doi:10.1016/S0022-2836(02)00589-2 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 321, 923–938

JMB

Computational *de novo* Design, and Characterization of an A_2B_2 Diiron Protein

Christopher M. Summa†, Michael M. Rosenblatt†, Jae-Kyoung Hong James D. Lear and William F. DeGrado*

Department of Biochemistry and Biophysics School of Medicine The University of Pennsylvania 1010 Stellar-Chance Bldg 421 Curie Blvd, Philadelphia PA 19104-6059, USA

*Corresponding author

Diiron proteins are found throughout nature and have a diverse range of functions; proteins in this class include methane monooxygenase, ribonucleotide reductase, Δ^9-acyl carrier protein desaturase, rubrerythrin, hemerythrin, and the ferritins. Although each of these proteins has a very different overall fold, in every case the diiron active site is situated within a four-helix bundle. Additionally, nearly all of these proteins have a conserved Glu-Xxx-Xxx-His motif on two of the four helices with the Glu and His residues ligating the iron atoms. Intriguingly, subtle differences in the active site can result in a wide variety of functions. To probe the structural basis for this diversity, we designed an A_2B_2 heterotetrameric four-helix bundle with an active site similar to those found in the naturally occurring diiron proteins. A novel computational approach was developed for the design, which considers the energy of not only the desired fold but also alternatively folded structures. Circular dichroism spectroscopy, analytical ultracentrifugation, and thermal unfolding studies indicate that the A and B peptides specifically associate to form an A_2B_2 heterotetramer. Further, the protein binds Zn(II) and Co(II) in the expected manner and shows ferroxidase activity under single turnover conditions.

Keywords: *de novo* design; coiled-coil; computational design; heterotetramer; retrostructural analysis

Introduction

The design of proteins *de novo*[1,2] has progressed significantly within the last decade. While the design of native-like proteins is by no means a routine endeavor, sufficient progress has been made that attention is beginning to focus on the design of functional proteins. One aspect of function currently being explored involves the design of a binding interaction between a *de novo* protein and a cofactor[3–13] or another protein.[14,15] A far more complex and subtle design objective is the rational design of *de novo* proteins with catalytic properties.

With the exception of some elegant demonstrations of the generation of novel enzyme activities through the mutagenesis of a naturally occurring protein,[6,16–26] most research seeking novel enzymatic activities has focused on combinatorial approaches such as *in vitro* evolution,[27–35] domain shuffling,[36] incremental truncation,[37] and phage display methods.[38–46] Recently, catalytic activities have been reported in some *de novo* designed systems.[47–49]

We have focussed on the design of diiron-binding proteins because of their extraordinary functional diversity. Representative members of this class of proteins include methane monooxygenase (MMO),[50–57] ribonucleotide reductase (RR),[58,59] Δ^9-acyl carrier protein desaturase,[60,61] hemerythrin,[62–64] rubrerythrin,[65] and the ferritins.[66] These proteins function as monooxygenases, radical generators, oxidoreductases, desaturases, ferroxidases, and oxygen and iron storage proteins. While the overall folds of the proteins differ, the structural subunits responsible for forming the diiron sites are remarkably similar four-helix bundles.[67,68] One of the goals of this project is to identify the determinants of metal binding and substrate activation through the study of simplified

† These authors contributed equally to this work.

Abbreviations used: MMO, methane monooxygenase; CD, circular dichroism.

E-mail address of the corresponding author: wdegrado@mail.med.upenn.edu

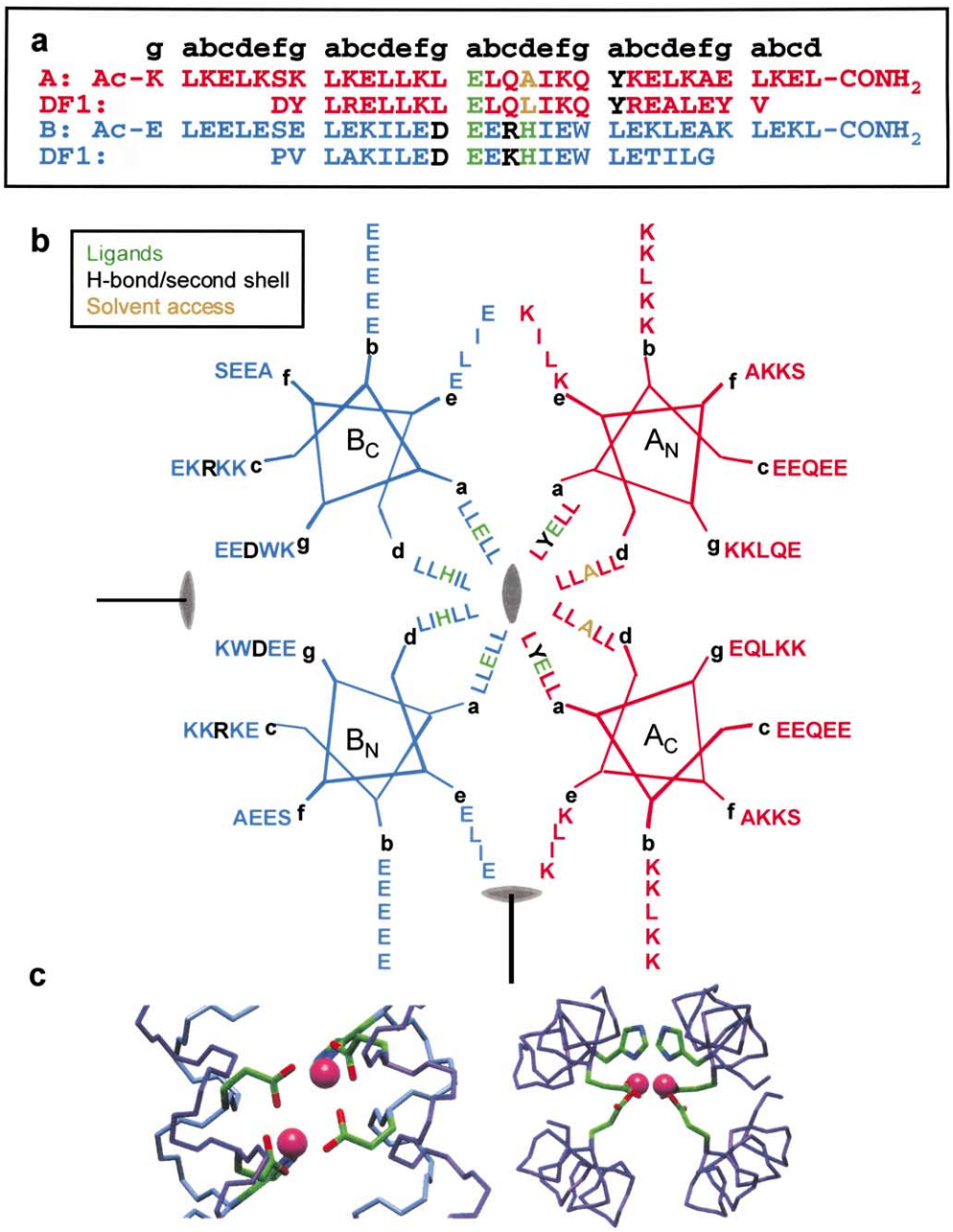


Figure 1. Sequence, helical wheel diagram, and active site of DFtet. (a) The sequence of DF1 is shown along with DFtet. DF1 has a helix-turn-helix topology; a KL link between the helices has been omitted from (a). The ligating residues (green), those residues that form second-shell hydrogen bonds to the ligating residues (black), and a residue that is important for solvent access (brown) are shown in the linear sequence and the helical wheel diagram (b). The position of the 2-fold axis running through the dimetal site is shown at the left oval. The positions of the additional quasi 2-fold axis relating the backbone atoms are shown at the center and bottom of the diagram. (c) Two views of the dimetal active site of di-Zn(II) DF1.

dimetal-binding systems. Interestingly, recent protein engineering studies of the diiron site in natural ribonucleotide reductase R2 subunits have resulted in the production of new catalytic intermediates.[69−71]

In previous work, antiparallel homo-dimeric four-helix bundle proteins were designed using a helix-turn-helix motif that binds metal ions including Zn, Fe, and Co ions.[47,72] At the active site of these proteins are four glutamate and two histidine residues, including two conserved Glu-Xxx-Xxx-His motifs found in nearly all diiron proteins of this class. These proteins were produced using previously established methods employing symmetry operators to produce C_n and D_n-symmetric backbone geometries.[68] Here, we have taken the conserved metal binding site of our DueFerro1 protein (DF1) and modified it to form a hetero-tetrameric assembly, DFtet.

DFtet was designed to self-assemble into a

heterotetrameric A_2B_2 bundle from four isolated peptide units.[73,74] The protein assembles with a precise topology allowing for the generation of a dimetal binding site by utilizing residues from each of the four independent peptides. Thus, a heterotetrameric system can bridge the gap between "rational design" and combinatorial approaches. However, while a heterotetrameric system is uniquely suited to the production of a library of proteins from a significantly smaller number of peptides, the design of such a system presents significant challenges for protein design. Specifying a unique topology is a more complex problem for a heterotetrameric system, as in DFtet, than it was in a homodimeric system, such as DF1 or DF2. In particular, it was important to destabilize both homooligomeric folds as well as undesired heterotetrameric topologies. Therefore, new computational design methods were used to solve this problem. While the concept of "negative design" (or designing against alternative structures) has been considered in the past[75–79] (and see DeGrado *et al.*[1] and references therein), this idea has not been explicitly codified into a computational design algorithm that was tested experimentally.

Results

Computational design of DFtet

DueFerro tetramer, or DFtet, is a set of two 33-residue peptides that have been designed to self-associate into a dimetal-binding A_2B_2 heterotetrameric protein. Because the target was a heterotetramer, we used negative design, to prevent alternate topologies from occurring. Also, we used a "minimalist" approach[80,81] in order to minimize extraneous structural variables, which might complicate the design and interpretation of results. The protein sequence was intended to be as simple as possible, while retaining the ability to tetramerize specifically into the desired topology and bind two metal ions.

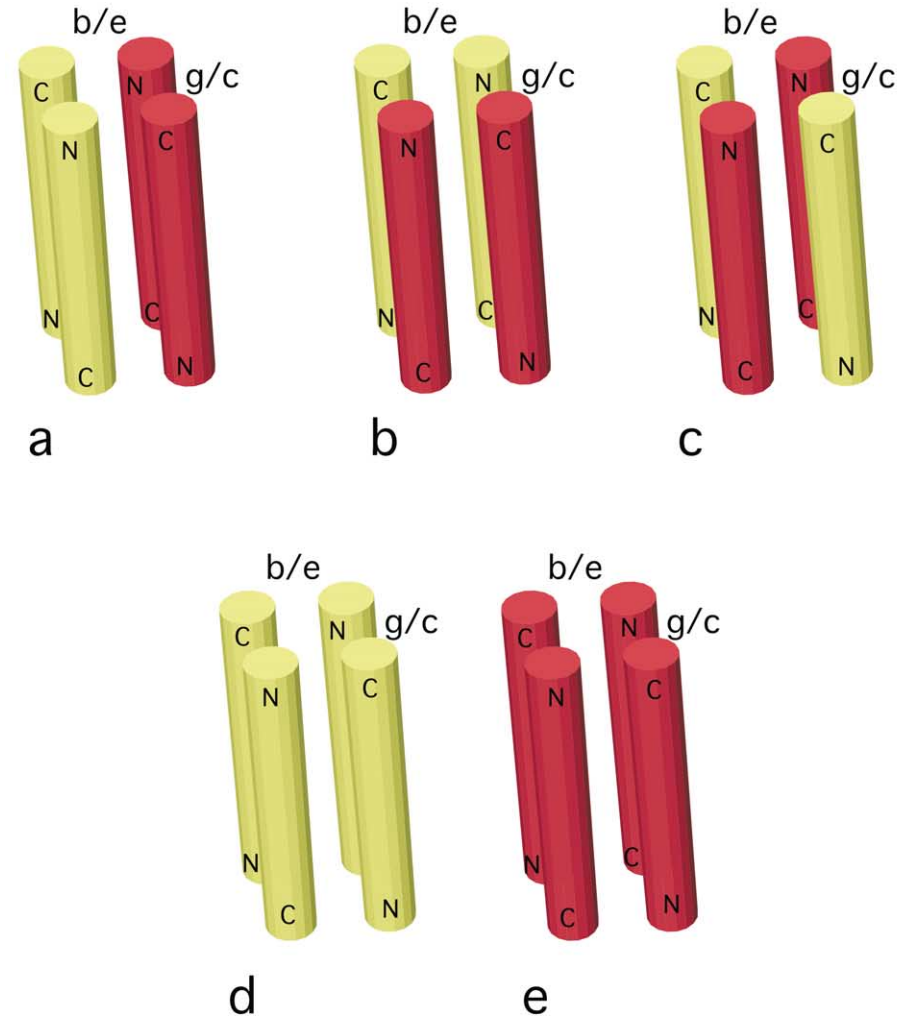
The first step in the design process involved the specification of the backbone. Although there are two different chains in DFtet, the backbone structure shows approximate D_2 symmetry with three orthogonal 2-fold rotational axes as shown in Figure 1. The parameters used to generate the four-helix bundle backbone of DFtet were obtained from an analysis of DF1, as described in Materials and Methods. The helices of DFtet were extended relative to those of DF1 (33 residues in DFtet *versus* 24 residues in DF1) to increase the stability of the DFtet system by increasing the size of the hydrophobic core.

While the backbone conformation of DFtet is approximately D_2 symmetric, the metal-binding site shows lower symmetry (C_2), because only two of the four helices donate His side-chains to the binding site (Figure 1(b) and (c)). Thus, in the second stage of the design, side-chains were asymmetrically placed onto the helices, resulting in an A_2B_2 heterotetramer with C_2 symmetry. The designation A was assigned to the peptide containing only the Glu ligands, and the designation B was assigned to the peptide containing the EXXH metal-binding motif. The residues within the region of the binding site are the same as the corresponding residues of DF1. Specifically, residues A12–A23 of DFtet are identical with residues 6–17 of DF1, and residues B12–B23 of DFtet correspond to residues 32–43 of DF1, with two exceptions. Residue 22 in DFtet-A was mutated from Leu (residue 13 in DF1) to Ala because Leu19 blocks access to the active site in the crystal structure of DF1.[47,72] Second, residue 18 in Dftet-B (corresponding to residue 38 of DF1) was mutated from Lys to Arg. Arg at this position might form a hydrogen bond with Asp15, which itself hydrogen bonds directly to the coordinating histidine. This Asp-Glu-Xxx-Arg-His motif is seen in diiron proteins such as MMO[82] and Δ^9-ACP desaturase;[60] the Arg might position the liganding His residue[83–85] and possibly also tune the redox potential of the metal site.[86,87]

The ligand-binding site, together with a requirement for low-energy side-chain rotamers entirely defined the backbone conformation. Residues at the remaining *a* and *d* positions were modeled as leucine because this side-chain effectively filled the interior volume of the bundle. This side-chain also has a high intrinsic helical propensity[88] and is known to stabilize antiparallel four-helix bundles, such as ROP.[77,89] Residues at *f* positions outside the vicinity of the binding site were chosen to maximize the overall helical propensity of the sequence, while minimizing hydrophobicity and electrostatic interactions with the interfacial residues. These positions were modeled as Ser or Ala in an alternating manner along the chain.

The nature of the residues at the remaining *e*, *g*, *b*, and *c* positions were chosen to specifically stabilize only one of the possible topologies for an antiparallel A_2B_2 heterotetramer (Figure 2). The desired topology is shown in Figure 2(a) with two alternate topologies (shown in Figure 2(b) and (c)) that may be considered. Clearly, there are other possible configurations for the protein than those shown, such as parallel structures or antiparallel structures with frame shifts along the central axis of the bundle. We made the assumption that these frame shifts would not occur, because of the energetic cost of exposing excess non-polar surface area. Also, parallel tetramers are possible, although modeling suggests that they would result in distorted metal-binding sites, which lacked some of the liganding interactions. In Figure 2(a), the *b*/*e* interface lies at the helix A/helix B boundary and the *g*/*c* interface lies between both the A/A and the B/B boundaries. The reverse situation holds for the other topology (Figure 2(b)); in Figure 2(c) both the *b*/*e* and *g*/*c* interfaces lie between helices A and B.

Interacting interfacial positions were allowed to



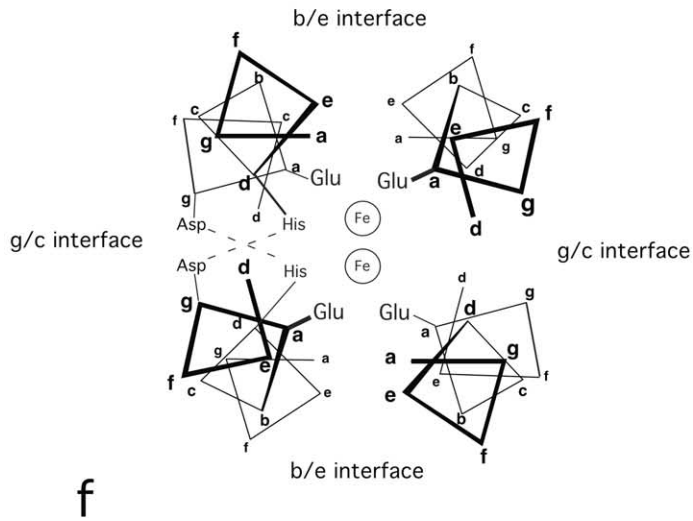


Figure 2. Possible topologies for an antiparallel tetrameric coiled-coil. The upper panel shows three different possibilities for an A_2B_2 heterotetramer. Helices of like color are identical (red denotes an A-helix and yellow denotes a B-helix) and the directionality of the chain is denoted by the markings denoting the N and C termini. The topology in (a) is the desired topology, having g/c interfaces between helices of identical sequence, and b/e interfaces between helices of different sequence. The topologies shown in (b) and (c) are undesirable in the present study, and a computational design algorithm was implemented to prevent these topologies from occurring in the design of DFtet. In (b), the g/c interface now lies between helices of different sequence, and the b/e interface lies between helices with identical sequence, opposite to what is found in (a). The topology shown in (c) is an intermediate between the previous two examples; both g/c and b/e interfaces occur between helices which differ in sequence. Configurations (d) and (e) show what would occur if a homotetramer were allowed to form. (f) The different interfaces of the intended antiparallel heterotetramer in detail.

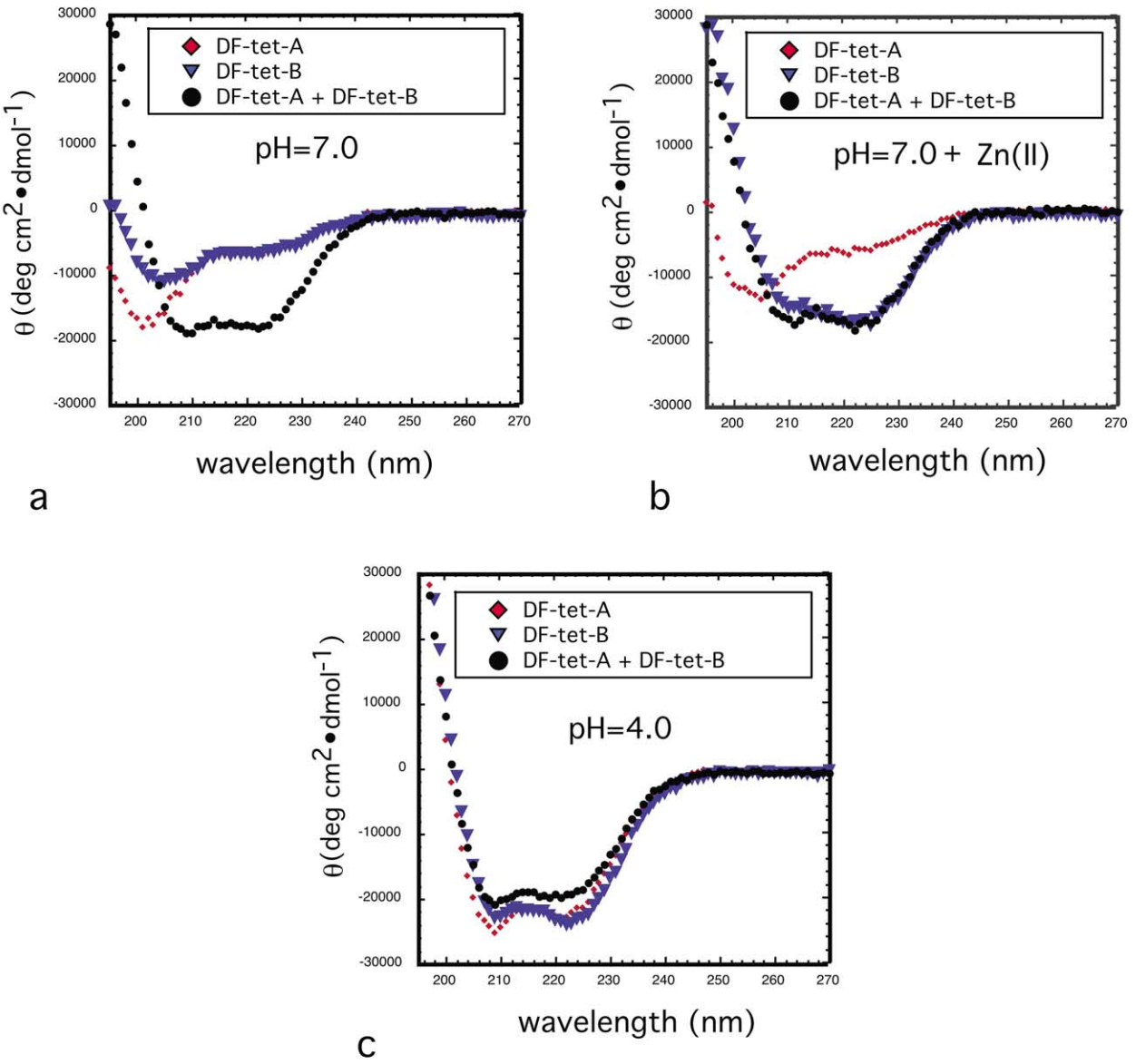


Figure 3. CD spectra of DFtet are shown at various pH values. (a) Spectra at pH 7.0 in the absence of exogenous metal, (b) spectra at pH 7.0 in the presence of one equivalent of Zn(II) and (c) spectra at pH 4.0 in the absence of exogenous metal. In each panel, red diamonds denote spectra of DFtet-A alone, blue triangles denote spectra of DFtet-B alone, and filled black circles show spectra for a 1:1 mixture of DFtet-A and DFtet-B.

assume an identity of either + or −. These designations would ultimately correspond to residue identities of Lys and Glu, respectively, in the amino acid sequence. We chose not to allow neutral or non-polar residues at these positions in order to maximize solubility in aqueous solution. Interacting pairs were scored as follows:

$$E_{\text{contact}} = \begin{bmatrix} +2, & +/+ \text{ interaction} \\ +3, & -/- \text{ interaction} \\ -1, & +/- \text{ interaction} \end{bmatrix} \quad (1)$$

A larger unfavorable "energy" for −/− over +/+ interactions was used because Glu has shorter side-chains than Lys, and hence is less able to maximize the distance between one another. A pair of residues was labeled interacting if the distance between their C^{α} atoms was less than 8.5 Å. These scores are a simplified representation of experimentally measured interhelical electrostatic interaction energies.[90] Since the energies were intended only as an approximation to find an optimal pattern of charge, and not to calculate binding energies, this simplification was reasonable.

The desired topology and a single alternate topology (Figure 2(b)) were explicitly considered in the calculation of the energy function. The energy function used was:

$$E_{\text{tot}} = E_{\text{desired}} - E_{\text{undesired}} \quad (2)$$

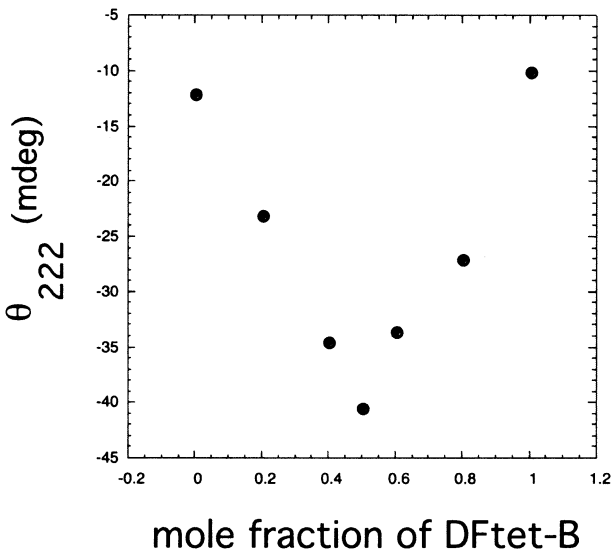


Figure 4. Titration of DFtet-A into DFtet-B measured by the mean residue ellipticity of the solution at 222 nm. The molar ratio of the two peptides was varied while keeping the total concentration constant at 5.0 μM (25 mM Mops, pH 7.0).

where:

$$E_x = \sum_{i=1}^{N_{\text{contacts}}} E_{\text{contact}}(i) \qquad (3)$$

where x designates a particular three-dimensional conformation for the sequence. The identities of all the variable positions were initially set to 0, representing no charge. At each step in the optimization protocol, a residue was chosen at random, and its identity was flipped to that of the opposite sign (or, in initial cases where the identity was unsigned, to +). At this point, the interaction energies for the new sequence were calculated. If the new sequence was lower than the lowest energy sequence yet encountered, the program would retain the sequence and its energy value. This process was run through 700,000 iterations for the DFtet sequence, and three-dimensional models were built for the sequences with the top four scores. A sequence was chosen (ranked second among all the top scorers) which showed the minimum number of unfavorable contacts in the desired topology (among the top four considered) and is shown in Figure 1. Visual inspection also showed that these sequences would show multiple unfavorable electrostatic interactions as a parallel homo or heterotetramers.

Solution characteristics of the protein

CD spectroscopy

At pH 7.0, the circular dichroic (CD) spectra of DFtet showed strong bands at 222 nm and 208 nm, indicating a high content of α-helical secondary structure (Figure 3(a)). By contrast, the individual DFtet-A and DFtet-B peptides showed a spectrum typical of a random-coil. This finding suggested that the DFtet-A and DFtet-B peptides are unstructured in the absence of their partners, and were able to form stable secondary structure only when mixed together at this pH. In the case of peptide DFtet-A, formation of a homooligomer at pH 7.0 would require the burial of one active-site glutamic acid residue per peptide. This is clearly an energetically unfavorable process, and likely inhibited association. A homooligomer of peptide DFtet-B at this pH would involve the burial of one His and one Glu per peptide.

To confirm that DFtet-A and DFtet-B peptides associate with a 1:1 stoichiometry, different molar ratios of the peptides were mixed and their signal at 222 nm was evaluated. A minimum (signifying maximal helical content) occurred at a molar ratio of precisely 0.5, indicating that the stoichiometry of the peptides was 1:1 (Figure 4).[91–93]

The next series of experiments were undertaken to determine the effects of transition metal ions on the secondary structure of the DFtet at neutral pH. Zn(II) was chosen for these experiments, because it binds strongly to the active sites of other diiron proteins.[47,72,94] The secondary structure of DFtet was not significantly altered by the addition of Zn(II) (Figure 3(b)). Similarly, the addition of Zn(II) had little effect on the secondary structure of the DFtet-A peptide, presumably because Zn(II) ions are rarely coordinated by carboxylate groups in the absence of other liganding groups in protein structures. DFtet-B, on the other hand, did appear to form significant secondary structure in the presence of the Zn(II) at pH 7.0, which was presumably due to coordination of the EXXH motifs. Histidine is commonly found to be an active-site ligand in Zn(II) proteins[95] and Zn(II)-assisted protein folding has been reviewed recently.[96]

Protonation of the active-site glutamate residues has been shown to stabilize the homodimeric DF2,[47] an analogous designed protein. Thus, we were interested to determine the effect of low pH on the individual components of the DFtet system. At pH 4.0, below the unperturbed pK_a of the Glu side-chain, DFtet-A, DFtet-B and a 1:1 molar ratio of these two peptides showed a spectrum typical of the α-helix (Figure 3(c)).

Sedimentation equilibrium ultracentrifugation

Equilibrium analytical ultracentrifugation was used to determine the molecular mass of DFtet. The proteins were centrifuged at 40,000, 45,000, and 48,000 rpm, and the data were globally analyzed to determine the mass averaged molecular mass. Analysis of a 1:1 mixture of DFtet-A and DFtet-B, at pH 7.0 indicated the formation of a tetramer with an apparent molecular mass of 15,400 Da (Figure 5(a)). Addition of one equivalent of Zn(II) to the 1:1 mixture resulted in an improved fit to the data (Figure 5(b)), and an observed

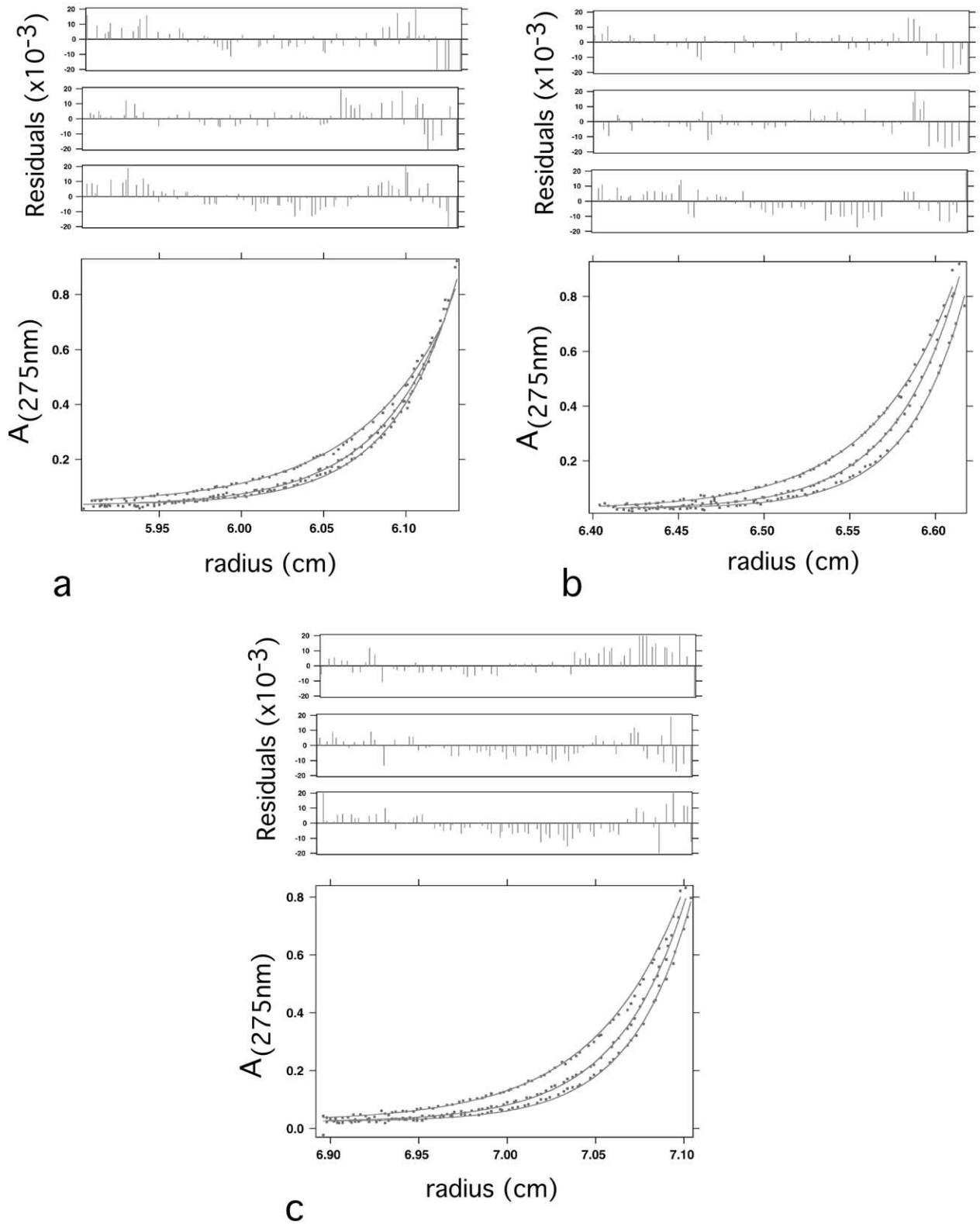


Figure 5. Equilibrium ultracentrifugation analysis of (a) DFtet-A + DFtet-B at pH 7.0 in the apo state, (b) in the presence of Zn(II) at pH 7.0, and (c) at pH 4.0 in the apo state. Data from 40,000, 45,000, and 48,000 rpm were used in a simultaneous fit of a fixed molecular mass model.

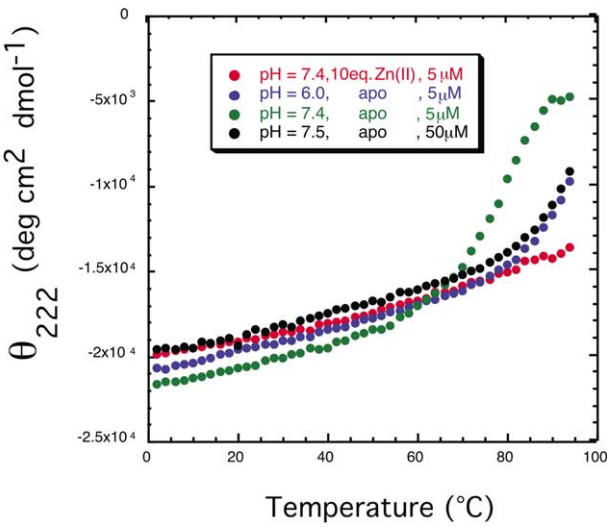


Figure 6. Thermal denaturation curves of DFtet with and without added metal.

molecular mass (16,500 Da) in excellent agreement with that expected for the tetramer 16,400 Da.

In contrast, DFtet-A was monomeric even at concentrations tenfold above the CD measurements (data not shown). The average molecular mass for DFtet-B was between that expected for a dimer and trimer, suggesting an intermediate degree of aggregation (data not shown).

The pH-dependence of the oligomerization state was also assessed using ultracentrifugation. Data collected for DFtet-A alone at pH 4.0 were well described by a fit to a single species with an apparent molecular mass of 11,500 Da, close to that expected for a trimer (data not shown). Numerous coiled-coil peptides have previously been reported to form trimers.[97] A monomer–tetramer equilibrium model provided the best fit to the sedimentation data collected for DFtet-B; attempts to use a single-species model resulted in poor fits. DFtet (a 1:1 mixture of DFtet-A and DFtet-B) data were best described by a fit to a single tetrameric species at pH 4.0 (Figure 5(c)), with an apparent molecular mass of 15,000 Da (theoretical mass is 16,292 Da).

Thermal unfolding

After confirming that the A_2B_2 system was indeed helical and tetrameric at neutral pH, the thermal stability of the designed protein was investigated. Figure 6 illustrates the thermal stability of the apo-protein at 5 μM and 50 μM total concentration near pH 7.5. As expected for a self-associating system, the protein was more stable at a higher concentration. Even at 5 μM, however, the protein was very stable and unfolded only at elevated temperatures, with a midpoint near 75 °C. Addition of Zn(II) led to a dramatic increase in stability, with the protein becoming entirely stable up 95 °C. A comparison of the unfolding curves measured at pH 7.4 *versus* 6.0 (5 μM protein concentration in each case), also showed the anticipated pH dependence. At pH 6.0 the energetic cost associated with proton condensation at the active site[47] was less unfavorable, and the stability of the protein was increased.

Co(II) binding

Co(II) provided an especially convenient spectroscopic probe since it has d–d transitions in the visible region (between 500 nm and 700 nm),[98] whose molar extinction coefficient depend on the coordination environment of the Co(II). The extinction coefficient increases from 10 $M^{-1}\,cm^{-1}$ to ~150 $M^{-1}\,cm^{-1}$ to ~400–600 $M^{-1}\,cm^{-1}$ as the coordination proceeds from octahedral to pentacoordinate to tetrahedral. Addition of $CoCl_2$ to a solution of the DFtet (36 μM in tetramer) produced a species with an absorption spectrum that is characteristic of Co(II) in a pentacoordinate environment. This result agrees with the design in which each Co(II) ion is ligated by a total of four carboxylate ligands (two from the two bridging glutamate residues and two from a chelating glutamate) and a single His ligand. Furthermore, the shape of the Co(II) spectrum (Figure 7 inset) was identical with that of DF2,[47] and the extinction coefficients were the same within experimental error (DF2: $\lambda = 520$ nm, $\varepsilon = 140\ M^{-1}\,cm^{-1}$; $\lambda = 550$ nm, $\varepsilon = 155\ M^{-1}\,cm^{-1}$; $\lambda = 600$ nm, $\varepsilon = 90\ M^{-1}\,cm^{-1}$; DFtet: $\lambda = 520$ nm, $\varepsilon = 125\ M^{-1}\,cm^{-1}$; $\lambda = 550$ nm, $\varepsilon = 140\ M^{-1}\,cm^{-1}$; $\lambda = 600$ nm, $\varepsilon = 80\ M^{-1}\,cm^{-1}$). The spectrum also showed a very close correspondence with that of bacterioferritin ($\lambda = 520$ nm, $\epsilon = 126\ M^{-1}\,cm^{-1}$; $\lambda = 555$ nm, $\varepsilon = 155\ M^{-1}\,cm^{-1}$; $\lambda = 600$ nm, $\varepsilon = 107\ M^{-1}\,cm^{-1}$),[94,99] whose active site resembles our intended design.[100] The correspondence of the spectra strongly suggests that the metal-binding site of DFtet adopts the intended geometry.

The stoichiometry of binding was probed by measuring the absorbance at 550 nm as a function of added Co(II) at a constant tetramer concentration of 36 μM (Figure 7). A linear increase in absorbance was observed, until a stoichiometry of two Co(II) per tetramer was reached. Beyond this point, no further significant increase in absorbance beyond that expected for hexa-aqua Co(II) was observed. These data indicate that the binding was tight and stoichiometric, and that the dissociation constant for binding to both sites was significantly lower than the total concentration of binding sites.

Ferroxidase reaction

A number of diiron proteins, such as bacterioferritin and rubrerythrin, catalyze the oxidation of Fe(II) to Fe(III). In these proteins, this ferroxidase reaction[65,101–104] results in the formation of a diferric oxo-bridged cluster characterized by a broad charge transfer transition centered near 300 nm.[105,106] The time-course of Fe(II) oxidation in the presence of DFtet (50 μM in tetramer, Figure

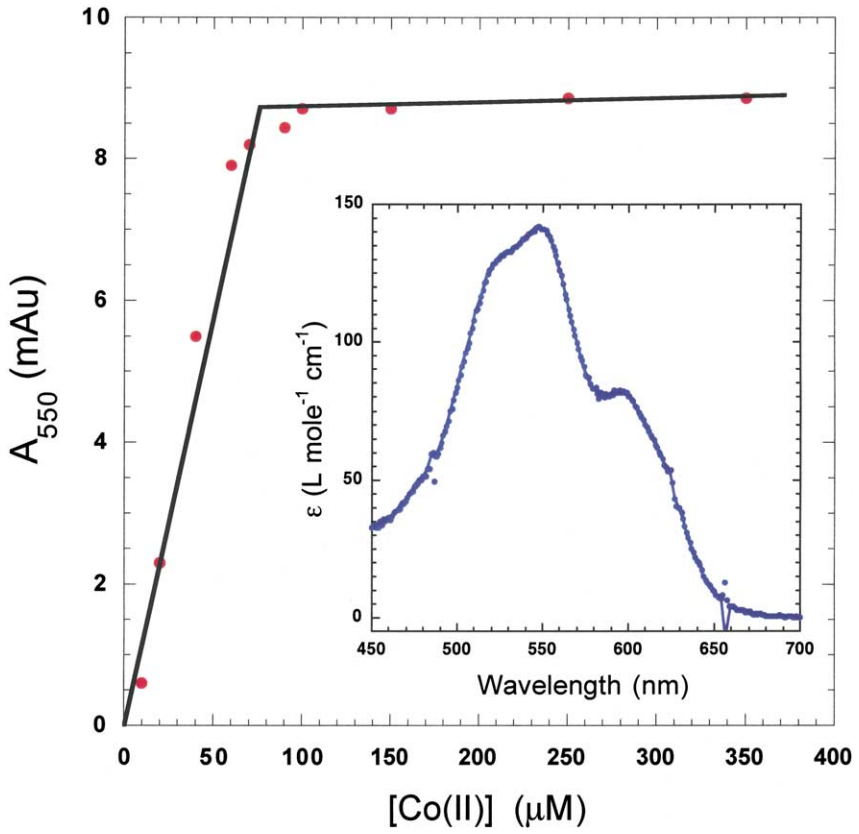


Figure 7. Co(II) titration of 36 μM DFtet-A_2B_2 (total tetramer concentration). The absorbance of free Co(II) has been subtracted. The two lines were obtained from a linear regression of the points up to two equivalents and beyond two equivalents of Co(II). The lines intersect at a concentration of 78 μM, *versus* the value of 72 μM expected for a stoichiometry of two Co(II) per tetramer. The inset shows the spectrum of a stoichiometric complex of $Co(II)_2$DFtet–A_2B_2.

8(a)) was monitored by the absorbance at 320 nm. Under single turnover conditions, the reactivity of the DFtet was of the same order of magnitude as DF2;[47] both were significantly more rapid than the uncatalyzed reaction, but nevertheless were about three orders of magnitude slower than the reaction catalyzed by bacterioferritin.

The ferroxidase reaction catalyzed by DF2 was first-order in both protein and Fe(II). Similarly, the initial rate of iron oxidation by DFtet depended linearly on the concentration of Fe(II) (Figure 8(b)). Furthermore, the individual time-courses at various iron concentrations successfully globally fit to an integrated rate equation which is first-order in both iron and protein (Figure 8(c)). The rate constant determined in this manner was 11.8 M^{-1} s^{-1}, and the extinction coefficient at 320 nm for the dimeric diiron center was 3730 M^{-1} cm^{-1}, which was within the range of extinction coefficients observed for diiron proteins including bacterioferritin[106] (3380 M^{-1} cm^{-1} at 300 nm), H-chain ferritin[104] (2990 M^{-1} cm^{-1} at 300 nm), horse spleen ferritin[104] (3540 M^{-1} cm^{-1} at 300 nm), ribonucleotide reductase[107](4700 M^{-1} cm^{-1} at 325 nm), and stearoyl-ACP Δ^9 desaturase[108] (2080 M^{-1} cm^{-1} at 325 nm).

Discussion

Backbone design

The method of retrostructural analysis[72] allows the construction of a consensus backbone from a group of related protein structures. More importantly, it provides data with which to set ranges of backbone parameters for a protein within that class. With this method it should be possible to access global parameters (such as helical tilt angle, interhelical distance) and test their effect on function in a way that was heretofore impossible with single mutation-based approaches. These design techniques will be used in future work on this system.

Sequence design

Three criteria must be satisfied before a design can be successful. First, the target structure must be the lowest-energy structure for the designed sequence. Second, there should exist a large energy gap between the sequence folded into the desired structure and that of any other, undesired structure. Finally, the ground state configuration should be non-degenerate. The last two criteria ensure that the desired structure will be significantly populated relative to any alternative conformations.

Historically, computational protein design approaches have focused on satisfaction of the first criterion alone.[109,110] This method has, in general, proven successful in many real-world protein designs. Gutin & Shakhnovich[111] have also shown that for a lattice model of a heteropolymer (based on the HP-model designed by Dill and co-workers) the probability of finding a degenerate

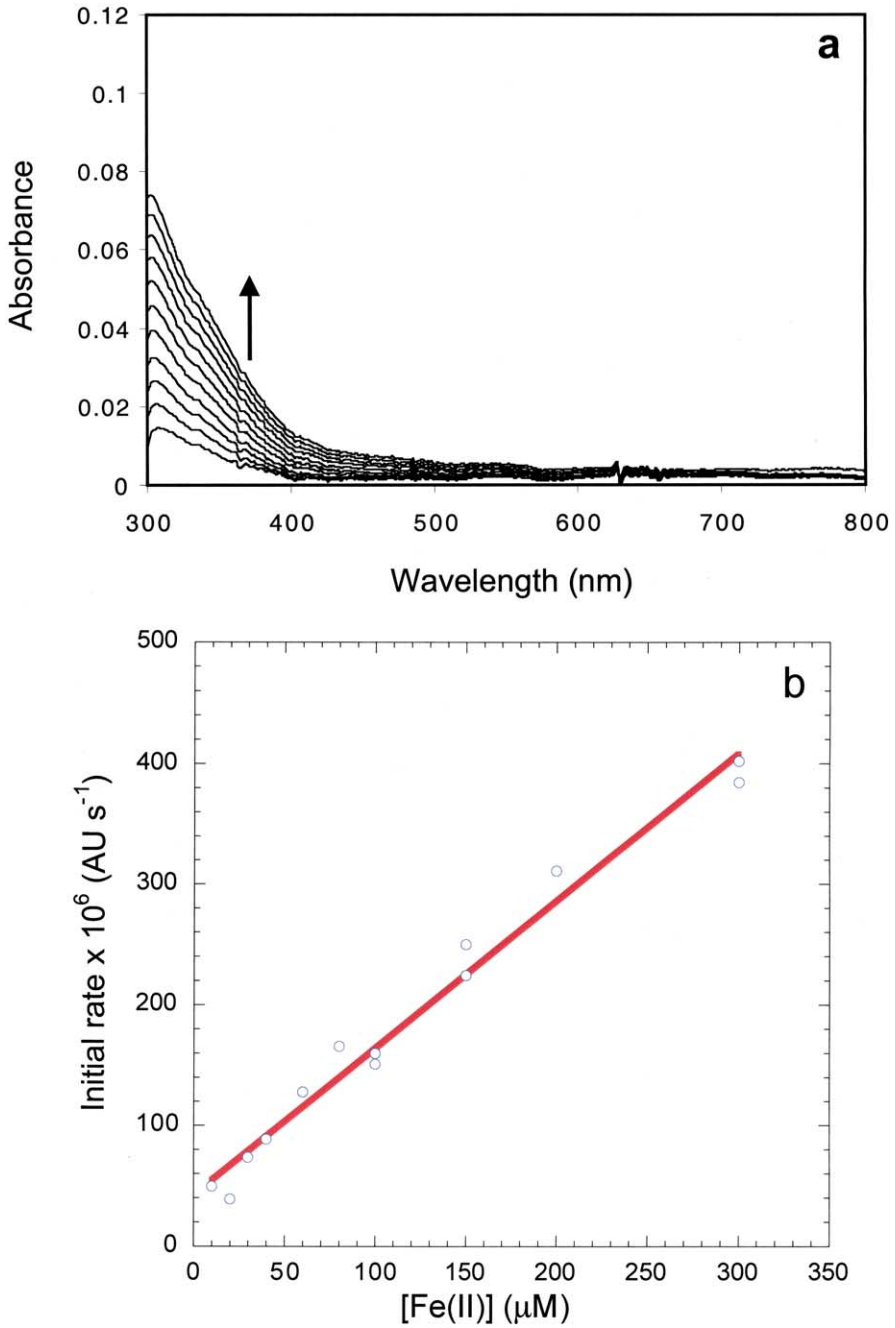


Figure 8 (a) and (b) (*legend opposite*)

ground state decreases exponentially with a decrease in the ground state energy. Therefore, the lower the energy of the design, the more likely that the designed structure will be non-degenerate.

However, it appeared unlikely that the present problem could be solved by simply optimizing the energetics of the intended structure. Shakhnovich & Gutin have suggested that explicit consideration of alternative configurations was an essential part of the optimization process for designing a uniquely folded protein.[111,112] Since exhaustively searching configurational space is impossible for all but the shortest peptide sequences, implementing such a design protocol in real-world protein design applications is non-trivial. In our case, we had a specific alternative configuration in mind; this may not be true of all protein design applications.

In our method, there were 28 unique residues which underwent optimization and each was allowed an identity of + or −. This represents 2^{28} or roughly 268 million possible sequences. While it would have been impossible to enumerate each possible sequence for comparison, our goal was to

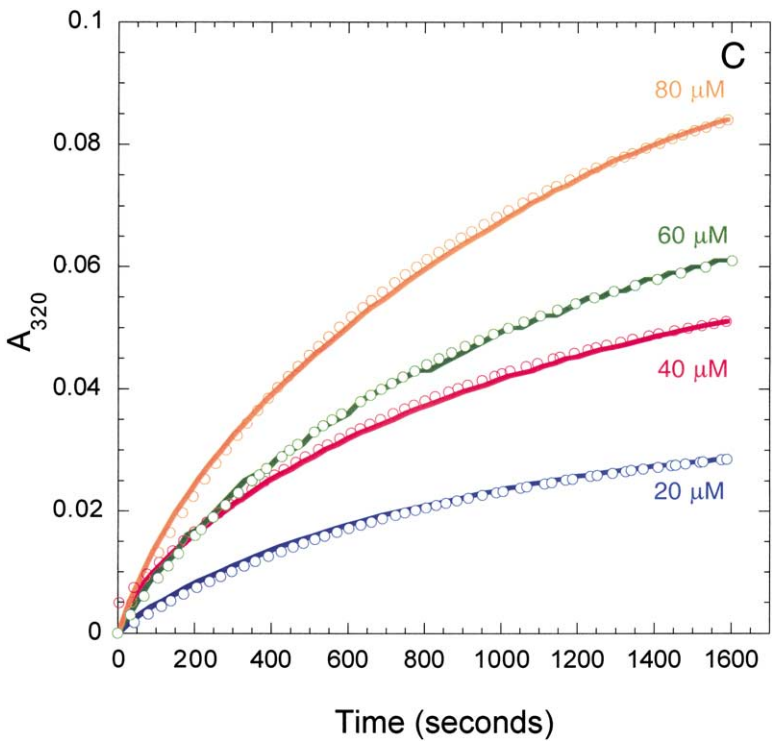


Figure 8. DFtet-catalyzed oxidation of Fe(II) to Fe(III). (a) The time-course of the ferroxidase reaction is followed by measuring the absorbance of the solution at 320 nm. (b) The initial rates are shown as a function of the initial Fe(II) concentrations at a total concentration of 50 μM tetramer. (c) Time-courses for the reaction (50 μM tetramer); the data were globally fit to a second-order rate equation $((1/([P]_i - [Fe^{2+}]_i)\ \{\ln([Fe^{2+}]_i/([Fe^{2+}]_i - [Fe^{3+}])) - \ln([P]_i/([P]_i - [Fe^{3+}]))\} = kt)$ in which $[P]_i$ is the initial protein concentration, $[Fe^{2+}]_i$ is the initial concentration of Fe(II), and $[Fe^{3+}] = (A_{320} - A_{320\ \text{initial}})/\Delta\varepsilon_{320}$. The rate constant was globally fit and found to be 11.8(±0.8) $M^{-1}\,s^{-1}$. The extinction coefficient at 320 nm was allowed to vary from run to run, but was nevertheless found to be approximately constant with a value of 1862(±153) $M^{-1}\,cm^{-1}$ (3729 M^{-1} for the dimeric diiron center). Concentrations of initial $[Fe^{2+}]$ are listed above each curve and are color-coded.

develop methodology that could be used in other design projects; this energy function was optimized with a random search algorithm. A simulated annealing algorithm, a genetic algorithm, or other optimization protocol could also be used successfully.

Here, a single alternative conformation was used as the "unwanted" configuration. Homooligomeric species of DFtet-A and DFtet-B were not considered in the sequence design process. As can be seen from the CD and sedimentation results on the isolated peptides, DFtet-B is capable of forming a homooligomer in the presence of Zn(II) in solution. Ideally, both peptides could be designed while taking the homooligomeric states into account as undesired alternate topologies. While this optimization alone is unlikely to completely abolish homooligomer formation in the isolated peptides in solution (due to the favorable energetics of hydrophobic core burial), it may be possible to significantly destabilize these states through the use of carefully designed interhelical electrostatic interactions.

Solution characteristics

The design of DFtet was successful: DFtet-A and DFtet-B are unfolded at neutral pH; however, when mixed in a 1:1 ratio they assemble into a helical bundle with the expected stoichiometry. Furthermore the peptide is tetrameric and stable over a wide temperature range, being thermally unfolded only at low concentrations and at an elevated pH. The main transitions of the thermal unfolding curves are steep, as would be expected for a native protein. Only at low pH do the individual peptides associate into helical oligomers, as expected. DFtet also binds metal ions in the proper stoichiometry, and the spectroscopic properties of the Co(II) and Fe(III) derivatives strongly suggest that the ligand environment is precisely as intended. The match of the Co(II) spectrum between DFtet and DF2 (whose crystal structure has been determined, unpublished results), and the identical kinetic scheme for Fe(II) oxidation provide good evidence that the metal-binding sites of these proteins are very similar. Thus, the intended topology appears to have been achieved.

Materials and Methods

Automated design

The parameters used to define the structure of DF1 were determined by fitting to the crystal structure of DF1[72] using a genetic algorithm and equations similar to those employed previously.[68,77,81] The best-fit D_2 symmetric coiled-coil parameters (Figure 9, legend) were used to produce a model whose backbone trace was within 1.1 Å r.m.s.d. over 68 C^{α} atoms of the template backbone. These parameters were then used to create an elongated bundle as follows: because the intended product was a coiled-coil model, the α-helix was slightly overwound to give a precise seven residue repeat; next the bundle was built using the parameters described above, and the entire bundle was then

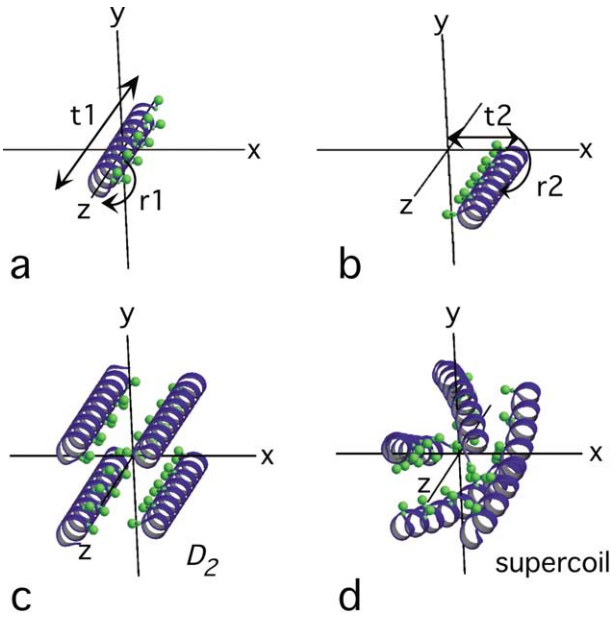


Figure 9. A D_2 symmetric coiled-coil can be fully specified by five parameters: the superhelical pitch, two rigid body translations, and two Eulerian rotations. An idealized α-helix was positioned with its axis coincident with the *z*-axis in Cartesian coordinate space. It was then given a right-handed twist (a). The "unwound" helix is then moved from the axis and positioned in Cartesian space (b) such that application of the D_2 symmetry operator would produce an antiparallel bundle (c) with the hydrophobic and liganding residues (shown in green) buried in the protein core. The supercoil operator was then re-applied to the entire bundle (this time with a left-handed twist), producing a coiled-coil (d). Values used for the generation of the DFtet coiled-coil model were: $r1 = -103.6°$, $t1 = 1.3$ Å, $t2 = 7.3$ Å, $r2 = -48.0°$, supercoil pitch = 190 Å. This Figure was generated using the programs MOLSCRIPT[113] and Raster3D.[114]

"rewound" using an idealized left-handed twist with a superhelical pitch of 190 Å.[68,77,81] The backbone was then relaxed to reduce strain *via* energy minimization with the molecular mechanics program Discover (Molecular Simulations, Inc., San Diego, CA) using the CVFF force field. Although the final model of DFtet was 33 residues in length, in the initial steps of model building three residues were added at the N terminus and C terminus to minimize end effects.

The computer program written to perform automated sequence design of the solvent-exposed residues of DFtet was implemented in F77 on an Apple Powerbook G3 running the LinuxPPC operating system. This same software package has been compiled and run on an SGI Indigo2 computer running the IRIX operating system.

Peptide synthesis and purification

Fmoc-protected amino acid residues (Fmoc, 9-fluorenylmethoxycarbonyl) and PAL resin were purchased from Applied Biosystems. HOBt (*N*-hydroxybenzotriazole) and HBTU (2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) were purchased from Novabiochem.

Peptides were synthesized on a 0.25 mmol scale using an Applied Biosystems model 433A solid-phase peptide synthesizer. Standard coupling conditions were used.[14] In order to minimize failure sequences, "capping" with acetic anhydride (100-fold excess) and DIEA (*N*,*N*-diisopropylethylamine) in NMP (*N*-methylpyrrolidinone) was performed after each coupling reaction. All peptides were acetylated at the amino terminus.

DFtet-A was cleaved using a mixture of 95.0% TFA (trifluoroacetic acid), 2.5% water, 2.5% TIS (triisopropylsilane)(by vol.). DFtet-B was cleaved with 81.5% TFA, 5% thioanisole, 5% phenol, 5% water, 2.5% EDT (ethanedithiol), 1% TIS. Both reactions were run at room temperature, under nitrogen, for two hours. The TFA was evaporated and the resins were precipitated with cold diethyl ether, centrifuged, and washed three times with ether. The mixture was extracted with water (80%)/acetonitrile (19.9%)/TFA(0.1%), centrifuged, and the supernatant lyophilized. The peptides were then purified by reverse-phase HPLC on a preparative C_4 column (Vydac, 2.2 cm × 25 cm; 10 μm particle size) and were determined to be at least 95% pure by analytical HPLC and MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometry.

CD spectroscopy

All CD measurements were performed on an AVIV 62A DS spectropolarimeter. Samples containing 5–30 μM of peptide were prepared in either 25 mM sodium acetate (pH 4.0) or 25 mM Mops (pH 7.0) and 100 mM NaCl. Stoichiometric amounts of $ZnCl_2$ were also added to selected samples and allowed to equilibrate overnight. All spectra were measured at 25 °C and were an average of six to eight scans. Thermal denaturation studies were performed by averaging the signal for 30 seconds, after five minutes of equilibration, and monitoring the signal at 222 nm between 2 °C and 94 °C.

Sedimentation analysis

Sedimentation equilibrium data were collected at 25 °C on a Beckman XLI analytical ultracentrifuge equipped with both adsorption and interference optics. Peptide concentrations were determined from either the tryptophan ($\varepsilon_{280} = 5500\ M^{-1}\ cm^{-1}$) or the tyrosine absorbance ($\varepsilon_{275} = 1490\ M^{-1}\ cm^{-1}$). Five samples of isolated A and B peptides and a 1:1 mixture (apo; Fe(II); Zn(II)) were prepared at 100 μM at pH 4.0 and 7.0 in the same buffers as used for the CD experiment. Samples were spun at three speeds (40,000, 45,000, and 48,000 rpm) and the data were analyzed using an in-house fitting routine in IGOR Pro (WaveMetrics, Inc.). Partial specific volumes were estimated from amino acid composition.

Metal ion-binding studies

Co(II) binding was conducted at 36 μM tetramer concentration using a Hewlett Packard model 8453 diode array spectrometer as described.[47] $CoCl_2$ was added in 0.2 eq (Co(II)/protein) increments to a solution of DFtet, and spectral data ranging from 400 to 800 nm were collected. To ensure proper equilibration, individual samples were allowed to equilibrate at room temperature for at least 12 hours.

Fe(II) oxidation studies

Kinetic studies of Fe(II) oxidation were conducted at 50 μM protein concentration as described.[47]

Acknowledgments

The authors thank Stephen Lippard for helpful discussion. This work was supported by the National Institutes of Health (grant no. GM54616) and the MRSEC program of NSF (award no. DMR0079909). M.M.R. was supported, in part, by an NRSA from the National Institutes of Health (grant no. GM63421-02).

References

1. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. (1999). *De novo* design and structural characterization of proteins and metalloproteins. *Annu. Rev. Biochem.* **68**, 779–819.
2. Moffet, D. A. & Hecht, M. H. (2001). *De novo* proteins from combinatorial libraries. *Chem. Rev.* **101**, 3191–3203.
3. Wisz, M. S., Garrett, C. Z. & Hellinga, H. W. (1998). Construction of a family of Cys_2His_2 zinc binding sites in the hydrophobic core of thioredoxin by structure-based design. *Biochemistry*, **37**, 8269–8277.
4. Hellinga, H. W. & Richards, F. M. (1991). Construction of new ligand-binding sites in proteins of known structure I. Computer-aided modeling of sites with pre-defined geometry. *J. Mol. Biol.* **222**, 763–785.
5. Coldren, C., Hellinga, H. W. & Caradonna, J. P. (1997). The rational design and construction of a cuboidal iron–sulfur protein. *Proc. Natl Acad. Sci. USA*, **94**, 6635–6640.
6. Benson, D. E., Wisz, M. S., Liu, W. & Hellinga, H. W. (1998). Construction of a novel redox protein by rational design: conversion of a disulfide bridge into a mononuclear iron–sulfur center. *Biochemistry*, **37**, 7070–7076.
7. Choma, C. T., Lear, J. D., Nelson, M. J., Dutton, P. L., Robertson, D. E. & DeGrado, W. F. (1994). Design of a heme-binding four-helix bundle. *J. Am. Chem. Soc.* **116**, 856–865.
8. Gibney, B. R., Mulholland, S. E., Rabanal, F. & Dutton, P. L. (1996). Ferredoxin and ferredoxin-heme maquettes. *Proc. Natl Acad. Sci. USA*, **93**, 15041–15046.
9. Johansson, J. S., Gibney, B. R., Rabanal, F., Reddy, K. S. & Dutton, P. L. (1998). A designed cavity in the hydrophobic core of a four-alpha-helix bundle improves volatile anesthetic binding affinity. *Biochemistry*, **37**, 1421–1429.
10. Mulholland, S. E., Gibney, B. R., Rabanal, F. & Dutton, P. L. (1999). Determination of nonligand amino acids critical to [4Fe–4S]2 + /+ assembly in ferredoxin maquettes. *Biochemistry*, **38**, 10442–10448.
11. Rabanal, F., DeGrado, W. F. & Dutton, P. L. (1996). Toward the synthesis of a photosynthetic reaction center maquette: a cofacial porphyrin pair assembled between two subunits of a synthetic four-helix bundle multiheme protein. *J. Am. Chem. Soc.* **118**, 473–474.
12. Robertson, D. E., Farrid, R. S., Moser, C. C., Urbauer, J. L., Mulholland, S. E., Pidikiti, R. *et al.* (1994). Design and synthesis of multi-haem proteins. *Nature*, **368**, 425–432.
13. Huffman, D. L., Rosenblatt, M. M. & Suslick, K. S. (1998). Synthetic heme-peptide complexes. *J. Am. Chem. Soc.* **120**, 6183–6184.
14. Ghirlanda, G., Lear, J. D., Lombardi, A. & DeGrado, W. F. (1998). From synthetic coiled coils to functional proteins: automated design of a receptor for the calmodulin-binding domain of calcineurin. *J. Mol. Biol.* **281**, 379–391.
15. Root, M. J., Kay, M. S. & Kim, P. S. (2001). Protein design of an HIV-1 entry inhibitor. *Science*, **291**, 884–888.
16. Pinto, A., Hellinga, H. W. & Caradonna, J. P. (1997). Construction of a catalytically active iron superoxide dismutase by rational design. *Proc. Natl Acad. Sci. USA*, **94**, 5562–5567.
17. Jez, J. M. & Penning, T. M. (1998). Engineering steroid 5 β-reductase activity into rat liver 3 α-hydroxysteroid dehydrogenase. *Biochemistry*, **37**, 9695–9703.
18. Erlanson, D. A., Braisted, A. C., Raphael, D. R., Randal, M., Stroud, R. M., Gordon, E. M. & Wells, J. A. (2000). Site-directed ligand discovery. *Proc. Natl Acad. Sci. USA*, **97**, 9367–9372.
19. Gerlt, J. A. (1994). Protein engineering to study enzyme catalytic mechanisms. *Curr. Opin. Struct. Biol.* **4**, 593–597.
20. Douglas, K. T. (1992). Alteration of enzyme specificity and catalysis. *Curr. Opin. Biotechnol.* **3**, 370–377.
21. Haering, D. & Distefano, M. D. (2001). Enzymes by design: chemogenetic assembly of transamination active sites containing lysine residues for covalent catalysis. *Bioconj. Chem.* **12**, 385–390.
22. Magnusson, A., Hult, K. & Holmquist, M. (2001). Creation of an enantioselective hydrolase by engineered substrate-assisted catalysis. *J. Am. Chem. Soc.* **123**, 4354–4355.
23. Hornung, E., Rosahl, S., Kuhn, H. & Feussner, I. (2000). Creating lipoxygenases with new positional specificities by site-directed mutagenesis. *Biochem. Soc. Trans.* **28**, 825–826.
24. Nilsson, L. O., Gustafsson, A. & Mannervik, B. (2000). Redesign of substrate-selectivity determining modules of glutathione transferase A1-1 installs high catalytic efficiency with toxic alkenal products of lipid peroxidation. *Proc. Natl Acad. Sci. USA*, **97**, 9408–9412.
25. Hennecke, J. & Glockshuber, R. (1998). Conversion of a catalytic into a structural disulfide bond by circular permutation. *Biochemistry*, **37**, 17590–17597.
26. Distefano, M. D., Kuang, H., Qi, D. & Mazhary, A. (1998). The design of protein-based catalysts using semisynthetic methods. *Curr. Opin. Struct. Biol.* **8**, 459–465.
27. Arnold, F. H. & Volkov, A. A. (1999). Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* **3**, 54–59.
28. Arnold, F. H., Wintrode, P. L., Miyazaki, K. & Gershenson, A. (2001). How enzymes adapt: lessons from directed evolution. *Trends Biochem. Sci.* **26**, 100–106.
29. Bornscheuer, U. T. & Pohl, M. (2001). Improved biocatalysis by directed evolution and rational protein design. *Curr. Opin. Chem. Biol.* **5**, 137–143.
30. Jermutus, L., Honegger, F., Schwesinger, F., Hanes,

J. & Pluckthun, A. (2001). Tailoring *in vitro* evolution for protein affinity or stability. *Proc. Natl Acad. Sci. USA*, **98**, 75–80.
31. Kuchner, O. & Arnold, F. H. (1997). Directed evolution of enzyme catalyts. *Trends Biotechnol.* **15**, 523–530.
32. Kurtzman, A. L., Govindarajan, S., Vahle, K., Jones, J. T., Heinrichs, V. & Patten, P. A. (2001). Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Curr. Opin. Biotechnol.* **12**, 361–370.
33. Petrounia, I. P. & Arnold, F. H. (2000). Designed evolution of enzymatic properties. *Curr. Opin. Biotechnol.* **11**, 325–330.
34. Reetz, M. T. & Jaeger, K. E. (2000). Enantioselective enzymes for organic synthesis created by directed evolution. *Chemistry*, **6**, 407–412.
35. Pluckthun, A., Schaffitzel, C., Hanes, J. & Jermutus, L. (2001). *In vitro* selection and evolution of proteins. *Advan. Protein Chem.* **55**, 317–366.
36. Hopfner, K. P., Kopetzki, E., Kresse, G. B., Bode, W., Huber, R. & Engh, R. A. (1998). New enzyme lineages by subdomain shuffling. *Proc. Natl Acad. Sci. USA*, **95**, 9813–9818.
37. Ostermeier, M., Nixon, A. E., Shim, J. H. & Benkovic, S. J. (1999). Combinatorial protein engineering by incremental truncation. *Proc. Natl Acad. Sci. USA*, **96**, 3562–3567.
38. Soumillion, P., Jespers, L., Bouchet, M., Marchand-Brynaert, J. & Winter, G. (1994). Selection of B-lactamase on filamentous bacteriophage by catalytic activity. *J. Mol. Biol.* **237**, 415–422.
39. Forrer, P., Jung, S. & Pluckthun, A. (1999). Beyond binding: using phage display to select for structure, folding, and enzymatic activity in proteins. *Curr. Opin. Struct. Biol.* **9**, 514–520.
40. Amstutz, P., Forrer, P., Zahnd, C. & Pluckthun, A. (2001). *In vitro* display technologies: novel developments and applications. *Curr. Opin. Biotechnol.* **12**, 400–405.
41. Atwell, S. & Wells, J. A. (1999). Selection for improved subtiligases by phage display. *Proc. Natl Acad. Sci. USA*, **96**, 9497–9502.
42. Hansson, L. O., Widersten, M. & Mannervik, B. (1997). Mechanism-based phage display selection of active-site mutants of human glutathione transferase A1-1 catalyzing SNAr reactions. *Biochemistry*, **36**, 11252–11260.
43. Jestin, J. L., Kristensen, P. & Winter, G. (1999). A method for the selection of catalytic activity using phage display and proximity coupling. *Angew. Chem.* **38**, 1124–1127.
44. Tanaka, F. & Barbas, C. F. (2001). Phage display selection of peptides possessing aldolase activity. *Chem. Commun.* **8**, 769–770.
45. Viti, F., Nilsson, F., Demartis, S., Huber, A. & Neri, D. (2000). Design and use of phage display libraries for the selection of antibodies and enzymes. *Methods Enzymol.* **326**, 480–505.
46. Widersten, M., Hansson, L. O., Trontad, L. O. & Mannervik, B. (2000). Use of phage display and transition-state analogs to select enzyme variants with altered catalytic properties: glutathione transferase as an example. *Methods Enzymol.* **328**, 389–404.
47. Pasternak, A., Kaplan, J., Lear, J. D. & DeGrado, W. F. (2001). Proton and metal ion-dependent assembly of a model diiron protein. *Protein Sci.* **10**, 958–969.
48. Benson, D. E., Wisz, M. S. & Hellinga, H. W. (2000). Rational design of nascent metalloenzymes. *Proc. Natl Acad. Sci.* **97**, 6292–6297.
49. Moffet, D. A., Certain, L. K., Smith, A. J., Kessel, A. J., Beckwith, K. A. & Hecht, M. H. (2000). Peroxidase activity in heme proteins from a designed combinatorial library. *J. Am. Chem. Soc.* **122**, 7612–7613.
50. Lipscomb, J. D. (1994). Biochemistry of soluble methane monooxygenase. *Annu. Rev. Microbiol.* **48**, 371–399.
51. Liu, K. & Lippard, S. J. (1995). Studies of the soluble methane monooxygenase protein system: structure, component interactions, and hydroxylation mechanism. *Advan. Inorg. Chem.* **42**, 263–289.
52. Que, L. & Dong, Y. (1996). Modeling the oxygen activation chemistry of methane monooxygenase and ribonucleotide reductase. *Accts Chem. Res.* **29**, 190–196.
53. Whittington, D. A., Valentine, A. M. & Lippard, S. J. (1998). Substrate binding and C–H bond activation in the soluble methane monooxygenase hydroxylase. *J. Biol. Inorg. Chem.* **3**, 307–313.
54. Feig, A. L. & Lippard, S. J. (1994). Reactions of non-heme iron(II) centers with diooxygen in biology and chemistry. *Chem. Rev.* **94**, 759–805.
55. Lipscomb, J. D. & Que, L., Jr (1998). MMO: P450 in wolf's clothing? *J. Biol. Inorg. Chem.* **3**, 331–336.
56. Que, L. (1997). Oxygen activation at non-heme diiron active sites in biology: lessons from model complexes. *J. Chem. Soc., Dalton Trans.* **21**, 3933–3940.
57. Wallar, B. J. & Lipscomb, J. D. (1996). Dioxygen activation by enzyme containing binuclear non-heme iron clusters. *Chem. Rev.* **96**, 2625–2657.
58. Stubbe, J. & van der Donk, W. A. (1996). Protein radicals in enzyme catalysis. *Chem. Rev.* **98**, 705–762.
59. Stubbe, J. & Riggs-Gelasco, P. (1998). Harnessing free radicals: formation and function of the tyrosyl radical in ribonucleotide reductase. *Trends Biochem. Sci.* **23**, 438–443.
60. Lindqvist, Y., Huang, W., Schneider, G. & Shanklin, J. (1996). Crystal structure of Δ^9-stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins. *Eur. Mol. Biol. J.* **15**, 4081–4092.
61. Whittle, E. & Shanklin, J. (2001). Engineering Δ-9-16:0-acyl carrier protein (ACP) desaturase specificity based on combinatorial saturation mutagenesis and logical redesign of the castor Δ-9-18:0—ACP desaturase. *J. Biol. Chem.* **276**, 21500–21505.
62. Brunold, T. C. & Solomon, E. I. (1999). Reversible dioxygen binding to hemerythrin. 1. Electronic structures of deoxy- and oxyhemerythrin. *J. Am. Chem. Soc.* **121**, 8277–8287.
63. Brunold, T. C. & Solomon, E. I. (1999). Reversible dioxygen binding to hemerythrin. 2. Mechanism of the proton-coupled two electron transfer to O_2 at a single iron center. *J. Am. Chem. Soc.* **121**, 8288–8295.
64. Stenkamp, R. E. (1994). Dioxygen and hemerythrin. *Chem. Rev.* **94**, 715–726.
65. DeMare, F., Kurtz, D. M., Jr & Nordlund, P. (1996). The structure of desulfovibrio vulgaris rubrerythrin reveals a unique combination of rubredoxin-like FeS4 and ferritin-like diiron domains. *Nature Struct. Biol.* **3**, 539–546.
66. Theil, E. C. (1995). Ferritin and iron biomineralization. In *Comprehensive Supramolecular Chemisty*

(Suslick, K. S., ed.), pp. 65–89, Elsevier Science, New York.
67. Weber, P. C. & Salemme, F. R. (1980). Structural and functional diversity in 4-α-helical proteins. *Nature*, **287**, 82–84.
68. Summa, C. M., Lombardi, A., Lewis, M. & DeGrado, W. F. (1999). Tertiary templates for the design of diiron proteins. *Curr. Opin. Struct. Biol.* **9**, 500–508.
69. Baldwin, J., Voegtli, W. C., Khidekel, N., Moenne-Loccoz, P., Krebs, C., Pereira, A. S. *et al.* (2001). Rational reprogramming of the R2 subunit of *Escherichia coli* ribonucleotide reductase into a self-hydroxylating monooxygenase. *J. Am. Chem. Soc.* **123**, 7017–7030.
70. Voegtli, W. C., Khidekel, N., Baldwin, J., Ley, B. A., Bollinger, J. M. & Rosenzweig, A. C. (2000). Crystal structure of the ribonucleotide reductase R2 mutant that accumulates a μ-1,2-peroxodiiron(III) intermediate during oxygen activation. *J. Am. Chem. Soc.* **122**, 3255–3261.
71. Moeenne-Loccoz, P., Baldwin, J., Ley, B., Loehr, T. M. & Bollinger, J. M. (1998). Oxygen activation by non-heme diiron proteins: identification of a symmetric μ-1,2 peroxide in a mutant of ribonucleotide reductase. *Biochemistry*, **37**, 14659–14663.
72. Lombardi, A., Summa, C. M., Geremia, S., Randaccio, L., Pavone, V. & DeGrado, W. F. (2000). Retrostructural analysis of metalloproteins: application to the design of a minimal model for diiron proteins. *Proc. Natl Acad. Sci. USA*, **97**, 6298–6305.
73. Fairman, R., Ghao, H. G., Mueller, L., Lavoie, T. B., Shen, L., Novotny, L. & Matsueda, G. R. (1995). Characterization of a new four-chain coiled-coil: Influence of chain length on stability. *Protein Sci.* **4**, 1457–1469.
74. Fairman, R., Ghao, H. G., Lavoie, T. B., Villafraca, J. J., Matsueda, G. R. & Novotny, J. (1996). Design of heterotetrameric coiled coils: evidence for increased stabilization by Glu–Lys^+ ion pair interactions. *Biochemistry*, **35**, 2824–2829.
75. Richardson, J. S. & Richardson, D. C. (1989). The *de novo* design of protein structures. *Trends Biochem. Sci.* **14**, 304–309.
76. Hellinga, H. W. & Richards, F. M. (1994). Optimal sequence selection in proteins of known structure by simulated evolution. *Proc. Natl Acad. Sci. USA*, **91**, 5803–5807.
77. Betz, S. F. & DeGrado, W. F. (1996). Controlling topology and native-like behavior of *de novo*-designed peptides: design and characterization of antiparallel four-stranded coiled coils. *Biochemistry*, **35**, 6955–6962.
78. Hill, R. B., Raleigh, D. P., Lombardi, A. & DeGrado, W. F. (2000). *De novo* design of helical bundles as models for understanding protein folding and function. *Accts Chem. Res.* **33**, 745–754.
79. Struthers, M., Ottesen, J. J. & Imperiali, B. (1998). Design and NMR analyses of compact, independently folded BBA motifs. *Fold. Des.* **3**, 95–103.
80. DeGrado, W. F., Wasserman, Z. R. & Lear, J. D. (1989). Protein design, a minimalist approach. *Science*, **243**, 622–628.
81. North, B., Summa, C. M., Ghirlanda, G. & DeGrado, W. F. (2001). D_n symmetrical tertiary templates for the design of tubular proteins. *J. Mol. Biol.* **311**, 1081–1090.
82. Rosenzweig, A. C., Brandstetter, H., Whittington, D. A., Nordlund, P., Lippard, S. J. & Frederick, C. A. (1997). Crystal structures of the methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath): implications for substrate gating and component interactions. *Proteins: Struct. Funct. Genet.*, **29**, 141–152.
83. Christianson, D. W. & Fierke, C. A. (1996). Carbonic anhydrase: evolution of a zinc binding site by nature and by design. *Accts Chem. Res.* **29**, 331–339.
84. Kuo, J. M., Chae, M. Y. & Raushel, F. M. (1997). Perturbation to the active site of phosphotriesterase. *Biochemistry*, **36**, 1982–1988.
85. Marino, S. F. & Regan, L. (1999). Secondary ligands enhance affinity at a designed metal-binding site. *Chem. Biol.* **6**, 649–655.
86. Goodin, D. B. & McRee, D. E. (1993). The Asp-His-Fe triad of cytochrome *c* peroxidase controls the reduction potential, electronic structure, and coupling of the tryptophan free radical to the heme. *Biochemistry*, **32**, 3313–3324.
87. Vance, C. K. & Miller, A. F. (1998). Simple proposal that can explain the inactivity of metal-substituted superoxide dismutases. *J. Am. Chem. Soc.* **120**, 461–467.
88. O'Neil, K. T. & DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science*, **250**, 646–651.
89. Banner, D. W., Kokkinidis, M. & Tsernoglau, D. (1987). Structure of the ColE1 rop protein at 1.7 Å resolution. *J. Mol. Biol.* **201**, 601–619.
90. Krylov, D., Mikhailenko, I. & Vinson, C. (1994). A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *EMBO J.* **13**, 2849–2861.
91. Schneider, J. P., Lear, J. D. & DeGrado, W. F. (1997). A designed buried salt bridge in a heterodimeric coiled coil. *J. Am. Chem. Soc.* **119**, 5742–5743.
92. Oakley, M. G. & Kim, P. S. (1998). A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry*, **37**, 12603–12610.
93. Davies, D. R. & Rich, A. (1958). The formation of a helical complex between polyisnosinic acid polycytidylic acid. *J. Am. Chem. Soc. USA*, **80**, 1003–1004.
94. Le Brun, N. E., Keech, A. M., Mauk, M. R., Mauk, A. G., Andrews, S. C. & Thomson, A. J. (1996). Charge compensated binding of divalent metals to bacterioferritin: H^+ release associated with cobalt(II) and zinc(II) binding at dinuclear metal sites. *FEBS Letters*, **397**, 159–163.
95. Lipscomb, W. N. & Strater, N. (1996). Recent advances in zinc enzymology. *Chem. Rev.* **96**, 2375–2433.
96. Cox, E. H. & McLendon, G. L. (2000). Zinc-dependent protein folding. *Curr. Opin. Chem. Biol.* **4**, 162–165.
97. Schneider, J. P., Lombardi, A. & DeGrado, W. F. (1998). Analysis and design of 3-stranded coiled coils and 3-helix bundles. *Fold. Des.* **3**, R29–R40.
98. Bertini, I. & Luchinat, C. (1984). High spin cobalt(II) as a probe for the investigation of metalloproteins. *Advan. Inorg. Biochem.* **6**, 71–111.
99. Keech, A. M., Le Brun, N. E., Wilson, M. T., Andrews, S. C., Moore, G. R. & Thomson, A. J. (1997). Spectroscopic studies of cobalt(II) binding to *Escherichia coli* bacterioferritin. *J. Biol. Chem.* **272**, 422–429.
100. Frolow, F., Kalb, A. J. & Yariv, J. (1994). Structure of

a unique twofold symmetric haem-binding site. *Nature Struct. Biol.* **1**, 453–460.
101. Lindsay, S., Brosnahan, D. & Watt, G. D. (2001). Hydrogen peroxide formation during iron deposition in horse spleen ferritin using oxygen as an oxidant. *Biochemistry*, **40**, 3340–3347.
102. Peceira, A. S., Small, W., Krebs, C., Tavares, P., Edmondson, D. E., Theil, E. C. & Huynh, B. H. (1998). Direct spectroscopic and kinetic evidence for the involvement of a peroxodiferric intermediate during the ferroxidase reaction in fast ferritin mineralization. *Biochemistry*, **37**, 9871–9876.
103. Treffry, A., Zhao, Z., Quail, M. A., Guest, J. R. & Harrison, P. M. (1997). Dinuclear center of ferritin: studies of iron binding and oxidation show differences in the two iron sites. *Biochemistry*, **36**, 432–441.
104. Yang, X., Chen-Barrett, Y., Arosio, P. & Chasteen, N. D. (1998). Reaction paths of iron oxidation and hydrolysis in horse spleen and recombinant human ferritins. *Biochemistry*, **37**, 9743–9750.
105. Tolman, W. B., Liu, S., Bentsen, J. G. & Lippard, S. J. (1991). Models of the reduced forms of polyiron-oxo proteins: an asymmetric, triply carboxylate bridged diiron(II) complex and its reaction with dioxygen. *J. Am. Chem. Soc.* **113**, 152–164.
106. Yang, X., Le Brun, N. E., Thomson, A. J., Moore, G. R. & Chasteen, N. D. (2000). The iron oxidation and hydrolysis chemistry of *Escherichia coli* bacterioferritin. *Biochemistry*, **39**, 4915–4923.
107. Bollinger, M. J., Edmondson, D. E., Huynh, B. H., Filley, H., Norton, J. R. & Stubbe, J. (1991). Mechanism of assembly of the tyrosyl radical-dinuclear iron cluster cofactor or ribonucleotide reductase. *Science*, **253**, 292–298.
108. Fox, B. G., Shanklin, J., Ai, J., Loehr, T. M. & Sanders-Loehr, J. (1994). Resonance Raman evidence for an Fe–O–Fe center in the stearoyl-ACP desaturase. Primary sequence identity with other diiron-oxo proteins. *Biochemistry*, **33**, 12776–12786.
109. Desjarlais, J. R. & Handel, T. M. (1995). *De novo* design of the hydrophobic cores of proteins. *Protein Sci.* **4**, 2006–2018.
110. Dahiyat, B. I. & Mayo, S. L. (1997). *De novo* protein design: fully automated sequence selection. *Science*, **278**, 82–87.
111. Gutin, A. M. & Shakhnovich, E. I. (1993). Ground state of random copolymers and the discrete random energy model. *J. Chem. Phys.* **98**, 8174–8177.
112. Shakhnovich, E. & Gutin, A. M. (1993). Engineering of stable and fast-folding sequences of model proteins. *Proc. Natl Acad. Sci. USA*, **90**, 7195–7199.
113. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
114. Merrit, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.

Edited by P. Wright

(Received 23 January 2002; received in revised form 4 June 2002; accepted 10 June 2002)