Protein Science (1999), *8*:394–403. Cambridge University Press. Printed in the USA.

Cloning, overexpression, purification, and physicochemical characterization of a cold shock protein homolog from the hyperthermophilic bacterium *Thermotoga maritima*

CHRISTINE WELKER,[1] GERALD BÖHM,[2] HARTMUT SCHURIG,[1] AND RAINER JAENICKE[1]

[1]Institut für Biophysik und physikalische Biochemie, Universität Regensburg, D-93040 Regensburg, Germany
[2]Institut für Biotechnologie, Martin-Luther-Universität Halle-Wittenberg, D-06120 Halle, Saale, Germany

(RECEIVED April 17, 1998; ACCEPTED July 20, 1998)

Abstract

Thermotoga maritima (*Tm*) expresses a 7 kDa monomeric protein whose 18 N-terminal amino acids show 81% identity to N-terminal sequences of cold shock proteins (Csps) from *Bacillus caldolyticus* and *Bacillus stearothermophilus*. There were only trace amounts of the protein in *Thermotoga* cells grown at 80 °C. Therefore, to perform physicochemical experiments, the gene was cloned in *Escherichia coli*. A DNA probe was produced by PCR from genomic *Tm* DNA with degenerated primers developed from the known N-terminus of *Tm*Csp and the known C-terminus of CspB from *Bacillus subtilis*. Southern blot analysis of genomic *Tm* DNA allowed to produce a partial gene library, which was used as a template for PCRs with gene- and vector-specific primers to identify the complete DNA sequence. As reported for other *csp* genes, the 5′ untranslated region of the mRNA was anomalously long; it contained the putative Shine–Dalgarno sequence. The coding part of the gene contained 198 bp, i.e., 66 amino acids. The sequence showed 61% identity to CspB from *B. caldolyticus* and high similarity to all other known Csps. Computer-based homology modeling allowed the conclusion that *Tm*Csp represents a β-barrel similar to CspB from *B. subtilis* and CspA from *E. coli*.

As indicated by spectroscopic analysis, analytical gel permeation chromatography, and mass spectrometry, overexpression of the recombinant protein yielded authentic *Tm*Csp with a molecular weight of 7,474 Da. This was in agreement with the results of analytical ultracentrifugation confirming the monomeric state of the protein. The temperature-induced equilibrium transition at 87 °C exceeds the maximum growth temperature of *Tm* and represents the maximal T_m-value reported for Csps so far.

Keywords: cold shock protein; hyperthermophiles; molecular modeling; protein stability; ribosomal protein L31; *Thermotoga maritima*

The cold shock response in microorganisms is a transient phenomenon affecting the growth rate of the cell and the saturation of fatty acids as well as the rates of DNA, RNA, and protein synthesis at temperatures significantly lower than the normal physiological temperature (Shaw & Ingraham, 1967). For most proteins, the expression under cold shock is decreased drastically, while for a few it is increased. Out of these, a group of small acidic proteins has met special interest because of their extremely high induction levels. They share high affinity to single-stranded nucleic acids and show extreme similarity with respect to their amino acid sequences. Therefore, they have been put together in the family of cold shock proteins (Csps) (for reviews, see Jones & Inouye, 1994; Graumann & Marahiel, 1996a; Thieringer et al., 1998). In the case of *Bacillus subtilis* (*Bs*), CspB has been shown to improve the viability at freezing temperature (Willimsky et al., 1992). Not all members of the Csp family are necessarily involved in the cold shock response. Some are also essential for growth under normal temperature conditions, others not cold-inducible at all (Lee et al., 1994; Yamanaka et al., 1994; Graumann et al., 1997; Yamanaka & Inouye, 1997). Apart from Csps, ribosomal binding factors were found to be induced to the same extent; they did not show any sequence similarity with the above-mentioned Csps (Jones et al., 1996; Jones & Inouye, 1996).

Csps are small globular proteins consisting of 65 to 70 amino acids corresponding to molecular masses of about 7.5 kDa. Struc-

Reprint requests to: Rainer Jaenicke, Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstraße 31, D-93040 Regensburg, Germany; e-mail: rainer.jaenicke@biologie.uni-regensburg.de.

Abbreviations: *Bc*, *Bacillus cadolyticus*; *Bs*, *Bacillus subtilis*; *Bst*, *Bacillus stearothermophilus*; Csp, cold shock protein; ORF, open reading frame; PCR, polymer chain reaction; RBS, robosome binding site; RP-HPLC, reversed-phase HPLC; *Tm*, *Thermotoga maritima*.

tures of the major Csps of *Escherichia coli* (*Ec*) and *Bs*, CspA and CspB, were solved by X-ray analysis and NMR (Schindelin et al., 1993, 1994; Schnuchel et al., 1993; Newkirk et al., 1994). Their three-dimensional structures were closely similar, with five β-strands combined in two antiparallel β-sheets, making up a β-barrel. Their native structure was not influenced by posttranslational modifications or cofactor binding. *Bs*CspB was found to undergo reversible folding/unfolding, obeying the two-state mechanism, without slow kinetic phases attributable to disulfide formation or cis-trans-prolyl isomerization (Schindler et al., 1995; Schindler & Schmid, 1996).

Csp homologs have been identified in a wide variety of microorganisms from psychrophiles to thermophiles. In eukaryotes, a group of proteins contained the so-called Y-box-domain, which revealed 43% identity to CspA (Wistow, 1990; Wolffe et al., 1992). Thus, from the evolutionary point of view, Csps and the respective domains belong to the most conserved proteins presently known (Graumann & Marahiel, 1996b). For this reason, they were used as models for comparative folding and thermodynamic studies, especially in connection with mechanisms of thermal stabilization. In the present study, the first Csp-homolog from a hyperthermophile, *Thermotoga maritima* (*Tm*), is described. The gene was cloned and the recombinant protein expressed in *Ec*. The physicochemical characterization of the protein showed that *Tm*Csp represents the most thermostable Csp reported so far.

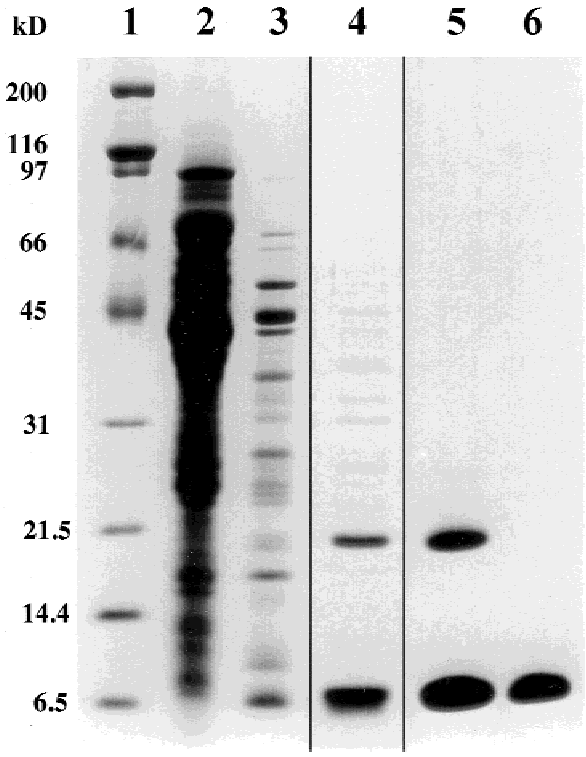

Fig. 1. Purification of natural *Tm*Csp: Coomassie-stained SDS-PAGE on a 10 to 22% acrylamide gradient gel. Lane 1, molecular weight marker; lane 2, *Tm* crude extract; lane 3, flow-through of Q-sepharose at pH 9.3; lane 4, eluate from Heparin-Sepharose; lane 5, eluate from SP-Sepharose at pH 6.0; lane 6, pure protein after gel filtration on Superdex 75 pg.

Results

Purification of the original Csp from Tm

*Tm*Csp was discovered as a low-molecular weight protein copurifying with a 25 kDa ribonuclease (Welker, 1996). Making use of the similar charge properties, the *Tm* crude extract was applied to an anion-exchange column at pH 9.3 to collect a flow-through containing basic proteins. Applying Heparin-Sepharose as a DNA/RNA analog, the *Tm*Csp eluted at an NaCl concentration between 160 and 210 mM, allowing the separation from the RNase-active fractions. It was purified to homogeneity by cation-exchange chromatography at pH 6.0 (elution at 280 to 330 mM NaCl) and gel permeation chromatography. Calibration of the Superdex 75 pg column with standard proteins revealed a molecular mass of ~7 kDa for the native protein. The efficiency of the purification protocol and the homogeneity of the protein were proven by SDS-PAGE (Fig. 1), spectroscopy, reversed-phase HPLC (RP-HPLC), and mass spectrometry (see below).

The sequence of the 18 amino-terminal amino acids, MRGKVK WFDSKKGYGFIT, was determined by Edman degradation. Database search with FASTA showed 81% identity with the N-termini of Csps from *Bacillus caldolyticus* (*Bc*) and *Bacillus stearothermophilus* (*Bst*).

The yield of purified Csp, starting from 28.5 g (wet weight) cells, was ~200 μg.

Cloning and sequencing

To obtain the protein in sufficient amounts for physicochemical measurements, it was cloned and expressed recombinantly. Degenerated oligonucleotides were developed to amplify part of the *csp* gene from genomic *Tm* DNA as well as from a complete gene library in pUN121. Primer N1 was chosen to be complementary to the N-terminal amino acid sequence of the original *Tm*Csp, while primer C1 was designed according to residues 53 to 63 of the amino acid sequence of *Bs*CspB (Table 1). Three PCR products, amplified independently from each other, were sequenced in both directions; they were identical, consisting of 185 bp, 127 of them coinciding with the natural *Tm* DNA sequence (see Fig. 2, underlined part of the coding region).

The amplified fragment was radiolabeled and used as a probe for southern blotting and colony hybridization. Southern blot hybridizations of this oligonucleotide to different restriction digests of genomic *Tm* DNA resulted in only one hybridization signal in each case (data not shown). Obviously, only one copy of the *csp*

Table 1. *Oligonucleotides used for cloning of TmCsp*[a]

N1	ATG (AC)G(GT) GG(GT) AAG GT(GT) AAG TGG TT(CT) GA
N2	CGG AAT TCA TAT GAG AGG AAA GGT TAA GTG G
C1	ACG TT(GT) GC(GT) GCT TG(GT) GG(GT) CC(GT) CTG TT(GT) CCT TC
C2	TAT GTG AGA CTA AAA ATC AAG TGG AGA AAG
C3	CGC CGG ATC CAA GCT TAT TAC TCA ACT ACT TTC ACG TGC GC
pUC1	CGA CGT TGT AAA ACG ACG GCC AGT
pUC2	CAC ACA GGA AAC AGC TAT GAC CAT

[a] N2 and C3 yield an *Nde*I and *Hind*III restriction site, respectively.

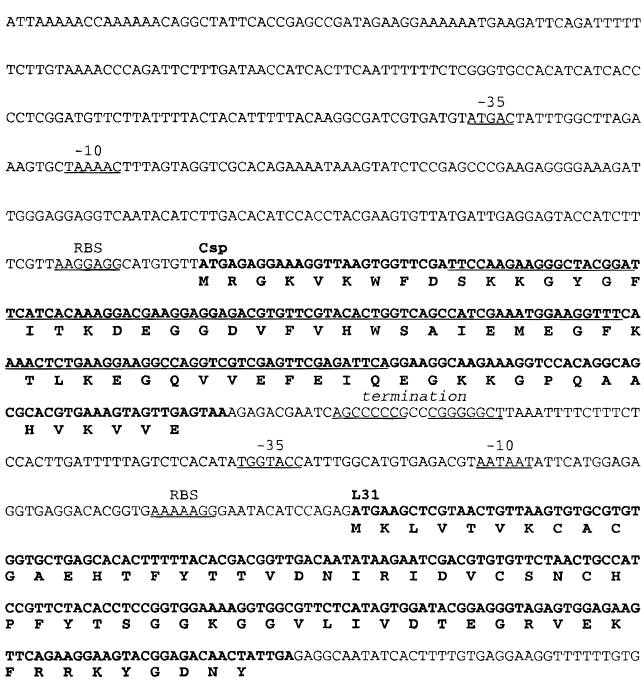

Fig. 2. DNA-sequence and deduced amino acid sequences of *Tm*Csp and ribosomal protein L31. The coding regions are marked by bold letters. The putative −35 and −10 regions as well as the putative ribosome binding sites (RBS) are underlined, the putative transcription termination site of *Tm csp* is indicated by reverse arrows. The underlined part of the coding region corresponds to the DNA probe used for southern blotting and colony hybridization (PCR primer sequences not included).

gene was present in the *Tm* genome. Similar to earlier reports for other *csp*s (Schröder et al., 1993), cloning the gene by colony hybridization failed, although about 20,000 *Ec* DH10b colonies were screened, 16,000 representing the complete gene library in pUN121 and 4,000 representing partial gene libraries in pUCBM20. Therefore, hybridizing fragments of *Eco*RV- and *Bam*HI-digested genomic *Tm* DNA were ligated in pUCBM20. The resulting plasmids were used as PCR templates. Primers were not only N1 and C1 but also pUC1 and pUC2, complementary to the regions of the plasmid next to the insert. PCR products with a vector-specific primer and N1 were expected to contain the 3′-sequence, those with C1, the 5′-sequence of *Tm csp*.

Indeed, pUC1/N1- and pUC2/N1-PCRs revealed the 3′-sequence of the *csp* gene as well as 375 additional basepairs (cf. Fig. 2). A translation termination codon (TAA) as well as a G/C-rich stem-loop structure serving as transcription termination signal was found. The plasmid also contained another open reading frame (ORF) coding for a 63-residues protein related to the 50S ribosomal protein L31. The ORFs of the *csp* and L31 genes were 145 bp apart. Within this distance, a putative ribosome binding site (RBS) (5′-AaAAAGG-3′), perfectly complementary to the 3′-end of *Tm* 16SrRNA (5′-UCUUUCCUCCA-3′), and a −10-region (aATAAT) could be clearly localized. The search for the −35-consensus sequence was less successful (TgGta or gTacC). The consensus sequences were TA(A/T)AAT for the −10, and TTGAC for the −35-region (Liao & Dennis, 1992). In this notation, uppercase and lowercase letters refer to identical and nonidentical bases, respectively. PCRs with the gene-specific primer C1 did not produce *csp*-containing fragments, independent of the $MgCl_2$ concentration and the annealing temperature. Therefore, a second C-terminal primer (C2), complementary to a sequence downstream of the *csp* gene, was designed. PCR with this gene-specific primer together with pUC2 yielded not only the complete coding sequence of *csp* but also 350 bp upstream of the initiation codon. At 15 bp upstream of this ATG codon, a putative ribosome binding site (5′-AAGGAGG-3′) was assigned. *Tm*-promoter-like motifs were localized 155 bp (5′-aTGAC-3′) and 130 bp (5′-TAAAAc-3′) upstream of the RBS.

The DNA and amino acid sequences of *Tm*Csp and ribosomal protein L31 are summarized in Figure 2. *Tm*Csp consists of 66 residues, resulting in a calculated molecular mass of 7,472.6 Da. This was in agreement with results from SDS-PAGE, gel permeation chromatography, mass spectrometry, and analytical ultracentrifugation. The alignment of the amino acid sequence with Csps from other organisms proved high sequence identity. The closest relationship holded for CspB from *Bc*: 61% identity/79% homology (data according to GeneDoc 2.2, Nicholas & Nicholas, 1997). Identities with Csps from mesophilic organisms were extremely high as well, e.g., *Bs*CspB: 61% identity/76% homology, and *Ec*CspA: 52% identity/71% homology. Sequences of eukaryotic cold shock domains exhibited differences attributable mainly to an insertion of four amino acid residues. However, identities/homologies were still significant, for example, in the case of the glycine-rich protein-2-precursor from *Nicotiana sylvestris*: 42%/56%. The given Csp sequences, including *Tm*Csp, had two highly conserved RNA binding motifs, RNP1 and RNP2, in common, suggesting a functional correlation of RNA-binding with the cold shock response.

All sequence data (including those for 50S ribosomal protein L31) are available under accession number Y11219 from the EMBL gene library.

Overexpression and purification of recombinant TmCsp

For overexpression of *Tm*Csp, the PCR product amplified from genomic *Tm* DNA with primers N2 and C3, containing an *Nde*I and a *Hind*III digestion site, was ligated into pET21a. The plasmid was sequenced and the coding region agreed with the results given in Figure 2.

Overexpression of *Tm*Csp in *Ec* BL21(DE3) was independent of the growth temperature and detectable even without induction. However, addition of 1 mM IPTG had a positive effect. For purification of the recombinant protein, cells were grown at 26 °C for 24 h with induction at OD 1. Since expression and vector stability tests showed that the plasmid was lost easily, the amount of ampicillin was increased to 500 μg/mL. In addition, the precultures were grown for at most 12 h at 37 °C and then washed twice with fresh medium to remove β-lactamase before inoculating the 2 L culture.

The purification followed basically the protocol that was used when the natural protein was first discovered. To remove bulk of the *Ec* proteins without significant coprecipitation of *Tm*Csp, the cell-free extract was diluted fivefold and then heated to 80 °C for 30 min. Pure *Tm*Csp was obtained by anion exchange at pH 9.3, cation exchange at pH 6.0, and gel permeation chromatography, leaving out the Heparine–Sepharose step used in the original Csp purification. The protein was 99% pure according to silver-stained SDS-PAGE, RP-HPLC, the ratio of the absorbance at 280 and 260 nm, and mass spectrometry (see below). N-terminal sequencing of the first 11 amino acids agreed with the determined DNA and protein sequences, including the initiator methionine. The total yield of *Tm*Csp was about 18 mg/L cell culture.

Physicochemical characterization and authenticity of recombinant Csp

Spectroscopic characteristics

The UV absorption spectra of recombinant and natural *Tm*Csp were essentially identical. They showed a maximum at 280 nm with a tryptophan shoulder at 288 nm and an A_{280}/A_{260} ratio of 1.6. The molar (specific) extinction coefficient for the native recombinant protein at 280 nm, determined according to Pace et al. (1995), was found to be 12,820 ± 70 M^{-1} cm^{-1} (1.72 ± 0.04 mL mg^{-1} cm^{-1}). Despite the fact that *Tm*Csp contains one additional tryptophan residue (Trp29) not present in other Csps, the change in absorbance in the course of GdmCl denaturation was insignificant.

The fluorescence emission spectra of both recombinant and natural *Tm*Csp were also indistinguishable (Fig. 3A). The maximum of the native protein at 341 nm, which was unaltered at λ_{exc} = 280 nm and λ_{exc} = 295 nm, showed only a slight red shift to 349 nm in the presence of 6 M GdmCl, which indicated that both the conserved Trp7 and the nonconserved Trp29 must be partially solvent-accessible in the native protein. The model structure confirmed this finding (see below). The 50% decrease in intensity upon unfolding was used to monitor the reversible denaturation of the protein.

The far-UV CD spectra of natural and recombinant *Tm*Csp again coincided, proving the secondary structure of both proteins to be identical (Fig. 3B). The maximum at 200 nm with a mean residue weight ellipticity of ~8,000 deg cm^2 $dmol^{-1}$ was indicative of a high degree of β-fold. At 210 to 250 nm, the signal was close to zero, with a shallow minimum and maximum at 219 and 228 nm, respectively. The missing negative contribution of α-helical structure to the dichroic absorption at 220 nm allowed positive signals of aromatic residues to take effect in this region (Perczel et al., 1992; Woody, 1994). The far-UV CD spectra in sodium and potassium phosphate as well as sodium cacodylate buffer at pH 7.0 were identical (data not shown). Figure 3B also depicts the spectral effects of GdmCl and temperature on *Tm*Csp. The spectra observed after heating to 98 °C clearly differed from those in 6 M GdmCl at 20 °C. At 98 °C, both the temperature- and GdmCl-dependent spectra coincided in the accessible wavelength range; obviously, the differences at 210–240 nm shown in Figure 3B are a temperature effect rather than reflecting a difference in conformation. The near-UV CD of the recombinant protein is shown in Figure 3C.

The fluorescence and far-UV CD spectra closely resembled those observed for BsCspB (Schindler et al., 1995) and EcCspA (Chatterjee et al., 1993; Reid et al., 1998).

Molecular mass and state of association

Mass spectrometry of natural and recombinant *Tm*Csp resulted in masses of 7,473.9 ± 0.4 Da and 7,475.1 ± 0.6 Da, respectively, in excellent agreement with the calculated molecular mass of the monomer (7,472.6 Da). Analytical gel permeation chromatography on Superdex 75 pg yielded a single peak at precisely the same retention time for both the natural and the recombinant protein. The calculated molecular mass of ~7,000 Da proved *Tm*Csp in dilute solution to be monomeric. Analytical ultracentrifugation confirmed this result for a wide concentration range and different buffer conditions: Sedimentation velocity experiments with recombinant *Tm*Csp at 0.2 to 13.2 mg/mL in both 50 mM sodium phosphate buffer pH 6.5, and 0.1 M sodium cacodylate buffer pH 7.0, yielded $s^0_{20,w}$ = 1.10 ± 0.09 S, high speed sedimentation equilibria at initial protein concentrations of 0.2 and 2.1 mg/mL, 7.3 ± 0.2 kDa, independent of the rotor speed and the buffer systems.

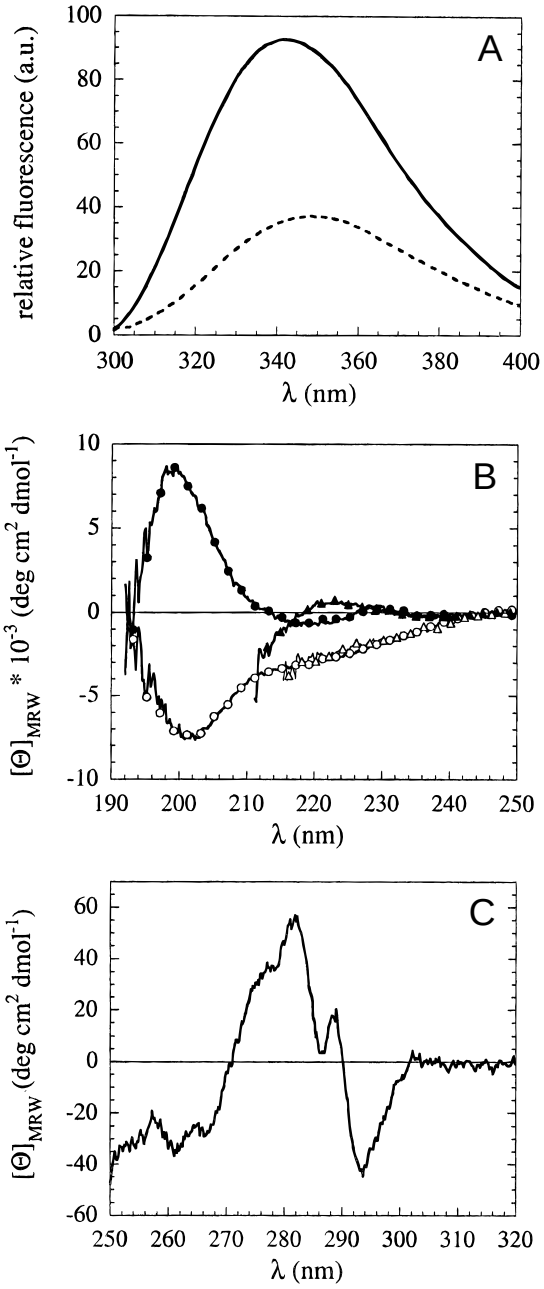

Fig. 3. Spectroscopic analysis of native and denatured *Tm*Csp. **A:** Fluorescence emission at 10 μg/mL Csp concentration, 1 cm pathlength, in 0.1 M sodium cacodylate pH 7.0 without (—) and with (- - -) 6 M GdmCl; excitation at 280 nm. **B:** Far-UV CD ($[\Theta]_{MRW}$ in deg cm^2 $dmol^{-1}$), 140 μg/mL Csp concentration, 0.1 cm pathlength, in 20 mM sodium phosphate pH 7.0, without (●) and with (▲) 6 M GdmCl at 20 °C and without (○) and with (△) 6 M GdmCl at 98 °C. **C:** Near-UV CD ($[\Theta]_{MRW}$ in deg cm^2 $dmol^{-1}$) at 335 μg/mL Csp concentration, 1 cm pathlength, in 20 mM sodium phosphate, pH 7.0.

Structure

The high homology of the sequences and the similarity of the fluorescence and far-UV CD spectra between *Tm*Csp and previ-

ously investigated Csps suggested the topologies of the proteins to be closely related. Therefore, the known X-ray structures of *Bs*CspB and *Ec*CspA, as well as the NMR structure of *Bs*CspB, were used to predict the three-dimensional structure of *Tm*Csp by homology modeling. Making use of the structure quality analysis in PROSAII, the residue Z-scores of the model based on the *Bs*CspB X-ray structure showed the highest reliability. Attempts to improve the model by molecular simulation protocols implemented in MODELLER were unsuccessful. Therefore, only minor adjustments regarding the stereochemical quality of the structure were performed. The three-dimensional model of *Tm*Csp did not deviate significantly from the reported experimental structures (Table 2). Compatibility included the putative nucleic-acid binding motif formed mainly by aromatic and positively charged residues clustering in exposed regions on one side of the protein; in the case of *Tm*Csp, this site is extended by two additional lysine residues. Compared to the known Csps from other organisms, the overall surface charge was increased drastically, reflecting the higher isoelectric point of *Tm*Csp. There are no obvious statistical differences in the number of salt bridges or hydrogen bonds (Table 3).

*Stability of recombinant Tm*Csp

The stability of recombinant *Tm*Csp was characterized by pH-, GdmCl-, and temperature-induced denaturation/renaturation, making use of fluorescence emission and far-UV CD. To avoid possible extrinsic effects of phosphate ions, measurements were carried out preferentially in 0.1 M sodium cacodylate buffer pH 7.0.

The pH-induced denaturation showed a bell-shaped profile with unaltered fluorescence emission between pH 3.5 and 8.5 (data not shown). GdmCl-induced equilibrium transitions at different temperatures are depicted in Figure 4; the corresponding results from nonlinear least-squares fits, based on the two-state model, are summarized in Table 4. All transitions were fully reversible with low cooperativity, as expected for a small protein. The cooperativity paremeter, $m = d\Delta G/dc_{\mathrm{GdmCl}}$, was unchanged over the measured temperature range. The half-concentrations, $(c_{\mathrm{GdmCl}})_{1/2}$, and free energies of stabilization at 0 M GdmCl, ΔG_{stab}, decreased with increasing temperature from 3.5 M and ~25 kJ/mol at 12 °C to 2.3 M and ~17 kJ/mol at 45 °C; full denaturation at 12 and 45 °C was achieved at 5 and 4 M GdmCl, respectively. The maximum of the thermodynamic stability was reached close to or below 12 °C.

Temperature-induced unfolding transitions were recorded at 202 nm in sodium phosphate and at 215 nm in sodium cacodylate buffer; in cacodylate, measurements were also performed at different GdmCl concentrations. Reversibility was examined by fast cooling immediately after denaturation and a subsequent scan of the same sample at a different heating rate. All transitions were found to be fully reversible (Fig. 5A). Because of the short post-transitional baseline, at zero GdmCl concentration the midpoint of the thermal transition, $T_m \approx 87$ °C, could only be estimated. The two buffer systems did not effect the thermal stability of *Tm*Csp (Fig. 5B). Far-UV CD-spectra measured at different temperatures showed an isodichroic point close to 193 nm (data not shown), providing evidence for a two-state equilibrium transition. As has been mentioned in connection with the GdmCl-induced denaturation, the chaotropic and thermal denaturation mutually effect each other. Increasing the GdmCl concentration from 0 to 0.28 M led to a decrease of the transition midpoint by 10 °C (Fig. 5C). A quantitative thermodynamic analysis combining spectroscopic data with calorimetric measurements is in progress.

Table 2. *RMS deviations of the structures discussed in this work*[a]

	*Tm*CSP (model)	*Bs*CspB X-ray (1csp)	*Bs*CspB NMR (1nmg)
Bs (1csp)	0.32 Å (65 aa)	—	—
Bs (1nmg)	1.39 Å (56 aa)	1.41 Å (57 aa)	—
Ec (1mjc)	0.59 Å (59 aa)	0.69 Å (58 aa)	1.46 Å (57 aa)

[a]The structure of *Tm*Csp was modeled according to the *Bs*CspB crystal structure. Numbers indicate the Cα atom deviation calculated with the "Magic Fit" option in the program SWISSPDB 2.6; numbers in brackets indicate the number of superimposable residues

Table 3. *Structural parameters calculated from the TmCsp model structure and the experimentally determined structures of cold shock proteins*[a]

	*Tm*Csp model	*Bs*CspB X-ray (1csp)	*Bs*CspB NMR (1nmg)	*Ec*CspA X-ray (1mjc)
Salt bridges ≤ 7 Å	5	5	4	1
Intramolecular H-bonds	35	35	29	37

[a]Salt bridges and hydrogen bonds were counted by the program INSIGHTII (Molecular Simulations, Inc.).

Discussion

In this work, the isolation, cloning, overexpression, and characterization of the first Csp homolog from a hyperthermophile has been

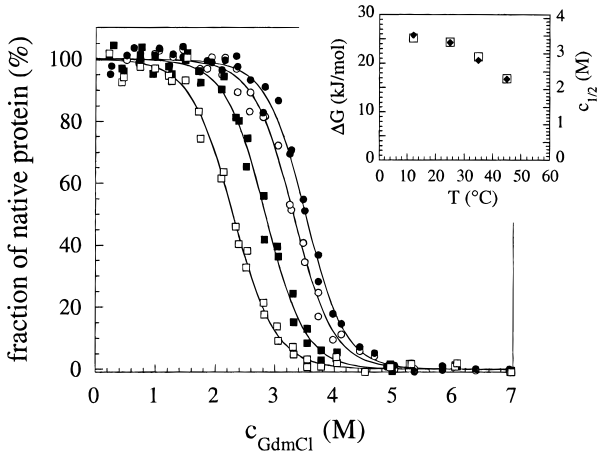

Fig. 4. GdmCl-induced equilibrium transitions of *Tm*Csp monitored by fluorescence emission at 337 nm (λ_{exc} = 280 nm) at 10 μg/mL protein concentration in 0.1 M sodium cacodylate pH 7.0: 12 °C (filled circles), 26 °C (open circles), 35 °C (filled squares), and 45 °C (open squares). The lines represent the least-squares fits assuming a two state model according to Santoro and Bolen (1988). Insert: Graphic representation of the temperature dependences of the free energy of stabilization at 0 M GdmCl (ΔG_{stab}) and half concentration of GdmCl ($c_{1/2}$) as obtained from those fits.

Table 4. *Results from nonlinear least-squares fits of GdmCl-induced equilibrium transitions of TmCsp at different temperatures, based on the two-state model according to Santoro and Bolen (1988)*

T (°C)	ΔG_{stab} (kJ/mol)	m (kJ/mol M)	$(c_{GdmCl})_{1/2}$ (M)
12	25.2	−7.16	3.52
26	24.3	−7.35	3.31
35	21.4	−7.53	2.82
45	16.8	−7.35	2.29

described. DNA sequencing of a variety of independently produced PCR fragments led to identical results. The deduced amino acid sequence of *Tm*Csp was in agreement with results from N-terminal sequencing of both the natural and recombinant protein. It was also highly homologous to Csps from other prokaryotes as well as the cold shock domain present in various eukaryotic proteins.

The *csp* gene showed all typical elements of standard genes. A putative −35-promoter sequence was located 170 bp upstream of the translation starting point. Other cold-inducible genes contained a 5′-untranslated region (5′UTR) of similar length; they included the gene of the ribosomal binding factor csdA apart from *csp*s (Jiang et al., 1996; Mitta et al., 1997). As shown previously, this 5′UTR is essential for the regulation of the cold shock response. In the *Ec cspA* gene, the 5′UTR had a strong effect on the stability of its own transcription product depending on temperature (Mitta et al., 1997); its overproduction led to a prolonged cold shock response. On the other hand, *csp* genes that are not cold inducible did not possess the long 5′-UTR. Therefore, it was suggestive to assume that *Tm*Csp is also cold inducible. Preliminary experiments connected with a possible cold shock response of *Tm* after a downshift in growth temperature have been inconclusive so far (C. Welker, unpubl. results). The observation that a low level of *Tm*Csp was present also in cells grown under optimal conditions proved that the protein was expressed constitutively. Undoubtedly, the question what "cold shock" means for a hyperthermophilic organism is of interest, but cannot be answered at the present moment.

At 145 bp downstream of the *csp* termination codon (TAA), an ORF was found that encoded the 50S ribosomal protein L31. It was shown that the limiting step in the translation of normal proteins at low temperature is the initiation of translation (Broeze et al., 1978). Some ribosomal proteins (L7, L12, S1, S6A, and S6B) were identified to be continually synthesized after cold shock (Jones & Inouye, 1994). Therefore, one might speculate that in *Tm* the L31 gene is co-regulated with *csp*. For L31, a ribosome binding site and a −10 box were localized; a sequence corresponding to the *Tm* −35-consensus could not be identified as distinctly, but there were also other *Tm* genes, whose promoter sequences differed from the consensus sequences (de la Tour et al., 1998). Furthermore, the transcription termination site of *csp* was clearly detectable. Thus, a direct connection between *Tm csp* and L31 at the genomic level is unlikely.

*Bs*CspB as well as CspA and CspE from *Bacillus cereus* were reported to dimerize in the absence of phosphate ions (Makhatadze & Marahiel, 1994; Mayr et al., 1996); similarly, *Bs*CspB crystallized as a dimer (Schindelin et al., 1993). In contrast, a variety of physical methods proved *Tm*Csp in solution to be monomeric.

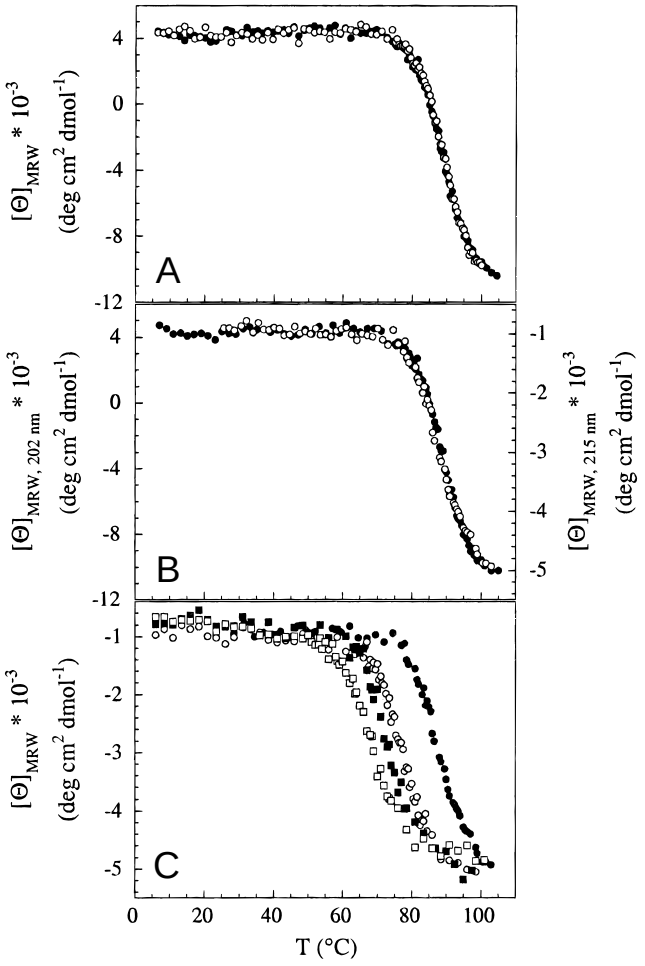

Fig. 5. Temperature-induced denaturation of *Tm*Csp monitored by far-UV CD at 0.1 cm pathlength. **A:** CD signal at 202 nm at 140 μg/mL protein concentration in 6 mM sodium phosphate pH 7.0, with 6 mM NaCl and 0.06 mM EDTA. Heating rates were 60 °C/h (filled circles) and 30 °C/h (open circles). **B:** Thermal transitions of *Tm*Csp measured as described in **A** (filled circles, left ordinate) and at 215 nm at 400 μg/mL protein concentration in 0.1 M sodium cacodylate pH 7.0 (open circles, right ordinate). **C:** Thermal transitions of *Tm*Csp monitored at 215 nm at 400 μg/mL protein concentration in 0.1 M sodium cacodylate pH 7.0, at 0 M (filled circles), 0.28 M (open circles), 0.58 M (filled squares), and 0.88 M (open squares) GdmCl. The transitions are independent of the heating rate (30 to 60 °C/h).

Regarding their relative intrinsic stabilities against GdmCl and temperature, bacterial Csps were found to reflect the optimum growth temperatures of their mesophilic, thermophilic, and hyperthermophilic origin: T_m(*Bs*CspB) ~ 52 °C, T_m(*Ec*CspA) ~ 60 °C, T_m(*Bc*CspB) ~72 °C, T_m(*Tm*Csp) ~87 °C (Chatterjee et al., 1993; Makhatadze & Marahiel, 1994; Perl et al., 1998; Reid et al., 1998). The observation that *Tm*Csp exhibits the highest thermal stability of all Csps presently known was confirmed by kinetic folding/unfolding experiments (Perl et al., 1998; Reid et al., 1998). Obviously, the melting point of *Tm*Csp exceeds the optimal growth temperature of the bacterium at ~80 °C (Huber et al., 1986), though the maximum of the stability profile clearly is far below the hyperthermophilic temperature range (cf. insert Fig. 4). The ΔG_{stab} vs. T profile is anomalously shallow, resembling the characteristics of other small proteins. As shown by

Alexander et al. (1992), proteins in the 7 to 10 kDa range tend to have high T_m values due to the small size of the folding unit and their specific thermodynamic properties. Examples comparable with *Bs*CspB showed denaturation temperatures around 80 °C. Based on the present data, we cannot answer the question why mesophilic Csps have low denaturation temperatures, or why *Tm*Csp is only marginally stable at physiological temperature except for the general notion that proteins are multifunctional and optimized for function rather than stability (Wetlaufer, 1980). Under physiological conditions they use to exhibit only marginal stabilities, in the order of 30–100 kJ/mol. The fact that ΔG_{stab} for isolated *Tm*Csp is found to be at the lower margin might be related to its function. In vivo, i.e., in the crowded cytosol and conjugated with its target-nucleic acid, it is expected to be stabilized significantly (Jaenicke, 1998).

Given the close similarity of the three-dimensional structures of Csps, one might ask, what is the mechanism by which *Tm*Csp gains its higher stability compared to its mesophilic and moderately thermophilic counterparts. The multiple alignment of 38 prokaryotic and two eukaryotic Csp sequences with *Tm*Csp allows the following conclusions: Regarding *aromatic* residues, two phenylalanines, that are highly conserved in other Csps, are substituted by tyrosine and tryptophan. The first is unique in Csps from thermophiles, while the latter one is solely found in *Tm*Csp. 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 *Tm*Csp compared to 19, 17, and 16 in *Bs*CspB, *Bc*CspB, and *Ec*CspA, respectively. The calculated isoelectric point (pI = 7.7) is exceedingly high compared to other Csps (pI = 4.3–5.0). *Tm*Csp contains the maximum Lys content: only 4 out of its 11 Lys residues are conserved. The single Arg is only present in *Bc*CspB, *Bst*CspB, and *Tm*Csp; apparently, it is a characteristic of thermophiles. His61 in *Tm*Csp (which replaces Asn in most other Csp sequences) stands out because it might participate in a peripheral ion pair. The same holds for a number of additional acidic residues in *Tm*Csp, which substitute polar, nondissociable residues in mesophilic Csps. With respect to hydrophobic residues, Val45 is unique in *Tm*Csp and Csp from *Streptomyces clavuligerus*; in most other sequences, polar residues prevail at its position. Similarly, Met34 is replaced by Gly/Ser/Thr in most Csps. Though these alterations may contribute to ΔG_{stab}, they do not provide a key to the stability of *Tm*Csp; rather they confirm that presently no general rules of thermal stabilization are available (Jaenicke, 1991).

To get insight into contributions of single residues to local interactions, the three-dimensional structure of *Tm*Csp is indispensable. Therefore, with no X-ray or NMR-data at hand, the attempt was made to predict the spatial structure by homology modeling based on the high sequence identity of *Tm*Csp to other Csps with known high-resolution structures. As summarized in Table 2, the resulting three-dimensional model turned out to be closely similar to the known Csp structures. The statistical analysis showed that the numbers of salt bridges and hydrogen bonds were practically unaltered, confirming that the molecular basis of the anomalous stability of *Tm*Csp is not easily detectable in terms of altered local interaction. Having in mind that the stability of proteins is a cumulative effect of small increments, the exact positions of the side chains would be required to discover the marginal differences responsible for thermal stabilization. They are currently under investigation by 2D-NMR.

Materials and methods

Strains, plasmids, cloning

Thermotoga maritima strain MSB8 (DSM strain 3109) was kindly provided by Drs. K.O. Stetter and R. Huber (University of Regensburg). Cells were grown at 80 °C (cf. Huber et al., 1986). *Ec* strains DH10b (Bethesda Research Laboratories, Bethesda, Maryland) and BL21(DE3) (Novagen, Madison, Wisconsin) were used for cloning and expression. Cells were grown in Luria Bertani medium (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride), supplemented with agar (1.3%) and ampicillin (500 μg/mL) or tetracycline (20 μg/mL) when appropriate. pUCBM20 (Vieira & Messing, 1982) was used for cloning. The expression vector pET21a is a pET-derivative according to Studier et al. (1990), developed by Novagen. Recombinant DNA techniques were based on Sambrook et al. (1989), using enzymes purchased from Boehringer Mannheim (Mannheim, Germany), Pharmacia (Uppsala, Sweden), and Gibco-BRL (Paisley, United Kingdom). Oligonucleotides (Table 1) were obtained in HPLC-purified state from MWG Biotech. Preparation of genomic *Tm* DNA was performed as reported earlier (Ostendorp et al., 1993). The *Tm* genomic library in pUN121 was provided by Dr. W. Liebl, University of Göttingen (cf. Ostendorp et al., 1993). DNA fragments were purified by separation on low melting-point agarose gels (Biozym) and subsequent isolation using the QIAquick Gel Extraction Kit from Qiagen (Hilden, Germany).

PCRs were carried out in a total volume of 100 μL, containing 2 mM $MgCl_2$, didesoxytrinucleotides at final concentrations of 200 μM each, 100 pmol of 3′- and 5′-terminal primer, about 200 ng template DNA, and 2.5 U Taq-DNA-Polymerase (Boehringer Mannheim). For amplification of the final fragment, cloned for recombinant expression of *Tm*Csp, High Fidelity Taq-DNA-Polymerase from Boehringer Mannheim was used.

Sequencing was performed using standard procedures with a Model 373A gene sequencer from Applied Biosystems (Foster, City, California). PCR products were sequenced directly by cycle sequencing, the primers were the same as for the preceding PCR reactions.

The DNA probe for southern blotting and colony hybridization was produced by PCR with degenerated primers N1 and C1 directly from genomic DNA, as well as from the complete genomic library in pUN121. It was labeled with [α-^{32}P]-dCTP (Amersham Buchler, Braunschweig, Germany) making use of the Random Primed Labeling Kit (Boehringer Mannheim).

For preparation of partial genomic *Tm* DNA libraries, genomic *Tm* DNA was digested with *Hind*III, *Eco*RI, *Bam*HI, *Eco*RV, and *Xho*I, respectively. DNA from the *csp*-containing fragments (determined by southern blot analysis) was cloned into pUCBM20. Both the complete and the partial genomic library were screened in *Ec* DH10b, making use of standard procedures for colony hybridization (Sambrook et al., 1989). Moreover, the partial gene libraries were used in form of plasmid pools as template DNA for PCRs with combinations of gene- and vector-specific primers, N1 and C2 and pUC1 and pUC2, respectively. The produced fragments allowed the 3′- and 5′-terminal regions of the *csp* gene to be sequenced and the primers N2 and C3 (including an *Nde*I and a *Hind*III restriction site) to be developed. These primers were used for cloning into pET21a.

Sequence analyses

For gene and protein database searches, statistical analysis of the protein sequence and multiple DNA and protein alignments, Gene-

Doc 2.2 (Nicholas & Nicholas, 1997) and the following programs from the HUSAR program package were used: FASTA, CLUSTAL, PEPSTATS, and MALIGN.

Homology modeling

Modeling was performed using MODELLER version 3 and version 4 (Sali & Blundell, 1993), with the coordinate set from the crystal structure of *Bs*CspB (PDB entry 1csp). Missing side chains in this parent structure (Glu3 Glu21, Glu36, and Glu66) were inserted according to standard side-chain rotamers. Alternative models were built based on the *Bs*CspB NMR structure (PDB entry 1nmg), and the *Ec*CspA structure (PDB entry 1mjc), respectively; structure quality analysis made use of the PROSAII subroutine in PROMETHEUS (Sippl, 1995). Superposition of the molecular structures was performed using the "Magic Fit" option from the program SwissPDB (Guex & Peitsch, 1997).

Protein purification

All chromatographic purification steps were performed aerobically at room temperature. Chromatographic materials were purchased from Pharmacia. Purification was followed by SDS-PAGE on 15% polyacrylamide gels after each step.

Purification of natural protein from Tm cells

Frozen cells (28.5 g wet weight) were thawed, resuspended in 40 mL 10 mM Tris-HCl buffer pH 8.0, containing 1 mM EDTA, and disrupted by passing two times through a French press (SLM Aminco Instruments) at 15,000-18,000 p.s.i., i.e., $(1.0–1.2) \times 10^8$ N/m^2. DNA and RNA were digested by incubation with 5 mM $MgCl_2$ plus 20 μg/mL DNaseI (Boehringer Mannheim) and RNaseA (Sigma, St. Louis, Missouri), respectively, for at least 1 h at 37 °C. After centrifugation (1 h at 48,000 g), the supernatant was adjusted to pH 9.3 with NaOH and applied to a 50 mL Q-Sepharose HP column (XK 26/20) equilibrated with 10 mM Tris-HCl buffer pH 9.3, 1 mM EDTA. The column was washed with this buffer until the baseline was reached again (flow rate 2 mL/min). The flow-through, containing proteins with high isoelectric points, was dialyzed against 10 mM sodium phosphate buffer pH 7.3, 1 mM EDTA, and subsequently applied to a 10 mL Heparin Sepharose CL-6B column equilibrated with the same buffer. Elution of bound proteins was achieved with a linear gradient of 0–0.5 M NaCl within eight column volumes (flow rate 1 mL/min). Csp-containing fractions were pooled, dialyzed against 50 mM MES pH 6.0, 10 mM NaCl, 1 mM EDTA, and loaded onto a 10 mL SP-Sepharose HR column under identical buffer conditions. The protein was eluted with a linear gradient from 0.01 to 1 M NaCl within 10 column volumes (flow rate 1 mL/min). Finally, the protein was subjected to gel permeation chromatography using a 180 mL XK 16/100 Superdex 75 pg column equilibrated with 50 mM Tris-HCl pH 7.9, containing 100 mM NaCl and 1 mM EDTA (flow rate 0.7 mL/min).

Purification of recombinant protein from Ec cells

2 to 4 L LB_{amp} medium were inoculated with 20 to 40 mL of a fresh *Ec* overnight culture, which had been washed twice with fresh medium to remove β-lactamase. Cells were grown at 26 °C to an OD_{546nm} of 1, and induced by IPTG at a final concentration of 1 mM. The cells were allowed to grow for further 16 h at 26 °C, and were then harvested by centrifugation (5,000 *g*, 15 min, 4 °C). The pellet was resuspended in 50 mM Tris-HCl buffer pH 8.0, containing 1 mM EDTA (buffer A), to a theoretical OD_{546nm} of 50 according to the starting density. Cells were disrupted, and DNA and RNA were digested as described for the natural protein. After centrifugation for 1 h at 48,000 *g*, the supernatant was diluted fivefold with buffer A. Most *Ec* proteins were denatured by a 30 min heat step at 80 °C, and removed by a further centrifugation for 1 h at 48,000 *g*. The supernatant was adjusted to a pH of 9.3 with NaOH and purification was carried on as described before, but without the Heparin–Sepharose step.

Characterization

Chemicals and buffers

10 to 20 mM sodium or potassium phosphate pH 7.0, or 0.1 M sodium cacodylate-HCl pH 7.0, were used as buffers. Guanidinium chloride (GdmCl) (ultrapure) was purchased from ICN, sodium cacodylate from Fluka (Buchs, Switzerland). All other chemicals were from Merck (Darmstadt, Germany). GdmCl concentrations were determined from refractive indices according to Pace (1986).

N-terminal amino acid sequencing

For N-terminal amino acid sequencing, the proteins were separated on a SDS-polyacrylamide gel and transferred onto a polyvinyldifluoride membrane (Millipore, Bedford, Massachusetts). The corresponding Coomassie stained protein bands were excised and sequenced by automated Edman degradation using an Applied Biosystems protein sequencer 477A.

Reversed-phase HPLC

Reversed-phase HPLC was performed by the standard procedure on a C4 column.

Electrospray ionization mass spectrometry

Electrospray ionization mass spectrometry was performed using a Finnigan MAT SSQ7000 mass spectrometer.

Analytical ultracentrifugation

Analytical ultracentrifugation was performed at room temperature in a Beckman Model E analytical ultracentrifuge making use of an AnH-Ti rotor and cells with 2 and 12 mm pathlength: Sedimentation velocity at 68,000 and 44,000 rpm, high-speed sedimentation equilibrium at 32,000 and 16,000 rpm. Initial protein concentrations: 0.2–13.2 mg/mL. The partial specific volume was calculated from the amino acid composition, the sedimentation coefficient from the slope in an $\ln(r)$ vs. time plot, the molecular mass distribution from $\ln(c)$ vs. r^2 plots.

Spectral analysis

UV absorption spectra of native and denatured (6 M GdmCl) protein were measured in a Cary 1 Spectrophotometer (Varian, Palo Alto, California). Protein concentrations were determined based on the molar absorption coefficient (Pace et al., 1995).

Fluorescence emission spectra were monitored using a Perkin Elmer MPF-L3 fluorescence spectrograph at an excitation wavelength (λ_{exc}) of 280 and 295 nm, and slit widths of 6 nm, and 4 nm for emission and excitation, respectively. GdmCl-induced equilib-

rium unfolding and refolding as well as pH stability made use of the change in fluorescence emission at 337 nm (λ_{exc} = 280 nm).

Circular dichroism spectra were recorded in a Jasco J715 spectropolarimeter using thermostated 1 cm or 0.1 cm cylindrical cuvettes in the near- or far-UV region, respectively. The spectra were accumulated 10 times at 1 nm band width, 1 s response time, 0.2 nm resolution, and 20 nm/min scan speed. The mean residue weight (113.3) was calculated based on the molecular mass determined by electrospray mass spectrometry.

Stability

To monitor the GdmCl-induced equilibrium unfolding/refolding transitions, a 1 mg/mL Csp solution in cacodylate buffer containing either 0 or 6 M GdmCl was incubated for 1 h at 25 °C. Fifteen microliter portions were pipetted into fluorescence cuvettes; de- and renaturation were induced by rapid 1:100 dilution with buffers containing 0 to 7 M GdmCl. Samples were overlayed with mineral oil (Sigma) and allowed to reach equilibrium at different temperatures between 12 and 45 °C. The ratio of native and unfolded protein was measured by changes in fluorescence (λ_{exc} = 280 nm, λ_{em} = 337 nm); data were evaluated according to Santoro & Bolen (1988).

pH-Stability of *Tm*Csp was measured by the same method, using 1:100 dilutions with the following buffers: 20 mM glycine/HCl pH 2.6, 3.2, 3.6; 20 mM sodium acetate pH 3.6, 4.4, 5.2; 20 mM sodium cacodylate pH 5.2, 6.0, 6.6; 20 mM sodium phosphate pH 6.6, 7.2, 8.0; 20 mM glycine/NaOH pH 8.6, 9.4, 10.4; 20 mM sodium phosphate/NaOH pH 11.0, 11.5.

For temperature-induced unfolding at temperatures between 5 ° and 105 °C, the cuvette was heated by a peltier element controlled by software provided by Jasco (Tokyo). Transitions were recorded by changes in the far-UV CD at 215 and 202 nm, 16 s response time, 0.5 °C resolution, and 30 to 60 °C/h heating rate. Samples were overlayed with mineral oil, and heated at least twice at different heating rates.

Acknowledgments

We should like to thank Dr. S. Fabry for his advice in connection with the cloning of the gene and for helpful discussions. The collaboration of Eduard Hochmuth in performing N-terminal sequencing, mass spectrometry, and HPLC analyses is gratefully acknowledged. This work was supported by grants of the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

Alexander P, Fahnestock S, Lee T, Orban J, Bryan P. 1992. Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: Why small proteins tend to have high denaturation temperatures. *Biochemistry 31*:3597–3603.

Broeze RJ, Carol JS, Daniel HP. 1978. Effects of low temperature on in vivo and in vitro protein synthesis in *E. coli* and *Pseudomonas fluorescens*. *J Bacteriol 134*:861–874.

Chatterjee S, Jiang W, Emerson SD, Inouye M. 1993. The backbone structure of the major cold-shock protein CS7.4 of *E. coli* in solution includes extensive β-sheet structure. *J Biochem 114*:663–669.

de la Tour CB, Portemer C, Kaltoum H, Duguet M. 1998. Reverse gyrase from the hyperthermophilic bacterium *Thermotoga maritima*: Properties and gene structure. *J Bacteriol 180*:274–281.

Fang L, Hou Y, Inouye M. 1998. Role of the cold-box region in the 5′ untranslated region of the cspA mRNA in its transient expression at low temperature in *E. coli*. *J Bacteriol 180*:90–95.

Graumann P, Marahiel MA. 1996a. Some like it cold: Response of microorganisms to cold shock. *Arch Microbiol 166*:293–300.

Graumann P, Marahiel MA. 1996b. A case of convergent evolution of nucleic acid binding modules. *Bioessays 18*:309–315.

Graumann P, Wendrich TM, Weber MHW, Schroeder K, Marahiel MA. 1997. A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Mol Microbiol 25*:741–756.

Guex N, Peitsch MC. 1997. Swiss-Model and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis 15*:2714–2723.

Huber R, Langworthy TA, König H, Thomm M, Woese CR, Sleytr UB, Stetter KO. 1986. *Thermotoga maritima sp. nov.* represents a new genus of unique extremely thermophilic eubacteria growing up to 90 °C. *Arch Microbiol 144*:324–333.

Jaenicke R. 1991. Protein stability and molecular adaptation to extreme conditions. *Eur J Biochem 202*:715–728.

Jaenicke R. 1998. What ultrastable globular proteins teach us about protein stabilization. *Biochemistry (Moscow) 63*:312–321.

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.

Jones PG, Inouye M. 1994. The cold shock response: A hot topic. *Mol Microbiol 11*:811–818.

Jones PG, Inouye M. 1996. RbfA, a 30S-ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol Microbiol 21*:1207–1218.

Jones PG, Mitta M, Kim Y, Jiang WN, Inouye M. 1996. Cold shock induces a major ribosomal-associated protein that unwinds double-stranded RNA in *E. coli*. *Proc Natl Acad Sci USA 93*:76–80.

Lee SJ, Xie A, Jiang W, Etchegaray J-P, Jones PG, Inouye M. 1994. Family of the major cold shock protein, CspA (CS7.4), of *E. coli*, whose members show a high sequence similarity with the eukaryotic Y-box binding proteins. *Mol Microbiol 11*:833–839.

Liao D, Dennis PP. 1992. The organization and expression of essential transcription translation component genes in the extremely thermophilic eubacterium *Thermotoga maritima*. *J Biol Chem 227*:22787–22797.

Makhatadze GI, Marahiel MA. 1994. Effect of pH and phosphate ions on self-association properties of the major cold-shock protein from *Bacillus subtilis*. *Protein Sci 3*:2144–2147.

Mayr B, Kaplan T, Lechner S, Scherer S. 1996. Identification and purification of a family of dimeric major cold shock protein homologs from the psychrotrophic *Bacillus cereus* WSBC 10201. *J Bacteriol 178*:2916–2925.

Mitta M, Fang L, Inouye M. 1997. Deletion analysis of cspA of *E. coli*: Requirement of the AT-rich UP element for cspA transcription and the downstream box in the coding region for its cold shock induction. *Mol Microbiol 26*:321–335.

Newkirk K, Feng W, Jiang W, Tejero R, Emerson SD, Inouye M, Montelione GT. 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.

Nicholas KB, Nicholas HB. 1997. GeneDoc: A tool for editing and annotating multiple-sequence alignments. Distributed by the author at http://www.cris.com/~ketchup/genedoc.shtml.

Ostendorp R, Liebl W, Schurig H, Jaenicke R. 1993. The L-lactate dehydrogenase gene of the hyperthermophilic bacterium *Thermotoga maritima* cloned by complementation in *Escherichia coli*. *Eur J Biochem 216*:709–715.

Pace CN. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol 131*:266–280.

Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci 4*:2411–2423.

Perczel A, Park K, Fasman GD. 1992. Deconvolution of the circular dichroism spectra of proteins: The circular dichroism spectra of the antiparallel β-sheet in proteins. *Proteins Struct Funct Genet 13*:57–69.

Perl D, Welker C, Schindler T, Schröder K, Marahiel MA, Jaenicke R, Schmid FX. 1998. Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat Struct Biol 5*:229–235.

Reid KL, Rodriguez HM, Hillier BJ, Gregoret L. 1998. Stability and folding of a model β-sheet protein, *E. coli* CspA. *Protein Sci 7*:470–479.

Sali A, Blundell TL. 1993. Comparative protein modeling by satisfaction of spatial restraints. *J Mol Biol 234*:779–815.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry 27*:8063–8068.

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.

Schindelin H, Marahiel MA, Heinemann U. 1993. Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature 364*:164–168.

Schindler T, Herrler M, Marahiel MA, Schmid FX. 1995. Extremely rapid protein folding in the absence of intermediates. *Nat Struct Biol 2*:663–673.

Schindler T, Schmid FX. 1996. Thermodynamic properties of an extremely rapid protein folding reaction. *Biochemistry 35*:16833–16842.

Schnuchel A, Wiltscheck R, Czisch M, Herrler M, Willimsky G, Graumann P, Marahiel MA, Holak TA. 1993. Structure in solution of the major cold shock protein from *B. subtilis*. *Nature 364*:169–171.

Schröder K, Zuber P, Willimsky G, Wagner B, Marahiel MA. 1993. Mapping of the *B. subtilis* cspB gene and cloning of its homologs in thermophilic, mesophilic and psychrophilic bacilli. *Gene 136*:277–280.

Shaw MK, Ingraham JL. 1967. Synthesis of macromolecules by *E. coli* near the minimal temperature for growth. *J Bacteriol 94*:157–164.

Sippl MJ. 1995. Knowledge based potentials for proteins. *Curr Opin Struct Biol 5*:229–235.

Studier WF, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol 185*:66–89.

Thieringer HA, Jones PG, Inouye M. 1998. Cold shock and adaptation. *Bioessays 20*:49–57.

Vieira J, Messing J. 1982. The pUC plasmids, an M13mp7-derived system for insertion, mutagenesis and sequencing with synthetic universal primers. *Gene 19*:259–268.

Welker C. 1996. Isolierung nucleotid-bindender Proteine aus dem hyperthermophilen Bakterium *Thermotoga maritima* [Thesis, University of Regensburg]. Regensburg, Germany.

Wetlaufer DB. 1980. Practical consequences of protein folding mechanisms. In: Jaenicke R, ed. *Protein folding*. New York: Elsevier/North-Holland. pp 323–329.

Willimsky G, Bang H, Fischer G, Marahiel MA. 1992. Characterization of cspB, a *B. subtilis* inducible cold shock gene affecting cell viability at low temperatures. *J Bacteriol 174*:6326–6335.

Wistow G. 1990. Cold shock and DNA binding. *Nature 344*:823–824.

Wolffe AP, Tafuri S, Ranjan M, Familari M. 1992. The Y-box factors: A family of nucleic acid binding proteins conserved from *E. coli* to man. *New Biol 4*:349–359.

Woody RW. 1994. Contributions of tryptophan side chains to the far-UV circular dichroism of proteins. *Eur Biophys J 23*:253–262.

Yamanaka K, Inouye M. 1997. Growth-phase-dependent expression of cspD, encoding a member of the CspA family in *E. coli*. *J Bacteriol 179*:5126–5130.

Yamanaka K, Mitani T, Ogura T, Niki H, Hiraga S. 1994. Cloning, sequencing, and characterization of multicopy suppressors of a mukB mutation in *E. coli*. *Mol Microbiol 13*:301–312.