FEBS Letters 569 (2004) 112–116 FEBS 28514

Determination of a high-precision NMR structure of the minicollagen cysteine rich domain from *Hydra* and characterization of its disulfide bond formation

Sebastian Meier[a], Daniel Häussinger[a], Elena Pokidysheva[b], Hans Peter Bächinger[c], Stephan Grzesiek[a,*]

[a]*Division of Structural Biology, Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland*
[b]*Division of Biophysical Chemistry, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland*
[c]*Shriners Hospital for Children, Department of Biochemistry and Molecular Biology, Oregon Health and Science University, Portland, OR 97239, USA*

Received 22 April 2004; accepted 17 May 2004

Available online 7 June 2004

Edited by Felix Wieland

Abstract A high-precision solution structure of the C-terminal minicollagen cysteine rich domain of *Hydra* has been determined using modern heteronuclear and weak alignment NMR techniques at natural isotope abundance. The domain consists of only 24 amino acids, six of which are prolines and six are cysteines bonded in disulfide bridges that constrain the structure into a new fold. The redox equilibrium of the structure has been characterized from a titration with glutathione. No local native structures are detectable in the reduced form. Thus, oxidation and folding are tightly coupled.

Keywords: Nematocyst; Disulfide bond; Heteronuclear NMR; Residual dipolar coupling

1. Introduction

Minicollagens from *Hydra* are the smallest known collagens to date [1]. They are found in the walls of nematocysts, which are explosive organelles formed from a post-Golgi vesicle in *Hydra*, jellyfish, corals and other *Cnidaria*. The explosive discharge of nematocysts is a specialized exocytosis process which achieves accelerations of >40 000 g and is one of the fastest natural processes [2]. This process is enabled by an estimated intramolecular osmotic pressure of 150 bar, which results from an enrichment of poly-γ-glutamate in the capsule matrix during maturation [3].

Nematocyst walls are stabilized by a collageneous matrix [4], which depends on disulfide crosslinking [5] to withstand the extreme osmotic pressure. The *Hydra* capsule wall mainly consists of the proteins minicollagen and NOWA [6–8] (Fig. 1). Both proteins contain highly homologous cysteine- and proline-rich domains, henceforward called minicollagen cysteine rich domains (MCRDs) [9]. Minicollagens are expressed as soluble precursor forms with intramolecular disulfide bonds in their MCRDs [6,8]. In nematocyst development a wall hardening occurs, during which the intramolecular MCRD disulfide bonds are most likely reshuffled to intermolecular bonds, thereby crosslinking molecules within the capsule wall [6]. The simultaneous appearance of MCRDs in minicollagen and NOWA points to a concerted function in wall hardening. A more detailed comparison and discussion of the biological context is given in an accompanying paper [10].

Here, we report the NMR determination of the solution structure of the C-terminal MCRD domain of minicollagen 1. The domain consists of only 24 amino acids, six of which are cysteines and six are prolines. The well-defined structure forms in high yields within 2 h when disulfide bond formation is catalyzed by glutathione. No local native structures are detectable in the reduced form of the MCRD. Thus, oxidation and folding are tightly coupled.

2. Materials and methods

The MCRD Minicol1C (Fig. 1, underlined) was chemically synthesized without isotope enrichment and purified after oxidation as described elsewhere [10]. NMR samples were prepared to 5 mM peptide concentration in 5 mM phosphate, 2 mM NaN_3, 95% H_2O/5% D_2O, and pH 6.5. A 2 mM NMR sample of identical pH and ionic strength was lyophilized and redissolved in D_2O. For the measurement of residual dipolar couplings (RDCs), a 3 mM peptide sample was prepared in 10 mg/ml Pf1 phage solution (Asla Biotech) with 5 mM phosphate of pH 6.5 and 350 mM NaCl to tune down electrostatic interactions of the positively charged MCRD (pI 8.74) with the phages [11], resulting in a maximal coupling $A_{ZZ}^{NH} = 11.5$ Hz.

2.1. NMR spectroscopy

All multidimensional NMR measurements were performed at 15 °C on Bruker DRX 600 and 800 spectrometers. ^{1}H, ^{13}C and ^{15}N assignments were obtained from TOCSY, NOESY as well as ^{15}N and ^{13}C natural abundance HSQC and HMBC spectra. $^{1}D_{HN}$ and $^{1}D_{CH}$ RDCs were obtained by non-decoupled HSQCs. Spectra were processed with NMRPipe [12] and analyzed with PIPP [13].

2.2. Structure determination

In addition to NOEs, the input for the structure calculations contained backbone $^{1}D_{HN}$, backbone and sidechain $^{1}D_{CH}$ RDCs as well as dihedral angle restraints obtained with TALOS [14] (Table 1). All five X-Pro peptide groups were found to be in *trans* conformation by specific, strong sequential $d_{\alpha\delta}$ NOEs. Structures were calculated in CNS [15] with a simulated annealing protocol using RDCs without a priori knowledge of the orientation tensor via the ISAC routine [16]. Structure representations were generated by MOLMOL [17].

* Corresponding author. Fax: +41-61-267-2109.
E-mail address: stephan.grzesiek@unibas.ch (S. Grzesiek).

0014-5793/$22.00
doi:10.1016/j.febslet.2004.05.034

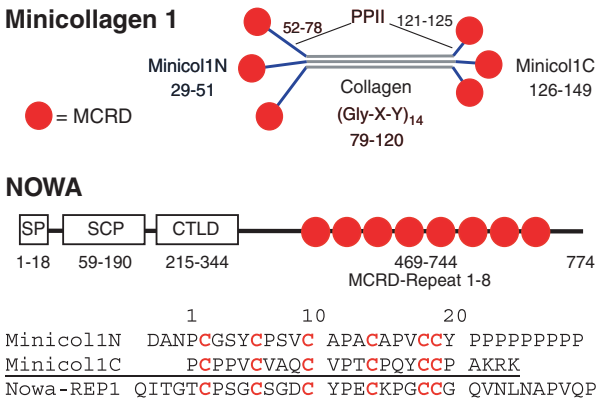

Fig. 1. Top: Domain organization of *Hydra* minicollagen 1 and NOWA. MCRDs are indicated as circles. Bottom: sequence alignment of minicollagen and NOWA MCRDs. For simplicity, residues of the Minicol1C MCRD used in this study are renumbered from 1 to 24.

Table 1
Statistics of the MCRD NMR structure[a]

RMSDs from experimental distance constraints [Å] all (331)[b]	0.044 ± 0.001
RMSDs from dipolar coupling constraints (37)[c]	1.2 Hz
NMR quality factor Q[d]	0.20
RMSDs from dihedral constraints [°] (15)[e]	0.83 ± 0.03
Deviation from the idealized covalent geometry	
Bonds [Å]	0.0053 ± 0.0002
Angles [°]	0.74 ± 0.01
Impropers[f] [°]	0.72 ± 0.01
Coordinate precision [Å][g]	
Backbone non-hydrogen atoms	0.16
All non-hydrogen atoms	0.56
Percentage of non-gly, non-pro residues in Ramachandran regions[h]	
Core	93.3
Allowed	6.7
Generous	0.0
Disallowed	0.0

[a] The statistics were obtained from a subset of the 10 lowest energy structures out of 100 calculated with a CNS [15] simulated annealing protocol using dipolar restraints implemented in the form of the ISAC routine [16]. The number of constraints is given in parentheses. Coordinate precision and Ramachandran plot quality are reported for the core of the structure (residues 2–21) excluding the flexible N- and C-terminus.
[b] NOEs comprise 60 intraresidual NOEs, 132 sequential NOEs ($|i-j|=1$), 114 short range NOEs ($1<|i-j|\leqslant 5$) and 25 long range NOEs ($|i-j|>5$). Hydrogen bond constraints were not imposed.
[c] RDC data comprise 13 ^{1}H–^{15}N, 18 ^{1}H–$^{13}C^{\alpha}$ and 6 sidechain ^{1}H–^{13}C methyl and methine dipolar one-bond couplings. Only RDCs of the non-flexible MCRD core residues are incorporated.
[d] The NMR quality factor Q is defined as the ratio of the rmsd between observed and calculated couplings and the rms of the observed couplings [27].
[e] The dihedral angle constraints comprise 4ϕ and 9ψ obtained for residues 2–21 with ^{1}H, ^{15}N and ^{13}C natural abundance chemical shift assignments as inputs for TALOS [14], as well as $2\chi^{1}$ angles obtained from NOESY peak intensities for two residues with very few long-range NOE data (Cys2 in a PCPP sequence and solvent exposed Gln9).
[f] The improper torsion angle restraints serve to maintain planarity and chirality.
[g] The coordinate precision is defined as the average root mean square distance between the individual simulated annealing structures and the mean coordinates. Values are reported for core residues[a].
[h] Values are calculated with the program PROCHECK-NMR [28] for the residues 2–21[a]. The inverse γ-turn geometry of Cys10 (Fig. 4) is the only one of these residues to fall outside the core region of the Ramachandran plot.

2.3. Reductive unfolding and oxidative folding

MCRD (0.3 mM) was reduced completely with 6 mM tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) at pH 7.5. Oxidative refolding was started by the addition of 40 mM oxidized glutathione (GSSG) in 5 mM phosphate buffer, pH 7.5. Refolding was monitored by 1D ^{1}H NMR experiments and fitted to monoexponentials with ProFit (Quantumsoft).

Equilibrium unfolding of 0.3 mM MCRD in 9 mM GSSG, and 100 mM phosphate (pH 7.0, 25 °C) was induced by the addition of varying amounts of reduced glutathione (GSH) in 100 mM phosphate (pH 7.0) and monitored by 1D ^{1}H NMR after 12 h incubation in airtight Shigemi tubes. Concentrations of GSH and GSSG were validated from the integrals of the respective ^{1}H NMR spectra.

3. Results

3.1. Assignment, structure determination and disulfide pattern of MCRD

Sequence specific ^{1}H, ^{13}C, and ^{15}N assignments were obtained from TOCSY and NOESY as well as heteronuclear HSQC and HMBC spectra acquired at natural isotope abundance (Fig. 2). All cysteine residues are bonded in disulfide bridges as evidenced by their $^{13}C^{\beta}$ shifts of 40–43.5 ppm [18]. The homogeneous spectra show only one single protein species (Fig. 2). All non-labile protons could be assigned for structure calculations. Water exchange results in broadening of amide protons of Cys2 and Gln16 (Fig. 2), which is abolished at lower pH values.

In total, 331 distance constraints were obtained from NOESY spectra in H_2O and D_2O (Table 1). In addition, 37 $^{1}D_{HN}$ and $^{1}D_{CH}$ RDCs for residues 2–21 were measured at natural isotope abundance in weakly oriented samples and incorporated in the structure calculations (Fig. 3, Table 1) to further improve the structural quality. The resulting high resolution structure with a heavy atom backbone RMSD of 0.16 Å as well as pairwise interresidual $d_{H\beta iH\beta k}$ NOEs in all disulfide bonds prove the disulfide pattern Cys2–Cys18, Cys6–Cys14 and Cys10–Cys19 (Fig. 4).

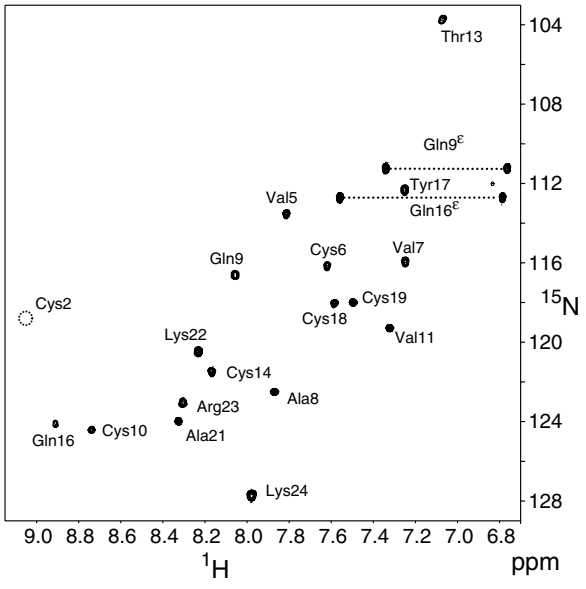

Fig. 2. Natural abundance ^{1}H–^{15}N HSQC spectrum of Minicol1C MCRD (5 mM, experimental time 1.5 h). Cys2 is broadened beyond detection due to solvent exchange but is identified at pH 4 at the indicated position.

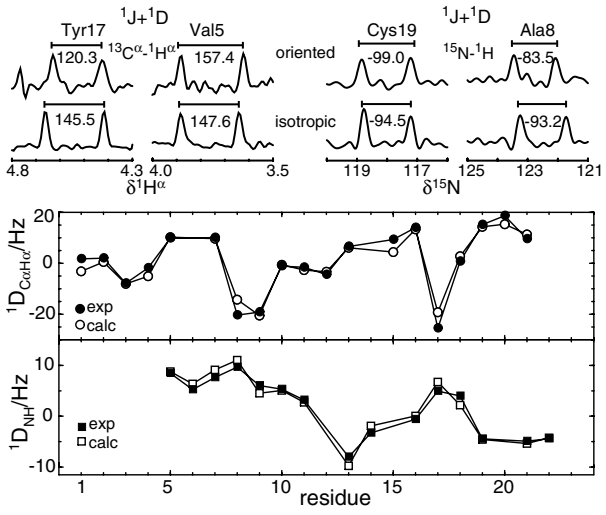

Fig. 3. Top: 1D traces of heteronuclear natural abundance spectra for the measurement of RDCs. Bottom: comparison of measured RDCs and RDCs derived from the calculated average structure of Minicol1C MCRD.

The N-terminus of MCRD (^{1}PCP^{3}) is an elongation of the preceding polyproline II helix in minicollagen 1 and adopts (ϕ/ψ)-values of (−60° to −90°/125–155°) similar to conformations in the ideal PPII helix (−75°/145°). The subsequent short α-helix (residues 4–8) is followed by an inverse γ-turn (9–11), a type I β-turn (11–14) and a type III β-turn (15–18). The highly conserved Pro12 stabilizes the βI turn topology as canonical residue $i+1$. The structure of MCRD (including the helix) can be viewed as a consecutive arrangement of turns, where each turn contains one cysteine that is directed away from the turn to form the aforementioned long-range disulfide pattern.

Backbone $H^N \rightarrow O'$ hydrogen bonds in the turns from residues $6 \rightarrow 3$, 8c4, $9 \rightarrow 5$, $11 \rightarrow 9$, $14 \rightarrow 11$, $18 \rightarrow 15$ are evident from the structure, deviations from random coil $^{1}H^N$ chemical shifts (Fig. 2) and reduced $^{1}H/^{2}H$ exchange. A schematic overview of turns, hydrogen bonds and disulfide bridges is shown in Fig. 4.

A three-dimensional similarity search of the PDB with the program DALI [19] did not yield any related structures. This indicates that the MCRD motif represents a new fold, reflecting its highly specialized function in the disulfide-mediated crosslinking of the nematocyst wall. It should also be noted that MCRD with six cysteines in a folded core of only 20 structured amino acids represents one of the sequences with the highest densities of disulfide bonds known so far [10].

$^{1}H^N$ T_2 relaxation times >120 ms of 2 mM MCRD at 25 °C show that MCRD is a monomer in solution and therefore should not contribute to the non-covalent stabilization of the minicollagen trimer by intermolecular interactions. Accordingly, the reduction of the MCRD disulfide bonds was found to have no influence on the transition temperature of 45 °C for minicollagen 1 trimerization [8].

3.2. Reductive unfolding and oxidative refolding

MCRD is readily reduced by an excess of TCEP or glutathione at pH 7.5. Upon reduction, all hydrogen bonds in the turns are abolished as evident from a change of amide $^{1}H^N$ resonance shifts to random coil values of >7.9 ppm for all residues (Fig. 5A). NOESY spectra give additional evidence that the structure is largely in extended or in random coil conformation, as the strong d_{HNHN} NOEs of the turns are abolished. Moreover, the ^{15}N resonance of Thr13 assumes a random coil value (114.3 ppm), whereas position $i+2$ of a βI turn usually has characteristic ^{15}N upfield shifts [14] (103.7 ppm in the native MCRD structure, Fig. 2). Thus, the disulfide pattern in MCRD obviously does not result from a non-covalently folded precursor, in which cysteines are already properly preoriented. Rather oxidation and folding occur simultaneously, and the disulfide bonds stabilize the natively folded structure.

Oxidative folding of completely reduced MCRD was started by the addition of a >100-fold excess of oxidized glutathione as its natural oxidant. Refolding was easily monitored by 1D NMR spectra of upfield shifted $^{1}H^N$ turn resonances (Fig. 5A) which do not overlap with the glutathione signal. The 1D signals apparently correspond either to the native folded species or a single unfolded state, whereas other intermediates are not highly populated. Folding proceeds to a steady state within about 2 h (Fig. 5A). Fitting of signal intensities to monoexponentials shows a decay rate of 1.13 ± 0.06 h^{-1} for resonances corresponding to the unfolded species, whereas the native signals are formed at a slightly slower rate of 0.78 ± 0.03 h^{-1}

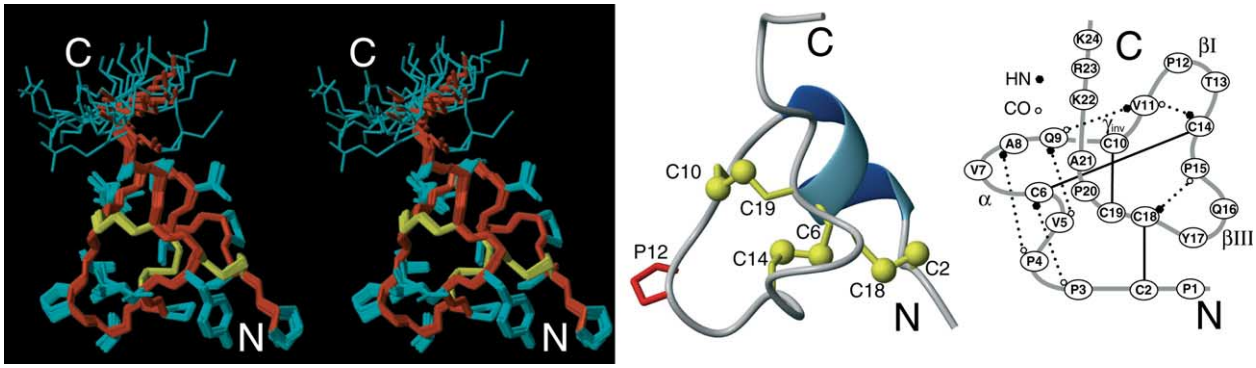

Fig. 4. The structure of Minicol1C MCRD. Left: stereo view in an all-heavy-atom representation of the 10 lowest energy conformers out of 100 structures calculated. The backbone is displayed in dark orange and sidechains in cyan, except for cysteine sidechains in yellow. The charged C-terminal sequence $K^{22}R^{23}K^{24}$ is disordered at pH 6.5 in the absence of additional salt. Center: ribbon representation in the same orientation as stereo view shows that disulfide bonds Cys10–Cys19 and Cys2–Cys18 are solvent accessible, whereas the central disulfide bond Cys6–Cys14 is inaccessible. Right: disulfide and hydrogen bond topology.

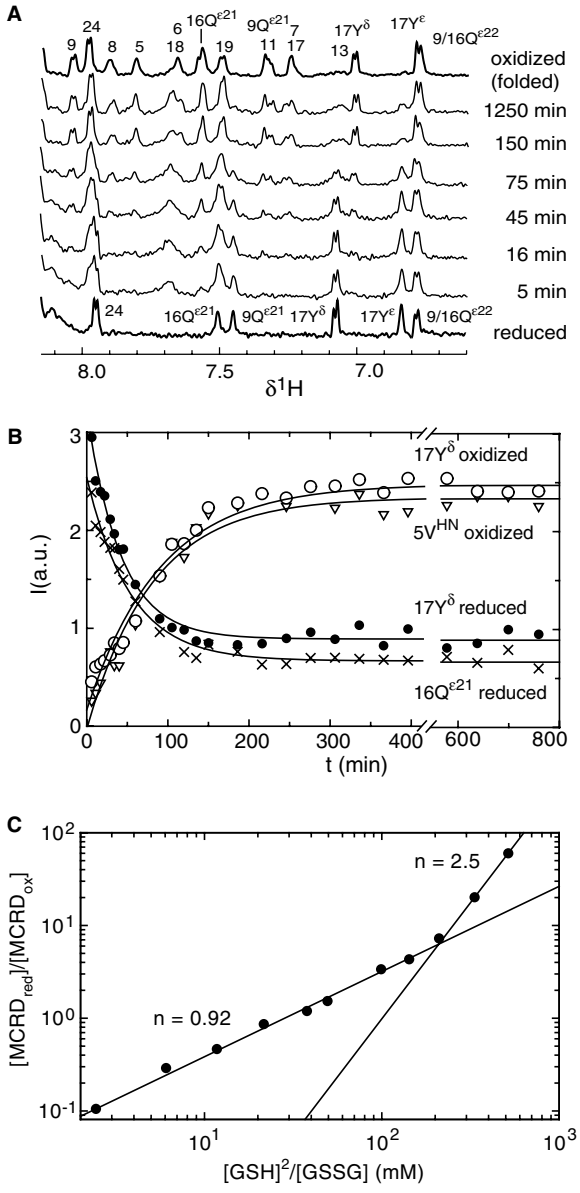

Fig. 5. Oxidative folding and reductive unfolding of MCRD. (A) 1D spectra of MCRD folding initiated by addition of 40 mM GSSG to 0.3 mM MCRD previously reduced with TCEP. Spectra were recorded at the indicated times after starting the oxidation reaction. Backbone $^1H^N$ (plain numbers) and sidechain resonances are indicated. (B) Time course of glutathione catalyzed refolding as monitored by the indicated resonance intensities derived from the spectra in (A). (C) Hill plot of equilibrium reductive unfolding monitored in a solution of varying GSH/GSSG content (see text). Solid lines correspond to linear fits of the low and high $[GSH]^2/[GSSH]$ regions of the correlation.

(Fig. 5B). This may indicate a hidden, albeit lowly populated intermediate, where for example a non-random coil and non-native oxidized ensemble rearranges slowly to the native oxidized state [20].

In order to obtain information on the redox potentials of the disulfides, equilibrium unfolding of MCRD was studied by a titration of the ratio of reduced (GSH) versus oxidized (GSSG) glutathione [21]. The ratio of reduced versus oxidized MCRD, $[MCRD_{red}]/[MCRD_{ox}]$, was obtained from an integration of the corresponding 1D $^1H^N$ and $^1H^\delta$ resonances of threonine-13 and tyrosine-17, respectively (Fig. 5C). Assuming that the loss of the folded oxidized structure is caused by the cooperative reduction of n disulfide bridges, the reaction can be described by

$$MCRD_{ox} + n\,GSSG \leftrightarrow MCRD_{red} + n(2GSH)$$

corresponding to an equilibrium of the form $K_{eq} = [MCRD_{red}]/[MCRD_{ox}] \cdot ([GSH]^2/[GSSH])^n$ [21]. The cooperativity n can be obtained from the slope of a Hill plot of $\log([MCRD_{red}]/[MCRD_{ox}])$ vs. $\log([GSH]^2/[GSSH])$ as shown in Fig. 5C. The transition from the folded, oxidized to the unfolded, reduced form is non-cooperative ($n = 0.92$) at low $[GSH]^2/[GSSG]$ (<100 mM), whereas it becomes more cooperative ($n = 2.5$) at higher $[GSH]^2/[GSSG]$ ratios. This can be explained by a model where at low reducing potential only one disulfide opens, whereas for stronger reducing potentials all three disulfides open. Apparently, the opening of the first disulfide is sufficient to account for the signal loss of the folded resonances. The midpoint of the redox reaction ($[MCRD_{ox}]/[MCRD_{red}] = 1$) is found at a $[GSH]^2/[GSSH]$ ratio of 30 mM. This corresponds to an apparent, effective redox potential E'_0(MCRD) of −185 mV, where the Nernst relation $E'_0(\mathrm{MCRD}) = E'_0(\mathrm{glutathione}) - RT/2F \ln([GSH]_2/[GSSH])_{midpoint}$ and a standard redox potential for glutathione E'_0(glutathione) of −230 mV at 25 °C [22] have been used. Judging from this effective equilibrium constant, MCRD should be predominantly folded with all disulfide bonds intact at physiological concentrations of GSH and GSSG in microsomes ($[GSH]^2/[GSSG] < 3$ mM [23,24]).

4. Discussion

The MCRD domain adopts a new fold that is highly defined by its three disulfide bridges. This structure arises during oxidative folding without populating other structured intermediates to high concentrations. The repeating cysteine pattern CXXXC in MCRD with three intervening residues between the cysteines has been shown to be highly unfavorable for nearest-neighbor disulfide bond formation [25]. Accordingly, the MCRD structure only contains long-range disulfide bonds, which presumably lock into the native conformation during oxidative refolding after the formation of the central, non-accessible disulfide bridge Cys6–Cys14. The wealth of proline residues as well as the formation of hydrogen bonds may restrict the conformational search and stabilize local structure in MCRD during this folding process.

The structure of oxidized MCRD is not unfolded in 8 M urea as evident from ^1H NMR spectra (data not shown), whereas even local structures are abolished by reduction. This underlines the high importance of the disulfide bonds relative to hydrophobic interactions in stabilizing the fold and local native structures. In consequence the cysteine pattern, but not the hydrophobic residues of MCRD domains, is highly conserved [7].

The MCRD domain fulfills its biological role by intermolecular disulfide bond reshuffling [7]. This process is most likely catalyzed [26] and probably involves at least one of the two solvent exposed disulfide bridges Cys2–Cys18 and Cys10–Cys19. Of these, the N-terminal Cys2–Cys18 should be particularly accessible in the N-terminal MCRDs of minicollagen. The equilibrium unfolding with the natural reductant GSH shows the presence of such a single susceptible disulfide bond in MCRD at weakly reducing conditions.

4.1. Data deposition

The atomic coordinates of the 10 lowest energy CNS conformers have been deposited in the RCSB Protein Data Bank (www.rcsb.org) under PDB Accession number 1SP7. Chemical shift assignments have been deposited in the BioMagResBank (www.bmrb.wisc.edu) under Accession number 6200.

Acknowledgements: We thank Jürgen Engel and Suat Özbek for initiating this project and for helpful discussions. This work was supported by SNF grant 31-61'757.00 to S.G.

References

[1] Engel, J. (1997) Science 277, 1785–1786.
[2] Holstein, T. and Tardent, P. (1984) Science 223, 830–833.
[3] Weber, J. (1990) J. Biol. Chem. 265, 9664–9669.
[4] Lenhoff, H.M., Kline, E.S. and Hurley, R. (1957) Biochim. Biophys. Acta 26, 204–205.
[5] Blanquet, R. and Lenhoff, H.M. (1966) Science 154, 152–153.
[6] Engel, U., Pertz, O., Fauser, C., Engel, J., David, C.N. and Holstein, T.W. (2001) Embo J. 20, 3063–3073.
[7] Engel, U., Ozbek, S., Streitwolf-Engel, R., Petri, B., Lottspeich, F. and Holstein, T.W. (2002) J. Cell Sci. 115, 3923–3934.
[8] Ozbek, S., Pertz, O., Schwager, M., Lustig, A., Holstein, T. and Engel, J. (2002) J. Biol. Chem. 277, 49200–49204.
[9] Kurz, E.M., Holstein, T.W., Petri, B.M., Engel, J. and David, C.N. (1991) J. Cell Biol. 115, 1159–1169.
[10] Pokidysheva, E., et al., 2004. J. Biol. Chem., in press.
[11] Hansen, M.R., Mueller, L. and Pardi, A. (1998) Nat. Struct. Biol. 5, 1065–1074.
[12] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) J. Biomol. NMR 6, 277–293.
[13] Garrett, D.S., Powers, R., Gronenborn, A.M. and Clore, G.M. (1991) J. Magn. Resonance 95, 214–220.
[14] Cornilescu, G., Delaglio, F. and Bax, A. (1999) J. Biomol. NMR 13, 289–302.
[15] Brunger, A.T. et al. (1998) Acta Crystallogr. D 54, 905–921.
[16] Sass, H.J., Musco, G., Stahl, S.J., Wingfield, P.T. and Grzesiek, S. (2001) J. Biomol. NMR 21, 275–280.
[17] Koradi, R., Billeter, M. and Wuthrich, K. (1996) J. Mol. Graphics 14, 51.
[18] Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) J. Biomol. NMR 6, 135–140.
[19] Holm, L. and Sander, C. (1994) Proteins 19, 165–173.
[20] Welker, E., Narayan, M., Wedemeyer, W.J. and Scheraga, H.A. (2001) Proc. Natl. Acad. Sci. USA 98, 2312–2316.
[21] Hawkins, H.C., de Nardi, M. and Freedman, R.B. (1991) Biochem. J. 275 (Pt 2), 341–348.
[22] Fasman, G.D. (1976) Handbook of Biochemistry and Molecular Biology. CRC Press, Cleveland.
[23] Hwang, C., Sinskey, A.J. and Lodish, H.F. (1992) Science 257, 1496–1502.
[24] Bass, R., Ruddock, L.W., Klappa, P. and Freedman, R.B. (2004) J. Biol. Chem. 279, 5257–5262.
[25] Zhang, R.M. and Snyder, G.H. (1989) J. Biol. Chem. 264, 18472–18479.
[26] Ozbek, S., Engel, U. and Engel, J. (2002) J. Struct. Biol. 137, 11–14.
[27] Cornilescu, G., Marquardt, J.L., Ottiger, M. and Bax, A. (1998) J. Am. Chem. Soc. 120, 6836–6837.
[28] Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) J. Biomol. NMR 8, 477–486.