2750–2757 Nucleic Acids Research, 2014, Vol. 42, No. 4 Published online 27 November 2013
doi:10.1093/nar/gkt1238

# Intramolecular binding mode of the C-terminus of *Escherichia coli* single-stranded DNA binding protein determined by nuclear magnetic resonance spectroscopy

**Dmitry Shishmarev[1], Yao Wang[2], Claire E. Mason[2], Xun-Cheng Su[1], Aaron J. Oakley[2], Bim Graham[3], Thomas Huber[1], Nicholas E. Dixon[2] and Gottfried Otting[1,*]**

[1]Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 0200, Australia, [2]Centre for Medical and Molecular Bioscience and School of Chemistry, University of Wollongong, New South Wales 2522, Australia and [3]Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Parkville Victoria 3052, Australia

Received October 13, 2013; Revised November 5, 2013; Accepted November 7, 2013

## ABSTRACT

**Single-stranded DNA (ssDNA) binding protein (SSB) is an essential protein to protect ssDNA and recruit specific ssDNA-processing proteins. *Escherichia coli* SSB forms a tetramer at neutral pH, comprising a structurally well-defined ssDNA binding domain (OB-domain) and a disordered C-terminal domain (C-domain) of ~64 amino acid residues. The C-terminal eight-residue segment of SSB (C-peptide) has been shown to interact with the OB-domain, but crystal structures failed to reveal any electron density of the C-peptide. Here we show that SSB forms a monomer at pH 3.4, which is suitable for studies by high-resolution nuclear magnetic resonance (NMR) spectroscopy. The OB-domain retains its 3D structure in the monomer, and the C-peptide is shown by nuclear Overhauser effects and lanthanide-induced pseudocontact shifts to bind to the OB-domain at a site that harbors ssDNA in the crystal structure of the SSB–ssDNA complex. $^{15}$N relaxation data demonstrate high flexibility of the polypeptide segment linking the C-peptide to the OB-domain and somewhat increased flexibility of the C-peptide compared with the OB-domain, suggesting that the C-peptide either retains high mobility in the bound state or is in a fast equilibrium with an unbound state.**

## INTRODUCTION

Single-stranded DNA (ssDNA) binding protein (SSB) is a key component in DNA replication, recombination and repair (1,2), protecting ssDNA from endonucleases and recombinases and preventing its reannealing (1), and forming a protein interaction hub that recruits ssDNA processing enzymes (3–7). The critical importance of SSB is highlighted by functional conservation across all kingdoms of life (1,3,8). SSB from *Escherichia coli* is the most thoroughly studied prokaryotic SSB. The protein functions as a tetramer, where tetramerization is mediated by a structurally well-defined oligonucleotide-binding domain (OB-domain) that is followed by a disordered polypeptide segment of ~64 residues (C-domain) (2).

The crystal structure of a complex of C-terminally truncated *E. coli* SSB (residues 1–135) with ssDNA shows that ssDNA wraps around the SSB tetramer by binding to the OB-domains, but no electron density was observed for residues in the C-domain beyond Arg115 (Figure 1) (9). Crystal structures determined of tetrameric SSBs from different organisms confirm the structural disorder of the C-domain, i.e. it is either conformationally variable or unobservable (9–14). Nonetheless, the C-domain is known to play an important biological role by mediating the interactions with all other proteins that have been identified to bind to SSB (3). More specifically, SSBs associate with the C-terminal eight residues of *E. coli* SSB (DFDDDIPF) (3), which are highly conserved in the amino acid sequences of bacterial SSBs (15). In the following, we refer to this highly negatively charged peptide as the 'C-peptide'.

Biochemical experiments with C-terminally truncated *E. coli* SSB accumulated strong evidence that the C-peptide binds to the OB-domain in competition with ssDNA (16–19). Competition between the C-domain of a distantly related phage SSB and ssDNA for binding to

*To whom correspondence should be addressed. Tel: +61 2 6125 6507; Fax: +61 2 6125 0750; Email: gottfried.otting@anu.edu.au
Present address:
Xun-Cheng Su, State Key Laboratory of Elemento-Organic Chemistry, Nankai University, Tianjin 300071, China.

© The Author(s) 2013. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

**Figure 1.** Crystal structure of the OB-domain tetramer of *E. coli* SSB in complex with ssDNA (PDB code 1EYG) (9). The four SSB monomers in the tetramer are shown in different colors. The backbones of the two $(dC)_{35}$ molecules in the structure are traced by orange lines. Red spheres mark the C-termini of the four OB-domains (carbonyl carbon of Leu112). The length of each 35-mer is ~75 Å, i.e. shorter than the average end-to-end distance of a random-coil peptide chain with 64 residues as in the C-domain (>90 Å).

an OB-domain has also been demonstrated for the phage T7 gene 2.5 protein (gp2.5), a homodimeric SSB homolog (20). This leads to a model in which, by displacing the C-peptide from the OB-domain, the binding of ssDNA would make the C-peptide available for binding to ssDNA processing enzymes.

At present, no structural details about the binding of the flexible C-domain to the OB-domain are known for any SSB protein. In particular, the study of gp2.5 involved chemical cross-linking and nuclear magnetic resonance (NMR) chemical shift mapping, but no specific cross-links were identified and no NMR resonance assignments established (20). To learn more about the C-domain–OB-domain interactions in *E. coli* SSB, we applied high-resolution NMR spectroscopy to isotope-labeled full-length SSB. Owing to limited solubility and a high molecular weight (75 kDa), the SSB tetramer is difficult to study by solution NMR spectroscopy. Serendipitously, however, we discovered that SSB is monomeric and highly soluble at acidic pH, allowing sequence-specific resonance assignments, the analysis of NOESY spectra and structure calculations. Our results identify the binding site of the C-peptide on the OB-domain and show that most of the C-domain is highly flexible.

## MATERIALS AND METHODS

### Protein expression and purification

The phage lambda promoter plasmid pND72 (19) was used for the overexpression of wild-type SSB at 42°C. The SSB E65C/E69D mutant was constructed by following the QuikChange protocol (Stratagene), with pND72 as template and oligonucleotides 694: 5′-TTCGGCAAACT GGCATGTGTGGCGAGCGATTATCTGCGTAAAG GT (mutated codons underlined) and its complement, 695. The sequence of the mutated gene in the new plasmid pYW1666 was confirmed by nucleotide sequence determination using vector primers 9 and 10 (21).

$^{15}N$- and $^{15}N/^{13}C$-labeled SSB and SSB E65C/E69D were purified as described (19), using cultures of *E. coli* BL21(λDE3)*recA* (22) containing pND72 or pYW16666 grown in M9 minimal medium with $^{15}NH_4Cl$ (and $^{13}C$-glucose) (Cambridge Isotope Laboratories) as the sole source of nitrogen (and carbon).

SSB samples were dialyzed against 0.4 mM HCl. When necessary, the pH was adjusted to 3.4 with HCl prior to NMR measurements. Deuterium oxide ($D_2O$) was added to all samples to a final concentration of 10% (v/v) before the NMR measurements, except for samples prepared in 100% $D_2O$, where the solvent was exchanged by lyophilization of SSB and by redissolving it in $D_2O$. Concentrations of SSB (expressed as subunits) were determined spectrophotometrically (19). SSB sequences are numbered as residues Ala1–Phe177.

The single-cysteine SSB mutant E65C/E69D was prepared for site-specific attachment of an IDA-SH tag (23). This tag immobilizes a lanthanide(III) ion by additional coordination to an aspartate side chain positioned four residues further toward the C-terminus (24). However, in the case of SSB at low pH, the attachment of the IDA-SH tag resulted in protein precipitation. Therefore, we resorted to the C1 lanthanide tag (25). C1 tags loaded with diamagnetic $Y^{3+}$, paramagnetic $Tb^{3+}$ or $Tm^{3+}$ were successfully attached by overnight incubation at 4°C. The reaction yields were almost quantitative with the C1-Y and C1-Tb tags and ~85% with the C1-Tm tag.

### NMR spectroscopy

All NMR measurements were conducted at pH 3.4 and 25°C. Experiments to assign the resonances of SSB used a Bruker 800 MHz NMR spectrometer equipped with a TCI cryoprobe. Backbone resonance assignments were obtained from the analysis of 3D HNCACB, CBCA(CO)NH, HNCA, HNCO, HN(CA)CO and HN(CO)CA spectra using a 0.37 mM sample of uniformly $^{15}N/^{13}C$-labeled SSB. Side chain resonances were assigned using H(CCCO)NH-TOCSY, (H)CC(CO)NH-TOCSY (26) and HC(C)H-TOCSY, (H)CCH-TOCSY (27) and NOESY-$^{13}C$-HSQC spectra. The same sample was used to record a 2D (H)CB(CGCC-TOCSY)$H^{ar}$ spectrum (28) to assist in the assignment of the side chains of phenylalanine residues. A 3D NOESY-$^{15}N$-HSQC spectrum was recorded of a 0.5 mM solution of $^{15}N$-labeled SSB. The 2D NOESY and TOCSY spectra used a 0.33 mM sample in $D_2O$. All NOESY spectra used a mixing time of 120 ms. $^{15}N$ relaxation data were recorded on a Bruker 600 MHz NMR spectrometer, using standard pulse sequences (29) and a 0.35 mM solution of $^{15}N$-SSB. The relaxation delays were 17, 34, 51, 68, 85, 102, 119 and 136 ms in the $R_2$ experiment and 10, 20, 40, 80, 160, 320, 640, 1280 and 2560 ms in the $R_1$ experiment. The $\tau_{cp}$ delay between the π pulses of the Carr-Purcell-Meiboom-Gill (CPMG) sequence was 900 μs. The relaxation data were fit using the program CCPNMR Analysis (30). Pseudocontact shifts (PCSs) were measured as the $^1H$ chemical shifts observed in $^{15}N$-HSQC spectra of 0.2 mM samples tagged with paramagnetic lanthanides ($Tm^{3+}$, $Tb^{3+}$) minus the

corresponding chemical shifts observed for the diamagnetic reference ($Y^{3+}$). The PCS data were recorded using the 800 MHz NMR spectrometer.

### Evaluation of PCSs

PCSs were evaluated only for $^{1}H$ spins to minimize residual anisotropic chemical shifts that can be associated with the chemical shifts of heteronuclei with large chemical shift anisotropies (31). The program Numbat (32) was used to fit magnetic susceptibility anisotropy ($\Delta\chi$) tensors to the crystal structure of the OB-domain [Protein Data Bank (PDB) code 1EYG] for minimal deviations between experimental and back-calculated PCSs. The equation

$$\Delta\delta^{PCS} = \frac{1}{12\pi r^3}[\Delta x_{ax}(3\cos^2\theta - 1) + \frac{3}{2}\Delta x_{rh}\sin^2\theta\cos 2\varphi] \quad (1)$$

was used to calculate the PCSs (measured in ppm), where $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ denote the axial and rhombic components of the $\Delta\chi$ tensor, respectively, and $r$, $\theta$ and $\phi$ are the polar coordinates of the nuclear spin with respect to the principal axes of the $\Delta\chi$ tensor.

### Structure calculations

The program PCS-Rosetta (33) was used to calculate full-length monomeric SSB models consistent with the experimental data. Rosetta 9- and 3-residue fragment libraries were generated using the Robetta server (34). The fragments were used in PCS-Rosetta low-resolution folding simulations to compute 2500 backbone-only models where the OB-domain (residues 2–111) was kept rigid while the structure of the C-domain was optimized, simultaneously using the PCSs and nuclear Overhauser effects (NOE) of the C-terminal residues (loosely mapped to the backbone). The lanthanide ion was initially placed at the $C^{\beta}$ atom of the tag ligation site and its position was dynamically optimized during the folding simulation, using the strategy implemented in PCS-Rosetta (33).

The 500 models with lowest scores were chosen for computationally expensive Rosetta all-atom refinement, carrying out five independent all-atom refinements for each backbone-only model. Final models were filtered to satisfy NOE distance restraints (<4 Å) and sorted by PCS score.

## RESULTS

### NMR resonance assignments

Complete resonance assignments were obtained for the backbone resonances except for the proline residues and residues 45, 52, 128 and 173. In addition, most of the side chain resonances were assigned. Figure 2 shows the $^{15}N$-HSQC spectrum with assignments.

### Structural integrity of the SSB monomer and binding site of the C-peptide

The chemical shift dispersion observed in the $^{15}N$-HSQC spectrum indicates that most of the protein assumes a defined 3D structure. To assess the structural integrity of the monomer relative to its structure in the tetramer, we used the coordinates of the OB-domain in the crystal structure of the tetramer (PDB code 1EYG) to identify intramolecular short $^{1}H$–$^{1}H$ distances expected to produce long-range NOEs with backbone amide protons. As far as they could be resolved, all of these long-range NOEs were observed in the 3D NOESY-$^{15}N$-HSQC, 3D NOESY-$^{13}C$-HSQC and 2D NOESY spectra, confirming the structural conservation of the OB-domain (Supplementary Table S1).

Independent confirmation of the structural conservation was obtained by measuring PCSs generated by paramagnetic C1-lanthanide tags attached to residue 65, which is located in the center of the alpha-helix (Figure 1) and exposed to the solvent. PCSs were measured for 111 of the 165 non-proline residues in SSB (Figure 3, Supplementary Table S2). Using the PCSs to fit a $\Delta\chi$ tensor to the crystal structure of SSB resulted in a good fit between back-calculated and experimental PCSs (Figure 4A). The tensor magnitudes were smaller than expected for rigidly attached lanthanides (Supplementary Table S3). It is well known that a single effective $\Delta\chi$ tensor can only imperfectly represent the range of overlapping $\Delta\chi$ tensors produced by a mobile tag (35). Therefore, the deviations from a perfect correlation between back-calculated and experimental PCSs in Figure 4A more likely reflect the mobility of the lanthanide tag than differences from the crystal structure.

The only NOEs between the C-domain and the OB-domain that could be assigned unambiguously were between side chain protons of Asp173 and Val29, and Phe171 and Val58 (Figure 5). The large PCSs observed for the C-peptide residues (Figure 3B) are in agreement with the binding site defined by the NOEs (Figure 4B). The C-peptide binding site identified by the NOEs and PCSs overlaps with the ssDNA-binding surface of SSB. Although the contacts identified by the NOEs are hydrophobic, the guanidinium groups of four arginine residues are within 8 Å of the side chains of these valine residues, whereas only a single carboxyl group (Glu100) is found in the same distance range. Therefore, electrostatic attraction between the highly negatively charged C-peptide and its positively charged binding site on the OB-domain clearly is an important component of the interaction. As the binding site remains unobstructed in the SSB tetramer, the C-peptide probably binds to the same site at neutral pH.

### $^{15}N$ relaxation analysis

Measurements of the longitudinal ($R_1$) and transverse ($R_2$) $^{15}N$ relaxation rates provided striking evidence for high mobility of the C-domain (Figure 6). Except for the N-terminal few residues that are highly mobile, the $R_1$ and $R_2$ relaxation rates were quite uniform across the entire OB-domain, confirming that the domain remains

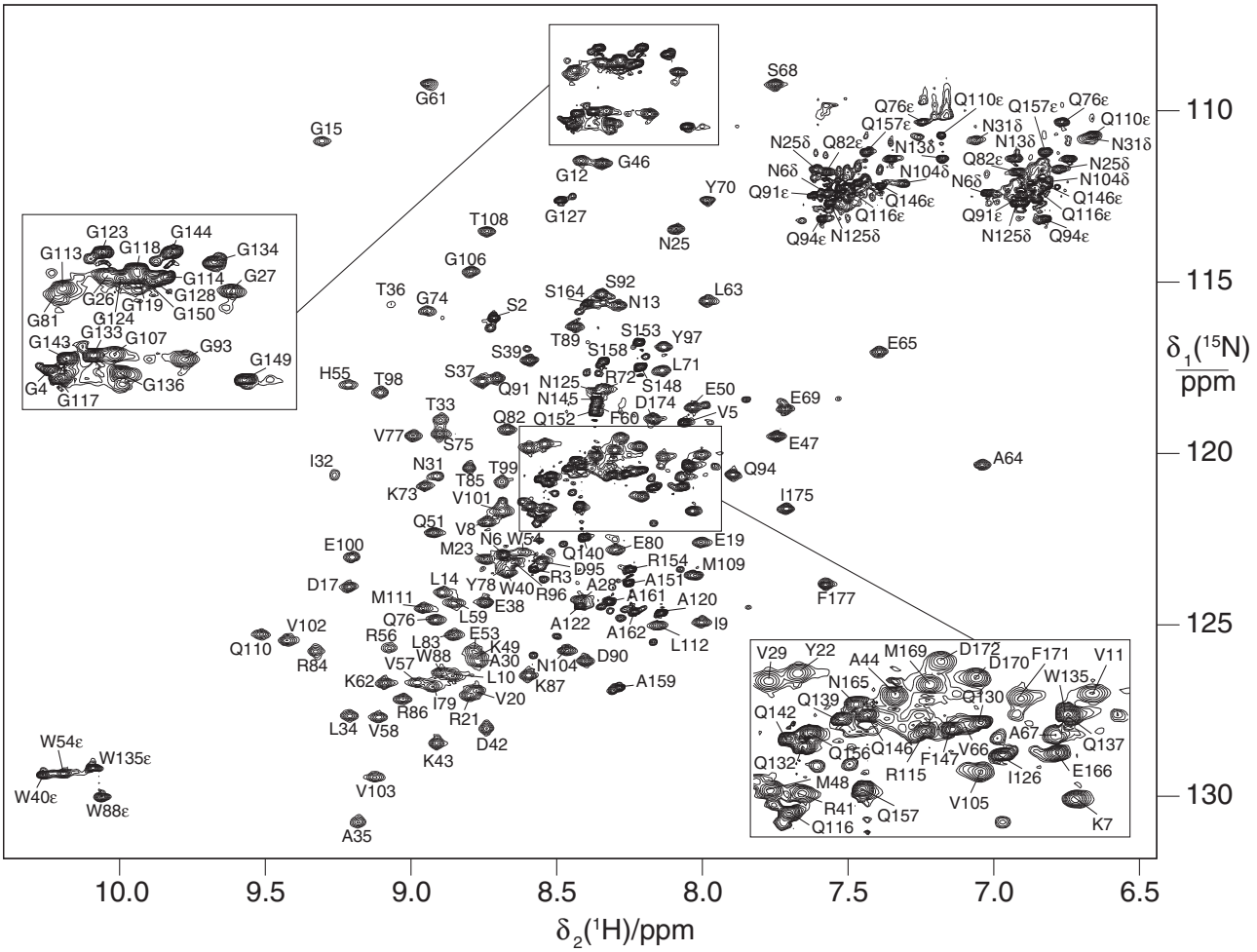


**Figure 2.** $^{15}$N-HSQC spectrum of a 0.5 mM solution of $^{15}$N-labeled *E. coli* SSB at pH 3.4 and 25°C. The cross-peaks of backbone amides are labeled with the amino acid type and sequence number. Cross-peaks of side chain NH groups are identified by the Greek characters of the side chain atoms.

folded at pH 3.4. In contrast, the transverse relaxation rates were greatly reduced for residues 112–166, highlighting the high mobility of the C-domain in solution, in agreement with the narrow chemical shift distribution observed for these residues. Notably, the C-terminal residues comprising the C-peptide showed enhanced longitudinal relaxation rates, and their transverse relaxation rates were comparable with the rates measured for the OB-domain. This result clearly demonstrates that the C-peptide associates with the OB-domain. The increased longitudinal relaxation rates of the C-peptide, however, show that the association is not entirely rigid, either allowing for somewhat enhanced mobility of the C-peptide in the bound state or reflecting an equilibrium between the bound and free states.

The high mobility of the polypeptide segment linking the OB-domain with the C-peptide explains the small magnitude of the PCSs in this segment (Figure 3B). If the polypeptide chain uniformly sampled the regions of positive and negative PCSs, the averaged PCSs would be expected to be 0. Owing to the length and flexibility of the C-domain, the only restraints to define the position of the C-peptide on the OB-domain are NOEs and PCSs of the C-peptide.

## Structure calculations

To obtain a structural model of the C-peptide in the complex with the OB-domain, we used PCS-Rosetta to identify a family of C-peptide conformations that fulfil the NOEs, PCSs and van der Waals restraints provided by the structure of the OB-domain, which was taken from the crystal structure 1EYG (9). The NOEs were translated into distance restraints of 4 Å, and the PCSs induced by two lanthanide ions, $Tb^{3+}$ and $Tm^{3+}$, were used to select the best-fitting structures from 2500 all-atom decoys calculated by PCS-Rosetta.

In view of the increased mobility of the C-peptide indicated by the $^{15}$N relaxation data, we also tested the hypothesis that the C-peptide is in fast exchange between bound and free states, with equal populations of both. In this case, the PCSs of the C-peptide in the bound state could be up to 2-fold larger than the experimentally measured values. Structures computed with 2-fold scaled PCS values better define the binding mode of the C-peptide. Among the best 20 computed structures, 19 adopt similar conformations in which the peptide occupies the ssDNA binding groove and the C-terminus points away from the dimer–dimer interface (Figure 7).

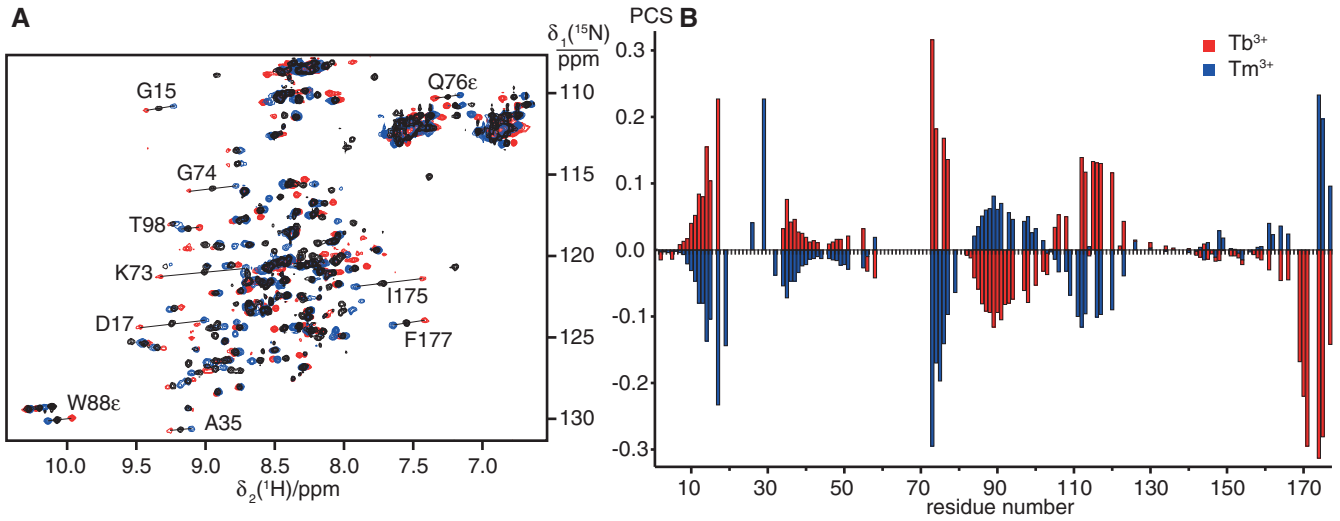


**Figure 3.** PCSs generated by paramagnetic lanthanide tags. (**A**) Overlay of $^{15}$N-HSQC spectra of $^{15}$N-labeled SSB E65C/E69D ligated with a C1 tag loaded with either diamagnetic $Y^{3+}$ (black) or paramagnetic $Tb^{3+}$ (red) or $Tm^{3+}$ (blue). The spectra were recorded of 0.2 mM solutions of protein at pH 3.4 and 25°C at a $^1$H NMR frequency of 800 MHz. The PCSs of selected residues are identified by lines and labeled with the residue assignment. (**B**) PCS values measured for the backbone amide protons from the spectra shown in (A), plotted versus the amino acid sequence.

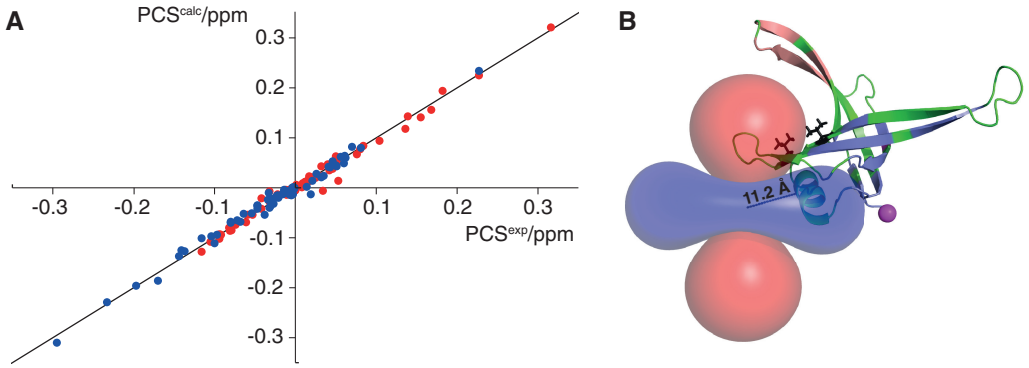


**Figure 4.** PCSs indicate structural conservation of the OB-domain. (**A**) Correlation between back-calculated and experimental PCSs for backbone amides of $^{15}$N-labeled SSB E65C/E69D with C1-lanthanide tags. The PCSs measured with $Tb^{3+}$ (red) and $Tm^{3+}$ (blue) were used to fit $\Delta\chi$ tensors to the chain A of the structure 1EYG (9), using the same metal position for both lanthanides. The good quality of the correlation indicates that the structure of the OB-domain is conserved between the crystal structure of the tetramer and the monomer in solution. (**B**) Cartoon representation of one of the SSB monomers in the tetramer of the crystal structure 1EYG (9), displaying the isosurfaces of constant PCS generated by the C1-Tb tag (red surface: −0.3 ppm, blue surface: +0.3 ppm). The dashed line highlights the distance between the metal ion and the $C^\beta$ atom of residue 65 to which the C1 tag is attached. The C-terminus of the OB-domain (Leu112) is identified by a red sphere. The side chains of Val29 and Val58 are shown in a stick representation. They are involved in NOEs with residues 171 and 173 in the C-peptide.

The structures also share a small helical turn, which orients the three consecutive aspartate residues 172–174 for strong electrostatic interactions with the highly positively charged ssDNA binding groove of the SSB OB-domain. Neither the chemical shifts nor the NOEs of the C-peptide, however, showed evidence for regular secondary structure (Supplementary Figure S1), in agreement with a reduced population of regular structure by fast exchange with unbound random coil conformations.

Fitting $\Delta\chi$ tensors from both lanthanide ions simultaneously to the best scoring structure resulted in a good fit between back-calculated and experimental PCSs for all structured residues (Supplementary Figure S2). The optimized metal ion location is positioned 11.3 Å from the $C^\alpha$ atom of the tagging site, which is in good agreement with the physical dimension of the C1 tag and with the distance found by the simple $\Delta\chi$ tensor fit to the OB-domain (Figure 4).

## DISCUSSION

### Binding mode of the C-peptide on the OB-domain

The present results show that *E. coli* SSB forms a monomer at pH 3.4 in which the OB-domain maintains the structure observed previously for the tetramer. Presumably, protonation of carboxylates at the subunit interfaces causes dissociation of the tetramer without affecting the OB-fold. The stability of the fold is underlined by the fact that the protein could be lyophilized and redissolved without change in structure. In contrast, the C-domain forms an unstructured highly mobile

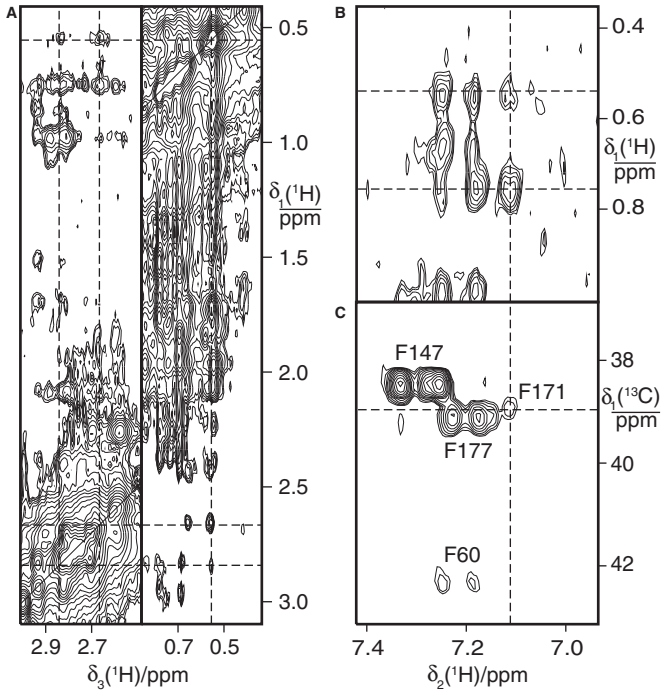


**Figure 5.** NOE cross-peaks between the C-peptide and the OB-domain. The spectra were recorded of a 0.37 mM solution of $^{15}N/^{13}C$-labeled SSB at pH 3.4 and 25°C at a $^1H$ NMR frequency of 800 MHz. (**A**) Selected spectral regions of a 3D NOESY-$^{13}C$-HSQC spectrum taken at $\delta_2(^{13}C) = 38.37$ (D173$C^{\beta}$) and 19.88 ppm (V29$C^{\gamma 1}$). The NOE cross-peaks between D173$H^{\beta}$ and V29$H^{\gamma}$ are at the intersection of the dashed lines drawn at $\delta_{1,3}(^1H) = 2.84$ (D173$H^{\beta 1}$), 2.67 (D173$H^{\beta 2}$) and 0.56 ppm (V29$H^{\gamma 1}$). (**B**) Spectral region of the 3D NOESY-$^{13}C$-HSQC spectrum taken at $\delta_2(^{13}C) = 130.50$ ppm (F171$C^{\delta}$). The NOE cross-peaks between F171$H^{\delta}$ and V58$H^{\gamma}$ are at the intersection of the dashed lines drawn at $\delta_2(^1H) = 7.11$ (F171$H^{\delta}$) and $\delta_1(^1H) = 0.76$ (V58$H^{\gamma 1}$) and 0.52 ppm (V58$H^{\gamma 2}$). The cross-peak between F171$H^{\delta}$ and V58$H^{\gamma 2}$ appears to be slightly shifted in the $\delta_1$-dimension because of an overlapping NOE cross-peak between F171$H^{\delta}$ and V29$H^{\gamma}$ or V101$H^{\gamma}$. (**C**) A 2D (H)CB(CGCC-TOCSY)$H^{ar}$ spectrum correlating the $^{13}C^{\beta}$ with the aromatic $^1H$ chemical shifts of phenylalanine residues (28). The dashed lines are drawn at $\delta_1(^{13}C) = 38.93$ (F171$C^{\beta}$) and $\delta_2(^1H) = 7.11$ ppm (F171$H^{\delta}$). This spectrum proves that the resonances marked by the dashed line belong to Phe171 rather than to Phe60, which also produces NOE cross-peaks to Val58. Only one of the aromatic signals of Phe171 is well resolved, whereas the others overlap with the stronger peaks from Phe147 and Phe177.

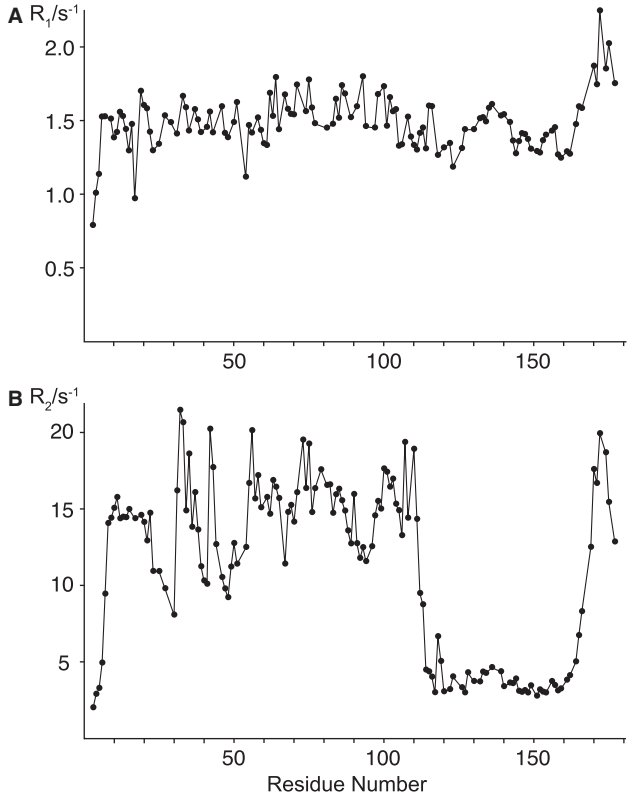


**Figure 6.** $^{15}N$ relaxation rates measured of a 0.35 mM solution of $^{15}N$-labeled SSB at pH 3.4 and 25°C at a $^1H$ NMR frequency of 600 MHz. (**A**) Longitudinal relaxation rates. (**B**) Transverse relaxation rates.

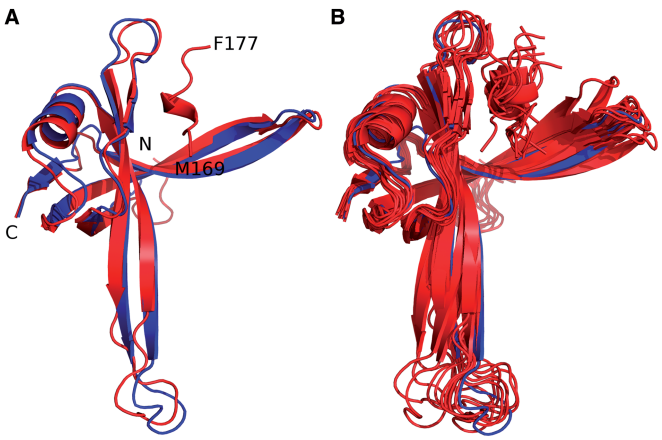


**Figure 7.** Final structures selected from PCS-Rosetta computations. (**A**) Superposition of the lowest scoring structure (red) with a SSB monomer from the crystal structure 1EYG (blue). The N- and C-termini of the OB-domain are labeled as well as residues 169 and 177 of the C-peptide. (**B**) Bundle of 10 computed lowest scoring structures from PCS-Rosetta simulations.

polypeptide chain with a C-terminal segment (C-peptide) that can bind to the ssDNA binding site of the OB-domain and therefore would be expected to compete with ssDNA binding. We were able to pinpoint the binding site by specific NOEs between the C-peptide and Val29 and Val58 of the OB-domain, which are near several positively charged arginine residues. Considering the high negative charge of the C-peptide, electrostatic attraction undoubtedly is an important component of the interaction.

A predominantly electrostatic interaction between the C-peptide and the OB-domain may explain the difficulty to define the structure of the bound C-peptide with atomic resolution. As electrostatic forces decrease considerably more slowly with distance than hydrophobic or hydrogen-bonding interactions, the binding site may accommodate different orientations of the C-peptide with similar binding energies, an example of a 'fuzzy' interaction (36). Structural disorder of the bound C-peptide may be suggested by the failure to detect the bound C-peptide in any of the crystal structures of SSBs from any organism determined to date. Alternatively, the bound state may be only partially populated. Both explanations, structural disorder and transient interactions, would explain the increased mobility of the C-peptide relative to the OB-domain provided by the $R_1(^{15}N)$ relaxation rates (Figure 6A) and the scarcity and weak

intensities of the NOEs between the C-peptide and the OB-domain (Figure 5). In calculating a model of the complex between OB-domain and C-peptide, we obtained better structural definition of the C-peptide by assuming that its PCSs needed to be doubled because the bound state is populated only half of the time. Although this assumption is arbitrary, using the PCSs at face value generated highly disordered structures of the C-peptide. Owing to the NOEs with the OB-domain, the C-peptide was still at the same site as in Figure 7, but with little definition of its structure or direction of approach to its binding site.

### Flexibility of the C-domain and biological implications

Previous biochemical experiments with C-terminally truncated *E. coli* SSB accumulated strong evidence that the C-peptide binds to the OB-domain in competition with ssDNA (16–19). Comparison of the C-peptide binding site identified in the present study with the crystal structure of the complex with ssDNA (9) shows that the competition between C-peptide and ssDNA is direct. To form the structure of the SSB–ssDNA complex, ssDNA has to displace the C-peptide from the OB-domain, increasing its availability for binding to ssDNA processing enzymes as hypothesized a long time ago (16).

In addition, the high flexibility of the polypeptide connecting the C-peptide to the OB-domain and its apparent lack of specific binding interactions with the OB-domain suggests that, in the tetramer at neutral pH, the C-peptide could easily bind to different OB-domains of the tetramer, explaining the increased rate of exchange between different SSB monomers in the tetramer of a SSB mutant devoid of the C-terminal eight residues compared with full-length *E. coli* SSB (19). These experiments also highlighted the electrostatic character of the C-peptide–OB-domain interactions by demonstrating their destabilization in the presence of 1 M ammonium acetate. Clearly, binding interactions between different SSB tetramers could explain the relatively poor solubility of SSB at neutral pH. At the pH of the present study (pH 3.4), it is likely that some of the carboxyl groups of the C-peptide have lost their negative charge, i.e. we expect binding to the OB-domain would be stronger at neutral pH.

Although tetramerization and a different pH could, in principle, affect the binding mode of the C-domain to the OB-domain, supporting evidence for the importance of the C-peptide in organizing the C-domain on the OB-domain tetramer arises from the crystal structure of full-length *E. coli* SSB (PDB code 1QVC) (12), which seemed to have lost its C-peptide by proteolysis during crystallization, resulting in different conformations for each of the four C-domain segments spanning residues 113–145. In agreement with a random coil conformation, the C-domain has long been known to be highly sensitive to proteolytic digestion (16). We find the evidence for extended C-domain segments in the 1QVC structure to be weak. Reinterpretation of the electron density map reveals an error in register of the sequence from residue 90 onward in the OB-fold, and discontinuous and weak density with no stabilizing interactions for any of the residues beyond Gly114; omission from the model of the segments from Arg115 onward and correction of the residue alignment resulted in a decrease of $R/R_{free}$ from 0.247/0.317 (12) to 0.210/0.262 (Supplementary Figure S3).

## CONCLUSIONS

By defining the binding site of the C-terminal peptide of *E. coli* SSB on the OB-domain, the present study strongly suggests that competition between ssDNA and the C-peptide is direct. The binding site confirms previous results indicating that electrostatic interactions play an important part in the C-peptide–OB-domain interaction. By populating a large conformational space, the intrinsically disordered C-domain is capable of interacting with different SSB monomers in the tetramer as well as with different tetramers, which could contribute to the reported cooperativity effects of ssDNA binding (37). Conceivably, the flexibility of the C-domain also facilitates capture of the multitude of proteins that bind to the C-peptide. Although the present studies were conducted at low pH to produce a monomeric state suitable for NMR analysis, the structural results are in excellent agreement with published biochemical data obtained for the SSB tetramer at neutral pH.

## ACCESSION NUMBERS

Biological Magnetic Resonance Data Bank: Resonance assignments for monomeric SSB at pH 3.4 (accession code 19446). Protein Data Bank: The revised 2.2 Å coordinates of *E. coli* SSB have been deposited with accession number 4MZ9.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online, including [38].

## ACKNOWLEDGEMENTS

The authors thank Mr Kala Bharath Pilla and Dr Hiromasa Yagi for helpful discussions.

## FUNDING

Australian Research Council [DP0877658 to N.E.D and A.J.O., DP0984797 to N.E.D. and T.H., FT0990287 to A.O., FT09917009 to T.H., DP110102737 to G.O., DP120100561 to T.H., G.O. and B.G.]. Funding for open access charge: Novartis Institute for Tropical Diseases.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Meyer,R.R. and Laine,P.S. (1990) The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.*, **54**, 342–380.

2. Lohman,T.M. and Ferrari,M.E. (1994) *Escherichia coli* single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu. Rev. Biochem.*, **63**, 527–570.
3. Shereda,R.D., Kozlov,A.G., Lohman,T.M., Cox,M.M. and Keck,J.L. (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.*, **43**, 289–318.
4. Richard,D.J., Bolderson,E. and Khanna,K.K. (2009) Multiple human single-stranded DNA binding proteins in genome maintenance: structural, biochemical and functional analysis. *Crit. Rev. Biochem. Mol. Biol.*, **44**, 98–116.
5. Costes,A., Lecointe,F., McGovern,S., Quevillon-Cheruel,S. and Polard,P. (2010) The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS Genet.*, **6**, e1001238.
6. Robinson,A., Brzoska,A.J., Turner,K.M., Withers,R., Harry,E.J., Lewis,P.J. and Dixon,N.E. (2010) Essential biological processes of an emerging pathogen: DNA replication, transcription and cell division in *Acinetobacter spp. Microbiol. Mol. Biol. Rev.*, **74**, 273–297.
7. Robinson,A., Causer,R.J. and Dixon,N.E. (2012) Architecture and conservation of the bacterial DNA replication machinery, an underexploited drug target. *Curr. Drug Targets*, **13**, 352–372.
8. Chase,J.W. and Williams,K.R. (1986) Single-stranded DNA binding proteins required for DNA replication. *Annu. Rev. Biochem.*, **55**, 103–136.
9. Raghunathan,S., Kozlov,A.G., Lohman,T.M. and Waksman,G. (2000) Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat. Struct. Biol.*, **7**, 648–652.
10. Raghunathan,S., Ricard,C.S., Lohman,T.M. and Waksman,G. (1997) Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength X-ray diffraction on the selenomethionyl protein at 2.9-Å resolution. *Proc. Natl Acad. Sci. USA*, **94**, 6652–6657.
11. Yang,C., Curth,U., Urbanke,C. and Kang,C.H. (1997) Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nat. Struct. Biol.*, **4**, 153–157.
12. Matsumoto,T., Morimoto,Y., Shibata,N., Kinebuchi,T., Shimamoto,N., Tsukihara,T. and Yasuoka,N. (2000) Roles of functional loops and C-terminal segments of a single-stranded DNA binding protein elucidated by X-ray structural analysis. *J. Biochem.*, **127**, 329–335.
13. Savvides,S.N., Raghunathan,S., Fütterer,K., Kozlov,A.G., Lohman,T.M. and Waksman,G. (2004) The C-terminal domain of full-length *E. coli* SSB is disordered even when bound to DNA. *Protein Sci.*, **13**, 1942–1947.
14. Kaushal,P.S., Singh,P., Sharma,A., Muniyappa,K. and Vijayan,M. (2010) X-ray and molecular-dynamics studies on *Mycobacterium leprae* single-stranded DNA-binding protein and comparison with other eubacterial SSB structures. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 1048–1058.
15. Lu,D. and Keck,J.L. (2008) Structural basis of *Escherichia coli* single-stranded DNA-binding protein stimulation of exonuclease I. *Proc. Natl Acad. Sci. USA*, **105**, 9169–9174.
16. Williams,K.R., Spicer,E.K., LoPresti,M.B., Guggenheimer,R.A. and Chase,J.W. (1983) Limited proteolysis studies on the *Escherichia coli* single-stranded DNA binding protein. Evidence for a functionally homologous domain in both the *Escherichia coli* and T4 DNA binding proteins. *J. Biol. Chem.*, **258**, 3346–3355.
17. Curth,U., Genschel,J., Urbanke,C. and Greipel,J. (1996) *In vitro* and *in vivo* function of the C-terminus of *Escherichia coli* single-stranded DNA binding protein. *Nucleic Acids Res.*, **24**, 2706–2711.
18. Kozlov,A.G., Cox,M.M. and Lohman,T.M. (2010) Regulation of single-stranded DNA binding by the C termini of *Escherichia coli* single-stranded DNA-binding (SSB) protein. *J. Biol. Chem.*, **285**, 17246–17252.
19. Mason,C.E., Jergic,S., Lo,A.T.Y., Wang,Y., Dixon,N.E. and Beck,J.L. (2013) *Escherichia coli* single-stranded DNA-binding protein: NanoESI-MS studies of salt-modulated subunit exchange and DNA binding transactions. *J. Am. Soc. Mass Spectrom.*, **24**, 274–285.
20. Marintcheva,B., Marintchev,A., Wagner,G. and Richardson,C.C. (2008) Acidic C-terminal tail of the ssDNA-binding protein of bacteriophage T7 and ssDNA compete for the same binding surface. *Proc. Natl Acad. Sci. USA*, **105**, 1855–1860.
21. Elvin,C.M., Thompson,P.R., Argall,M.E., Hendry,P., Stamford,N.P.J., Lilley,P.E. and Dixon,N.E. (1990) Modified bacteriophage lambda promoter vectors for overproduction of proteins in *Escherichia coli*. *Gene*, **87**, 123–126.
22. Williams,N.K., Prosselkov,P., Liepinsh,E., Line,I., Sharipo,A., Littler,D.R., Curmi,P.M.G., Otting,G. and Dixon,N.E. (2002) *In vivo* protein cyclization promoted by a circularly-permuted *Synechocystis sp.* PCC6803 DnaB mini-intein. *J. Biol. Chem.*, **277**, 7790–7798.
23. Swarbrick,J.D., Ung,P., Chhabra,S. and Graham,B. (2011) An iminodiacetic acid based lanthanide binding tag for paramagnetic exchange NMR spectroscopy. *Angew. Chem. Int. Ed.*, **50**, 4403–4406.
24. Yagi,H., Maleckis,A. and Otting,G. (2013) A systematic study of labeling an α-helix in a protein with a lanthanide using IDA or NTA tags. *J. Biomol. NMR*, **55**, 157–166.
25. Graham,B., Loh,C.T., Swarbrick,J.D., Ung,P., Shin,J., Yagi,H., Jia,X., Chhabra,S., Pintacuda,G., Huber,T. *et al.* (2011) A DOTA-amide lanthanide tag for reliable generation of pseudocontact shifts in protein NMR spectra. *Bioconjug. Chem.*, **22**, 2118–2125.
26. Grzesiek,S., Anglister,J. and Bax,A. (1993) Correlation of backbone amide and aliphatic side-chain resonances in $^{13}C/^{15}N$-enriched proteins by isotropic mixing of $^{13}C$ magnetization. *J. Magn. Reson. Ser. B*, **101**, 114–119.
27. Kay,L.E., Xu,G.Y., Singer,A.U., Muhandiram,D.R. and Forman-Kay,J.D. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain $^{1}H$ and $^{13}C$ correlations in $H_2O$ samples for proteins. *J. Magn. Reson. Ser. B*, **101**, 333–337.
28. Löhr,F., Hänsel,R., Rogov,V.V. and Dötsch,V. (2007) Improved pulse sequences for sequence specific assignment of aromatic proton resonances in proteins. *J. Biomol. NMR*, **37**, 205–224.
29. 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. and Kay,L. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by $^{15}N$ NMR relaxation. *Biochemistry*, **33**, 5984–6003.
30. Vranken,W.F., Boucher,W., Stevens,T.J., Fogh,R.H., Pajon,A., Llinas,M., Ulrich,E.L., Markley,J.L., Ionides,J. and Laue,E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, **59**, 687–696.
31. John,M., Park,A.Y., Pintacuda,G., Dixon,N.E. and Otting,G. (2005) Weak alignment of paramagnetic proteins warrants correction for residual CSA effects in measurements of pseudocontact shifts. *J. Am. Chem. Soc.*, **127**, 17190–17191.
32. Schmitz,C., Stanton-Cook,M.J., Su,X.C., Otting,G. and Huber,T. (2008) Numbat: an interactive software tool for fitting $\Delta\chi$-tensors to molecular coordinates using pseudocontact shifts. *J. Biomol. NMR*, **41**, 179–189.
33. Schmitz,C., Vernon,R., Otting,G., Baker,D. and Huber,T. (2012) Protein structure determination from pseudocontact shifts using ROSETTA. *J. Mol. Biol.*, **416**, 668–677.
34. Kim,D.E., Chivian,D. and Baker,D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.*, **32**, W526–W531.
35. Shishmarev,D. and Otting,G. (2013) How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study. *J. Biomol. NMR*, **56**, 203–216.
36. Tompa,P. and Fuxreiter,M. (2007) Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.*, **33**, 2–8.
37. Roy,R., Kozlov,A.G., Lohman,T.M. and Ha,T. (2007) Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J. Mol. Biol.*, **369**, 1244–1257.
38. Wishart,D.S., Sykes,B.D. and Richards,F.M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, **31**, 1647–1651.