Structural determinants of protein stabilization by solutes

The importance of the hairpin loop in rubredoxins

Tiago M. Pais¹, Pedro Lamosa¹, Wagner dos Santos¹, Jean LeGall¹,², David L. Turner¹,³ and Helena Santos¹

¹ Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Portugal
² Department of Biochemistry, University of Georgia, Athens, GA, USA
³ Department of Chemistry, University of Southampton, UK

Keywords
compatible solutes; hairpin structure; NMR; rubredoxin; thermostability

Correspondence
H. Santos, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Apartado 127, 2780-156 Oeiras, Portugal
Fax: +351 21 4428766
Tel: +351 21 4469828
E-mail: santos@itqb.unl.pt

(Received 22 June 2004, revised 7 December 2004, accepted 17 December 2004)

In spite of the extensive accumulation of data on protein structure, the molecular determinants of protein thermal stability remain elusive. Also, the beneficial stabilizing effects exerted by various compatible solutes have been known for a long time, yet the mechanisms responsible for this stabilization are a matter of intense discussion [1–4]. One of the reasons for this apparent lack of success is that many different factors, both intrinsic and extrinsic, seem to contribute to the thermostability of any given protein [5]. Protein stability appears as the result of a delicate balance of stabilizing and destabilizing interactions, with the thermodynamic

Despite their high sequence homology, rubredoxins from Desulfovibrio gigas and D. desulfuricans are stabilized to very different extents by compatible solutes such as diglycerol phosphate, the major osmolyte in the hyperthermophilic archaeon Archaeoglobus fulgidus [Lamosa P, Burke A, Peist R, Huber R, Liu M Y, Silva G, Rodrigues-Pousada C, LeGall J, Maycock C and Santos H (2000) Appl Environ Microbiol 66, 1974–1979]. The principal structural difference between these two proteins is the absence of the hairpin loop in the rubredoxin from D. desulfuricans. Therefore, mutants of D. gigas rubredoxin bearing deletions in the loop region were constructed to investigate the importance of this structural feature on protein intrinsic stability, as well as on its capacity to undergo stabilization by compatible solutes. The three-dimensional structure of the mutant bearing the largest deletion, Δ17–29, was determined by ¹H-NMR, demonstrating that, despite the drastic deletion, the main structural features were preserved. The dependence of the NH chemical shifts on temperature and solute concentration (diglycerol phosphate or mannosylglycerate) provide evidence of subtle conformational changes induced by the solute. The kinetic stability (as assessed from the absorption decay at 494 nm) of six mutant rubredoxins was determined at 90 °C and the stabilizing effect exerted by both solutes was assessed. The extent of protection conferred by each solute was highly dependent on the specific mutant examined: while the half-life for iron release in the wild-type D. gigas rubredoxin increased threefold in the presence of 0.1 m diglycerol phosphate, mutant Δ23–29 was destabilized. This study provides evidence for solute-induced compaction of the protein structure and occurrence of weak, specific interactions with the protein surface. The relevance of these findings to our understanding of the molecular basis for protein stabilization is discussed.

Abbreviations
DGP, diglycerol phosphate; MG, mannosylglycerate; Rd, rubredoxin; RdDd, rubredoxin from Desulfovibrio desulfuricans; RdDg, rubredoxin from Desulfovibrio gigas.
stability of the native state emerging as a small difference of large numbers [6]. Similarly, the stabilizing effect conferred by compatible solutes will be the result of a plethora of direct and/or indirect, weak interactions between the solute (or the changes that the solute causes in the solvent properties) and the several chemical groups present on the protein surface, rendering the magnitude of this effect subtly dependent on the particular solute/protein pair examined and, therefore, extremely difficult to predict.

One of the strategies used to explore this maze of interactions and try to rationalize them is to investigate series of homologous proteins in order to unravel the structural determinants of protein stabilization by compatible solutes. In a previous study we compared the action of a compatible solute, diglycerol phosphate (DGP), on the stability of rubredoxins from three bacterial sources [7]. These small metalloproteins display a wide variation in thermal stability, despite having a considerable degree of sequence and structural similarity. Typically, rubredoxins are composed of about 52–54 residues and include a three-stranded β-sheet, a metal centre comprising one iron atom tetrahedrally coordinated by four cysteine sulfur atoms, and a small hydrophobic core, which is shielded from solvent access by a hairpin loop [8]. Despite the structural similarity between rubredoxins, the degree of stabilization conferred by DGP was diverse. Although having almost no effect on the thermal stability of the rubredoxin (Rd) from Desulfovibrio desulfuricans (RdDd), DGP was able to triple the half-life for thermal denaturation of the other two rubredoxins examined. RdDd is the least heat-stable of the several rubredoxins investigated, and is the only one not stabilized by DGP. Conversely, the Rd from D. gigas (RdDg) is the most stable and strongly stabilized by this solute. The main structural difference between RdDd and other rubredoxins is the lack of seven amino acids in the hairpin loop.

In order to investigate why this structural feature (the presence of the loop region) seemed to have such a profound effect on stability and stabilization of rubredoxins, we constructed a series of mutants of RdDg with different extents of deletion in the original hairpin loop. The determination of the NMR solution structure was deemed important, first, to ensure that the deletion had not substantially altered the protein structure (except in the loop region); and second, to provide the structural detail needed to elucidate the molecular basis of protein stabilization by solutes. Three point mutants were also studied to assess the importance of total surface charge or changes in the most exposed hydrophobic residue.

DGP and mannosylglycerate (MG), two negatively charged compatible solutes that we isolated from hyperthermophiles, were used in this study. The effect of these solutes on the thermal stability of six mutants was investigated. Moreover, as chemical shifts are good indicators of changes in protein structure or dynamics, the changes of the proton chemical shifts with temperature and solute concentration were analysed to extract information on protein/solute interactions.

**Results**

**Thermal stability of rubredoxins**

Mutant iron rubredoxins show the same characteristic bands of the UV–visible absorption spectrum as the native protein with maxima centred at 380, 494 and 570 nm. These bands are bleached, due to the disruption of the iron centre when the protein undergoes denaturation. Monitoring the loss of the metal centre through the decrease in absorbance at 494 nm provides an expeditious way to evaluate the kinetic stability of rubredoxins [9–11]. The half-life (t₁/₂) for iron release of the native and mutant rubredoxins was measured at 90 °C.

All rubredoxins examined exhibited mono-exponential behaviour in regard to the decay of absorbance at 494 nm (data not shown). Complete bleaching of spectral features at 380 and 494 nm occurred without formation of detectable precipitates, either from protein precipitation or insoluble ferric oxides. The spectral features did not recover on cooling, which indicates that protein denaturation under these conditions is an irreversible process, in agreement with previous studies regarding thermal denaturation of rubredoxins [7,9,11].

Recombinant RdDg presented a half-life for disruption of the iron centre (t₁/₂) of 96 min; all mutations resulted in a decrease of this parameter. The mutants bearing deletions in the loop region showed a dramatic decrease (between 69 and 89%) in their half-lives relative to the native form (Table 1; Fig. 1). Interestingly, mutant Δ23/29 had a half-life comparable with that of the RdDd, but the two other mutants lost iron at an even higher rate. The larger the deletion, the shorter the half-life became, with mutant Δ17/29 showing the lowest value for this parameter. In general, single mutations had a smaller effect on the rate of iron loss, except for V8N, which showed a rate comparable with that of mutant Δ23/29.

The effect induced by DGP in native RdDg was impressive with at least a threefold increase of the half-life [7]. However, the effect observed for the
mutant rubredoxins was lower. Mutants D2K, K17E and Δ17/29 showed a clear increase in the half-life for iron loss (between 52 and 94%), but a minor change was observed with mutants Δ17/26 and V8N (Table 1; Fig. 1). Most surprisingly, the half-life of mutant Δ23/29 was reduced in the presence of DGP. It is also interesting to note that, for the point mutants, the added stabilization follows the intrinsic stability, with the larger increases occurring in the proteins with higher intrinsic stability. This trend, however, was not observed in the case of loop deletions, where the most stable mutant (with respect to the iron loss), Δ23/29, was actually destabilized by addition of DGP. In contrast, the presence of MG caused a consistent retardation on the rates of all rubredoxins examined; in the case of RdDg the increment of half-life induced by MG was much lower than that of DGP, but MG stabilized the deletion-mutants to a much higher degree, including Δ23/29. Because K⁺ was the counter-ion for the negative charge of DGP and MG, the effect of KCl on the rate of iron release was also determined. We found that KCl had no significant effect on the half-life of the proteins examined (Table 1).

### Table 1. Effect of solute addition on the half-life values (min) for the thermal denaturation of native rubredoxins and mutants.

<table>
<thead>
<tr>
<th>Protein</th>
<th>No additions</th>
<th>Diglycerol phosphate 0.1 M</th>
<th>Mannnosylglycerate 0.2 M</th>
</tr>
</thead>
<tbody>
<tr>
<td>RdDg*</td>
<td>96.2 ± 9.4</td>
<td>295.0 ± 7.1</td>
<td>129.6 ± 5.2</td>
</tr>
<tr>
<td>Δ17</td>
<td>29b</td>
<td>10.5 ± 1.5</td>
<td>16.0 ± 5.0</td>
</tr>
<tr>
<td>Δ17</td>
<td>26</td>
<td>14.1 ± 1.4</td>
<td>15.9 ± 2.1</td>
</tr>
<tr>
<td>Δ23</td>
<td>29</td>
<td>29.7 ± 3.8</td>
<td>15.1 ± 2.5</td>
</tr>
<tr>
<td>D2K</td>
<td>77.6 ± 5.5</td>
<td>150.7 ± 4.6</td>
<td></td>
</tr>
<tr>
<td>K17E</td>
<td>55.5 ± 4.1</td>
<td>98.7 ± 6.7</td>
<td></td>
</tr>
<tr>
<td>V8N</td>
<td>33.5 ± 2.1</td>
<td>36.5 ± 2.1</td>
<td></td>
</tr>
<tr>
<td>RdDd*</td>
<td>30.0 ± 4.0</td>
<td>35.7 ± 4.0</td>
<td></td>
</tr>
</tbody>
</table>

* The half-life in the presence of 0.2 M KCl is 104 ± 13 min. *b The half-life values in the presence of 0.2 M KCl and 0.4 M trehalose are 11.2 ± 2.1 min and 19 ± 1.7 min, respectively. *c Values from Lamosa et al. [7].

### Structure determination of mutant Δ17|29 by NMR

Proton signal assignment was performed using the classical approach described by Wüthrich [12a]. Analysis of TOCSY and COSY spectra allowed the identification of the spin systems. Sequence-specific assignment was achieved using NOESY spectra and identifying connectivities between NH protons and between the NH and H protons of adjacent spin systems. The spin-systems for Met1 and Asp19 could not be identified, probably because mobility of the N-terminus and the loop region leads to weak signals. Spin diffusion was taken into account and a value of 6.2% was used to loosen all NOESY-derived constraints. Stereo-specific assignments were obtained using preliminary calculated structures with the aid of program GLOMSA; of these, 16 were derived from stereopairs with non-degenerate chemical shifts and 50 NOESY cross-peaks could be pseudo-stereospecifically assigned to one or the other side of the fast-flipping aromatic side chain rings.

The program INDIANA was used to generate 500 conformers from which the 20 structures with the lowest target functions were selected. A schematic representation of the 20 superimposed structures showing the backbone, aromatic side chains and cysteine sulfur atoms, is presented in Fig. 2A and a statistical analysis is given in Table 2. The metal centre conserves both the geometry and the chirality of the native protein and is well defined, with the heavy atoms of the four coordinating cysteines (residues 6, 9, 26 and 29) having an RMSD < 0.55 Å (Fig. 3). Analysis of the secondary structure with MOLMOL v. 2.6 [12] and PROCHECK-NMR showed the presence of a three-stranded β-sheet similar to that of the native protein (Fig. 2B). The Ramachandran plot shows that most of the residues (94.7%) fall in the most favoured and additionally allowed regions; however, 5.2% appear in the generously allowed and one residue (Asp19) appears in the disallowed region in one of the 20 structures. This residue is located in the residual loop of the mutant and, if only well-defined regions are considered (Table 2), no residue appears in the disallowed region. The deviation is probably a consequence of the large deletion (25% of the residues were deleted) straining the protein.
Overall, the structure of the mutant retains the main features of the native structure with the obvious exception of the loop region. The RMSD between the backbones of the mean structures for the native and mutant rubredoxins is 2.24 Å. However, if residues 16–22 (sequence numbering of the mutant), which make up the shortened loop region in the mutant, are excluded, the deviation decreases to 0.82 Å, showing that this large deletion left the remaining structure virtually unaltered (Fig. 2B). The optimal hydrogen bond network was calculated for each of the 20 structures and it is also similar to that displayed by the native protein [13]. However, the average exposure to water increased, especially in segment 16–23, with values over 40% observed for some of these residues (Fig. 3).

In particular, the exposure of the residues that comprise the lower part (relative to the orientation depicted in Fig. 2) of the hydrophobic core of the native protein, namely, Y4, Y13, F17, L20 and W24 (numbering according to the mutant) increased substantially.

**Table 2.** Restraint violations and quality analysis for the rubredoxin Δ17I29 mutant structure.

<table>
<thead>
<tr>
<th>DYANA target function</th>
<th>Average total (Å)</th>
<th>0.21 ± 0.021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function Range</td>
<td>0.16–0.23</td>
<td></td>
</tr>
<tr>
<td>Violated Constraints</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistent violations (&gt; 0.2 Å)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>van der Waals violations (&gt; 0.2 Å)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Precision (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean global backbone RMSD</td>
<td>0.98 ± 0.21</td>
<td></td>
</tr>
<tr>
<td>Mean global heavy atom RMSD</td>
<td>1.72 ± 0.26</td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favoured</td>
<td>58.3 (54.5)</td>
<td></td>
</tr>
<tr>
<td>Additionally allowed</td>
<td>36.5 (40.2)</td>
<td></td>
</tr>
<tr>
<td>Generously allowed</td>
<td>5.2 (5.2)</td>
<td></td>
</tr>
<tr>
<td>Disallowed</td>
<td>0 (0.2)</td>
<td></td>
</tr>
<tr>
<td>Nonredundant distance restraints (lower limits)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intraresidual</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>Sequential (</td>
<td>i</td>
<td>= 1)</td>
</tr>
<tr>
<td>Medium range (2 &lt;=</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Long range (</td>
<td>i</td>
<td>&gt; 5)</td>
</tr>
<tr>
<td>Total redundant and nonredundant</td>
<td>734</td>
<td></td>
</tr>
</tbody>
</table>

* Residues with $\phi$ < 50° and $\psi$ < 120° were not included for the Ramachandran plot calculation; the values obtained using all residues are shown in parentheses.
The structure of mutant Δ17/29 shows considerable similarities with RdDd, a protein naturally truncated in the loop region. In fact, excluding residues 16–22 (sequence numbering of the mutant), the RMSD between the backbones of the X-ray model of RdDd and the mutant rubredoxin is only 1.22 Å. However, if the residues corresponding to the residual loop region of mutant Δ17/29 are included, the RMSD between its backbone and that of RdDd increases to 2.64 Å. The most striking difference between mutant Δ17/29 and RdDd is the absence of a histidine residue in the mutant protein and the 6.2 Å shift of Phe17 (30 in the RdDd sequence).

Temperature dependence of amide chemical shifts

In general, the proton chemical shifts depended linearly on temperature, with the smallest coefficients observed in the metal binding loops (Fig. 6). The binding sequences X-Cys-X-X-Cys-Gly-X (X = variable amino acid) are largely conserved among rubredoxins, and comprise residues Val5 to Tyr11, and Ala25 to Ala31 in the Δ17/29 mutant. We found a reasonably good correlation between the existence of hydrogen bonds and amide protons with small absolute temperature dependence (values more positive than –4.5 ppm K⁻¹) have been proposed to be a reliable indicator of H-bonding especially if combined with slow exchange rates [14]. In the structure of mutant Δ17/29, among the residues with high probability of being involved in H-bonds according to analysis with the WHATIF software, 79% have temperature coefficients above –4.5 ppm K⁻¹ (Fig. 6A). The presence of DGP in the sample produced a generally small, but consistent increase in the temperature coefficients of amide protons, with the exception of Lys33 and Ala35, which
Fig. 4. Dependence of the amide proton chemical shifts for mutant A17I29 with solute concentration. Bars are arranged from left to right with increasing solute concentration: 1 mM (yellow), 2 mM (orange), 10 mM (green), 20 mM (grey), 50 mM (red), 100 mM (black), 200 mM (cyan) and 400 mM (purple). Experiments were performed at 303K. Concentration of DGP or KCl: 1, 10 (only for DGP), 50, 100 and 200; trehalose concentration: 2, 20, 100, 200 and 400; MG concentration: 2, 20, 100 and 200 mM. The horizontal axis represents the residue number in the protein sequence and the vertical axis has a total range of 0.14 ppm.
showed a much larger increase. A similar pattern characterizes the MG effect except that the temperature coefficient of the amide proton in Asp14 shows a large increase (threefold) that is not observed in the presence of DGP (Fig. 6).

**Discussion**

**Framework for data interpretation**

The aim of this study was to obtain knowledge about the molecular basis of protein stabilization by charged compatible solutes. A series of mutants of RdDg was constructed to investigate the importance of these mutations on the degree of stabilization rendered by solutes. It is well-documented that the thermal denaturation of rubredoxins occurs via a thermodynamically irreversible process [9–11]. Therefore, thermodynamic stability parameters are not accessible for these iron-proteins and the stability data reported here refer to kinetic stability, estimated from the half-life for release of the iron at 90 °C.

A link between kinetic stability (half-lives) and the conformational stability of the native forms needs to be established to provide a framework for interpreting our data. As the precise mechanism of denaturation in these proteins is unknown, that link has to be made on the basis of reasonable assumptions within a model for irreversible denaturation, and the simplest form of the Lumry–Eyring model [15] seems appropriate for this purpose. It is reasonable to suppose that the state recently identified in rubredoxins by LeMaster et al. [16], in which the hydrophobic core is clearly disrupted, represents the unfolded state that denatures irreversibly by loss of the metal centre. Exchange between this state and the native conformation was shown to be fast [16], in which case a shift in the equilibrium will change the half-life. Because the structure of the metal centre is unchanged in the mutants examined, we may assume that the rate constant for the irreversible step is similar in each case and, hence, the half-lives should correlate with the relative stability of the native and ‘unfolded’ states. This is far from providing a quantitative relationship between half-life and the

Fig. 5. Variation of the amide proton chemical shifts for mutant Δ17|29 with the concentration of DGP and MG after correction for the ionic strength effect. Chemical shift values obtained with KCl were subtracted from those obtained at the same concentration of DGP and MG. The bars shown are organized from left to right in increasing solute concentrations and refer to 1 (DGP), 2 (MG), 50 (DGP), 100 and 200 mM. The colour code for the solute concentration is: 1 mM (yellow), 2 mM (orange), 50 mM (red), 100 mM (black) and 200 mM (cyan). The horizontal axis represents the residue number in the protein sequence and the vertical axis has a total range of 0.14 ppm.
Fig. 6. Temperature coefficients of amide protons for mutant Δ17/29. (A) Temperature coefficient values for each amide proton of the protein in water (○), 200 mM DGP (■), 200 mM KCl (○) and 400 mM Mg (△). The lower plots show the difference of the coefficients in the presence of (B) DGP or (C) Mg relative to water. The points on the horizontal axis cover the entire sequence of residues of the protein and the vertical axis refers to the temperature coefficients expressed in ppb·K⁻¹. The horizontal line at -4.5 ppb·K⁻¹ was drawn to indicate the cut-off value proposed by Baxter and Williamson [14] for hydrogen-bonded amide protons.
intrinsic stability of the native conformations. Never-
theless, it provides the justification for seeking a link
between changes in half-life in the presence of solutes
and structural features of the native forms.

Intrinsic kinetic stability of native rubredoxins
and mutants

The deletion of the hairpin loop in RdDg induced a
strong decrease in thermal stability. Moreover, the
progressive increase in the number of deleted residues
of mutants Δ23/29, Δ17/26 and Δ17/29 was accompa-
nied by a progressive decrease of the intrinsic stability,
showing that this structural motif is particularly
important for the stability of the tertiary structure of
rubredoxins, probably by protecting the protein hydro-
phobic core from solvent access. Corroborating evi-
dence for this view emerges from the large increase
(2.3-fold) in the exposure of the hydrophobic core of
mutant Δ17/29 compared with the native structure. It
is worth pointing out here that the opening of the mid-
dle loop with the concomitant increase in the solvation
of the hydrophobic core has been proposed to trigger
the loss of the metal ion and subsequent unfolding of
Clostridium pasteurianum Rd [17].

The similarity of stability shown by mutant Δ23/29
and the native RdDd, which naturally lacks part of
the loop (Table 1), seems more than a coincidence,
and reinforces the positive contribution of this hairpin
structure to the stability of this family of proteins. In
apparent contrast, the shortening of loops observed in
thermophilic proteins compared with mesophilic coun-
terparts has been often proposed as a general strategy
for thermostabilization [18,19]. Most likely, this con-
tradiction arises from the fact that the favourable
effect of a shorter, rigid hairpin is outweighed in
rubredoxins by an increased solvent exposure of the
hydrophobic core, with the overall system becoming
less stable.

The calculated structure of mutant Δ17/29 revealed
notable features, such as a deep cavity in the molecule,
and extensive exposure of the aromatic side chains.
The minimal hairpin region of this mutant is respon-
sible for the cavity formation. RdDd also has a very
short loop but does not show this feature because a
histidine ring partially fulfils the structural role of the
loop [20]. It is remarkable that mutant Δ17/29 is able
to fold despite the drastic deletion; this reveals the
structural importance of the other unaffected motifs in
directing folding. Even with a severe disruption of the
middle loop, the rest of the characteristic features of
the protein structure remained virtually unchanged as
shown by the small RMSD value of 0.82 Å obtained
for the superimposition of the mutant and native mean
structures.

The progressive decrease in protein stability connec-
ted with the shortening of the loop region cannot be
ascribed to the size of the loop alone because all the
mutants examined, including those with point muta-
tions, showed a clear decrease in their intrinsic stabil-
ity. In the case of mutant V8N, the exchange of a
highly conserved aliphatic side chain for an uncharged
polar group had a striking effect, reducing the half-life
for iron release by 65%. Another study reporting
mutations on Val8 (V8A and V8D) [21] for more polar
residues demonstrated that the absence of a nonpolar
residue at this position dramatically decreased the pro-
tein stability. The aliphatic nature of this residue along
with three others (5, 38 and 41; numbering in the
sequence of RdDg) and their spatial positioning enable
them to pack together, thereby preventing exposure of
the metal centre to solvent water. The hydrophobic
cluster created by these residues is largely conserved
among rubredoxins [8,22] and acts like a cap on the
tetrahedrally coordinated iron centre, which probably
increases its rigidity and compactness, properties gen-
erally associated with highly thermostable proteins
[23–25].

The considerable decrease of stability (42%) caused
by mutation K17E is probably connected with the
addition of an extra negative charge in an already neg-
atively charged patch. The mutation D2K, by contrast,
caused only a small decrease of stability (20%).
Although this residue does not appear to interact spe-
cifically with any other region of the protein [13], it
has been hypothesized that the termini could play a
decisive role in the unzipping of the β-sheet [26] and
this may explain the observed decrease in stability.

Altogether, the decrease in the intrinsic stability of
all tested mutants, even in the case of single-residue
mutations shows that the native conformation of
RdDg is remarkably well designed for thermal stabil-
ity. Moreover, we showed that rubredoxin stability is
clearly dependent on the size of the loop region, but it
also depends, to a lesser extent, on subtle individual
contributions dispersed throughout the protein struc-
ture.

Stabilization by compatible solutes

To obtain insight into the mode of action of compat-
ible solutes, we examined the impact of several muta-
tions of RdDg on the degree of stabilization rendered
by DGP and MG. In addition, NMR was used to
characterize possible interactions of these solutes with
the native form of the most perturbed mutant, Δ17/29.
Structural determinants of protein stabilization

T. M. Pais et al.

The NMR structure calculation of mutant Δ17/29 has shown that, except for the original loop region, the rest of the protein backbone was virtually unchanged, making it reasonable to assume that the same applies to the other mutants where the deletion was less severe. This assumption is supported by the observation that all mutants retained the UV–visible spectrum displayed by the wild-type Rd, indicating that the metal centre geometry and basic structure was preserved in all the engineered proteins.

DGP exerted a remarkable stabilization on the wild-type Rd, but surprisingly, was extremely inefficient for the stabilization of the mutants with different size-deletions in the loop region. Given the fact that DGP is a charged solute it is pertinent to analyse the alterations in the electric charge distribution of the loop region associated with the engineering of the loop size. Mutant Δ17/26 has a net charge identical to that of the parent Rd, whereas mutant Δ23/29 shows a decrease of two positive charges and mutant Δ17/29 has a net loss of one positive charge. In the case of mutant Δ23/29, which is destabilized by the solute, the deletion of seven residues led to the formation of a cluster of four negatively charged residues (DPDSFED), not present in the other mutants that are stabilized by DGP. We hypothesize that the repulsive forces originated from this sequence could contribute to the negative effect exerted by DGP on the stability of this mutant. In agreement with this view, the RdDd, which naturally has a deletion of seven residues in the loop region but lacks this cluster of charged residues, is stabilized by DGP [7]. However, the explanation is surely more complex because MG, which is also negatively charged, does not destabilize mutant Δ23/29, and actually increased its half-life for iron release by 50%. The contrasting behaviour of these equally charged solutes is clear evidence for the distinct nature of the mechanisms underlying protein stabilization by MG and DGP. The differences are not restricted to this mutant. For example, the stability of the wild-type Rd was strongly enhanced by DGP but only modestly improved by MG. By contrast, our work demonstrates that minimal alterations in the protein sequence (single mutations) produce considerable differences in the extent of stabilization rendered by a given solute (Fig. 1). Altogether, these results consistently support the view that the effect induced by solutes on protein stability is strongly dependent on the specific protein/compatible solute system examined.

Given the observed specificity of the stabilizing effect, one could hypothesize the existence of specific interactions, or loci for preferential binding on the protein molecule. Proton chemical shifts are very sensitive probes of local fluctuations of the average chemical environment and therefore, were used to look for evidence of preferential interaction sites of the solute with the protein. The pattern of NH shifts induced by the three charged solutes (DGP, MG and KCl) was broadly similar. However, when the effect of ionic strength was discounted, the differences between DGP and MG became apparent (Fig. 5). The three stabilizing solutes (DGP, MG and trehalose) produce different patterns of chemical shift variation but of similar magnitude, which reinforces the idea of small, but distinct structural alterations, probably due to specific interactions with the protein surface. Solutes are generally regarded as causing no major change in protein structure given the low magnitude of chemical shift variations observed in the few studies available [27,28]. Although our results corroborate this broad view, we looked for evidence at a much finer level and found some evidence for the presence of small conformational changes. These changes may be large enough to improve the protein stability, and yet, as reflected by the low magnitude of the chemical shifts, too small to affect the overall structure, and probably the physiological function. We should bear in mind, however, that the NMR data was obtained at a temperature lower than the stability data and the solute/protein interactions could change with temperature.

Preferential sites for solute action are not clearly apparent and probably the interactions are spread throughout the protein surface; however, residues Cys9, Leu20 and Asp22 exhibit shifts that are well above the average, these features being common to DGP and MG. Ala25 also experiences a notable shift which is induced by MG only. Most of these residues (20–25) are located in the poorly structured residual loop and the large effect observed in Cys9 could indicate that this cysteine has the least stable conformation among the iron ligands.

Overall the stabilizing solutes produce mainly negative NH shifts, which is generally associated with stronger hydrogen bonds and therefore a tighter protein structure. The same general effect on chemical shifts is observed upon lowering the temperature of protein solutions. Further evidence for a more compact structure in the presence of solutes is provided by the observed increase in the temperature coefficients of NH groups (Fig. 6). In fact, the signals of amide protons involved in hydrogen bonds generally shift less with temperature [14]. Therefore, the tendency to increase the coefficients in the presence of the stabilizing solutes reflects the strengthening of the hydrogen bond network. These findings are in line with an earlier study about the effect of DGP on the dynamics of
wild-type RdDg, which showed a restriction of large-scale motions induced by solute addition [28].

In summary, this study provides indication for the existence of at least two effects that could play a role in the complex strategy by which solutes confer a higher stability on proteins: an induced overall compaction of the native protein, and specific, weak interactions of the solutes with the protein surface.

**Experimental procedures**

**Mutagenesis of Rd and protein production in* Escherichia coli***

Plasmid pRPPL1 [7], harbouring the gene encoding RdDg under the control of a heat sensitive promoter, was constructed from pCYTEXP1 [29] and used as a template for all mutations. Mutants Δ17/29, Δ23/29 and Δ17/26 were obtained using the ExSite™ PCR-Based Site-Directed Mutagenesis Kit according to the procedures outlined in the respective kit instructions (Stratagene, La Jolla, CA, USA). Mutants D2K and K17E were constructed to evaluate the effect of total surface charge on the stabilizing efficacy of the solutes. A third mutant (V8N) was designed to assess the effect of changing the most exposed hydrophobic residue of the native structure. These mutants were constructed using the QuikChange™ Site-Directed Mutagenesis Kit (Stratagene). The sequence of the coding unit of each mutant was confirmed by restriction analysis of DNA isolated from positive recombinant clones. The resulting plasmids, encoding mutant proteins D2K, K17E and V8N, were cloned in *Escherichia coli* strain XL1BLUE. This system led to very low yields for the expression of mutants Δ23/29, Δ17/26 and Δ17/29, hence the respective coding units were transferred to an isopropyl thio-β-d-galactoside-inducible plasmid, p7-7 [7].

Mutant proteins were purified using three chromatographic steps as previously described [7]. After the last column in the purification procedure (Resource Q; Pharmacia, Uppsala, Sweden), iron (red) and zinc (colourless) forms of rubredoxin mutants were separated and judged pure by silver-staining native PAGE electrophoresis.

**Thermal stability assays**

The kinetics for the disruption of Rd structure at 90 °C was monitored by UV–visible absorption spectroscopy in a Shimadzu UV-1601 spectrophotometer equipped with a thermostated cell. A rubber septum was adapted to a quartz cell to allow measurements under anaerobic conditions. The assay solution consisted of 50 mM Tris/HCl buffer, pH 7.6, and the desired concentration of a given solute. The temperature of the solution in the spectrophotometric cell was measured with a thermocouple. Once thermal equilibrium was reached, the desired amount of protein solution was rapidly added and spectral scanning started. Spectra were recorded for each time point and baseline corrected. The values of absorbance measured at 494 nm (A494) as a function of time were fitted to a single exponential decay. The effect of solute concentration on the degree of stabilization of the native RdDg was studied in the range 0–0.5 M. DGP, at a concentration of 0.1 M, stabilized the protein to the same extent as 0.5 M. We also observed that MG at a concentration of 0.2 M exerted an effect comparable to that at 0.5 M. Therefore, unless otherwise stated, the final concentrations of DGP and MG in the assay solutions were 0.1 and 0.2 M, respectively.

**NMR structure calculation and analysis**

1H-NMR spectra of zinc rubredoxin Δ17/29 were acquired on a Bruker DRX500 (Bruker, Rheinstetten, Germany) spectrometer equipped with a 5 mm probe head with internal B0 gradient coils. Protein NMR samples (4 mM) were prepared in 10% 2H2O and the pH adjusted to 7.6. Assignment of the proton signals was performed in spectra acquired at 303 K, but additional spectra were obtained at 313 K to help resolve peak overlap, especially in the aromatic region. NOESY spectra (mixing times of 35, 70, 100 ms), TOCSY [30] using the clean total correlation spectroscopy pulse sequence with spin lock times of 70 and 100 ms, and COSY spectra were recorded. Raw data was processed using standard XWIN-NMR software (Bruker). Polynomial baseline corrections were applied in both dimensions of each spectrum. The software XEASY (v. 1.3.10) was used for assignment and integration of NOESY cross-peak volumes. NOEs were measured at 303 K in the 80 ms NOESY spectrum.

The NOESY cross-peak volumes were used to calculate upper (upl) and lower (lol) limit volumes using the program INDYANA [31]. Nonstereospecifically assigned protons, degenerate protons, overlapping peaks, and flexible proline rings were treated as previously described [31,32]. Stereospecific assignments were determined with the help of the program GLOMSA [33]. Upper and lower distance limits for each pair of sulfur atoms involved in zinc coordination (C6, C9, C26 and C29) were fixed between 3.9 and 3.5 Å [8,13,34]. This range of distances allows a significant distortion from tetrahedral geometry and does not specify the chirality of the centre. The experimental distance restraints were then used as input to generate protein conformers using the program for restrained dynamics and simulated annealing, DYANA v. 1.4 with modifications (INDYANA) to optimize scaling factors for calibrating NOE intensities [32]. In the final refinement stages, each batch of structures was checked for the existence of short distances (< 2.5 Å) between protons for which no NOE had yet been measured. If an expected NOE was not visible in a clear region of the
spectra, the volume at that frequency was measured and used to provide a lower limit distance in subsequent structure calculations, thus reducing the possibility of incorrect short interproton distances. A complete relaxation matrix analysis was applied to an ensemble of 10 structures to estimate the error that might be introduced via spin-diffusion; this value was used to loosen all distance constraints in subsequent calculations [32].

The program MOLMOL v. 2.6 [12] was used for superimposition and visual inspection of the final family of structures. The NMR structures were analysed with respect to experimental constraints using standard procedures of the DYANA program. The quality of structures with respect to dihedral angles was evaluated using the program PROCHECK-NMR v. 3.4.4 [35] and Ramachandran plots were generated. The optimal hydrogen bond network was calculated for each structure using program WHAT-IF v. 5.0 [36]. The atomic coordinates and constraint files were deposited in the Protein Data Bank at the Research Collaboratory for Structural Informatics-Rutgers with the Accession code 1SPW.

Proton chemical shift variation

Samples of mutant Zn rubredoxin Δ17/29 with a concentration of 0.4 mM and pH 7.6 were prepared in 10% 2H2O. Chemical shifts were followed in TOCSY spectra [10] obtained with water presaturation and phase-sensitive mode using TPP1. Spectra were acquired in the presence of different concentrations of DGP, KCl, trehalose or MG. Temperature dependence of proton chemical shifts was evaluated between 273 and 303 K at 2.5-degree intervals in the absence of solutes or in the presence of 200 mM DGP, 200 mM KCl or 400 mM MG. This dependence was expressed by the respective temperature coefficient (10⁻³ ppm K⁻¹).

Proton chemical shifts are referenced to internal 3-(trimethylsilyl)propanesulfonic acid (sodium salt), and analysed using XEASY (version 1.3.10) software. Secondary shifts were obtained by subtraction of random coil chemical shifts [12a] from the experimental ones.

Acknowledgements

This work was supported by the European Commission (Contracts QLK3-CT-2000-00640 and COOP-CT-2003-508644 and by Fundação para a Ciência e a Tecnologia (FCT), Portugal, and FEDER (POCTI/BME/35131/2000).

References

5 Petsko GA (2001) Structural basis of thermostability in hyperthermophilic proteins, or ‘there’s more than one way to skin a cat’. Methods Enzymol 334, 469–478.