



Complete sequential assignment and secondary structure prediction of the cannulae forming protein CanA from the hyperthermophilic archaeon *Pyrodictium abyssi*

Raphael Kreitner¹ · Claudia E. Munte¹ · Katrin Singer¹ · Karl O. Stetter² · Gudrun Horn¹ · Werner Kremer¹ · Hans Robert Kalbitzer¹

Received: 12 December 2019 / Accepted: 27 January 2020 / Published online: 12 February 2020
© The Author(s) 2020

Abstract

CanA from *Pyrodictium abyssi* forms a heat-resistant organic hollow-fiber network together with CanB and CanC. An N-terminally truncated construct of CanA (K₁-CanA) gave NMR spectra of good quality that could be assigned by three-dimensional NMR methods on ¹⁵N and ¹³C-¹⁵N enriched protein. We assigned the chemical shifts of 96% of all backbone ¹H^N atoms, 98% of all backbone ¹⁵N atoms, 100% of all ¹³C^α atoms, 100% of all ¹H^α atoms, 90% of all ¹³C^γ atoms, and 100% of the ¹³C^β atoms. Two short helices and 10 β-strands are estimated from an analysis of the chemical shifts leading to a secondary structure content of K₁-CanA of 6% helices, 44% β-pleated sheets, and 50% coils.

Keywords *Pyrodictium abyssi* · CanA · Cannulae · Extremophile · Solution NMR · Assignments

Abbreviations

CD	Circular dichroism
DSS	2,2-Dimethyl-2-silapentane-5-sulfonate
DTE	1,4-Dithioerythritol
GdmCl	Guanidinium hydrochloride

Biological context

The members of the archaeal family *Pyrodictiaceae* have typical growth temperatures between 75 and 110 °C (348 to 383 K) (Stetter 1982; Stetter et al. 1983; Pley et al. 1991). Three different species of *Pyrodictium* are known, *P. occultum*, *P. Brockii* and *P. abyssi*. *P. abyssi* was isolated from a black smoker, a hydrothermal deep-sea vent 3600 m below sea level (Deininger 1994). As in most of the *Crenarchaeota*

the surface of the cells is covered with a protein layer (S-layer) with hexagonal symmetry. A specific feature of the *Pyrodictiaceae* is the formation of a complex extracellular matrix connecting the cells. The matrix consists of interconnected hollow fibers (cannulae) consisting of helically arranged subunits (König et al. 1988; Rieger et al. 1995). The cannulae are formed from glycoproteins and connect the periplasmic space of different cells (Nickell et al. 2003). While the biological function of cannulae is still unknown, their arrangement suggests that they serve to the exchange of molecular components (metabolites, genetic information, signaling) between the cells in the network (Horn et al. 1999). Optical microscopy under *in-vivo* conditions shows that the polymerisation of these tubules is coupled to the cell division and the daughter cells stay connected after division. The tubules grow with a velocity of 1–1.5 μm/min (Horn et al. 1999). The heat stability of the cannulae is remarkable; even after exposure at 413 K for one-hour intact cannulae are found (Rieger et al. 1995). The cannulae consist of three helically arranged highly homologous glycoprotein subunits CanA, CanB, and CanC (Mai 1998). CanA is composed of 182 amino acids with a total molecular mass of 19.8 kDa. For the transport to the periplasm, a signaling sequence of 25 amino acids is required. No proteins with similar sequences are found in the protein databases. In *Escherichia coli*, expressed CanA monomers spontaneously

✉ Hans Robert Kalbitzer
hans-robert.kalbitzer@biologie.uni-regensburg.de

¹ Institute of Biophysics and Physical Biochemistry, Biophysics I and Centre of Magnetic Resonance in Chemistry and Biomedicine (CMRCB), University of Regensburg, Universitätsstr. 31, 93053 Regensburg, Germany

² Lehrstuhl für Mikrobiologie und Archaeenzentrum, University of Regensburg, Universitätsstr. 31, 93053 Regensburg, Germany

form stable tubules with the same characteristics as native cannulae.

The heat resistance of the hollow-fibers promises interesting biotechnological applications. In nanotechnology, carbon nanotubes are a central research topic. CanA nanotubes could take a similar role in nanobiophysics, since they are heat stable, self-organizing, and can simply be modified by site-directed mutagenesis. A prerequisite for the structure determination of CanA is the complete sequential assignment of uniformly ^{13}C - ^{15}N enriched protein that is presented here.

Materials and experiments

Protein expression and purification

K₁-CanA is an N-terminally truncated construct of CanA whose sequence is depicted in Fig. 1. This truncated version of CanA has been selected for assignments after limited proteolysis experiments (see below). It exhibits NMR spectra with much higher quality than the full-length protein.

The pET17b-CanA plasmid (Amp^r, T7) contains the 549 bp long sequence for CanA without a 75 bp long leading sequence. The pET17b-K1CanA plasmid corresponds to the pET17b-CanA plasmid but with a coding sequence shortened by 30 bp. For obtaining unlabeled CanA or the N-terminally truncated K₁-CanA *E. coli* BL21(DE3)pLysS transformed with pET17b-CanA and pET17b-K1CanA were grown overnight and then inoculated 1:100 into fresh LB-medium containing chloramphenicol (34 mg/L) and carbenicillin (50 mg/L), and grown at 310 K to an OD of 0.8 at 600 nm. Protein expression was induced by adding

IPTG (final concentration 0.8 mM). After 16 h, the cells were harvested by centrifugation. To prepare uniformly ^{15}N - or ^{13}C - ^{15}N -labeled proteins, cells were grown in isotopically labeled New Minimal Medium (NMM) (Budisa et al. 1995), pH 7.4, containing 1 g/L $^{15}\text{NH}_4\text{Cl}$ or 1 g/L $^{15}\text{NH}_4\text{Cl}$ and 2 g/L ^{13}C -glucose, respectively. The yield of pure unlabeled CanA or K₁-CanA was approximately 30 mg protein/L medium, of isotope labeled approximately 10 mg/L. Cells were centrifuged down, washed, and resuspended in 50 mM TRIS, pH 7.4, 50 mM NaCl, 5 mM EDTA. Cells were lysed by sonicating them 2-times for 30 s in an ice bath. Debris was spun down at 14×10^3 g for 15 min at 277 K. The supernatant was shaken in a water bath with a temperature of 353 K for 15 min. After the heat denaturation, the sample was cooled in an ice bath for 10 min and the denatured protein was removed by centrifugation. The supernatant was transferred to Q-sepharose (HiLoad 26/10 Q Sepharose Fast Flow) column equilibrated with 50 mM TRIS-HCl, pH 7.4. The protein was eluted with a linear NaCl gradient using an Äkta chromatography system (ÄKTA™ FPLC + Frac-900) with flow rate of 4 mL/min and gradient mixing time (0 to 500 mM NaCl). All fractions were monitored by SDS/PAGE on 13% acrylamide gels. The fractions containing the protein were pooled where CanA (or K₁-CanA) eluted at a NaCl concentration of approximately 150 mM. After concentrating the obtained CanA sample with a Vivaspinn system (Sartorius Stedim Biotech, Göttingen) it was transferred to a gel filtration column (HiLoad 26/60 Superdex 75 prep grade) equilibrated with 50 mM TRIS/HCl pH 7.4, 50 mM NaCl and eluted isocratically. The purity of the sample was checked by SDS-PAGE, the integrity of the protein was confirmed by mass spectrometry. Amino acid sequencing

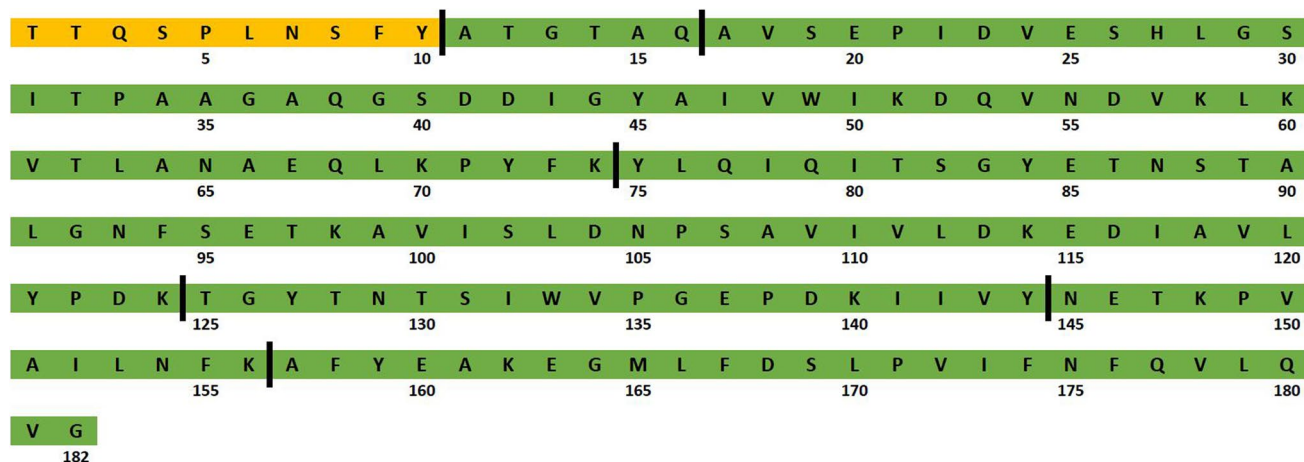


Fig. 1 Amino acid sequence of CanA from *P. abyssi*. The sequence of K₁-CanA is depicted in green. In orange, the additional amino acids of the full length CanA after removal of the signaling sequence are

shown. The cleavage sites obtained by limited proteolysis of CanA with trypsin for 3 h are labeled with black bars

Table 1 Configuration of the peptide bond in front of proline residues^a

Amino acid	C ^β (ppm)	C ^γ (ppm)	Δ _{βγ} (ppm)	Strong NOE	Isomer
P21	32.11	27.27	4.84	– ^b	trans
P33	31.65	26.96	4.69	H ^α (32)–H ^δ (33)	trans
P71	31.33	27.74	3.59	H ^N (70)–H ^δ (71)	trans
P106	32.74	26.49	6.25	H ^α (105)–H ^δ (106)	trans
P122	32.27	26.80	5.47	– ^b	trans
P135	31.33	– ^c	–	– ^b	?
P138	31.18	26.80	4.38	H ^α (137)–H ^δ (138)	trans
P149	31.78	– ^c	–	– ^b	?
P171	32.43	26.33	6.10	H ^α (170)–H ^δ (171)	trans

^aThe peptide bond configuration was determined from the chemical shift difference Δ_{βγ} of the C^β and C^γ shifts of the proline residues according to (Schubert et al. 2002) and/or from the NOEs between the amino acid X preceding the proline residue and the proline residue according to Wüthrich et al. (1984)

^bResonances could not be assigned

^cNo unambiguous NOEs could be found

HNCO, HNCA, CBCA(CO)NH, CBCANH, HCAN, [¹H, ¹⁵N]-NOESY-HSQC and HCCH-TOCSY experiments. Aromatic side chains were assigned with (HB)CB(CGCC-TOCSY)Har and (HB)CB(CGCD)HD experiments. Assignments were performed with the program AUREMOL (Gronwald and Kalbitzer 2004; <https://www.auremol.de>). Backbone dihedral angles and secondary structure

propensities were predicted with the program TALOS-N (Shen and Bax 2013).

Extent of assignment and data deposition

The NMR data of K₁-CanA were recorded typically at a protein concentration of 0.5 mM in 20 mM sodium phosphate buffer (Na₂HPO₄/NaH₂PO₄) pH 6.6 containing 0.1 mM EDTA, 0.4 mM NaN₃, 0.4 mM DSS and either 90% H₂O/10% D₂O or 100% D₂O, at 323 K. Only at this elevated temperature, spectra with that high quality could be obtained. Figure 2 shows an [¹H, ¹⁵N]-HSQC-spectrum of K₁-CanA at this temperature. Note that *P. abyssi* is growing in a temperature range between 340 to 378 K (“physiological temperature”) (Marteinsson et al. 1997).

We assigned the chemical shifts of 96% of all backbone ¹H^N atoms, 98% of all backbone ¹⁵N atoms, 100% of all ¹³C^α atoms, 100% of all ¹H^α atoms, 90% of all ¹³C^γ atoms, and 100% of the ¹³C^β atoms. The side chain resonances were almost completely assigned including all aromatic side chains. Only the proton resonances of side chain amino groups of the 12 lysine residues were broadened beyond detection and could therefore not be assigned. Under favorable conditions, the configuration of the peptide bond in front of proline residues can be determined from the chemical shift difference Δ_{βγ} of the ¹³C^β and ¹³C^γ resonances. For the *trans*- and *cis*-isomers one obtains Δ_{βγ} values of 4.51 ± 1.37 and 9.64 ± 1.62 ppm, respectively (Schubert et al. 2002). For 7 of the 9 proline residues the *trans*-configuration can be confirmed by this method, for P135 and P149 the method cannot be used since the ¹³C^γ

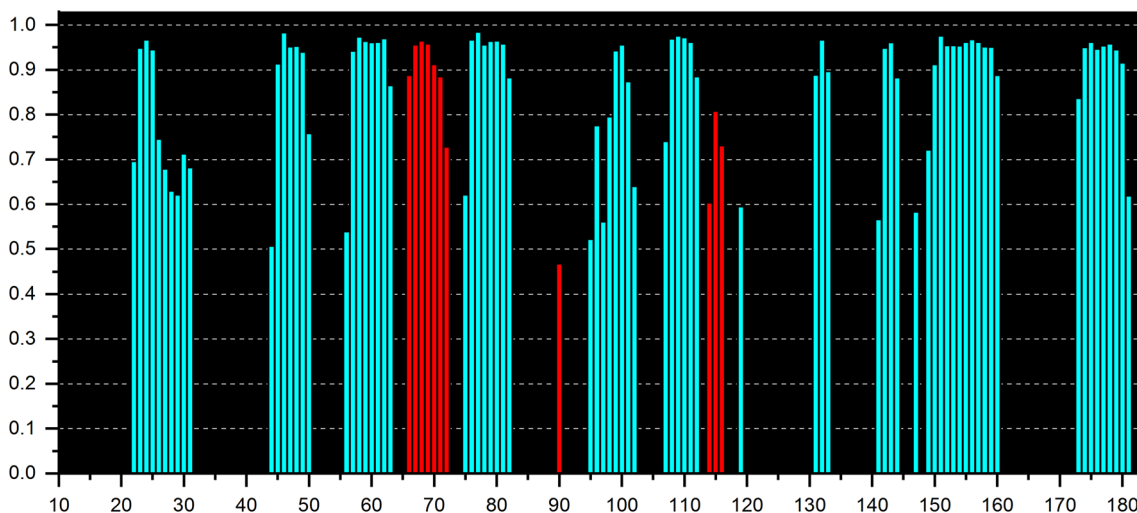


Fig. 3 Secondary structure propensities of K₁-CanA from chemical shifts. The secondary structure propensities were obtained from the chemical shifts of C^γ, C^α, C^β, N, H^N and H^α atoms by the program

TALOS-N (Shen and Bax 2013), (blue) propensities for β-strands, (red) propensities for helices

Table 2 Location of secondary structure elements in K₁-CanA

Secondary structure	Amino acid	Helix	Strand	Coil
L1	11–21			×
E1	22–31		×	
L2	32–43			×
E2	44–50		×	
L3	51–55			×
E3	56–63		×	
L4	64–65			×
H1	66–72	×		
L5	73–74			×
E4	75–82		×	
L6	83–94			×
E5	95–102		×	
L7	103–106			×
E6	107–112		×	
L8	113			×
H2	114–116	×		
L9	117–130			×
E7	131–133		×	
L10	134–140			×
E8	141–144		×	
L11	145–148			×
E9	149–160		×	
L12	161–172			×
E10	173–181		×	
L13	182			×

The secondary structure propensities were obtained from the chemical shifts of C', C^α, C^β, N, H^N and H^α atoms by the program TALOS-N (Shen and Bax 2013)

Li coil regions, *Ei* β-strand regions, *Hi* helix

resonances could not be identified (Table 1). Alternatively, the *cis*–*trans* isomers can be identified by sequential NOEs between the amino acid X preceding the proline residue and the proline. The *trans*-isomer is characterised by strong H^α_(X)–H^{δ*}_(Pro) and H^N_(X)–H^{δ*}_(Pro) NOEs, the *cis*-isomer by strong H^α_(X)–H^α_(Pro) and H^N_(X)–H^α_(Pro) NOEs (Wüthrich et al. 1984). With these data, the conclusions from the chemical shift analysis could be confirmed. Stereospecific assignments of side chain amide protons of asparagine and glutamine residues were performed with the program AssignmentChecker contained in AUREMOL that is based on the chemical shift difference between the two geminal protons (Harsch et al. 2017). Asparagine and glutamine amide proton resonances can be stereochemically assigned if chemical shift differences are ≥ 0.40 ppm for asparagine and ≥ 0.42 ppm for glutamine with a confidence level > 95%. In this case, the downfield shifted resonance lines can be assigned to H^{δ21} and H^{ε21}, respectively. These assignments were confirmed, if possible, by

the analysis of the corresponding NOEs (Wüthrich 1986). In this manner, the side chain amide protons stereospecific assignment of all 8 glutamine and of 7 from 9 asparagine could be performed. In one asparagine (N154) the amide proton resonances are degenerate and their mean is shifted significantly upfield (6.30 ppm). The resonance assignments were deposited in the BioMagResBank (<https://www.bmrb.wisc.edu>) under accession number 50124.

Secondary structure prediction

From an analysis of the C', C^α, C^β, N, H^N and H^α chemical shifts by the program TALOS-N the secondary structure propensities can be estimated. They are depicted in Fig. 3 showing high propensities for two helices and 10 β-strands. The corresponding regions in the sequence are presented in Table 2. In total, from the chemical shifts a secondary structure content of K₁-CanA of 6% helices, 44% β-pleated sheets and 50% coils is predicted. Since the first 10 amino acids of CanA most probably are also disordered, similar values are obtained for the complete protein (6% helices, 41% β-pleated sheets and 53% coils). These experimental values are surprisingly close to the secondary structure prediction obtained by PredictProtein (<https://www.predictprotein.org>) (Rost et al. 2004) that predicts 7% α-helices, 43% β-pleated sheets and 50% coils from the amino acid sequence.

Acknowledgements Open Access funding provided by Projekt DEAL. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie (FCI).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Budisa N, Steipe B, Demange P, Eckerskorn C, Kellermann J, Huber R (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. Eur J Biochem 230:788–796
- Deininger W (1994) Untersuchungen zur hitzeabhängigen expression des Thermosoms bei *Pyrodictium* sowie Isolierung und Charakterisierung eines neuen Vertreters der Art *Pyrodictium* (TAG11) aus der Tiefsee. Diploma thesis, University of Regensburg.

- Gronwald W, Kalbitzer HR (2004) Automated structure determination of proteins by NMR spectroscopy. *Progr NMR Spectr* 44:33–96
- Harsch T, Schneider P, Kieninger B, Donaubaue H, Kalbitzer HR (2017) Stereospecific assignment of the asparagine and glutamine side chain amide protons in proteins from chemical shift analysis. *J Biomol NMR* 67:157–164
- Horn C, Paulmann B, Kerlen G, Junker N, Huber H (1999) In vivo observation of cell division of anaerobic hyperthermophiles by using a high-intensity dark-field microscope. *J Bacteriol* 181:5114–5118
- König H, Messner P, Stetter KO (1988) The fine structure of the fibers of *Pyrodictium occultum*. *FEMS Microbiol Lett* 49:207–212
- Mai B (1998) In vitro Untersuchungen zum extrazellulären Netzwerk von *Pyrodictium abyssi* TAG11. PhD thesis, University of Regensburg.
- Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich K (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *Pure Appl Chem* 70:117–142
- Marteinsson NT, Moulin P, Birrien J-L, Gambacorta A, Vernet M, Prieur D (1997) Physiological responses to stress conditions and barophilic behavior of the hyperthermophilic vent archaeon *Pyrococcus abyssi*. *Appl Environ Microbiol* 63:1230–1236
- Nickell S, Hegerl R, Baumeister W, Rachel R (2003) *Pyrodictium* canulae enter the periplasmic space but do not enter the cytoplasm, as revealed by cryo-electron tomography. *J Struct Biol* 141:34–42
- Pley U, Schipka J, Gambacorta A, Jannasch H, Fricke H, Rachel R, Stetter KO (1991) *Pyrodictium abyssi* sp. nov. represents a novel heterotrophic marine archaeal hyperthermophile growing at 110 °C. *Syst Appl Microbiol* 14:245–253
- Rieger G, Rachel R, Hermann R, Stetter KO (1995) Ultrastructure of the hyperthermophilic archaeon *Pyrodictium abyssi*. *J Struct Biol* 115:78–87
- Rost B, Yachdav G, Liu J (2004) The PredictProtein server. *Nucl Acids Res* 32:W321–W326
- Schubert M, Labudde D, Oschkinat H, Schmieder P (2002) A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ¹³C chemical shift statistics. *J Biomol NMR* 24:149–154
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241
- Stetter KO (1982) Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105 °C. *Nature* 300:258–260
- Stetter KO, König H, Stackebrandt E (1983) *Pyrodictium* gen. nov., a new genus of submarine disc-shaped sulphur reducing archaeobacteria growing optimally at 105 °C. *Syst Appl Microbiol* 4:535–551
- Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York
- Wüthrich K, Billeter M, Braun W (1984) Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances. *J Mol Biol* 180:715–740

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.