# Solution NMR structure of the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima*

## Werner Kremer, Benjamin Schuler, Stefan Harrieder, Matthias Geyer, Wolfram Gronwald, Christine Welker, Rainer Jaenicke and Hans R. Kalbitzer

Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Germany

Cold-shock proteins (Csps) are a subgroup of the coldinduced proteins preferentially expressed in bacteria and other organisms on reduction of the growth temperature below the physiological temperature. They are related to the cold-shock domain found in eukaryotes and are some of the most conserved proteins known. Their exact function is still not known, but translational regulation, possibly via RNA chaperoning, has been discussed.

Here we present the structure of a hyperthermophilic member of the Csp family. The NMR solution structure of *Tm*Csp from *Thermotoga maritima*, the hyperthermophilic member of this class of proteins, was solved on the basis of

The cold-shock response in micro-organisms is a transient phenomenon affecting the growth rate of the cell and the saturation of fatty acids as well as the rate of synthesis of DNA, RNA, and protein at temperatures significantly lower than the normal physiological temperature [1]. For most proteins, the expression under cold shock is dramatically decreased, whereas for a few it is increased. Of the latter, one group of small acidic proteins shows an extremely high induction level and high affinity to single stranded nucleic acids [2]. Because of their high sequence homology, they are classified together in the family of cold-shock proteins (Csps) (reviewed in [3–5]). The biological function is still not known, but translational regulation, possibly via RNA chaperoning, is the possibility currently favored [6].

The first hyperthermophilic member of this family has been cloned from *Thermotoga maritima* (*Tm*Csp); its sequence and physicochemical characteristics were recently reported [7]. *Tm*Csp is a small globular protein of 66 amino acids with a molecular mass of 7474 Da. It exhibits extreme intrinsic stability making it the most thermostable Csp known at present [7].

The Csps with known 3D structure are Greek-key  $\beta$  barrel proteins and belong to the 'OB-fold family' [8],

1015 conformational constraints. It contains five  $\beta$  strands combined in two antiparallel  $\beta$  sheets making up a  $\beta$  barrel structure, in which  $\beta$  strands 1–4 are arranged in a Greekkey topology. The side chain of R2, which is exclusively found in thermophilic members of the Csp family, probably participates in a peripheral ion cluster involving residues D20, R2, E47 and K63, suggesting that the thermostability of *Tm*Csp is based on the peripheral ion cluster around the side chain of R2.

*Keywords*: cold-shock protein; hyperthermophiles; NMR spectroscopy; structure determination; *Thermotoga maritima*.

which includes the known structures of CspA from Escherichia coli [9-11], CspB from Bacillus subtilis [12,13], and CspB from Bacillus caldolyticus [14], as well as several oligonucleotide-binding and/or oligosaccharide-binding proteins. The structural analysis of CspA from E. coli revealed putative RNP1 and RNP2 sequence motifs commonly found in ssRNA-binding proteins [15,16]. These structural features suggest that all Csps may function as single stranded nucleic acid-binding proteins [9-14] and possibly have a regulatory role. It has been shown that CspA activates transcription of the hns and gyrA genes encoding H-NS proteins and subunit A of DNA gyrase, which are also cold-inducible [17,18]. The single stranded nucleic acid interaction has been verified by chemical-shift perturbation analysis of complexes between CspA and ssDNA [10], ssRNA-binding gel-shift assays [19,20], tryptophan fluorescence-quenching studies [21], and sitedirected mutagenesis of the ssDNA-binding function of E. coli CspA [21] and B. subtilis CspB [22]. For EcCspA, the binding specificity has not yet been determined; for ssDNA the sequences CCAAT and ATTGG seem to be preferred while ssRNA is bound unspecifically. A thorough analysis of the interaction of BsCspB with ssDNA templates revealed that BsCspB preferentially binds to polypyrimidine but not polypurine ssDNA templates [2]. Thymine-based ssDNA templates bind with high-affinity and salt-independently, whereas binding of cytosine-based ssDNA templates is strongly salt-dependent, indicating that a large electrostatic component is involved in the interactions [2]. On binding, each BsCspB seems to cover a stretch of 6-7 thymine bases on T-based ssDNA [2]. The binding of BsCspB to T-based ssDNA template is enthalpically driven, indicating the possible involvement of interactions between aromatic side chains on the protein with the thymine bases [2]. The precise role of the single stranded nucleic acid-binding function in the cell biology of

*Correspondence to* H. R. Kalbitzer, Institut für Biophysik und physikalische Biochemie, Universität Regensburg, D-93040 Regensburgh, Germany. Fax: + 49 941 9432479,

Tel.: + 49 941 9432594,

E-mail: hans-robert.kalbitzer@biologie.uni-regensburg.de

Abbreviations: Bc, Bacillus caldolyticus; Bs, Bacillus subtilis; Bst, Bacillus stearothermophilus; Tm, Thermotoga maritima; Ec,

*Escherichia coli*; Csp, cold-shock protein; HSQC, heteronuclear single quantum coherence.

<sup>(</sup>Received 14 November 200, revised 12 February 2001, accepted 23 February 2001)

the cold-shock response is still not fully understood (reviewed in [1,23,24].

Regarding their relative intrinsic stabilities against guanidinium hydrochloride and temperature, bacterial Csps were found to reflect the optimum growth temperatures of their mesophilic, thermophilic, and hyperthermophilic origin. The corresponding melting points are:  $T_m(BsCspB) \approx 325$  K,  $T_m(EcCspA) \approx 333$  K,  $T_m(BcCspB) \approx 345$  K,  $T_m(TmCsp) \approx 360$  K [25–28]. In this respect, the recently published crystal structure of BcCspB [14] showed that the location, rather than the number of charged residues, is important for stability. Given the close sequence similarity of Csps, what is the mechanism by which TmCspgains its higher stability compared with its mesophilic and moderately thermophilic counterparts? A prerequisite for understanding its thermostability is the determination of its structure.

### MATERIALS AND METHODS

### **Protein purification**

Protein expression of unlabeled protein was performed as described [7]. For the expression of <sup>15</sup>N isotopically enriched protein, E. coli BL21 (DE3) cells containing the plasmid coding for TmCsp were grown in M9 minimal medium containing 0.5 g·L<sup>-1</sup> <sup>15</sup>NH<sub>4</sub>Cl and 1 g·L<sup>-1</sup> glucose at 310 K (37 °C) in the presence of 500  $\mu$ g mL<sup>-1</sup> ampicillin to an  $D_{550}$  of 1. Expression was induced by adding 1 mM isopropyl  $\beta$ -D-thiogalactopyranoside, and bacterial growth was continued at 299 K (26 °C) for about 16 h. Purification of labeled protein was performed as described [7]: to remove the bulk of the E. coli proteins without significant coprecipitation of TmCsp, the cell-free extract was diluted fivefold and then heated to 80 °C for 30 min. Pure TmCsp was obtained after hydrophobic interaction chromatography at pH 8 and size exclusion chromatography. The final yields were between 5 and 20 mg pure <sup>15</sup>N-enriched *Tm*Csp per liter of cell culture. NMR samples of unlabeled and 15N-labeled TmCsp contained 1.5 mm protein in 50 mm sodium phosphate buffer (pH 6.5), 20 mM NaCl, and 0.2 mM EDTA (sodium salt) in  ${}^{1}\text{H}_{2}\text{O}/{}^{2}\text{H}_{2}\text{O}$  (92%/8%) or  ${}^{2}\text{H}_{2}\text{O}$ .

#### NMR spectroscopy

The NMR experiments were carried out on Bruker DMX 500-MHz, Bruker DRX 600-MHz and Bruker DMX 800-MHz spectrometers. <sup>1</sup>H-<sup>1</sup>H 2D NOESY spectra were recorded at 500 and 800 MHz as described by Jeener et al. [29] with a mixing time of 80, 100, and 150 ms at temperatures 288 K and 303 K. At 800 MHz, <sup>1</sup>H-<sup>1</sup>H 2D TOCSY spectra were recorded as described by Braunschweiler and Ernst [30] using an MLEV-17 decoupling sequence [31] of 60 ms for isotropic mixing at temperatures of 288 K, 303 K and 318 K and <sup>1</sup>H-<sup>1</sup>H 2D DQF-COSY spectra at 303 K. Phase-sensitive detection in  $\omega_1$  was obtained by the TPPI method [32]. For echo suppression, a pair of orthogonal spin lock pulses of 5 ms and 2.5 ms were applied before the low-power water suppression pulse. <sup>15</sup>N-resolved 3D NOESY-heteronuclear single-quantum coherence (HSQC) (mixing time 150 ms) and TOCSY-HSQC experiments (mixing time 80 ms) using TPPI in  $t_1$ ,

and echo/anti-echo gradient selection in  $t_2$  [33], and <sup>15</sup>N GARP decoupling in  $t_3$  were carried out on a Bruker DMX-500 operating at a proton frequency of 500 MHz and a <sup>15</sup>N frequency of 50.68 MHz at 303 K. Proton resonances of the side chains were assigned from the 2D-TOCSY spectrum recorded with 60 ms mixing time at 800 MHz both in <sup>1</sup>H<sub>2</sub>O and <sup>2</sup>H<sub>2</sub>O and from the <sup>15</sup>N-resolved TOCSY-HSQC spectrum recorded with 80 ms mixing time at 500 MHz in  ${}^{1}\text{H}_{2}\text{O}$ . In addition, sequence-specific resonance assignments were carried out on the basis of HNCA, CBCA(CO)NH, HNCO all of them recorded on the 600-MHz spectrometer in  ${}^{1}\text{H}_{2}\text{O}$  at 303 K. For the spinsystem recognition, a HCCH-TOCSY spectrum in <sup>2</sup>H<sub>2</sub>O was recorded at the same magnetic field strength and 303 K. Distance restraints were obtained from the two homonuclear 2D NOESY spectra with 80 ms mixing time at 800 MHz in  ${}^{1}\text{H}_{2}\text{O}$  and  ${}^{2}\text{H}_{2}\text{O}$  and from the  ${}^{15}\text{N}$ -resolved NOESY-HSQC with 150 ms mixing time at 500 MHz in <sup>1</sup>H<sub>2</sub>O. The NOE evaluation from the 3D-NOESY spectra was iteratively refined using NOE back-calculation methods based on the complete relaxation matrix formalism [34,35]. Dihedral  $\phi$  angle restraints were based on the  ${}^{3}J_{H^{N}H^{\alpha}}$  coupling constants measured in high-resolution DQF-COSY experiments (digital resolution of 8192 complex points in  $t_2$ ). Slowly exchanging amide protons were

Table 1. Structural statistics.

Constraints	Number
NOEs	933
Intraresidual (i, i)	358
Sequential $(i, i + 1)$	283
Backbone–backbone	125
Backbone-side chain	153
Side chain-side chain	5
Intermediate range $(i, i + n; n \le 5)$	47
Backbone-backbone	8
Backbone-side chain	29
Side chain-side chain	10
Long range $(i,j; j > i + 5)$	245
Backbone-backbone	89
Backbone-side chain	98
Side chain-side chain	58
φ-Angle constraints	28
Hydrogen bonds	54

Structural statistics for the 21 lowest-energy structures (from 50 calculated)

E <sub>total</sub>	$5140.6 \pm 258.8 \text{ kJ} \cdot \text{mol}^{-1}$
E <sub>NOE</sub>	$2269.2 \pm 190.6 \text{ kJ} \cdot \text{mol}^{-1}$
NOE violations $> 0.05$ nm	$6.95 \pm 1.65$
Edihed	$190.0 \pm 29.7 \text{ kJ} \cdot \text{mol}^{-1}$
Ebond	$70.4 \pm 15.1 \text{ kJ} \cdot \text{mol}^{-1}$
Eangle	$1200.4 \pm 82.9 \text{ kJ} \cdot \text{mol}^{-1}$
$E_{ m vdW}$	$952.9 \pm 66.6 \text{ kJ} \cdot \text{mol}^{-1}$
$E_{ m imp}$	$231.5 \pm 23.9 \text{ kJ} \cdot \text{mol}^{-1}$
rmsds <sup>a</sup>	
Amino acid 1-66 (whole protein)	0.094 nm (0.151 nm)
Core regions	0.025 nm (0.087 nm)

<sup>a</sup> All backbone atoms, values in parentheses all non-hydrogen atoms.

identified from a series of  ${}^{1}\text{H}{}^{-15}\text{N}$  HSQC spectra recorded in  ${}^{2}\text{H}{}_{2}\text{O}$  buffer.

The proton chemical shifts were referenced to sodium-2,2-dimethyl-2-silapentane-5-sulfonate (DSS) used as internal reference. The <sup>15</sup>N chemical shifts were indirectly referenced to DSS [36] using the frequency ratio given by Wishart *et al.* [37]. Spectral analysis, peak picking, volume integration, and relaxation matrix back calculation were performed using the program AURELIA [38].

### Structure calculations

Approximate interproton distances were obtained from the NOE cross-peaks in the 2D  $^{1}$ H- $^{1}$ H 2D NOESY spectra deriving classifications of very strong (up to 0.2 nm), strong (up to 0.25 nm), medium (up to 0.35 nm), weak (up to 0.45 nm), and very weak (up to 0.55 nm) distances with upper bound distance restraints set to 0.35 nm, 0.40 nm, 0.50 nm, and 0.60 nm, respectively, as described elsewhere [39]. In the case of spectral overlap, NOE distance restraints were derived from the 3D  $^{15}$ N-resolved NOESY-HSQC spectra, referencing the NOE intensities to a known distance of 0.23 nm and setting an upper bound distance restraint of an additional 40%. The minimum sum of the van der Waals radii (0.18 nm) was used as the lower bound distance restraints.

The amide hydrogen-bond donors were identified as slowly exchanging amide protons, and acceptors were identified from the constraints imposed by interstrand NOEs. In the final structure calculations, the protein backbone hydrogen-bond restraints were included, using upper limits of 0.19 nm for NH-O and 0.29 nm for N-O distances.

J-coupling constants were determined by 2D DQF-COSY experiments at 500 MHz and 303 K in <sup>1</sup>H<sub>2</sub>O, using a sinebell window apodizaton. Peak-to-peak separation was determined with the program AURELIA [38], and J-coupling was taken as the average distance in the  $\omega_2$ direction of the multiplet components. The backbone torsion angle  $\phi$  was determined using the Karplus relation, and an estimated error of 20% was added in the structure calculation.

Calculations were performed using the program XPLOR, version 3.851 [40] and the dynamic simulated annealing protocol for extended strand starting structures [41]. In the structure calculation, 1015 conformational NMR constraints were included (Table 1). The initial structure was energy minimized with 1000 cycles of Powell minimization. High-temperature dynamics were run for 30 ps at an initial temperature of 1000 K. The system was then slowly cooled to a temperature of 1000 K in 50-K steps over a period of 20 ps. At 100 K, a final stage of 2000 steps of Powell minimization was performed to yield the final simulated annealing structures. The stereochemical quality



Fig. 1. Sequence alignment of 30 prokaryotic Csp sequences. BACST, B. stearothermophilus; LISMO, Listeria monocytogenes; BACSU, B. subtilis; BACCE, B. cereus; LACPL, Lactobacillus plantarum; PSEAE, Pseudomonas aeruginosa; STRCO, Streptomyces coelicolor; SALTY, Salmonella typhimurium; STIAU, Stigmatella aurantiaca, PSEFR, Pseudomonas fragi; ECOLI, E. coli; ARTGO, Arthrobacter globiformis; MYCTU, Mycobacterium tuberculosis; STRCL, S. clavuligerus; HAEIN, Haemophilus influenzae. Residues conserved more than 80% are colour coded in red, residues conserved more than 60% are colour coded in yellow, and residues conserved more than 40% are colour coded in light blue. The pairwise identity of TmCsp with BcCspB is 63%, TmCsp with BsCspB is 64%, TmCsp with EcCspA is 53%, BcCspB with BsCspB is 82%, BcCspB with EcCspA is 54% and BsCpB with EcCspA is 57%.

of the structures was examined using the program PROCHECKNMR [42].

### Analysis of backbone torsion angles

We used the program cyclist [43] to calculate the mean of the backbone torsion angles in our ensemble of 21 structures. For each single structure, the backbone angles were calculated using the program AURELIA [38]. To compare the backbone torsion angles of TmCsp with the other known structures of Csps, the alignment of BsCspB, EcCspA, BcCspB and TmCsp displayed in Fig. 1 was used to adjust the positions of the amino acids. The gaps between the different amino-acid sequences were filled manually with x for the torsion angle values to keep them unprocessed at this position [44]. In the case of the NMR structure of EcCspA [11], for which an ensemble of 16 structures has been deposited in the Protein Data Bank (PDB), we calculated the mean of the backbone torsion angles of these.

### BioMagResBank and PDB accesion numbers

A full list of <sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>C chemical shifts has been deposited in the BioMagResBank under accession number 4895. The co-ordinates of *Tm*Csp have been deposited with the PDB under accession number 1G6P.

### RESULTS

### Assignment of the NMR resonance lines

Figure 2 shows the 2D  $^{15}$ N-<sup>1</sup>H HSQC spectrum of *Tm*Csp at pH 6.5 and 303 K. As shown by MS, recombinant *Tm*Csp expressed in *E. coli* contains 66 amino acids, with





Fig. 3. Summary of sequential and intermediate-range NOE connectivities, amide hydrogen life times,  ${}^{3}J_{\rm NH\alpha}$  coupling constants, and deviations of  ${}^{1}\!H^{\alpha}\!,\,{}^{13}\!C^{\alpha}$  and 13C1 chemical shifts from random-coil values. The NOE correlations were determined from 2D NOESY and 3D <sup>15</sup>N-resolved NOESY-HSQC spectra, all recorded at pH 6.5 and 303 K (30 °C). Sequential NOE intensities d(i, i + 1)between amino acid i and i + 1 are represented by the bar heights as strong, medium, or weak. For medium-range NOEs, lines denote the pairs of protons correlated. The life times  $\ensuremath{\mathsf{NH}_{\mathsf{ex}}}$  of slowly exchanging backbone amides are indicated by the height of the bars: high intensity, not exchanged after 17 h; medium intensity, exchanged after 6 h; weak intensity, exchanged after 30 min. Values of the conformation-dependent secondary shifts  $\Delta \delta^1 H^{\alpha}$ ,  $\Delta \delta^{13} C^{\alpha}$  and  $\Delta \delta^{13} C'$ are plotted with solid bars. Secondarystructure elements deduced from the NOESY data are indicated below the NOE data:  $\beta$  strands by arrows, an  $\alpha$ -helical turn by a circle, and unstructured regions by lines.

Solution NMR structure of Csp from T. maritima (Eur. J. Biochem. 268) 2531



methionine as N-terminal amino acid; this is also found in the natural protein [7]. All expected backbone amide crosspeaks for residues R2–E66 (except for P57 which contains only an imino group) are observed. As usual, the cross-peak of the N-terminal amino group is not visible because the protons exchange too rapidly with the bulk water. There are 58 well-resolved backbone <sup>15</sup>N-<sup>1</sup>H<sup>N</sup> signals and two totally overlapping backbone <sup>15</sup>N-<sup>1</sup>H<sup>N</sup> signals, which are assigned to residues K63/Q51 and F48/V46, and two partially overlapping backbone <sup>15</sup>N-<sup>1</sup>H<sup>N</sup> signals assigned to H28/V65 and E47/F8 in the 2D <sup>15</sup>N-<sup>1</sup>H<sup>N</sup> HSQC spectrum. The N<sup>e</sup>H resonance of the side chain of R2 is folded into the spectrum. The NMR samples contained a minor amount of Csp with the N-terminal methionine formylated, as shown directly by NMR spectroscopy and independently confirmed by MS. This leads to a couple of doubled signals in the NMR spectra which could be sequentially assigned (see e.g. the signal of the amide proton of formylated M1, labeled as form. M1 in Fig. 2). The sequential assignment was performed using a set of HNCA and CBCACONH



Fig. 4. Schematic diagram of the  $\beta$  sheet topology of *Tm*Csp. Arrows denote unambiguous interstrand NOEs. Dashed lines represent hydrogen bonds derived from the hydrogen exchange and NOE data. In the lower left corner, arrows indicate the direction and location of the five strands of antiparallel  $\beta$  sheet.



Fig. 5. Plot of the number of <sup>1</sup>H-<sup>1</sup>H NOE constraints as a function of the position of amino acid *i* in the sequence. Heavily shaded bar, intraresidual NOEs d(i,i); solid bar, sequential NOEs d(i,i + 1); open bar, intermediate-range NOEs  $d(i,i + n; 1 < n \le 5)$ ; lightly shaded bars, long-range NOEs d(i, i + n; n > 5).

experiments. The assigned primary structure was identified in the 2D spectra (2D <sup>1</sup>H-<sup>1</sup>H TOCSY, and 2D <sup>1</sup>H-<sup>1</sup>H NOESY) in <sup>1</sup>H<sub>2</sub>O and <sup>2</sup>H<sub>2</sub>O and 3D spectra (<sup>15</sup>N-<sup>1</sup>H TOCSY-HSQC and <sup>15</sup>N-<sup>1</sup>H NOESY-HSQC) in <sup>1</sup>H<sub>2</sub>O as well as in the HCCH-TOCSY spectra. All proton, nitrogen, carbon  $\alpha$ , carbon  $\beta$ , and carbonyl resonances could be assigned (see Materials and methods). The secondarystructure elements and the long-range NOEs extracted from the 2D and 3D NOESY are depicted in Fig. 3.

## Characterization of secondary-structure elements in *Tm*Csp

Secondary-structure elements of TmCsp were identified using a combination of data, including medium-range and longe-range interstrand NOEs,  ${}^{3}J_{\text{HN-H}\alpha}$  coupling constants, amide <sup>1</sup>H-<sup>2</sup>H exchange rates and deviations of backbone atom shifts from random-coil values derived from tetrapeptides [45]  $\Delta\delta H^{\alpha}$ ,  $\Delta\delta^{13}C^{\alpha}$  and  $\Delta\delta^{13}C'$  values, where  $\Delta \delta H^{\alpha} = \delta H^{\alpha}_{obs} - \delta H^{\alpha}_{rc} \text{ and } \Delta \delta^{13} C^{\alpha,\prime} = \delta^{13} C^{\alpha,\prime}_{obs} - \delta^{13} C^{\alpha,\prime}_{rc}.$ From these data (Fig. 3), five  $\beta$  strands corresponding to segments containing the residues R2-D9, G13-K19, D24-W29, G43-Q51, and Q58-V65 can be identified. Interstrand hydrogen bonds were determined on the basis of the interstrand backbone-backbone NOEs and the locations of slowly exchanging amide protons. From the interstrand NOEs the five  $\beta$  strands can be arranged into an antiparallel  $\beta$  pleated sheet (Fig. 4), in which  $\beta$  strands 1–4 form a Greek-key topology. Strands 1 and 5 contain  $\beta$  bulges at positions 6 and 61, respectively. In general, the amide <sup>1</sup>H-<sup>2</sup>H exchange rates for *Tm*Csp at pH 6.5 and 303 K were found to be rather low. In particular, the amide protons of residues G3-W7, Y14-T18, V25, V27, Q44, and V46-E49 did not exchange after 17 h at 303 K, suggesting high stability of the  $\beta$  sheet. Only the amide protons of the residues at the termini of the protein (R2, E66), near the bulge position in  $\beta$ 1 (F8) and  $\beta$ 5 (H61) and in the loop regions S10-K12, D20-D24, S30, A31, E33-L40, E42, G43, I50, and E52–P57, exchanged so rapidly that they were not detectable in the HSQC spectrum after the sample had been dissolved in  $D_2O$  (that is after 20 min).

### 3D structure of TmCsp

The 3D structure of *Tm*Csp was calculated with the molecular dynamics program XPLOR as described in Materials and methods. The constraints used for the simulated annealing approach and the structural statistics are summarized in Table 1. Figure 5 shows the distribution of the NOE-derived distance restraints along the amino-acid sequence of the protein with an average number of 13.9 distance constraints per residue.

The overall solution structure of TmCsp is well defined by our NMR data as shown in Fig. 6A. It depicts superposition of the backbone atoms of a family of the 21 lowestenergy conformers of TmCsp generated with x-plor. Large structural differences are only found for a part of the long surface loop between  $\beta$  strands 3 and 4, ranging from E35 to T39. As Fig. 5 shows, only a relatively small number of NOEs were observed in this part of the structure, indicating that this part of the structure may be mobile or disordered in solution. The structures calculated by X-PLOR had the tendency to cluster into two groups with different conformations for loop L<sub>3</sub>; however, it was not possible to verify unambiguously the existence of two local conformational states on the basis of NOE patterns. As the use of a closely related molecular dynamics program (CNS) with the same restraint files as used for the X-PLOR calculations could not reproduce two well-separated conformational states, at present the occurrence of the two states can be considered consistent with the experimental data but not proven. Backbone rmsds are 0.094 nm for all residues to the mean structure (Table 1). The ordered regions within the  $\beta$  sheets are much better defined and have a backbone rmsd of 0.025 nm to the mean structure. Apart from residues in poorly defined regions of the protein, all  $\phi$  and  $\psi$  values occur in low-energy regions of the Ramachandran plot. Figure 6B shows a representative conformer of the family



Fig. 6. Stereoviews of TmCsp. (A) Stereoview of a superposition of the backbone atoms (N, C<sup> $\alpha$ </sup>, and C') of the 21 accepted structures of TmCsp. The corresponding rmsd value is 0.094 nm. (B) Stereoview of a representative ribbon diagram of TmCsp. The secondary-structure elements identified by the program MOLMOL are indicated. (C) Stereoview of the solution NMR structure of TmCsp (ribbon diagram, side view rotated). Location of a possible peripheral ion cluster. The side-chain atoms of the basic residues K63 and R2 are shown in blue, and the side-chain atoms of the acidic residues E47 and D20 are shown in red.

of *Tm*Csp structures indicating the secondary-structure elements identified by the program MOLMOL [46].

A characteristic and useful parameter for the description and characterization of protein structures determined by X-ray crystallography or NMR spectroscopy is the backbone torsion angle. We used the program cyclist [43] to visualize the differences between the known structures of Csps and our ensemble of 21 structures. Figure 7A shows a comparison of TmCsp with the NMR-derived structure of EcCspA [11], Fig. 7B with the NMR-derived structure of BsCspB [13], Fig. 7C with the crystal structure of EcCspA [9], Fig. 7D with the crystal structure of BsCspB [12], and Fig. 7E with the crystal structure of BcCspB [14]. Generally, the backbone torsion angles in the  $\beta$  strand regions of our TmCsp structure are similar to those observed in the  $\beta$  strand region of the other Csp structures. In the loops, especially the long loop L<sub>3</sub>, there is great variability between the backbone torsion angles of all the Csp structures. The angular standard deviation of backbone dihedral angle  $\phi$  of L40 is rather large in *Tm*Csp, indicating that the value of this angle is a main determinant for the possible existence of two states of L<sub>3</sub>.

### DISCUSSION

### Solution structure of TmCsp

From the NMR structure determination we conclude that *Tm*Csp forms a closed  $\beta$  barrel consisting of five  $\beta$  strands:  $\beta_1$ , R2–D9;  $\beta_2$ , G13–K19;  $\beta_3$ , D24–W29;  $\beta_4$ , G43–Q51;



Fig. 7. Comparison of the backbone torsion angles  $\phi$  with those in different Csps. The mean values of the  $\phi$  angles of the *Tm*Csp (21 conformers) are represented as bars; the standard error is indicated. The backbone torsion angles of the NMR structures of (A) *Ec*CspA (16 conformers [11]) and (B) *Bs*CspB [13] and the crystal structures of (C) *Ec*CspA [9], (D) *Bs*CspB [9], and (E) *Bc*CspB [14] are indicated by black diamonds. The numbers indicate the position of amino acids in the *Tm*Csp sequence. Secondary-structure elements deduced from the NOESY data are indicated below:  $\beta$  strands by arrows, an  $\alpha$ -helical turn by a circle, and unstructured regions by lines.

 $\beta_5$ , Q58–V65. The first four,  $\beta_1$ – $\beta_4$ , show a Greek-key fold.  $\beta_1$  and  $\beta_5$  contain  $\beta$  bulges including the residues V5-K6-W7-F8 in  $\beta_1$  and A59-A60-H61-V62 in  $\beta_5$ . They break  $\beta_1$  into  $\beta_1'$  and  $\beta_1''$ , and  $\beta_5$  into  $\beta_5'$  and  $\beta_5''$ . As a result, two  $\beta$  sheet surfaces are formed,  $\beta$  sheet 1 by strands  $\beta_1'' - \beta_2 - \beta_3 - \beta_5'$ , and  $\beta$  sheet 2 by strands  $\beta_1' - \beta_4 -$  $\beta_5^{\prime\prime}$  (in *Ec*CspA similar sheets are defined by strands  $\beta_1^{\prime\prime}$ - $\beta_2 - \beta_3$  and  $\beta_1' - \beta_4 - \beta_5' - \beta_5''$  [11]). The core region of TmCsp showed very slowly exchanging backbone amide protons; the amide protons of residues G3-W7, Y14-T18, V25, V27, Q44, and V46-E49 did not exchange at all in 17 h at 303 K. This indicates that these amide groups are involved in strong hydrogen bonds well shielded against the bulk water. In line with this observation, the structurally important hydrophobic residues V5, I17, V25, V27, V46, and F48 are located in these regions. Their side chains point

to the interior of the protein and form a hydrophobic core together with I50 and V62. The  $\beta$  barrel is stabilized by the hydrophobic interactions between the two antiparallel  $\beta$  sheets and the hydrogen-bond network pattern within the  $\beta$  sheets.  $\beta$  strands 1 and 2 as well as 2 and 3 are connected by very short (three and four residues in length) surface loops L<sub>1</sub> and L<sub>2</sub>, respectively (L<sub>1</sub>, S10–K12; L<sub>2</sub>, D20–G23). Both loops form very tight turns.  $\beta$  strands 3 and 4 are connected by a very long surface loop (L<sub>3</sub>) 13 residues in length (S30–E42), which contains at the beginning a very short  $\alpha$ -helical turn at positions S30, A31, I32 and E33. Loop L<sub>4</sub> between  $\beta$  strands 4 and 5 is also located on the surface of the protein and encompasses six residues (E52–P57).

One important structural feature of *Tm*Csp is the large number of charged residues; this implies a relatively high



Fig. 7. Continued.

probability of ion-pair formation, which could contribute to the high stability of the protein [47]. A possible peripheral ion cluster of positively and negatively charged residues includes the side chain of R2: D20-R2-E47-K63 (Fig. 6C). In all our structures, these side chains are spatially very close to each other (i.e. the distance from the nitrogen atom of one guanidinium group of R2 to the  $\gamma$ -carbon atom of D20 is 6.15  $\pm$  0.09 Å, the distance from the nitrogen atom of the other guanidinium group of R2 to the  $\delta$ -carbon atom of E47 is 6.37  $\pm$  0.13 Å, and the distance from the  $\delta$ -carbon atom of E47 to the  $\zeta$ -nitrogen atom of K63 is 5.46  $\pm$  1.46 Å), which allows the formation of a peripheral ion cluster. The arginine side chain plays an important role in the geometry of this ion cluster and is located in the center. It is only conserved in thermophilic or hyperthermophilic members of the Csp family: BstCspB, BcCspB, and TmCsp. In BcCspB, the arginine corresponding to R2 of TmCsp is at position 3 in the amino-acid sequence. R3 was shown to be very close to the glutamate side chain of E46 in BcCspB, indicating a strong localized ionic interaction between the two side chains [14].

In a site-directed mutagenesis study, Perl et al. [48] showed that, in BcCspB, only the two side chains of R3 and L66 were important for thermostability. Changing the two corresponding positions in the mesophilic protein BsCspB (E66L and E3R) resulted in a thermostable protein similar to BcCspB, and inverting these two residues in BcCspB (L66E and R3E) to the BsCspB sequence decreased the thermostability to the value for the mesophilic protein BsCspB [48]. In TmCsp, a valine residue (V65) is found at the corresponding position of L66 in BcCspB, which forms a hydrophobic patch with V45, a hydrophobic residue unique to TmCsp and Csp from Streptomyces clavuligerus. As seen in BcCspB, the thermostability is based on the arginine R2 and the valine V65. In addition, the side chains of E49 and H61, as well as K6 and D24, which are also close together (i.e. 5.47  $\pm$  2.47 Å for the  $\delta$ -carbon atom of E49 to the  $\varepsilon_2$ -nitrogen atom of H61, and 3.98  $\pm$  0.19 Å for the  $\zeta$ -nitrogen atom of K6 to the  $\gamma$ -carbon atom of D24), could form single peripheral ion pairs. The mutagenesis of H61N had no effect on protein stability in TmCsp [49]. Therefore, H61 does not contribute to a stable ion cluster which is displayed in the high standard deviation of 50% for the distance between H61 and E49 in our family of 21 structures. In contrast, the mutagenesis of D9 to an asparagine residue showed a considerable destabilizing effect [49], which is due to the fact that D9 is involved in the  $\beta$  sheet 2 formed by  $\beta$  strands  $\beta_1^{\prime\prime} - \beta_2 - \beta_3 - \beta_5^{\prime}$ .

Csps are thought to bind to nucleic acids. The corresponding binding regions include the RNA-binding motifs RNP1 and RNP2 commonly found in ssRNA-binding proteins [15,16]. As taken from the NMR structure, in *Tm*Csp the side chains of the residues involved in the RNP1 (G13-Y14-G15-F16-I17) and RNP2 motif (V25-F26-V27-H28-W29) are located on the same side of the surface of the protein, supporting the hypothetical complex formation. Chemical-shift perturbation analysis of complexes between CspA from *E. coli* and ssDNA [9,10] showed that residues K10, W11, K16, F18, F20, F31 H33, F34, and K60 were involved in nucleic acid binding in CspA. In *Tm*Csp, these residues correspond to K6, W7, K12, Y14, F16, F26, H28, W29, and K55. Of these, again K12 and K55 are located in loop regions on the surface of the protein; their interactions may be complemented by two further nearby lysine residues (K11 and K54).

Another property typical of Csps is the presence of a considerable number of solvent-exposed aromatic side chains. In TmCsp, these are W7, Y14, F16, F26, W29, and F37, five forming a cluster on one side of the protein surface (W7, Y14, F16, F26, and W29). Mutational analysis of solvent-exposed aromatic side chains in BsCspB showed that the mutation F38A has no effect on protein stability whereas F15A, F17A, and F27A decreased the stability dramatically [50]. In TmCsp, the corresponding residues are F37, Y14, F16, and F26. In our NMR-derived structure, residues Y14, F16, and F26 form a hydrophobic patch within the presumed nucleic acid-binding surface area, and are likely to be important for both the function and stability of the protein. F37 is located in a different region and does not seem to be essential for either function or stability. In all, the results derived from the mutagenesis in BsCspB [50] are in good agreement with the NMR-derived 3D structure of TmCsp. This is in line with the extreme phylogenetic conservation of the Csp family [6].

# Comparison of the *Tm*Csp structure with the NMR-derived structures of *Ec*CspA and *Bs*CspB and the crystal structure of *Bc*CspB

A comparison of the surface charge distribution between the NMR-derived structures of EcCspA [10,11] and BsCspB [13] and TmCsp is shown in Fig. 8. TmCsp, the hyperthermophilic protein, shows a larger number of positive charges on the surface compared with its mesophilic counterparts. The high temperature leads to a decrease in the affinity for nucleic acids; this effect can be compensated for by a higher number of charges involved in the interaction. In EcCspA, the hydrophobic core is formed by the eight residues V9, I21, V30, V32, V51, F53, I55, and V67 [11]. In TmCsp, the corresponding residues (V5, I17, V25, V27, V46, F48, I50, and V62) are also present. One important feature of the EcCspA, BsCspB, BcCspB and TmCsp structures is the large number of solvent-exposed aromatic residues. As shown in Fig. 1, these aromatic residues are highly conserved in the family of Csps. In the case of TmCsp, the RNP1 and RNP2 RNA-binding motifs, K–G–Y–G–F–I and V–F–V–H–W, in  $\beta$  strands 2 and 3 differ slightly from those found in EcCspA and BsCspB: K-G-F-G-F-I (RNP1) and V-F-V-H-F (RNP2). It is worth mentioning that the tyrosine residue is only found in RNP1 from thermophilic micro-organisms and that tryptophan is only found in RNP2 from T. maritima. All these residues (including W7 and W29) are exposed as potential candidates for specific interactions with single stranded nucleic acids. As mentioned above, TmCsp contains two more basic residues in this region. The differences in the binding surface between the hyperthermophilic and mesophilic proteins may be essential for the specific regulatory functions of the proteins in their different natural environments.

A recently published crystal structure of the thermophilic member of the Csp family, *Bc*CspB, showed strong localized ionic interaction between  $N^{\eta 1}$  of the arginine side chain R3 and the  $O^{\epsilon 1}$  of the glutamate side chain E46 [14]. The corresponding residues in *Tm*Csp are R2 and V45 (Fig. 1). Our NMR-derived structure shows the possibility

### mesophilic



### hyperthermophilic

**TmCsp** 

Fig. 8. Comparison of the surface charge distribution of the solution NMR structures of *EcCspA* [10,11] and *BsCspB* [13] with the surface charge distribution of the solution NMR structure of *TmCsp*. Basic amino acids are shown in blue, acidic amino acids in red, and hydrophobic amino acids in white. Surface charges were calculated in MOLMOL [46].

of a peripheral ion cluster involving residues D20–R2– E47–K63. The arginine side chain exclusively found in thermophilic and hyperthermophilic Csps seems to play a major role in the thermostability of these proteins. In *Bc*CspB, the arginine residue located in strand  $\beta$ 1 seems to interact with the side chain of E46, the first residue located in strand  $\beta$ 4 in *Bc*CspB [14]. In contrast, the arginine side chain located in strand  $\beta$ 1 in *Tm*Csp seems to interact with the glutamate E47 located in the center of strand  $\beta$ 4 in *Tm*Csp. This may allow the involvement of more charged residues in localized ionic interactions, forming a whole peripheral ion cluster centered around the arginine side chain (D20–R2–E47–K63).

### Conclusions

Given the close similarity of the known 3D structures of Csps, what is the mechanism by which TmCsp gains its

higher stability compared with its mesophilic and moderately thermophilic counterparts? The multiple alignment of 30 prokaryotic Csp sequences (Fig. 1) with TmCsp allows the following conclusions. Regarding aromatic residues, two phenylalanine residues, which are highly conserved in other Csps, are replaced by tyrosine and tryptophan. The first is unique in Csps from thermophiles, while the latter is found only in *Tm*Csp. As we pointed out, they are located in the single stranded nucleic acid-binding site and thus may account for the specific regulatory roles of the proteins. In general, the number of aromatic residues is high: five Phe, one Tyr, two Trp. At neutral pH, the number of charged residues is increased to more than one-third of the total number of amino acids; 24 in TmCsp compared with 19, 17, and 16 in BsCspB, BcCspB, and EcCspA, respectively. TmCsp contains the maximum Lys content; only four of its 11 Lys residues are conserved. The single Arg is only present in BcCspB, BstCspB, and TmCsp; this is apparently a characteristic of thermophiles. In BcCspB, the single Arg3 was shown to confer the thermostability [48] by interacting with Glu46 [14]. In TmCsp, this Arg seems to participate not only in a single ion pair but in a peripheral ion cluster together with Asp20, Glu47 and Lys63. With respect to hydrophobic residues, Val45 is unique in TmCsp and Csp from S. clavuligerus; in most other sequences, polar residues prevail at this position. In TmCsp, it forms a hydrophobic patch with Val65, where the corresponding side chain in BcCspB (Leu66) has been shown to be important for thermostability [48].

Bearing in mind that the stability of proteins is a cumulative effect of small increments [51], we can conclude that, in the case of TmCsp, a single peripheral ion cluster around the side chain of R2, as well as increased hydrophobic stacking of side chains at the surface area, are used to increase the thermal stability of TmCsp.

### ACKNOWLEDGEMENTS

We are grateful to the Deutsche Forschungsgemeinschaft (Ja 78/34) and the European Union (Biomed Program) for supporting this work. M. G. acknowledges support by the Peter and Traudl Engelhorn Stiftung. We thank E. Hochmuth and D. Deutzmann for the mass spectroscopy.

### REFERENCES

- Shaw, M.K. & Ingraham, J.L. (1967) Synthesis of macromolecules by *E. coli* near the minimal temperature for growth. *J. Bacteriol.* 94, 157–164.
- Lopez, M.M., Yutani, K. & Makhatadze, G.I. (1999) Interactions of the major cold shock protein CspB with single-stranded DNA templates of different base composition. *J. Biol. Chem.* 274, 33601–33608.
- Jones, P.G. & Inouye, M. (1994) The cold shock response: a hot topic. *Mol. Microbiol.* 11, 811–818.
- Graumann, P. & Mahariel, M.A. (1996) Some like it cold: response of microorganisms to cold shock. *Arch. Microbiol.* 166, 293–300.
- Thieringer, H.A., Jones, P.G. & Inouye, M. (1998) Cold shock and adaptation. *Bioessays* 20, 49–57.
- Graumann, P.L. & Marahiel, M.A. (1998) A superfamily of proteins that contain the cold shock domain. *Trends Biochem. Sci.* 23, 286–290.
- Welker, C., Böhm, G., Schurig, H. & Jaenicke, R. (1999) Cloning, overexpression, purification, and physicochemical characterization

of a cold shock protein homolog from the hyperthermophilic bacterium *Thermotoga maritima*. *Protein Sci.* **8**, 394–403.

- Murzin, A.G. (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* 12, 861–867.
- Schindelin, H., Jiang, W., Inouye, M. & Heinemann, U. (1994) Crystal structure of CspA, the major cold shock protein of *E. coli*. *Proc. Natl Acad. Sci. USA* 91, 5119–5123.
- Newkirk, K., Feng, W., Jiang, W., Tejero, R., Emerson, S.D., Inouye, M. & Montelione, G.T. (1994) Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. *Proc. Natl Acad. Sci. USA* 91, 5114–5118.
- Feng, W., Tejero, R., Zimmerman, D.E., Inouye, M. & Montelione, G.T. (1998) Solution NMR structure and Backbone dynamics of the major cold-shock protein (CspA) from *Escherichia coli*: evidence for conformational dynamics in the single-stranded RNA-binding site. *Biochemistry* 37, 10881–10896.
- Schindelin, H., Mahariel, M.A. & Heinemann, U. (1993) Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold shock protein. *Nature (London)* 364, 164–168.
- Schnuchel, A., Wiltschek, R., Czisch, M., Herrler, M., Willimsky, G., Graumann, P., Mahariel, M.A. & Holak, T.A. (1993) Structure in solution of the major cold shock protein from *B. subtilis. Nature* (*London*) 364, 169–171.
- Mueller, U., Perl, D., Schmid, F.X. & Heinemann, U. (2000) Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein. *J. Mol. Biol.* 297, 975–988.
- Bandziluis, R.J., Swanson, M.S. & Dreyfuss, G. (1989) RNA-binding proteins as developmental regulators. *Genes Dev.* 3, 431–437.
- Burd, C.G. & Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* 265, 615–621.
- La Teana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C.L. & Gualerzi, C.O. (1991) Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. *Proc. Natl Acad. Sci. USA* 88, 10907–10911.
- Jones, P.G., Krah, R., Tafuri, S.A. & Wolffe, A.P. (1992) DNA gyrase, CS7.4, and the cold shock response in *Escherichia coli*. *J. Bacteriol.* 174, 5798–5802.
- Jiang, W., Fang, L. & Inouye, M. (1996) The role of the 5'-end untranslated region of the mRNA for CspA, the major cold shock protein of *E. coli*, in cold shock adaptation. *J. Bacteriol.* 178, 4919–4925.
- Jiang, W., Hou, Y. & Inouye, M. (1997) CspA, the major coldshock protein of *Escherichia coli*, is an RNA chaperone. *J. Biol. Chem.* 272, 196–202.
- Hillier, B.J., Rodriguez, H.M. & Gregoret, L.M. (1998) Coupling protein stability and protein function in *Escherichia coli* CspA. *Folding Des.* 3, 87–93.
- 22. Schröder, K., Graumann, P., Schnuchel, A., Holak, T.A. & Mahariel, M.A. (1995) Mutational analysis of the putative nucleic acid-binding surface of the cold-shock domain, CspB, revealed an essential role of aromatic and basic residues in binding of singlestranded DNA containing the Y-box motif. *Mol. Microbiol.* 16, 699–708.
- Yamanaka, K., Fang, L. & Inouye, M. (1998) The CspA family in Escherichia coli: multiple gene duplication for stress adaptation. Mol. Microbiol. 27, 247–256.
- Phadtare, S., Alsina, J. & Inouye, M. (1999) Cold-shock response and cold-shock proteins. *Curr. Opin. Microbiol.* 2, 175–180.
- Chatterjee, S., Jiang, W., Emerson, S.D. & Inouye, M. (1993) The backbone structure of the major cold-shock protein CS4.7 of

*E. coli* in solution includes extensive  $\beta$ -sheet structure. *J. Biochem.* (*Tokyo*) **114**, 663–669.

- Makhatadze, G.I. & Mahariel, M.A. (1994) Effect of pH and phosphate ions on self-association properties of the major coldshock protein from *Bacillus subtilis*. *Protein Sci.* 3, 2144–2147.
- Perl, D., Welker, C., Schindler, T., Schröder, K., Mahariel, M.A., Jaenicke, R. & Schmid, F.X. (1998) Conservation of rapid twostate folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol.* 5, 229–235.
- Reid, K.L., Rodriguez, H.M., Hillier, B.J. & Gregoret, L. (1998) Stability and folding of a model β-sheet protein, *E. coli* CspA. *Protein Sci.* 7, 470–479.
- Jeener, J., Meier, B.H., Bachmann, P. & Ernst, R.R. (1979) Investigation of exchange processes by 2D NMR spectroscopy. *J. Chem. Phys.* 71, 4546–4553.
- Braunschweiler, L. & Ernst, R.R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson.* 53, 521–528.
- Levitt, M.-H., Freeman. R. & Frenkiel, T. (1982) Broad-band heteronuclear decoupling. J. Magn. Reson. 47, 328–330.
- Marion, D. & Wüthrich, K. (1983) Application of phase-sensitive 2D COSY spectroscopy for measurements of <sup>1</sup>H–<sup>1</sup>H spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* 113, 967–974.
- Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S.J., Sorensen, O.W. & Griesinger, C. (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR* 4, 301–306.
- Görler, A. & Kalbitzer, H.R. (1997) Relax, a flexible program for the back calculation of NOESY spectra based on completerelaxation-matrix formalism. J. Magn. Reson. 124, 177–188.
- 35. Görler, A., Gronwald, W., Neidig, K.-P. & Kalbitzer, H.R. (1999) Computer assisted assignment of <sup>13</sup>C or <sup>15</sup>N edited 3D-NOESY-HSQC spectra using back calculated and experimental spectra. *J. Magn. Reson.* 137, 39–45.
- 36. Live, D.H., Davis, D.G., Agosta, W.C. & Cowburn, D. (1984) Longe range hydrogen bond mediated effects in peptides: <sup>15</sup>N NMR study of gramicidin S in water and organic solvents. *J. Am. Chem. Soc.* **106**, 1939–1941.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. & Sykes, B.D. (1995b) <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* 6, 135–140.
- Neidig, K.-P., Geyer, M., Görler, A., Antz, C., Saffrich, R., Beneicke, W. & Kalbitzer, H.R. (1995) AURELIA, a program for computer-aided analysis of multidimensional NMR spectra. *J. Biomol. NMR* 6, 255–270.
- 39. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids.* John Wiley & Sons, New York.
- Brünger, A.T. (1992) X-PLOR Manual, Version 3.1: a System for X-Ray Crystallography and NMR. Yale University Press, New Haven, CT.
- Nilges, M., Clore, G.M. & Gronenborn, A.M. (1988) Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* 229, 317–324.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.
- Döker, R., Maurer, T., Kremer, W., Neidig, K.-P. & Kalbitzer, H.R. (1999) Determination of mean and standard deviation of dihedral angles. *Biochem. Biophys. Res. Commun.* 257, 348–350.
- Döker, R. (2000) Cyclist Manual. http://www.biologie.uni-regensburg.de/Biophysik/Kalbitzer/software/index\_engl.html.
- Wishart, D.S., Bigam, C.G., Arne, H., Hodges, R.S. & Sykes, B.D. (1995) <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N random coil NMR shifts of the common amino

acids. I. Investigation of nearest neighbor effects. J. Biomol. NMR 5, 67–81.

- Koradi, R., Billeter, M. & Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14, 51–55.
- Perutz, M.F. & Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature* (*London*) 255, 256–259.
- Perl, D., Mueller, U., Heinemann. U. & Schmid, F.X. (2000) Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat. Struct. Biol.* 7, 380–383.
- Frankenberg, N., Welker, C. & Jaenicke, R. (1999) Does the elimination of ion pairs affect the thermal stability of cold shock protein from the hyperthermophilic bacterium *Thermotoga maritima FEBS Lett.* **454**, 299–302.
- Schindler, T., Perl, D., Graumann, P., Sieber, V., Marahiel, M.A. & Schmid, F.X. (1998) Surface-exposed phenylalanine in the RNP1/ RNP2 motif stabilizes the cold-shock protein CspB from *Bacillus* subtilis. Proteins: Structure, Function, Genetics 30, 401–406.
- 51. Jaenicke, R. & Böhm, G. (1998) The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738–748.