J|A|C|S ARTICLES

Published on Web 10/20/2010

Optimized Measurement Temperature Gives Access to the Solution Structure of a 49 kDa Homohexameric β-Propeller

Ilka Varnay,[†] Vincent Truffault,[§] Sergej Djuranovic,[‡] Astrid Ursinus,[‡] Murray Coles,*[,‡] and Horst Kessler*[,†]

Institute for Advanced Study and Center of Integrated Protein Science, Department Chemie, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany, Department of Protein Evolution, Max-Planck-Institute for Developmental Biology, Spemannstr. 35, 72076 Tübingen, Germany, Department of Biochemistry, Max-Planck-Institute for Developmental Biology, Spemannstr. 35, 72076 Tübingen, Germany

Received July 21, 2010; E-mail: horst.kessler@ tum.de; murray.coles@tuebingen.mpg.de

Abstract: Ph1500 is a homohexameric, two-domain protein of unknown function from the hyperthermophilic archaeon *Pyrococcus horikoshii*. The C-terminal hexamerization domain (Ph1500C) is of particular interest, as it lacks sequence homology to proteins of known structure. However, it resisted crystallization for X-ray analysis, and proteins of this size (49 kDa) present a considerable challenge to NMR structure determination in solution. We solved the high-resolution structure of Ph1500C, exploiting the hyperthermophilic nature of the protein to minimize unfavorable relaxation properties by high-temperature measurement. Thus, the side chain assignment (97%) and structure determination became possible at full proton density. To our knowledge, Ph1500C is the largest protein for which this has been achieved. To minimize detrimental fast water exchange of amide protons at increased temperature, we employed a strategy where the temperature was optimized separately for backbone and side chain experiments.

Introduction

Ph1500 is a hexameric, two-domain protein from the hyperthermophilic archaeon *Pyrococcus horikoshii*. Although the function is unknown, both domains are of structural interest. The N-domain (Ph1500N) shows clear similarity to a β-clam domain to date only observed in the substrate recognition domains of various AAA ATPases.[1] In contrast, the C-terminal domain (Ph1500C) shows no significant sequence similarity to proteins of known structure. While the N-terminal domain is monomeric in isolation (8.5 kDa) and could be solved by conventional NMR methods (to be published elsewhere), the 49 kDa hexameric C-domain presented a considerable challenge, representing one of the largest protein structures undertaken ab initio by NMR. Here, we present the full (97%) resonance assignment and the high-resolution structure of the Ph1500C homohexamer, solved at full proton density.

Structural studies of large proteins by NMR remain a challenge, as molecular correlation time increases approximately linearly with molecular weight, leading to increased line widths that reduce the effective signal-to-noise and exacerbate signal overlap. This limits the application of NMR in solving structures of larger proteins. In recent years, advances in sample preparation and measurement techniques have addressed this problem, substantially increasing the molecular mass of proteins accessible to NMR analysis. These methods principally rely on deuteration[2] in combination with transverse relaxation-optimized experiments (TROSY[3] and methyl-TROSY[4−7]) or direct ^{13}C-excited/detected experiments.[8−10] Deuteration aims to reduce relaxation rates by dilution of the proton density; thus, it necessarily lowers the number of proton−proton NOE contacts available to define the structure. To circumvent this problem, selective labeling strategies are used to reintroduce protons either site-specifically, for example, for isoleucine, leucine, and valine methyl groups,[11,12] or stereospecifically.[13] However, the labeled precursors required for these strategies are expensive, and the resolution obtainable depends on the density and distribution of the labels.

[†] Department Chemie, Technische Universität München.
[‡] Department of Protein Evolution, Max-Planck-Institute for Developmental Biology.
[§] Department of Biochemistry, Max-Planck-Institute for Developmental Biology.

(1) Djuranovic, S.; Rockel, B.; Lupas, A. N.; Martin, J. *J. Struct. Biol.* **2006**, *156*, 130–138.
(2) LeMaster, D. M.; Richards, F. M. *Biochemistry* **1988**, *27* (1), 142–150.
(3) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366–12371.
(4) Yang, D.; Zheng, Y.; Liu, D.; Wyss, D. F. *J. Am. Chem. Soc.* **2004**, *126*, 3710–3711.
(5) Tugarinov, V.; Kay, L. E. *J. Biomol. NMR* **2004**, *28*, 165–172.
(6) Tugarinov, V.; Sprangers, R.; Kay, L. E. *J. Am. Chem. Soc.* **2004**, *126*, 4921–4925.
(7) (a) Tugarinov, V.; Kay, L. E. *ChemBioChem.* **2005**, *6*, 1567–1577. (b) Religa, T. I.; Kay, L. E. *J. Biomol. NMR* **2010**, *47*, 163–169.
(8) Bertini, I.; Felli, I. C.; Kümmerle, R.; Moskau, D.; Pierattelli, R. *J. Am. Chem. Soc.* **2004**, *126*, 464–465.
(9) Bermel, W.; Bertini, I.; Felli, I. C.; Kümmerle, R.; Pierattelli, R. *J. Magn. Reson.* **2006**, *178*, 56–64.
(10) Hu, K.; Vögeli, B.; Clore, G. M. *J. Biomol. NMR* **2006**, *36*, 259–266.
(11) Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 627–636.
(12) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. *J. Biomol. NMR* **1999**, *13*, 369–374.
(13) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Ono, A. M.; Guentert, P. *Nature* **2006**, *440*, 52–57.

 10.1021/ja1064608

An alternative approach for larger proteins is to reduce relaxation rates by directly reducing the correlation time. This is most easily achieved by measuring at increased temperature, while the emerging technique of encapsulating proteins in reverse micelles suspended in low viscosity solvents[14] takes this principal to another level. If the correlation time can be decreased sufficiently, deuteration is no longer necessary, and measurement can be carried out at full proton density, promising to extend high-resolution structure determination to much larger proteins. However, even in these cases, spectral complexity remains an issue, due to the sheer number of resonances involved. For this reason, and the fact that the various existing techniques that allow mapping of interdomain or complexation surfaces are less sensitive to molecular size, most of the larger protein structures solved to date have been assembled from multiple domains or subunits that represented tractable problems in isolation. A notable exception is for symmetric homo-oligomers and cases where segmental labeling[15] is possible. Here, the spectral complexity corresponds to that of the monomer or segment, respectively. The Ph1500C homohexamer is a good example of this, as its spectral complexity corresponds to the 71 residue monomer, and we reasoned it would be feasible to solve its structure by NMR. Initially, we had planned to apply the existing standard techniques for large proteins; however, measurement at high temperature (353 K) reduced the effective correlation time to about a third of that at room temperature, allowing structure determination at full proton density. To date, the number of NMR structures measured at elevated temperature is limited. The current study offers an example how potential problems, such as faster water exchange of amide protons, can be overcome by optimizing the measurement temperature.

Materials and Methods

Ph1500C was expressed in *E. coli* C41 (DE3) at 37 °C after induction with 1 mM IPTG at OD600 ~0.6. Minimal media supplement contained ^{15}N-labeled NH_4Cl as the sole nitrogen source and ^{13}C-labeled glucose as the sole carbon source, while ^{2}D/^{13}C-labeled glucose and D_2O were used as deuterium sources. After purification, using a combination of Ni-NTA affinity and gel filtration chromatography, the protein was dialyzed against a 20 mM sodium phosphate buffer (pH 7.4) containing 250 mM NaCl.

Two samples were prepared, a 0.6 mM uniformly ^{2}H,^{13}C/^{15}N-labeled sample on a T101P mutant of Ph1500C and a 1.2 mM uniformly ^{13}C/^{15}N-labeled sample of the wild type protein. The mutation was inadvertent and first traced during sequential assignment. To determine the practical thermostability, a 40 μL volume of the protein solution proved sufficient. First, a 20 μL aliquot was heated slowly in a clear 200 μL Eppendorf-cap, with denaturation detected by visible aggregation (~95 °C, 268 K). Another 20 μL aliquot was heated for one week at 85 °C. This was then diluted, and intact folding was checked with a 1D NMR spectrum. In contrast, the ^{2}H,^{13}C/^{15}N-labeled sample showed aggregation at ~323 K. Different thermostabilities of the double- and triple-labeled samples may be due to different purification levels, rather than the effect of deuteration. An effect of the mutation can be excluded, since a previously prepared 0.6 mM uniformly ^{13}C/^{15}N-labeled sample on the T101P mutant was also stable at ~95 °C.

All experiments were recorded at temperatures between 318 and 353 K on a Bruker DMX600 and a Bruker DMX750 spectrometer with a TXI probehead equipped with triple-axis gradients. Due to technical limitations, transverse axis gradients were not used at temperatures over 330 K.

Backbone and side chain assignment of Ph1500C was first performed on the T101P mutant of Ph1500C. Later, the chemical shifts were reassigned on the wild type sample by comparing triple resonance and ^{15}N-HSQC NOESY slices. Backbone sequential assignments were completed to 96% (with exception of K83, V135, and I136) using TROSY-based versions of standard triple-resonance experiments (trHNCO, trHN(CA)CO, trHNCA, and trHNCACB) recorded on the 85% deuterated mutant sample at 318 K. The latter two residues were later assigned by analyzing trHNCO, trHNCA, and HN(CA)HA experiments recorded at 353 K on the double-labeled sample of wild type Ph1500C, thus completing the backbone assignment to 99%. Aliphatic side chain assignments were completed to 97% (missing residues are E122 and R130) using a combination of H(C)CH-TOCSY and (H)CCH-TOCSY using the relaxation-optimized MOCCA mixing sequence[16] (mixing time 16 ms) and (H)CCH-COSY experiments measured at 353 K on the double-labeled wild type sample. As a consequence of the different measurement temperatures, it was necessary to reassign backbone and C^{β} signals at 353 K during the side chain assignment process. As the ^{13}C chemical shifts are less temperature dependent than amide protons, it was sufficient to use trHNCO, trHNCA, and HN(CA)HA spectra recorded at the higher temperature, together with a series of TROSY spectra recorded at intermediate temperatures (323, 333, and 343 K) to trace chemical shift changes. ^{15}N-HSQC-based experiments were run using the extended aliphatic flipback scheme for fast pulsing,[17] demonstrating the advantage of this technique over selective excitation schemes for large proteins.

Assignments of the aromatic residues were made by linking the aromatic spin systems to the respective $C^{\beta}H_2$ protons in a 2D NOESY spectrum, supported by a ^{13}C-HSQC of the aromatic region. Side chain χ_1 rotamer determination and stereospecific assignment were made for 15 of 46 prochiral $C^{\beta}H_2$ protons. Determinations of χ_1 rotamers were also available for 8 of 9 isoleucine residues, 2 of 3 threonine residues, and 6 of 10 valine residues, leading to a stereospecific assignment of the $C^{\gamma}H_3$ groups. Determinations of χ_2 rotamers were made for 7 of 9 isoleucine residues. All rotamer assignments were made by a hypothesis-test approach where expectation spectra back-calculated from a local conformational hypothesis were compared to experimental NOESY spectra.

Distance restraints were derived manually from the conventional ^{15}N- and ^{13}C-HSQC NOESY spectra and the heteronuclear edited trNNH-, CCH-, and CNH-NOESY spectra.[18] A 2D-NOESY spectrum was recorded with destructive filtering of ^{15}N-bound protons, thus revealing well-resolved traces for the three aromatic residues. In addition, a folded ^{13}C-HSQC NOESY spectrum was recorded of the methyl region from 13 to 31 ppm in the indirect carbon dimension, which was essential for identifying intra- and especially intermolecular methyl−methyl contacts.

The structure calculation with XPLOR (NIH version 2.9.3) was performed using the data described in Table 1. Restraints were classified into four categories corresponding to upper interproton distance restraints of 2.7, 3.2, 4.0, and 5.0 Å. Lower distance restraints were included for very weak or absent intraresidue or sequential H^N-H^{α}, $H^N(i)-H^{\beta}$ and H^N-H^N crosspeaks using a minimum distance between 2.5 and 4.0 Å. Hydrogen-bond restraints were applied via inclusion of pseudocovalent bonds.[19] Chi1 and Chi2 rotamer restraints were defined with a tolerance of ±30°. Dihedral angle restraints were derived for backbone Φ and Ψ angles based on C^{α}, C^{β}, C′, N, and H^{α} chemical shifts using the programs

(14) Wand, A. J.; Ehrhardt, M. R.; Flynn, P. F. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15299–15302.
(15) Xu, R.; Ayers, B.; Cowburn, D.; Muir, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 388–393.
(16) Furrer, J.; Kramer, F.; Marino, J. P.; Glaser, S. J.; Luy, B. *J. Magn. Reson.* **2004**, *166*, 39–46.
(17) Diercks, T.; Daniels, M.; Kaptein, R. *J Biomol NMR* **2005**, *33*, 243–259.
(18) Diercks, T.; Coles, M.; Kessler, H. *J. Biomol. NMR* **1999**, *15*, 177–180.
(19) Truffault, V.; Coles, M.; Diercks, T.; Abelmann, K.; Eberhardt, S.; Lüttgen, H.; Bacher, A.; Kessler, H. *J. Mol. Biol.* **2001**, *309*, 949–960.

Table 1. Structural Statistics and Atomic R.M.S. Deviations[a]

A. Structural Statistics

	SA	<SA>$_r$
R.M.S.D. from distance restraints (Å)[b]		
all (1124)	0.007 ± 0.0005	0.007
intraresidue (66)	0.000 ± 0.0000	0.000
inter-residue sequential (159)	0.006 ± 0.0006	0.007
medium range (82)	0.002 ± 0.0007	0.003
long-range (232)	0.010 ± 0.0006	0.010
intermolecular (448)	0.007 ± 0.0004	0.006
H-bond (128)	0.002 ± 0.0008	0.001
persistent violations threshold (Å)[c]	0.10	
R.M.S.D. from dihedral restraints (°) (245)	0.13 ± 0.008	0.12
persistent violations threshold (°)	1.1	
H-bond restraints (Å/deg)[d] (64)	2.21 ± 0.10/9.9 ± 6.4	2.22 ± 0.10/9.8 ± 6.4
deviations from ideal covalent geometry		
bonds (Å × 10^{-3})	8.9 ± 0.04	8.9
angles (°)	0.50 ± 0.01	0.49
impropers (°)	0.89 ± 0.01	0.88
Ramachandran map regions[e] (%)	98.5/100.0	97.9/100.0

B. Atomic R.M.S. Differences (Å)[f]

	SA vs <SA>		SA vs <SA>$_r$	
	backbone	all	backbone	all
ordered residues (hexamer)	0.10 ± 0.05	0.25 ± 0.05	0.38 ± 0.03	0.67 ± 0.02
ordered residues (monomer)	0.08 ± 0.02	0.24 ± 0.05	0.25 ± 0.02	0.60 ± 0.02
<SA> vs <SA>$_r$	0.37	0.62		

[a] Structures are labeled as follows: SA, the set of 17 final simulated annealing structures; <SA>, the mean structure calculated by averaging the coordinates of SA structures after fitting over secondary structure elements; <SA>$_r$, the structure obtained by regularising the mean structure under experimental restraints. [b] Numbers in brackets indicate the number of restraints of each type. [c] Persistent violations are defined as those violating in over 75% of structures. [d] H bonds were restrained by treating them as pseudocovalent bonds (see Materials and Methods section). Deviations are expressed as the average distance/average deviation from linearity for restrained H bonds. [e] Determined using the program MOLPROBITY.[22] Percentages are for residues in 98% and >99.8% regions of the Ramachandran map. [f] Based on heavy atoms superimpositions over ordered residues (V77-R156).

TALOS[20] and SimShiftDB.[21] Refinement was carried out by comparison of experimental and back-calculated 2D and 3D NOESY spectra (in-house software).

Fifty structures were calculated in an initial simulated annealing protocol and refined in two further slow cooling stages. The first included a conformational database potential, while in the second, the force constant on peptide bond planarity was relaxed to 50 kcal/mol/rad^2. Seventeen structures were selected for the final set on the basis of lowest restraint violations. An average structure was calculated and regularized to give a structure representative of the ensemble (used in all figures). Validation parameters were obtained from the program Molprobity.[22]

Results

Data Collection. For Ph1500C, we adopted the strategy of optimizing the measurement temperatures with respect to the type of experiment. ^{13}C-HSQC-based NOESY and side chain assignment experiments, where amide protons are not involved, could be carried out at the highest temperature tolerated by the hardware (353 K). Figure 1 shows a comparison of the H$^{\alpha}$ region of ^{13}C-HSQC spectra at four temperatures. The spectral quality was continuously improved with increasing temperature due to reduced molecular correlation time. For experiments where amide protons are involved in the magnetization transfer pathway, temperature selection was less straightforward due to increased water exchange of amide protons at higher temperature. Accordingly, signals showed either increased or decreased intensity, depending on which effect was dominant. In experiments measured at 318 K, peaks due to three amide protons (K83, V135, and I136) were missing, due to either very weak intensity or overlap. Although V135 and I136 reappeared at 353 K, seven additional signals (K85, N95, E103, K120, E122, N123, and N131) had disappeared and further signals showed strongly decreased intensities. In ^{15}N-HSQC-based spectra, all signals arising from the highly solvent exposed NH_2-groups have disappeared at 353 K (Figure 2).

The 3D triple resonance experiments employing multiple INEPT-based magnetization transfer were more exposed to the effects of rapid transverse relaxation, and increasing temperature was especially advantageous (Figure 3). In contrast, NOESY experiments were less sensitive to transverse relaxation, as the magnetization is longitudinal during NOESY mixing but were more exposed to fast exchange of solvent exposed protons with water (Figure 2).

In optimizing measurement temperature for collection of structural data, the optimal temperature for side chain assignment and NOESY experiments, which necessarily need to be recorded at the same temperature, seemed to be the highest possible temperature (353 K). We considered that the disadvantage of missing NOESY traces was outweighed by a more complete methyl group assignment in combination with a higher quality of ^{13}C-HSQC-NOESY experiments, especially since the protein has a high content (42%) of isoleucine, leucine, and valine residues, whose methyl−methyl contacts were anticipated to be vital to define the structure. The optimal temperature for

(20) Cornilescu, G.; Delagio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289–302.
(21) (a) Ginzinger, S. W.; Fischer, J. *Bioinformatics* **2006**, *22* (4), 460–465. (b) Ginzinger, S. W.; Coles, M. *J. Biomol. NMR* **2009**, *43*, 179–185.
(22) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., III; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*, 375–383.

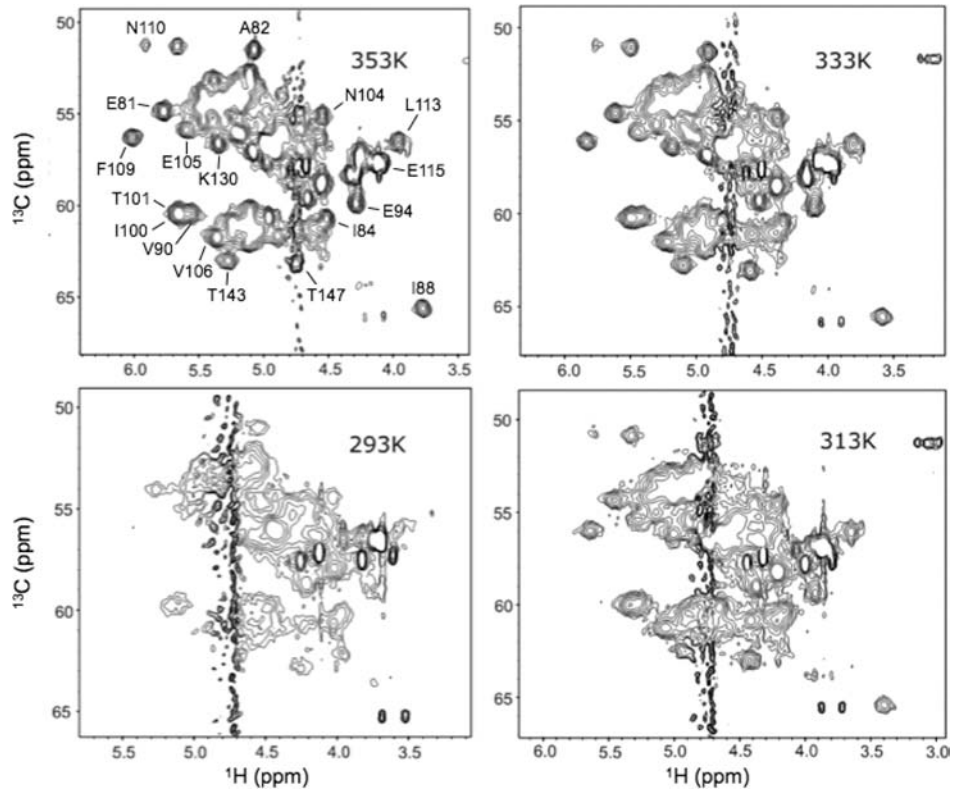

Figure 1. Effect of increasing temperature on the quality of spectra of Ph1500C. Comparison of the H^{α} region of ^{13}C-HSQC spectra recorded at 750 MHz on a uniformly $^{13}C/^{15}N$-labeled sample of Ph1500C. A constant increase in signal-to-noise ratio and resolution is visible with increasing temperature.

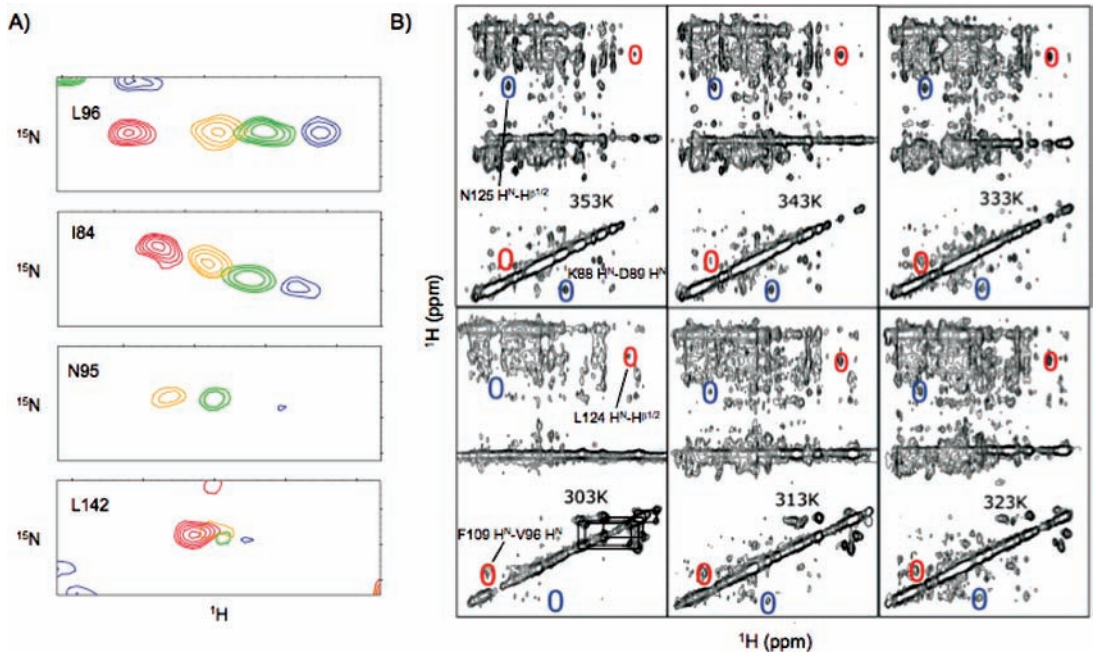

Figure 2. Effects of temperature on amide proton signals. Spectra measured at 750 MHz on a uniformly $^{13}C/^{15}N$-labeled sample are shown. (A) Selected signals from TROSY spectra at four temperatures (323 K, blue; 333 K, green; 343 K, orange; 353 K, red). The top two show examples of average temperature behavior, i.e., a gradual increase in intensity, reaching a maximum close to the maximum temperature. Two further examples show examples where signals are missing at high temperature (N95) or where intensity increases markedly over the temperature range (L142). (B) 2D H^N-H projections of a 3D ^{15}N HSQC-NOESY experiment. The general line narrowing effect with increasing temperature is visible. In contrast, the signal-to-noise ratio of individual signals is either decreased or increased. Blue circles mark examples of solvent protected amide protons where the effect of faster water exchange is negligible. Red circles mark examples of solvent exposed amide protons where the effect of a faster tumbling at higher temperatures does not compensate for faster exchange with water. Side chain NH_2 groups (connected by lines in the low-temperature spectrum) disappear completely at high temperature. For the sake of clarity, chemical shift axes are omitted; the areas from ~6.8−10.2 ppm in the direct dimension and from ~0.2−10.2 ppm in the indirect dimension are shown.

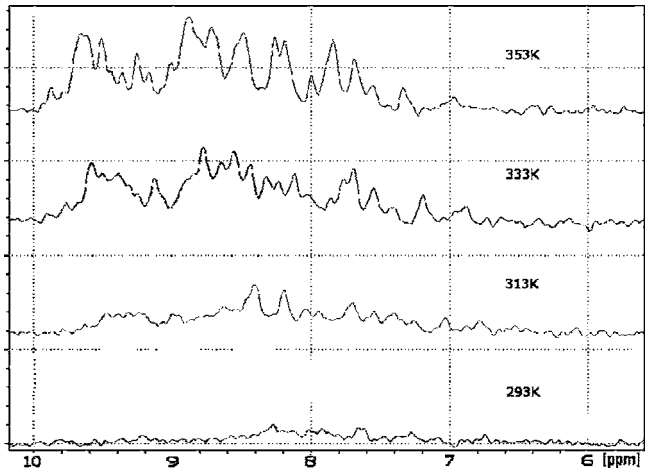

Figure 3. Comparison of the first processed FID of 3D HNCA spectra measured at four temperatures showing pronounced line narrowing with increasing temperature.

sequential assignment was around 328 K, as several Cβ resonances, which are essential for the sequential assignment, disappeared in the less sensitive trHNCACB experiment at higher temperature. However, since the triple-labeled sample proved less thermostable, the temperature chosen for the backbone experiments measured on this sample was 318 K. As a consequence, it was necessary to reassign backbone amide signals in the TROSY spectrum and Cα and Cβ resonances at 353 K during the side chain assignment process. As the ^{13}C chemical shifts are less temperature dependent than amide protons, it was sufficient to use trHNCA, trHNCO, and HN(CA)HA spectra recorded at the higher temperature, together with a series of TROSY spectra recorded at intermediate temperatures (323, 333, and 343 K) to trace chemical shift changes. A positive feature of the strategy described here follows from the fact that deuteration has a negative effect on ^{13}C-HSQC-based experiments, while it has a solely positive effect on triple-resonance experiments. Thus, by using a 85% deuterated sample only for backbone assignment, it was possible to compensate for the lower optimal measurement temperature of the experiments involved.

Structure Determination. Secondary structure analysis of the NMR data revealed the Ph1500C monomer to consist of eight β-strands. These could be unambiguously linked by interstrand contacts into two similar, four-stranded β-sheets, one formed of strands β2-β3-β4-β5 and the second of β6-β7-β8-β1, with interstrand loops of five residues linking β1-β2 and β5-β6 (Figure 4). Initial model building showed both faces of both sheets to be highly hydrophobic, implying that all four should be involved in building the hydrophobic core of the domain. This is in line with preliminary electron microscopy data on the full Ph1500 hexamer, which showed a single ring of subunits with a 6-fold symmetry axis arranged around a central pore (data not shown). To bury these hydrophobic faces, the two sheets have four possible orientations. However, a number of unambiguously assigned contacts of well resolved methyl groups and aromatic side chains between the sheets and between the β5-β6 interstrand loop and one sheet interface excluded all but one orientation and thus allowed the construction of a model for the Ph1500 monomer.

In constructing the hexamer, two topological possibilities remained: one based on completely intramolecular β-sheet contacts, and a second based on circular domain swapping of the β1-strand (Figure 4). The latter was suggested by the similarity of each sheet to "blades" of seven-bladed, monomeric β-propellers, identified on the basis of chemical shifts and sequence using the program SimShiftDB.[21] While the first possibility placed both interstrand loops at one end of the structure (Figure 4), which was both sterically unfavorable and poorly supported by the experimental data, the circular domain swap proved consistent, particularly explaining contacts between the β1-β2 interstrand loop and side chains at the largely intermolecular sheet interface.

The final structure can be described as a 12-bladed, homohexameric β-propeller (Figure 5). In total, 52 and 28 distance restraints were assigned at each interblade interface; the higher number of restraints corresponds to the mostly intramolecular interface, with the difference due to the larger number of well-resolved aromatic and aliphatic side chains. The structure was confirmed in an iterative process by extensive back-calculation of expectation NOESY spectra. The high quality of NOESY experiments and the almost complete side chain assignment resulted in a highly resolved structure of very good stereochemical quality (Table 1).

The structure agrees well with the preliminary EM data in terms of its architecture and overall dimensions. Particularly, the diameter of the inner pore (2 nm) is in good agreement. Significantly, it is larger than observed in seven-bladed propellers as consequence of insertion of an additional five blades into the ring. The outer cross-section diameter could not be measured accurately due to a diffuse outline of the averaged EM pictures, consistent with the flexible attachment of the N-domain to the periphery of the hexameric ring.

The final structure ensemble is very well-defined, with an rmsd for superimposition of backbone residue of 0.10 Å for superimposition over the hexamer. Residual restraint violations were very low, with an rmsd for distance restraints of 0.007 Å and no persistent violation greater than 0.10 Å (see Table 1). The structure reflects the internal symmetry observed in the secondary structure, with the two blades having an rmsd of 1.9 Å (Figure 5). To our knowledge, Ph1500C is the first instance of both a natively oligomeric and a 12-bladed β-propeller.

Discussion

The global tumbling correlation time τ_c of an isotropically reorienting spherical molecule of volume V_h relates to the bulk solvent viscosity η through the well-known Stokes−Einstein relation:

$$\tau_c = \frac{\eta V_h}{kT}$$

At first sight, increasing the absolute temperature in the denominator would seem to decrease the correlation time only marginally. However, the viscosity of water itself has a strong, nonlinear dependence on temperature. Accordingly, the correlation times at 328 and 353 K are expected to be about half and one-third of that at room temperature. This is reflected in the observed linewidths (Figure 6). The line width of water is also highly correlated with its viscosity. The parallel in the temperature behavior of the water and protein linewidths (Figure 6) therefore demonstrates that solvent viscosity is indeed the dominant effect in determining protein line widths in this system.

Examination of the structure can explain most of the outliers in temperature behavior among backbone amides (Figure 5C). First, amide groups where protein signals were either not

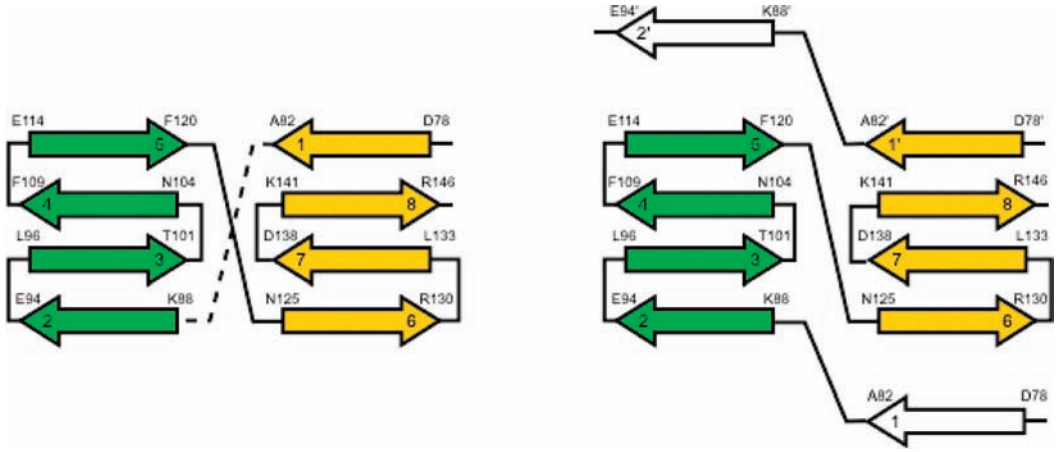

Figure 4. Topology of Ph1500C. The structure is built of two very similar four-stranded β-sheets linked by two intersheet loops. A simple topology (left) requires these loops to cross at one end of the structure (dashed line). A domain-swapped topology, where β1 is donated to the sheet of a neighboring protomer, avoids this crossing.

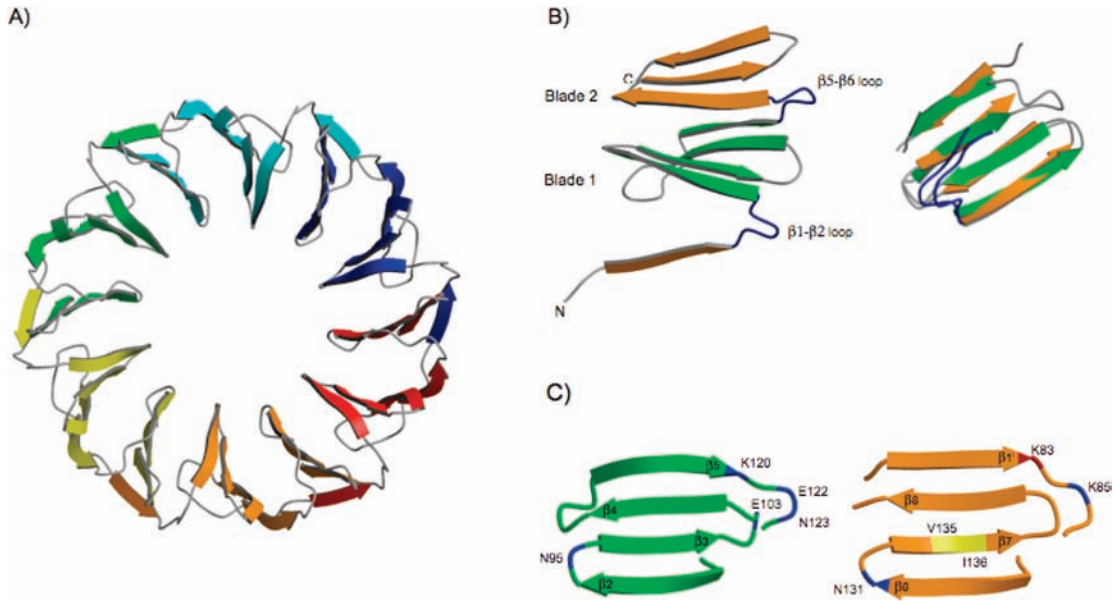

Figure 5. Solution structure of Ph1500C. (A) The Ph1500C hexamer. A cartoon representation is shown, with the view on the larger opening of a funnel-shaped ring. Each color represents one monomer unit. (B) The Ph1500C monomer. A monomer unit is shown with the first blade (β2-β3-β4-β5) in green and the second (β6-β7-β8-β1) in orange. The intersheet loops are in blue. A superimposition of two blades drawn from the hexamer is shown on the right. (C) Temperature behavior of amide protons. Amide protons with extreme temperature behavior are mapped onto the structure. The blades are shown separately for clarity. Residues that disappear at higher temperature are shown in blue, those that appear at high temperature are in yellow, and K83, which was not observed at any temperature, is in red.

observed at any temperature or disappeared at higher temperature are almost exclusively located in the hairpin turns linking β-strands or in the intersheet loops. The structure shows that none of these amide groups participate in main chain hydrogen bonds, and only one (N131) has a potential side chain hydrogen-bonding partner. Interestingly, each β-sheet has two edge strands, which have similarly unpaired hydrogen-bond donors. However, in this case, all remain observable at the highest temperature. Also notable are two residues in the β7 strand (V135 and I136), where amide signals were unexpectedly absent at room temperature but appeared at higher temperature. It is likely that the broad lines for these signals at room temperature are the result of chemical exchange due to internal motions on the millisecond time scale. The above average narrowing of these lines at higher temperatures could therefore be attributed to a shift in these motions to faster time scales. The size of Ph1500C however precludes quantitative relaxation studies to investigate this in more detail.

NOESY spectra on uniformly labeled samples are the richest source of information for protein structure determination. The current study raises the question as to how generally high temperature measurement can be used to access this information for larger proteins. At first glance, Ph1500C appears to be an exception, perhaps representing the size limit for structure determination at full proton density. However, our results indicate that this is not the case. First, the hexamer forms a ring with a diameter around three times that of each subunit, resulting in a considerably longer effective correlation time than would be expected for a spherical or rod-shaped protein of similar molecular mass or for multiple domains connected by flexible linkers. Second, although Ph1500C is extremely thermostable, 353 K is the highest measurement temperature

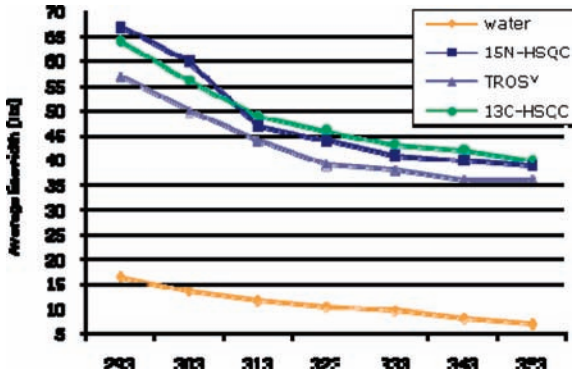

Figure 6. Temperature dependence of the line width of water and protein signals. The average line width of various protein signals (T101, G119, and L124 in ^{15}N-HSQC and TROSY; I100 $H^{\gamma 12}$, I136 $H^{\delta 1}$, I137 H^{β}, and I45 $H^{\gamma 2}$ in ^{13}C-HSQC) measured in the proton dimension on a 750 MHz spectrometer on the uniformly ^{13}C/^{15}N-labeled sample of Ph1500C.

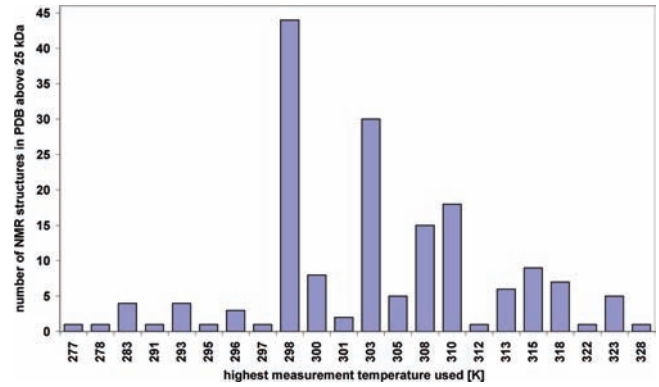

Figure 7. Reported measurement temperatures of NMR structures above 25 kDa.[25] Prior to this study, 328 K was the highest measurement temperature reported.

reported for a NMR structure determination, it was advantageous, but not necessary, to measure at the highest possible temperature. Adequate NOESY spectra could have been obtained at intermediate temperatures (e.g., 323 K, see Figure 2), where many proteins of thermophilic and even some of mesophilic origin would be stable.

A prerequisite for interpreting NOESY information on uniformly labeled samples is a largely complete side chain assignment, and this is a limiting factor. For Ph1500C, the highest measurement temperature was most advantageous for obtaining side chain assignments with standard (H)CCH-TOCSY experiments. For proteins that are less thermostable, experiments less sensitive to transverse relaxation or suitable for deuterated samples would be needed. In recent years several strategies involving direct ^{13}C detection have been proposed, for example, ^{13}C-detected HCC-TOCSY[23] and ^{13}C$-^{13}$C-NOESY spectra,[9−11,24] which should allow full assignment at lower temperatures.

Assuming that the upper limit of a structure determination based on a highly protonated, uniformly labeled sample would be 25 kDa at room temperature (298 K), then the limit at 328 K would be 50 kDa, neglecting the greater spectral complexity. Figure 7 shows a histogram of measurement temperatures for NMR structure determinations of proteins above 25 kDa.[25] The majority of these proteins were solved at temperatures between 298 and 303 K. Therefore, an increase in measurement temperature of 10−20 K would place many proteins in a temperature range where structure determination at high, uniform proton density would be possible. Despite the fact that manipulation of solution conditions and protein constructs, including the introduction of mutations, has been widely used to improve the solubility and general solution behavior of NMR samples, we are unaware of any instance where similar efforts have been aimed at improving thermostability.

The results presented here suggest that optimizing the measurement temperature, possibly by improving thermostability,[26,27] is potentially as important for larger proteins as optimizing deuteration levels and labeling schemes. Accordingly, we suggest that further improvements in engineering proteins for thermostability should have an impact on NMR structure determination, allowing larger proteins to be solved at high, uniform protein density.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft.

JA1064608

(23) Hu, K.; Vögeli, B.; Pervushin, K. *J. Magn. Reson.* **2005**, *174* (2), 200–208.
(24) Matzapetakis, M.; Turano, P.; Theil, E. C.; Bertini, I. *J. Biomol. NMR* **2007**, *38*, 237–242.
(25) RCSB Protein Data Bank; http://www.pdb.org/pdb/home/home.do (accessed July 2010).
(26) Lehmann, M.; Pasamontes, L.; Lassen, S. F.; Wyss, M. *Biochim. Biophys. Acta* **2000**, *1543*, 408–415.
(27) Lehmann, M.; Wyss, M. *Curr. Opin. Biotechnol.* **2001**, *12*, 371–375.