Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the single-stranded DNA-binding protein from *Streptomyces coelicolor*

Zoran Štefanić, Dušica Vujaklija and Marija Luić*

Rudjer Bošković Institute, PO Box 180, Bijenička cesta 54, Zagreb, Croatia

Correspondence e-mail: marija.luic@irb.hr

Received 18 May 2009
Accepted 19 June 2009

PDB Reference: single-stranded DNA-binding protein, 3eiv, r3eivsf.

The crystal structure of the single-stranded DNA-binding protein (SSB) from *Streptomyces coelicolor*, a filamentous soil bacterium with a complex life cycle and a linear chromosome, has been solved and refined at 2.1 Å resolution. The three-dimensional structure shows a common conserved central OB-fold that is found in all structurally determined SSB proteins. However, it shows variations in quaternary structure that have previously only been found in mycobacterial SSBs. The strand involved in the clamp mechanism characteristic of this type of quaternary structure leads to higher stability of the homotetramer. To the best of our knowledge, this is the first X-ray structure of an SSB protein from a member of the genus *Streptomyces* and it was predicted to be the most stable of the structurally characterized bacterial or human mitochondrial SSBs.

1. Introduction

Single-stranded DNA-binding (SSB) proteins are involved in the replication, recombination and repair of DNA. SSB proteins bind ssDNA and protect transiently formed DNA single strands from nucleolytic digestion and the formation of unproductive secondary structures. The importance of SSBs in maintaining genomic integrity makes them indispensable for all cellular life (Mushegian & Koonin, 1996). Although all SSB proteins perform similar functions, they show very little sequence similarity. Nevertheless, some common structural features that unite this class of proteins across evolution can be recognized. The first is a structurally conserved folding motif called an oligonucleotide–oligosaccharide–oligopeptide-binding fold (OB-fold) defined by Murzin (1993). The OB-fold constitutes approximately the first 110 amino-acid residues of the N-terminal domain of the protein and is involved in DNA binding. The C-terminal domain is usually rich in glycine and proline residues and is responsible for interactions with enzymes of DNA metabolism.

The second common feature of most SSBs is obligate oligomerization. The structure of most bacterial and eukaryotic mitochondrial SSBs in their stable oligomeric form represents an arrangement of four OB-folds which act together in binding ssDNA. They are homotetramers, except for the SSB proteins from the thermophilic genera *Thermus* (Jędrzejczak *et al.*, 2006; Fedorov *et al.*, 2006) and *Deinococcus* (Bernstein *et al.*, 2004), which are homodimers. Since each monomer in such homodimers consists of two OB-folds fused together by a linker, they also contain four OB-folds.

In 1997, the first SSB crystal structures were determined: after extensive crystallization efforts, the three-dimensional X-ray structure of the *Escherichia coli* SSB protein was solved

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	*I*222
Unit-cell parameters (Å)	a = 101.4, b = 104.8, c = 163.3
Multiplicity	8
No. of molecules in ASU	1 (tetramer)
Wavelength (Å)	0.97
Resolution range (Å)	31.18–2.14 (2.21–2.14)
Total No. of reflections	226705
Unique reflections	48148
Completeness (%)	99.7 (99.9)
Mean $I/\sigma(I)$	13.4 (1.5)
$R_{\rm merge}$	0.074 (0.35)
Refinement	
R factor	0.230
$R_{\rm free}$	0.259
Reflections in working set/test set	45711/2437
No. of protein atoms	3146
No. of solvent atoms	288
Average B factor (Å^2)	46.8
R.m.s.d. bonds (Å)	0.006
R.m.s.d. angles (°)	0.915
Ramachandran plot: No. of residues	
Favoured	346 [95.8%]
Allowed	15 [4.2%]
Outliers	0

by Raghunathan and coworkers (PDB code 1kaw; Raghunathan *et al.*, 1997), while Yang and coworkers published the three-dimensional structure of the closely analogous human mitochondrial SSB (PDB code 3ull; Yang *et al.*, 1997). Since then, the crystal structures of SSBs from the following sources have been determined: *Mycobacterium tuberculosis* (PDB code 1ue1; Saikrishnan *et al.*, 2003), the archeon *Sulfolobus solfataricus* (PDB code 1o7i; Kerr *et al.*, 2003), *Deinococcus radiodurans* (PDB code 1se8; Bernstein *et al.*, 2004), *M. smegmatis* (PDB code 1x3e; Saikrishnan *et al.*, 2005), *Thermatoga maritima* (PDB code 1z9f; DiDonato *et al.*, 2006), *Thermus aquaticus* (PDB code 2fxq, Jędrzejczak *et al.*, 2006; PDB code 2ihe, Fedorov *et al.*, 2006), *Mycoplasma pneumoniae* (PDB code 2hql; Das *et al.*, 2007) and *Helicobacter pylori* complexed with ss-DNA (PDB code 2vw9; Chan *et al.*, 2009).

We have previously reported (Mijaković *et al.*, 2006) a eukaryotic type of post-translational modification of distantly related bacterial SSB proteins. As part of our deep interest in understanding the biological significance of this modification in antibiotic-producing bacteria with high G+C content (Hopwood, 2006), we have determined the crystal structure of *Streptomyces coelicolor* SSB. This is the first three-dimensional structure of an SSB from a member of the genus *Streptomyces* and provides a structural framework for further biochemical and genetic investigations.

2. Materials and methods

2.1. Crystallization and data collection

Expression of the gene encoding the SSB protein, purification and the preliminary crystallization conditions obtained using Crystal Screen Cryo (Hampton Research) have been described previously (Štefanić *et al.*, 2007). A crystal grown at 291 K in a drop composed of 2 µl protein solution (15.5 mg ml^{-1} in Tris–HCl pH 7.0) and 2 µl reservoir solution [0.085 *M* Na HEPES pH 7.5, 1.7%(*v*/*v*) PEG 400, 1.7 *M* ammonium sulfate] was used for data collection on the DND-CAT ID5 beamline of the Advanced Photon Source synchrotron-radiation facility, Argonne National Laboratory, Chicago. A complete data set to a resolution of 2.1 Å was collected at 100 K using 0.97 Å radiation and a MAR CCD 165 detector. A total of 120 images of 1° rotation were collected. The protein crystallized in the orthorhombic space group *I*222, with unit-cell parameters a = 101.4, b = 104.8, c = 163.3 Å. The *XGEN* package (Howard, 2000) was used for data processing and scaling. The program *TRUNCATE* (French & Wilson, 1978) was used to convert intensities to structure factors. The resulting diffraction data statistics are presented in Table 1.

Table 2
Summary of the molecular-replacement procedure for *S. coelicolor* SSB.

Resolution range (Å)	20–4
Rotation and translation function	
Best solution (first dimer)	θ = 132.1, φ = 144.1, χ = 162.5, t_x = 0.85, t_y = 0.04, t_z = 0.37
Correlation coefficient	0.247
R factor	0.611
Best solution (second dimer)	θ = 37.3, φ = 144.1, χ = 162.5, t_x = 0.65, t_y = 0.74, t_z = 0.37
Correlation coefficient	0.598
R factor	0.475

2.2. Structure determination and refinement

The structure was solved by molecular replacement with *MOLREP* (Vagin & Teplyakov, 1997) using SSB from *M. smegmatis* as a search model (PDB code 1x3e; Saikrishnan *et al.*, 2005). The search model was a dimer (two monomers: chains *A* and *B*) pruned of the loops extending from the core of the protein. The best solution for the first dimer was fixed before the search for the location of the second dimer. A summary of the molecular-replacement procedure is given in Table 2. After rigid-body refinement using the *CNS* program (Brünger *et al.*, 1998), the R factor fell from 47.7% to 44.9%. In order to improve the phases, solvent flattening and NCS density averaging were performed by *CNS* (Brünger *et al.*, 1998). To minimize the model bias, the phases obtained by molecular replacement were input into *ARP*/*wARP* (Langer *et al.*, 2008). The program automatically built around 120 residues in four subunits and also generated water molecules. This model was then refined with *PHENIX* (Adams *et al.*, 2002) in a number of iterations guided by $2F_{\rm o} - F_{\rm c}$ and $F_{\rm o} - F_{\rm c}$ electron-density maps to a final R factor of 23.0% and an $R_{\rm free}$ of 25.9%. The Ramachandran plot shows the good quality of the refined protein model.

3. Results and discussion

3.1. Monomer structure

The crystal structure of *S. coelicolor* SSB was solved by molecular replacement using the *M. smegmatis* SSB molecule

(PDB code 1x3e) as a search model. The structure of the subunit is similar to those of all other four-OB-fold SSBs, while the quaternary structure is similar to the recently described unique quaternary structure of mycobacterial SSBs. There are four subunits, *i.e.* one tetramer, in the asymmetric unit of the orthorhombic *I*222 space group.

S. coelicolor SSB encodes a 199-residue single-stranded DNA-binding protein with a molecular mass of 19.9 kDa (Štefanić *et al.*, 2007). The 120 N-terminal residues form a common OB-fold constituting the DNA-binding domain (Fig. 1).

The structure of this domain is very similar in all bacterial SSB proteins. It is characterized by three long β-hairpin loops extending out of a five-stranded β-barrel capped by an α-helix. These three loops show extreme mobility (Saikrishnan *et al.*, 2003). In *S. coelicolor*, loop 1 consists of residues 22–26, loop 2 of residues 36–52 and loop 3 of residues 90–94. The C-terminus of the OB-fold domain extends to a strand (strand 9, amino acids 113–119) which forms a β-sheet with the equivalent strand from the neighbouring subunit, thus making a clamp mechanism (Saikrishnan *et al.*, 2003). This feature is observed only in two other structures, the SSBs from *M. smegmatis* and *M. tuberculosis*. The stretch of seven amino acids (ATAKVTK in *S. coelicolor*) forming strand 9 is conserved in these three structures. Only one amino acid, a threonine, is replaced by asparagine in the *Mycobacterium* proteins. A multiple sequence alignment of the DNA-binding domains of the SSBs with solved crystal structures from bacteria and mitochondria shows that β-strand 9 is an insertion at the C-terminus of the OB-fold in mycobacterial and *Streptomyces* SSBs. This insertion is characteristic of high-G+C Gram-positive bacteria (Saikrishnan *et al.*, 2005).

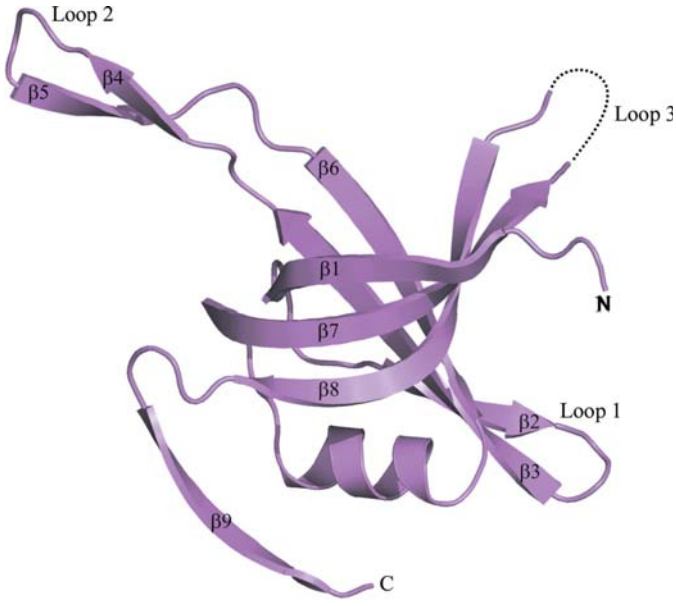

Figure 1
Ribbon representation of the crystal structure of the *S. coelicolor* SSB. The nomenclature of the secondary-structure elements follows that originally proposed for the OB-fold (Murzin, 1993). There are seven β-strands, one α-helix and three loops forming a β-barrel capped by the α-helix. All figures were prepared using *PyMOL* (DeLano, 2002).

The C-terminal domain extends from residues 120 to 199 and is not seen in the electron-density maps, although the full-length protein was crystallized. The C-terminal domain is common to all bacterial SSBs but has so far eluded structural characterization. According to *FoldIndex* (Prilusky *et al.*, 2005; http://bioportal.weizmann.ac.il/fldbin/findex) this part of the protein is defined as unstructured, *i.e.* unfolded, which is not surprising since *S. coelicolor* SSB has 50% glycine residues in the C-terminal domain (Štefanić *et al.*, 2007). Similarly to other SSBs (Chan *et al.*, 2009), this domain also possesses an acidic C-terminus (^{194}DEPPF). The C-terminus mediates protein–protein interactions and probably becomes structured upon interacting with its protein partners. The fact that the *R* and $R_{\rm free}$ values for the refined model are very reasonable (Table 1) although a large part of the protein is missing also confirms that this part is disordered and does not contribute significantly to the diffraction intensities.

Although none of the four chains could be located in the electron density in its entire length, every amino acid from the OB-fold domain is visible in at least one of the chains. Structural superpositions of C^{α} atoms from four chains were carried out and the r.m.s. deviations were calculated using the *SUPERPOSE* program (Krissinel & Henrick, 2004). These values ranged from 0.32 to 0.56 Å.

3.2. Architecture of tetramers

SSB tetramers have 222 molecular symmetry. The interfaces between individual monomers of the SSBs fall into two prin-

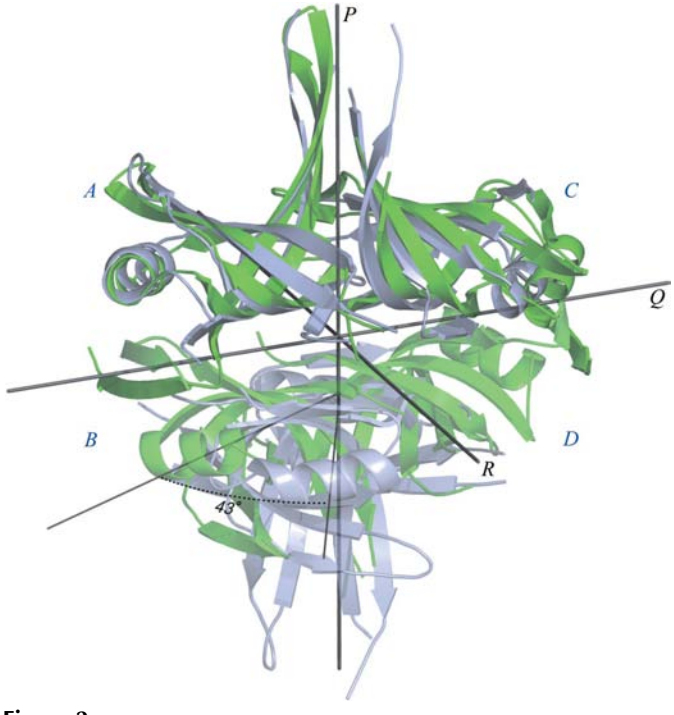

Figure 2
Superposition of the quaternary structures of *S. coelicolor* (green) and *E. coli* (violet) SSBs. Dyad axes of molecular 222 symmetry are marked *P*, *Q* and *R* according to Saikrishnan *et al.* (2003). The angle of 43° corresponds to the rotation of *BD* subunits of *E. coli* with respect to the *BD* subunits of *S. coelicolor*. The standard naming of the monomers is given.

Table 3
Interface areas between two monomers with their corresponding free energies of assembly dissociation (ΔG^{diss} values obtained from the *PISA* service; Krissinel & Henrick, 2007) for some bacterial and human mitochondrial SSBs.

For the whole tetramers, the values of the solvent-accessible surface area, the solvent-accessible surface area of monomeric units buried upon assembly formation and ΔG^{diss} values are given. Crystallographic symmetry is explicitly given when it coincides with molecular symmetry and the names of monomers are changed accordingly. 1 kcal mol^{-1} = 4.186 kJ mol^{-1}.

	Interface (1)		Interface (2)		Tetramer		
Source	Area (Å^2)	ΔG^{diss} (kcal mol^{-1})	Area (Å^2)	ΔG^{diss} (kcal mol^{-1})	Surface area (Å^2)	Buried area (Å^2)	ΔG^{diss} (kcal mol^{-1})
E. coli (PDB code 1sru)	1014 (*AC*)	3.7	204 (*AB*)	−6.8	19270	6300	4.6
	1106 (*BD*)	3.9	192 (*CD*)	−6.8			
Human mitochondrial (PDB code 2dud)	1218 (*AB*)	1.6	238 (*AA*^{i})	−8.3	19890	7170	2.4
	1218 (*A*^{i}*B*^{i})	1.6	244 (*BB*^{i})	−7.4			
M. smegmatis (PDB code 1x3e)	790 (*AA*^{ii})	−6.2	1380 (*AB*)	15.9	24520	9090	6.0
	876 (*BB*^{ii})	−5.6	1380 (A^{ii}B^{ii})	16.0			
M. tuberculosis (PDB code 1ue1)	852 (*AA*^{iii})	−4.5	1383 (*AB*)	13.7	25830	9120	7.1
	773 (*BB*^{iii})	−4.5	1383 (*A*^{iii}*B*^{iii})	13.7			
S. coelicolor (PDB code 3eiv)	830 (*AC*)	−2.0	1471 (*AB*)	21.1	20450	9320	11.7
	931 (*BD*)	3.0	1415 (*CD*)	20.3			

Symmetry codes: (i) $-x+2, -x+y+1, -z+1/3$, (ii) $y, x, -z+1$, (iii) $x-y+1, -y+2, -z+2/3$.

cipal groups: (1) the interface between chains *A* and *C* (or *B* and *D*) and (2) the interface between *A* and *B* (or *C* and *D*; Fig. 2). According to the literature (Saikrishnan *et al.*, 2003, 2005), the SSB tetramers can be viewed as consisting of two dimers, where the dimer is defined as those two monomers which share the largest interface area between them (Table 3). In the case of the type (1) interfaces there is almost no difference between *E. coli* and mycobacterial and *S. coelicolor* SSBs. In contrast, the type (2) interfaces areas in these two groups differ significantly (Table 3). To further quantify the stability of dimers with type (1) and (2) interfaces, we have calculated their free energies of assembly dissociation (ΔG^{diss}). These values correlate very well with interface-area values.

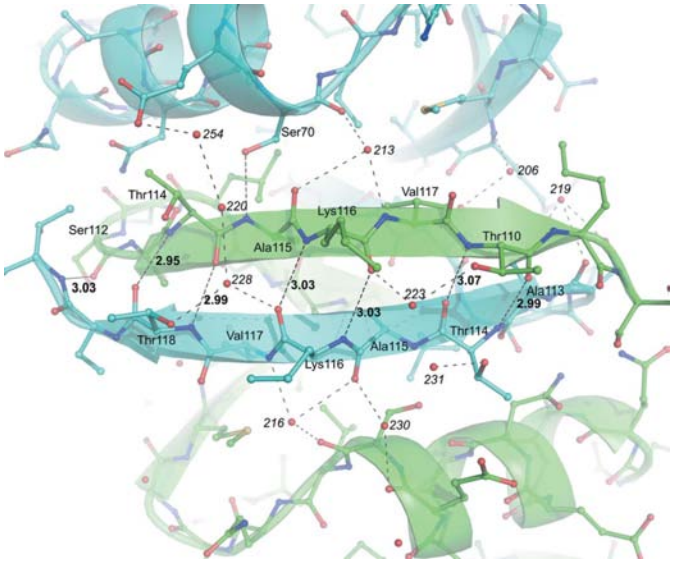

Figure 3
Possible hydrogen bonds between strands 9 from monomer *A* (turquoise) and *B* (green). The view is along the *Q* dyad axes. Distances are given for some of the contacts between the two chains (in bold). Water molecules are labelled with numbers in italic and their contacts are shown in grey.

Therefore, we can say that in the *E. coli* SSB protein subunits *A* and *C* form a dimer, while in mycobacterial and *S. coelicolor* SSBs the dimer is formed by the *A* and *B* subunits (Fig. 2). Their quaternary structures are related by rotating together subunits *B* and *D* by 43° (Saikrishnan *et al.*, 2005) while keeping subunits *A* and *C* fixed. The axis of rotation is the molecular dyad axis designated *P* in Fig. 2. The quaternary structure of SSB homodimers from *Thermus* and *Deinococcus* could be described as an intermediate between the *E. coli* and mycobacterial structures. For example, rotation of subunit *B* in *D. radiodurans* SSB by −25° would give the *E. coli* arrangement of OB-folds and rotation by 21° would give that of *M. smegmatis* SSB (Saikrishnan *et al.*, 2005).

The solvent-accessible surface area of the monomeric units buried upon oligomer formation is a measure of its stability. As previously mentioned, the *E. coli* tetramers consist of *AC* (*BD*) dimers and the *Mycobacterium* and *Streptomyces* tetramers of *AB* (*CD*) dimers. Comparison of all interface areas given in Table 3 shows that the *Mycobacterium*/*Streptomyces*-type dimers display the largest area. Moreover, a clamping mechanism in these dimers fixes the orientation of the subunits and probably adds to their stability. Two strands 9 from one dimer make an antiparallel β-sheet sandwiched between two α-helices (Fig. 3). Strand 9 from one monomer is tightly bound between strand 9 and a helix from the other monomer of the same dimer. Possible hydrogen bonds between strand 9 and the neighbouring amino acids are depicted in Fig. 3. The dimer is further stabilized by two salt bridges formed between Arg76 and Glu105 from both monomers.

Comparison of the surfaces buried on tetramerization in bacterial and mitochondrial SSBs shows a significant increase from the *E. coli* and human mitochondrial type of quaternary structure toward the *Mycobacterium* and *Streptomyces* type (Table 3). The same tendency is evident from the values of the free energy of assembly dissociation. The larger the ΔG^{diss} value, the more thermodynamically stable the assembly is. On

the basis of these two criteria, it could be predicted that *S. coelicolor* tetramers are more stable than other SSB tetramers described to date.

Mycobacterium spp. and *Streptomyces* spp. belong to distantly related genera of actinobacteria. The first are slow-growing widespread bacteria with some pathogenic species, while the second are soil-inhibiting filamentous bacteria best known as antibiotic producers and as such of significant biotechnological interest. As mentioned previously, their SSB structures share significant similarity, most likely owing to a special requirement of high-G+C-content genomes. There have been many studies on the biochemical, DNA-binding (Purnapatre & Varshney, 1999; Handa *et al.*, 2000, 2001; Reddy *et al.*, 2001; Sikder *et al.*, 2001; Acharya & Varshney, 2002) and structural (Saikrishnan *et al.*, 2003, 2005) properties of mycobacterial SSBs. In contrast, *S. coelicolor* SSB still remains to be characterized. We have previously demonstrated tyrosine phosphorylation of the *Bacillus subtilis*, *S. coelicolor* and *E. coli* SSBs. Tyr82 has been identified as the phosphorylation site in the *B. subtilis* SSB (Mijaković *et al.*, 2006). This residue is highly conserved in the *S. coelicolor* SSB as well as in SSBs from other Gram-positive bacteria, while it is absent from *E. coli* SSB. We have compared the crystal structures of the *S. coelicolor* and *E. coli* SSBs and the predicted structure of *B. subtilis* SSB (A. Kriško, unpublished data). It was found that Tyr82 in *B. subtilis*, Tyr88 in *S. coelicolor* and Tyr98 in *E. coli* are all located in loop 3 and that they all occupy similar positions in the structure. This strongly indicates the presence of phosphorylation sites in the *S. coelicolor* and *E. coli* SSBs.

To date, there are only two known single-stranded DNA–SSB complexes: those from *E. coli* (PDB code 1eyg, Raghunathan *et al.*, 2000; PDB code 1sru, Savvides *et al.*, 2004) and a very recently reported SSB–DNA complex from *Helicobacter pylori* (PDB code 2vw9; Chan *et al.*, 2009). Owing to the variation in the quaternary structure between *E. coli* SSB on one hand and mycobacterial and *Streptomyces* SSBs on the other, the path that DNA adopts to wrap around it is expected also to be different. The DNA-binding surface in the case of *E. coli* SSB is an approximate spheroid, while that in *Mycobacteria* and *Streptomyces* SSBs is an ellipsoid, suggesting that the path of DNA wrapping is shorter in the latter structures. A different mutual orientation between the *AC* and *BD* subunits in these two types of quaternary structures also indicates large differences in DNA wrapping around the SSB tetramer. In the *H. pylori* complex different ssDNA-binding paths and binding surfaces have been found compared with the *E. coli* complex. According to the authors, this discrepancy in binding could be partially ascribed to the different rotation angles between the dimers. Since the *H. pylori* SSB–DNA complex structure is not available in the PDB at the time of writing, a direct comparison with the title structure is not possible. However, the genome of *H. pylori* contains a relatively low GC content (an average of 32.5%), compared with 65.6% (Cole *et al.*, 1998) and 72.1% (Bentley *et al.*, 2002) in *M. tuberculosis* and *S. coelicolor*, respectively, and the sequence alignment of its SSB with *S. coelicolor* and *M. tuberculosis* SSBs displays very low sequence identity (15% and 21%, respectively). Furthermore, the stretch of seven amino acids that forms strand 9 in high-G+C-content bacteria is absent from *H. pylori* SSB. Taking all these facts into account, a new mode of DNA binding to SSB is expected for the *S. coelicolor* tetramer.

Very recently, Katre and Suresh analyzed 650 unique homotetrameric protein structures from the PDB at a sequence-identity cutoff of 70% in order to better understand tetrameric associations and their possible biological relevance (Katre & Suresh, 2009). According to this analysis, homotetramers can be divided into four categories: (i) tetramers with a nearly perfect tetrahedral arrangement of monomers, (ii) tetramers composed of two dimers perpendicular to each other, (iii) tetramers whose monomers lie in one plane and the corresponding sides of neighbouring monomers face in opposite directions and (iv) tetramers with monomers lying in one plane but arranged in a head-to-tail manner. The analysis of SSBs with known crystal structures shows that they fall into the second and third groups. Interestingly, some DNA-binding proteins are among the very few homotetrameric proteins that could not be grouped into the four mentioned categories (Katre & Suresh, 2009).

Using the same approach of representing one monomer by its centre of mass, we have calculated the angles by which *BD* subunits are turned with respect to *AC* subunits in representative SSB structures (Fig. 4). This angle ranges from 12° in *M. tuberculosis* to a quite high 95° in *Thermatoga maritima* (PDB code 1z9f; DiDonato *et al.*, 2006). The angles also roughly cluster around low and high values in accordance with the second and third groups from the Katre and Suresh analysis. Various methods of calculating this angle have been used in the past and all depend on some features of secondary

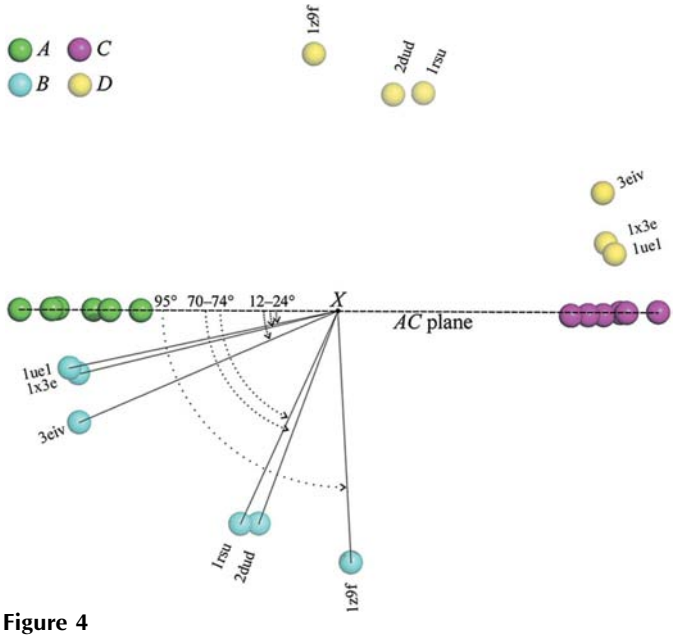

Figure 4
The positions of the centres of mass of four monomers (*A*, *B*, *C* and *D*) that constitute various SSB tetramers (denoted by PDB codes) are shown. The centres of mass of all tetramers are labelled *X*. All structures are aligned such that their respective *AXC* planes coincide and the view is down that plane. The angles represent rotations made by the *BXD* with respect to the *AXC* planes.

structure, such as the angle between the average directions of three β-strands on the interface between the *AC* and *BD* subunits. In our opinion, our way of calculating this angle as the angle formed between planes *AXC* and *BXD* (where *X* is the centre of mass of the whole tetramer) is more general and less structure-specific and additionally facilitates comparison with other protein families.

4. Summary

The first crystal structure of the single-stranded DNA-binding protein from a member of the genus *Streptomyces* has been solved and refined at 2.1 Å resolution. The mycobacterial type of quaternary structure is present in *S. coelicolor* SSB, showing a large similarity in tetramer architecture between SSBs belonging to distantly related genera of actinobacteria. A more general way of calculating the angle between subunits in tetramers is suggested.

This work was supported by the Ministry of Science, Education and Sports of the Republic of Croatia (grant Nos. 098-1191344-2943 and 098-0982913-2877). We are grateful to Professor D. Hopwood for critical reading of the manuscript.

References

Acharya, N. & Varshney, U. (2002). *J. Mol. Biol.* **318**, 1251–1264.
Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Bentley, S. D. *et al.* (2002). *Nature (London)*, **417**, 141–147.
Bernstein, D. A., Eggington, J. M., Killoran, M. P., Misic, A. M., Cox, M. M. & Keck, J. L. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 8575–8580.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chan, K. W., Lee, Y. J., Wang, C. H., Huang, H. & Sun, Y. J. (2009). *J. Mol. Biol.* **388**, 508–519.
Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Das, D., Hyun, H., Lou, Y., Yokota, H., Kim, R. & Kim, S.-H. (2007). *Proteins*, **67**, 776–782.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*, http://www.pymol.org.
DiDonato, M. *et al.* (2006). *Proteins*, **63**, 256–260.
Fedorov, R., Witte, G., Urbanke, C., Manstein, D. J. & Curth, U. (2006). *Nucleic Acids Res.* **34**, 6708–6717.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Handa, P., Acharya, N., Thanedar, S., Purnapatre, K. & Varshney, U. (2000). *Nucleic Acids Res.* **28**, 3823–3829.
Handa, P., Acharya, N. & Varshney, U. (2001). *J. Biol. Chem.* **276**, 16992–16997.
Hopwood, D. A. (2006). *Annu. Rev. Genet.* **40**, 1–23.
Howard, A. J. (2000). In *Crystallographic Computing 7*, edited by P. E. Bourne & K. D. Watenpaugh. Oxford University Press.
Jędrzejczak, R., Dauter, M., Dauter, Z., Olszewski, M., Długołęcka, A. & Kur, J. (2006). *Acta Cryst.* D**62**, 1407–1412.
Katre, U. V. & Suresh, C. G. (2009). *Acta Cryst.* D**65**, 1–10.
Kerr, I. D., Wadsworth, R. I. M., Cubeddu, L., Blankenfeldt, W., Naismith, J. H. & White, M. F. (2003). *EMBO J.* **22**, 2561–2570.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Mijaković, I., Petranović, D., Maček, B., Čepo, T., Mann, M., Davies, J., Jensen, P. R. & Vujaklija, D. (2006). *Nucleic Acids Res.* **34**, 1588–1596.
Murzin, A. G. (1993). *EMBO J.* **12**, 861–867.
Mushegian, A. R. & Koonin, E. V. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 10268–10273.
Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I. & Sussman, J. L. (2005). *Bioinformatics*, **21**, 3435–3438.
Purnapatre, K. & Varshney, U. (1999). *Eur. J. Biochem.* **264**, 591–598.
Raghunathan, S., Kozlov, A. G., Lohman, T. M. & Waksman, G. (2000). *Nature Struct. Biol.* **7**, 648–652.
Raghunathan, S., Ricard, C. S., Lohman, T. M. & Waksman, G. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 6652–6657.
Reddy, M. S., Guhan, N. & Muniyappa, K. (2001). *J. Biol. Chem.* **276**, 45959–45968.
Saikrishnan, K., Jeyakanthan, J., Venkatesh, J., Acharya, N., Sekar, K., Varshney, U. & Vijayan, M. (2003). *J. Mol. Biol.* **331**, 385–393.
Saikrishnan, K., Manjunath, G. P., Singh, P., Jeyakanthan, J., Dauter, Z., Sekar, K., Muniyappa, K. & Vijayan, M. (2005). *Acta Cryst.* D**61**, 1140–1148.
Savvides, S. N., Raghunathan, S., Fuetterer, K., Kozlov, A. G., Lohman, T. M. & Waksman, G. (2004). *Protein Sci.* **13**, 1942–1947.
Sikder, D., Unniraman, S., Bhaduri, T. & Nagaraja, V. (2001). *J. Mol. Biol.* **306**, 669–679.
Štefanić, Z., Vujaklija, D., Andrišić, L., Mikleušević, G., Andrejašič, M., Turk, D. & Luić, M. (2007). *Croat. Chem. Acta*, **80**, 35–39.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Yang, C., Curth, U., Urbanke, C. & Kang, C. (1997). *Nature Struct. Biol.* **4**, 153–157.