Thioredoxin Fold as Homodimerization Module in the Putative Chaperone ERp29: NMR Structures of the Domains and Experimental Model of the 51 kDa Dimer

E. Liepinsh,[1] M. Baryshev,[2,3] A. Sharipo,[3]
M. Ingelman-Sundberg,[2] G. Otting,[1,4]
and S. Mkrtchian[2]
[1] Department of Medical Biochemistry
  and Biophysics
[2] Division of Molecular Toxicology
Institute of Environmental Medicine
Karolinska Institute
171 77 Stockholm
Sweden
[3] Biomedical Research and Study Center
  of Latvian University
LV-1067 Riga
Latvia

Summary

Background: ERp29 is a ubiquitously expressed rat endoplasmic reticulum (ER) protein conserved in mammalian species. Fold predictions suggest the presence of a thioredoxin-like domain homologous to the *a* domain of human protein disulfide isomerase (PDI) and a helical domain similar to the C-terminal domain of P5-like PDIs. As ERp29 lacks the double-cysteine motif essential for PDI redox activity, it is suggested to play a role in protein maturation and/or secretion related to the chaperone function of PDI. ERp29 self-associates into 51 kDa dimers and also higher oligomers.

Results: 3D structures of the N- and C-terminal domains determined by NMR spectroscopy confirmed the thioredoxin fold for the N-terminal domain and yielded a novel all-helical fold for the C-terminal domain. Studies of the full-length protein revealed a short, flexible linker between the two domains, homodimerization by the N-terminal domain, and the presence of interaction sites for the formation of higher molecular weight oligomers. A gadolinium-based relaxation agent is shown to present a sensitive tool for the identification of macromolecular interfaces by NMR.

Conclusions: ERp29 is the first eukaryotic PDI-related protein for which the structures of all domains have been determined. Furthermore, an experimental model of the full-length protein and its association states was established. It is the first example of a protein where the thioredoxin fold was found to act as a specific homodimerization module, without covalent linkages or supporting interactions by further domains. A homodimerization module similar as in ERp29 may also be present in homodimeric human PDI.

Introduction

Protein maturation in the luminal compartment of the endoplasmic reticulum (ER) is accomplished by the tight network of molecular chaperones and folding enzymes assisting in the import, folding, and assembly of the nascent polypeptides and securing the proper quality control of the proteins exiting from the ER [1, 2]. The most abundant folding enzyme of the ER is protein disulfide isomerase (PDI), which catalyzes the formation and rearrangement of disulfide bonds but may also act as a general chaperone [3]. The family of eukaryotic PDIs includes ERp72, ERp57, P5, PDIp, and PDIR [4]. A recent and unusual addition to this family is a redox-inactive and ubiquitously expressed rat endoplasmic reticulum protein, ERp29 [5–7].

After cleavage of the N-terminal signal peptide, mature ERp29 contains 228 residues (Figure 1). It forms a 51 kDa homodimer and also higher oligomers both in vivo and in vitro [8]. Amino acid sequence comparisons predict the presence of a thioredoxin-like N-terminal domain in ERp29 and a helical C-terminal domain of unknown fold. The sequences of both domains align with sequences found in the group of P5-like PDIs (also termed PDI-Dα [4]) from plants (*Medicago sativa* [alfalfa], *Arabidopsis thaliana*, *Nicotiana tabacum*, and others) and amoebae (*Dictyostelium discoideum*) (Figure 1). In contrast to these proteins, ERp29 does not contain the active-site double-cysteine motif which is a hallmark of redox-active PDIs. ERp29 is conserved among mammalians, including humans [9, 10]. In addition, the *windbeutel* gene product from *Drosophila* is highly similar to ERp29 (>30% sequence identity) and lacks the double-cysteine motif at the canonical location, while a double-cysteine motif of unknown function appears near the amino terminus (Figure 1) [11]. The Windbeutel protein has been shown to act as a dedicated chaperone in facilitating the function and specific Golgi targeting of Pipe, a putative oligosaccharide-modifying enzyme essential for embryonic development [12].

While the function of ERp29 cannot be predicted from its amino acid sequence, growing evidence suggests its involvement in the protein maturation and/or secretion processes in the ER, in analogy to the *Drosophila* homolog Windbeutel. It is induced in certain cell types under stress conditions characterized by the accumulation of unfolded proteins in the ER and elevated levels of molecular chaperones [7]. In rat enamel cells and other secretory cells, ERp29 levels may reach those of the major chaperones and folding enzymes of the ER [10]. The human ortholog of ERp29, ERp28 (Figure 1), can be coimmunoprecipitated with overexpressed hepatitis B small surface antigen [9]. Similarly, ERp29 has been found to associate with the misfolded κ isoform of immunoglobulin light chains (S.M., unpublished data) and with the transport-incompetent form of immunoglobulin

[4] Correspondence: gottfried.otting@mbb.ki.se

(a)

```
                α1              ß1                          α2          ß2
                40          50              60          70      80          90
ERp29     33 LHTKGALPLDTVTFYKVIPKSK-FVLVKFDTQYPY--GEKQDEFKRLAENSA-SSDDLLVAEVGISDY    96
ERp28     33 LHTKGALPLDTVTFYKVIPKSK-FVLVKFDTQYPY--GEKQDEFKRLAENSA-SSDDLLVAEVGISDY    96
Wind      20 VTCTGCVDLDELSFEKTVERFP-YSVVKFDIAYPY--GEKHEAFTAFSKSAHKATKDLLIATVGVKDY    86
Aa-PDI    26 VSADDVVVLTEENFEKEVGHDK-GALVEFYAPWCGHCKKLAPEYEKLPNSFK-KAKSVLIAKVD----   87
Aa-PDI   144 TAPSHVVVLTPETFNEVVLDGTKDVLVEFYAPWCGHCKSLAPIYEKVAAVFK-SEDDVVIANLD----  206
Dd-PDI    19 SAEGNVVVLSPDNFDTVVDGS-KTVFVKFYAPWCGHCKKLAPDFEILADTFAPVSNKVVIAKVD---C   82
Dd-PDI   139 KAPSNVVDLSPSNFDSVVLDKSKNVLVEFYAPWCGHCKKLMPDYEILGNTYA-NEKDVVIAKIDC---  202
PDI-a      4 EEEDHVLVLRKSNFAEALAAHK-YLLVEFYAPWCGHCKALAPEYAKAAGKLKAEGSEIRLAKVDATEE   70
PDI-b    120 -----ATTLPDGAAAESLVESSEVAVIGFFK----DVESDSAKQFLQAEAIDDI-PFGI-TSNS---  173
PDI-b´   218 ---PLVIEFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLSNFKTAAESFKGKIL-FIFIDSDHTDN  282
PDI-a´   352 -----VKVLVGKNFEDVAFDEKKNVFEFYAPWCGHCKQLAPIWDKLGETYKDHEN-IVIAKMDST--  416

             α3              ß3                                 α4
                100         110         120         130             140         150
ERp29     97 GDKLNMELSEKYKLDKESYPVFYLFRDGDFENPV-----PYSGAVKVGAIQRWLKGQGVYLGMPG-C  157
ERp28     97 GDKLNMELSEKYKLDKESYPVFYLFRDGDFENPV-----PYTGAVKVGAIQRWLKGQGVYLGMPG-C  157
Wind      87 GELENKAIGDRYKVDDKNFPSIFLFKGNADEYVQ----LPSHVDVTLDNLKAFVSANTPLYIGRDGC  149
Aa-PDI    88 -CDEHKSVCSKY--GVSGYPTIQWFPKGSLEPKK------FEGPRTAESLAEFVNTEGGTNVKIA   143
Aa-PDI   207 -ADKYRDLAEKY--DVSGFPTLKFFPKGN-KAGE-----DYGGGRDLDDFVAFINEKS--------  255
Dd-PDI    83 DQADNKAICSKY--DVSGYPTLKIFDKST-TAKD------YNCARSVDELLTYINNHAKTNVKVK   138
Dd-PDI   203 DAADNKAICSKY--GVTGFPTLKWFGKQS-KDGE-----KYEQGRDLDTFINYINKQA--------  252
PDI-a     71 -----SDIAQQYGVR--GYPTIKFFRNGDTASPK-----EYTAGREADDIVNWLKKR        115
PDI-b    174 ------DVFSKYQLD--KDGVVLFKKF-D-EGRN-----NEEGEVTKENLLDFIKHNQ        217
PDI-b´   283 QRILEFFGLKKEE-----CPAVRLITLEEEMTKYKPESEELTAERITEFCHRFLEGKI        334
PDI-a´   417 -----ANEVEAVKVH--SFPTLKFFPASADRTVI-----DYNGERTLDGFKKFLESGG        462
```

(b)

```
              α5              α6                      α7
                160         170         180         190         200         210
ERp29    158 LPAYDALAGQFIEASSREARQAILKQGQDGLSGVKETDKKWASQYLKIMGKILDQGE  214
ERp28    158 LPVYDALAGEFIRASGVEARQALLKQGQDNLSSVKETQKKWAEQYLKIMGKILDQGE  214
Wind     150 IKEFNEVLKNYANIPDAEQLKLIEKLQAKQEQLTDPEQQQNARAYLIYMRKIHEVGY  206
Aa-PDI   274 --DLDELVKEFVAAND-EEKKAVFARIEEEVKKLEGSASRYGKIYLKVSKKYLEKGS  327
Dd-PDI   270 -EQLDTIATEFIAAAA-EVRKELVKKAQTVVDSLPEELRIEGSYYVKVMKTIAEKSI  324
                  ■ ■■ ■   ■     ■          □□ ■       ■   □  □■ □■■■ ■■   ■

              α8                      α9
                220         230         240         250             260
ERp29    215 DFPASELARISKLIEN-KMSEGKKEELQRSLNILTAFRKKGA-----EKEEL  260
ERp28    215 DFPASEMTRIARLIEKNKMSDGKKEELQKSLNILTAFQKKGA-----EKEEL  261
Wind     207 DFLEEETKRLLRLKAG-KVTEAKKEELLRKLNILEVFRVHKVTKTAPEKEEL  257
Aa-PDI   328 DYAKNEIQRLERLLEK-SISPAKADELTLKKNILSTYA            364
Dd-PDI   325 DFVTTEIARITKLVSG-SMSGKKADEFAKKLNILESFKSK          363
                □■  □□  ■   □      □        ■□ □ ■■■■■      □
```
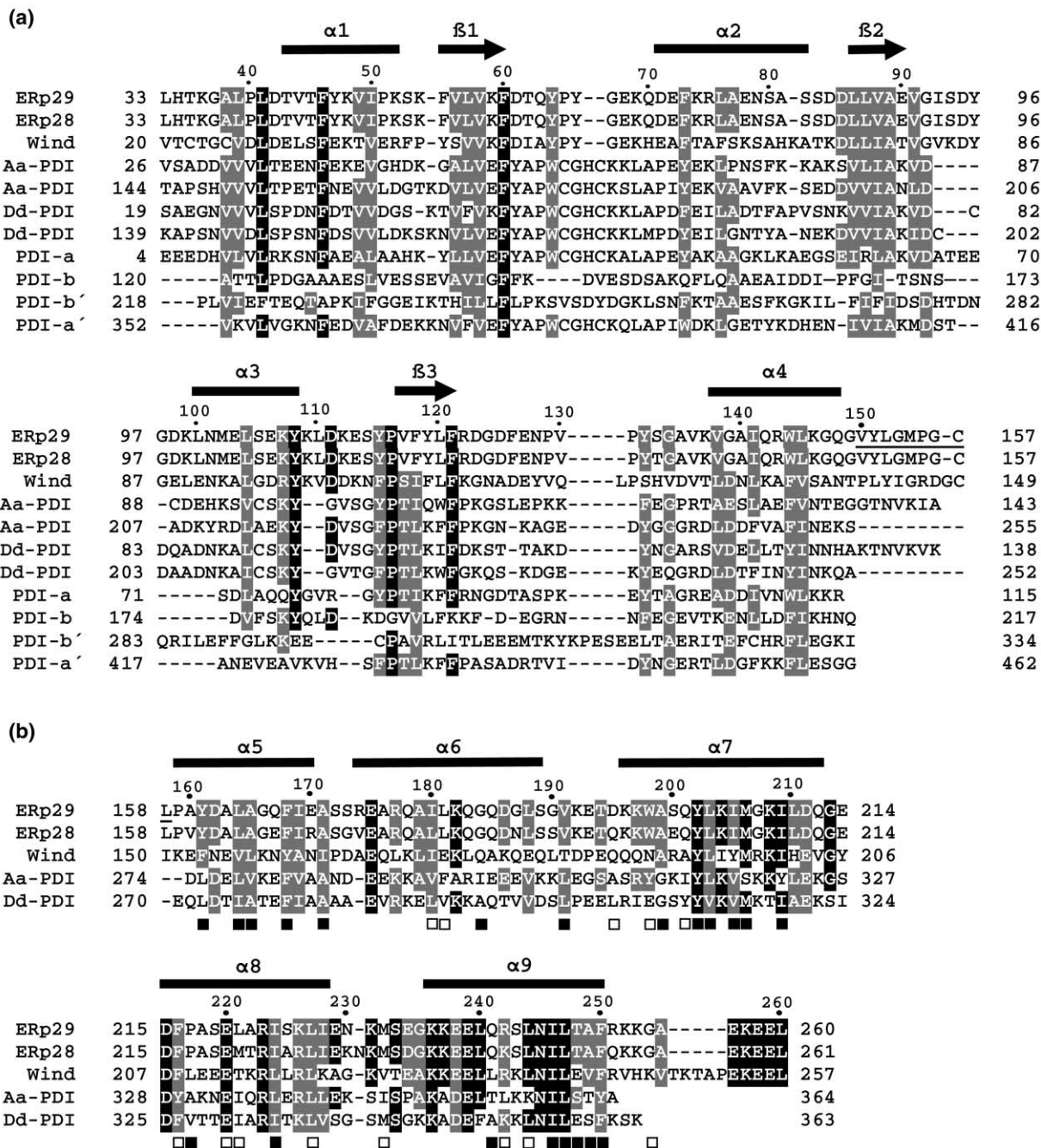
Figure 1. Amino Acid Sequence of ERp29 and Alignment with Selected Proteins

Bars at the top indicate the positions of regular secondary structure elements in ERp29. The residue numbering of ERp29 is shown at the top. The linker residues between the N- and C-terminal domains are underlined. Alignments for both domains are shown for the ERp29 homologs, human ERp28 (GenBank XM007009), *Drosophila* Windbeutel (AF025408), and the P5-like PDIs from alfalfa (Aa-PDI; P38661) and *Dictyostelium discoideum* (Dd-PDI; AAB86685). The latter two proteins have two thioredoxin-like domains, both of which were included in the alignment. Dark shading indicates identical residues; lighter shading indicates conservative changes. (a) N-terminal domain of ERp29. Additional alignments are with the *a*, *b*, *b´*, and *a´* domains of human PDI. (b) C-terminal domain of ERp29. Filled and open boxes at the bottom identify buried residues for which side chain solvent accessibilities of less than 10% and 20%, respectively, were calculated from the NMR conformers, when compared to the solvent accessibility of the corresponding side chain in a conformation with maximum solvent exposure.

heavy chains in the multipartite complex containing all major ER chaperones (L. Hendershot, personal communication). Furthermore, ERp29-BiP complexes were observed in FAO rat hepatoma cells [7]. Finally, ERp29 seems to be involved in thyroglobulin processing, as its mRNA expression is enhanced 3-fold in rat thyrocytes upon induction by thyroid-stimulating hormone [13]. These data demonstrate the capability of ERp29 to bind
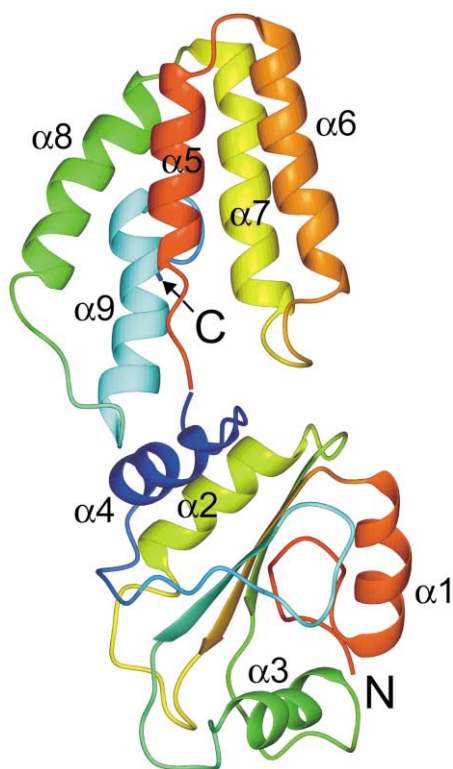
Figure 2. Ribbon Diagram of an ERp29 Monomer as Determined by NMR Spectroscopy

The structures of the N-terminal domain (bottom) and the C-terminal domain (top) were determined individually. In both domains, the colors change from red (N-terminal ends) to blue (C-terminal ends). The linker segment, interrupted between residues Met154 and Pro155 in the drawing, is continuous in the full-length protein.

to a range of different proteins and suggest a role similar to human PDI, which can assist protein folding in a disulfide-independent manner [14, 15].

Only limited structural information is available on PDIs. Crystal structures have been determined of the PDI from *Pyrococcus furiosus* [16] and of *E. coli* DsbA [17] and DsbC [18]. The latter two proteins do not possess all of PDI's activities but are involved in disulfide bond processing in a similar way as eukaryotic PDIs. Besides human PDI, only DsbC displays additional chaperone activity [19]. The structures of the *a* and *b* domains of human PDI have been solved individually by NMR spectroscopy [20–22], and progress toward structure determinations of the *a′* and *b′* domains has been reported [23, 24].

To provide a basis for further functional studies, we determined the three-dimensional structure of ERp29 (Figure 2). Initial attempts to crystallize ERp29 failed. Consequently, we have determined the three-dimensional structures of the N- and C-terminal domains separately by NMR spectroscopy. In addition, the full-length protein was studied to identify the dimerization domain and interface, interdomain mobility, and sites involved in the formation of higher oligomers. ERp29 presents one of the largest systems that has been studied by NMR spectroscopy at this level of detail.

Results and Discussion

Protein Expression

Full-length ERp29 (residues 33–260) was expressed as a 27 kDa construct with the N-terminal His tag sequence MRGSHHHHHHGS. Two additional constructs were designed, comprising the individual domains with N-terminal His-tag. The N-terminal domain (Leu33–Met154) was expressed with the same tag as the full-length protein. The construct of the C-terminal domain (Met154–Leu260) was preceded by the tag sequence MRGS HHHHHHGIR. Whereas the C-terminal domain was readily expressed in *E. coli*, the N-terminal domain expressed poorly and was prone to irreversible precipitation. Therefore, NMR samples of the N-terminal domain were prepared by the expression of full-length ERp29 followed by chemical cleavage at Cys157 by the cysteine-specific reagent NTCB and purification as described in Experimental Procedures. The presence of the His-tag did not interfere much with the NMR analysis, as no clear cross peaks could be observed for any of the histidine protons.

Structure of the N-Terminal Domain of ERp29

The structure of the N-terminal domain of ERp29 was determined in aqueous solution at pH 4.9, 31°C, using protein concentrations of about 0.5 mM. Although the NMR signals were broader than expected for a monomeric protein (Figure 3), no intermolecular nuclear Overhauser effects (NOEs) could be identified without ambiguity, as the analysis of weak NOEs was hampered by spectral overlap encountered in the two-dimensional (2D) NOESY spectra, the presence of impurities, and insufficient sensitivity in the 3D NOESY-HSQC spectra. Thus, all distance restraints were taken as arising from a monomer, and the structure at the dimer interface may be distorted as a result.

The NMR structure of the N-terminal domain of ERp29 (Figures 2 and 4) resembles the *a* domain of human PDI, with helices for residues 43–52, 71–82, 100–108, 138–148, and strands for residues 55–60, 86–90, and 117–121. The three-stranded β sheet is extended on either side by hydrogen bonds between residues 39 and 88 and between 120 and 130, but too few residues of these outer peptide segments align to define a proper five-stranded sheet as in thioredoxin or the PDI *a* domain. The structure is well defined for residues 35–152. Broad line shapes and no long-range NOEs were observed for the residues outside this region, suggesting increased mobility.

All peptide bonds of the NMR structure are in *trans* orientation. In particular, Pro116 was modeled with a *trans* peptide bond, although only one of the two sequential $d_{\alpha\delta}$ NOEs seemed to be intense, and the absence of a sequential $d_{\alpha\alpha}$ NOE could not be verified, due to overlap with t_1 noise from the residual water resonance. This residue is conserved between ERp29 and PDIs (Figure 1) and forms a *cis* peptide bond in most thioredoxin-like domains. *trans* peptide bonds have, however, also been observed. Examples are glutathione peroxidase [25] and the spliceosomal protein U5 [26], where the corresponding residues are nonproline
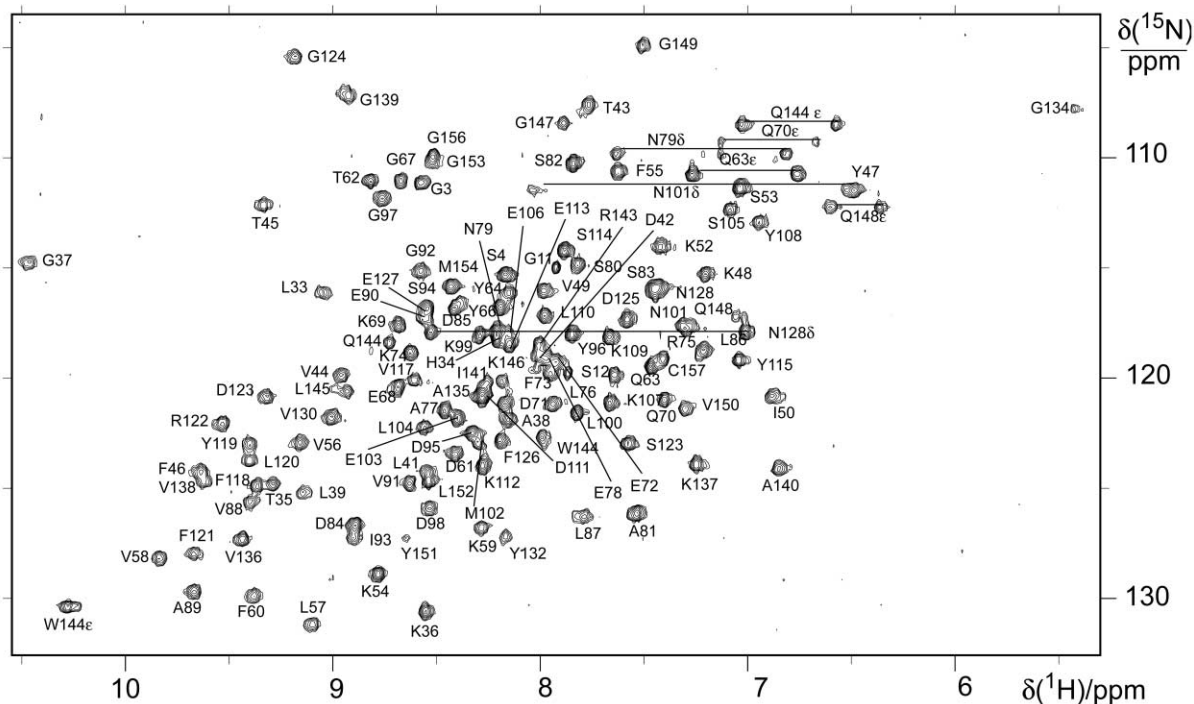
Figure 3. ¹⁵N-HSQC Spectrum of the N-Terminal Domain of ERp29, Showing the Resonance Assignments of the Amide Protons

Side chain amides and the indol NH of Trp144 are identified by Greek characters. The sequence numbers used are those of Figure 1, while residues from the N-terminal tag are numbered 1–12.

residues, and the third domain of calsequestrin, which has a proline at this position [27]. Like in ERp29, the thioredoxin-like domains of all these proteins are devoid of the redox-active Cys-X-X-Cys motif. When thioredoxin and glutaredoxin bind peptides via a disulfide bridge to the Cys-X-X-Cys sequence, additional contacts take place at a site located between this motif and the *cis* proline corresponding to Pro116 in ERp29 [28]. There is no evidence that the corresponding site in ERp29 is involved in peptide binding or dimerization.

Structure of the C-Terminal Domain of ERp29

The structure determination of the C-terminal domain of ERp29 was performed under the same conditions as that of the N-terminal domain, except that the protein concentration was about 2 mM. The widths of the NMR signals were characteristic of a monomeric protein (Figure 5). The higher solubility and narrower line widths greatly facilitated the structure determination compared to the N-terminal domain, resulting in more restraints per residue and better Ramachandran statistics (Table 1). The structure (Figures 2 and 6) contains five helices and presents a novel fold, as indicated by the failure to find any protein structure matching all helices in a search with DALI [29]. The helices comprise residues 159–170, 174–189, 196–212, 215–228, and 236–250. Very narrow resonances were observed for residues 254–260, indicating increased mobility on the subnanosecond time-scale for these C-terminal residues that comprise the ER-retrieval signal peptide sequence KEEL (Figure 1). No significant line narrowing was observed for the N-terminal residues after the His-tag. Furthermore, the first

long-range NOE was observed already for Leu158, indicating a relatively rigid conformation, despite the variation observed between the different NMR conformers (Figure 6b).

In many proteins with all-helical domains, the helices tend to arrange in a right-handed superhelix which can persist over many helical turns as in the structure of lytic transglycosylase (Protein Data Bank [PDB] code 1QSA) [30]. The structure of the C-terminal domain of ERp29 is unusual in that the helices arrange in a super-secondary structure that is partly right handed and partly left handed. The structure presents a stable fold, as slowly exchanging amide protons could be observed for all helices after dissolving the lyophilized protein in D_2O.

The sequence homology between the C-terminal domains of different PDI-like proteins also suggests structural conservation, which is strongly supported by the present NMR structure. Virtually all of the buried residues of the C-terminal domain of ERp29 are hydrophobic in the homologous domains (Figure 1b), and also the salt bridges observed in ERp29 between Lys208 and Asp211, Glu220 and Arg223, Lys226 and Glu229, and Lys236 and Glu239 seem to be conserved. Glu240 in the Windbeutel protein (corresponding to the buried residue Thr248 in ERp29) may participate in a salt bridge with Arg218 (Arg223 in ERp29).

Homodimerization of ERp29

Cross-linking experiments showed that ERp29 forms homodimers both in vivo and in vitro [8]. The line widths observed for the NMR signals of the individual domains
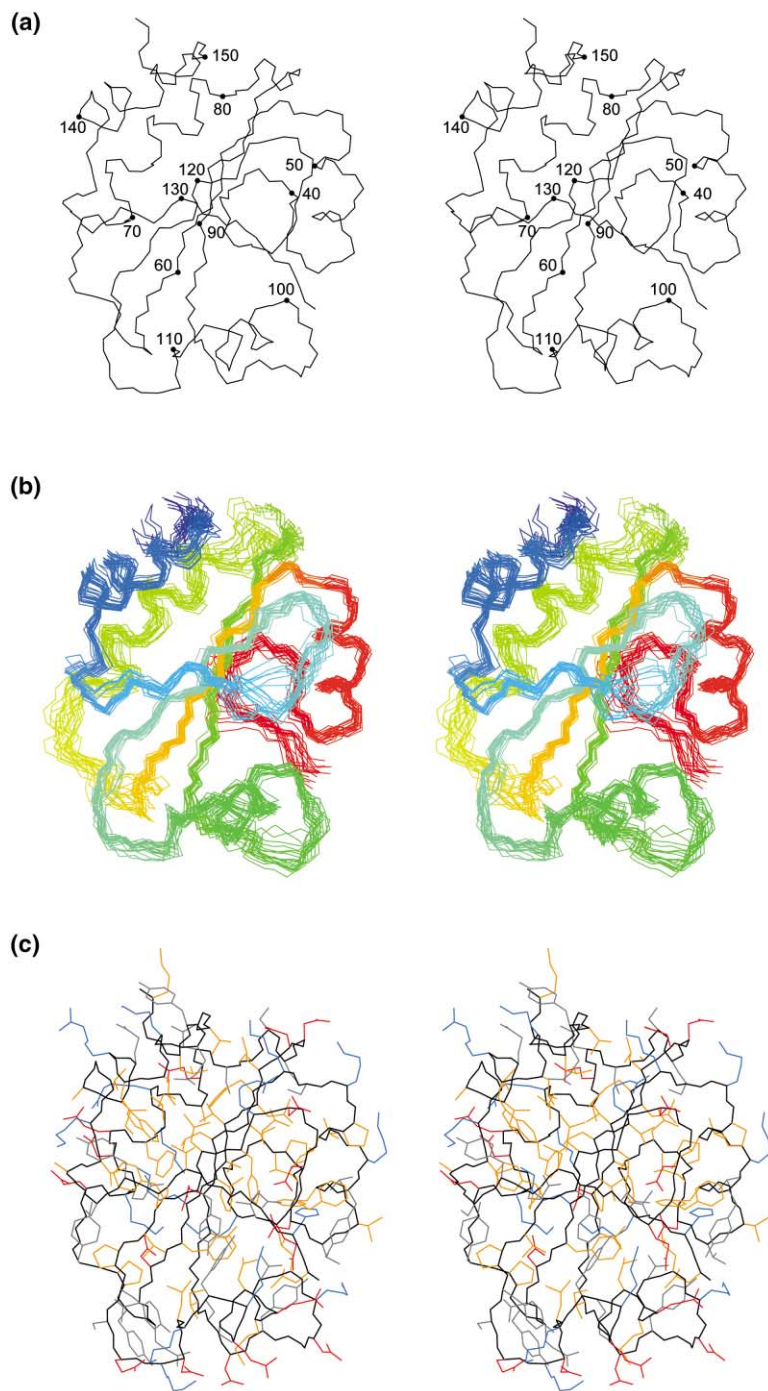
(a)

(b)

(c)

Figure 4. Different Stereorepresentations of Residues 33–154 of the NMR Solution Structure of the N-Terminal Domain of ERp29

All three representations show the structure in the same orientation as in Figure 2. (a) Backbone trace of the conformer with the lowest energy after restrained energy minimization. (b) Ensemble of 20 conformers, using the backbone atoms of residues 35–152 for superposition. (c) Heavy-atom display of the conformer of (a). The following color code was used: black, protein backbone; yellow, hydrophobic side chains (Ala, Ile, Leu, Met, Phe, Pro, Trp, Val); gray, polar side chains (Asn, Gln, His, Ser, Thr, Tyr); blue, positively charged side chains (Arg, His, Lys); red, negatively charged side chains (Asp, Glu).

indicated that homodimerization is exclusively mediated by the N-terminal domain. This view was confirmed by a more detailed cross-linking experiment, where the cross-linked product was chemically cleaved at Cys157 and the fragments analyzed by immunoblotting using domain-specific antibodies (Figure 7a). Uncleaved monomeric and dimeric species and an additional ∼42 kDa band are visible on both blots (lanes 3 and 3′), while the band of ∼31 kDa size is absent from the sample developed by the C-terminal domain-specific antibodies (lane 3′). The latter band corresponds to the dimer consisting of two N-terminal domains (2 × 15.5 kDa), whereas the 42 kDa fragment recognized by both antibodies probably contains two cross-linked N-terminal domains and one C-terminal domain, as expected for incomplete cleavage of a cross-linked dimer. No evidence for two cross-linked C-terminal domains (23 kDa) could be found. The bands of the monomeric cleaved domains were clearly resolved on the gels, confirming the specificity of the antibodies used.

Several NMR experiments were performed to check for the involvement of the C-terminal domain in homodi-
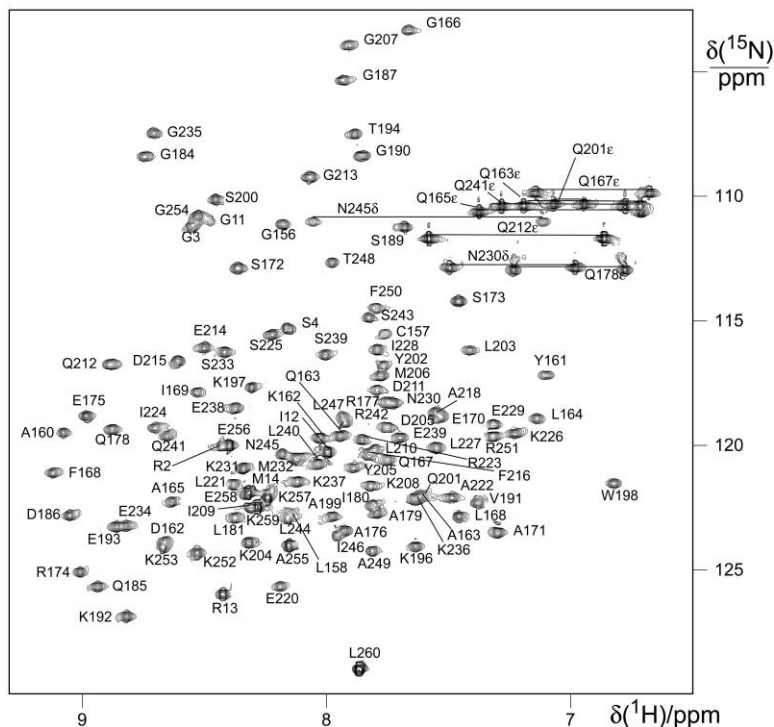
Figure 5. ^{15}N-HSQC Spectrum of the C-Terminal Domain of ERp29, Showing the Resonance Assignments of the Amide Protons

Side chain amides and the indol NH of Trp198 are identified by Greek characters. The sequence numbers used are those of Figure 1, while residues from the N-terminal tag are numbered 1–13.

merization. ^{15}N-HSQC spectra recorded of full-length ^{15}N/^{13}C/^{2}H-labeled ERp29 at pH 4.9 showed weaker resonances for the N-terminal than for the C-terminal domain, indicating that the C-terminal domain is considerably more mobile in the dimer than the N-terminal domain. These results were confirmed by the $T_1(^{15}$N) relaxation times measured for the amide nitrogens of the full-length protein, which showed on average higher values for the N- than for the C-terminal domain (Figure 8a). Furthermore, the backbone chemical shifts hardly changed between the individual domains and the full-length protein, except for residues of the linker region. Finally, upon addition of unlabeled C-terminal domain to ^{15}N-labeled N-terminal domain or addition of unla-

Table 1. Structural Statistics for the NMR Structures of the N- and C-Terminal Domains of ERp29

Parameter	Value	
	N Domain	C Domain
Assigned NOE cross peaks	1845	1956
Nonredundant NOE upper distance limits	1279	1492
Stereospecific assignments	65	121
Scalar coupling constants[a]	185	340
Dihedral angle restraints	358	328
AMBER-energy (kcal/mol)	-5338 ± 166	-5759 ± 89
Residual NOE restraint violations (Å)		
Sum	18.1 ± 0.7	20.7 ± 0.3
Maximum	0.11 ± 0.10	0.11 ± 0.00
Residual dihedral-angle restraint violations (°)		
Sum	128.0 ± 11.0	69.9 ± 5.4
Maximum	2.9 ± 1.2	2.7 ± 0.2
Rmsd[b,c] (Å)		
Backbone atoms N,C$^\alpha$, C′	0.83 ± 0.22	0.61 ± 0.22
All heavy atoms	1.34 ± 0.17	0.98 ± 0.20
Ramachandran plot appearance[c,d]		
Most favored regions (%)	66.3	86.0
Additionally allowed regions (%)	26.0	12.9
Generously allowed regions (%)	7.7	0.0
Disallowed regions (%)	0.0	1.1

[a] N-terminal domain: 54 ^{3}J(H^N,H$^\alpha$), 131 ^{3}J(H$^\alpha$,H$^\beta$); C-terminal domain: 97 ^{3}J(H^N,H$^\alpha$), 101 ^{3}J(H$^\alpha$,H$^\beta$), 142 ^{3}J(N,H$^\beta$).

[b] To the mean structure.

[c] N-terminal domain, residues 35–152; C-terminal domain, residues 160–250.

[d] From PROCHECK-NMR [72].
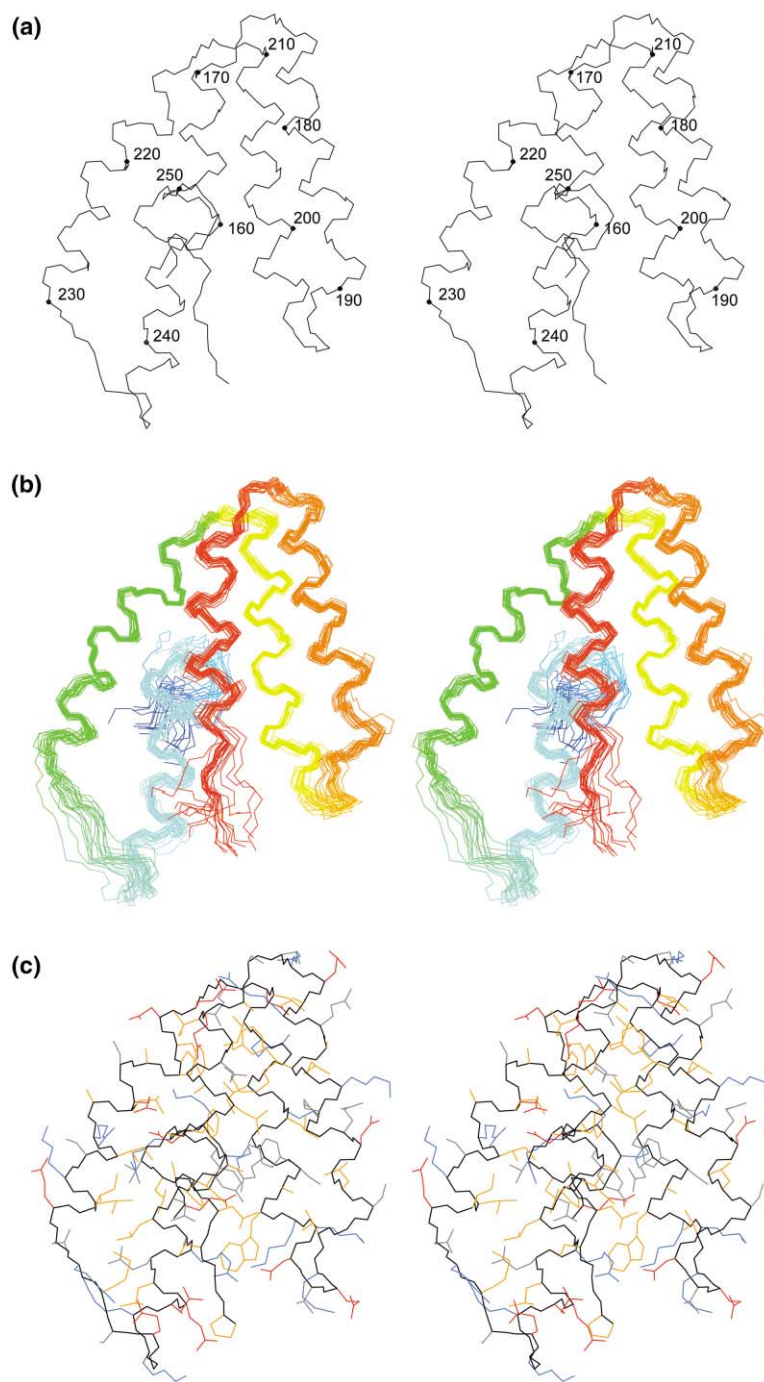
(a)

(b)

(c)

Figure 6. Different Stereorepresentations of Residues 155–256 of the NMR Solution Structure of the C-Terminal Domain of ERp29

All three representations show the structure in the same orientation as in Figure 2. (a) Backbone trace of the conformer with the lowest energy after restrained energy minimization. (b) Ensemble of 20 conformers, using the backbone atoms of residues 160–250 for superposition. (c) Heavy-atom display of the conformer of (a). The same color code was used as in Figure 4.

beled N-terminal domain to ^{15}N-labeled C-terminal domain, only minimal chemical shift changes were observed that did not indicate a unique binding site. Taken together, these data indicate a flexible linker between the two domains and no involvement of the C-terminal domain in the dimer interface.

The affinity of the homodimer was determined by ESI mass spectrometry to be about 1 μM at pH 4.5, based on equal peak intensities of the monomer and the dimer at a total protein concentration of 3.5 μM. This affinity corresponds to less than 5% of monomeric protein at NMR concentrations (0.4–0.7 mM), explaining the con-

servation of chemical shifts between the isolated N-terminal domain and the full-length protein. At the same time, the tight dimer prevented the identification of the dimerization interface by monitoring chemical shift changes as a function of protein concentrations.

An initial indication of the location of the dimer interface came from the observation of slowly exchanging amide protons in the N-terminal domain, where the backbone amides of Asp71, Phe118, Arg122, and Asp123 did not yield exchange cross peaks with the water resonance in a 3D NOESY-^{15}N-HSQC experiment, although they are solvent accessible in the NMR struc-

(a)

(b)

(c)
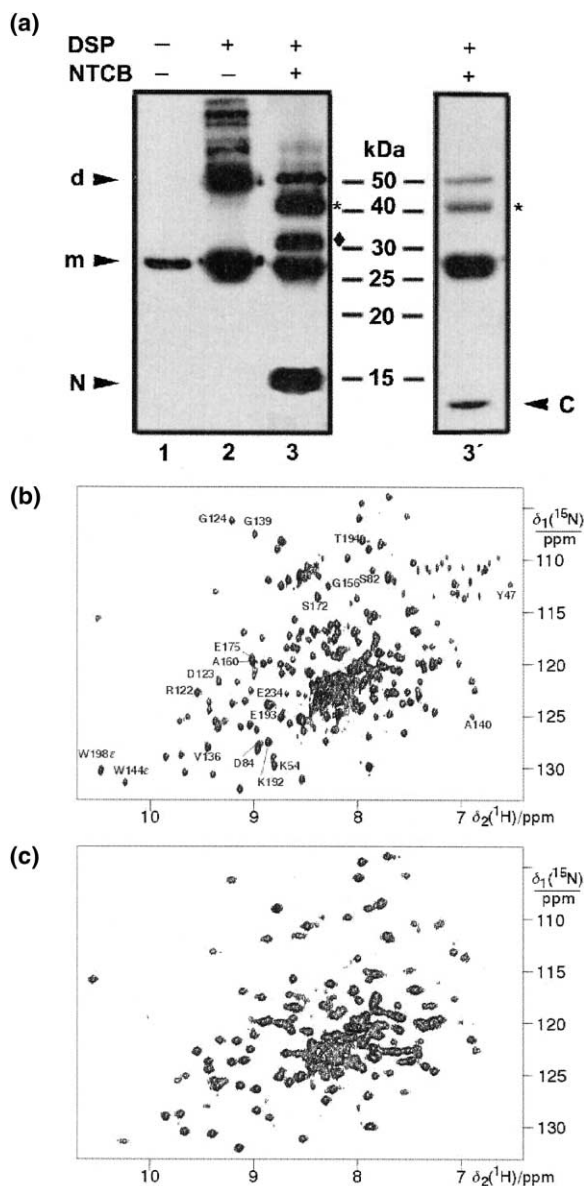
Figure 7. Experiments for the Identification of the Dimerization Domain and Dimer Interface of ERp29

(a) Chemical cross-linking with DSP followed by specific chemical cleavage between the N- and C-terminal domains using NTCB. The fragments were analyzed by reducing SDS-PAGE and immunoblotting with domain-specific antibodies. Samples in the lanes 1–3 were developed with antibodies against the N-terminal domain, while the sample in lane 3′ was developed with an antibody against the C-terminal domain. Arrows labeled "N," "m," "d," and "C" identify the bands of monomeric N-terminal domain, monomeric ERp29, dimeric ERp29, and monomeric C-terminal domain, respectively. A star identifies the position of an ∼42 kDa dimer consisting of the N-terminal domain cross-linked to a full-length ERp29 molecule, while the diamond indicates the band of the ∼31 kDa dimeric N-terminal domain. (b) ^{15}N-TROSY spectrum recorded of a 0.3 mM solution of uniformly ^{15}N/^{13}C/^{2}H-labeled full-length ERp29 in 90% H_2O/10% D_2O (pH 4.9), 31°C, using a ^{1}H NMR frequency of 600 MHz, $t_{1max} = 26$ ms, $t_{2max} = 115$ ms, and a total experimental time of 7 hr. Selected cross peaks are assigned. (c) Same as (b) but with 8 mM Gd(DTPA-BMA) present. Total experimental time 42 hr.

ture (Figure 8c). Except for Asp71, all these residues map onto the same face of the protein which is devoid of charged residues (Figures 4, 9a, and 9b). As expected for a protein-protein interface, no rapidly exchanging amide protons were identified on this surface. However, the data are of limited value, as only a total of 14 cross peaks with the water could be identified for backbone amide protons. A more significant observation may be that the only ^{1}H NMR resonance that could be observed for a tyrosine hydroxyl proton was that of Tyr132, indicating slow exchange with the water, although this proton is solvent exposed in the monomeric structure (Figure 9a).

Independent confirmation of the dimer interface came from a novel experiment where we used the paramagnetic relaxation agent Gd(DTPA-BMA) to probe amide-proton solvent accessibility in full-length ERp29. Compared to TEMPOL, which is most frequently used in studies of solvent exposure, Gd(DTPA-BMA) is effective at about 20-fold lower concentration [31]. Because of the lower concentrations needed, chances for binding of the relaxation agent to the protein backbone are reduced. Consequently, the chemical shifts are more likely to be preserved in the presence of the relaxation agent. In the case of ERp29, the H^N chemical shift changed by less than 0.04 ppm in the presence of 8 mM Gd(DTPA-BMA), which was important for tracking the resonances in crowded spectral regions. At this concentration of relaxation agent, some of the cross peaks disappeared. For example, the $N^\varepsilon H$ cross peak from the side chain of Trp198 disappeared, while that of Trp144 was still visible (Figures 7b and 7c). As both side chain protons are solvent exposed in the structures of the monomeric domains, the protection of Trp144 suggests its participation in the dimerization interface.

To assist with the interpretation of the experimentally observed effects of Gd(DTPA-BMA) on the backbone amides (Figure 8b), the water accessibility of the amide protons and a relaxation enhancement parameter R were calculated, using the NMR structures of the N- and C-terminal domains to predict the relaxation enhancement. Figure 8c plots the water accessibility of the amide protons, and Figure 8d presents the predicted relaxation rate enhancement of the amide protons for a uniform concentration of Gd(DTPA-BMA) in the solvent. Comparison of Figures 8b and 8c shows that water-exposed amide protons are generally relaxed more efficiently by the relaxation agent than amide protons in regular secondary structure elements. For example, all helices of the C-terminal domain are clearly separated by loop regions that are sensitive to the relaxation agent. The same holds, in principle, for the N-terminal domain. Although the amide protons in the loop between helix 2 and strand 2 are not very water accessible, an increased relaxation rate is expected based on Figure 8d. The polypeptide segment between helices 3 and 4 is the only region, where Gd(DTPA-BMA) does not completely suppress any of the amide-proton cross peaks. The amide protons between helix 3 and strand 3 are not very solvent exposed, but Figure 8d would predict a similar relaxivity as for the loop between helix 2 and strand 2, based on a monomeric model. These data thus indicate protection of residues 109–137 in the dimer.
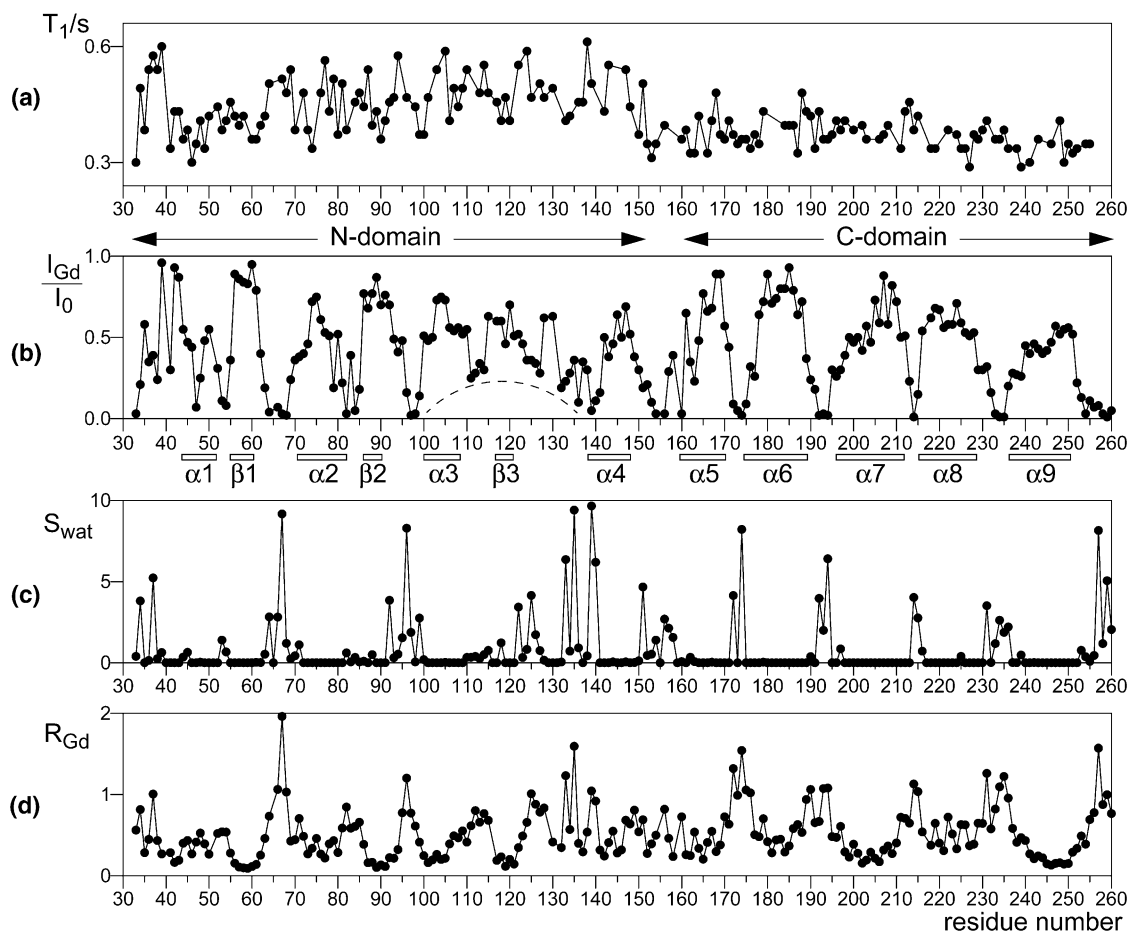
While the correlation between experimental and pre-

Figure 8. Overview of NMR Data Demonstrating Increased Mobility of the C-Terminal Domain versus the N-Terminal Domain and Identifying the Homodimerization Interface of ERp29

(a) T_1 relaxation times of the backbone amide protons. The increased scatter of the data for the N-terminal domain results from the broader line widths observed for these residues. (b) Ratios of cross-peak intensities, measured as peak heights in the ^{15}N-TROSY spectra of Figure 7, showing the signal attenuation by Gd(DTPA-BMA). The location of regular secondary structure elements is shown underneath. The dashed line highlights a region of the N-domain, where amide cross peaks were less affected by Gd(DTPA-BMA) than expected for a monomer. (c) Solvent-accessible surface area of the amide protons, calculated with a circular probe of 1.4 Å radius and averaged over all NMR conformers. The vertical scale is in arbitrary units. (d) Relaxation enhancement of amide protons by Gd(DTPA-BMA) predicted by a grid search algorithm, where the Gd complex was represented by a sphere of 3.5 Å radius with equal residence probability on each grid point in the solution. The relaxation contribution of each accessible grid point was weighed by d^{-6}, where d is the distance between the grid point and the amide proton, and the effects from all grid points were averaged. The figure shows the average of the predictions calculated for all NMR conformers in arbitrary units.

dicted relaxation enhancement provides good qualitative criteria for the identification of interaction sites, the correlation is not quantitative. Notably, the experimental intensity ratios (Figure 8b) vary more strongly between exposed and protected amides than predicted for a simply distance-dependent effect (Figure 8d). Therefore, water accessibility of the amide protons (Figure 8c) is a helpful additional parameter for the interpretation of relaxation enhancement data, as the signals from water-exposed amide protons seem to be reliably attenuated by the relaxation agent. An exception is Gly92 H^N, which is water accessible (Figure 8c) but not easily approached by Gd(DTPA-BMA) (Figure 8d). Another exception is the linker residues between the N- and C-terminal domains, where discrepancies arise from the fact that the solvent exposure was calculated for the individual domains rather than for a model of full-length ERp29. Similarly,

the discrepancy between predicted and observed relaxation enhancement for the amides of Ser189 and Lys231 may reflect proximity of the N-terminal domain in the full-length protein.

The proton exchange data and amide protons for which little relaxation enhancement by Gd(DTPA-BMA) was observed despite water accessibility predicted from monomeric structures (Figure 9a) identify the same protein surface as the dimerization interface. Unfortunately, the dimer structure could not be refined by residual dipolar couplings, as ERp29 was precipitated by lamellar phases [32] and associated tightly with Pf1 phages [33]. Similarly, saturation transfer experiments in mixtures of unlabeled and ^{15}N/^{2}H-labeled ERp29 [34] were unsuccessful due to insufficient sensitivity.

The overall charge distribution of the N-terminal domain and shape complementarity suggests a dimer
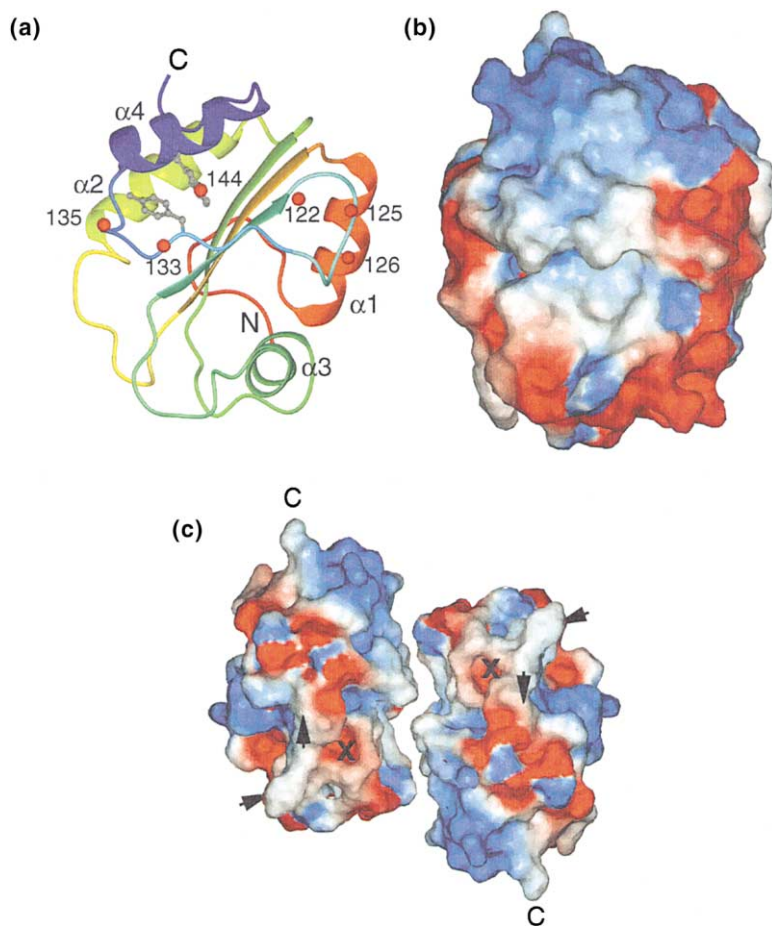
Figure 9. Dimer and Multimer Formation of ERp29

(a) Ribbon drawing of the N-terminal domain of ERp29. Red spheres indicate the locations of NHs that are protected from access to Gd(DTPA-BMA), although significant solvent accessibility would be predicted from the NMR structure (Figure 8). In addition, the side chains of Tyr132 and Trp144 are shown. The view is centered on the putative dimer interface. (b) Surface display of the N-terminal domain in the same orientation as in (a). Red and blue colors indicate regions of negative and positive electrostatic potential, respectively. (c) Model of the dimer between two N-terminal domains, using a representation as in (b), scaled down in size. The molecule on the left is in the same orientation as in (b), except for a rotation by 90° around a vertical axis. The resulting view is along the two-fold symmetry axis of the complex. Arrows identify the approximate locations of the amide groups of residues 67, 68, 97, and 98, for which the cross peaks were broadened in the presence of ERp29 spin labeled at Cys157 in the linker between the N- and C-terminal domains. Crosses mark the locations corresponding to peptide and glutathione binding sites in thioredoxin and glutaredoxin, which, in these proteins, are located between the Cys-X-X-Cys motif and the *cis*-proline corresponding to Pro116 in ERp29 [28].

structure as shown in Figure 9c. Additional support for placement of the C-terminal domains at opposite ends comes from results using spin-labeled ERp29 (see below).

Multimerization of ERp29
Dynamic light scattering (data not shown) as well as cross-linking experiments (Figure 7a) [8] indicated multimerization of a fraction of ERp29 into high-molecular weight complexes. For further characterization of the self-association of ERp29, ^{15}N/^{13}C/^{2}H-labeled ERp29 was mixed with an excess of unlabeled ERp29 that had been spin-labeled at Cys157 by the stable nitroxyl radical 4-maleimido-TEMPO. While the amide cross-peak intensities varied less than 1.5-fold between the ^{15}N-HSQC spectra with and without the spin-labeled ERp29, the cross peaks of residues 67, 68, 97, and 98 were attenuated more than 2.5-fold in the presence of spin-labeled ERp29. Figure 9c shows that these residues are located near the ends of a groove that does not participate in the dimer interface, suggesting that this groove is involved in multimerization. Furthermore, residues 67 and 68 are located more than 20 Å apart from residues 97 and 98, indicating the presence of more than one binding mode, as expected for nonspecific self-association. The absence of any other significant relaxation enhancement by spin-labeled ERp29 further supports the dimer model of Figure 9c, where a spin

label at residue 157 would be positioned too far from the N- and C-terminal domains of the other molecule in the dimer to exert any significant relaxation effect.

Structure-Function Comparison between ERp29 and PDI
In Figure 1a, the amino acid sequences of the N-terminal domain of ERp29 and *a* and *b* domains of human PDI were aligned using the corresponding secondary structure elements observed in the respective three-dimensional structures. This alignment results in about 21% sequence identity between the N-terminal domain of ERp29 and the *a* domain of PDI. This similarity extends to their 3D structures; a combinatorial extension search of the PDB [35] yielded the *a* domain of human PDI as the protein that is structurally most similar to the N-terminal domain of ERp29, with a rmsd of 2.8 Å for 105 aligned residues.

The C-terminal domain of ERp29 revealed a novel five-helix fold that is apparently conserved also in the group of P5-like PDIs and the Windbeutel protein from *Drosophila*. For the P5-like PDIs, which lack the C-terminal ER-retrieval signal KDEL, the importance of the C-terminal domain for retention in the ER was demonstrated by deletion mutations disrupting its structural integrity [36]. The C-terminal domain of ERp29 could play a similar role, although its C-terminal tetrapeptide KEEL may be sufficient for effective retention of the protein in the ER.

The amino acid sequence homology of the N-terminal domain of ERp29 is even greater to the thioredoxin-like domains of the P5-like PDIs than to the *a* domain of human PDI (23%–25% identity; Figure 1). While the functions of ERp29 and P5-like PDIs very likely overlap, only limited biochemical data are available for these PDIs.

The most studied protein disulfide isomerase is human PDI. It is a multifunctional enzyme that catalyzes the formation of correct disulfide bonds in oxidation/reduction equilibria [3]. In addition, PDI has been found as a subunit of larger protein complexes in prolyl-4-hydroxylase [37] and triglyceride transfer protein [38], where its role seems to be purely structural, similar to bacterial thioredoxin in the T7 DNA polymerase complex [39]. Finally, human PDI has been shown to possess a general chaperone activity that is independent of its redox activity [14, 15]. Like all PDIs, the redox-active domains of human PDI contain a Cys-X-X-Cys motif near the amino terminus of helix 2. This motif is absent in ERp29. Correspondingly, ERp29 did not display any thiol/disulfide oxido-reductase activity in an assay for the oxidative refolding of reduced RNase (S.M., unpublished data). It was also inactive in a CDNB (1-chloro-2,4-dinitrobenzene) assay, which probes for glutathione-S-transferase (GST) activity (S.M., unpublished data). To date, neither ERp29 nor its human homolog ERp28 have been found as integral structural components of stable multiprotein complexes. Therefore, the chaperone role of PDI remains as the only plausible function that ERp29 may share with PDI.

This assumption is supported by the chaperone role reported for the ERp29 homolog Windbeutel from *Drosophila* [12]. While the expression of Windbeutel is spatially and temporally restricted, the mammalian protein ERp29 is expressed more widely and is capable of interacting with a broader range of proteins [7, 9], suggesting that broad substrate specificity is an essential component of the putative chaperone function of ERp29.

Little is known about the functional mechanism of the chaperone activity of human PDI or the polypeptide binding sites involved. The protein comprises two redox-active (*a* and *a'*) and two redox-inactive (*b* and *b'*) thioredoxin-like domains in the sequence *a-b-b'-a'* [21]. Under physiological conditions, human PDI forms homodimers [40]. When expressed individually, all domains seem to be monomers, except for the *b'* domain, which aggregates in aqueous solutions [23]. The major peptide binding site seems to be exclusively located in the *b'* domain, although presence of all domains is required for full activity [41, 42]. In particular, mutations in the C-terminal part of the *a'* domain have been shown to affect peptide binding, suggesting an intramolecular interaction between the *a'* and *b'* domains [43, 44].

Several of the characteristic features reported for the functionally important *b'* domain of PDI seem to be shared by the thioredoxin-like domain of ERp29: neither domain contains a Cys-X-X-Cys motif, while both domains harbor a site for nonspecific peptide binding, which, in the case of the N-terminal domain of ERp29, causes oligomerization by interaction with the linker region between the N- and C-terminal domains. Both domains are especially prone to self-aggregation in the absence of the other domains of the wild-type protein.

Unfortunately, the sequence alignment of Figure 1 is too uncertain for the *a'* and *b'* domains to model their structures and homo- or heterodimers with any reliability. If human or P5-like PDIs formed a homodimer like in the dimer model of ERp29, the Cys-X-X-Cys motif would be near the interface yet accessible to substrates (Figure 9c).

Homodimerization and Chaperone Function

ERp29 is the first example of a protein where a thioredoxin-like domain was found to act as a homodimerization module without supporting interactions by covalent bonds or additional contacts by other domains. The dimerization interface identified in ERp29 has so far not been observed in the thioredoxin superfamily. For example, human thioredoxin requires an intermolecular disulfide bridge for dimer formation [45, 46], and the intermolecular contacts observed in the single crystal of calsequestrin involve more than two thioredoxin-like domains [27]. The intramolecular contacts observed between the two covalently linked thioredoxin-like domains of *Pyrococcus furiosus* PDI are also different from the interface identified for ERp29 [16]. GST [47] and DsbC [18] also form dimers but show no extensive contacts between the thioredoxin-like domains. All these examples demonstrate that the thioredoxin fold presents a module highly suitable for protein-protein interactions while avoiding noncovalent homodimerization.

Homodimerization and multimer formation has been observed for many chaperones (BiP [48], GRP94 [49], HSP70 [50], small heat shock proteins [51], DsbC [17]). The role of aggregation for chaperone function is still uncertain, although oligomerization obviously increases the surface area available for the binding of extended polypeptide chains and may generate new binding clefts at the dimer interface. In the case of bacterial HSP70, DnaK, an all-helical C-terminal domain is connected to the substrate binding domain via a flexible segment that allows coverage of the peptide binding site by the helical domain [52]. The linker between the two domains of ERp29 is different in that it is flexible but too short for any intramolecular contacts between the two domains that are not in the immediate vicinity of the linker (Figure 2).

In the search for a potential peptide binding site, we noticed that the C-terminal domain of ERp29 partially exposes a number of hydrophobic residues to the solvent (Leu158–Ala160, Val191, Ile228) that form a surface area of uncharged residues (presenting toward the reader in Figure 6c). This feature, however, is not conserved between related C-terminal domains (Figure 1b), and there is no evidence that it supports ligand binding in ERp29. In particular, we observed no chemical shift changes after addition of model peptides (mastoparan and somatostatin) to a sample of the C-terminal domain (data not shown). A much more likely ligand binding site is presented by the multimerization sites in the N-terminal domain, which were detected by the TEMPO-labeling experiment. Binding to these sites was nonspecific and involved residues from the flexible linker segment, suggesting that other unfolded peptide chains might associate to the same sites. Interestingly, some

chaperones have been found to self-associate via their peptide binding sites and to dissociate upon binding to unfolded substrates [53, 54]. A competition between homoassociation between dimers of ERp29 and hetero-association between ERp29 dimers and unfolded peptide chains would bear strong similarities to the association behavior of other chaperones.

While chaperone function requires the capability of binding a range of different polypeptides with little selectivity, the protein binding capacity must not result in irreversible self-aggregation. In the case of ERp29, the isolated N-terminal domain is much more prone to aggregation and irreversible precipitation than the full-length protein. Possibly, solubilization of the N-terminal domain is one of the most important functions of the all-helical, highly soluble C-terminal domain. In addition, the flexibility of the hinge between the N- and C-terminal domains may have functional importance for polypeptide binding and release.

Biological Implications

ERp29 is the first eukaryotic PDI-related protein for which the structures of all domains have been determined. The data underline the structural similarity between ERp29 and the class of P5-like PDIs. The N-terminal domain of ERp29 is also similar to the thioredoxin-like *a* domain of human PDI. The structural homologies to PDIs suggest overlapping functions between these and ERp29. ERp29 lacks, however, the redox-active Cys-X-X-Cys motif of conventional PDIs, and there is no evidence that ERp29, in analogy to human PDI, acts as a structural subunit of larger multiprotein complexes. Consequently, only the redox-independent chaperone function of human PDI may be shared by ERp29, which is further corroborated by the substrate-specific chaperone properties reported for the *Drosophila* homolog of ERp29, Windbeutel. The function of ERp29 is probably broader, as suggested by its ubiquitous expression in parallel to molecular chaperones, high concentration in many secretory cells, upregulation under conditions when misfolded protein chains accumulate in the ER, and association with a number of transport-incompetent secretory proteins.

The present study revealed a noncovalent homodimer for the N-terminal domain with additional binding sites involved in multimerization of the full-length protein. These multimerization sites associate with the flexible linker region between the two domains of ERp29 in a nonspecific manner and might thus serve as a binding locus also for unfolded polypeptides. Self-association via their peptide binding sites and dissociation upon binding to unfolded substrates has been reported previously for other chaperones [53, 54].

The thioredoxin fold has not previously been observed as a homodimerization module without covalent linkages or additional contacts by other domains. Specific homodimerization and nonspecific multimerization may also account for the self-association reported for the *b'* domain of human PDI, which has been shown to be the most important domain for peptide binding.

Experimental Procedures

Gene Expression and Protein Purification

His-tagged ERp29 was expressed in *E. coli* as described earlier [7]. Subcloning of the coding region of ERp29 cDNA (excluding the leader sequence) into the pQE30 expression vector (Qiagen) resulted in the N-terminal addition of a dodeca-peptide including six histidines. Recombinant ERp29 was expressed in the *E. coli* strain *JM109* and purified under nondenaturing conditions on a Ni-NTA-agarose affinity column according to the manufacturer's recommendations (Qiagen), followed by extensive dialysis against 10 mM potassium phosphate buffer at pH 7.4.

For expression of the His-tagged C-terminal domain of ERp29, the corresponding gene segment was excised as a SmHI-HindIII fragment from the recombinant pQE30 vector and inserted into the pQE32 expression vector. Uniform isotope labeling was achieved by expression in *JM109* in ^{15}N/^{13}C double-labeled Celtone medium (Martek) and purification as described above. ^{15}N/^{13}C-labeled N-terminal domain was obtained by expression of full-length, His-tagged ERp29 in double-labeled Celtone medium, followed by purification of the protein and selective chemical cleavage N-terminal of Cys157. For chemical cleavage at this unique Cys residue, the protein was incubated for 2 hr at 37°C with a 100-fold molar excess of 2-nitro-5-thiocyanobenzoic acid (NTCB) and subsequently incubated overnight under mild alkaline conditions (pH 9.5 and 37°C) [55]. The His-tagged N-terminal domain was separated from the untagged C-terminal fragment using the Ni-NTA-agarose affinity column. Remaining uncleaved full-length protein was removed by adsorption chromatography on a hydroxylapatite column, which retained the full-length protein but not the N-terminal domain. ^{15}N/^{13}C/^{2}H-labeled ERp29 was expressed in the *E. coli* strain *M15* using M9 minimal medium containing ^{15}NH$_4$Cl, ^{13}C-glucose, and D$_2$O.

Chemical Cross-Linking and Protein Cleavage

A 10 μM solution of ERp29 was incubated in 50 mM HEPES (pH 7.4), with 200 μg/ml of the homobifunctional cleavable cross-linker DSP (Pierce) at room temperature for 30 min. The reaction was stopped by 50 mM Tris-HCl (pH 7.4), and the samples were incubated on ice for 10 min. Subsequently, the protein was cleaved by NTCB as described above. The samples were analyzed by reducing SDS-PAGE and immunoblotting. Rabbit polyclonal anti-ERp29 raised against purified N-domain and anti-C-terminal peptide antibodies [7] were used for detection of the N- and C-terminal domains, respectively.

Spin Labeling

A 1 mM solution of ERp29 was incubated with 1 mM DTT for 1 hr at 37°C. DTT was removed thereafter by centrifugation of the reaction mixture through a 4 ml G-25 Sepharose spin column (Pharmacia) for 5 min at 200 × g. Subsequently, 4-maleimido-TEMPO (Sigma) in DMSO was added at a final concentration of 100 mM and incubated for 2 hr at ambient temperature [56]. Excess spin label was removed as above and the sample further dialyzed overnight against 10 mM potassium phosphate buffer (pH 7.4).

NMR Spectroscopy

All NMR spectra of the N and C-terminal domains of ERp29 and of the full-length protein were recorded at 31°C and pH 4.9, using either 95% H$_2$O/5% D$_2$O or 100% D$_2$O as a solvent. No NMR spectra of the N-terminal domain were recorded in 100% D$_2$O solution, because of the strong tendency of the protein to precipitate during ultrafiltration or lyophilization. In contrast, a sample of the C-terminal domain in D$_2$O was readily prepared by lyophilization and redissolving in 100% D$_2$O. Protein concentrations were about 0.5 and 2 mM for the N- and C-terminal domain, respectively, and 0.3 mM for full-length ERp29.

All heteronuclear NMR spectra were recorded on Bruker DRX 500 and DMX 600 NMR spectrometers. The homonuclear NOESY spectrum of unlabeled N-terminal domain used for collection of NOE restraints (40 ms mixing time) was accumulated at 800 MHz on a Varian Unity NMR spectrometer for 1 week. Backbone-resonance assignments were achieved by three-dimensional HNCO, HNCA, and HN(CO)CA spectra [57], which were recorded with TROSY schemes [58, 59] in the case of full-length ERp29. Further ^{1}H reso-

nance assignments were obtained from 2QF-COSY, clean TOCSY (60 ms mixing time), and NOESY (100 ms and 40 ms mixing time) spectra recorded with unlabeled samples of the domains. Ambiguities in ^{1}H resonance assignments were further resolved by ^{15}N-HSQC and 3D NOESY-^{15}N-HSQC spectra [60]. Stereospecific resonance assignments of the β protons in the C-terminal domain were assisted by a 3D HNHB spectrum [61].

NOE distance restraints were collected from two-dimensional NOESY spectra recorded with 40 ms mixing time in H_2O, using $t_{1max} = 63$ ms and $t_{2max} = 128$ ms, where the water signal was suppressed by a spinlock purge pulse before acquisition [62]. Zero-quantum coherences were suppressed by a published procedure [63], and the spectra were corrected for the nonuniform excitation profile in the F_2 dimension using PROSA [64].

$^3J_{HN,H\alpha}$ couplings of the N-terminal domain were measured by a CT-HMQC-HN experiment [65]. $^3J_{H\alpha,H\beta}$ couplings were estimated as 11.0 and 4.0 Hz, respectively, when COSY, TOCSY, and NOESY cross peaks indicated the presence of large and small couplings, respectively, together with staggered conformations around the C^α-C^β bond. For the C-terminal domain, $^3J_{HN,H\alpha}$ and $^3J_{H\alpha,H\beta}$ couplings of valine, threonine, and isoleucine residues were measured by line fitting from the in-phase cross peaks observed in ^{15}N-HSQC and NOESY spectra, respectively [66]. Error ranges of ±2 and ±3 Hz were assumed for all $^3J_{HN,H\alpha}$ and $^3J_{H\alpha,H\beta}$ couplings, respectively.

$T_1(^{15}N)$ relaxation measurements of full-length ERp29 were recorded with five ^{15}N-HSQC-type spectra [67], using relaxation delays of 1, 100, 200, 400, and 1200 ms. SG-^{15}N-TROSY spectra [68] were recorded to measure the relaxation enhancements of the amide resonances by the paramagnetic compound gadolinium-diethylenetriamine penataacetic acid-bismethylamide [Gd(DTPA-BMA), Nycomed] in full-length ERp29.

^{15}N-HSQC spectra were recorded to monitor the effect of TEMPO-labeled ERp29 at natural isotopic abundance on ^{15}N/^{13}C/^{2}H-labeled ERp29.

Structure Calculations

The cross peaks in the NOESY spectra were assigned and integrated using the program XEASY [69]. The NMR structure was calculated using the program DYANA (version 1.5) and the associated routines CALIBA, HABAS, and GLOMSA [70]. Fifty random conformers were annealed in 30,000 steps using torsion-angle dynamics. The twenty conformers with the lowest residual restraint violations were energy minimized using the program OPAL (version 2.6) with standard parameters [71].

Table 1 shows an overview of the restraints used and structural statistics. The Ramachandran plot was analyzed using PROCHECK-NMR (version 3.4) [72]. For the N-terminal domain, no ϕ, ψ pair was found in forbidden regions for all 20 final conformers. Thr194 of the C-terminal domain was the only residue consistently found in a forbidden region. It is located in a highly solvent-exposed, irregular segment with relatively low NOE density.

Secondary structure elements and rmsd values were calculated using the program MOLMOL (version 2.6) [73]. MOLMOL was also used to generate all molecular graphics figures.

References

1. Gething, M.J., and Sambrook, J. (1992). Protein folding in the cell. Nature *355*, 33–45.

2. Stevens, F.J., and Argon, Y. (1999). Protein folding in the ER. Semin. Cell. Dev. Biol. *10*, 443–454.

3. Freedman, R.B., Hirst, T.R., and Tuite, M.F. (1994). Protein disulphide isomerase: building bridges in protein folding. TIBS *19*, 331–336.

4. Ferrari, D.M., and Söling, H.D. (1999). The protein disulphide-isomerase family: unravelling a string of folds. Biochem. J. *339*, 1–10.

5. Fang, C., Mkrtchian, S., and Ingelman-Sundberg, M. (1997). Combination of direct DNA sequencing with degenerate primer-mediated PCR and 5′-/3′-RACE to screen novel cDNA sequences. Biotechniques *23*, 52–58.

6. Demmer, J., Zhou, C., and Hubbard, M.J. (1997). Molecular cloning of ERp29, a novel and widely expressed resident of the endoplasmic reticulum. FEBS Lett. *402*, 145–150.

7. Mkrtchian, S., Fang, C., Hellman, U., and Ingelman-Sundberg, M. (1998). A stress-inducible rat liver endoplasmic reticulum protein, ERp29. Eur. J. Biochem. *251*, 304–313.

8. Mkrtchian, S., et al., and Ingelman-Sundberg, M. (1998). Oligomerization properties of ERp29, an endoplasmic reticulum stress protein. FEBS Lett. *431*, 322–326.

9. Ferrari, D.M., Nguyen Van, P., Kratzin, H.D., and Söling, H.D. (1998). ERp28, a human endoplasmic-reticulum-lumenal protein, is a member of the protein disulfide isomerase family but lacks a CXXC thioredoxin-box motif. Eur. J. Biochem. *255*, 570–579.

10. Hubbard, M.J., McHugh, N.J., and Carne, D.L. (2000). Isolation of ERp29, a novel endoplasmic reticulum protein, from rat enamel cells. Evidence for a unique role in secretory-protein synthesis. Eur. J. Biochem. *267*, 1945–1957.

11. Konsolaki, M., and Schupbach, T. (1998). *windbeutel*, a gene required for dorsoventral patterning in Drosophila, encodes a protein that has homologies to vertebrate proteins of the endoplasmic reticulum. Genes Dev. *12*, 120–131.

12. Sen, J., Goltz, J.S., Konsolaki, M., Schupbach, T., and Stein, D. (2000). Windbeutel is required for function and correct subcellular localization of the Drosophila patterning protein Pipe. Development *127*, 5541–5550.

13. Kwon, O.Y., Park, S., Lee, W., You, K.H., Kim, H., and Shong, M. (2000). TSH regulates a gene expression encoding ERp29, an endoplasmic reticulum stress protein, in the thyrocytes of FRTL-5 cells. FEBS Lett. *475*, 27–30.

14. Wang, C.C. (1998). Protein disulfide isomerase assists protein folding as both an isomerase and a chaperone. Ann. NY Acad. Sci. *864*, 9–13.

15. Puig, A., and Gilbert, H.F. (1994). Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme. J. Biol. Chem. *269*, 7764–7771.

16. Ren, B., Tibbelin, G., de Pascale, D., Rossi, M., Bartolucci, S., and Ladenstein, R. (1998). A protein disulfide oxidoreductase from the archaeon Pyrococcus furiosus contains two thioredoxin fold units. Nat. Struct. Biol. *5*, 602–611.

17. Martin, J.L., Bardwell, J.C., and Kuriyan, J. (1993). Crystal structure of the DsbA protein required for disulphide bond formation in vivo. Nature *365*, 464–468.

18. McCarthy, A.A., Haebel, P.W., Torronen, A., Rybin, V., Baker, E.N., and Metcalf, P. (2000). Crystal structure of the protein disulfide bond isomerase DsbC from Escherichia coli. Nat. Struct. Biol. *7*, 196–199.

19. Chen, J., Song, J.L., Zhang, S., Wang, Y., Cui, D.F., and Wang, C.C. (1999). Chaperone activity of DsbC. J. Biol. Chem. *274*, 19601–19605.

20. Kemmink, J., Darby, N.J., Dijkstra, K., Nilges, M., and Creighton, T.E. (1996). Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear ^{13}C/^{15}N NMR spectroscopy. Biochemistry *35*, 7684–7691.

21. Kemmink, J., Darby, N.J., Dijkstra, K., Nilges, M., and Creighton, T.E. (1997). The folding catalyst protein disulfide isomerase is constructed of active and inactive thioredoxin modules. Curr. Biol. *7*, 239–245.

22. Kemmink, J., et al., and Darby, N.J. (1999). The structure in solution of the b domain of protein disulfide isomerase. J. Biomol. NMR *13*, 357–368.

23. Darby, N.J., van Straaten, M., Penka, E., Vincentelli, R., and Kemmink, J. (1999). Identifying and characterizing a second structural domain of protein disulfide isomerase. FEBS Lett. *448*, 167–172.

24. Dijkstra, K., et al., and Kemmink, J. (1999). Assignment of ¹H, ¹³C and ¹⁵N resonances of the a' domain of protein disulfide isomerase. J. Biomol. NMR *14*, 195–196.

25. Epp, O., Ladenstein, R., and Wendel, A. (1983). The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. Eur. J. Biochem. *133*, 51–69.

26. Reuter, K., Nottrott, S., Fabrizio, P., Luhrmann, R., and Ficner, R. (1999). Identification, characterization and crystal structure analysis of the human spliceosomal U5 snRNP-specific 15 kD protein. J. Mol. Biol. *294*, 515–525.

27. Wang, S., Trumble, W.R., Liao, H., Wesson, C.R., Dunker, A.K., and Kang, C.H. (1998). Crystal structure of calsequestrin from rabbit skeletal muscle sarcoplasmic reticulum. Nat. Struct. Biol. *5*, 476–483.

28. Nordstrand, K., Åslund, F., Holmgren, A., Otting, G., and Berndt, K.D. (1999). NMR structure of Escherichia coli glutaredoxin 3-glutathione mixed disulfide complex: implications for the enzymatic mechanism. J. Mol. Biol. *286*, 541–552.

29. Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

30. van Asselt, E.J., Thunnissen, A.M., and Dijkstra, B.W. (1999). High resolution crystal structures of the Escherichia coli lytic transglycosylase Slt70 and its complex with a peptidoglycan fragment. J. Mol. Biol. *291*, 877–898.

31. Petros, A.M., Mueller, L., and Kopple, K.D. (1990). NMR identification of protein surfaces using paramagnetic probes. Biochemistry *29*, 10041–10048.

32. Rückert, M., and Otting, G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J. Am. Chem. Soc *122*, 7793–7797.

33. Hansen, M.R., Hanson, P., and Pardi, A. (2000). Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions. Methods Enzym. *317*, 220–240.

34. Takahashi, H., Nakanishi, T., Kami, K., Arata, Y., and Shimada, I. (2000). A novel NMR method for determining the interfaces of large protein-protein complexes. Nat. Struct. Biol. *7*, 220–221.

35. Shindyalov, I.N., and Bourne, P.E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng. *11*, 739–747.

36. Monnat, J., Neuhaus, E.M., Pop, M.S., Ferrari, D.M., Kramer, B., and Soldati, T. (2000). Identification of a novel saturable endoplasmic reticulum localization mechanism mediated by the C-terminus of a dictyostelium protein disulfide isomerase. Mol. Biol. Cell. *11*, 3469–3484.

37. Pihlajaniemi, T., et al., and Kivirikko, K.I. (1987). Molecular cloning of the beta-subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. EMBO J. *6*, 643–649.

38. Wetterau, J.R., Combs, K.A., Spinner, S.N., and Joiner, B.J. (1990). Protein disulfide isomerase is a component of the microsomal triglyceride transfer protein complex. J. Biol. Chem. *265*, 9801–9807.

39. Huber, H.E., Tabor, S., and Richardson, C.C. (1987). Escherichia coli thioredoxin stabilizes complexes of bacteriophage T7 DNA polymerase and primed templates. J. Biol. Chem. *262*, 16224–16232.

40. Lambert, N., and Freedman, R.B. (1983). Structural properties of homogeneous protein disulphide-isomerase from bovine liver purified by a rapid high-yielding procedure. Biochem. J. *213*, 225–234.

41. Darby, N.J., Penka, E., and Vincentelli, R. (1998). The multidomain structure of protein disulfide isomerase is essential for high catalytic efficiency. J. Mol. Biol. *276*, 239–247.

42. Klappa, P., Ruddock, L.W., Darby, N.J., and Freedman, R.B. (1998). The b' domain provides the principal peptide-binding site of protein disulfide isomerase but all domains contribute to binding of misfolded proteins. EMBO J. *17*, 927–935.

43. Koivunen, P., et al., and Kivirikko, K.I. (1999). The acidic C-terminal domain of protein disulfide isomerase is not critical for the enzyme subunit function or for the chaperone or disulfide isomerase activities of the polypeptide. EMBO J. *18*, 65–74.

44. Klappa, P., et al., and Freedman, R.B. (2000). Mutations that destabilize the a' domain of human protein-disulfide isomerase indirectly affect peptide binding. J. Biol. Chem. *275*, 13213–13218.

45. Weichsel, A., Gasdaska, J.R., Powis, G., and Montfort, W.R. (1996). Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer. Structure *4*, 735–751.

46. Qin, J., Clore, G.M., and Gronenborn, A.M. (1994). The high-resolution three-dimensional solution structures of the oxidized and reduced states of human thioredoxin. Structure *2*, 503–522.

47. Armstrong, R.N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. Chem. Res. Toxicol. *10*, 2–18.

48. Carlino, A., Toledo, H., Skaleris, D., DeLisio, R., Weissbach, H., and Brot, N. (1992). Interactions of liver Grp78 and Escherichia coli recombinant Grp78 with ATP: multiple species and disaggregation. Proc. Natl. Acad. Sci. USA *89*, 2081–2085.

49. Wearsch, P.A., and Nicchitta, C.V. (1996). Endoplasmic reticulum chaperone GRP94 subunit assembly is regulated through a defined oligomerization domain. Biochemistry *35*, 16760–16769.

50. Schlossman, D.M., Schmid, S.L., Braell, W.A., and Rothman, J.E. (1984). An enzyme that removes clathrin coats: purification of an uncoating ATPase. J. Cell. Biol. *99*, 723–733.

51. Arrigo, A.P., Suhan, J.P., and Welch, W.J. (1988). Dynamic changes in the structure and intracellular locale of the mammalian low-molecular-weight heat shock protein. Mol. Cell. Biol. *8*, 5059–5071.

52. Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., and Ogata, C.M., Gottesman, M.E., and Hendrickson, W.A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. Science *272*, 1606–1614.

53. Benaroudj, N., Triniolles, F., and Ladjimi, M.M. (1996). Effect of nucleotides, peptides, and unfolded proteins on the self-association of the molecular chaperone HSC70. J. Biol. Chem. *271*, 18471–18476.

54. Chevalier, M., King, L., Wang, C., Gething, M.J., Elguindi, E., and Blond, S.Y. (1998). Substrate binding induces depolymerization of the C-terminal peptide binding domain of murine GRP78/BiP. J. Biol. Chem. *273*, 26827–26835.

55. Jacobson, G.R., Schaffer, M.H., Stark, G.R., and Vanaman, T.C. (1973). Specific chemical cleavage in high yield at the amino peptide bonds of cysteine and cystine residues. J. Biol. Chem. *248*, 6583–6591.

56. Vlasova, I.I., and Kuprin, S.P. (1996). Study of yeast phosphoglycerate kinase by electron paramagnetic resonance. 1. Modification of protein SH-group by spin labels. Conformational changes of enzyme. Biofizika *41*, 11193–11200.

57. Grzesiek, S., and Bax, A. (1992). Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. J. Magn. Reson. *96*, 432–440.

58. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (1998). TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc. Natl. Acad. Sci. *95*, 13585–13590.

59. Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wüthrich, K. (1999). TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. J. Am. Chem. Soc. *121*, 844–848.

60. Talluri, S., and Wagner, G. (1996). An optimized 3D NOESY-HSQC. J. Magn. Reson. B *112*, 200–205.

61. Archer, S.J., Ikura, M., Torchia, D.A., and Bax, A. (1991). An alternative 3D NMR technique for correlating backbone ¹⁵N with side chain Hβ resonances in larger proteins. J. Magn. Reson. *95*, 636–641.

62. Otting, G., Liepinsh, E., and Wüthrich, K. (1991). Proton exchange with internal water molecules in proteins in aqueous solution. J. Am. Chem. Soc. *113*, 4363–4364.

63. Otting, G., Orbons, L., and Wüthrich, K. (1990). Suppression of zero-quantum coherence in NOESY and soft-NOESY. J. Magn. Reson. *89*, 423–430.

64. Güntert, P., Dötsch, V., Wider, G., and Wüthrich, K. (1992). Pro-

cessing of multi-dimensional NMR data with the new software PROSA. J. Biomol. NMR *2*, 619–629.

65. Ponstingl, H., and Otting, G. (1998). Rapid measurement of scalar three-bond $^1H^N$-$^1H^{\alpha}$ spin coupling constants in ^{15}N-labeled proteins. J. Biomol. NMR *12*, 319–324.

66. Szyperski, T., Güntert, P., Otting, G., and Wüthrich, K. (1992). Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets. J. Magn. Reson. *99*, 552–560.

67. Andersson, P., Gsell, B., Wipf, B., Senn, H., and Otting, G. (1998). HMQC and HSQC experiments with water flip-back optimized for large proteins. J. Biomol. NMR *11*, 279–288.

68. Weigelt, J. (1998). Single scan, sensitivity and gradient enhanced TROSY for multidimensional NMR experiments. J. Am. Chem. Soc. *120*, 10778–10779.

69. Bartels, C., Xia, T., Güntert, P., Billeter, M., and Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral analysis. J. Biomol. NMR *5*, 1–10.

70. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. J. Mol. Biol. *273*, 283–298.

71. Luginbühl, P., Güntert, P., Billeter, M., and Wüthrich, K. (1996). The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. J. Biomol. NMR *8*, 136–146.

72. Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

73. Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph *14*, 51–55.

Accession Numbers

The coordinates of the 20 energy-refined DYANA conformers of the N- and C-terminal domains of ERp29 were deposited in the Protein Data Bank with the accession codes 1G7E and 1G7D, respectively. The NMR chemical shifts were deposited at the BioMagResBank (BMRB) under the accession codes 4919 and 4920, respectively.