*Proc. Natl. Acad. Sci. USA*
Vol. 91, pp. 5119–5123, May 1994
Biochemistry

# Crystal structure of CspA, the major cold shock protein of *Escherichia coli*

(nucleic acid binding/cold shock domain/x-ray crystallography/Y-box factors)

HERMANN SCHINDELIN*, WEINING JIANG†, MASAYORI INOUYE†, AND UDO HEINEMANN*‡

*Forschungsgruppe Kristallographie, Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle-Strasse 10, D-13122 Berlin, Germany; and †Department of Biochemistry, University of Medicine and Dentistry of New Jersey, 675 Hoes Lane, Piscataway, NJ 08854-5635

*Communicated by Richard E. Dickerson, February 22, 1994*

ABSTRACT **The major cold shock protein of *Escherichia coli*, CspA, produced upon a rapid downshift in growth temperature, is involved in the transcriptional regulation of at least two genes. The protein shares high homology with the nucleic acid-binding domain of the Y-box factors, a family of eukaryotic proteins involved in transcriptional and translational regulation. The crystal structure of CspA has been determined at 2-Å resolution and refined to $R = 0.187$. CspA is composed of five antiparallel β-strands forming a closed five-stranded β-barrel. The three-dimensional structure of CspA is similar to that of the major cold shock protein of *Bacillus subtilis*, CspB, which has recently been determined at 2.45-Å resolution. However, in contrast to CspB, no dimer is formed in the crystal. The surface of CspA is characteristic for a protein interacting with single-stranded nucleic acids. Due to the high homology of the bacterial cold shock proteins with the Y-box factors, *E. coli* CspA and *B. subtilis* CspB define a structural framework for the common cold shock domain.**

The cold shock response in *Escherichia coli* follows an abrupt shift in growth temperature from 37°C to 10°C, inducing a lag phase in cell growth of 4–5 hr. It is accompanied by a severe reduction in protein synthesis (1). As a consequence of this cold shock, the relative rate of production of at least 14 cold shock proteins is increased. For 13 out of the 14 proteins the increase is 2- to 10-fold whereas synthesis of the major cold shock protein, CspA (CS 7.4), increases at least 100-fold, reaching a level of >10% of total protein synthesis at 10°C (2). Other proteins expressed as part of the cold shock response of *E. coli* include NusA, RecA, polynucleotide phosphorylase, translation initiation factors 2α and 2β, pyruvate dehydrogenase (lipoamide), dihydrolipoamide acetyltransferase of pyruvate dehydrogenase, the nucleoid protein H-NS, and subunit A of DNA gyrase (1, 3, 4).

CspA is a small hydrophilic protein consisting of 70 amino acids. It has striking similarity, at the level of 43% sequence identity (Fig. 1), with one domain of the Y-box factors, which is referred to as the cold shock domain (5, 6). The Y-box factors are a family of eukaryotic nucleic acid-binding proteins that preferentially bind to the Y box, an element of sequence CTAATTGGYYAA found in the promoter regions of mammalian major histocompatibility complex class II genes (6). Within this sequence the underlined pentamer is especially conserved. Members of this family have also been found to bind to mRNA and to regulate translation in germ cells (7, 8).

CspA was shown to act as a transcriptional activator of the *hns* and *gyrA* genes encoding two other cold shock proteins (3, 4). The promoter of *hns* contains one ATTGG element, whereas the promoter of *gyrA* contains three such elements, one of which is required for specific CspA–DNA interaction (4). ATTGG elements have been identified also in the promoter regions of genes encoding RecA, NusA, and polynucleotide phosphorylase, suggesting a common mechanism for induction of these proteins at reduced temperature. However, CspA seems not to regulate its own biosynthesis at the level of transcription, although an ATTGG element is present in the promoter region of the *cspA* gene (9). Recently, three *E. coli* genes homologous to *cspA* have been found, one of which is also transcriptionally induced by cold shock (10).

CspB, the analog of CspA in *Bacillus subtilis*, consists of 67 amino acids, and the two proteins share 61% sequence identity (11). The structure of CspB has been determined from two crystal forms and also in solution (12, 13). Here we present the x-ray structure analysis of *E. coli* CspA at high resolution. In contrast to *B. subtilis* CspB, this protein is a monomer in the crystal. A comparison of the three-dimensional structures of the bacterial cold shock proteins and consideration of data on nucleic acid binding derived from gel retardation and NMR studies (12, 14) [ref. 14 is the preceding paper in this issue] lead to the assignment of a surface area as the probable nucleic acid-binding portion of the cold shock domain.§

## MATERIALS AND METHODS

Crystals were grown by vapor diffusion of a solution containing purified CspA at 10 mg/ml in 10 mM Tris·HCl, pH 7.5/50 mM NaCl against a reservoir solution containing 26% PEG 1500 as precipitant. The crystals belong to the orthorhombic space group $P2_12_12_1$ with one molecule per asymmetric unit and 38% solvent. Diffraction data were collected on a MAR-Research (Hamburg, Germany) imaging plate detector mounted on an Enraf-Nonius (Delft, The Netherlands) FR 571 rotating anode operating at 45 kV and 80 mA and were processed with MOSFLM (15). The structure was solved by molecular replacement with XPLOR (16) using the coordinates of the trigonal crystal form of CspB (12) as starting model. The cross-rotation search followed by Patterson correlation refinement identified one solution with a correlation coefficient of 0.143 (the next highest correlation coefficient was 0.134). The translation search gave a solution which was 5.3 standard deviations ($\sigma$) above the mean score with an $R$ value of 0.533 in the resolution range from 10 to 3 Å. The next highest solution which did not correspond to the same translation vector was 3.4 $\sigma$ above the mean.

Within 20 cycles of rigid-body refinement in the resolution range from 10 to 3 Å the $R$ value dropped to 0.455. The coordinate file was then modified according to the CspA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: OB fold, oligonucleotide/oligosaccharide-binding fold.

‡To whom reprint requests should be addressed.

§The atomic coordinates and structure amplitudes have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973.

```
                          10        20        30        40        50        60        70
                          .         .         .         .         .         .         .
CspA  E. coli   MSGKMTGIVKWFNADKGFGFITPDDGSKDVFVHFSAIQNDG::::YKSLDEGQKVSFTIESGAKGPAAGNVTSL
CspB  E. coli   --N----L-------------S-V----------------N::::-RT-F-----T-S----------A--IITD        80%
CspC  E. coli    MA-IK-Q-----ES--------A-------------GN-::::F-T-A---N-E-E-QD-Q-----V---AI         68%
CspD  E. coli      MEK-T-----NA------C-EG-GE-I-A-Y-T--M--::::-RT-KA--S-Q-DVHQ-P--NH-SVIVPVEVEAAVA  48%
CspB  B. sub.      MLE-K-----SE------EVE:-QD----------GE-::::F-T-E---A---E-VE-NR--Q-A---KEA        61%
Csp7  S. clav.     MA--T------E------AQ-G-GP-----Y---NAT-::::FR--E-N-V-N-DVTH-E:--Q-E--SPA         56%
CbfR  Rat       IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--AE-A---GPG      43%
CbfM  Mouse     IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--AE-A---GPG      43%
CbfH  Human     IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--AE-A---GPG      43%
DbpA  Human     LAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--AE-A---GPD      43%
Yb1H  Human     IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--EE-A---GPG      43%
Yb1X  Xenopus   IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--AE-A---GPE      43%
Yb2X  Xenopus   LATQVQ-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKFLR-VGD-ET-E-DVVE-E--AE-A---GPG      41%
Yb3X  Xenopus   IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KKNNPRKYLR-VGD-ET-E-DVVE-E--AE-A---GPG      43%
```

FIG. 1. Sequence alignment of the cold shock proteins and representatives of the Y-box factors. In the alignment, identical residues are indicated by dashes and insertions by colons. The percent identity with CspA shown on the right was calculated by one-by-one comparisons of the aligned sequences. *B. sub., Bacillus subtilis; S. clav., Streptomyces clavuligerus.*

sequence. Simulated annealing refinement with XPLOR (17, 18), first at 2.8-Å resolution and then at 2.0-Å resolution, followed by stereochemically restrained least-squares refinement using PROFFT (19, 20), resulted in a model of CspA without the N-terminal methionine and including 36 water oxygens.

## RESULTS

**Structure of CspA.** Structure refinement converged at an $R$ value of 0.187 with good model stereochemistry (Table 1). The final model comprises 514 protein atoms and 36 solvent sites treated as water oxygens with unit occupancies. Due to disorder the electron density is ill defined in the region from residue 39 to residue 42, where the highest thermal ($B$) factors are observed, with an average value of 72 Å$^2$ for the main-chain atoms. In contrast to this polypeptide segment the average main-chain $B$ factor is 27.5 Å$^2$ for the entire model. According to a PROCHECK (21) analysis, there are no outliers in the Ramachandran diagram. The mean coordinate error according to Luzzati (22) is <0.25 Å. The N-terminal methionine is not present in the crystal, and partial N-terminal sequencing of CspA has shown that the major fraction of the protein used for crystallization does not contain Met[1].

CspA is a compact molecule and is folded into five antiparallel β-strands with connecting turns and loops (Fig. 2). The five β-strands form a β-barrel which is created by strands β1–β3 on one side and by strands β4 and β5 on the other side and is closed on top by a long loop of 16 residues connecting strands β3 and β4. As a consequence of this folding, strands β1–β4 are arranged in the form of a Greek key motif (24). The precise location of the β-strands according to Kabsch and Sander (25) is as follows: β1 encompasses residues 5 to 13; β2, residues 18–23; β3, residues 30–33; β4, residues 50–56; and β5, residues 63–69. On the surface of strands β1–β3 we find an array of solvent-exposed aromatic residues including Trp[11], Phe[18], Phe[20], Phe[31], and Phe[34]. The folding of CspA as seen in the crystal is consistent with structural data derived from NMR spectroscopy (14).

**Comparison of the Structures of CspA and CspB.** The folding of *E. coli* CspA and *B. subtilis* CspB is similar, and the structural differences are mainly restricted to loop regions (Fig. 3). The CspA crystal structure can be superimposed with CspB, excluding loop regions, by least-squares refinement using the main-chain atoms of 49 residues with root-mean-square (rms) distances between corresponding atoms of 0.47 Å for the trigonal crystal form and 0.63 Å for the tetragonal crystal form of CspB, respectively. Thus, not only

Table 1. Data collection and refinement

| | |
|---|---|
| Data collection | |
| Space group | $P2_12_12_1$ |
| Lattice constants ($a$, $b$, $c$), Å | 47.12, 39.92, 30.92 |
| No. of measured reflections | 15,912 |
| No. of unique reflections | 3785 |
| Completeness (all data/2.13 to 2 Å) | 0.889/0.878 |
| Mean $I/\sigma$ | 5.99 |
| $R_{sym}$ (all data/2.13 to 2 Å) | 0.063/0.113 |
| Refinement | |
| No. of protein atoms | 514 |
| No. of solvent molecules | 36 |
| Total no. of distances restrained | 1422 |
| Bond distances, Å | 0.014 (0.020) |
| Angle distances, Å | 0.047 (0.040) |
| Planar distances, Å | 0.049 (0.050) |
| No. of planar groups restrained | 78 |
| Deviation from plane, Å | 0.009 (0.015) |
| No. of chiral centers restrained | 74 |
| Chiral volumes, Å$^3$ | 0.106 (0.100) |
| No. of nonbonded contacts restrained | 374 |
| Single torsion contacts, Å | 0.133 (0.150) |
| Multiple torsion contacts, Å | 0.139 (0.150) |
| Possible (X···Y) H bonds, Å | 0.124 (0.150) |
| No. of torsion angles restrained | 162 |
| Planar (0°, 180°) | 1.7 (3.0) |
| Staggered (±60°, 120°) | 19.6 (15.0) |
| Orthonormal (±90°) | 17.7 (20.0) |
| Isotropic thermal factor restraint | |
| Main-chain bonds, Å$^2$ | 2.31 (2.00) |
| Main-chain angles, Å$^2$ | 4.13 (3.00) |
| Side-chain bonds, Å$^2$ | 3.43 (4.50) |
| Side-chain angles, Å$^2$ | 5.29 (4.50) |
| H bonds, Å$^2$ | 17.17 (15.00) |
| $R$ factor | 0.187 |
| Correlation coefficient | 0.951 |

$R_{sym} = \Sigma_i\Sigma_j|\langle I_i\rangle - I_{i,j}|/\Sigma_i\Sigma_j I_{i,j}$, where $I_{i,j}$ are the individual measurements contributing to the mean reflection intensity, $\langle I_i\rangle$. $R = \Sigma|F_o - F_c|/\Sigma F_o$, the crystallographic residual. The correlation coefficient is $\Sigma[(F_o - \langle F_o\rangle)(F_c - \langle F_c\rangle)]/[\Sigma(F_o - \langle F_o\rangle)^2(F_c - \langle F_c\rangle)^2]^{1/2}$. For stereochemical parameters, the left number gives the rms deviation from ideal values and the right number (in parentheses) is the target variance used in refinement.

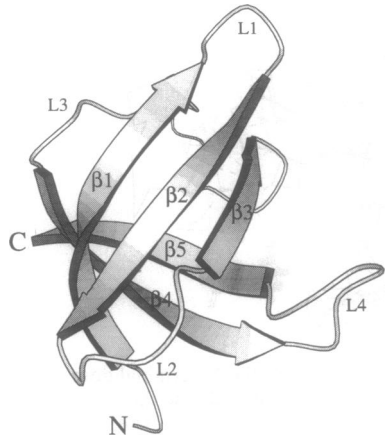


FIG. 2. Three-dimensional structure of CspA. β-Strands are given as curved arrows and are numbered β1–β5. Loops between β-strands are numbered L1–L4. This figure and Figs. 3, 4, and 6 were drawn by using MOLSCRIPT (23).

the folding topology but also the core structure of the cold shock proteins are nearly identical within the experimental limits of accuracy. Larger deviations between the two molecules occur in loop L2, connecting strands β2 and β3, where one residue is inserted in CspA, and in the long loop L3 around residues 40 and 41, where the electron density is ill defined in the CspA structure (see above). This loop is stabilized in both crystal forms of CspB by a stacking interaction of $Phe^{38}$, corresponding to $Tyr^{42}$ in CspA, with $Arg^{56}$ of a symmetry-related molecule in the crystal.

The largest deviations, up to 5 Å, are seen in the second half of strand β4 and the turn between strands β4 and β5. They cause a functionally important difference between the two structures; the dimerization observed in both crystal forms of CspB (12) is impossible for CspA. This is mainly due to the N-terminal extension in CspA and, in addition, due to the slightly different conformations of strands β4 along which dimerization occurs through antiparallel interaction in both crystal forms of CspB. It should be noted that the sequence match between residues in strands β4 of CspA and CspB is the lowest of all β-strand pairs in the two molecules.

**Cold Shock Domain and the Oligonucleotide/Oligosaccharide-Binding (OB) Fold.** The basic folding pattern of the cold shock domain as defined by the crystal structures of CspA and CspB is remarkably similar to the recently characterized OB fold (26, 27). Nucleic acid-binding representatives of the OB fold include the staphylococcal nuclease (28) and the N-terminal domain of yeast aspartyl-tRNA synthetase (29). In contrast to the cold shock domain, strands β3 and β4 in the OB fold are connected by an α-helix and not by the long loop seen in CspA. The structure of CspA can be superimposed upon the structures of staphylococcal nuclease and the N-terminal domain of yeast aspartyl-tRNA synthetase with rms deviations of 1.56 Å or 1.59 Å, respectively, for 38 pairs of equivalent $C^{\alpha}$ atoms in each case (Fig. 4). Additional representatives of the OB fold are the single-stranded-DNA-binding proteins (gene *5* proteins) of filamentous phage fd and M13, but, like CspA, gene *5* protein lacks the α-helix connecting strands β3 and β4 (30, 31).

All members of the OB-fold protein family, including gene *5* protein, have roughly the same oligomer binding site which is located on one side of the barrel and is formed by the surface of strands β1–β3 and the loops L1, L3, and L4. A part of this surface in the CspA crystal structure is shown in Fig. 5 with $2F_o - F_c$ electron density.

**Electrostatic Considerations.** That the side of the barrel which corresponds to the ligand binding site in the OB-fold proteins is the only region of CspA suited for nucleic acid binding is evident from a consideration of the electrostatic properties of the molecules (Fig. 6). At neutral pH CspA, as well as CspB, carries an overall negative charge. The isoelectric point of CspA has been experimentally determined as 5.0 (data not shown). Yet an attractive potential for nucleic acids is created by the arrangement of positive charges mainly on the above mentioned surface of the molecule, originating from $Lys^{10}$ in β1, $Lys^{16}$ in loop L1, $His^{33}$ in β3, $Lys^{43}$ in L3, and $Lys^{60}$ in L4 in the case of CspA and from $Lys^{7}$ in β1, $Lys^{13}$ in L1, $His^{29}$ in β3, $Lys^{39}$ in L3, and $Arg^{56}$ in L4 in the case of CspB. These surfaces of the two molecules also include the cluster of solvent-exposed aromatic residues.

## DISCUSSION

Under the conditions of gel electrophoresis CspA binds preferentially to the CCAAT containing strand of a 35-mer oligonucleotide derived from the *E. coli gyrA* promoter region and neither to the double strand nor to the complementary single strand (data not shown). *B. subtilis* CspB shows the same specificity toward this oligonucleotide (12) and, in addition, was found to interact with both single strands derived from a 54-mer Y-box oligonucleotide containing either the ATTGG or the CCAAT motif (32). A DNA single strand of sequence AACGGTTTGACGTACAGACCATTA, corresponding to the first 24 nucleotides of the untranslated part of the CspA mRNA, gives rise to defined chemical shifts in the NMR spectrum of CspA when bound to the protein (14). This combined evidence suggests that

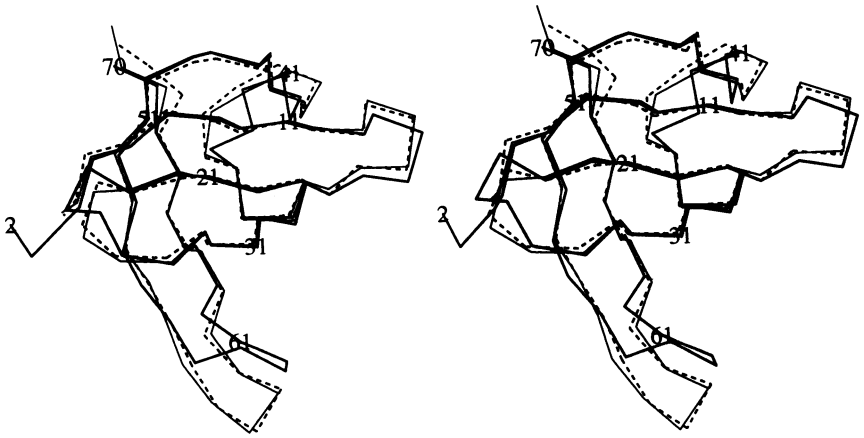


FIG. 3. Least-squares superposition of CspA (thick line) with the trigonal (thin line) and tetragonal (broken line) crystal forms of CspB. Every 10th $C^{\alpha}$ position in CspA is labeled with its residue number.

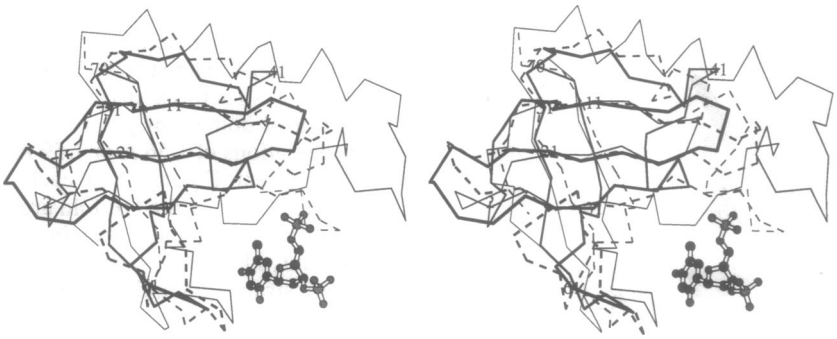

FIG. 4. Least-squares superposition of CspA (thick line), staphylococcal nuclease (thin line), and the N-terminal domain of aspartyl-tRNA synthetase (broken line). All three proteins are shown as $C^{\alpha}$ traces; every 10th $C^{\alpha}$ position in CspA is labeled, and the inhibitor of staphylococcal nuclease, thymidine 3',5'-bisphosphate, is shown in ball-and-stick representation to mark the oligomer binding site.

bacterial cold shock proteins bind preferentially to single-stranded nucleic acids.

The specificity of the binding, however, remains unclear. A simple common sequence motif of these nucleic acids cannot be deduced, raising the question whether a specific tertiary structure of these single strands is recognized. A secondary-structure prediction reveals hypothetical stem–loop structures with 5 bases in the loop and without sequence conservation for the 54-mer and the 24-mer oligonucleotides, but not for the 35-mer. On this basis, it appears impossible at present to construct atomic models of the CspA–DNA complex. On the other hand, binding to RNA also appears possible, since the cold shock proteins contain the RNP1 sequence motif typical for RNA-binding proteins (33–35) and a rudimentary RNP2 motif which are part of the proposed nucleic acid-binding surface. This surface also shows exposed residues characteristic for RNA-binding proteins as present in the ribosomal proteins S17, L6, and L9 (36–38). It contains aromatic and basic side chains that can interact directly with the bases and the backbone of RNA, as well as residues with small side chains and glycines that would allow a close approach between protein and RNA. In this context

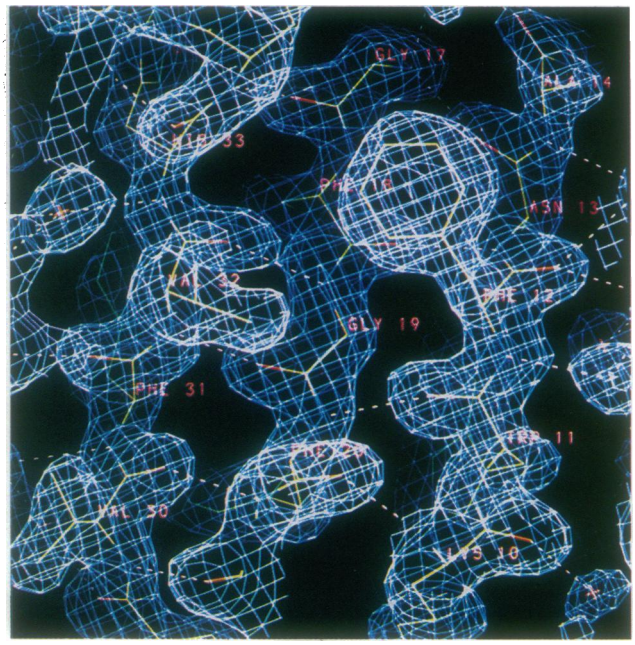


FIG. 5. Electron density ($2F_o - F_c$) of the crystal showing parts of strands $\beta$1–$\beta$3. The map is contoured at 1 times the rms density above the mean density and is superimposed with the refined model. Broken lines indicate hydrogen bonds.

it may be noted that the molecular architecture of S17 (36) has striking similarity with that of the cold shock domain.

With the identification of the nucleic acid-binding site by considering the electrostatic properties of the molecule and by comparison with the OB fold, we suggest that the side chains of $Trp^{11}$, $Phe^{18}$, $Phe^{20}$, $Phe^{31}$, $Phe^{34}$, and $Tyr^{42}$ together with the basic side chains of $Lys^{10}$, $Lys^{16}$, $His^{33}$, and $Lys^{60}$ of CspA, as well as the corresponding side chains in *B. subtilis* CspB, are involved in DNA binding and possibly also in RNA binding. All these residues are located at the binding site and most of them are highly conserved within the cold shock proteins (Fig. 1): $Lys^{10}$, $Trp^{11}$, $Phe^{20}$, $Phe^{31}$, and $His^{33}$ are entirely conserved, $Phe^{18}$ is replaced by Tyr in the Y-box factors, and $Lys^{60}$ is replaced by Arg in CspB, whereas a residue in this position is omitted in the sequence of *Streptomyces clavuligerus* Csp7. Within the cold shock proteins (but not within the Y-box factors) $Lys^{16}$ is conserved and either Phe or Tyr is found in positions 34 and 42 of these proteins. This high conservation underlines the functional importance of these residues.

Based on the identification by NMR spectroscopy of residues that show a large chemical shift upon binding of single-stranded DNA (14), the location of these residues in the three-dimensional structure of CspA has been analyzed. It is obvious that nearly all lie within the region that has been identified as the binding site (Figs. 4–6). It should be noted that shifts in protein resonances are not necessarily due to direct interaction with DNA but could also reflect structural changes occurring upon binding. Residue $Phe^{12}$, for instance, which appears to engage in direct contact with DNA in the NMR study, is shielded from solvent in the crystal structure of CspA, as is the corresponding $Phe^{9}$ in CspB. This residue might thus sense a small structural perturbation caused by DNA binding.

Similar function of the bacterial cold shock proteins and the cold shock domain of the eukaryotic Y-box factors has been described (39), indicating a remarkable conservation of function during evolution. Therefore, our conclusions derived for the bacterial proteins might, at least in part, also be valid for the Y-box factors. An aspect that warrants further examination is the dimerization observed in both crystal forms of *B. subtilis* CspB, but not for *E. coli* CspA. Dimerization and oligomerization as described for the Y-box factors (39) could be essential for all members of the cold shock domain family to achieve full biological activity.

Most of the experimental work described here was carried out at the Institut für Kristallographie at the Freie Universität, Berlin. Prof. W. Saenger of the Freie Universität is thanked for providing equipment and continued support. Work in the laboratory of U.H. was supported by the Deutsche Forschungsgemeinschaft through Grant SFB 344 D/3 and by the Fonds der Chemischen Industrie, and work

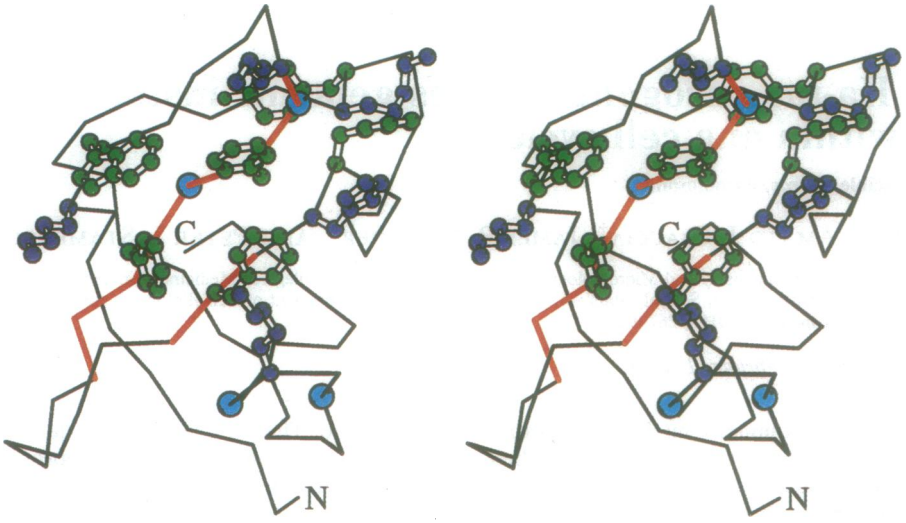


FIG. 6. Stereoview of the presumed nucleic acid-binding surface of CspA. The RNP1 motif and the rudimentary RNP2 motif are highlighted in red, conserved glycine residues are indicated by cyan dots at their $C^{\alpha}$ positions, basic side chains are blue, and aromatic side chains are green.

in the laboratory of M.I. was supported by the National Institutes of Health (GM19043).

1. Jones, P. G., VanBogelen, R. A. & Neidhardt, F. C. (1987) *J. Bacteriol.* **169**, 2092–2095.
2. Goldstein, J., Pollitt, N. S. & Inouye, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 283–287.
3. La Teana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C. L. & Gualerzi, C. O. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10907–10911.
4. Jones, P. G., Krah, R., Tafuri, S. A. & Wolffe, A. P. (1992) *J. Bacteriol.* **174**, 5798–5802.
5. Wistow, G. (1990) *Nature (London)* **344**, 823–824.
6. Wolffe, A. P., Tafuri, S., Ranjan, M. & Familari, M. (1992) *New Biol.* **4**, 290–298.
7. Murray, T. M., Schiller, D. L. & Franke, W. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11–15.
8. Tafuri, S. R., Familari, M. & Wolffe, A. P. (1993) *J. Biol. Chem.* **268**, 12213–12220.
9. Tanabe, H., Goldstein, J., Yang, M. & Inouye, M. (1992) *J. Bacteriol.* **174**, 3867–3873.
10. Lee, S. J., Xie, A., Jiang, W., Etchegaray, J.-P., Jones, P. G. & Inouye, M. (1994) *Mol. Microbiol.*, in press.
11. Willimsky, G., Bang, H., Fischer, G. & Marahiel, M. A. (1992) *J. Bacteriol.* **174**, 6326–6335.
12. Schindelin, H., Marahiel, M. A. & Heinemann, U. (1993) *Nature (London)* **364**, 164–168.
13. Schnuchel, A., Wiltscheck, R., Czisch, M., Herrler, M., Willimsky, G., Graumann, P., Marahiel, M. A. & Holak, T. A. (1993) *Nature (London)* **364**, 169–171.
14. Newkirk, K., Feng, W., Jiang, W., Tejero, R., Emerson, S. D., Inouye, M. & Montelione, G. T. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5114–5118.
15. The SERC Collaborative Computing Project (1979) *CCP4* (Science and Engineering Research Council, Warrington, U.K.), Project #4.
16. Brünger, A. T. (1990) *Acta Crystallogr. A* **46**, 46–57.
17. Brünger, A. T., Kuriyan, J. & Karplus, M. (1987) *Science* **235**, 458–460.
18. Brünger, A. T., Krukowski, A. & Erickson, J. (1990) *Acta Crystallogr. A* **46**, 585–593.
19. Hendrickson, W. A. (1985) *Methods Enzymol.* **115**, 252–270.
20. Finzel, B. C. (1987) *J. Appl. Crystallogr.* **20**, 53–55.
21. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291.
22. Luzzati, V. (1952) *Acta Crystallogr.* **5**, 802–810.
23. Kraulis, J. P. (1991) *J. Appl. Crystallogr.* **24**, 946–950.
24. Hutchinson, E. G. & Thornton, J. M. (1993) *Protein Eng.* **6**, 233–245.
25. Kabsch, W. & Sander, C. (1983) *Biopolymers* **22**, 2577–2637.
26. Murzin, A. G. (1993) *EMBO J.* **12**, 861–867.
27. Murzin, A. G. & Chothia, C. (1992) *Curr. Opin. Struct. Biol.* **2**, 895–903.
28. Hynes, T. R. & Fox, R. O. (1991) *Proteins Struct. Funct. Genet.* **10**, 92–105.
29. Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C. & Moras, D. (1993) *Nature (London)* **362**, 181–184.
30. Brayer, G. D. & McPherson, A. (1983) *J. Mol. Biol.* **169**, 565–596.
31. Folkers, P. J. M., van Duynhofen, J. P. M., Jonker, A. J., Harmsen, B. J. M., Konings, R. N. H. & Hilbers, C. W. (1991) *Eur. J. Biochem.* **202**, 349–360.
32. Graumann, P. & Marahiel, M. A. (1994) *FEBS Lett.* **338**, 157–160.
33. Landsman, D. (1992) *Nucleic Acids Res.* **20**, 2861–2864.
34. Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. (1990) *Nature (London)* **348**, 515–520.
35. Mattaj, I. W. (1993) *Cell* **73**, 837–840.
36. Golden, B. L., Hoffman, D. W., Ramakrishnan, V. & White, S. W. (1994) *Biochemistry* **32**, 12812–12820.
37. Golden, B. L., Ramakrishnan, V. & White, S. W. (1993) *EMBO J.* **12**, 4901–4908.
38. Hoffman, D. W., Davies, C., Gerchman, S. E., Kycia, J. H., Porter, S. J., White, S. W. & Ramakrishnan, V. (1994) *EMBO J.* **13**, 205–212.
39. Tafuri, S. R. & Wolffe, A. P. (1992) *New Biol.* **4**, 349–359.