

Structural variability of *Mycobacterium tuberculosis* SSB and susceptibility to inhibition

Srikalaivani Raja^{1,†}, Anju Paul^{1,†}, Sriram Raghavan^{1,4}, Sibi Narayanan^{1,5}, Somnath Shee^{2,3}, Amit Singh^{2,3}, Umesh Varshney², Balasubramanian Gopal¹ and Mamannamana Vijayan^{1,*}

¹Molecular Biophysics Unit, Indian Institute of Science, Bengaluru 560 012, India

²Department of Microbiology and Cell Biology, Indian Institute of Science, Bengaluru 560 012, India

³Centre for Infectious Disease Research, Indian Institute of Science, Bengaluru 560 012, India

⁴Present address: RIKEN Center for Computational Science, Japan

⁵Present address: Sitaram Ayurveda Private Limited, Thrissur 680 007, India

Single-stranded DNA is formed at various stages of DNA metabolism. It is protected from degradation by single-stranded DNA-binding proteins (SSBs). Structural variability has been observed in the quaternary arrangement of tetrameric SSBs from mycobacteria and other sources. Here we describe two novel crystal forms which illustrate the extent of structural variability. Docking studies carried out with inhibitors identified from DNA-binding assays allowed the characterization of eight distinct potential binding regions or grooves on each tetramer that circumvent structurally variable regions. Compounds known to inhibit certain bacterial SSBs were tested against *Mycobacterium tuberculosis* SSB (*MtSSB*) using DNA-binding and cellular assays. We report two compounds that inhibit *MtSSB* and growth of the bacterium. Together, this structural analysis reveals a strategy to exploit the variability of *MtSSB* for the design of inhibitors to this protein. The variability in structure of *MtSSB* could contribute to its susceptibility to inhibition.

Keywords: Binding regions, crystal structure, docking, inhibitor development, *Mycobacterium tuberculosis*, structural plasticity.

MAJOR processes of DNA metabolism such as replication, recombination and repair involve unwinding of the double-stranded DNA (dsDNA) to form transient, single-stranded DNA (ssDNA). In the cell, ssDNA is highly susceptible to attack by nucleases and other chemically reactive groups. They are also prone to form secondary structures. Single-stranded DNA-binding proteins (SSBs) are designed to protect vulnerable ssDNA from chemical attacks and aberrant secondary structure formation. SSBs are ubiquitous proteins found in viruses, archaea, bacteria

and eukaryotes^{1,2}. Though the overall structure and function of SSBs across all lifeforms are conserved, they share very low sequence similarity. They bind to DNA with high affinity and in a sequence-independent manner. In addition to maintaining the chemical integrity of DNA, SSBs are also involved in binding and controlling the function of other proteins involved in various stages of DNA metabolism. Most bacterial SSBs are homo-tetrameric, with each protomer containing an N- and a C-domain. The N-domain folds into a highly conserved oligonucleotide-binding (OB) fold responsible for coating by ssDNA, whereas the C-terminal domain interacts with nearly a dozen other proteins that together constitute the SSB interactome^{3,4}.

The structure of the 164-residue long SSB from *Mycobacterium tuberculosis* (*MtSSB*), determined as part of a concerted national and international effort⁵⁻⁸, showed the same tertiary structure as that seen in previously characterized SSBs from *Escherichia coli* (*EcSSB*)⁹ and human mitochondria (*HMtSSB*)¹⁰. The *MtSSB* tetramer consists of an N-terminal DNA-binding domain and a C-terminal disordered tail. The N-terminal domain consists of three long β -hairpin loops extending out of a globular core formed by a five-stranded β -barrel capped by an α -helix, characteristic of the OB fold¹¹. As shown in Figure 1, *MtSSB* tetramer is a dimer of dimers with molecular dyads along *P*, *Q* and *R*. While the quaternary structures of *EcSSB* and *HMtSSB* are comparable, the quaternary arrangement of *MtSSB* is different from the other homologues. One important feature that sets it apart from other well-characterized SSBs is the presence of a hook-like structure at the end of the N-domain. This peptide stretch is largely responsible for the unique quaternary structure adopted by the mycobacterial SSBs. In addition to the canonical SSB (SSBa), a paralogous protein (SSBb) has been found in several bacteria including *M. tuberculosis*. Structures of both paralogues of SSB from *M. smegmatis* and that of SSBa from *M. leprae* (*MISSB*) have also been

*For correspondence. (e-mail: mv@iisc.ac.in)

†Contributed equally to the work.

determined in the laboratory^{12–14}. The quaternary structural variability in SSBs brought forth by these studies has been reviewed earlier¹⁵.

Given their vital role in DNA metabolism, SSBs make a promising drug target. The surface contour of SSBs that contains several cavities and grooves also makes it suitable to be targeted by small-molecule inhibitors. Broadly speaking, two approaches have been adopted towards inhibitor development against SSBs. While one approach aims to disrupt protein–protein interactions at the C-terminal tip using peptide-like molecules, the other is aimed at inhibiting ssDNA binding to the N-terminal domain using small molecules^{16,17}. Recently, SSBs purified from both Gram-positive and Gram-negative bacteria were assessed using high-throughput screening for inhibition of DNA binding by small-molecule inhibitors¹⁷. A few compounds identified to inhibit SSBs were further validated by EMSA in this study¹⁷. Huang *et al.*¹⁸ reported the crystal structure of SSB from *Pseudomonas aeruginosa* with a flavonol, myricetin. It was desirable to explore the effect of the reported compounds on *Mt*SSB. Here we describe results from biochemical studies and preliminary *in situ* studies carried out on *Mt*SSB with these compounds. Furthermore, docking studies were carried out with compounds identified from these functional studies to characterize their potential binding sites on the tetramer. We also describe novel crystal forms of *Mt*SSB identified during these efforts, further highlighting the plasticity of the protein and susceptibility to inhibition owing to this structural variability.

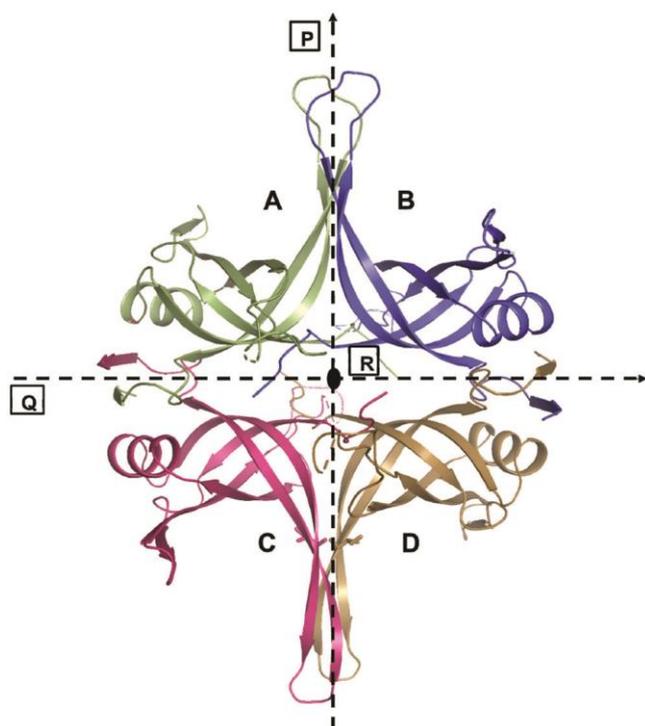


Figure 1. *Mycobacterium tuberculosis* SSB (*Mt*SSB) tetramer with its subunits and molecular symmetry axes *P*, *Q* and *R* marked.

Materials and methods

Bacterial strains and culture conditions

M. tuberculosis strain H37Rv was cultured in 7H9 broth supplemented with 0.2% glycerol, 0.1% Tween-80 and ADS (0.5% albumin, 0.2% dextrose and 0.085% NaCl) with shaking at 180 rpm in a rotary shaker incubator (Lab Therm LT-X, Kuhner, Switzerland) at 37°C.

Chemicals

Ten compounds identified as SSB inhibitors by Glanzer *et al.*¹⁷ were supplied by the National Cancer Institute, National Institutes of Health (NIH), USA. Myricetin, luteolin and quercetin were purchased from Sigma-Aldrich. Buffers and salts routinely used for protein expression and purification were also obtained from Sigma-Aldrich (USA). Other analytical-grade chemicals were procured from local vendors.

Protein expression and purification

*Mt*SSB was purified following the standard protocol for Ni-NTA affinity chromatography using *E. coli* BL21 (λ DE3) as host cells. In brief, competent cells were transformed with the recombinant clone and transformed colonies were selected on an LB agar plate containing 30 mg/ml kanamycin. The transformed cells were picked and primary and secondary cultures were grown at 37°C using LB medium. The culture was induced at A_{600} of 0.6 with 0.5 mM Isopropyl β -D-1-thiogalactopyranoside and was further incubated for 6 h at 37°C. Cells were harvested by centrifugation at 6000 rpm and resuspended in buffer I (30 mM *Tris*-HCl (pH 7.6), 1 M NaCl, 10% glycerol and 10 mM imidazole). Resuspended cells were lysed by sonication and cell debris was removed by centrifugation at 14,000 rpm. The supernatant was then allowed to bind to pre-equilibrated Ni-NTA resin followed by washing with buffer II (20 mM *Tris*-HCl (pH 7.6), 1 M NaCl, 10% glycerol and 30 mM imidazole). The protein fractions were eluted with a gradient (50–300 mM) of imidazole. The eluted fractions were pooled and further purified by size exclusion chromatography using a pre-equilibrated Superdex 200, 10/300 column in buffer III (20 mM *Tris*-HCl (pH 7.6), 300 mM NaCl and 5% glycerol). Purity of the protein was confirmed by SDS-PAGE.

Electrophoretic mobility shift assay

DNA binding affinity of purified *Mt*SSB was measured using electrophoretic mobility shift assay (EMSA) with two different lengths of 5'-fluorescent-tagged poly-dT 25mer and 35mer ssDNA. The 10 μ l reaction mixture

Table 1. Data collection and refinement statistics

	Form III (PDB ID: 7F5Y)	Form IV (PDB ID: 7F5Z)
Space group	P3 ₂ 12	P6 ₁ 22
Unit cell dimensions ($a = b, c$; Å)	59.55, 124.98	110.04, 104.12
Resolution (Å)	1.92	3.00
No. of subunits/asymmetric unit	2	2
Unique reflections	18,944 (2650)	7920 (1127)
Multiplicity	5.3 (5.3)	9.8 (9.9)
Completeness (%)	96.6 (94.1)	100.0 (100.0)
$\langle I/\sigma(I) \rangle$	10.9 (2.2)	8.7 (2.4)
R_{sym} (%) [†]	8.8 (74.4)	18.3 (88.3)
Refinement and model statistics		
R_{work} (%)	19.12	24.84
R_{free} (%)	22.86	28.13
RMS deviation from ideal values		
Bond length (Å)	0.012	0.008
Bond angle (°)	1.4	1.3
Ramachandran map statistics		
Favoured	93.4	85.5
Additionally allowed	3.6	11.0
Generously allowed	2.0	2.5
Disallowed	1.0	1.0

consisted of 2 μM ssDNA and increasing concentrations of *MtSSB* (5–50 μM) in 20 mM *Tris*-HCl and 50 mM NaCl at pH 7.6. The reaction mixture was incubated for 10 min at room temperature and for an additional 5 min on ice before electrophoresis at 4°C. To test the affinity of inhibitors to *MtSSB*, 10 μl reaction mixtures containing 20 μM *MtSSB*, 2 μM poly-dT 35mer ssDNA and inhibitor concentrations ranging from 50 μM to 3.5 mM were used. The samples were loaded onto an 8% native PAGE in TBE buffer and visualized under UV fluorescence in filter in a BioRad gel documentation system. Chemical structures of inhibitors tested are provided in the [Supplementary Table 1](#).

Resazurin microtitre assay

Minimum inhibitory concentration (MIC) of compounds against *M. tuberculosis* H37Rv was quantified using Resazurin microtitre assay (REMA). Resazurin is a redox-sensitive dye routinely used as an indicator of cell viability. Briefly, 100 μl 7H9-ADS was dispensed in each well of a sterile, flat-bottom 96-well plate and serial two-fold dilutions of compounds were prepared directly in the plate. The log-phase bacteria (approximately 1×10^5 bacterial cells/well, 100 μl) were added in triplicate. Mixtures containing standard tuberculosis (TB) drugs, moxifloxacin and rifampicin, were used as positive controls in the assay. Wells containing no *M. tuberculosis* cells were used for autofluorescence control and those containing only cells formed an additional control. Plates were incubated for five days at 37°C before 30 μl resazurin (0.02% w/v stock solution) was added. The plates were further incu-

bated for 48 h to detect colour change (blue to pink). Fluorescence intensity was measured in a SpectraMax M3 plate reader (Molecular Device) in bottom-reading mode with excitation at 530 nm and emission at 590 nm. The intensity of the pink colour directly correlates to the extent of bacterial growth. Per cent inhibition was calculated based on the relative fluorescence units to growth (only cells) control. MIC was defined as the lowest concentration of the molecule that prevented 90% reduction in fluorescence compared to an untreated growth control.

Crystallization and structure determination

Various crystallization conditions were screened using commercially available screens from Hampton employing the microbatch under-oil method. The crystallization condition (1 μl) was mixed with an equal quantity of purified *MtSSB* in each drop. Diffraction quality crystals were obtained in two conditions, namely Salt Rx-31 (3.5 M sodium formate and 0.1 M Bis-*Tris* propane, pH 7.0) and Crystal Screen II-4 (35% v/v 1,4-dioxane). These were diffracted and data were collected on an image plate (Rigaku *R*-axis IV⁺⁺) mounted on a Cu-K α rotating-anode X-ray generator (RIGAKU ULTRA18) up to 1.89 Å and another image plate (MAR 345) mounted on a CuK α rotating-anode X-ray generator (Bruker MICROSTAR ULTRA II) up to 3.00 Å respectively. Diffraction data were processed using iMosflm¹⁹ and scaled using Scala²⁰ of the CCP4 suite. Matthews²¹ coefficient indicated the likely presence of two molecules, accounting for a dimer, in the asymmetric unit in both crystal forms. The resultant model from Phaser was subsequently subjected to

rigid-body refinement followed by B-factor and positional refinement using Refmac5 of the CCP4 suite along with manual model-building using COOT^{22–24}. Water molecules were added at the later stages of refinement to peaks corresponding to 1σ and 3σ in the $2F_o - F_c$ and $F_o - F_c$ maps respectively. Table 1 presents the data collection and refinement statistics.

Docking

System preparation, modelling and initial docking calculations were performed employing the Maestro interface of Schrödinger Suite molecular modelling package (version 2015–4), using default parameters. Four *MtSSB* models were used to carry out docking to ensure that the variability seen in the crystal structures is sampled and accounted for in inhibitor binding. The protein tetramer models were prepared using the protein preparation wizard²⁵. Here, force-field atom types and bond orders were assigned, missing atoms were added, tautomer/ionization states were assigned, water orientations were sampled and residues Asn, Gln, and His had their tautomers adjusted to optimize the hydrogen-bond network. A constrained energy minimization was then performed. SiteMap was employed to identify potential binding sites on the protein surface of various *MtSSB* models²⁶. Results from the various models were compared and binding sites were identified for docking. Glide XP (extra precision) was employed for docking of compounds at these sites using the prepared structures as the receptor²⁷. Epik, included in the same software package, was used to prepare the ligands²⁸. Docking was performed using default XP settings with flexible ligand sampling and post-docking minimization.

Analysis of structures

Structural alignments were carried out using ALIGN²⁹. NACCESS was employed to compute buried surface areas³⁰. Interatomic and hydrogen-bond distances were calculated employing CONTACTS from CCP4 (ref. 20). Chimera and PyMOL were used for viewing and analysing structures, and for figure generation^{31,32}. Multiple sequence alignment was performed using the Clustal Omega server and the figure was generated using ESPript^{33,34}. CABS flex 2.0 server was used to generate an ensemble of structures to model structure flexibility in different crystal forms³⁵.

Results

Determination of IC_{50} value of inhibitors

DNA binding was detected at 20 μM *MtSSB* concentration using EMSA (Figure 2a). Further inhibitor-binding

studies were carried out at this protein concentration. Thirteen compounds were tested to check their inhibitory action against *MtSSB* (Supplementary Table 1). Of these, 10 were reported to variably inhibit SSBs from different (both Gram-positive and Gram-negative) sources, while myricetin was shown to bind and inhibit *PaSSB*^{17,18}. Myricetin analogues, namely luteolin and quercetin, were also tested for inhibition of *MtSSB*. EMSA was used as the assay of choice in the present study, given that a small set of compounds was being screened. Among the tested compounds, myricetin and purpurogallin inhibited DNA binding to *MtSSB* (Figure 2b and c; Supplementary Figure 1). Inhibitory concentration (IC_{50}) was calculated based on the band intensities from the gel. The IC_{50} values of myricetin and purpurogallin were 596 and 238 μM respectively.

Effect of inhibitors on *M. tuberculosis* cell growth

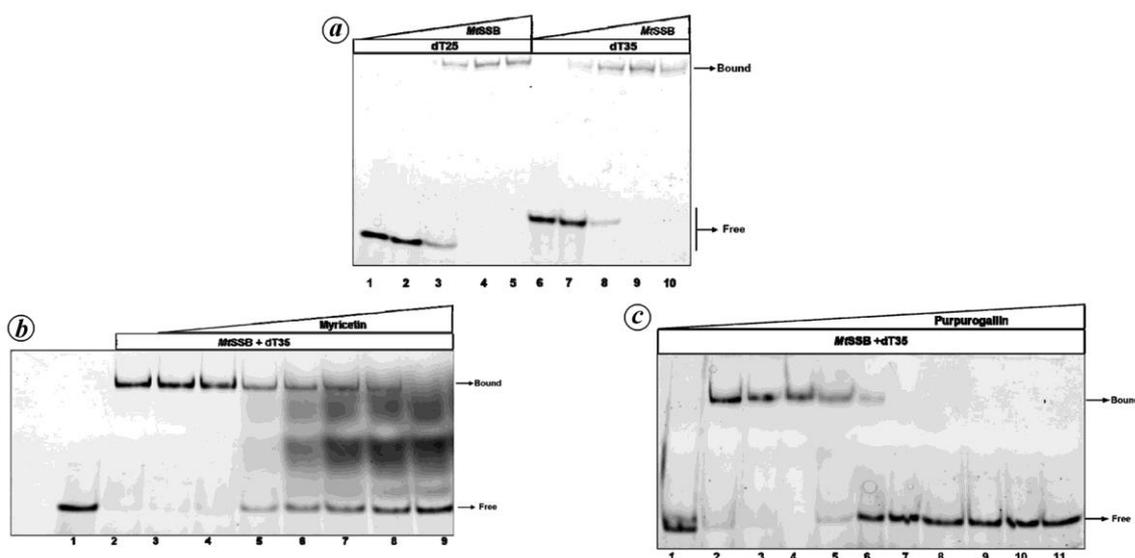
Having tested the inhibitory activity of the selected compounds on *MtSSB*, we evaluated their effect on *M. tuberculosis* cell growth. To this end, known TB drugs were tested along with *MtSSB* inhibitors identified by EMSA using the REMA assay. Both myricetin and purpurogallin were able to inhibit *M. tuberculosis* growth with MIC values in the range 0.125–0.25 mM and 0.06–0.125 mM respectively. The inhibition by these two compounds was significant, although much lower than that by moxifloxacin and rifampicin (MIC values in the sub- μM range).

Crystal structures

The structure of *MtSSB* has been determined in the laboratory in two crystal forms³⁶. All the subunits in the two forms exhibited the globular OB-fold for the ordered N-terminal DNA-binding domain. Three flexible β -hairpin loops, often containing disordered residues, extended from the globular component (Figure 3). All the subunits were about 120 residues long. Therefore, the DNA-binding domain may be considered to consist of residues 1–120. All the remaining residues, which make up the C-terminal domain, are disordered. The last few residues in the DNA-binding domain, ordered only in SSBs of mycobacterial and closely related species, form a hook which plays an important role in the quaternary association. Although all the subunits have similar structure, the quaternary structures of the molecules in the two forms exhibit significant differences (Table 2). The molecule is located on a crystallographic twofold axis which coincides with the molecular dyad P in form I. Subunits A and C constitute the Asymmetric Unit (ASU). Form II has subunits A and B from two independent tetramers in the ASU, and the tetramers were built around the crystallographic axes coinciding with axis Q in one and R in the other. The molecule thus appeared to exhibit some variability expressed in terms of differences in quaternary association. This

Table 2. Details of differences among the various crystal forms of *Mycobacterium tuberculosis* SSB (*MtSSB*)

Form	Details of molecular symmetry	Solvent content (%)	Subunits comprising ASU
Form I – 1UE1	Space group – $P3_12_1$ Crystallographic dyad – P	37.6	A, C
Form II – 1UE6 tetramer 1	Space group – $I2_12_12_1$ Crystallographic dyad – R	44.7	A, B
Form II – 1UE6 tetramer 2	Space group – $I2_12_12_1$ Crystallographic dyad – Q		A, B
Form III	Space group – $P3_21_2$ Crystallographic dyad – R	32.7	A, C
Form IV	Space group – $P6_12_2$ Crystallographic dyad – P	52.5	A, C

**Figure 2.** Electrophoretic mobility shift assays of *MtSSB*. (a) Analysis of dT25 and dT35 binding to 0, 5, 10, 20 and 50 μM of *MtSSB* in lanes 1–5 and lanes 6–10 respectively. Analysis of release of free DNA from *MtSSB* + dT35 complex using (b) 0, 50, 100, 250, 500, 750, 1000 and 1500 μM of myricetin in lanes 2–9 and (c) 0, 50, 100, 250, 500, 750 and 1000 μM of purpurogallin in lanes 1–7 respectively. All gels are appropriately cropped for clarity.

variability possibly could have implications for the development of inhibitors. Therefore, an effort was made to further explore the variability by growing more crystal forms of the molecule.

The above-mentioned effort resulted in two novel forms (III and IV). In both these forms, subunits A and C are in the ASU and the tetramer is built by applying crystallographic symmetry coinciding with molecular axes *R* and *P* in forms III and IV respectively (Figure 4). In form III, density corresponding to residues 39–47, 89–91 and 120 is missing in one subunit, while all the residues up to and including 120 are defined in the other subunit. In form IV, density for residues 40, 41 and 120 is missing in one subunit while that for residues 41 and 42 is missing in the other. The subunit structure in all four forms of *MtSSB* is similar. The globular core of the DNA-binding N-terminal domain is made up of a six-stranded β -barrel capped by an α -helix. Pairwise superposition of the cores

in all the independent subunits in the four crystal forms yields root mean square deviation (RMSD) in the $C\alpha$ position ranging from 0.23 to 0.49 \AA . Three loops, L_{12} , L_{23} and L_{45} , extend out of the globular core. The hook region is followed by a disordered C-terminal domain. None of the rotations about the non-crystallographic twofold axes is less than 178° . Therefore, the molecules in all the forms are nearly 222 symmetric. The variability is primarily in the composition of the ASUs. The flexibility of the loops is also likely to contribute to the variability.

Insights from docking studies

Repeated efforts to prepare crystals of inhibitor complexes of the protein were unsuccessful. However, the various crystal forms of *MtSSB* encompassing different loop conformations and quaternary structures provided a basis to

explore inhibitor binding computationally. Four models were chosen to carry out docking studies, each with a distinct quaternary structure (Table 2). Thus, one tetramer from form I, two distinct tetramers from form II and one tetramer from form III were chosen. The tetramer resulting from form IV is essentially the same as that from form I and hence not used for these studies. Each model was prepared as described earlier in the text, to ensure uniformity across models. Each tetramer was used as a single entity. To avoid confusion, residues were numbered as 1, 2, 3, ..., 201, 202, 203, ..., 401, 402, 403, ... and 601, 602, 603, ... in subunits A, B, C and D respectively.

As described earlier, the models were first explored through SiteMap to identify sites with favourable geometric and electrostatic features. Results obtained from the various models were compared to identify potential binding sites on the tetramer. The SiteMaps obtained for the four models are depicted in the [Supplementary Figure 2](#).

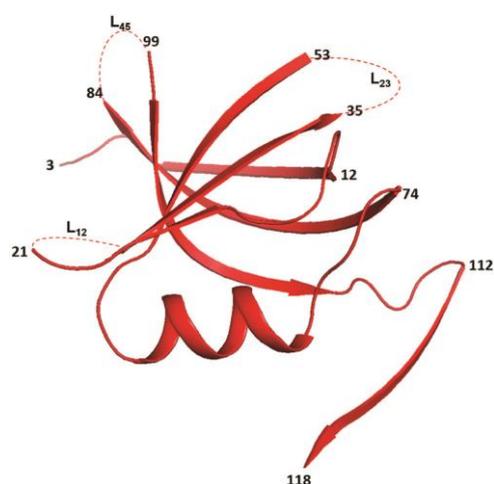


Figure 3. Globular core structure of the *MtSSB* subunit. Locations of the three flexible loops are indicated as dashed curves.

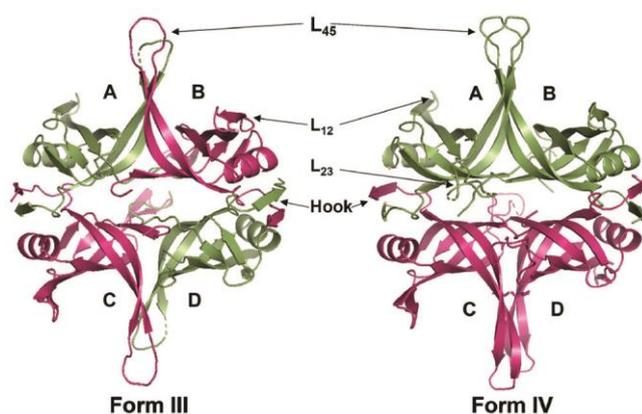


Figure 4. Representation of form III and form IV of *MtSSB* crystal structure and resultant tetramers. Subunits A and C constitute the asymmetric unit in both forms. Subunits A and D (likewise, B and C) are related by crystallographic dyad along R in form III, while subunits A and B (likewise, C and D) are related by dyad along P. Subunits are coloured to indicate their molecular symmetry. The three major loops protruding from the globular core as well as the hook region are marked.

It is noteworthy that even slight differences in molecular structure led to notable differences in the SiteMap results. This is probably indicative of how broad the binding sites on the SSB tetramer can be.

The identified sites were roughly related by molecular symmetry. These roughly symmetry-related sites were grouped as sub-sites. Thus, two sites (1 and 2) having four sub-sites (a to d) each were identified, resulting in eight sites per SSB tetramer. Both sites have an average SiteScore over 0.8, an indicator that has been suggested to accurately distinguish between drug-binding and non-drug-binding sites²⁶. A third site-type was also found in two of the models. However, this site is ill-defined comprising residues bridging sites 1 and 2 and was therefore excluded from further analysis. Site 1 is located close to the base of loop L₄₅ in each subunit, while site 2 is near the hook-like peptide stretches of two neighbouring subunits (for example, subunits A and C). Figure 5 is a schematic illustration of these sites. Ligands identified from functional studies (myricetin and purpurogallin) were prepared before being docked into the models, as discussed earlier in the text. Docking was carried out in the Glide XP mode using the eight identified sites as receptors for both molecules in a 20 Å × 20 Å grid box. At site 1, the ligands were found to bind near the groove of loop L₄₅ with a majority of the interactions with residues from one subunit (for example, subunit A) and a few interactions with residues from the neighbouring subunit (subunit B). At site 2, the ligand was found straddled between two subunits, near the clamp-like structure formed by the hook-like peptides. The resulting top two ligand poses at each site were used for analysis and allowed the delineation of residues involved in inhibitor binding. Residues in contact with the docked ligand were identified by examining all atoms within 4 Å radius and

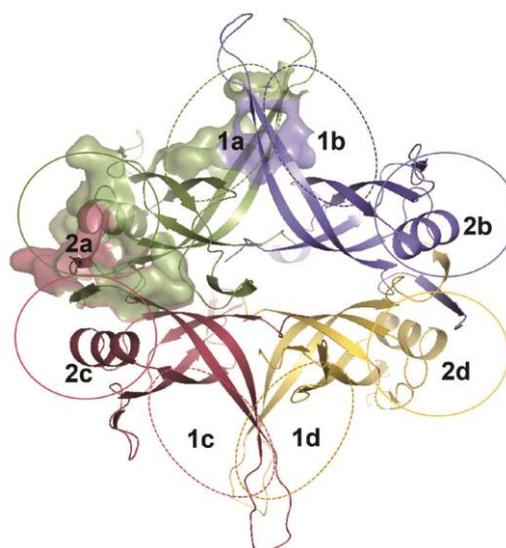


Figure 5. Schematic representation of the eight docking sites identified in *MtSSB*. Sites 2a and 2c (likewise, 2b and 2d) lie on opposite faces of the tetramer.

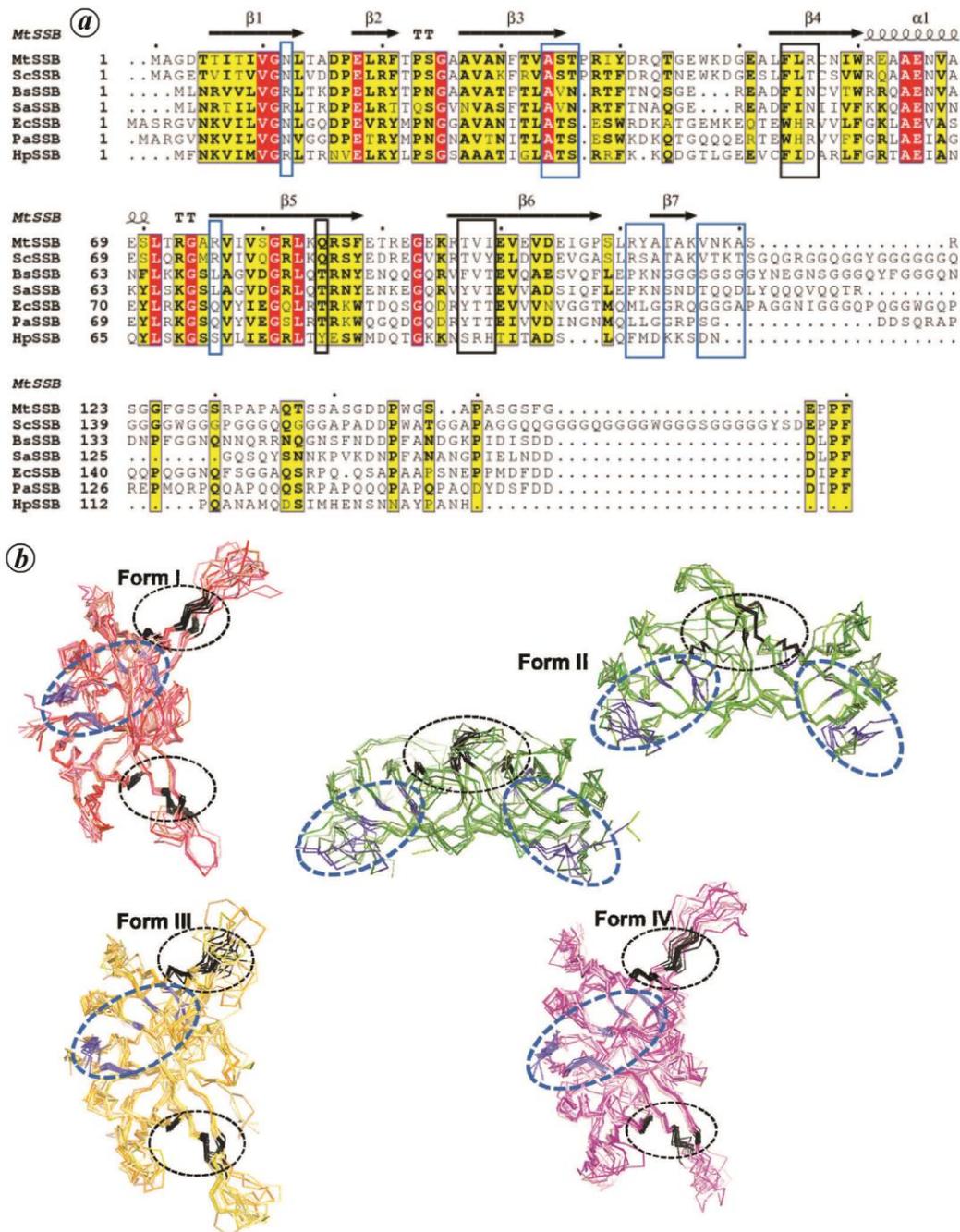


Figure 6. Conserved sequence and structural features of the inhibitor-binding regions. *a*, Sequence alignment of a few bacterial SSBs using ClustalO and ESPrpt. The regions comprising sites 1 and 2 identified by docking are marked in black and blue boxes respectively. *b*, The corresponding regions in all four crystal forms of *MtSSB* marked as dashed circles.

further confirmed by calculating the change in buried surface area of the residues with and without the docked ligand.

Comparison of all equivalent sites in and among the various models showed that the residues involved in ligand binding were scattered and that the binding sites were better described as binding regions or binding grooves. The interactions were mostly polar in nature, probably owing to the many hydroxyl substituents on the

inhibitors and the electrostatic nature of *MtSSB*. Interestingly, despite the molecular symmetry of the tetrameric SSB, it was found that neither the SiteMap results nor the docking results followed symmetry strictly. In spite of the scatter, it was fairly straightforward to arrive at a consensus of residues that are involved in ligand binding at these two binding regions (Table 3). A comparison of various bacterial SSB sequences demonstrates that many of these consensus residues comprising the binding regions

Table 3. Consensus of residues involved in ligand binding at various docking sites

Site	A	B	C	D
1A	F54, L55, R56, V98, I99	Q285, T297, V298, I299		
1B	Q285, T297, V298, I299	F254, L255, R256, T297, V298, I299		
1C			F454, L455, R456, T497, V498, I499	T697, V698, I699
1D			T497, V498, I499	F654, L655, R656, T697, V698, I699
2A	N12, A34, S35, T36, R76, R111, Y112, A113		V517, N518, K519, A120	
2B		N212, A234, S235, T236, R276, R311, Y312, A313	V717, N718, K719, A720	
2C	V117, N118, K119, A120		N412, A434, S435, T436, R476, R511, Y512, A513	
2D		V317, N318, K319, A320		N612, A634, S635, T636, R676, R711, Y712, A713

are reasonably well conserved (Figure 6a). Mapping the location of these residues on an ensemble of structures corresponding to each crystal form also indicates that they occur in the relatively less dynamic stretches of the protein (Figure 6b). Phe54 at site 1 is homologous to Trp54 of *EcSSB* which has been suggested to be crucial for DNA-binding. Studies on DNA-bound *Bacillus subtilis* SSB (*BsSSB*) and *EcSSB* structures also led to the identification of a potential ‘bridge’ interface between neighbouring SSB tetramers³⁷. This bridge interface is formed by loop L₁₂. Site 1 in fact lies between L₁₂ and the base of L₄₅. Residues in and around site 1 form a part of the bridge interface as well. Given its role in DNA binding and relatively higher SiteScore suggested by SiteMap, it would seem that, in *MtSSB*, residues at site 1 may be responsible for the inhibitor-binding affinity, while site 2 with its distinct hook may provide specificity.

Summary and conclusion

The present study builds upon the detailed structural analyses we have carried out on different mycobacterial SSBs. Several compounds which inhibit SSBs from other bacterial species have been reported in the literature. Biochemical studies showed that two of them, namely, myricetin and purpurogallin, inhibit *MtSSB* with IC₅₀ values of 596 and 238 μM respectively. They also inhibit the growth of *M. tuberculosis* with MIC values in the range 0.125–0.25 and 0.06–0.125 mM respectively. This finding is consistent with observations from dynamic light-scattering experiments suggesting that these inhibitors modulate the quaternary structure of *MtSSB*. The availability of the structures of three independent copies of the *MtSSB* molecule distributed in two crystal forms made the study of the binding of myricetin and purpurogallin to the protein molecule feasible. Each subunit of the tetrameric protein molecule is made up of a globular core, with three flexible hairpin loops extending from it. The flexibility of the loops and subtle differences in quaternary association endow the molecule with some structural variability. The variability was further elaborated by

determining the structures of two additional crystal forms of the protein reported here.

Efforts to prepare crystals of the complexes of the protein with the compounds were not successful. However, protein models obtained from the four crystal forms provided a structural scaffold for computationally exploring the binding of the compounds to the protein. Two binding sites and their symmetry equivalents could be identified. The sites were broad and exhibited some differences among the different models. Each of the two compounds, when docked into the sites, exhibited some scatter in the location among the symmetry equivalents as well as the models. However, it was not difficult to identify consensus residues involved in binding. Close examination of the sites indicated that one of them is in a region involved in DNA binding. The other straddles two hooks from neighbouring subunits. It may be recalled that these hooks are characteristic of SSBs from mycobacterial and related species. It is possible that contribution to affinity primarily comes from the first site, while the second site is primarily responsible for specificity. The breadth and variability of the sites and the scatter in the locations of the ligands are such that the binding locations are more appropriately described as binding regions or grooves than as binding sites. This is not surprising as SSBs bind DNA non-specifically. This observation is also in consonance with the known diffusional migration of SSBs along the DNA³⁸. However, the conservation of residues in the inhibitor-binding regions across species gives confidence in this approach to further develop inhibitors against *MtSSBs* in a targeted manner.

1. Chase, J. W. and Williams, K. R., Single-stranded DNA binding proteins required for DNA replication. *Annu. Rev. Biochem.*, 1986, **55**, 103–136.
2. Lohman, T. M. and Ferrari, M. E., *Escherichia coli* single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu. Rev. Biochem.*, 1994, **63**, 527–570.
3. Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. and Keck, J. L., SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.*, 2008, **43**(5), 289–318.

4. Costes, A., Lecointe, F., McGovern, S., Quevillon-Cheruel, S. and Polard, P., The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks. *PLoS Genet.*, 2010, **6**(12), e1001238.
5. Terwilliger, T. C. *et al.*, The TB structural genomics consortium: a resource for *Mycobacterium tuberculosis* biology. *Tuberculosis*, 2003, **83**(4), 223–249.
6. Vijayan, M., Structural biology of mycobacterial proteins: the Bangalore effort. *Tuberculosis*, 2005, **85**(5), 357–366.
7. Arora, A. *et al.*, Structural biology of *Mycobacterium tuberculosis* proteins: the Indian efforts. *Tuberculosis*, 2011, **91**(5), 456–468.
8. Arcus, V. L., Lott, J. S., Johnston, J. M. and Baker, E. N., The potential impact of structural genomics on tuberculosis drug discovery. *Drug Discov. Today*, 2006, **11**(1–2), 28–34.
9. Raghunathan, S., Ricard, C. S., Lohman, T. M. and Waksman, G., 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*, 1997, **94**(13), 6652–6657.
10. Yang, C., Curth, U., Urbanke, C. and Kang, C., Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nature Struct. Biol.*, 1997, **4**(2), 153–157.
11. Saikrishnan, K. *et al.*, Crystallization and preliminary X-ray studies of the single-stranded DNA-binding protein from *Mycobacterium tuberculosis*. *Acta Crystallogr., Sect. D*, 2002, **58**(2), 327–329.
12. Saikrishnan, K. *et al.*, Structure of *Mycobacterium smegmatis* single-stranded DNA-binding protein and a comparative study involving homologous SSBs: biological implications of structural plasticity and variability in quaternary association. *Acta Crystallogr. Sect. D*, 2005, **61**(8), 1140–1148.
13. Kaushal, P. S., Singh, P., Sharma, A., Muniyappa, K. and Vijayan, M., X-ray and molecular-dynamics studies on *Mycobacterium leprae* single-stranded DNA-binding protein and comparison with other eubacterial SSB structures. *Acta Crystallogr., Sect. D*, 2010, **66**(10), 1048–1058.
14. Singh, A., Vijayan, M. and Varshney, U., Distinct properties of a hypoxia specific paralog of single stranded DNA binding (SSB) protein in mycobacteria. *Tuberculosis*, 2018, **108**, 16–25.
15. Arif, S. M. and Vijayan, M., Structural diversity based on variability in quaternary association. A case study involving eubacterial and related SSBs. *Methods Mol. Biol.*, 2012, **922**, 23–35.
16. Voter, A. F., Killoran, M. P., Ananiev, G. E., Wildman, S. A., Hoffmann, F. M. and Keck, J. L., A high-throughput screening strategy to identify inhibitors of SSB protein–protein interactions in an academic screening facility. *SLAS Discov.*, 2018, **23**(1), 94–101.
17. Glanzer, J. G., Endres, J. L., Byrne, B. M., Liu, S., Bayles, K. W. and Oakley, G. G., Identification of inhibitors for single-stranded DNA-binding proteins in eubacteria. *J. Antimicrob. Chemother.*, 2016, **71**(12), 3432–3440.
18. Huang, C. Y., Crystal structure of SSB complexed with inhibitor myricetin. *Biochem. Biophys. Res. Commun.*, 2018, **504**(4), 704–708.
19. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. and Leslie, A. G. W., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D*, 2011, **67**(4), 271–281.
20. Winn, M. D. *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D*, 2011, **67**(4), 235–242.
21. Matthews, B. W., Solvent content of protein crystals. *J. Mol. Biol.*, 1968, **33**(2), 491–497.
22. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J., Phaser crystallographic software. *J. Appl. Crystallogr.*, 2007, **40**(4), 658–674.
23. Murshudov, G. N. *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D*, 2011, **67**(4), 355–367.
24. Emsley, P. and Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D*, 2004, **60**(12 and 1), 2126–2132.
25. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. and Sherman, W., Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.*, 2013, **27**(3), 221–234.
26. Halgren, T. A., Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.*, 2009, **49**(2), 377–389.
27. Friesner, R. A. *et al.*, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *J. Med. Chem.*, 2006, **49**(21), 6177–6196.
28. Greenwood, J. R., Calkins, D., Sullivan, A. P. and Shelley, J. C., Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided Mol. Des.*, 2010, **24**(6–7), 591–604.
29. Cohen, G., ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallogr.*, 1997, **30**(6), 1160–1161.
30. Hubbard, S. J. and Thornton, J. M., ‘NACCESS’, computer program. Department of Biochemistry and Molecular Biology, University College, London, UK, 1993.
31. Pettersen, E. F. *et al.*, UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 2004, **25**(13), 1605–1612.
32. DeLano, W. L., *PyMOL*, 2002, <http://www.pymol.org>.
33. Madeira, F. *et al.*, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.*, 2019, **47**(W1), W636–W641.
34. Robert, X. and Gouet, P., Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, 2014, **42**(W1), W320–W324.
35. Kuriata, A. *et al.*, CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res.*, 2018, **46**(W1), W338–W343.
36. Saikrishnan, K. *et al.*, Structure of *Mycobacterium tuberculosis* single-stranded DNA-binding protein. Variability in quaternary structure and its implications. *J. Mol. Biol.*, 2003, **331**(2), 385–393.
37. Dubiel, K. *et al.*, Structural mechanisms of cooperative DNA binding by bacterial single-stranded DNA-binding proteins. *J. Mol. Biol.*, 2019, **431**(2), 178–195.
38. Waldman, V. M., Weiland, E., Kozlov, A. G. and Lohman, T. M., Is a fully wrapped SSB–DNA complex essential for *Escherichia coli* survival? *Nucleic Acids Res.*, 2016, **44**(9), 4317–4329.

ACKNOWLEDGEMENTS. X-ray data were collected at the X-ray facility for Protein X-ray Structure Determination and Protein Design at the Indian Institute of Science, Bengaluru, supported by Science and Engineering Research Board, Department of Science and Technology (IRHPA Grant no. IR/SO/LU-003/2010). M.V. is a NASI Platinum Jubilee Senior Scientist. Financial support for his work was provided by the Office of the Principal Scientific Adviser to the Government of India.

Received 1 October 2021; revised accepted 17 December 2021

doi: 10.18520/cs/v122/i3/281-289