J Biomol NMR (2009) 43:255–259
DOI 10.1007/s10858-009-9298-y

NMR STRUCTURE NOTE

Solution structure of GSP13 from *Bacillus subtilis* exhibits an S1 domain related to cold shock proteins

Wenyu Yu · Jicheng Hu · Bingke Yu · Wei Xia · Changwen Jin · Bin Xia

Received: 2 December 2008 / Accepted: 22 December 2008 / Published online: 17 January 2009
© Springer Science+Business Media B.V. 2009

Abstract GSP13 encoded by gene *yugI* is a σ^B-dependent general stress protein in *Bacillus subtilis*, which can be induced by heat shock, salt stress, ethanol stress, glucose starvation, oxidative stress and cold shock. Here we report the solution structure of GSP13 and it is the first structure of S1 domain containing protein in *Bacillus subtilis*. The structure of GSP13 mainly consists of a typical S1 domain along with a C-terminal 50-residue flexible tail, different from the other known S1 domain containing proteins. Comparison with other S1 domain structures reveals that GSP13 has a conserved RNA binding surface, and it may function similarly to cold shock proteins in response to cold stress.

Keywords GSP · GSP13 · *yugI* · S1 domain · Cold shock protein

Biological context

As a soil bacterium, *Bacillus subtilis* often faces energy starvation and a variety of sudden environmental stresses, such as heat, cold, acid, ethanol, and salt stress etc. To survive under such conditions, *Bacillus subtilis* has developed complicated adaption networks, in which the alternative transcription factor σ^B plays a particularly important role (Haldenwang and Losick 1980). σ^B is activated in response to stress by complicated signal transduction pathways, and then it coordinates the expression of a large number of proteins (Hecker et al. 2007). These proteins are called σ^B-dependent general stress proteins (GSPs). So far, more than 150 σ^B-dependent genes have been identified by proteome studies and transcription profiling. However, only less than 20 GSPs have been found to have defined functions (Hecker et al. 2007).

GSP13, encoded by *Bacillus subtilis* gene *yugI*, is a typical member of GSPs (Bernhardt et al. 1997). Homology search with BLAST shows that GSP13 seems to be specific for gram-positive bacteria. It was reported that GSP13 protein level is up-regulated after heat shock, salt stress, ethanol stress, glucose starvation, oxidative stress and ammonium starvation (Antelmann et al. 1997; Bernhardt et al. 1997). Also, the mRNA level and the protein level of GSP13 show significant increases after cold shock (Kaan et al. 2002). In addition, GSP13 was found to be degraded when *Bacillus subtilis* enters the stationary phase induced by glucose-starvation, indicating that the protein is tightly regulated in vivo. Moreover, GSP13 was reported to associate with ribosome only during the exponential growth phase, while its function in ribosome remains unclear (Nanamiya et al. 2004).

GSP13 has 130 amino acid residues, and sequence analysis suggests that it contains an S1 domain (residues 8–77) and a C-terminal charge-rich region (residues 78–130 with 10 Lys and 7 Glu). The S1 domain is found in a large number of RNA-associated proteins. Most of the S1 domain containing proteins are multi-domain proteins with one or more S1 domains in addition to other functional domains, while small proteins ($\sim$70 residues) like bacterial and chloroplastic translation initiation factor 1 (IF1) are constituted of only a single S1 domain. The C-terminal charge-rich region shows no homology with other proteins.

W. Yu · J. Hu · B. Yu · W. Xia · C. Jin · B. Xia (✉)
Beijing Nuclear Magnetic Resonance Center, College of Life Sciences, Peking University, Beijing 100871, China
e-mail: binxia@pku.edu.cn

C. Jin · B. Xia
College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Here, we report the solution structure of GSP13. It is the first structure of S1 domain containing proteins in *Bacillus subtilis*. The functional implication of the structure is also discussed.

Methods and results

NMR spectroscopy

The protein was expressed and purified, and the assignments of ^{1}H, ^{15}N and ^{13}C chemical shifts of GSP13 were finished as reported (Yu et al. 2008). The NMR sample containing 1.5 mM $^{15}N/^{13}C$ labeled GSP13 was in 50 mM potassium phosphate buffer (pH 7.0) with 90% H_2O/10% D_2O, along with 20 mM NaCl, 1 mM EDTA, 20 mM DTT, 10 mM PMSF, and 0.01% sodium azide. Nuclear Overhauser effects (NOEs) constraints for structural calculations were obtained from the following experiments: 3D ^{15}N NOESY-HSQC (mixing time of 110 ms), 3D ^{13}C NOESY-HSQC (mixing time of 110 ms) selective for aromatic carbons and 3D ^{13}C NOESY-HSQC (mixing time of 110 ms) selective for aliphatic carbons. All NMR spectra were acquired on Bruker Avance 600 MHz spectrometers at 298 K. All spectra were processed with NMRPipe (Delaglio et al. 1995) and analyzed with NMRView (Johnson and Blevins 1994). The ^{1}H dimension was referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), and the ^{13}C and ^{15}N dimensions were indirectly referenced to DSS.

Structure determination

The structures were calculated using the program package CYANA (Herrmann et al. 2002) and refined with AMBER (Pearlman et al. 1995). The distance restraints were obtained from the analysis of NOESY data using SANE (Duggan et al. 2001). Dihedral angles restraints were obtained based on the analysis of the backbone chemical shifts with TALOS (Cornilescu et al. 1999). 200 structures were first calculated by CYANA, and the 100 structures with the lower target function values were selected as the initial structures and further refined using AMBER. Finally, an ensemble of 20 structures with lowest energies that had no NOE distance violations >0.2 Å and dihedral angle violations >5° was obtained. The chemical shift assignments of the protein have been finished as reported. The structure was finally calculated based on 2777 NOE restraints, 83 dihedral angle restraints, and 468 chirality restraints. Structural statistics and the ramachandran analysis of the ensemble of 20 structures are listed in Table 1. The mean structure of the 20 structures was calculated with the software SUPPOSE (Maiti et al. 2004). The ensemble of the structures was analyzed by PROCHECK_NMR (Laskowski et al. 1996) and program MOLMOL (Koradi et al. 1996).

Table 1 Restraints and structure statistics

Conformational restraints		
NOE restraints	2777	
Intra-residue	760	
Sequencetial	461	
Medium	185	
Long-range	453	
Ambigous	918	
Dihedral angle restraints	83	
phi	40	
psi	43	
Structure statistics[a]		
Violations	Numbers	Maximum violation
Distance restraints (>0.2 Å)	0	0.16 Å
Dihedral angle restraints(>5°)	0	
RMSD (Å)	Residues 10–75	Regular secondary structure regions
Backbone heavy atoms (Å)	0.45	0.26
All heavy atoms (Å)	1.07	0.75
Ramachandran plot		
Most favored regions (%)	81.3	
Additionally allowed regions (%)	17.8	
Generously allowed regions (%)	0.7	
Disallowed regions (%)	0.3	

[a] Structure statistics is based on the 20 conformers with lowest AMBER energies selected from 100 calculated structures

Overall structure description

The solution structure of the GSP13 is showed in Fig. 1a and b, and structure parameters are summarized in Table 1. Coordinates of the ensemble of twenty structures have been deposited in the Protein Data Bank (http://www.rcsb.org/pdb/) with accession code 2k4k. The overall fold of GSP13 consists of a S1 domain and a long C-terminal flexible tail (Fig. 1a). The GSP13 S1 domain contains a five-stranded antiparallel β-sheet (β_1: residues 10–19; β_2: residues 22–28; β_3: residues 31–36; β_4: residues 57–66; β_5: residues 71–75) and two small 3_{10} helices (Fig. 1b). The five β-strands are arranged in a classical 1-2-3-5-4-1 topology to form a closed β-barrel, which is the characteristic of the OB-fold. Two β-bulges are located around residues 15/16 and 62/63, allowing β_1 and β_4 to form a closed barrel. Two 3_{10} helices, α_1 (residues 37–39) and α_2 (residues 48–50), exist between β_3 and β_4. The difference between GSP13 S1 domain and classic OB-fold is that it has an extra 3_{10} helix

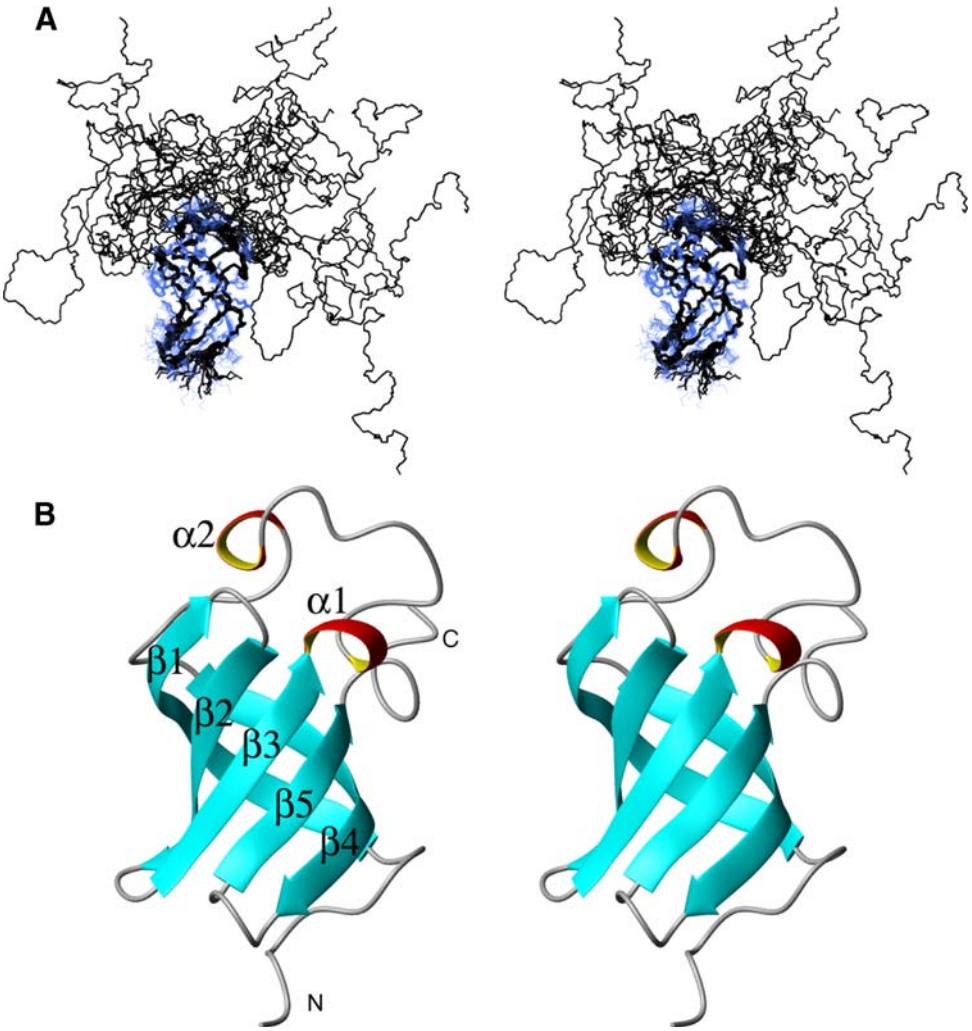

Fig. 1 Stereo-view representation of the solution structure of the GSP13. **a** Superposition of twenty lowest energy structures of GSP13 with side chains of residues 1–75 superimposed. **b** Ribbon diagram of the mean structure of the twenty lowest energy structures (residues 1–82)

right after strand β_3. The long C-terminal tail (residues 76–130) after β_5 is rather flexible, as indicated by the fact that no inter-residue NOE is observed for most of the residues. So far, several structures of S1 domain containing proteins have been determined, however, the structure of GSP13 is different from all of them in terms that it contains a single S1 domain along with a long flexible tail (Bycroft et al. 1997; Johnson et al. 2008).

Discussion and conclusions

RNA binding site

During the purification of GSP13, it was found that GSP13 is eluted from the anion-exchange column as a single peak when the bacterial lysate is pre-treated with RNase A. Without RNase A treatment, GSP13 was found in multiple peaks of the elution profile (data not shown). This indicates that the protein should bind RNA with different lengths in the cell lysate.

Considerable work has been carried out to investigate on the RNA binding mechanism of the S1 domains. In *E. coli* PNPase S1 domain, several residues F19, F22, H34, D64 and R68, clustered on the surface that are conserved in many S1 domains, are believed to be involved in RNA binding (Bycroft et al. 1997). In Tex S1 domain, mutation experiment showed that F668, F671, H683, and R718 are involved in RNA binding (Johnson et al. 2008). Structure alignment shows that the corresponding residues in GSP13, Y21, F24, H34, D66 and K71, are conserved, and they should be involved in RNA binding of GSP13 (Fig. 2a and b).

Functional implication

Among the superfamily of OB-fold proteins, cold shock proteins (CSPs) have the most similar structures to the S1 domain. Different from other OB-fold proteins, they all have a 3_{10} helix at the end of the β_3 strand, indicating they are likely to be the closest relatives in the superfamily (Bycroft et al. 1997). Moreover, the experiment performed in *E. coli*, in which the cold sensitivity by deletion of CspA

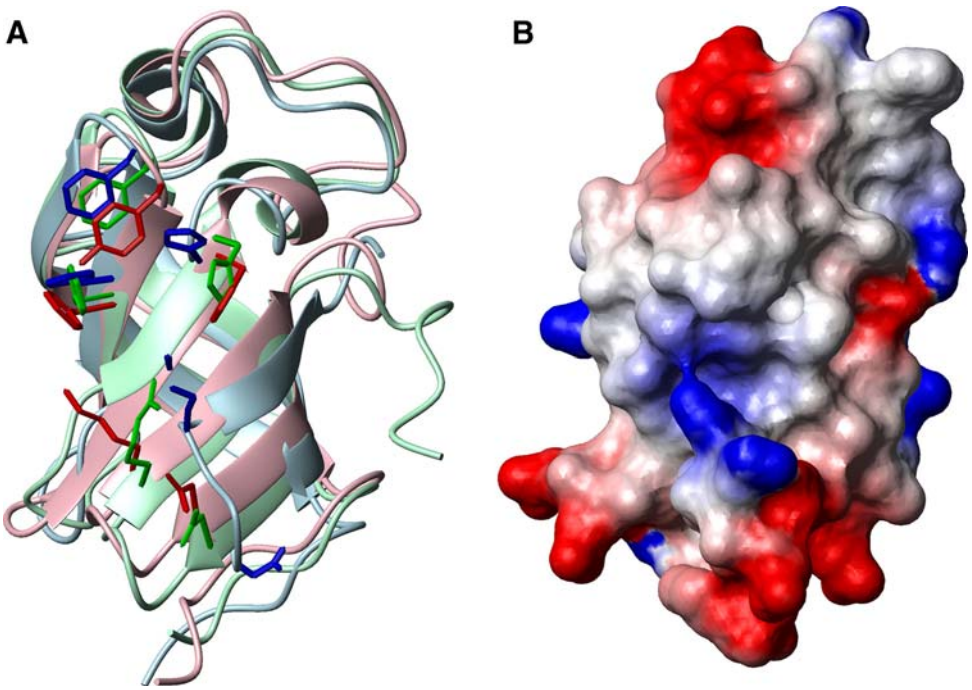

Fig. 2 **a** Overlay of ribbon diagrams of GSP13 S1 domain (pink), PNPase S1 domain (light blue) and Tex S1 domain (pale green), with side chains of residues Y21, F24, H36, D66 and K71 of GSP13 (red), which are believed to be involved in RNA binding shown. The corresponding residues in PNPase (blue) and Tex S1 domains (green) are also displayed. **b** Electrostatic potential surface of GSP13 S1 domain with the same orientation to the ribbon show

family was suppressed by PNPase S1 domain, indicated that the S1 domain and CSPs may have common function properties in addition to similar tertiary structures (Xia et al. 2001). Considering the high sequence identity and structure similarity between PNPase S1 and GSP13 S1, it is quite possible that the GSP13 S1 domain also has similar function to CSPs as PNPase S1 domain does. In addition, *E. coli* IF1, which is actually a single S1 domain (total 72 residues), can also complement the function of CSPs when expressed in *B. subtilis* (Weber et al. 2001). Therefore, we propose that the GSP13 can act similarly to cold shock proteins in response to cold stress.

Acknowledgments All NMR experiments were carried out at the Beijing NMR Center. We thank Dr. Xiao-Dong Su for providing us the genome library of *Bacillus subtilis*. This research was supported by Grant 30125009 to BX from the National Science Foundation of China and grant 2006AA02A323 to CJ from the 863 Program.

References

Antelmann H, Bernhardt J, Schmid R, Mach H, Völker U, Hecker M (1997) First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. Electrophoresis 18:1451–1463

Bernhardt J, Völker U, Völker A, Antelmann H, Schmid R, Mach H, Hecker M (1997) Specific and general stress proteins in *Bacillus subtilis*—a two-deimensional protein electrophoresis study. Microbiology 143:999–1017

Bycroft M, Hubbard TJ, Proctor M, Freund SM, Murzin AG (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. Cell 88:235–242

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13:289–302

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe—a multidimensional spectral processing system based on unix pipes. J Biomol NMR 6:277–293

Duggan BM, Legge GB, Dyson HJ, Wright PE (2001) SANE (Structure assisted NOE evaluation): an automated model-based approach for NOE assignment. J Biomol NMR 19:321–329

Haldenwang WG, Losick R (1980) Novel RNA polymerase sigma factor from *Bacillus subtilis*. Proc Natl Acad Sci USA 77:7000–7004

Hecker M, Pane-Farre J, Völker U (2007) SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. Annu Rev Microbiol 61:215–236

Herrmann T, Guntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 319:209–227

Johnson BA, Blevins RA (1994) NMRView: a computer program for the visualization and analysis of NMR data. J Biomol NMR 4:603–614

Johnson SJ, Close D, Robinson H, Vallet-Gely I, Dove SL, Hill CP (2008) Crystal structure and RNA binding of the Tex protein from *Pseudomonas aeruginosa*. J Mol Biol 377:1460–1473

Kaan T, Homuth G, Mader U, Bandow J, Schweder T (2002) Genome-wide transcriptional profiling of the *Bacillus subtilis* cold-shock response. Microbiology 148:3441–3455

Koradi R, Billeter M, Wüthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14(51–55):29–32

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486

Maiti R, Van Domselaar GH, Zhang H, Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Res 32:W590–594

Nanamiya H, Akanuma G, Natori Y, Murayama R, Kosono S, Kudo T, Kobayashi K, Ogasawara N, Park SM, Ochi K, Kawamura F (2004) Zinc is a key factor in controlling alternation of two types of L31 protein in the *Bacillus subtilis* ribosome. Mol Microbiol 52:273–283

Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. Comp Phys Commun 91:1–41

Weber MH, Beckering CL, Marahiel MA (2001) Complementation of cold shock proteins by translation initiation factor IF1 in vivo. J Bacteriol 183:7381–7386

Xia B, Ke H, Inouye M (2001) Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in *Escherichia coli*. Mol Microbiol 40:179–188

Yu W, Yu B, Hu J, Xia W, Jin C, Xia B (2008) ^{1}H, ^{13}C, and ^{15}N resonance assignments of a general stress protein GSP13 from *Bacillus subtilis*. Biomolecular NMR Assignments 2:163–165