Solution Structure and Backbone Dynamics of the Biotinylation Domain of *Helicobacter pylori* Biotin-carboxyl Carrier Protein

Jinwon Jung,[a] Chul-Jin Lee, Young Ho Jeon,[†] Chaejoon Cheong,[†] and Weontae Lee[*]

*Department of Biochemistry and Protein Network Research Center, College of Science, Yonsei University,
Seoul 120-749, Korea. *E-mail: wlee@spin.yonsei.ac.kr
[†]Magnetic Resonance Team, Korea Basic Science Institute (KBSI), Ochang, Chungbuk 363-883, Korea
Received November 25, 2007*

Acetyl-CoA carboxylase (ACC) is an excellent candidate for antibiotics drug target, which mediates malonyl-CoA synthesis from acetyl-CoA through acetylation process. It is also involved in the committed step of fatty acid synthesis which is essential for living organisms. We have determined the three dimensional structure of C terminal domain of HP0371, biotin-carboxyl carrier protein of *H. pyroli*, in solution state using heteronuclear multi-dimensional NMR spectroscopy. The structure of HP0371 shows a flatten β-sheet fold which is similar with that of *E. coli*. However, the sequence and structure of protruding thumb are different with that of *E. coli* and the thumb shows different basis of structural rigidity based on backbone dynamics data.

Key Words : BCCP, Backbone dynamics, Solution structure, NMR spectroscopy

Introduction

Helicobacter pylori (*H. pylori*) is a gram-negative micro-aerophilic bacterium that lives in the stomach and duodenum. Since it survives in the harsh acidic environment of the stomach and causes the duodenal ulcer, gastric cancer and peptic ulcer disease, it has been known as a possible pathogen of many gastric diseases. Recently, the genome sequence of *H. pylori* has been determined and the genomic function could be deduced from sequence information.[1] Recently, more evidences about relationship between *H. pylori* and gastric cancer have been reported.[2] Therefore, it is of interest to develop a drug to treat *H. pylori* to prevent cancer and other gastric disease.

Acetyl-CoA carboxylase (ACC)[3] is considered as a candidate for antibiotics drug target. Acetyl-CoA carboxylase (ACC) mediates malonyl-CoA synthesis from acetyl-CoA through acetylation process and it is the committed step of fatty acid synthesis which is essential for living organisms. It has been determined that the molecular property of bacterial ACC is different from that of eukaryote. For many bacteria, ACC is a multi-subunit enzyme whereas it is a single-chained multi-domain enzyme for human, plant, and fungi. Therefore, it becomes a candidate of antibiotic target.

ACC is functionally divided into three parts: biotin carboxyl carrier protein (BCCP), biotin carboxylase (BC), and carboxyl transferase (CT). For bacteria, these parts are separated in three distinct polypeptide chains. Biotin carboxyl carrier protein is a subunit of ACC, which is covalently attached to biotin, the coenzyme that carry activated carboxyl group and transfer it to acetyl-CoA. The structure of C terminal domain BCCP of *E. coli* was determined by X-ray crystallography[4] and NMR spectrsocopy.[5,6]

H. pylori BCCP shares 33% of sequence homology with that of *E. coli*. Because BCCP is essential protein and has a few interacting proteins, it could be a possible antibiotic target. To develop drugs that interfere interactions of BCCP, it is necessary to know its molecular structure and understand its molecular behavior. In this work, we present the solution structure and dynamics of the biotinylation domain of *H. pylori* BCCP by nuclear magnetic resonance (NMR) spectroscopy.

Materials and Methods

Protein Purification and Sample Preparation. Biotinylation domain of hpBCCP (81-156) was constructed into pET-15b vector (Novagen Inc.) and expressed in *E. coli* BL21(DE3) cells with isotope-enriched M9 media (1 g/L $^{15}NH_4Cl$, 1 g/L ^{13}C-glucose). Protein was induced by addition of 1 mM IPTG at the mid of exponential growth phase, and cells were grown for additional 4 hrs at 37 °C. Harvested cells were lysed by sonication. The lysate was loaded on Ni-NTA column (Qiagen). After washing column with 20 volumes of buffer solution (pH 7.4, 50 mM sodium phosphate, 1 mM b-mercaptoethanol) containing 5 mM imidazole. 200 unit of bovine thrombin was loaded for on-column cleavage of histag. Protein solution was circulated with a peristaltic pump for 24 hrs at 25 °C and followed by gel filtration using Superdex G-75 column. Protein sample was concentrated to 1-1.5 mM in 250 uL with 50 mM Sodium phosphate buffer (pH 6.5) containing 150 mM NaCl, 1 mM DTT, 100 μM DSS, and 0.01% NaN3.

NMR Spectroscopy. All NMR spectra were recorded at 25 °C on Bruker DRX500 spectrometer with triple-resonance probes equipped with triple axis gradient coil or Varian 500 INOVA Unity[+] spectrometer. HNCACB, CBCA(CO)NH, and HNCA experiments were collected to assign backbone resonances. HNHA, HCCH-TOCSY, and 15N-edited TOCSY

[a]Present address: Department of Structural Biology, School of Medicine, University of Pittsburgh, 3501 fifth avenue, Pittsburgh, PA, USA, 15260

experiments were performed for sidechain resonance assignment. To extract distance information, ^{15}N-edited 3D NOESY-HSQC (τ_m = 200 ms) and ^{13}C-edited 3D HMQC-NOESY (τ_m = 200 ms) data were collected. For 3D HCCH-TOCSY and ^{13}C-edited HMQC-NOESY experiment. the protein sample was dissolved in 99% $D_2O/10\%$ H_2O solution.

Relaxation Measurement. Longitudinal (R_1) and transversal (R_2) relaxation data for the backbone ^{15}N nuclei for hpBCCP were recorded as 1024 × 128 data sets with 16 scans per point, using a relaxation delay of 3s. Eight different values for the relaxation time were recorded with different durations of the T delay. T_1 = 52. 111, 178, 255. 346. 458, 601 and 804 ms and T_2 = 16.51. 33.02, 49.5. 66.05. 82.56. 115.58. 132.10, 148.61 and 165.12 ms. To permit estimation of noise levels, duplicate spectra were recorded for T = 255 ms (T_1 spectra) and T = 66.05 ms (T_2 spectra). To remove the cross correlation effects between ^{15}N-^{1}H dipolar and ^{15}N CSA relaxation mechanisms. ^{1}H 180° pulses were inserted during the T relaxation times.[7] ^{15}N-{^{1}H} steady-state heteronuclear NOE (XNOE)[8] data were obtained using a relaxation delay of 5 s. yielding data sets of 1024 × 128 data sets after accumulation of 128 scans per point. Both saturated and unsaturated spectrum were acquired in interleaved manner and separated with XWINNMR program.

Data Processing and Calculation. All data were processed with NMRPipe[9] and analyzed by Sparky.[10] Backbone assignment was performed by AutoAssign program.[11] Automated NOESY assignment for structure calculation was done by program CYANA version 2.0[12] on a Linux cluster with 15 nodes. NOE cross-peaks were assigned in the aforesaid NOESY spectra by a combination of interactive procedures and the automated method resulting in a total number of 671 unambiguously assigned upper distance limits. TALOS program[13] was used to generate PHI and PSI angle constraints from chemical shift list and 52 backbone torsion angle constraints (26 for ϕ and 26 for φ) were derived. The presence of hydrogen bond was detected from H/D exchange experiment. and donors of hydrogen bond were found by NOE patterns and initial fold of structures.[14]

Relaxation Analysis. For T_1 and T_2 relaxation measurements. cross-peak intensities were measured from peak heights and fit to an exponential decay function by nonlinear least-squares analysis using the Levenburg-Marquardt method implemented by CurveFit program from Arthur G Palmer (http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer). NOE values were determined as the ratios of the peak intensities measured from spectra acquired with and without saturation during the recycle delay. The uncertainties in the measured peaks were set equal to the root-mean-square baseline noise in the spectra.[15] The uncertainties of the NOE value was determined on the basis of measured background noise levels using the following relationship:[8] σ_{NOE}/NOE = $((\sigma_{Isat}/I_{sat})^2 + (\sigma_{Iunsat}/I_{unsat})^2)^{1/2}$ where I_{sat} and I_{unsat} represent the measured intensities of a resonance in the presence and absence of proton saturation. respectively. The uncertainties of these values were calculated from the root-mean-square noise of background regions and represented by σ_{Isat} and

σ_{Iunsat}. The ^{15}N heteronuclear relaxation rates were interpreted using the program TENSOR2.[16,17] which uses the description of the molecular diffusion derived by Woessner, in combination with the Lipari-Szabo model-free analysis of local flexibility.[18] The τ_m and rotational diffusion tensor were determined after 500 step Monte-Carlo simulation. Before determining the τ_m and the rotational diffusion tensor, the R_2/R_1 ratios are trimmed to exclude either 10% of all extreme values (5% highest and 5% lowest values). Motional parameters were determined after 500 steps of Monte-Carlo simulation and followed by model selection using TENSOR2 program.

Results

Solution Structure. Solution structure of biotinylation domain of hpBCCP was calculated by NMR data. A total of 671 distance restraints were assigned by CYANA program. 46 hydrogen bond constraints were derived and 52 dihedral angle constraints were calculated by TALOS analysis. The 20 lowest target function structures show a mean backbone
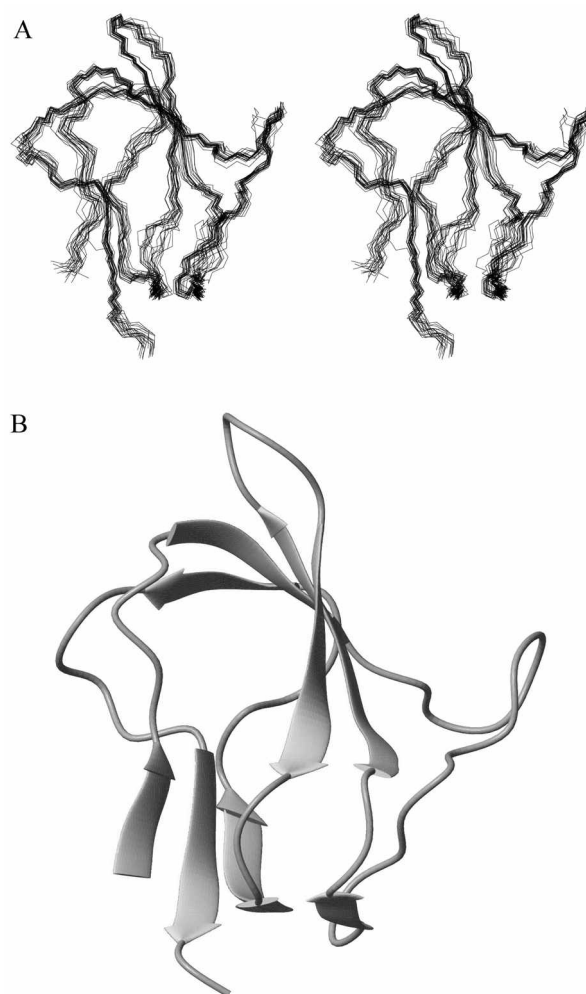
Figure 1. Structure of the biotinyl domain of *H. pylori* acetyl-CoA carboxylase biotin carboxyl carrier protein (hpBCCP). (a) Superposition of backbone traces from the final ensemble of 20 solution structures. (b) Ribbon diagram of lowest target function structure.

Table 1.

Distance restraints	
All	671
Short ($|i-j|<=1$)	382
Medium ($1<|i-j|<5$)	45
Long-range ($|i-j|>=5$)	244
Hydrogen bonds[a]	46
Dihedral restraints	
All	52
Phi	26
Psi	26
Mean CYANA target function	0.91 ± 0.07
RMS deviations from the average coordinate[b]	
Backbone atoms (N,Ca,C',O)	$0.66 \text{ Å} \pm 0.13 \text{ Å}$
All heavy atoms	$1.13 \text{ Å} \pm 0.12 \text{ Å}$
Ramachandran statistics	
Residues in most favored region	77.5%
Residues in additional allowed regions	20.6%
Residues in generously allowed regions	1.6%
Residues in disallowed regions	0.2%

[a]Two restraints per one hydrogen bond. [b]Secondary structured region only.

RMSD of 0.66 Å (Figure 1A) and the value of mean target function is 0.91 ± 0.07 Å (Table 1). The structure of biotinylation domain of hpBCCP forms a flattened β-barrel structure with two antiparallel β-sheets comprised by 8 β-strands: $\beta1(79-81)$, $\beta2(87-90)$, $\beta3(106-107)$, $\beta4(111-117)$, $\beta5(123-127)$, $\beta6(131-136)$, $\beta7(142-144)$ and $\beta8(151-154)$ (Figure 1B). First b-sheet comprised by $\beta2$, $\beta4$, $\beta5$, and $\beta7$ includes Lys120, which is the biotinylation site. Second β-sheet is composed by four strands ($\beta1$, $\beta3$, $\beta7$ and $\beta8$). Two β-sheets are stabilized by hydrophobic cluster composed by residues, Val81, Val113, Ile151, Phe89, Leu150 and Ile125. By primary sequence analysis, most of these hydrophobic residues are well conserved for both hpBCCP and *E. coli*. BCCP. Lys120 is highly conserved among BCCP homologues and known as the biotinylation site (Figure 2) and it is located at the tip of tight β-turn. Especially, most of highly conserved residues among biotin carrier proteins are found on β-hairpin between $\beta4$ and $\beta5$. Two bulge structures are found near Val113 and Val134.

There is a protruding thumb from A92 to P99. This loop starts with highly conserved Pro-Ser-Pro sequence and ends with a proline residue. In the determined structure, the thumb shows a well defined structure despite the lack of regular secondary structure. Backbone RMSD is 0.191 Å in

this region (A92-P99).

Comparison of Biotinylation Domain of *H. pylori* and *E. coli* BCCP. When it was compared with the structure of *E. coli* BCCP (PDB:1BDO, ecBCCP). Ca RMSD was 2.699 Å. Overall structure was similar but protruding thumb structures were different. If RMSD is calculated without protruding thumb, the value was 1.941 Å. On ecBCCP, first half of thumb structure with highly conserved Pro95-Ser96-Pro97 sequence stretched inside and it is toward the tip of biotinylation site. However, the corresponding part in hpBCCP, Pro93-Ser94-Pro95, faced outside oppositely. There is a difference in hydrophobic core. Cys116 of ecBCCP is replaced by Gly114 in hpBCCP. The empty space that is caused by replacing with glycine that has no sidechain is filled and compensated in hpBCCP by changing neighbor Leu115 to Val113, which is shorter but bulkier near by backbone. This change of residue from longer leucine to shorter valine is compensated again by changing Leu153 to Ile151 in hpBCCP.

Backbone Dynamics of hpBCCP. The NMR relaxation parameters, longitudinal relaxation rate constants (R_1), transverse relaxation rate constants (R_2) and steady-state $\{^1H\}$-^{15}N nuclear Overhauser effects (NOE) were obtained by the analysis of two-dimensional proton-detected heteronuclear NMR spectroscopy for 69 residues. The relaxation data are plotted in Figure 3A-C. The mean values for R_1, R_2, and NOE were 2.68 s^{-1}, 7.2 s^{-1}, and 0.72, respectively. Using 10% trimmed R_2/R_1 ratio, global τ_m was determined as 4.78 ns. The model-free approach was used to analyze the relaxation data. The calculation of the model-free parameters from the measured relaxation data was performed using TENSOR2 software. Of the 69 backbone ^{15}N nuclear spins, relaxation data for a total of 34, 9, 5, 4, and 0 backbone ^{15}N nuclei were fit by models 1-5 where model 1 fits only the generalized order parameter (S^2); model 2 fits S^2 and the effective internal correlation time; model 3 fits S^2 and the chemical exchange term (R_{ex}); model 4 fits S^2, τ_e, and R_{ex} terms; and model 5 fits order parameters of internal motions on both fast and slow time scales (S_f^2 and S_s^2), and also τ_e. 6 residues (Phe89, Gly96, Asp105, Met119, Asn123, Gly147, Lys155) were not fit to any model. Fitted values of the model-free parameters of BCCP are presented in Figure 3D-E. Gly96, Asp105, Met119, and Lys155 were assigned to model 4. Phe89 and Asn123 were assigned to model 5, and Gly147 to model 1. However, these residues did not satisfy statistical significance.

Motional parameters were obtained from 63 residues (76
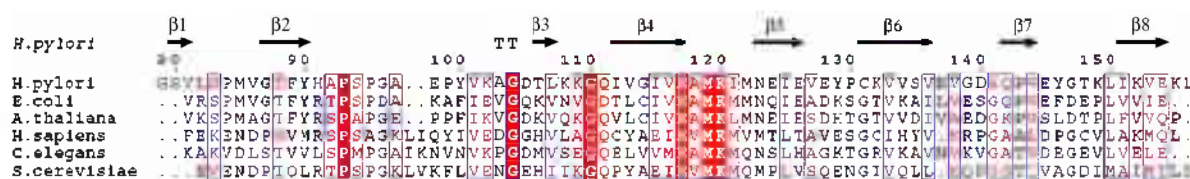
Figure 2. Sequence alignment of hpBCCP with homologues from various species. Red box is identical sequence and red character is similar sequence. Secondary structure of hpBCCP is shown on top of the alignment. Arrows are β-strand and 'T' is turn. This alignment is made by ClustalW[22] and this figure is generated with ESpript.[23]
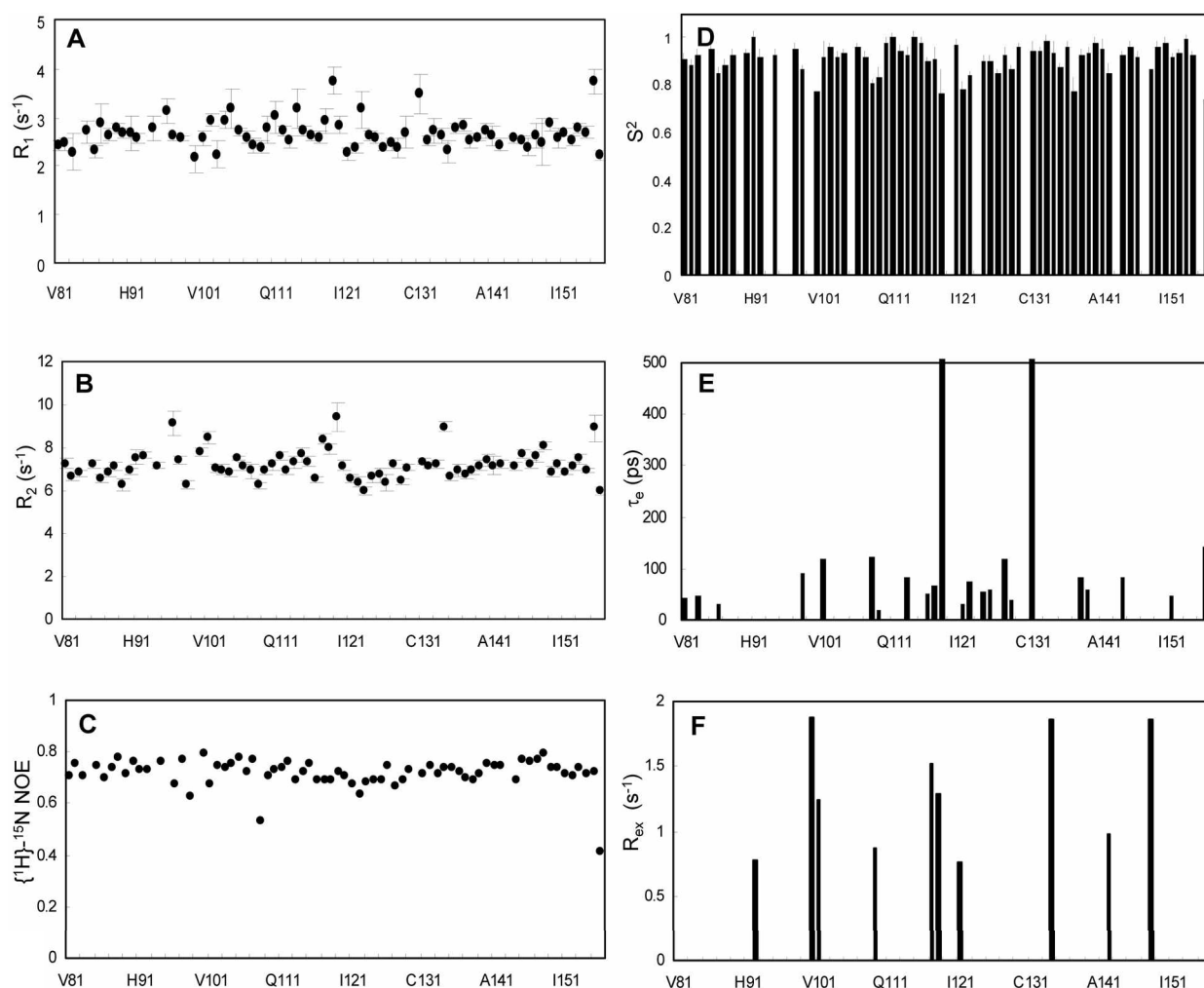
Figure 3. Plot of ^{15}N Relaxation and dynamics parameter of hpBCCP at 500 MHz. (a) R_1, (b) R_2, (c) steady-state NOE, (d) S^2, (e) τ_e, (f) R_{ex}.
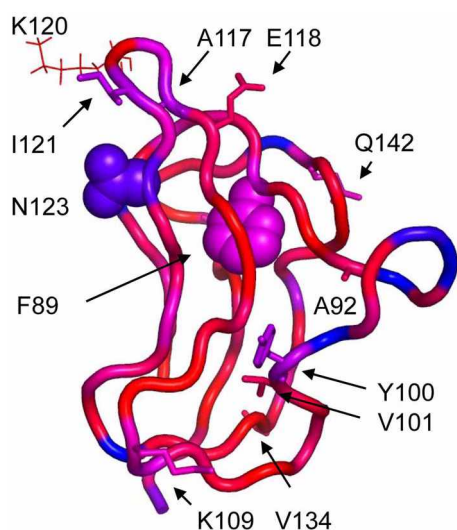
Figure 4. Schematic representation of the spatial distributions of dynamics parameters. S^2 values of each residue are expressed in blue-magenta-red. Blue for 0.5 or less, magenta for 0.78, and red for 1. Residues that require R_{ex} term to fit the measured relaxation data are shown in stick and residues that were possibly fitted model 5 are shown in sphere. Lys120, the biotinylation site, is shown in thin stick.

minus 7 prolines and 6 unfitted residues). The average value of S^2 was 0.908(+/− 0.061) for all residues and 0.911(+/− 0.058) for the secondary structured region only. The distribution of S^2 values was visualized on the ribbon-diagram structure (Fig. 4). As seen in the figure, centers of both beta sheets show a highly rigid structure with relatively large S^2 values. In contrast to the central region, the edge side of β-sheet showed a relative low S^2 value of 0.8604.

Discussion

Residues around the hairpin were well conserved. This means that these residues have functional roles. There are two possible roles. One is a hydrophobic core for folding and the other is a protein binding site. The inner face of the hairpin may act as hydrophobic core and the outer face as binding site of biotin or partner proteins. The inner face is comprised with Val113, Val116 and Ile125. They make hydrophobic core of HP0371. All of them can be seen in various homologue proteins. The outer face has Glu117, Glu124, and Ile115. Many residues on β-hairpin showed distinct dynamics parameters. Glu117, Ala118, and Ile121 were fitted to model 4 and had R_{ex} values around 1-2 s^{-1}. It

may be concerned with the interaction with biotinylation enzyme. Reche *et al.* have reported which residue of BCCP interact with BirA. *E. coli* biotinyl protein ligase, by heteronuclear NMR.[19] In that report, many residues on the edge of β5 showed large perturbation of chemical shift in the presence of BirA. On the expansion of this finding. Weaver *et al.* suggested that the β-hairpin is involved in recognition of BirA *via* forming of intermolecular β-sheet.[20] If this model is correct, R_{ex} values of these residues may be related with the protein recognition of BCCP.

One of unique feature of BCCP is the protruding thumb structure. Other types of bacterial biotinyl protein or lipoyl proteins do not have thumb structure and eukaryotic ACC and other kinds of biotinylation proteins including pyruvate dehydrogenase or transcarboxylase do not have thumb sequence either. Cronan demonstrated that this thumb is essential for the function of BCCP in the ACC reaction.[21] He also suggested that the thumb structure may be involved in the interaction with other subunit of ACC. Moreover, eukaryotic ACC is a single-chain protein but that of prokaryotic is a multisubunit protein complex. Hence, the thumb structure may be involved in the interaction with other subunit of ACC.

As mentioned before, the thumb structure is not flexible. It showed lower backbone RMSD values in structural ensemble. Averaged S^2 value for this region was 0.91 and this value was not smaller than overall average value. This structural rigidity seems to be caused by the contacts between residues on the thumb. There are 25 distance restraints on the thumb. Pro95 is located on the tip of loop and seems to form a rigid turn and contribute to rigidity. Moreover, *helicobacter* species have one more proline, Pro99, on the end of the thumb instead of alanine in comparison with *E. coli* (Fig. 2). This proline makes additional contact with Ala92 and may make the thumb more rigid. This thumb structure may make contribution to the interaction with other ACC subunit because it has small entropy change caused by rigid structure and large enthalpy change caused by hydrophobic interaction of prolines and hydrogen bonding of Ser94 or Glu90 with residues on partner proteins.

References

1. Kwok. T.; Zabler. D.; Urman. S.; Rohde. M.; Hartig. R.; Wessler. S.; Misselwitz. R.; Berger. J.; Sewald. N.; Konig. W.; Backert. S. *Nature* **2007**, *449*. 862-6.
2. Cronan, J. E., Jr.; Waldrop. G. L. *Prog. Lipid Res.* **2002**. *41*. 407-35.
3. Tomb, J. F.; White, O.; Kerlavage, A. R.; Clayton. R. A.; Sutton. G. G.; Fleischmann, R. D.; Ketchum. K. A.; Klenk, H. P.; Gill. S.; Dougherty, B. A.; Nelson. K.; Quackenbush, J.; Zhou. L.; Kirkness. E. F.; Peterson. S.; Loftus. B.; Richardson. D.; Dodson. R.; Khalak. H. G.; Glodek. A.; McKenney. K.; Fitzegerald. L. M.; Lee. N.; Adams. M. D.; Hickey. E. K.; Berg. D. E.; Gocayne. J. D.; Utterback. T. R.; Peterson. J. D.; Kelley. J. M.; Cotton. M. D.; Weidman. J. M.; Fujii. C.; Bowman. C.; Watthey. L.; Wallin. E.; Hayes. W. S.; Borodovsky. M.; Karp. P. D.; Smith. H. O.; Fraser. C. M.; Venter, J. C. *Nature* **1997**, *388*. 539-47.
4. Athappilly, F. K.; Hendrickson, W. A. *Structure* **1995**. *3*. 1407-19.
5. Yao. X.; Wei, D.; Soden, C.. Jr.; Summers, M. F.; Beckett. D. *Biochemistry* **1997**, *36*. 15089-100.
6. Roberts, E. L.; Shu. N.; Howard, M. J.; Broadhurst, R. W.; Chapman-Smith. A.; Wallace. J. C.; Morris. T.; Cronan. J. E.. Jr.; Perham. R. N. *Biochemistry* **1999**. *38*. 5045-53.
7. Boyd. J.; Hommel. U.; Campbell. I. D. *Chemical Physics Letters* **1990**. *175*. 477-82.
8. Farrow. N. A.; Muhandiram. R.; Singer. A. U.; Pascal. S. M.; Kay. C. M.; Gish. G.; Shoelson. S. E.; Pawson. T.; Forman-Kay. J. D.; Kay. L. E. *Biochemistry* **1994**. *33*. 5984-6003.
9. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer. J.; Bax. A. *J. Biomol NMR* **1995**. *6*. 277-93.
10. Goddard, T. D.; Kneller, D. G.; 3.110 ed.; University of California: San Francisco. 2004.
11. Zimmerman. D. E.; Kulikowski. C. A.; Huang. Y.; Feng. W.; Tashiro. M.; Shimotakahara. S.; Chien. C.; Powers. R.; Montelione. G. T. *J. Mol. Biol.* **1997**. *269*. 592-610.
12. Herrmann. T.; Guntert. P.; Wuthrich. K. *J. Mol. Biol.* **2002**. *319*. 209-27.
13. Cornilescu. G.; Delaglio. F.; Bax. A. *J. Biomol. NMR* **1999**. *13*. 289-302.
14. Fernandez, C.; Hilty. C.; Wider. G.; Guntert, P.; Wuthrich, K. *J. Mol. Biol.* **2004**, *336*. 1211-21.
15. Palmer. A. G.; Rance, M.; Wright. P. E. *J. Am. Chem. Soc.* **1991**. *113*, 4371-4380.
16. Dosset. P.; Hus. J. C.; Blackledge. M.; Marion. D. *J. Biomol. NMR* **2000**. *16*. 23-8.
17. Tsan. P.; Hus. J. C.; Caffrey. M.; Marion. D.; Blackledge. M. *J. Am. Chem. Soc.* **2000**. *122*. 5603-12.
18. Lipari. G.; Szabo. A. *J. Am. Chem. Soc.* **1982**. *104*. 4546-4559.
19. Reche. P. A.; Howard. M. J.; Broadhurst. R. W.; Perham. R. N. *FEBS Lett.* **2000**. *479*, 93-8.
20. Weaver, L. H.; Kwon, K.; Beckett. D.; Matthews. B. W. *Protein Sci.* **2001**, *10*, 2618-22.
21. Cronan. J. E., Jr. *J. Biol. Chem.* **2001**. *276*. 37355-64.
22. Wilbur, W. J.; Lipman. D. J. *Proc. Natl. Acad. Sci. USA* **1983**. *80*. 726-30.
23. Gouet, P.; Courcelle. E.; Stuart, D. I.; Metoz, F. *Bioinformatics* **1999**. *15*. 305-8.