Biomolecular NMR Assignments (2021) 15:159–164
https://doi.org/10.1007/s12104-020-09999-8

ARTICLE

Resonance assignments and secondary structure of thermophile single-stranded DNA binding protein from *Sulfolobus solfataricus* at 323K

Min June Yang[1] · Woonghee Lee[2] · Chin-Ju Park[1]

Received: 1 November 2020 / Accepted: 22 December 2020 / Published online: 6 January 2021

Abstract

Single-stranded DNA (ssDNA)-binding proteins (SSBs) are essential for DNA replication, recombination, and repair processes in all organisms. *Sulfolobus solfataricus* (*S. solfataricus*), a hyperthermophilic species, overexpresses its SSB (*S. solfataricus* SSB (SsoSSB)) to protect ssDNA during DNA metabolisms. Even though the crystal structure of apo SsoSSB and its ssDNA-bound solution structure have been reported at room temperature, structural information at high temperature is not yet available. To find out how SsoSSB maintains its structure and ssDNA binding affinity at high temperatures, we performed multidimensional NMR experiments for SsoSSB at 323K. In this study, we present the backbone and side chain chemical shifts and predict the secondary structure of SsoSSB from the chemical shifts. We found that SsoSSB is ordered, even at high temperatures, and has the same fold at high temperature as at room temperature. Our data will help improve structural analyses and our understanding of the features of thermophilic proteins.

Keywords Thermophile · Single-stranded binding · OB-fold · High Temperatures NMR spectroscopy

Biological context

Single-stranded DNA (ssDNA) binding proteins (SSBs) are necessary to all living organisms for maintaining their genome integrity (Ashton et al. 2013; Mushegian and Koonin 2002). It is important to protect ssDNA from SSBs during DNA replication and transcription. The common feature of the oligonucleotide/oligosaccharide binding fold (OB-fold) domain is conserved in many SSBs. OB-fold superfamilies interact with staphylococcal nucleases, bacterial enterotoxins, inorganic pyrophosphatases, and nucleic acids (Theobald et al. 2003). SSBs such as human replication protein A (hRPA; Arunkumar et al. 2003), *Escherichia coli* SSB (Raghunathan et al. 2000), and human mitochondrial SSB (Qian and Johnson 2017) can bind to ssDNA tightly (Murzin 1993). All eukaryotes use a heterotrimeric SSB known as RPA with six OB-folds. The hRPA complex has six OB-folds, and four domains (RPA70 DNA binding domains A, B, and C, RPA 32) are involved in DNA binding (Arunkumar et al. 2003; Theobald et al. 2003). OB-folds are important motifs that interact with DNA and maintain genome stability.

Sulfolobus is one of the genera described by the Brock group (Brock et al. 1972), a well-known hyperthermophilic archaebacterium. This genus includes eight species of *Sulfolobus*, of which *Sulfolobus solfataricus* is one. Its cell wall does not contain peptidoglycan, unlike other bacteria that have a typical cell wall structure. They survive at pH 0.9–5.8, temperatures of 55–80 °C, and optimal pH values of 2–3, and their optimal temperature is 70–75 ℃ (Brock et al. 1972; Grogan and Palm 1990; Quehenberger et al. 2017; Sakai and Kurosawa 2018; Zillig et al. 1980). Whole-genome sequencing of *S. solfataricus* was completed in 2001, and the genome of this species has 2,992,245 bp on a single chromosome that encodes 2977 proteins. *S. solfataricus* has 40% archaeal-specific features and 12% bacterial features, 2.3% of eukaryote features (She et al. 2001). Several catabolic enzymes have been studied in detail in

✉ Woonghee Lee
woonghee.lee@ucdenver.edu

✉ Chin-Ju Park
cjpark@gist.ac.kr

[1] Department of Chemistry, Gwangju Institute of Science and Technology, Gwangju 61005, Korea

[2] Department of Chemistry, University of Colorado Denver, Denver, CO 80217-3364, USA

S. solfataricus, and these enzymes are metabolically very diverse (Brasen et al. 2014).

For organisms that inhabit severe environments, DNA damage occurs more frequently. ssDNA is much more sensitive to damage than dsDNA, so it needs a specific strategy for protection (Ashton et al. 2013; Dickey et al. 2013). Unstable ssDNA is more likely to exist at high temperatures (Lindahl and Nyberg 1972). Overexpression of SSB has been observed in many hyperthermophilic archaebacterial species for protecting ssDNA from high temperatures and harsh environments. In addition to DNA, SSB also binds to RNA with high affinity to protect RNA from high temperature (Morten et al. 2017). *S. solfataricus* SSB (SsoSSB) consists of 148 amino acids and has well-conserved residues with other SSBs that have OB-folds. Most OB-fold SSBs interact with DNA by aromatic residue base stacking interactions. Two aromatic residues are preserved among hRPA70A, hRPA70B, hSSB1, and hSSB2. SsoSSB has one additional aromatic sidechain that is known to form a stacking interaction with ssDNA (Gamsjaeger et al. 2015; Kerr et al. 2003). The structure of SsoSSB was solved by X-ray crystallography (PDB ID: 1O7I; Kerr et al. 2003) and nuclear magnetic resonance spectroscopy (NMR; PDB ID: 2MNA; Gamsjaeger et al. 2015) at room temperature. The DNA-SsoSSB complex structure has also been determined by NMR (Gamsjaeger et al. 2015). SsoSSB is monomeric and binds to ssDNA as a monomer, with five bases of ssDNA. SsoSSB also recognizes DNA directionality (Gamsjaeger et al. 2015; Morten et al. 2017).

The Gamsjaeger group deposited backbone assignment of $SsoSSB_{1-117}$ at 298K (BioMagResBank, http://www.bmrb.wisc.edu/, accession number 19,095; Gamsjaeger et al. 2014), and the protein was briefly tested for high-temperature stability with HSQC and other NMR experiments at 330K (Gamsjaeger et al. 2015). Although the structure at room temperature is available, the detailed structural information at high temperature has not yet been studied. Here, we present here the backbone and side chain assignment and secondary structure prediction of $SsoSSB_{1-114}$ at 323K. Our data will help improve our understanding of thermophilic protein features, such as structure and dynamics at high temperatures. Our results with allso provide insights into how an exceptionally stable protein structure can be maintained under harsh conditions.

Method and experiments

Protein expression and purification

DNA encoding SsoSSB residues 1–114 with a C-terminal His-tag ($SsoSSB_{1-114}$) was cloned into a pET C-terminal TEV His6 cloning vector with BioBrick polycistronic restriction sites (9Bc) and transformed into BL21(DE3) cells. C-terminal His-tag consists of the linker (9 amino acids) and 6X His (ENLYFQSGSHHHHHH). Residues 1-114 of SsoSSB contain the entire OB-fold, and the C-terminal His-tag allowed the protein to bind to a Ni-NTA column to facilitate protein purification. We cultured cells in 10 mL LB medium (25 g/L) with ampicillin (0.3 mM, final concentration) at 37 °C for more than 12 h. A small number of cells were poured into 1L of LB medium with ampicillin. Cells were grown until OD_{600} reached 0.5–0.6 at 37 °C. At an OD_{600} of 0.5–0.6, (isopropyl β-D-1-thiogalactopyranoside, IPTG, 0.5 mM final concentration) was added to the culture. Cells were cultured for 14–18 h at 18 °C and then harvested by centrifugation for 15 min at 7500 rpm at 4 °C. Cells were resuspended and sonicated in binding buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0). The supernatant was separated by centrifugation for 15 min at 13,000 rpm at 4 °C and applied to a Ni-NTA column (ELPIS Bio tech, Korea). The proteins were purified using wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 40 mM imidazole, pH 8.0) and elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 300 mM imidazole pH 8.0). Gel filtration chromatography was performed using an AKTA pure with a Hi-Load 16/600 75 pg column (GE Healthcare, US) and high-salt buffer (2 M NaCl, 20 mM MES, pH 6.5) for RNA detachment. After purification, we exchanged the buffer with NMR buffer (50 mM NaCl, 10 mM MES, pH 6.0).

For chemical shift assignments, we prepared ^{15}N- and ^{13}C-labeled proteins. Cells were grown in M9 minimal media that contained $^{15}NH_4Cl$ and $^{13}C_6H_{12}O_6$ (Cambridge Isotope Laboratories, Inc., US) for the expression of ^{15}N - and ^{13}C labeled proteins. The composition of M9 minimal media is 870 mL of DW, 1 g of $^{15}NH_4Cl$, 100 mL of M9 10X salt, 20 mL of 10% glucose (^{13}C labeled) solution, 2 mL of 1 M $MgSO_4$ solution, 0.3 mL of 1 M $CaCl_2$ solution, 0.33 mL of vitamin solution, and 10 mL of trace metal solution.

NMR spectroscopy

We prepared fresh ^{15}N- and ^{13}C-labeled $SsoSSB_{1-114}$ for chemical shift assignments. The sample was dissolved to a final protein concentration of 1.2 mM in NMR buffer and transferred to a salt-tolerant Shigemi NMR tube (Shigemi, Allison Park, PA) for subsequent NMR experiments. NMR spectra were acquired on a Bruker AVANCE Neo spectrometer operating at 600 MHz (^1H), equipped with a cryogenic triple-resonance probe (GIST, Gwangju). All experiments were performed at 323 K, and for the backbone assignments, we acquired 2D ^1H-^{15}N HSQC, 3D HNCO, 3D HN(CA)CO, 3D CBCA(CO)NH, 3D HNCACB, and 3D NOESY-^{15}N HSQC spectra. The NOESY experiment utilized a mixing time of 150 ms. More information about the NMR experiments is shown in Table 1. NMRPipe (Delaglio et al. 1995)

Table 1 2D and 3D NMR spectra information for backbone and side-chain assignments

Experiment	Number of scans	Spectral width (ppm) $^{1}H \times {}^{13}C(^{1}H) \times {}^{15}N$	Number of points $^{1}H \times {}^{13}C(^{1}H) \times {}^{15}N$	Offset (ppm) ^{1}H, $^{13}C(^{1}H)$, ^{15}N
^{1}H-^{15}N HSQC[a]	4	15.15 × 34	1024 × 192	4.7, 118
HNCO[a]	8	15.15 × 16 × 36	1024 × 128 × 50	4.705, 172, 118
HN(CA)CO[a]	16	13.66 × 14 × 35	2048 × 128 × 40	4.7, 173.5, 117
CBCA(CO)NH[a]	8	15.15 × 80 × 36	1024 × 256 × 70	4.705, 40, 118
HNCACB[a]	16	15.15 × 80 × 36	1024 × 128 × 60	4.705, 40, 118
NOESY-^{15}N HSQC[a]	8	13.66 × 13.66 × 34	2048 × 192 × 40	4.7, 4.7, 118
^{1}H-^{13}C HSQC[b]	16	13.02 × 80	1024 × 512	4.7, 40
C(CO)NH[b]	16	16.02 × 80 × 36	2048 × 128 × 40	4.7, 40, 118
HBHA(CO)NH[b]	16	16.02 × 16.02 × 36	2048 × 192 × 40	4.7, 4.7, 118
H(CCO)NH[b]	16	16.02 × 16.02 × 35	2048 × 128 × 40	4.701, 4.701, 117
H(C)CH–TOCSY[b]	8	13.66 × 13.66 × 80	2048 × 128 × 40	4.7, 4.7, 40

All experiments were performed at 323K using Bruker AVANCE NEO 600 MHz

[a]Experiments used for backbone assignments

[b]Experiments used for side chain assignments

software and TOPSPIN (Bruker) software were used for NMR data processing, and the NMRFAM-SPARKY software package (Lee et al. 2015) was used to analyze all processed spectra. Our 2D ^{1}H-^{15}N HSQC exhibited large peak shifts and disappearances from the $SsoSSB_{1-117}$ construct at 298K (Gamsjaeger et al. 2014); therefore, we had to assign it from scratch. (Semi-)automation and verification tools provided by NMRFAM-SPARKY (Lee et al. 2016) were used. The 2D ^{1}H-^{15}N HSQC and triple-resonance through-bond experiments described in Table 1 were mainly used to adapt the sequential walking strategy for backbone assignments. Similarities of NOE cross-peak patterns from 3D NOESY were used to confirm sequential connectivities. The APES automation plugin *(*two-letter-code *ae*) was used for initial peak-picking, and peaks were manually verified by two strip plots (two-letter-codes *sp* and *SP*). To verify the CA/CB peaks, we created strips of CBCA(CO)NH and HNCACB spectra. For CO verification, we created strips of HNCO and HN(CA)CO spectra. Verified peaks from the backbone spectra were used as inputs to the PINE-SPARKY.2 plugin (two-letter-code *ep*; Lee and Markley 2018) to operate the I-PINE web server (Lee et al. 2019) for the automated assignments. Chemical shifts obtained from the I-PINE web server were validated and confirmed with PINE Graph Assigner and PINE Assigner from the PINE-SPARKY package (Lee et al. 2009) with the support of the 3D NOESY-^{15}N HSQC experiment verifying sequential connectivities from NOE cross peak patterns. For the side chain assignments, we collected additional spectra (2D constant-time ^{1}H-^{13}C-HSQC, 3D C(CO)NH, 3D HBHA(CO)NH, 3D H(CCO)NH, and 3D H(C)CH-TOCSY) and used the "transfer and simulate assignments" tool available in NMRFAM-SPARKY (two-letter-code *ta*) to perform the *predict-and-confirm* method that extends the backbone assignments to the side chains.

Extent of assignments and data deposition

Assigned chemical shifts of backbone and side chain ^{1}H, ^{15}N, and ^{13}C resonances for $SsoSSB_{1-114}$ have been deposited in the BioMagResBank ((http://bmrb.io/, accession number 50,523). The 2D ^{1}H-^{15}N HSQC spectrum of $SsoSSB_{1-114}$ with assigned peaks annotated with the sequence and the residue number is illustrated in Fig. 1. We assigned the major peaks when split peaks were observed for a residue. Despite the large dynamic range of peak intensities, peak splitting, and peak overlaps, we could assign 116 of 120 backbone amide groups, excluding the four proline residues. Only residues at the terminal (M1, E2, and H124) and S89 were not assigned. Assignment statistics: 116 of 120 backbone ^{15}N-$^{1}H^{N}$ (96.7%), 121 of 124 $^{13}C^{\alpha}$ (97.6%), 110 of 113 $^{13}C^{\beta}$ (97.4%), 120 of 124 $^{13}C^{o}$ (96.8%), and total 1223 of 1280 (95.6%) backbone aliphatic side chain atoms were assigned. Several tools in NMRFAM-SPARKY were utilized to analyze the NMR data. Chemical shifts were re-referenced using linear analysis of chemical shifts (LACS; two-letter-code *lv*; Wang et al. 2005). PECAN (Wang et al. 2005) and TALOS-N (Shen and Bax 2013) were performed to analyze the assigned and referenced backbone chemical shifts (two-letter-codes *n6* and *tl*). NDP-PLOT was used to plot the secondary structure prediction and random coil index order parameter results from PECAN and TALOS-N (Fig. 2). Nine β-strands 1−9 were identified by TALOS-N and PECAN. Two additional 3^{10}-helices were also suggested by TALOS-N (5–7 and 104–106), which we assume resulted from the algorithmic discrepancy between the two programs. This will be elucidated with the three-dimensional solution structure of SsoSSB, which we are pursuing. We believe that β-Strands 2–5 and 6–9 form an antiparallel β-sheet, considering the proximity in the sequence, which

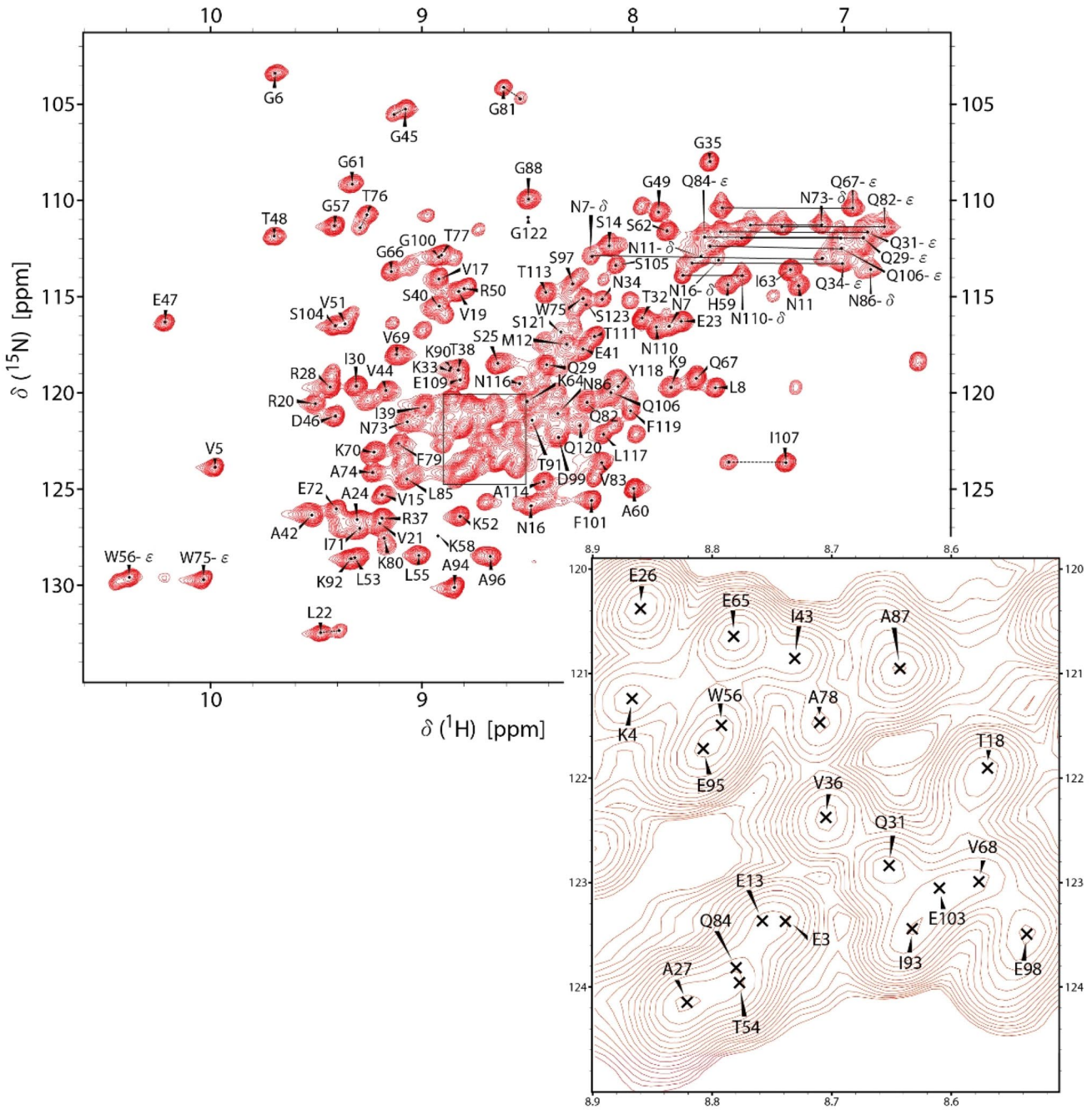

Fig. 1 2D ^{1}H-^{15}N HSQC spectrum of 1.2 mM SsoSSB$_{1\text{-}114}$with NMR buffer (50 mM NaCl, 10 mM MES, pH 6.0) and 10% D_2O. Thisexperiment was performed at 323K on a BRUKER ADVANCE NEO 600MHz. Assignmentsare labeled on the spectrum. Splitting peaks areconnected with dotted line. Side chain peaks from same amino acid are connectedwith solid line

defines this domain as a β barrel protein similar to the room temperature structure. According to the RCI-S^2 values calculated from TALOS-N (Fig. 2c), the overall structure is rigid and only the C-terminal His-tag residues are disordered (residues 115−123). From the results above, we confirmed that SsoSSB$_{1\text{-}114}$ maintains its structure at high temperatures, in line with previous X-ray and NMR structures at room temperature.

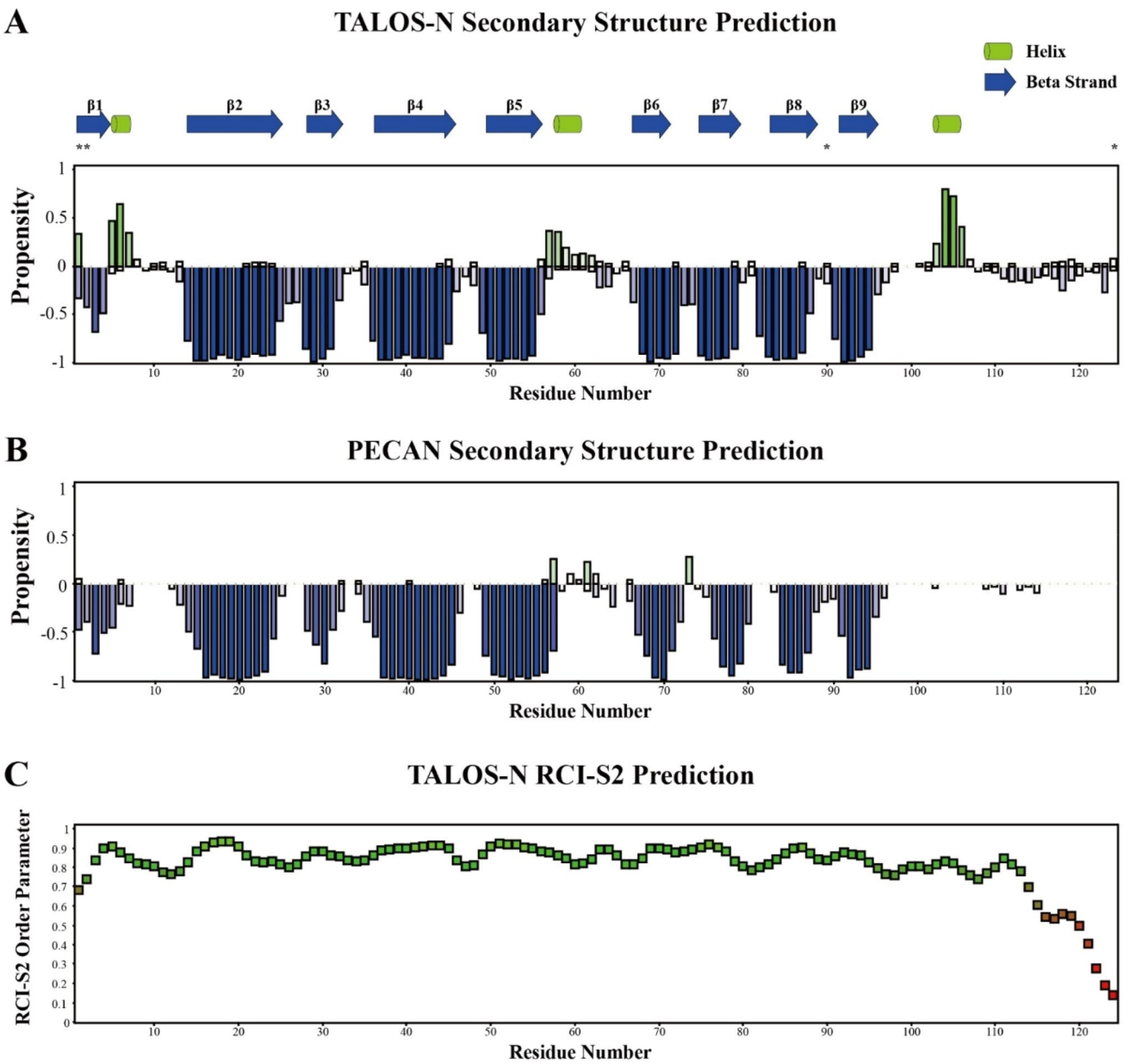

Fig. 2 TALOS-N and PECAN results calculated frombackbone chemical shift data of $SsoSSB_{1\text{-}114}$. Nine strands werepredicted from both models, and two short helices were additionally predictedby TALOS-N. Unassigned residues are labeled by asterisks. a TALOS-N secondary structurepropensities prediction (Green: helix, Blue: beta-strand). b PECAN secondary structurepropensities prediction (Green: helix, Blue: beta-strand). c Random coil index order parameters.($RCI\text{-}S^2$) calculated from TALOS-N (Red to green, ordered to rigid)

Acknowledgements We thank GIST Central Research Facilities (GIST, Gwangju) for allowing us to use their NMR spectrometer. This work was supported by the National Research Foundation (NRF) of Korea [Grant No. 2018R1A2B6004388 to C.-J.P], which is funded by the Korean government (MSIT); by a GIST Research Institute grant funded by GIST in 2020; the National Science Foundation (NSF) [DBI-2051595 to W.L.] and the University of Colorado Denver [the start-up support to W.L.].

Author contributions CP and WL designed the research. MY performed the experiments. MY and WL analyzed the data. MY, WL, and CP wrote the paper.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest with the contents of this article.

References

Arunkumar AI et al (2003) Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. J Biol Chem 278(42):41077–41082

Ashton NW et al (2013) Human single-stranded DNA binding proteins are essential for maintaining genomic stability. BMC Mol Biol 14:1–20

Brasen C, Esser D, Rauch B, Siebers B (2014) Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol Mol Biol Rev 78(1):89–175. https://doi.org/10.1128/MMBR.00041-13

Brock TD, Katherine M, Brock RT, Belly, Weiss RL (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84(1):54–68

Delaglio F et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6(3):277–293

Dickey TH, Sarah E, Altschuler, Wuttke DS (2013) Single-stranded DNA-binding proteins: multiple domains for multiple functions. Structure 21(7):1074–1084. https://doi.org/10.1016/j.str.2013.05.013

Gamsjaeger R et al (2014) Backbone and side-chain 1H, 13C and 15N resonance assignments of the OB domain of the single stranded DNA binding protein from sulfolobus solfataricus and chemical shift mapping of the DNA-binding interface. Biomol NMR Assign 8(2):243–246

Gamsjaeger R et al (2015) The structural basis of DNA binding by the single-stranded DNA-binding protein from sulfolobus solfataricus. Biochem J 465(2):337–346

Grogan D, Palm P (1990) Isolate B12, which harbours a virus-like element, represents a new species of the archaebacterial genus sulfolobus, sulfolobus shibatae, Sp. Nov. Arch Microbiol 154(6):594–599. https://doi.org/10.1007/BF00248842

Kerr ID et al (2003) Insights into SsDNA recognition by the OB fold from a structural and thermodynamic study of sulfolobus SSB protein. EMBO J 22(11):2561–2570

Lee W et al (2009) PINE-SPARKY: graphical interface for evaluating automated probabilistic peak assignments in protein NMR spectroscopy. Bioinformatics 25(16):2085–2087

Lee W et al (2016) Integrative NMR for biomolecular research. J Biomol NMR 64(4):307–332

Lee W, Markley JL (2018) PINE-SPARKY.2 for automated NMR-based protein structure research. Bioinformatics 34(9):1586–1588

Lee W et al (2019) I-PINE web server: an integrative probabilistic NMR assignment system for proteins. J Biomol NMR 73(5):213–222. https://doi.org/10.1007/s10858-019-00255-3

Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31(8):1325–1327

Lindahl T (1972) Rate of depurination of native deoxyribonucleic acid. Biochemistry 11(19):3610–3618. https://doi.org/10.1021/bi00769a018

Morten MJ et al (2017) High-affinity RNA binding by a hyperthermophilic single-stranded DNA-binding protein. Extremophiles 21(2):369–379

Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J 12(3):861–867

Mushegian AR, Koonin EV (2002) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci 93(19):10268–10273

Qian Y, Johnson KA (2017) The human mitochondrial single-stranded DNA-binding protein displays distinct kinetics and thermodynamics of DNA binding and exchange. J Biol Chem 292(31):13068–13084

Quehenberger J et al (2017) Sulfolobus-a potential key organism in future biotechnology. Front Microbiol 8(DEC):1–13

Raghunathan S et al (2000) Structure of the DNA binding domain of *E . Coli* SSB bound to ssDNA. Nat Struct Biol 7(8):648–652

Sakai HD, and Norio Kurosawa (2018) Saccharolobus Caldissimus Gen. Nov., Sp. Nov., a Facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of sulfolobus solfataricus as saccharolobus solfataricus Comb. Nov. And. Int J Syst Evol Microbiol 68(4): 1271–1278. https://doi.org/10.1099/ijsem.0.002665

She Q et al (2001) The complete genome of the crenarchaeon sulfolobus solfataricus P2. Proc Natl Acad Sci 98(14):7835–7840. https://doi.org/10.1073/pnas.141222098

Shen Y (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J Biomol NMR 56(3):227–241

Theobald DL, Rachel M, Mitton-Fry, Deborah SWuttke (2003) Nucleic acid recognition by OB-fold proteins. Annu Rev Biophys Biomol Struct 32(1):115–133. https://doi.org/10.1146/annurev.biophys.32.110601.142506

Wang L, Eghbalnia HR, Bahrami A, Markley JL (2005) Linear analysis of carbon-13 chemical shift differences and its application to the detection and correction of errors in referencing and spin system identifications. J Biomol NMR 32(1):13–22

Zillig W et al (1980) The sulfolobus-"Caldariella" group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Arch Microbiol 125(3):259–269