NMR spectroscopic characterization of a eukaryotic rubredoxin from the cryptomonad alga *Guillardia theta*

Kristian Schweimer[a], Silke Hoffmann[a], Jürgen Wastl[b], Uwe-G. Maier[b], Paul Rösch[a] and Heinrich Sticht[a]

[a]Lehrstuhl für Biopolymere, University of Bayreuth, 95440 Bayreuth, Germany

[b]Cell Biology and Applied Botany, Philipps-University Marburg, D-35032 Marburg, Germany

Universität Bayreuth

Introduction

Rubredoxins (Rd) are small iron proteins (~ 6 kDa), which contain one iron atom tetrahedrally coordinated by four cysteines. Despite of the fact that all rubredoxins show very similar redox potentials of approximately 0 mV numerous different physiological functions have been reported for them (Sieker et al., 1994). Up to present, the structures of five rubredoxins from various bacteria and archaea have been determined either by NMR spectroscopy or X-Ray crystallography revealing a similar overall structure of these prokaryotic rubredoxins (Sieker et al., 1994; Goodfellow and Macedo, 1999). Recently, the first eukaryotic rubredoxins have been identified by analyzing the genome sequences of *Arabidopsis thaliana* and of the cryptomonad *Guillardia theta* (SWISS-Prot. accession number AAD25628 and CAC40406, respectively).

These Rd's consist of a core fold that shares a significant sequence homology to prokaryotic Rd's including the conserved pattern of iron-ligating cysteines but they differ at several sequence positions that are strictly conserved among prokaryotic Rd's (Sieker et al., 1994). In contrast to the structurally characterized prokaryotic Rd's, the core fold is extended at its N- and C-terminal end by several amino acids of yet unknown function and a membrane anchor is present at the very C-terminus.

Structural investigation of *G. theta* Rd thus should be helpful in identifying the physiological function of eukaryotic rubredoxins (e.g. by verifying surface complementarity to putative redox partners) and in defining the role of the residues extending the core fold. In addition, a structure of this protein will provide the basis for the identification of evolutionary relationships between prokaryotic and eukaryotic Rd's.

We assigned most of the ^{1}H, ^{13}C, and ^{15}N resonances and determined the solution structure from multidimensional heteronuclear NMR data. In order to avoid paramagnetic effects arising from the presence of an iron atom, a zinc substituted form was used for structural investigation. This strategy appears justified by the fact that iron- and zinc-substituted forms of *C. pasteurianum* Rd share an almost identical fold as evidenced from X-Ray data (Dauter et al., 1996).

Methods and results

For the NMR studies *G. theta* Rd, spanning amino acids 57 to 126 of the rubredoxin precursor was cloned into pET28a and overexpressed in *E. coli* BL21(DE3). This 70-residue protein includes the typical Rd fold as well as approximately five amino-terminal and fifteen carboxy-terminal flanking residues. The N-terminal signal sequence and the C-terminal membrane anchor were not included in the construct.

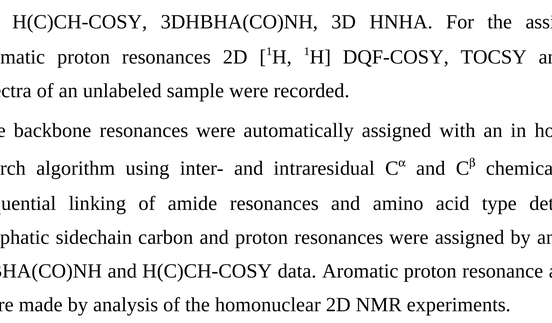
^{15}N-/^{13}C-labeled protein samples were prepared from cells grown in M9-medium NMR samples conditions were: 2.6 mM protein, 10 mM potassium phosphate, pH 6.5, in H_2O / D_2O (9:1).

All NMR experiments were acquired on a Bruker DRX 600 NMR spectrometer at 25 °C. The following NMR spectra were recorded for backbone and aliphatic resonance assignment: 3DHNCO, 3D HNCA, 3D HNCACB, 3D CBCA(CO)NH, 3D H(C)CH-COSY, 3DHBHA(CO)NH, 3D HNHA. For the assignment of aromatic proton resonances 2D [^{1}H, ^{1}H] DQF-COSY, TOCSY and NOESY spectra of an unlabeled sample were recorded.

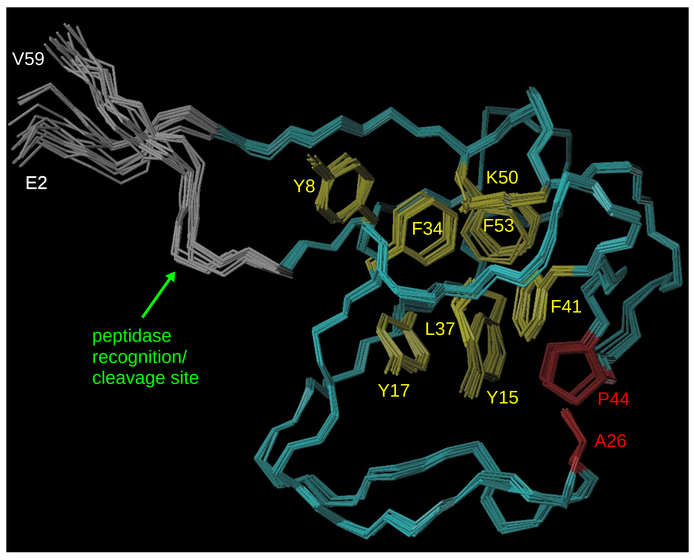
The backbone resonances were automatically assigned with an in house written search algorithm using inter- and intraresidual C^{α} and C^{β} chemical shifts for sequential linking of amide resonances and amino acid type determination. Aliphatic sidechain carbon and proton resonances were assigned by analyzing the HBHA(CO)NH and H(C)CH-COSY data. Aromatic proton resonance assignments were made by analysis of the homonuclear 2D NMR experiments.

Distance restraints for the structure calculation were taken from the ^{13}C-NOESYHSQC, ^{15}N-NOESYHSQC and 2D-NOESY. The solution structures were calculated with XPLOR 3.851 using a modified simulated anneling protocol.

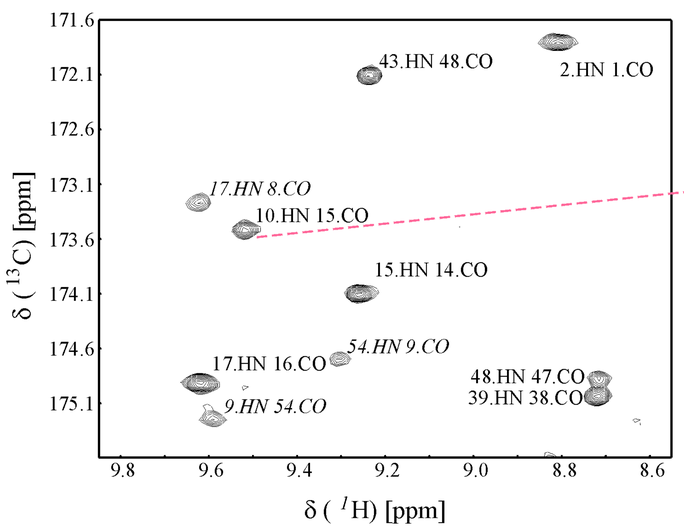
Hydrogen bonds were identified by measuring correlations via the ^{3h}J(^{15}N,^{13}CO) using the long range HNCO experiment (Cordier and Grzesiek, 1999) as well as observing slowly exchanging amide protons in ^{15}N-HSQC spectra after dissolving the labeled sample in D_2O.



^{15}N-HSQC of *G. theta* rubredoxin. The peaks are labeled with the corresponding sequence positions. Peaks marked with an asterisk are aliased along the ^{15}N-dimension. Sidechain NH_2-resonances are marked with a connecting line.



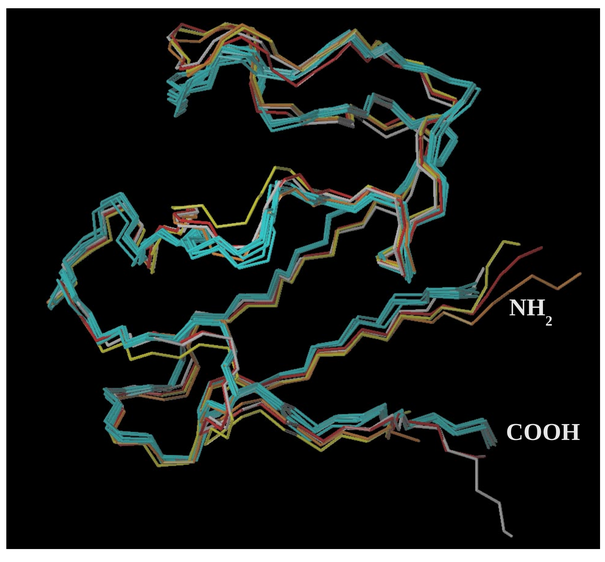
Backbone overlay of a family of 15 structures for residues 2-59. The protein backbone is shown in blue for the rubredoxin core fold and in white for the flanking residues. The sidechains of the residues forming the hydrophobic core are shown in yellow. The amino-terminal residues containing the putative signal peptidase recognition/cleavage site show an increased flexibility compared to the core fold, but still adopt a defined three-dimensional orientation, which is mainly stabilized by non-local interactions to residues of the carboxy-terminal region. This orientation might reflect the structural elements and charge pattern necessary for correct signal peptidase recognition of the *G. theta* rubredoxin precursor.



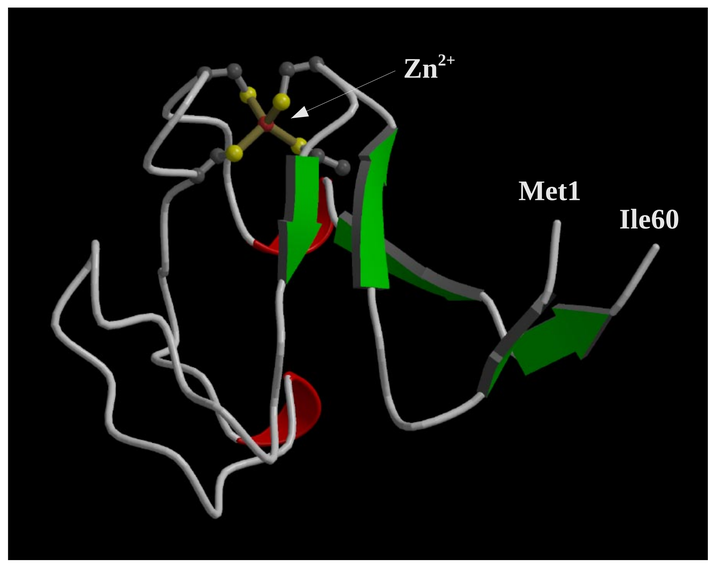
Selected region of the long range 2D H(N)CO. The H(N)CO was recorded with a dephasing delay of 133 ms for the build-up and refocusing of carbon-nitrogen antiphase coherence. The cross peaks showing correlations across the the hydrogen bonds via the $^{3h}J_{NC'}$ are marked in bold, and the residual signals caused by the $^{1}J_{NC'}$ are given in italics.

Summary of structure calculation

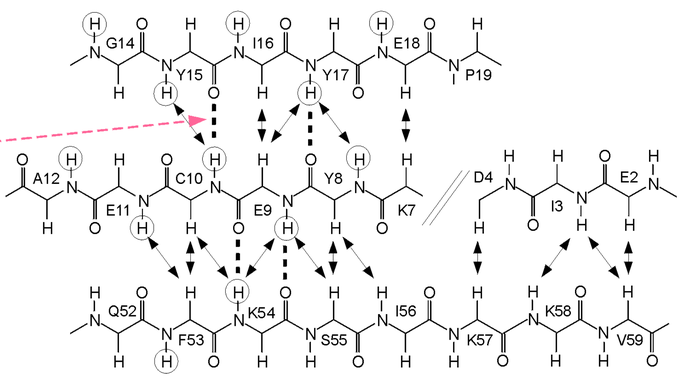
Experimental restraints for the final structure calculation		Atomic r.m.s.d. of 21 calculated structures (Å)		
NOE-restraints		Residues	Backbone	Heavy atoms
sequential (|i-j|=1)	161	6-56 (Rd fold)	0.21	0.72
medium range (|i-j| <5)	107	2-60	0.63	1.07
long range (|i-j| >5)	227			
intraresidual NOEs	42			
Dihedral angle restraints				
^{3}J(H^N,H^{α})	25	Hydrogen bonds		18



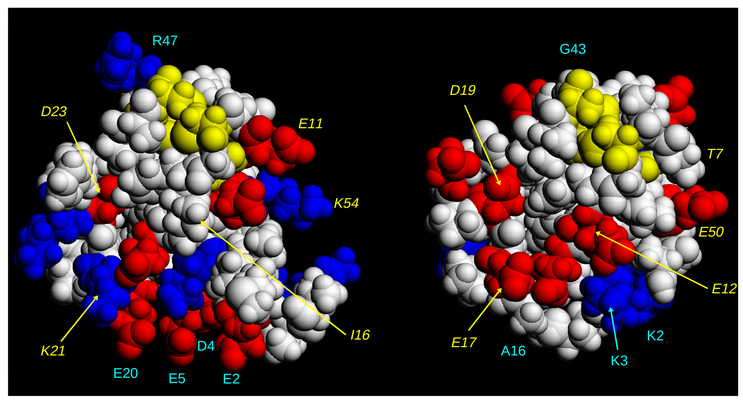
Backbone overlay of the six lowest-energy structures of *G. theta* rubredoxin (blue) with the high resolution X-ray structures of four prokaryotic rubredoxins: *Clostridium pasteurianum* rubredoxin (red), *Desulfovibrio vulgaris* rubredoxin (yellow), Desulfovibrio *gigas* rubredoxin (orange) and *Pyrococcus furiosus* rubredoxin (white). The average pairwise r.m.s.d. of the backbone atoms of the solution structures to the Xray structures is in the range from 1.08 - 1.29 Å.



MOLSCRIPT drawing of the *G. theta* rubredoxin. The additional short beta sheet formed by the five amino- and carboxyterminal residues flanking the core fold is shown. This short beta sheet is deduced from several typical interstrand NOEs (see below), but the fast amide proton exchange as well as rotameric averaging of the $^{3}J(H^N,H^{\alpha})$ coupling constant indicate an increased flexibility compared to the core fold. The ten carboxyterminal residues (61-70) remain unstructured in solution.



Schematic representation of the extended regions in *G. theta* rubredoxin. Inter-strand NOEs and hydrogens bonds unambiguously measured from a long range 2D H(N)CO spectrum are indicated by dashed lines. Slowly exchanging backbone amide protons are marked by open circles.



Electrostatic surface properties of the rubredoxins from *G. theta* (left), *Desulfovibrio vulgaris* (right). Acidic residues are shown in red, basic residues in blue, and the metal-coordinating cysteines in yellow. Key residues, which emphasize the difference between both proteins are labelled. Those polar residues that were proposed to be important for the rubredoxin - cytochrome c_3 interaction in *D. vulgaris* (Stewart et. al., 1989) are emphasized in italics. The significant differences with respect to the charge distribution at the protein surface suggest that *G. theta* rubredoxin exerts a different physiological function compared to the structurally characterized prokaryotic rubredoxins. This hypothesis was confirmed recently by the observation that *G. theta* rubredoxin is associated with photosystem II (Wastl et al., 2000).

References

Cordier, F. and Grzesiek, S. (1999) *J. Am. Chem. Soc.*, **121**, 1601-1602.

Dauter, Z., Wilson, K.S., Sieker, L.C., Moulis, J.M. and Meyer, J. (1996) *Proc. Natl. Acad. Sci. USA*, 93, 8836-8840.

Goodfellow, B.J. and Macedo, A.L. (1999) *Annual Reports on NMR spectroscopy*, 119-177.

Sieker, L.C., Stenkamp, R.E. and Le Gall, J. (1994) *Meth. Enzymol.*, 243, 203-216.

Schweimer, K., Hoffmann, S., Wastl., J., Maier, U.-G., Rösch, P. & Sticht, H. (2000) *Protein Sci.*, in press

Stewart, D.E., Legall, J., Moura, I., Moura, J.J., Peck, H.D., Xavier, A.V., Weiner, P.K., Wampler, J.E. (1989) *Eur J Biochem 185*, 695-700.

Wastl, J., Iuzzolino, L., Dörner, W., Duin, E., Link, T., Hoffmann, S., Sticht, H., Dau, H., Lingelbach, K. & Maier, U.-G. (2000) *J. Biol. Chem.*, in press