The de novo design of a rubredoxin-like Fe site

EDGARDO FARINAS AND LYNNE REGAN

Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, Connecticut 06520

(RECEIVED April 1, 1998; ACCEPTED June 11, 1998)

Abstract

A redox center similar to that of rubredoxin was designed into the 56 amino acid immunoglobulin binding B1 domain of *Streptococcal* protein G. The redox center in rubredoxin contains an iron ion tetrahedrally coordinated by four cysteine residues, $[\text{Fe}(\text{S-Cys})_4]^{-1,-2}$. The design criteria for the target site included taking backbone movements into account, tetrahedral metal-binding, and maintaining the structure and stability of the wild-type protein. The optical absorption spectrum of the Co(II) complex of the metal-binding variant is characteristic of tetrahedral chelation by four cysteine residues. Circular dichroism and nuclear magnetic resonance measurements reveal that the metal-free and Cd(II)-bound forms of the variant are folded correctly and are stable. The Fe(III) complex of the metal-binding mutant reproduces the optical and the electron paramagnetic resonance spectra of oxidized rubredoxin. This demonstrates that the engineered protein chelates Fe(III) in a tetrahedral array, and the resulting center is similar to that of oxidized rubredoxin.

Keywords: de novo design; metalloprotein; protein engineering; rubredoxin

Protein engineering and de novo design approaches have begun to afford a more detailed understanding of the balance of the forces that determine protein structure and function (Regan & DeGrado, 1988; Regan, 1993; Bryson et al., 1995; Regan, 1995; Dalal et al., 1997; Smith & Regan, 1997; Hellinga, 1998). To date, there are just a few examples of designed proteins that adopt a specified fold, and number which display novel activities is still smaller (Dill, 1990; Schafmeister et al., 1993; Jackson et al., 1994; Quinn et al., 1994; Yan & Erickson, 1994; Baltzer et al., 1996; Dahiyat & Mayo, 1997; Hill & DeGrado, 1998; Quéneéneur et al., 1998). A particular focus of functional design efforts has been the engineering of novel metal-binding sites. Metal-binding sites represent an attractive target because they play a variety of functions from stabilization of protein structure to catalytic roles such as the activation of water for nucleophilic attack, electron transfer, and redox activity. More recently, several groups have focused strongly on the design of metal-binding sites with particular activities in mind. For example, two metal-binding histidine residues have been engineered into rat trypsin to “switch” the catalytic activity on or off by disrupting the catalytic triad. The protein can be turned off by Cu(II) chelation in the designed site and reactivated upon removal of Cu(II) with EDTA (Halfon & Craik, 1996). In another design, a metal-binding site, based on the metal center of carbonic

anhydrase, was built into a single-chain antibody with a goal to enhance catalytic activity (Crowder et al., 1995). A variety of metal sites have been designed into thioredoxin. For example, one mimics the active site of iron-superoxide dismutase (Pinto et al., 1997), while another introduces a $[\text{Fe}_4\text{S}_4]$ cuboidal cluster into the hydrophobic core (Coldren et al., 1997). Several other designs have incorporated $[\text{Fe}_4\text{S}_4]$ cuboidal clusters into the loops of four-helix bundle proteins (Gibney et al., 1996; Scott & Biggins, 1997). Here, we describe a novel design approach that we have used to successfully incorporate a tetrahedrally coordinated rubredoxin-like iron-binding site into a simple protein framework.

Previous work in this laboratory used computer-assisted searches to design tetrahedrally coordinated Zn(II)-binding sites that mimic structural and catalytic Zn(II) sites found in natural metalloproteins (Klemba & Regan, 1995; Klemba et al., 1995). The first step in identifying potential tetrahedrally coordinated metal-binding sites in proteins of known structure involves a computer search. The program Metal Search (Clarke & Yuan, 1995) uses backbone coordinates of the crystal or average NMR structure as input, and it replaces the natural amino acids with His and Cys residues. The program then searches all possible side-chain rotamers and stores them in an array. Finally, the program outputs sets of residues that potentially could form tetrahedrally coordinated metal-binding sites. Computer models are then built and assessed using a variety of criteria. The ultimate test is to make the protein and characterize its metal-chelating properties. An attractive scaffold for these designs is a variant of the B1 domain of IgG-binding protein G. The structure of this stable 56 residue domain is known at high resolution by both X-ray crystallography and NMR. Conveniently, the B1 domain contains neither Cys nor His residues, which could complicate metal chelation to the designed sites resulting from

Reprint requests to: Lynne Regan, Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, Connecticut 06520; e-mail: Lynne@nero.csb.yale.edu.

Abbreviations: CD, circular dichroism; Cd, cadmium; Co, cobalt; Cu, copper; EPR, electron paramagnetic resonance; Fe, iron; HEPES, 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; TCEP, tricarboxylethylphosphine.

their potential interaction with the added metal. In earlier studies, we described the successful introduction of a number of tetrahedrally coordinated metal binding sites, which utilized different combinations of His and Cys residues (Klemba et al., 1995; Klemba, 1996) into this framework. Here we describe a modification of the original design strategy that enables us to identify additional sites that were missed in the initial search. The new sites are less disruptive to protein structure than the previous ones, and move from designs for structural Zn-binding sites to a Fe-binding motif with the potential for redox and electron transfer activity.

In previous designs we, and other groups, have typically used the backbone coordinates (derived from an X-ray or average NMR structure) of the target framework to search for compatibility of substituted side chains with the desired structure or function (Heltinga & Richards, 1991; Desjarlais & Handel, 1995; Levitt et al., 1995). The limitation of using the static backbone coordinates is somewhat unrealistic and restrictive. In several examples where the hydrophobic core of a protein has been repacked, it has been shown that small backbone movements provide slight changes necessary to accommodate the newly introduced side chains in favorable conformations (Baldwin et al., 1993; Lim et al., 1994). In the study reported here, we have focused on the design of a rubredoxin-like site. The metal center in rubredoxin involves a tetrahedrally coordinated iron site ligated by the cysteinato sulfur residues, $[\text{Fe}(\text{S-Cys})_4]^{-1,-2}$. A particular feature of the design described here is that we attempted to include a more realistic description of backbone flexibility in our design strategy. Rather than using the average NMR structure as the search target, we selected at random 6 of the 60 calculated NMR structures of the B1 domain for the search. The aim was to identify positions not found in the average structure in order to provide a more comprehensive list of tetrahedral-metal-binding sites. The results were interesting, and they underscore the strength of this approach. We found that certain sites, particularly those in the well-defined hydrophobic core of the protein, are identified in both the average and the individually calculated structures. In contrast, there are also a large number of sites that are found in some, but not all, of the individual structures, and not in the average structure. These are sites that

would have been missed if only the average coordinates had been used. Several of these involve residues in the less well-defined (perhaps intrinsically more mobile) regions of the protein. For example, ^{15}N relaxation studies of the backbone dynamics of the B1 domain reveal that residues in the turns and loops display a higher degree of mobility relative to residues elsewhere in the protein, and that the inner two β -strands are the least mobile (Barich et al., 1994). Figure 1 displays the backbone trace of the individual NMR structure (green) used in our design overlaid with the average NMR structure (red). As expected, the backbones match well, especially in the β -sheet region with the most variability in the loop region connecting the helix to the beta strand 3. We chose to construct a metal-binding site in a loop region that was identified in one of the individual structures, but not in the average structure, and see if it truly provided a viable tetrahedral coordinating site. The site that is the subject of this report is an S-Cys₄ site that is intended to mimic the tetrahedrally coordinated Fe site of the electron transfer protein, rubredoxin. The site is formed by replacing residues Gly9, Gly38, Val39, and Glu56 with Cys. Residues 9, 38, and 39 are residues in the flexible loop region discussed above. E56 is the C-terminal residue of the protein. We named this variant [β 1-Cys4] (Fig. 2).

A computer model (MidasPlus) of this site was built using the individual NMR structure of the wild-type B1 domain and the side-chain orientations of the metal-binding residues obtained from the Metal Search output. Energy minimization calculations were performed using X-PLOR (Brünger, 1992). Finally, the computer program WHAT IF (Vriend, 1990) was used to identify any possible steric clashes that the metal-binding residues may encounter with the surrounding side chains and to assess the overall quality of the model; no unfavorable interactions were detected.

Results and discussion

We created the gene encoding the designed rubredoxin mimetic using a combination of PCR and cassette mutagenesis and purified the protein under reducing conditions, in a soluble form from *Escherichia coli*.

Fig. 1. Overlay of the backbone traces of the average NMR structure (red) and the individual NMR structure (green). The diagram was generated using MidasPlus.

Fig. 2. Model of the metal-binding site. The side chains of the metal-binding cysteine residues in yellow are shown together with the metal ion in green. The iron atom is not drawn to scale. The model was created using MidasPlus.

Characterization

The optical absorption spectra of Co(II)-substituted metalloproteins provide an excellent means by which to characterize the geometry and nature of the protein ligands in metal-binding site. The optical absorption spectrum of Co(II)-bound [β 1-Cys4] is shown in Figure 3. A key feature of the spectrum is the envelope of d-d transitions at 694 nm with $\epsilon = 380 \text{ M}^{-1} \text{ cm}^{-1}$. The intensity of this peak is characteristic of Co(II) coordination in a tetrahedral environment. Trigonal bipyramidal or octahedral coordination would give peaks at least 100-fold less intense and at shorter wavelengths. The λ_{max} of the d-d transitions is sensitive to the identity of the ligands in the first coordination sphere: λ_{max} moves to lower energy (i.e., longer wavelength) as the number of thiolate ligands increases (Moura et al., 1991; Krizek et al., 1993). This is consistent with the relative positions of histidine and cysteine in the spectrochemical series. We have

Fig. 3. Optical absorption spectrum of [β 1-Cys4] in the presence of $0.4 \mu\text{M}$ Co(II). Protein concentration, 0.36 mM ; buffer, 20 mM HEPES pH = 7.55, $T = 25^\circ\text{C}$.

seen this property with previously studied designed metal-binding sites in our laboratory (Regan & Clarke, 1990; Klemba et al., 1995; Klemba & Regan, 1995) with λ_{max} varying with the nature of the first shell of coordinating ligands as follows: His₃Cys ($\lambda_{\text{max}} = 592 \text{ nm}$) and His₂Cys₂ ($\lambda_{\text{max}} = 617 \text{ nm}$). The value of 694 nm we observe for the Co(II)-bound [β 1-Cys4] is consistent with an all Cys coordination environment. A further feature of the Co(II) spectrum to note is the intense charge transfer at 320 nm which also indicates thiolate ligation.

Effect of metal-binding on protein structure and stability

Having established that the designed site coordinates metal ions with tetrahedral geometry, it was important to determine the effect of metal-binding on protein structure and stability. Our expectation was that the introduction of the site into the more flexible loops rather than the tightly packed areas of secondary structure would result in less disruption of protein stability than we had observed in some of our earlier designs (Klemba et al., 1995). CD and NMR provide sensitive spectroscopic methods by which to determine the effect of introducing the coordinating residues and the consequence of metal binding on protein structure and stability. We chose to use Cd(II) in the CD and NMR investigations for several reasons. Cd(II) binds tightly to cysteine-rich systems due to the principles of hard and soft acids and bases (Pearson, 1997). Cd(II) is also known to retain tetrahedral geometry when substituted into metalloproteins that have tetrahedral metal chelating sites, for example, the S-Cys₄ site in rubredoxin (Henehan et al., 1993). Furthermore, cadmium exists almost exclusively in the Cd(II) oxidation state and, therefore, does not have the potential for complications associated with varying oxidation states as may be encountered with Fe. Before making the CD and NMR measurements, we characterized the binding properties of Cd(II) to [β 1-Cys4]. The absorption spectrum of the Cd(II) complex displays a characteristic charge transfer band at $\sim 250 \text{ nm}$ ($\epsilon = 23,000 \text{ M}^{-1} \text{ cm}^{-1}$). In a titration experiment monitoring the formation of this band, saturation occurs at 1 equivalent of Cd(II) to [β 1-Cys4]. The extinction coefficient observed is comparable to reported Cd(II) complexes

of S-Cys₄ containing peptides (Krizek et al., 1993) and Cd(II)-substituted rubredoxin (Henehan et al., 1993).

Figure 4 shows a CD wavelength scan comparing the structure of the apo and Cd(II)-bound forms of the [β 1-Cys4]. Both spectra are similar to that of the parent B1 domain, indicating that neither the introduction of the metal-binding residues, nor the chelation of metal by the protein causes a significant disruption of protein secondary structure.

To further verify the folded state of the proteins, we obtained 1D-NMR spectra of the Cd(II)-bound and metal-free forms of the protein. Figure 5A shows the spectrum of the Cd(II)-bound [β 1-Cys4] and Figure 5B that of the metal-free protein. Both spectra show the dispersion characteristic of the folded parent protein. Metal addition results in broadening of the resonances and it is most likely due to conformational exchange in the intermediate regime between the Cd(II)-bound and the apo protein forms. The kinetics of metal-binding will be a topic for future studies.

In addition to comparing the CD wavelength scans and 1D-NMR of the apo and Cd(II)-bound [β 1-Cys4], we also compared their thermal stabilities. Figure 6 shows thermal denaturation curves, monitored by CD, of the apo and Cd(II)-bound protein. The T_m of the apo protein is $77 \pm 1^\circ\text{C}$ and that of the Cd(II)-bound form is $82 \pm 1^\circ\text{C}$. Both display reversible, cooperative denaturation transitions. These stabilities compare favorably with that of the parent variant of the B1 domain which under similar conditions has a T_m of 82°C (Smith et al., 1996). Clearly, protein structure and stability have not been disrupted in either the metal-free or Cd(II)-bound form of [β 1-Cys4].

Fe(III)-bound [β 1-Cys4]

Rubredoxins are proteins of ~ 6 kD molecular weight. They are found primarily in anaerobic bacteria, and are believed to be involved in electron transfer reactions. The redox potential for various rubredoxins range from -57 to 6 mV (Lovenberg & Sobel, 1965; Moura et al., 1979; LeGall et al., 1988; Adams, 1992). The variations observed in the redox potential with proteins such as the

Fig. 5. One-dimensional ^1H NMR spectra between 6–10 ppm of (A) Cd(II)-bound [β 1-Cys4] and (B) metal-free [β 1-Cys4] in 90% H_2O –10% D_2O at 40 and 25°C , respectively.

rubredoxins is not well understood, and it has been suggested that tuning of the redox potential is modulated by the protein environment. The protein matrix can alter the redox potential through the ligand field produced by the chelating residues and the electrostatic potential surrounding the metal center. Such structural considerations include protein hydrogen bonding to the iron center, solvent accessibility to the site, and the number of charged residues (Stephens et al., 1996). Several high resolution crystal structures of

Fig. 4. Structural changes of metal-free [β 1-Cys4] and Cd(II)-bound [β 1-Cys4]. CD spectra of metal-free [β 1-Cys4] (open circles) and [β 1-Cys4] in the presence of 1 equivalent of Cd(II) (closed circles). The spectra were recorded at 25°C in a 2 mm pathlength cell with a protein concentration of $10 \mu\text{M}$ in 1 mM MOPS containing $100 \mu\text{M}$ tricarboxylethylphosphine at pH 8.

Fig. 6. The effect of metal binding on the stability of [β 1-Cys4]. Thermal denaturation of metal-free [β 1-Cys4] (open circles) and [β 1-Cys4] in the presence of 1 equivalent Cd(II) (closed circles). The spectra were recorded with a protein concentration of $50 \mu\text{M}$ in 20 mM MOPS containing 1 mM tricarboxylethylphosphine at pH 8.0 in a 2 mm pathlength cell at 220 nm.

rubredoxins have been solved (Watenpaugh et al., 1979; Day et al., 1992), and the iron ion is coordinated by four cysteinato sulfur atoms in a tetrahedral array $[\text{Fe}(\text{S-Cys})_4]^{-1,-2}$ with an average Fe-S bond distance of 2.3 Å. The $[\text{Fe}(\text{S-Cys})_4]^{-1,-2}$ site is located near the surface of the protein and is flanked by two hydrophobic arms. The structure of rubredoxin reveals several hydrogen bonds from the protein main chain to the ligands directly ligated to Fe (i.e., N-H...S-Fe). The protein milieu about the metal center typically consists of 52–54 amino acid residues folded into three-stranded antiparallel β -sheets and a number of connecting loops. Finally, the spectroscopic properties of rubredoxin display distinct characteristics associated with the tetrahedral $[\text{Fe}(\text{S-Cys})_4]^{-1}$ center.

The visible spectra of oxidized rubredoxin from *Clostridium pasteurianum* and other reported rubredoxins (Moura et al., 1978; Papavassiliou & Hatchikian, 1985; Blake et al., 1991) have absorption maxima at 380 nm and 490 nm with a less intense shoulder at ~565 nm. The tetrahedral Fe(III) center has an $S = 5/2$ spin state, and no allowed ligand field transitions are expected. Therefore, the observed red color is due to the ligand-to-metal charge transfer transitions. The Fe(III) complex of $[\beta 1\text{-Cys4}]$ also displays a red color, and the optical spectrum is dominated by two bands at 350 and 495 nm with a shoulder at ~575 nm (Fig. 7). The optical spectrum supports our model for $[\beta 1\text{-Cys4}]$ with the Fe(III) center that is tetrahedrally coordinated with an $S = 5/2$ spin state, and the red color here is also due to the ligand-to-metal charge-transfer transition (Eaton & Lovenberg, 1973; Bennett & Johnson, 1987; Gebhard et al., 1990).

Additional verification of the coordination geometry of the Fe(III) complex of $[\beta 1\text{-Cys4}]$ is seen in the EPR spectrum. EPR spectra of various oxidized rubredoxins display g values of 4.3 and 9.5, which arise from the 3/2 and 1/2 Kramers' doublets, respectively, of an $S = 5/2$ multiplet (Peisach et al., 1971). EPR spectra of Fe(III) complex of $[\beta 1\text{-Cys4}]$ at 4 K similarly exhibits a sharp g

value at 4.4 and 9.6 (Fig. 8). Again this provides support for $[\beta 1\text{-Cys4}]$ functioning as a mimetic of the rubredoxin Fe site.

The Fe-binding protein we have reported incorporates minimal design features: tetrahedral geometry at the metal site, absence of unfavorable interactions between the metal-binding residues and the rest of the protein, and retention of secondary structure and stability. The $[\beta 1\text{-Cys4}]$ variant was designed based on these criteria, and the Fe(III) complex of $[\beta 1\text{-Cys4}]$ reproduces optical and EPR spectra similar to that of rubredoxin. This demonstrates that the designed protein chelates Fe(III) in a tetrahedral array, and the resulting center is similar to that of rubredoxin.

The Fe(III) complex of $[\beta 1\text{-Cys4}]$ is able to reproduce some of the spectroscopic properties of rubredoxin based on our minimal design constraints. However, our designs did not incorporate several of the features found in rubredoxin. For example, the $[\text{Fe}(\text{S-Cys})_4]^{-1,-2}$ site of rubredoxin is buried between two hydrophobic arms, and the metal center is almost completely shielded from solvent. A noteworthy characteristic of rubredoxin is that it is isolated as a stable Fe(III)-tetrathiolate. Indeed, one of the factors that may attribute to the stability of the Fe(III) center is that the protein matrix shields the center from solvent. In the crystal structure of the oxidized form of rubredoxin from *C. pasteurianum* (Watenpaugh et al., 1979, 1980; Yelle et al., 1995), the nearest water molecule is 7.1 Å from the metal center, and only two waters molecules are within 8 Å. Furthermore, the crystal structure of rubredoxin from *Pyrococcus furiosus* also reveals that the Fe(III) center is inaccessible to solvent.

It should be noted that the absorption bands of the Fe(III) complex of $[\beta 1\text{-Cys4}]$ decrease with time with the concomitant formation of a rust-colored precipitate and attempts to measure the redox potential thus far have been unsuccessful. Including solvent accessibility into the design constraints may be necessary to stabilize the metal center and to reproduce all the features of the

Fig. 7. Optical absorption spectrum of $[\beta 1\text{-Cys4}]$ in the presence of 1 equivalent of Fe(III). The spectrum was recorded at 20 °C with a protein concentration of 100 μM in 20 mM HEPES at pH 7.5. Inset: Optical absorption spectrum of oxidized rubredoxin. The spectrum of rubredoxin was recorded with a protein concentration of 50 μM in 20 mM HEPES at pH 7.5.

Fig. 8. Electron paramagnetic resonance spectrum of the Fe(III) complex of [β 1-Cys4] in 20 mM HEPES at pH 7.5. Spectrometer settings: microwave frequency, 9.24 GHz; microwave power, 2 mW; magnetic field modulation amplitude, 10 G; temperature, 10 K. Inset: EPR spectrum of oxidized rubredoxin.

rubredoxin site. In a previous report, peptides designed to mimic the Fe-binding site of *C. pasteurianum* rubredoxin (Z-Cys-X-X-Cys-OMe (X-X = Thr-Val, Ala-Ala, or Pro-Leu) bind Fe(III) but are unstable decomposing with half lives ranging from 7–25 min. A reversible redox couple is not observed for these complexes in aqueous solutions (Christou et al., 1977; Ueyama et al., 1981; Nakamura & Ueyama, 1989). It is known that Fe(III) notoriously reacts with simple thiolates via an auto redox reaction to afford Fe(II) and disulfide (Millar et al., 1996). This mechanism may also account for the instability of our designed protein. Furthermore, the metal-binding site was engineered in a loop region of the B1 domain, and in this position the protein matrix will provide minimal protection of the metal center from solvent. Therefore, the Fe(III) center would be vulnerable to water coordination, which may facilitate the formation of insoluble hydroxide species.

Conclusions

We have successfully engineered a protein with a designed tetrahedral site based on the active site of rubredoxin. Our design aimed to take backbone movements into account in order to find new metal-binding sites that are not found in the average structure of the protein. Using this approach, a more comprehensive list of

potential tetrahedral sites were identified. This is important because suitable sites that are not identified in the average NMR or crystal structure may be found in areas of the protein with different dielectric constants. These varied environments may prove essential to the stability of the metal center and the tuning of the redox potential. When comparing the average and the individual NMR structure, the greatest differences are located in the flexible loop region. Comparison of the crystal (Gallagher et al., 1994) and NMR structure of the wild-type B1 domain also reveal differences in the loop regions. This is also reflected in ^{15}N relaxation measurements. Furthermore, by designing the site in the loops and not disturbing the hydrophobic core of the protein, we were able to preserve the secondary structure and high thermal stability of the protein.

Taking into account minimal design constraints, we were able to reproduce several of the spectroscopic properties of the rubredoxin. However, including more design criteria should enhance the stability of the metal center. From these studies, it appears that the burial of the iron center within the hydrophobic core may prevent the auto redox reaction by shielding the Fe(III) center from solvent. It appears there is a subtle balance between retaining both protein stability and metal center stability when designing [$\text{Fe}(\text{S-Cys})_4$] $^{-1,-2}$ sites into proteins. Perhaps, the 56 amino acid scaffold provided by the B1 domain is too small to accommodate metal chelating residues in the hydrophobic core, and maintain the overall folded structure and stability of the wild-type protein. Our second generation of designed [$\text{Fe}(\text{S-Cys})_4$] $^{-1,-2}$ containing sites, which incorporate additional design constraints based on these observation, are currently underway.

Materials and methods

Design of metal site

The Metal Search program (Clarke & Juan, 1995) is the first step in our design of tetrahedral metal-binding sites. We randomly chose model number 56 (Gronenborn et al., 1991) of the individual NMR structures as input for the program Metal Search. We chose a site that was not present in the average NMR structure. The metal-binding site was located in the loop region of the B1 domain in order to keep the hydrophobic core intact. Therefore, the secondary structure and high thermal stability would be maintained. The site was modeled using MidasPlus, and energy minimization calculations were performed using X-PLOR (Brünger, 1992). The computer program What If (Vriend, 1990) was utilized to identify steric clashes between the side chains of the designed site with surrounding residues, and no steric clashes were identified in the model.

Cloning, expression, and purification

The gene encoding for [β 1-Cys4] was constructed from polymerase chain reaction derived cassettes into a synthetic gene encoding B1T2Q, a modified form of the wild-type B1 domain in which the mutation T2Q has been incorporated to eliminate incomplete N-terminal processing (Smith et al., 1996). The identity of the genes were confirmed by dideoxy sequencing. All molecular biology manipulations were performed using standard techniques (Maniatis et al., 1982). Protein expression and purification were performed by literature procedures (Klemba & Regan, 1995). The

extinction coefficient at 280 nm, based on the amino-acid analysis, was determined to be $11,900 \text{ M}^{-1} \text{ cm}^{-1}$. The entire protein sequence of [β 1-Cys4] is: QYKLILNCKTLKGETTTEAVDAATA EKVKQYANDNCCDGEWTWDDATKTFTVTC.

Reconstitution of mutants with transition metals

All manipulations were performed with thoroughly degassed buffers at 20 °C in a nitrogen-filled glove box. The metal-free protein was reduced by diluting a concentrated stock solution of protein in an equal volume of 1 M dithiothreitol and 6 M guanidine hydrochloride in 100 mM phosphate buffer pH 7.3, and the solution was allowed to incubate for 30 min. The solution was desalted over a Sephadex G-25 M column, which was equilibrated with 20 mM HEPES at pH 7.5. Final protein concentrations were determined by the reaction of free cysteine with 3,5'-dithiobis-(2-nitrobenzoic acid) (Riddles et al., 1983). A freshly prepared stock solution 5 mM CoCl_2 in 20 mM HEPES at pH 7.5 was added in excess to metal free protein to screen for tetrahedral metal binding.

The reconstitution of the iron complex was accomplished by adding an excess amount of a freshly prepared 5 mM solution of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ in 20 mM HEPES at pH 7.5 to the fully reduced protein. The resulting solution was desalted on a Sephadex G-25 M column to remove excess Fe(II). The Fe(III) complexes of the mutants were either generated by air or $\text{K}_3\text{Fe}(\text{CN})_6$ oxidation of the Fe(II) complexes.

Circular dichroism studies

All CD data were collected on an Aviv 62DS Circular Dichroism Spectropolarimeter. CD spectra for the wavelength scans were recorded on 10 μM protein in 1 mM MOPS containing 100 μM TCEP pH 8 at 25 °C in a 2 mm path length cell. Thermal denaturation data were taken on samples containing 50 μM protein in 20 mM MOPS containing 1 mM TCEP at pH 8.0. The thermal denaturation was measured at 220 nm over a range of 4–98 °C with a temperature step of 1 °C and a 1 min equilibration time and a 30 s signal averaging time.

Electron paramagnetic resonance spectroscopy

Low temperature electron paramagnetic resonance spectra were collected on a Varian E-line EPR spectrometer equipped with an Oxford Instruments ESR 900 liquid helium cryostat. The spectrometer conditions were microwave frequency, 9.24 GHz; microwave power, 2 mW; magnetic field modulation amplitude, 10 G; temperature, 10 K. Sample preparations of the Fe(III) complexes of the mutants were similar to the reconstitution procedures described.

Nuclear magnetic resonance spectroscopy

Spectra were collected on a Varian 500 MHz Unity instrument. Protein concentrations were $\sim 250 \mu\text{M}$ in 20 mM HEPES buffer at pH 7.5 with 10% D_2O . The one-dimensional spectra of metal-free and Cd(II)-bound proteins in 90% H_2O –10% D_2O comprised of 128 and 1,000 transients, respectively, with a spectra width of 7,000 Hz. Water suppression was achieved by presaturation of the solvent for 1 s during the recycle delay. The samples were prepared as described in the reconstitution procedure.

Other physical measurement

Absorption spectra were recorded with a Hewlett Packard HP 8453 UV-Visible spectrometer.

Note added in proof

Benson et al. (1998) recently described the introduction of a Cys4 Fe-binding site into thioredoxin.

Acknowledgments

We thank James Vergis for cloning the [β 1-Cys4] variant. We are indebted to Dr. Suganthi Balasubramanian for her expertise in obtaining the NMR data. We are extremely grateful to Ryan MacArthur for experimental assistance with the EPR spectrometer. We are also appreciative to Olaf Kievit for assistance with cyclic voltametry. We thank Pat Flemming for invaluable assistance with computer modeling. We thank Chris Coldren for stimulating discussions regarding his interest in the project. We are also appreciative of the Regan lab members for help during the project, and gratefully acknowledge the support from the Office of Naval Research grant N00014951-0913 and National Institutes of Health grant GM49146-04.

References

- Adams M. 1992. Novel iron-sulfur centers in metalloenzymes and redox proteins from extreme thermophilic bacteria. *Adv Inorg Chem* 38:341–396.
- Baldwin EP, Hajiseyedi O, Basse WA, Matthews BW. 1993. The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme. *Science* 262:1715–1718.
- Baltzer L, Lundh A-C, Broo K, Olofsson S, Ahlberg P. 1996. Polypeptides with supersecondary structures as templates in rational catalyst design. Catalysis of self functionalization by designed helix-loop-helix motifs. *J Chem Soc Perkin Trans 2*:1671–1676.
- Barich JJ, Brasberger B, Gronenborn AM, Clore GM. 1994. Investigations of the backbone dynamics of the IgG-binding domain of streptococcal protein G by heteronuclear two-dimensional ^1H - ^{15}N nuclear magnetic resonance spectroscopy. *Protein Sci* 3:15–21.
- Bennett DE, Johnson MK. 1987. The electronic and magnetic properties of rubredoxin: A low-temperature magnetic circular dichroism study. *Biochim Biophys Acta* 911:71–90.
- Benson et al. 1998. *Biochemistry* 37:7070–7076.
- Blake PR, Park J, Bryant FO, Aono S, Magnuson JK, Eccleston E, Howard JB, Summers MF, Adams MWW. 1991. Determinants of protein hyperthermostability: Purification and amino acid sequence of rubredoxin from the hyperthermophilic archaeobacterium *Pyrococcus furiosus* and secondary structure of the zinc adduct by NMR. *Biochemistry* 30:10885–10895.
- Brünger AT. 1992. *X-PLOR Version 3.1: A system for X-ray crystallography and NMR*. New Haven: Yale University Press.
- Bryson JW, Betz SF, Lu HS, Suich DJ, Zhou HX, O'Neil KT, DeGrado WF. 1995. Protein design: A hierarchic approach. *Science* 270:935–941.
- Christou G, Ridge B, Rydon HN. 1977. Direct formation of peptide analogues of rubredoxin and four-iron ferredoxins from their components. *JCS Chem Comm* (24):908–909.
- Clarke ND, Yuan S-M. 1995. Metal Search: A computer program that helps design tetrahedral metal-binding sites. *Proteins Struct Funct Genet* 23:256–263.
- Coldren CD, Hellinger HW, Caradonna JP. 1997. The rational design and construction of a cuboidal iron-sulfur protein. *Proc Natl Acad Sci* 94:6635–6640.
- Crowder MW, Stewart JD, Roberts VA, Bender CJ, Tevelakh E, Peisach J, Getzoff ED, Gaffney BJ, Benkovic SJ. 1995. Spectroscopic studies on the designed metal-binding sites of the 43C9 single chain antibody. *J Am Chem Soc* 117:5627–5634.
- Dahiyat BI, Mayo SL. 1997. De novo protein design: Fully automated sequence selection. *Science* 278:82–87.
- Dalal S, Balasubramanian S, Regan L. 1997. Transmuting α helices and β sheets. *Folding Design* 2:71–79.
- Day MW, Hsu BT, Joshua-Tor L, Park JB, Zhou ZH, Adams MWW, Rees DC. 1992. X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Sci* 1:1494–1507.
- Desjarlais JR, Handel TM. 1995. New strategies in protein design. *Curr Opin Biotechnol* 6:460–466.

- Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* 29:7133–7155.
- Eaton WA, Lovenberg W. 1973. The iron-sulfur complex in rubredoxin. In: Spiro TG, eds. *Iron-sulfur proteins*, Vol. II. New York: Academic Press. pp 131–161.
- Gallagher T, Alexander P, Bryan P, Gilliland GL. 1994. Two crystal structures of the B1 immunoglobulin-binding domain of *Streptococcal* protein G and comparison with NMR. *Biochemistry* 33:4721–4729.
- Gebhard MS, Deaton JC, Koch SA, Millar M, Solomon EI. 1990. Single-crystal spectral studies of $\text{Fe}(\text{SR})_4^-$ [$\text{R} = 2,3,5,6\text{-(Me)}_4\text{C}_6\text{H}$]; the electronic structure of the ferric tetrathiolate active site. *J Am Chem Soc* 112:2217–2231.
- Gibney BR, Mulholland SE, Rabanal F, Dutton PL. 1996. Ferredoxin and ferredoxin-heme maquettes. *Proc Natl Acad Sci USA* 93:15041–15046.
- Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. 1991. The immunoglobulin binding domain of *Streptococcal* protein G has a novel and highly stable polypeptide fold. *Science* 253:657–661.
- Halfon S, Craik CS. 1996. Regulation of proteolytic activity by engineered tridentate metal binding loops. *J Am Chem Soc* 118:1227–1228.
- Hellings HW. 1998. The construction of metal centers in proteins by rational design. *Folding Design* 3:R1–R8.
- Hellings HW, Richards FM. 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.
- Henehan CJ, Pountney DL, Zerbe O, Vasak M. 1993. Identification of cysteine ligands in metalloproteins using optical and NMR spectroscopy: Cadmium-substituted rubredoxin as a model $[\text{Cd}(\text{CysS})_4]^{2-}$ center. *Protein Sci* 2:1756–1764.
- Hill BR, DeGrado WF. 1998. Solution structure of a $\alpha 2\text{D}$, a native-like de novo designed protein. *J Am Chem Soc* 120:1138–1145.
- Jackson DT, Burnnier J, Quan C, Stanely M, Tom J, Wells JA. 1994. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* 266:243–247.
- Klemba M. 1996. Design and characterization of de novo zinc(II) sites in two proteins [Dissertation]. Yale University.
- Klemba M, Gardner KH, Marino S, Clarke ND, Regan L. 1995. Novel metal-binding proteins by design. *Nature Struct Biol* 2:368–373.
- Klemba M, Regan L. 1995. Characterization of metal binding by a designed protein: Single ligand substitutions at a tetrahedral Cys_2His_2 site. *Biochemistry* 34:10094–10100.
- Krizek BA, Merkle DL, Berg JM. 1993. Ligand variation and metal ion binding specificity in zinc finger peptides. *Inorg Chem* 32:937–940.
- LeGall J, Prickril I, Moura I, Xavier AV, Moura JJ, Huynh BH. 1988. Isolation and characterization of rubredoxin, a non-heme iron protein from *Desulfovibrio vulgaris* that contains rubredoxin centers and a hemerythrin-binuclear iron cluster. *Biochemistry* 27:1636–1642.
- Levitt M, Gerstein M, Huang E, Subbiah S, Tsai J. 1995. Protein folding: The endgame. *Annu Rev Biochem* 66:549–579.
- Lim WA, Hodel A, Sauer RT, Richards FM. 1994. The crystal structure of a mutant protein with altered but improved hydrophobic core packing. *Proc Natl Acad Sci USA* 91:423–427.
- Lovenberg W, Sobel B. 1965. Rubredoxin: A new electron transfer protein from *Clostridium pasteurianum*. *Proc Natl Acad Sci USA* 54:193–199.
- Maniatis T, Fritsch EF, Sambrook J. 1982. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Millar M, Lee JF, O'Sullivan T, Koch SA, Fikar R. 1996. Models for the iron-sulfur protein rubredoxin: The use of sterically hindered thiolate ligands to stabilize $[\text{Fe}(\text{SR})_4]^-$ complexes; some considerations of the structure of the $[\text{Fe}(\text{S-Cys})_4]$ centers in the oxidized rubredoxin. *Inorg Chim Acta* 243:333–343 (and references within).
- Moura I, Moura JIG, Santos MH, Xavier AV, LeGall J. 1979. Redox studies on rubredoxin from sulphate and sulfur reducing bacteria. *FEBS Lett* 107:419–421.
- Moura I, Teixeira M, LeGall J, Moura JIG. 1991. Spectroscopic studies of cobalt and nickel substituted rubredoxin and desulfuredoxin. *J Inorg Biochem* 44: 127–139.
- Moura I, Xavier AV, Cammack R, Bruschi M, LeGall J. 1978. A comparative spectroscopic study of two non-haem iron proteins lacking labile sulphide from *Desulphovibrio gigas*. *Biochim Biophys Acta* 533:156–162.
- Nakamura A, Ueyama N. 1989. Cysteine-containing oligopeptide model complexes of iron-sulfur proteins. In: Sykes AG, ed. *Advances in inorganic chemistry*, Vol. 33. New York: Academic Press, Inc. pp 39–67.
- Papavassiliou P, Hatchikian EC. 1985. Isolation and characterization of rubredoxin and a two-(4Fe-4S) ferredoxin from *Thermodesulfobacterium commune*. *Biochim Biophys Acta* 810:1–121.
- Pearson RG. 1997. *Chemical hardness*. Weinheim: Wiley-VCH.
- Peisach J, Blumberg WE, Lode ET, Coon MJ. 1971. An analysis of the electron paramagnetic resonance spectrum of *Pseudomonas oleovorans* rubredoxin. *J Biol Chem* 246:5877–5881.
- Pinto AL, Hellings HW, Caradonna JP. 1997. Construction of a catalytically active iron superoxide dismutase by rational protein design. *Proc Natl Acad Sci USA* 94:5562–5567.
- Quénéneur E, Moutiez M, Charbonnier J-B, Ménez A. 1998. Engineering cyclophilin into a proline-specific endopeptidase. *Nature* 1998:301–304.
- Quinn TP, Tweedy NB, Williams RW, Richardson JS, Richardson DC. 1994. Betadoublet: De novo design, synthesis, and characterization of a β -sandwich. *Proc Natl Acad Sci USA* 91:8747–8751.
- Regan L. 1993. The design of metal-binding sites in protein. *Ann Rev Biophys Biomol Struct* 22:257–281.
- Regan L. 1995. Protein design: Novel metal-binding sites. *TIBS* 20:280–285.
- Regan L, Clarke ND. 1990. The tetrahedral zinc(II)-binding site introduced into a designed protein. *Biochemistry* 29:10878–10883.
- Regan L, DeGrado WF. 1988. Characterization of a helical protein designed from first principles. *Science* 241:976–978.
- Riddles PW, Blakeley, RL, Zerner B. 1983. Reassessment of Ellman's reagent. *Methods Enzymol* 91:49–60.
- Schafmeister CE, Miercke LJW, Stroud RM. 1993. Structure at 2.5 Å of a designed peptide that maintains solubility of membrane. *Protein Sci* 262:734–738.
- Scott MP, Biggins J. 1997. Introduction of a $[\text{4Fe-4S}(\text{S-cys})_4]^{+1,+2}$ iron-sulfur center into a four- α helix protein using design parameters from the domain of the F_x cluster in the photosystem I reaction center. *Protein Sci* 6:340–346.
- Smith CK, Regan L. 1997. The construction and design of beta sheets. *Acc Chem Res* 30:153–161.
- Smith CK, Withka JM, Regan L. 1996. A thermodynamic scale for the β -sheet forming tendencies of the amino acids. *Biochemistry* 33:5510–5517.
- Stephens PJ, Jollie DR, Warshel A. 1996. Protein control of redox potentials of iron-sulfur proteins. *Chem Rev* 96:2491–2513.
- Ueyama N, Nakata M, Nakamura A. 1981. *Bull Chem Soc Jpn* 54:1727–1730.
- Vriend G. 1990. WHAT IF: A molecular modelling and drug design program. *J Mol Graph* 8:52–56.
- Watenpugh KD, Sieker LC, Jensen LH. 1979. The structure of rubredoxin at 1.2 Å resolution. *J Mol Biol* 131:509–522.
- Watenpugh KD, Sieker LC, Jensen LH. 1980. Crystallographic refinement of rubredoxin at 1.2 Å resolution. *J Mol Biol* 138:615–633.
- Yan Y, Erickson BW. 1994. Engineering of betabellin 14D: Disulfide-induced folding of a β -sheet protein. *Protein Sci* 3:1069–1073.
- Yelle RB, Park N-S, Ichiye T. 1995. Molecular dynamics simulations of rubredoxin from *Clostridium pasteurianum*: Changes in structure and electrostatic potential during redox reactions. *Proteins Struct Funct Genet* 22:154–167.