a significant improvement, unless an unduly large number of adjustable variables were included.

Having parameterized the backbone geometries of the active sites, it is interesting to observe the location of the center of mass of the diiron site relative to the origin of the Cartesian coordinate system (whose axes define the three symmetry axes of the model). For an idealized *222*-symmetrical bundle, one might expect the center of mass to be coincident with the origin. The center of mass of the diiron site indeed lies within approximately 1 Å of the x–y plane (Figure 5). The iron atoms are, however, displaced by approximately 1 Å towards the histidine ligands, which are asymmetrically located on helices *2* and 4. Interestingly, the metal ions in rubrerythrin show a smaller displacement, arising from an alternative coordination geometry involving a glutamic acid residue that is not observed in the other proteins.

**Figure 5**

A superposition of all six of the model 222-symmetrical four-helix bundles, viewed down **(a)** the z axis and **(b)** the y axis. The small spheres represent the positions of the centers of mass of the dinuclear centers. The locations of the iron centers were determined by first superimposing the actual three-dimensional structures onto the appropriate model structure. For reference, the positions of two histidine residues have been included in one model, to illustrate the movement of the metal centers away from the origin and towards the histidine ligands.




## Sidechain conformations

The finding of quasi-*222* symmetry in the backbone implies that the conformations and orientations of the four helices are approximately equivalent. This symmetry extends to the conformations of the chelating glutamic acid sidechains in the structures of the proteins in both the apo state and complexed with divalent metal ions. In apoferritin, in the diferrous forms of MMO [21] and R2 [22], and in the di-$Mn^{2+}$ forms of bacterioferritin [26] and R2 [25], a total of 16 out of 18 of the chelating glutamic acid residues adopt a ($g^+$,$g^+$) conformation (using the nomenclature of [37]). This clustering of rotamers is particularly significant given that only 15% of the glutamic acid residues in helices adopt the ($g^+$,$g^+$) rotamer.

The diferric states of MMO, R2 and rubrerythrin are less symmetrical, however, and show a greater range of sidechain conformations for the chelating glutamic acid residues — the number of sidechains in the ($g^+$,$g^+$), ($g^+$,t) and (t,$g^-$) conformations are 7, 2 and 2, respectively, out of a total of 11. These conformational changes allow carboxylate shifts, which are of importance for stabilizing the higher valency Fe(III) oxidation state of the metal ion. By contrast, the histidine sidechains show a single ($g^+$,t) conformation, irrespective of the oxidation state of the diiron centers. Again, this is significant given that only approximately 18% of the histidine residues in the helices adopt this conformation. Thus, the histidine residues appear to represent fixed points about which the iron atoms and carboxylates move in response to changes in oxidation state [11].

## Functional implications

The structural uniformity and symmetry of the active sites of diiron proteins belie their functional diversity. Subtle, but functionally important, variations in their geometries may be reflected in the parameters describing their geometries. The proteins with high ferroxidase activity (bacterioferritin, ferritin and rubrerythrin) tend to show smaller helical tilt angles and helical rotations than the more complex proteins that utilize oxygen for alkane oxidation, desaturation or radical formation.

A more clearly important mechanism for generating functional diversity from a simple primordial *222*-symmetrical bundle is to lift the symmetry through the introduction of asymmetric changes to the sequences of the individual helices. The most obvious deviation from *222*-symmetry involves the inclusion of a chelating histidine in only two of the four helices. Other sequence-specific variations are apparent in Table 2, which lists the sequences of the four-helix bundles in the vicinity of the active sites of this class of proteins. Beyond the chelating residues, there is no absolute conservation of the properties or identities of any other position in the sequence — a remarkable finding in view of the structural similarity of the six bundles. Indeed, even the chelating residues show some variation, which may be functionally significant. For example, H ferritin lacks one of its histidine ligands and one chelating glutamic acid is changed to glutamine. These changes may be important for decreasing the affinity of the protein for Fe(III), thereby facilitating the diffusion of the ion into the mineral core. Rubrerythrin shows an additional liganding glutamic acid in H3 at a 'd' position that is usually occupied by small residues in other proteins. In the structure of the diferric form of rubrerythrin, one ferric ion has moved away from its canonical location and the additional glutamic acid chelates in place of one of the histidine residues. It has been speculated that the iron atom moves to the neutral histidine sidechain in the diferrous form of the protein [27]. Finally, in R2, one of

**Table 3**

**Backbone parameterization of antiparallel four-helix bundles*.**

| Protein name | α angle | β angle | γ angle | X disp. | Y disp | Z disp | rmsd | Number of atoms superimposed |
|---|---|---|---|---|---|---|---|---|
| ROP[a] | 158.07 | −12.81 | −0.59 | 4.33 | 5.03 | 8.18 | 0.40 | 48 |
| DHP1[b] | 176.31 | −16.30 | 0.98 | 5.16 | 4.43 | 8.35 | 2.56 | 47 |
| α2D[c] | 181.80 | −4.45 | 7.52 | 4.71 | 4.08 | 6.24 | 1.21 | 48 |
| LacZ[d] | 140.47 | −8.82 | 3.74 | 3.96 | 5.19 | 7.41 | 0.36 | 48 |

*Structural parameterization of natural and *de novo* designed four-helix bundles. Using the same methods as in Table 1, it can be seen that the parametric fit of a 222-symmetrical model to several designed and natural four-helix bundles produces parameters similar to those seen in fits to the diiron sites above. ROP, or repressor of primer, and LacZ are naturally occurring four-helix bundles. The proteins DHP1 and α2D are four-helix bundles that are the products of *de novo* design efforts. (a) The PDB code for ROP is 1ROP [30], residues 15–26, 34–45, 108–119 and 141–152. (b) The PDB code for DHP1 is 4HBI [50], residues 11–22 and 104–115. Although DHP1 forms a twofold-symmetric homodimer, the fit to a 222-symmetrical tetramer is relatively poor, as can be seen from the high rmsd value. (c) The PDB code for α2D is 1QP6 [51], residues A3–A14, B3–B14, A22–A33 and B22–B33. Note that α2D is not a true antiparallel four-helix bundle; two of its helix–helix interfaces are antiparallel and two are parallel. This may explain the relatively poor best-fit rmsd using the 222-symmetrical model. (d) The PDB code for LacZ is 1LBI [52], residues 342–353.

the carboxylate ligands is Asp84, rather than a glutamic acid, as in the other proteins. This aspartic acid also comes within contact distance of the radical precursor Tyr122 in the diferrous form of the protein and the breakdown of the symmetric μ-1,2-peroxide intermediate is strongly retarded in the Asp84Glu mutant [22]. Other sequence-specific changes have been discussed above and are expected to be important to the creation of binding sites for tuning the redox properties of the metal ions and for positioning appropriate groups to facilitate radical formation and electron transfer.

Deviations from ideal α-helical geometry may provide yet another source of functional diversity in this class of proteins. One or more residues are inserted into one of the four active-site helices in each of the diiron proteins, except bacterioferritin. These insertions occur near the metal-ion-binding sites and cause local formation of mixed α/π-helical geometries. Indeed, H1 of $\Delta^9$ ACP desaturase and much of H3 of MMO are almost fully π helices. These deviations from α-helical geometry may be functionally important. For example, a switch in the redox state from ferric to ferrous in R2 causes very small changes in the α-helical portions of the active site (Table 1); however, large changes are observed in the geometries of residues 161 to 166 (of the 1R2F and 2R2F sequences [38]), which surround the inserted residue. Finally, it is important to emphasize that the active sites of these proteins are not isolated entities, but rather function within the context of much larger tertiary and quaternary structures.

## Implications for *de novo* protein design

An important goal of *de novo* protein design is the creation of model systems that are simpler in structure and function than their natural counterparts and, yet, nevertheless contain the essential features required for folding and function [34,39]. The availability of a parametric model for diiron active sites should greatly assist in the construction of mimics of this class of protein. Although the active sites of these proteins show a high degree of uniformity, their overall folds vary significantly. Thus, a variety of protein architectures may support the formation of this site, including simple homodimeric and homotetrameric four-helix bundles. We therefore examined the middle portions of the experimentally determined structures of several natural and designed four-helix bundles. These proteins included four-stranded antiparallel coiled coils, as well as classical four-helix bundles. Table 3 lists the parameters obtained by fitting their structures to the 222-symmetrical four-helix model described above. Significantly, the parameters observed for these four-helix bundles bracket the values obtained from fitting the diiron protein structures (Table 1). Thus, a variety of well-understood and easily designed proteins are capable of presenting the geometry required for diiron binding.

Perhaps the simplest system for investigating the design of these proteins would be a symmetrical homodimer of helix-loop-helix units. A prototype for this class of proteins is the RNA-binding protein ROP [30] and several groups have designed proteins that are intended to adopt this fold [40–43]. Furthermore, the principles required for the design of homodimeric and heterodimeric four-helix bundles with a variety of geometries are well understood [34,44]. Thus, a diiron site might be introduced into the structure of a highly simplified four-helix bundle. Such a simple protein would provide attractive opportunities for determining how systematic variations in amino acid sequence and bundle geometry affect the physical and spectroscopic properties of the dinuclear site and, ultimately, its function.

## References

1. McDonald DQ, Still WC: **Molecular mechanics parameters and conformational free energies of proline-containing peptides.** *J Org Chem* 1996, **61**:1385-1391.

2. Waller BJ, Lipscomb JD: **Dioxygen activation by enzymes containing binuclear non-heme iron clusters.** *Chem Rev* 1996, **96**:2625-2657.

3. Nordlund P, Sjöberg B-M, Eklund H: **Three-dimensional structure of the free radical protein of ribonucleotide reductase.** *Nature* 1990, **345**:593-598.
4. Lange SJ, Que L Jr: **Oxygen activating nonheme iron enzymes.** *Curr Opin Chem Biol* 1998, **2**:159-172.
5. Andersson KK, Gräslund A: **Diiron-oxygen proteins.** *Adv Inorganic Chem* 1995, **43**:359-403.
6. Stenkamp RE: **Dioxygen and hemerythrin.** *Chem Rev* 1994, **94**:715-726.
7. Powell AK: **Ferritin. Its mineralization.** *Met Ions Biol Syst* 1998, **35**:515-561.
8. Chasteen ND: **Ferritin. Uptake, storage, and release of iron.** *Met Ions Biol Syst* 1998, **35**:479-514.
9. Harrison PM, Hempstead PD, Artymiuk PJ, Andrews SC: **Structure-function relationships in the ferritins.** *Met Ions Biol Syst* 1998, **35**:435-477.
10. Briat JF, Lobreaux S: **Iron storage and ferritin in plants.** *Met Ions Biol Syst* 1998, **35**:563-584.
11. Nordlund P, Eklund H: **Di-iron-carboxylate proteins.** *Curr Opin Struct Biol* 1995, **5**:758-766.
12. Gallagher SC, George A, Dalton H: **Sequence-alignment modelling and molecular docking studies of the epoxygenase component of alkene monooxygenase from *Nocardia corallina* B- 276.** *Eur J Biochem* 1998, **254**:480-489.
13. Bouvier F, Keller Y, d'Harlingue A, Camara B: **Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum L.*).** *Biochim Biophys Acta* 1998, **1391**:320-328.
14. Lee M, Lenman M, Banas A, Bator M, Singh S, Schweizer M, Nilsson R, Liljenberg C, Dahlquist A, Gummeson PD *et al.*: **Identification of non-heme diiron proteins that catalyze triple bond and epoxy group formation.** *Science* 1998, **280**:915-918.
15. Bollinger JM Jr, Tong WH, Ravi N, Huynh BH, Edmondson DE, Stubbe JA: **Use of rapid kinetics methods to study the assembly of the diferric-tyrosyl radical cofactor of *E. coli* ribonucleotide reductase.** *Methods Enzymol* 1995, **258**:278-303.
16. Broadwater JA, Ai J, Loehr TM, Sanders-Loehr J, Fox BG : **Peroxodiferric intermediate of stearoyl-acyl carrier protein delta 9 desaturase: oxidase reactivity during single turnover and implications for the mechanism of desaturation.** *Biochemistry* 1998, **37**:14664-14671.
17. Möenne-Loccoz P, Baldwin J, Ley BA, Loehr TM, Bollinger JM Jr: **$O_2$ activation by non-heme diiron proteins: identification of a symmetric μ-1,2-peroxide in a mutant of ribonucleotide reductase.** *Biochemistry* 1998, **37**:14659-14663.
18. Yang X, Chen-Barrett Y, Arosio P, Chasteen ND: **Reaction paths of iron oxidation and hydrolysis in horse spleen and recombinant human ferritins.** *Biochemistry* 1998, **37**:9743-9750.
19. Möenne-Loccoz P, Krebs C, Herlihy K, Edmondson DE, Theil EC, Huynh BE, Loehr TM: **The ferroxidase reaction of ferritin reveals a diferric mu-1,2 bridging peroxide intermediate in common with other $O_2$-activating non-heme diiron proteins.** *Biochemistry* 1999, **38**:5290-5295.
20. Shu L, Nesheim JC, Kauffmann K, Münch E, Lipscomb JD, Que L Jr : **An $Fe_2^{IV}O_2$ diamond core structure for the key intermediate Q of methane monooxygenase.** *Science* 1997, **275**:515-518.
21. Rosenzweig AC, Nordlund P, Takahara PM, Frederick CA, Lippard SJ: **Geometry of the soluble methane monooxygenase catalytic diiron center in two oxidation states.** *Chem Biol* 1995, **2**:409-418.
22. Logan DT, Su XD, Aberg A, Regnström K, Hajdu J, Eklund H, Nordland P: **Crystal structure of reduced protein R2 of ribonucleotide reductase: the structural basis for oxygen activation at a dinuclear iron site.** *Structure* 1996, **4**:1053-1064.
23. Andersson ME, Högbom M, Rinaldo-Matthis A, Andersson KK, Sjöberg BM, Nordlund P: **The crystal structure of an azide complex of the diferrous R2 subunit of ribonucleotide reductase displays a novel carboxylate shift with important mechanistic implications for diiron-catalyzed oxygen activation.** *J Am Chem Soc* 1999, **121**:2346-2352.
24. Lindqvist Y, Huang W, Schneider G, Shanklin J: **Crystal structure of $\Delta^9$ stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins.** *EMBO J* 1996, **15**:4081-4092.
25. Atta M, Nordlund P, Aberg A, Eklund H, Fontecave M: **Substitution of manganese for iron in ribonucleotide reductase from *Escherichia coli*. Spectroscopic and crystallographic characterization.** *J Biol Chem* 1992, **267**:20682-20688.
26. Frolow F, Kalb (Gilboa) AJ, Yariv J: **Structure of a unique twofold symmetric haem-binding site.** *Nat Struct Biol* 1994, **1**:453-460.
27. deMare F, Kurtz DM Jr, Nordlund P: **The structure of *Desulfovibrio vulgaris* rubrerythrin reveals a unique combination of rubredoxin-like FeS4 and ferritin-like diiron domains.** *Nat Struct Biol* 1996, **3**:539-546.
28. Lawson DM, Artymiuk PJ, Yewdall SJ, Smith JM, Livingstone JC, Treffry A, Luzzago A, Levi S, Arosio P, Cesareni G *et al.*: **Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts.** *Nature* 1991, **349**:541-544.
29. Weber PC, Salemme FR: **Structural and functional diversity in 4-α-helical proteins.** *Nature* 1980, **287**:82-84.
30. Banner DW, Kokkinidis M, Tsernoglou D: **Structure of ColE1 rop protein at 1.7 Å resolution.** *J Mol Biol* 1987, **196**:657-675.
31. Crick FHC: **The Fourier transform of a coiled-coil.** *Acta Crystallogr* 1953, **6**:685-689.
32. Crick FHC: **The packing of alpha-helices: simple coiled-coils.** *Acta Crystallogr* 1953, **6**:689-697.
33. Cohen C, Parry DAD: **α-Helical coiled coils and bundles: how to design an α-helical protein.** *Proteins* 1990, **7**:1-15.
34. Bryson JW, Betz SF, Lu HS, Suich DJ, Zhou HY, O'Neil KT, DeGrado WF: **Protein design: a hierarchic approach.** *Science* 1995, **270**:935-941.
35. Cheesman MR, le Brun NE, Kadir FH, Thomson AJ, Moure GR, Andrews SC, Guest JR, Harrison PM, Smith JM, Yewdall SJ: **Haem and non-haem iron sites in *Escherichia coli* bacterioferritin: spectroscopic and model building studies.** *Biochem J* 1993, **292**:47-56.
36. Gernert KM, Surles MC, Labean TH, Richardson JS, Richardson DC: **The Alacoil: a very tight, antiparallel coiled-coil of helices.** *Protein Sci* 1995, **4**:2252-2260.
37. Schrauber H, Eisenhaber F, Argos P: **Rotamers: to be or not to be?** *J Mol Biol* 1993, **230**:592-612.
38. Eriksson M, Jordan A, Eklund H: **Structure of *Salmonella typhimurium* nrdF ribonucleotide reductase in its oxidized and reduced forms.** *Biochemistry* 1998, **37**:13359-13369.
39. DeGrado W, Summa CM, Pavone V, Nastri F, Lombardi A: ***De novo* design and structural characterization of proteins and metalloproteins.** *Annu Rev Biochem* 1999, **68**:779-819.
40. Betz SF, Liebman PA, DeGrado WF: ***De novo* design of native proteins: characterization of proteins intended to fold into antiparallel, rop-like, four-helix bundles.** *Biochemistry* 1997, **36**:2450-2458.
41. Predki PF, Regan L: **Redesigning the topology of a four-helix bundle protein: monomeric Rop.** *Biochemistry* 1995, **34**:9834-9839.
42. Vlassi M, Steif C, Weber P, Tsernoglou D, Wilson KS, Hinz HJ, Kokkinidis M: **Restored heptad pattern continuity does not alter the folding of a four-alpha-helix bundle.** *Nat Struct Biol* 1994, **1**:706-716.
43. Munson M, O'Brien R, Sturtevant JM, Regan L: **Redesigning the hydrophobic core of a four-helix bundle protein.** *Protein Sci* 1994, **3**:2015-2022.
44. Betz SF, Bryson JW, DeGrado WF: **Native-like and structurally characterized designed α-helical bundles.** *Curr Opin Struct Biol* 1995, **5**:457-463.
45. Hempstead PD, Yewdall SJ, Fernie AR, Lawson DM, Artymiuk PJ, Rice DW, Ford GC, Harrison PM: **Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution.** *J Mol Biol* 1997, **268**:424-448.

46. Nordlund P, Eklund H: **Structure and function of the *Escherichia coli* ribonucleotide reductase protein R2.** *J Mol Biol* 1993, **232**:123-164.

47. Rosenzweig AC, Brandstetter H, Whittington DA, Nordlund P, Lippard SJ, Frederick CA: **Crystal structures of the methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath): implications for substrate gating and component interactions.** *Proteins* 1997, **29**:141-152.

48. Logan DT, deMaré F, Persson BO, Slaby A, Sjöberg BM, Nordlund P: **Crystal structures of two self-hydroxylating ribonucleotide reductase protein R2 mutants: structural basis for the oxygen-insertion step of hydroxylation reactions catalyzed by diiron proteins.** *Biochemistry* 1998, **37**:10798-10807.

49. Ormö M, Regnström K, Wang Z, Que L Jr, Sahlin M, Sjöberg BM: **Residues important for radical stability in ribonucleotide reductase from *Escherichia coli*.** *J Biol Chem* 1995, **270**:6570-6576.

50. Schafmeister CE, LaPorte SL, Miercke LJ, Stroud RM: **A designed four helix bundle protein with native-like structure.** *Nat Struct Biol* 1997, **4**:1039-1046.

51. Hill RB, DeGrado WF: **Solution structure of α2D, a nativelike *de novo* designed protein.** *J Am Chem Soc* 1998, **120**:1138-1145.

52. Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, Schumacher MA, Brennan RG, Lu P: **Crystal structure of the lactose operon repressor and its complexes with DNA and inducer.** *Science* 1996, **271**:1247-1254.