frontiers in Microbiology

ORIGINAL RESEARCH
published: 29 June 2021
doi: 10.3389/fmicb.2021.699140

Unlimited Cooperativity of *Betatectivirus* SSB, a Novel DNA Binding Protein Related to an Atypical Group of SSBs From Protein-Primed Replicating Bacterial Viruses

Ana Lechuga[1†], Darius Kazlauskas[2], Margarita Salas[1‡] and Modesto Redrejo-Rodríguez[1,3,4]*

[1] Centro de Biologiìa Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain, [2] Institute of Biotechnology, Life Sciences Center, Vilnius University, Saulėtekio Av. 7, Vilnius, Lithuania, [3] Departamento de Bioquímica, Universidad Autónoma de Madrid (UAM), Madrid, Spain, [4] Instituto de Investigaciones Biomédicas Alberto Sols (CSIC-UAM), Madrid, Spain

Bam35 and related betatectiviruses are tail-less bacteriophages that prey on members of the *Bacillus cereus* group. These temperate viruses replicate their linear genome by a protein-primed mechanism. In this work, we have identified and characterized the product of the viral ORF2 as a single-stranded DNA binding protein (hereafter B35SSB). B35SSB binds ssDNA with great preference over dsDNA or RNA in a sequence-independent, highly cooperative manner that results in a non-specific stimulation of DNA replication. We have also identified several aromatic and basic residues, involved in base-stacking and electrostatic interactions, respectively, that are required for effective protein–ssDNA interaction. Although SSBs are essential for DNA replication in all domains of life as well as many viruses, they are very diverse proteins. However, most SSBs share a common structural domain, named OB-fold. Protein-primed viruses could constitute an exception, as no OB-fold DNA binding protein has been reported. Based on databases searches as well as phylogenetic and structural analyses, we showed that B35SSB belongs to a novel and independent group of SSBs. This group contains proteins encoded by protein-primed viral genomes from unrelated viruses, spanning betatectiviruses and Φ29 and close podoviruses, and they share a conserved pattern of secondary structure. Sensitive searches and structural predictions indicate that B35SSB contains a conserved domain resembling a divergent OB-fold, which would constitute the first occurrence of an OB-fold-like domain in a protein-primed genome.

Keywords: ssDNA, OB-fold, DNA replication, Bam35, *Salasvirus*

OPEN ACCESS

Edited by:
Andrew S. Lang,
Memorial University of Newfoundland, Canada

Reviewed by:
Rodrigo Reyes,
McGill University, Canada
Cynthia Haseltine,
Washington State University, United States

****Correspondence:***
Modesto Redrejo-Rodríguez
modesto.redrejo@uam.es

†Present address:
Ana Lechuga,
Laboratory of Gene Technology, Department of Biosystems, KU Leuven, Belgium

‡Deceased

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *22 April 2021*
Accepted: *08 June 2021*
Published: *29 June 2021*

Citation:
Lechuga A, Kazlauskas D, Salas M and Redrejo-Rodríguez M (2021) Unlimited Cooperativity of Betatectivirus SSB, a Novel DNA Binding Protein Related to an Atypical Group of SSBs From Protein-Primed Replicating Bacterial Viruses. Front. Microbiol. 12:699140. doi: 10.3389/fmicb.2021.699140

HIGHLIGHTS

- Bam35 ORF 2 product encodes a viral single-stranded DNA binding protein (B35SSB).
- B35SSB binds ssDNA in a highly cooperative manner but with no sequence specificity.
- B35SSB-ssDNA binding is mediated by base-stacking and ionic interactions.
- Bam35 and Φ29-related SSBs form a novel group of SSBs from protein-primed viruses.
- The B35-Φ29 SSBs group shares a highly divergent OB-fold-like domain.

INTRODUCTION

The *Bacillus* virus Bam35 is the model virus of the genus *Betatectivirus*, a group of temperate *Tectiviridae* members infecting Gram-positive animal and human pathogens from the *Bacillus cereus* group (Ackermann et al., 1978; Gillis et al., 2018). Besides preying on pathogens of economic and global health relevance, this group of bacteriophages has also raised interest in the last years after the suggestion of a possible evolutionary link between betatectiviruses and the origin of several groups of large DNA viruses (Koonin et al., 2015; Krupovic and Koonin, 2015).

The genome of Bam35 is replicated by a protein-priming DNA replication process (Berjón-Otero et al., 2016), a widespread mechanism for the initiation of genome replication in a number of linear genomes of viruses and linear plasmids. By this mechanism, a specific amino acid of the so-called terminal protein (TP) primes the replication providing a hydroxyl group for the incorporation of the first nucleotide by the viral DNA polymerase and thus it becomes covalently linked to the 5′ genome ends. Among the characterized models of protein-primed DNA replication, the *Bacillus* virus Φ29 from the *Podoviridae* family has been extensively characterized (Salas and de Vega, 2016; Salas et al., 2016). Bam35 genome replication can be carried out *in vitro* with only two proteins, the DNA polymerase (B35DNAP) and the TP (B35TP), as is in the case of genome replication of Φ29 or PRD1, a well-characterized lytic tectivirus from the genus *Alphatectivirus* infecting Gram-negative hosts (Savilahti et al., 1991). However, although not identified in Bam35, a number of accessory DNA binding proteins, such as single-stranded DNA binding proteins (SSBs), increase the efficiency of genome replication *in vitro* and are essential *in vivo* in the case of Φ29 and other systems (Pakula et al., 1990; Salas, 1991; Blanco et al., 1994).

SSBs are ubiquitous factors that protect single-stranded DNA (ssDNA) intermediates required for genetic information metabolism. These proteins not only protect against nucleases attack, breakage and chemical mutagens, but also can play active roles in preventing secondary structures in the ssDNA, recruiting enzymes and stimulating DNA replication by enhancing the processivity, rate and fidelity of DNA synthesis (Pestryakov and Lavrik, 2008; Byrne and Oakley, 2019; Hernandez and Richardson, 2019; Naue et al., 2013). Warranted by all those features, SSBs have also proved to be suitable for diverse molecular biology and analytical applications in biotechnology (Perales et al., 2003; Kur et al., 2005; Chisty et al., 2018).

Although SSBs span a wide diversity of protein groups with little sequence similarity, the vast majority of them share a common structural domain, called oligonucleotide/oligosaccharide-binding fold (OB-fold). This domain consists of a five-stranded β-sheet coiled to form a β-barrel capped by an α-helix (Murzin, 1993; Hyland et al., 2003; Kazlauskas and Venclovas, 2012). There are very few exceptions of SSB structures with no clear similarity to an OB-fold, such as those in adenovirus (Tucker et al., 1994), the hyperthermophilic archaea representing clade Thermoproteales (Paytubi et al., 2012) and Drc protein from the N4-related bacterial viruses (Boon et al., 2020).

Generally, OB-fold domains interact with ssDNA by base stacking with aromatic residues situated in strands 2 and 3 of the β-barrel. Also, cation-π stacking, hydrophobic and hydrogen-bonding with base and ribose moieties contribute to the SSB-bases binding without sequence specificity. Meanwhile, the phosphate backbone is often exposed to the solvent, although, it can also contribute to ssDNA binding through salt bridges and hydrogen bonds (Dickey et al., 2013).

As each OB-fold unit can bind a very short ssDNA tract, SSBs present different modular organization, either by the presence of more than one OB-fold in the same polypeptide, as in the case of eukaryotic RPA, or by oligomerization of independent OB-fold monomers, often by interaction among their C-terminal tails (Pestryakov and Lavrik, 2008). In other cases, the N-terminal mediates oligomerization, as in the case of GA-1 viral SSB, which is a hexamer with a highly efficient DNA binding ability, whereas the Φ29SSB is a monomer with lower DNA binding proficiency (Gascón et al., 2000a). Furthermore, SSB-ssDNA interaction may be required to trigger multimerization complex formation. This can result in a cooperative DNA binding mechanism, common for SSBs whose main function is related to DNA replication, such as recA, the SSBs from phage T4 or *E. coli* (EcoSSB), among others (Jose et al., 2015; Dubiel et al., 2019). However, cooperativity can be also low in some SSBs, as T7 gp2.5 and eukaryotic RPA (Kim et al., 1992; Kumaran et al., 2006). Further, some SSBs, like EcoSSB, can show either limited or unlimited cooperative binding depending on the salt concentration (Antony and Lohman, 2019).

In this work, we identified and characterized the product of the Bam35 ORF 2 (hereafter B35SSB) as a novel viral SSB. Biochemical characterization of B35SSB reveals high specificity for ssDNA binding in a greatly cooperative manner, which results in the stimulation of processive DNA replication. Site-directed mutagenesis also allowed us to disclose some aspects of the DNA binding mechanism, similar to that reported for other SSBs with a canonical OB-fold. Further, phylogenetic analyses and consensus structural predictions showed that this protein belongs to a diverse clade of SSBs that contains, besides betatectivirus orthologs, Φ29-related podoviruses and diverse bacterial proteins that might be also coded by uncharacterized or overlooked viral genomes. This group of

viral SSBs would constitute a novel clade of SSBs from protein-primed replicating genomes that would contain an OB-fold-like conserved domain.

RESULTS

B35SSB Is a Proficient Single-Stranded DNA Binding Protein

Characterization of the protein P2 from the betatectivirus-related *B. cereus* plasmid pBClin15, orthologous to the Bam35 ORF2 product, showed DNA binding capacity (Stabell et al., 2009). Other works proposed that Bam35 ORF2 encodes an SSB similar to other bacteriophage SSBs based on HHpred searches (Jalasvuori et al., 2013). To confirm that the product of the Bam35 ORF2 is a viral SSB protein, we analyzed its nucleic acid binding ability by electrophoretic mobility shift assay (EMSA) using a variety of nucleic acid substrates. First, a 50-mer ssDNA and dsDNA oligonucleotides were incubated with increasing amounts of purified protein (hereafter B35SSB). The SSB from *E. coli* (EcoSSB) and the TP from Φ29 were used as positive controls (**Figure 1A**), as the EcoSSB binds more efficiently ssDNA (lanes 2 and 8) while the TP shows higher affinity for dsDNA (lanes 3 and 9). In presence of B35SSB, the ssDNA substrate showed slower mobility due to the formation of B35SSB-ssDNA stable complexes, being the whole probe bound by 500 nM B35SSB (lanes 4–6). The lack of intermediate bands suggests a cooperative DNA binding. Moreover, although the EcoSSB monomer has a similar molecular weight to B35SSB (19 and 18.5 kDa, respectively), EcoSSB-ssDNA complexes (lane 2) showed higher mobility than B35SSB-ssDNA complexes (lane 6), which indicates that a higher number of B35SSB molecules are bound to the ssDNA fragment.

Participation of SSBs in DNA replication initiation steps has been related to a binding capacity to RNA–DNA or to duplex dsDNA (Pakula et al., 1990; Shereda et al., 2008). Nevertheless, dsDNA binding of B35SSB seems insignificant as only a very faint shifted band was observed (lane 12′) at high a concentration of B35SSB (1000:1, B35SSB:dsDNA molar ratio), whereas B35SSB could shift nearly all ssDNA substrate at a 125:1 (B35SSB:ssDNA) molar ratio (lane 5). Recent works have also highlighted the RNA binding capacity of several SSBs from bacteria and viruses, leading to RNA metabolism regulation (Shi et al., 2013; Boon et al., 2020). However, when RNA binding capacity of B35SSB was analyzed, only a single weak band of about 10% of the substrate shifted at a very high concentration (5 μM) of B35SSB (**Figure 1B** , lanes 7–12) indicating an inefficient binding to RNA, as compared with ssDNA of the equivalent sequence (**Figure 1B**, lanes 1–6). While EcoSSB-ssDNA and EcoSSB-RNA complexes migrate a similar distance, B35SSB-RNA migrated faster than B35SSB-ssDNA complexes. This suggests a different mode of binding of B35SSB to RNA where a smaller number of molecules could bind to the RNA substrate with less affinity than to ssDNA. Therefore, we conclude that B35SSB binds ssDNA proficiently, with a great preference over dsDNA or RNA substrates.

We also analyzed different ssDNA substrates to determine the sequence-specificity in DNA binding of B35SSB. Overall, no differences were observed when two different 50-mer ssDNA substrates were used (**Figure 1C**, lines 2–4 vs. 6–8). However, when we compared homopolymeric 33-mer substrates, we found that, while in the presence of B35SSB poly-dT oligonucleotides mobility was shifted in a smeared gradient (lines 12–14), indicating partial or unstable DNA binding, B35SSB binding capacity is severely impaired on poly-dA homopolymeric ssDNA (lines 9–11). This indicates a preference for pyrimidines over purines, as previously reported for other OB-fold containing SSBs (Pestryakov and Lavrik, 2008), suggesting a similar nature of interactions involved in B35SSB binding mechanism to ssDNA.

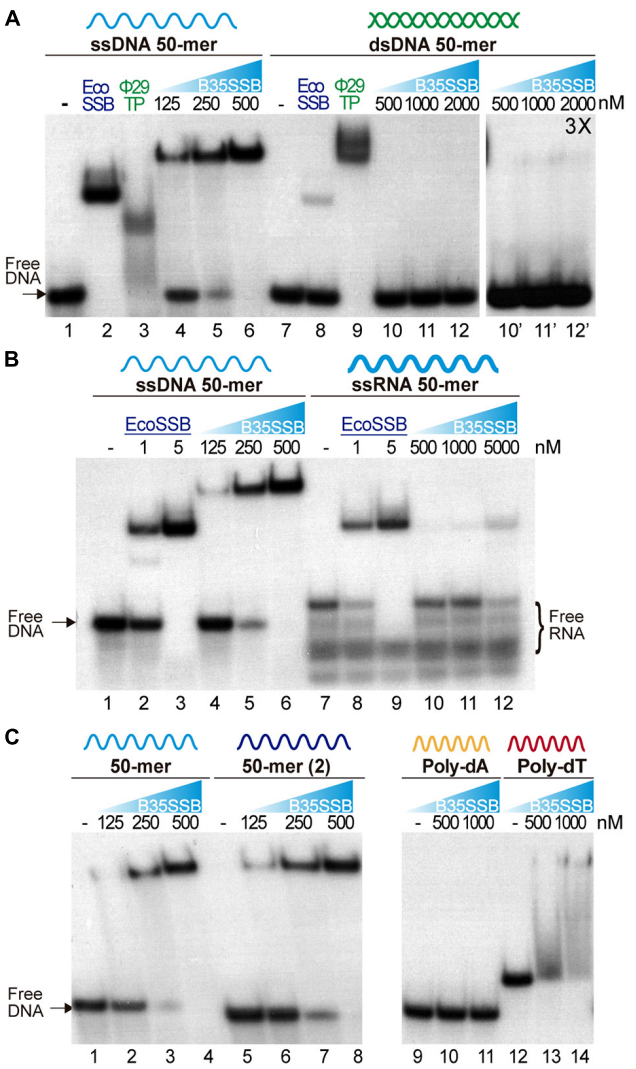

FIGURE 1 | Nucleic acid binding capacity of B35SSB. DNA binding capacity of B35SSB is analyzed by comparison of mobility shift of diverse substrates, ssDNA vs. dsDNA **(A)**, ssDNA vs. ssRNA **(B)** or different sequence contexts in ssDNA **(C)**. EMSA were performed in presence of the indicated amounts of B35SSB and 2 nM of radiolabeled ssDNA, dsDNA, or RNA as indicated. Negative controls, without protein, were loaded in the lanes indicated by a dash. Where indicated, 1 μM EcoSSB and 20 nM Φ29 terminal protein (Φ29TP) were used as controls.

Altogether, these results confirm the product of Bam35 ORF 2 as a viral SSB, with an intrinsic high affinity for ssDNA and low sequence specificity, but a negligible capacity of binding to other substrates.

B35SSB Stimulates Processive DNA Replication

To ascertain the role of B35SSB in DNA replication, we next determined its influence during processive DNA replication in a rolling circle replication assay *in vitro*. Thus, we used a singly primed ssDNA of M13 as a template, which can be replicated by Bam35 DNA polymerase (B35DNAP), giving rise to a replication product larger than full-length M13 DNA thanks to its ability to couple processive DNA replication and strand displacement capacity (**Figure 2A**; Berjón-Otero et al., 2015). When increasing amounts of SSB were added to the reaction, a stimulatory effect of B35SSB was observed in terms of a higher amount of DNA replication product (**Figure 2B**, lanes 1–3) being about 1.5-2X fold in the presence of 2 μM B35SSB (lane 3 vs. lane 1). However, in the presence of very high concentrations of B35SSB (lanes 4–5), a decrease of ssDNA product up to 65% as compared with control conditions was observed (lane 5 vs. 1), resulting in the strong reduction of product larger than full-length M13 DNA. Thus, B35SSB stimulates DNA replication by B35DNAP in these conditions, although a fine-tuned protein:DNA ratio may be required. This stimulatory effect was also observed in a time-course experiment where B35SSB effect could be detected at all the reaction times analyzed, obtaining approximately a twofold increase in the DNA synthesis rate (**Figure 2C**). Interestingly, rather than specific, the stimulation of its cognate DNA polymerase by B35SSB seems generalized. As such, B35SSB can stimulate also Φ29 DNA polymerase and the B35DNAP can be stimulated by *E. coli* and Φ29 SSBs, although the optimum SSB:DNA ratio might be different (**Figures 2D,E**).

As in the case of Φ29SSB (Soengas et al., 1995), the stimulation of processive DNA replication by B35SSB seems to be facilitated by a DNA helix-destabilizing activity (**Supplementary Figure 1A**), which further supports a role in genome replication. Although this DNA unwinding capacity can be also required at replication initiation steps (Matsumoto and Ishimi, 1994), B35SSB seems irrelevant for protein-primed DNA replication initiation *in vitro* (**Supplementary Figures 1B–D**). Therefore, we conclude that rather than in early steps of protein-primed genome replication, B35SSB would play a role in the processive genome replication.

B35SSB Is a Monomer That Binds ssDNA in a Highly Cooperative Manner

To analyze the DNA binding mechanism of B35SSB, we first checked the effect of ssDNA length. As shown in **Figures 3A,B**, whereas in the presence of B35SSB the 15-mer DNA fragment only gave rise to a minor smear (<10%) at high protein concentration (1 μM, lane 3), 25 and 50% of the 33-mer substrate was stably bound with 250 and 500 nM, respectively (lanes 6 and 7) and 250 nM of protein was enough to stably bind 60 and 70% of the 50 and 80-mer probe, respectively (lanes 11 and 15). Minimal binding site of most SSBs is usually 4–6 nucleotides per monomer (Murzin, 1993), which downplays the possibility that a minimal DNA length for effective interaction at the DNA binding site is longer than 15 nucleotides. These results indicate that B35SSB binds ssDNA in a highly cooperative way that would be enhanced by the length of the substrate. Moreover, contrary to other SSBs (Lohman et al., 1986; Pant et al., 2018), B35SSB cooperativity is not affected by the ionic strength, as high salt concentrations impair DNA binding moderately without affecting the pattern of retarded DNA or leading to smeared or intermediate products (**Figure 3C**).

Cooperative DNA binding may also depend on protein-protein interactions that facilitate the formation of long, stable protein-DNA complexes. We examined the oligomerization state of B35SSB in solution by analytical ultracentrifugation in a linear 15–30% glycerol gradient. The sedimentation peak of B35SSB revealed that it is homogeneously a monomer in solution (**Figure 4A**). This suggests that oligomerization would occur and be triggered by DNA binding.

We also performed glutaraldehyde cross-linking assays to disclose the cooperative DNA binding mechanism. Glutaraldehyde covalently cross-links the amino groups of lysine residues and DNA bases with a very close carbonyl group of amino acids, allowing the analysis of protein-protein and protein-DNA complexes in denaturing SDS-PAGE (Jaenicke and Rudolph, 1986; Gamsjaeger et al., 2015). In the absence of ssDNA, and, in agreement with the previous result, cross-linked B35SSB is still a monomer (**Figure 4B**, lane 2). Only a minor band corresponding to dimeric species was detected in presence of the cross-linker, likely due to the stabilization of weak or transient protein–protein interactions that disappear in the presence of large ssDNA fragments. On the other hand, in the presence of ssDNA of increasing lengths, diverse oligomeric species can be detected (lanes 3–6). Longer substrates resulted in higher bands that would correspond to oligomers beyond the resolution of the gel. A detailed analysis with DNA fragments from 20 to 25-mer, increasing one nucleotide at a time, allowed us to determine the minimal binding site of B35SSB (**Figure 4C**). While with 20–23 mer substrates the main oligomeric species corresponded to a tetramer (lanes 3–6), when 24 and 25-mer substrates were added, the tetramer band lost intensity and the band corresponding to the pentameric species was detected (lanes 7 and 8). This indicates that B35SSB has a binding site that can span 4–5 nt. Altogether, these results show that B35SSB is a monomer in solution, and it forms oligomers as it binds to ssDNA cooperatively with a minimal binding site of 4–5 nucleotides per monomer, being cooperativity stimulated when longer substrates are present, in agreement with the EMSA results.

To further assess the role of protein-protein interactions in cooperative B35SSB-ssDNA binding, we performed cross-linking assays adding ssDNA in excess to the reaction. Although the B35SSB:ssDNA ratio was increased to 1:100, the intensity of the high molecular weight band corresponding to oligomeric species remained unchanged (**Figure 4D**). This indicates that, although B35SSB nucleoprotein complexes are only formed by interaction with DNA, they would be stabilized by protein-protein interactions that hold monomers close to each other.

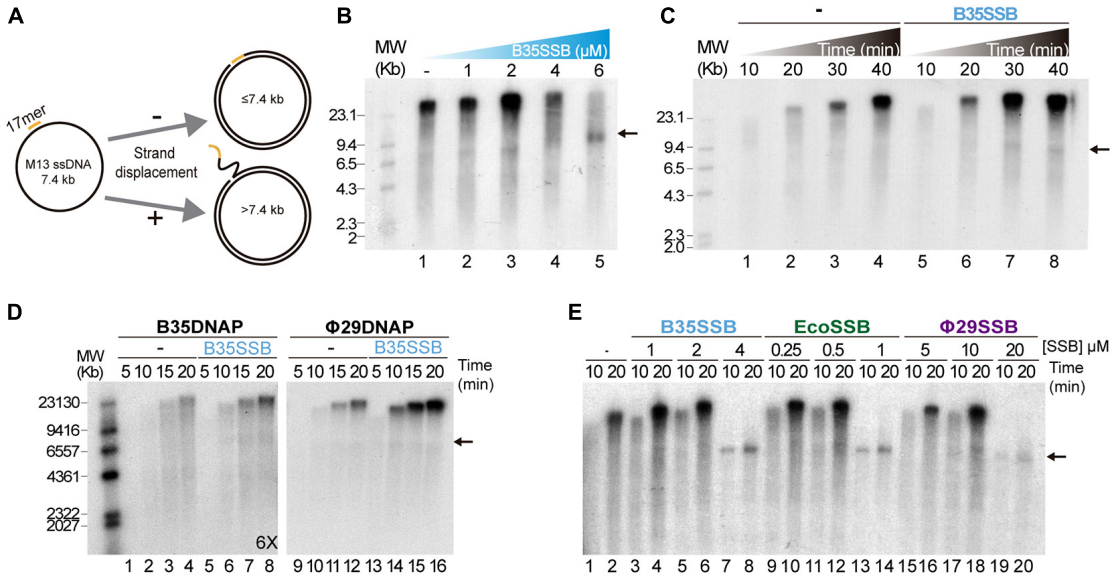

FIGURE 2 | B35SSB stimulation of processive DNA replication. **(A)** Schematic representation of M13 ssDNA rolling-circle replication. In strand-displacement conditions, a product much larger than the full-length M13-DNA (7.4 kb) is obtained. **(B)** Effect of B35SSB concentration in ssDNA M13 replication. Reactions were carried out for 20 min as detailed in Methods. **(C)** Time-course analysis of addition of B35SSB. Reactions in absence or presence of 2 μM B35SSB were carried out and stopped at the indicated times. **(D)** Stimulatory capacity of B35SSB (2 μM) on M13 DNA replication performed by Bam35 and Φ29 DNA polymerases. Longer autoradiography exposition time is indicated. **(E)** Effect of different SSBs on B35DNAP processive DNA replication. Reactions were carried out using primed M13 circular ssDNA as template and the indicated DNA polymerase. After incubation at 37°C, the length of the synthesized DNA was analyzed by alkaline 0.7% agarose gel electrophoresis alongside a λ DNA (MW) and autoradiography. M13 ssDNA unit length is indicated by a black arrow. See section "Materials and Methods" for details.

This oligomer formation occurs even at very high ssDNA concentration which indicates that the SSB binds in a non-distributive manner where the binding of one single SSB leads to the binding of the next monomer, and so forth, until the ssDNA molecule is completely covered. These and the EMSA results confirm that B35SSB binds to ssDNA in a highly cooperative mechanism where both protein–DNA and protein–protein interactions play a fundamental role.

DNA Binding Is Mediated by Ionic and Hydrophobic Interactions

As outlined in the Introduction, SSBs mainly bind ssDNA through the nucleotide bases, *via* base-stacking interactions to aromatic residues and cation-stacking contacts (Dickey et al., 2013). We aimed to identify the main molecular interactions involved in the ssDNA-binding mechanism of B35SSB and, likely, other B35SSB-related proteins. We selected several residues with diverse chemical properties and conservation degrees in order to analyze their contribution to DNA binding. A total of 13 variants were generated (F48A, S49A, Y63A, Y63F, D85A, V87A, Y117A, Y117F, V124A, K130A, Y150A, Y150F, and K156A), but two of them (D85A and Y117A) gave rise to insoluble protein in the inclusion bodies (**Supplementary Figure 2A**). This suggests that these modifications have a severe impairment in the structural stability of the protein. Solubility and expression were also impaired, to a lesser extent in the case of Y63A, whose purification resulted in low yield.

The ssDNA-binding ability was strongly impaired for all the obtained variants, except for Y150A and Y150F (**Figure 5A**). These two variants showed an ssDNA-binding capacity similar to the WT. In turn, V87A and Y117F and especially the variants F48A and V124A formed mainly unstable ssDNA-protein complexes detected as a high proportion of quantified smear in EMSA gels (**Figures 4A,B**). These variants correspond to hydrophobic residues that may be involved in base-stacking/hydrophobic interactions. Also, the ssDNA-binding capacity of Y63A is strongly impaired while the conservative substitution in the conservative Y63F showed a milder defect on DNA binding, which suggests an important role of Y63 in ssDNA binding by base-stacking interactions. Finally, K130A and K156A ssDNA binding is strongly reduced, indicating that ionic interactions can be involved in B35SSB–ssDNA interaction.

Interestingly, a clear increase in the intensity of protein–DNA complexes was detected for large ssDNA fragments (50-mer) over shorter oligonucleotides (33-mer). As shown in **Figure 5B**, even the proficient Y150A and Y150F variants formed more stable complexes with the 50-mer substrate, which suggests that Y150 residue may have some role in the interaction with the ssDNA. This effect was stronger for S49A and V87A, although a higher SSB concentration was required to detect stable complexes (lanes 4 and 7). When we use a 33-mer substrate, cooperative binding would be less efficient, hence these results would imply that these variants are not defective in cooperative binding. On the other hand, V124A and F48A could not form stable protein-DNA complexes even with the longer substrate and at high

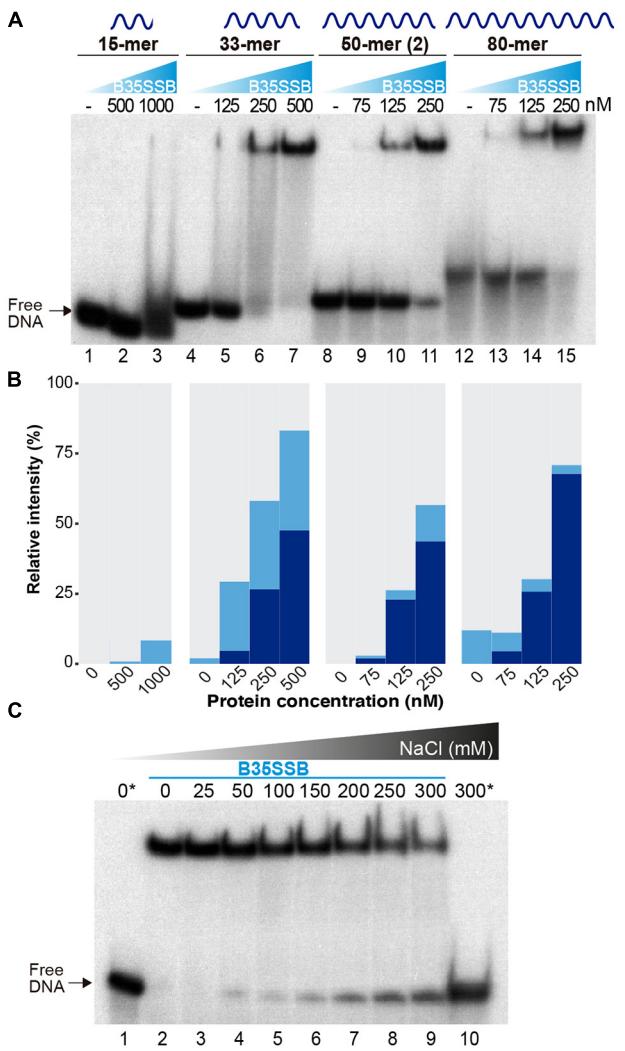

FIGURE 3 | Characterization of cooperative DNA binding. The effect of substrate DNA length in ssDNA binding as determined by oligonucleotides of the indicated length was analyzed by EMSA **(A)** and quantified by gel densitometry with ImageJ **(B)**. Note that a lower B35SSB concentration range was used for longer ssDNA fragments. **(C)** DNA binding assay in the presence of increasing concentrations of NaCl. Samples were prepared in a sample buffer without salt and were supplemented with NaCl to obtain the indicated extra salt concentration. The reactions also contained a basal concentration (12.7 mM NaCl) contributed by the reaction components. Samples with no protein are indicated with an *.

protein concentration (lanes 3 and 9), suggesting a very unstable binding to ssDNA which would not be stabilized in conditions that facilitate cooperative binding. In line with this observation, the use of cross-linking agent could stabilize DNA complexes for Y63F, V87A, Y117F, K130A, Y150A, and Y150F variants (**Figure 5C**), although only Y150A and Y150F complexes showed a similar intensity to the WT ones. On the other hand, for F48A, S49A, and K156A, oligomers formation was observed as a faint band only when a 40-mer oligo was used as substrate and no oligomers were detected for V124A. These results would mean a severe defect in ssDNA binding, in agreement with the poor binding detected by EMSA.

Finally, the increase of the ssDNA:protein ratio does not affect the results of the cross-linking experiment for any of the analyzed variants (**Figure 5D**), downplaying a highly specific role in cooperativity. Nevertheless, the loss of charge variants K130A and K156A seems to increase the dimer band, which cannot be detected in the case of the WT and other variants when a large substrate was used, which also may indicate a role of these basic residues in protein–protein interactions and, therefore, in cooperativity.

A Novel, Highly Divergent OB-Fold-Related Domain in a New Group of Protein-Primed Viral SSBs

As mentioned above, protein-primed replicating genomes code for divergent single stranded DNA binding proteins, with non-OB-fold conformation or without structural characterization (Kazlauskas et al., 2016). The latter is the case of proteins P12 and P19 from PRD1, which are atypical single-stranded DNA binding proteins that would play a key role in the protein-primed viral genome replication, protecting DNA intermediaries and stimulating DNA synthesis (Pakula et al., 1990, 1993). These proteins are conserved in PRD1-related viruses, but no ortholog has been detected in other *Tectiviridae* members, including Bam35 (Ravantti et al., 2003). Although no OB-fold protein has been identified in Bam35 proteome, we have showed here that the product of ORF2 is a DNA binding protein with a great specificity for ssDNA, whose binding mechanisms entails base-stacking and ionic interactions. Thus, we can conclude that it shares most of the biochemical properties of OB-fold containing SSBs.

As mentioned above, previous reports had already annotated Bam35 ORF2 product as a putative SSB (Jalasvuori et al., 2013). In line with this, B35SSB and betatectivirus orthologs are included in a small Pfam family of viral SSBs, along with Φ29 SSB and other proteins (PF17427). We aimed to clarify both the sequence similarity of B35SSB with known SSBs and the presence of a conserved OB-fold. For the first task, we performed Jackhammer (Potter et al., 2018) iterative searches seeded with the amino acid sequence of B35SSB against the Uniref90 database. We obtained a similar but extended dataset, compared with the PF17427 group. Thus, although we did not detect any sequence corresponding to known OB-fold structures, we obtained several hits corresponding to characterized or predicted SSBs from Φ29 and diverse podoviruses, like GA-1 or asccphi28, other possible phage proteins, as well as diverse protein sequences from bacteria, including *Bacillus sp.* and other *Firmicutes*, which may correspond to misannotated bacteriophage or prophage proteins rather than chromosomal bacterial ORFs. Similar to the Pfam cluster, we also obtained a few metagenomic sequences as well as some eukaryotic proteins from nematodes and insects that might have an origin on horizontal transfer events, though sample contamination cannot be ruled out.

As expected from such a diverse dataset of short sequences, phylogeny reconstruction not always yielded confident clades (**Figure 6**). However, it should be highlighted that B35SSB and

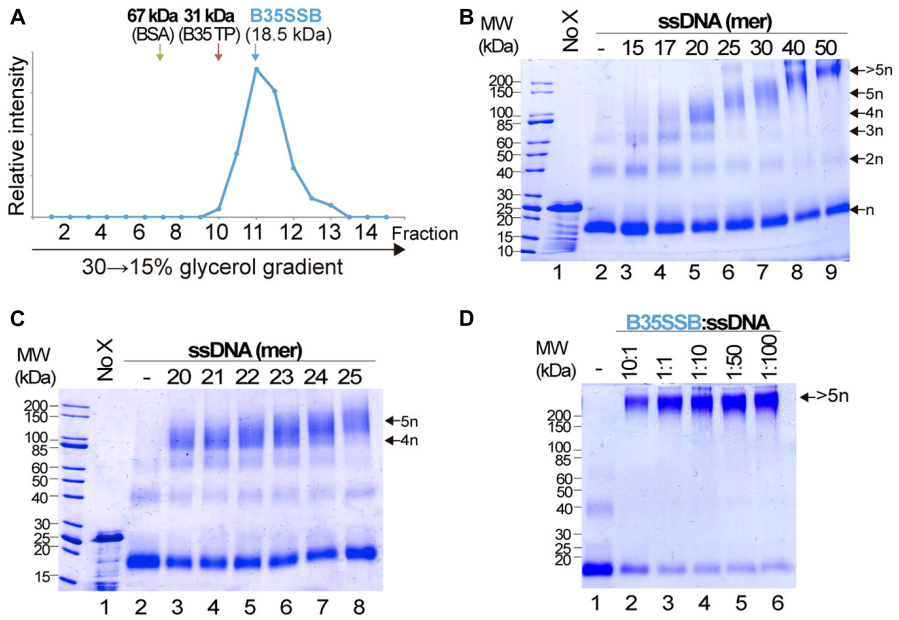

FIGURE 4 | Analysis of B35SSB modularity and cooperativity. **(A)** Determination of the oligomerization state of B35SSB in solution. Recombinant B35SSB was subjected to sedimentation in 15–30% (w/v) glycerol gradients in the presence of molecular weight markers as indicated. The fractions at which the maximal amount of each protein appears are indicated by arrows. **(B)** Cross-linking protein interaction analysis using different ssDNA sizes (**Supplementary Table 1**). B35SSB (1 μM, equivalent to 2 μg) was incubated with 165 ng of the indicated oligonucleotides. After glutaraldehyde cross-linking, total protein was precipitated and analyzed by SDS-PAGE. The number of SSB cross-linked species is indicated on the right. **(C)** Determination of the B35SSB minimal binding site by crosslinking assays. **(D)** Effect of B35SSB:ssDNA ratio in ssDNA-SSB binding. Assays contained 1 μg B35SSB and increasing amounts of the 50-mer oligonucleotide to obtain the indicated protein:DNA ratio. See Methods for details. MW, molecular weight; No X and dash (–) stand for no crosslinker and no ssDNA lanes, respectively.

its betatectivirus orthologs form a very compact group, within a well-supported clade (bootstrap value of 98) of more diverse sequences annotated as *Bacillus* proteins, which suggests a single event of acquisition of this sequence in an ancestral betatectivirus.

The combination of the multiple sequence alignment with secondary structure predictions allowed us to identify a core of about 100 residues that overall share a common fold. Remarkably, the C-terminal part of this fold might be a reminiscent of an OB-fold, as it is made up of five β-strands and one α-helix. Next, we did HHpred search using B35SSB as a query. Among hits, we found the profile of an OB-fold protein (ECOD_001121967_e4m8oA1), albeit with a very low probability (32%) (**Supplementary Figure 4**).

To further test the presence of an OB-fold in the Bam35-Φ29 SSBs group, we generated an extended dataset by the addition of sequences from previous phylogenetic analysis of diverse groups of OB-fold-containing cellular and viral SSBs (Murzin, 1993; Ravantti et al., 2003), updated with our own searches (see Methods), which attempts to span all known SSBs groups. To improve the sensitivity of comparisons, we built HHsearch profiles for all sequences of SSBs, did their all-against-all comparison and clustered the results using CLANS (Potter et al., 2018) at *P*-value of 1e-08. Clustering analysis revealed that the new SSBs clade form a compact and independent group, without significant relationship to any particular OB-fold family (**Supplementary Figure 5A**). Only when we raised the *P*-value to 1e-04, connections between the B35-Φ29 SSBs group and other OB-fold groups appeared (**Supplementary Figure 5B**), indicating a very remote similarity.

Since diverse sensitive sequence comparisons did not find a clear similarity between the B35-Φ29 SSBs and known OB-folds, we decided to perform structural comparisons. To do so, we first obtained structural models of B35SSB with diverse comparative modeling and contact-based approaches. Our notion that B35SSB has an OB-fold-like domain was supported by three different methods [RaptorX (Källberg et al., 2012), trRossetta (Yang et al., 2020), and RosettaCM (Kim et al., 2004; Song et al., 2013)], which yielded similar models, of intermediate-good quality, as assessed by VoroMQA and ModFold8 (Maghrabi and McGuffin, 2017; Olechnovič and Venclovas, 2017; **Figure 7A**). B35SSB structural models consisted of an N-terminal motif with a disordered region between two α-helixes and a C-terminal domain made up of a five-stranded β-sheet and one α-helix. In agreement with the secondary structure predictions, the structural model of B35SSB somewhat resembles an OB-fold, particularly in the C-terminal two-thirds portion (residues 76–167).

We then performed structural searches of the three models using Dali Server (Holm, 2020). The best hits obtained (**Supplementary Figure 6**) included several DNA binding proteins, but only one SSB, the T4 SSB (gp32 protein). The structural alignment between the RaptorX model of B35SSB and the T4 SSB structure (Shamoo et al., 1995; **Figure 7B**) indicates an overall similarity between the conserved fold of B35-Φ29 SSBs group and the T4 SSB, spanning the β-sheets backbone from

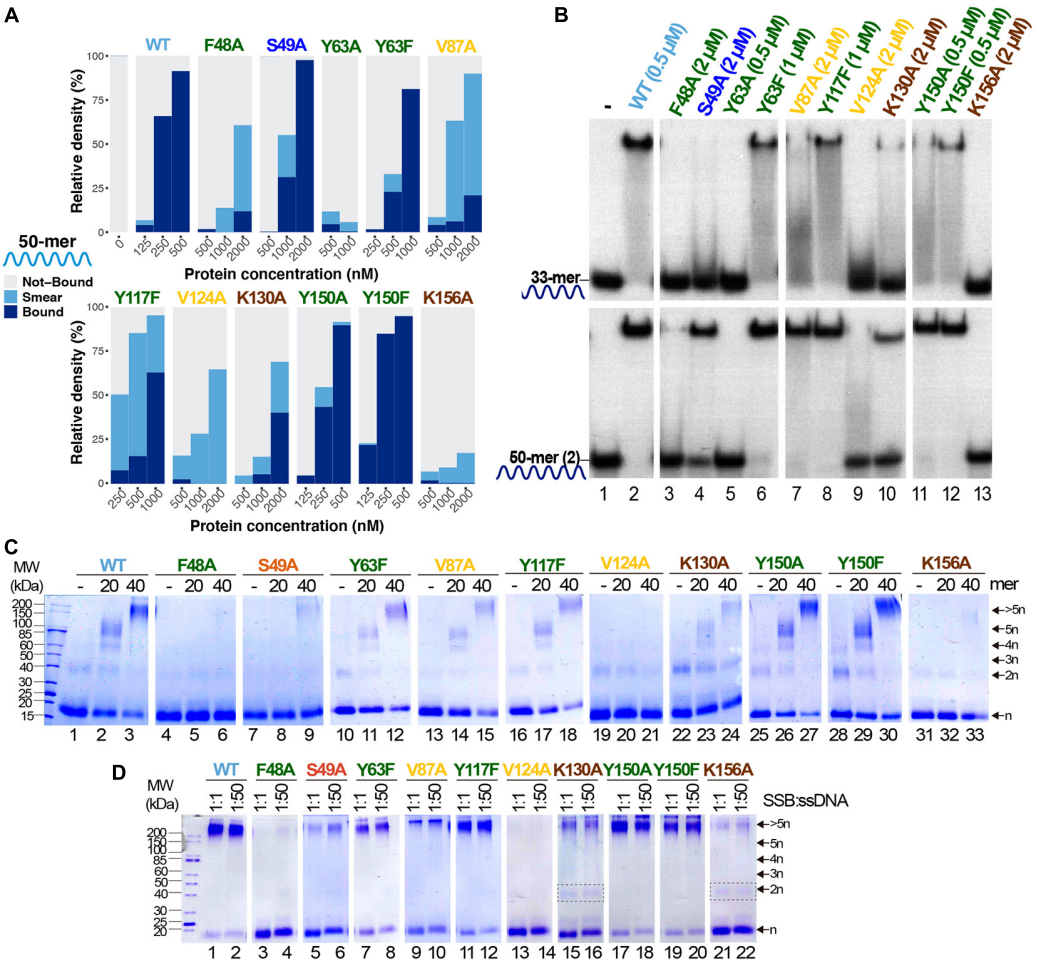

FIGURE 5 | Study of ssDNA-binding ability of B35SSB variants. **(A)** Representative graphic of EMSA assay with B35SSB variants quantified with ImageJ. EMSA was performed using the 50-mer ssDNA probe (**Supplementary Table 1**) and the indicated concentration of B35SSB. Note that the B35SSB concentration range was adapted for each variant. **(B)** Effect of ssDNA size in the ssDNA-binding capacity of B35SSB variants. EMSA was performed incubating the indicated concentration of B35SSB or its variants with 33-mer or 50-mer(2) substrates as previously described. **(C)** B35SSB variants cross-linking analysis using different ssDNA sizes. **(D)** Effect of SSB:ssDNA ratio in binding. One μM B35SSB (or variants) was incubated either with 1 or 50 μM 50-mer(2) ssDNA, according to the ratio protein:DNA indicated. Dimeric species for K130A and K156A are indicated with a gray rectangle. See section "Materials and Methods" for details.

subdomains I (Zn finger) and II (DNA binding) and the α-helix that caps the OB-fold.

In conclusion, the conserved betatectiviral SSB contains a predicted alpha-beta complex conserved fold, only distantly related to an OB-fold. This domain seems to be conserved among SSBs from protein-primed replicating viral genomes, spanning Φ29 and related *Picovirinae* proteins, thus constituting the first OB-fold-like protein in a protein-primed genome.

DISCUSSION

Extremely Cooperative ssDNA Binding Capacity in B35SSB

The biochemical characterization of B35SSB confirmed some features common to all SSBs, and some specific ones shared only with other protein-primed replication systems. Thus, we found that B35SSB has a high specificity for ssDNA over dsDNA and RNA. Both B35SSB and Φ29SSB display negligible dsDNA-binding capacity (Martin and Salas, 1988). This contrasts with other DNA-binding proteins of protein-primed replicons such as the PRD1 P12 and the adenovirus DBP in which dsDNA-binding ability is reduced but clearly better than in B35SSB (Pakula et al., 1990; Stuiver and van der Vliet, 1990). Interestingly, bacteriophage T4 SSB also shows a high specificity for ssDNA over dsDNA, due to the narrow DNA binding groove (Shamoo et al., 1995), which would have similar width in B35SSB according to RosetaCM and RaptorX models and thus could also explain the negligible binding to dsDNA of B35SSB (**Figure 7**).

B35SSB binds to ssDNA in a non-sequence dependent manner, which also correlates with the non-specific DNA binding by most SSBs as well as the related pBClin15 ORF 2 gp (85% amino acid identity with B35SSB) (Murzin, 1993; Stabell et al., 2009). B35SSB does show a preference for pyrimidines, which has

FIGURE 6 | B35SSB and Φ29 related viruses belong to the same clade of SSBs from protein-primed genomes. Maximum-likelihood tree **(left)** generated using the trimmed multiple sequence alignment of B35SSB-related proteins **(right)** and visualized with the ggtree package for R software (Yu et al., 2018). Strain names of phages belonging to *Picovirinae* (blue) or *Betatectivirus* (orange) groups are indicated. Secondary structure prediction was obtained with the Jpred4 (Drozdetskiy et al., 2015) and is indicated above the alignment, with the conserved C-terminal region (residues 76–167 in B35SSB) boxed. Residues are colored by their biochemical properties: aromatic (yellow), hydrophobic (orange), uncharged (purple), positively charged (blue), negatively charged (red), others (dark blue).

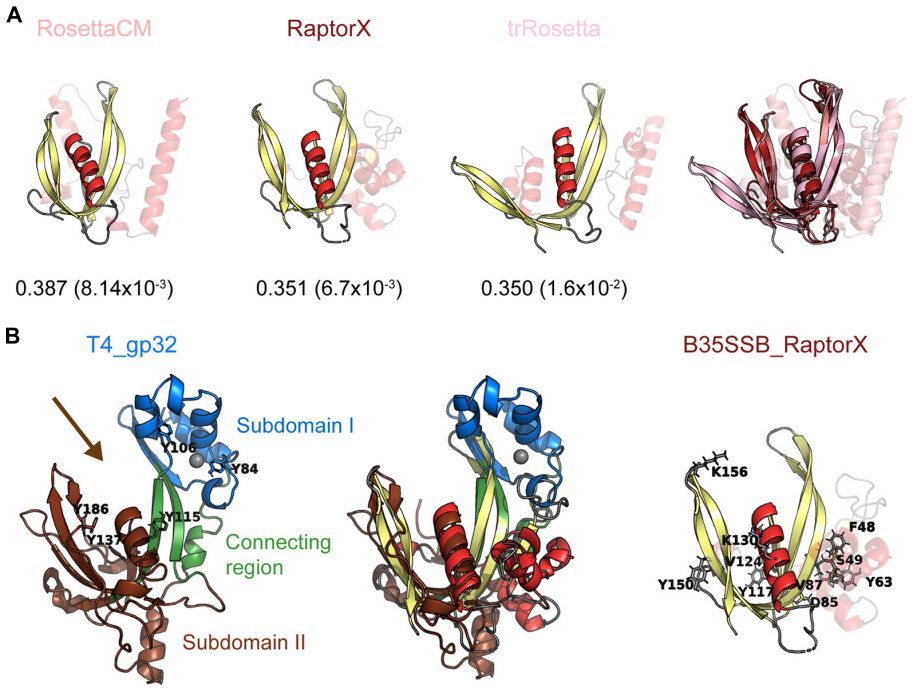

FIGURE 7 | B35SSB and related proteins can constitute a novel SSB fold, divergent to known OB-folds. **(A)** Best structural models of B35SSB, obtained with three different methods as indicated. The N-terminal 76 residues are displayed as semitransparent to show clearly the C-terminal domain rich in β-sheets that might be reminiscent of an OB-fold. The VoroMQA Score and the model confidence *E*-value from ModFold8 server are also shown. Note that VoroMQA considers "likely good models" those with Score > 0.4 and "likely bad models" when the Score is less than 0.3. The rightmost figure represent a structural alignment of the three models. **(B)** Structural alignment of T4 SSB and RaptorX model of B35SSB (middle). Cartoon representation of T4 gp32 (PDB ID 1GPC) is shown on the left. Protein subdomains are colored as in Shamoo et al. (1995) and tyrosine residues involved in DNA binding (Shamoo et al., 1989) are highlighted as sticks. The zinc cation bound to T4 gp32 Subdomain I is represented as a gray sphere. The arrow marks the DNA binding cleft. The right image represents the B35SSB model with the residues analyzed displayed as sticks.

also been described for some OB-fold proteins and explained by a less thermodynamically favorable interaction of aromatic side chains with the purines due to steric hindrance and inefficient base-stacking (Kozlov and Lohman, 1999; Touma et al., 2016).

Thus, our results point to a DNA-binding mechanism involving base-stacking interactions with aromatic residues, similar to that of OB fold-containing SSBs.

In line with the results of the Y2H Bam35 intraviral interactome (Berjón-Otero et al., 2017), where no self-interaction of this protein was observed, B35SSB is a monomer in solution. The podoviral Φ29SSB and Nf SSBs are also monomers but GA-1 SSB is generally a hexamer in solution, although at high salt concentration (200 mM NaCl) it is detected as a monomer (Gascón et al., 2002). Although a similar salt concentration was used in our ultracentrifugation assays, a lower concentration was used in the crosslinking assays, confirming the monomeric state of B35SSB regardless of the salt concentration.

The B35SSB very high cooperativity was further supported by an enhanced binding to longer ssDNA substrates and, more importantly, by cross-linking assays in which, even at an extremely low B35SSB:ssDNA ratio (1:100), only full-length covered substrates could be detected (**Figure 4D**). In contrast, *S. solfataricus* SSB monomers can be cross-linked at a ratio of 1:1 but not in the presence of a somewhat molar excess of ssDNA (1:4), indicating that in non-processive binding the monomers distribute evenly along and between the nucleotides (Touma et al., 2016). Positive cooperativity can be mediated by protein-protein direct contacts between the nearest neighbors such as *E. coli*, *T. thermophilus,* and T4 SSBs (Casas-Finet et al., 1992; Raghunathan et al., 1997; Witte et al., 2008), which can be dependent on salt concentration (Lohman et al., 1986; Pant et al., 2018). However, in the case of B35SSB cooperative binding is unaffected by salt concentration (**Figure 5B**). The results on B35SSB cooperativity show a virtually unlimited cooperativity based on nearest-neighbor interaction, similar to the case of the 35-nt mode of *E. coli* SSB (Lohman and Ferrari, 1994). Comparing B35SSB and EcoSSB, a higher concentration of B35SSB is needed to shift all the 50 nt ssDNA probe, suggesting a smaller binding site. Similar to other monomeric SSBs, B35SSB occludes a small binding site of 4–5 nt (Gascón et al., 2000a; Wadsworth and White, 2001). This binding site suggests that rather than wrapping around the SSB, the ssDNA is bound initially through an exposed binding surface and B35SSB then multimerizes sequentially along the substrate in a contiguous fashion, as proposed for T4, *S. solfataricus* or Enc34 SSBs (Gamsjaeger et al., 2015; Jose et al., 2015; Cernooka et al., 2017). B35SSB binds ssDNA in a highly cooperative

manner, suggesting that both protein–protein and protein–DNA interactions contribute to the formation of stable ssDNA-SSB complexes.

In line with their cooperative binding, a lower concentration of B35SSB than of Φ29SSB is required to obtain proficient DNA binding (Gascón et al., 2000a). This lower affinity in Φ29SSB has been linked to the necessity of the SSB dissociation to allow the DNA polymerase advance during processive replication (Gascón et al., 2000b), which would require a fine-tuned SSB-DNA ratio (**Figure 4**). The SSB of GA-1 is considerably different in sequence, oligomerization, and ssDNA affinity (Gascón et al., 2000a,b). Interestingly, Φ29SSB shares less similarity and identity with GA-1 SSB than GA-1 SSB with B35SSB (17 and 21% identity and 42 and 59% similarity respectively). GA-1 SSB and B35SSB are similar in the N-terminal region but different to Φ29SSB (**Supplementary Figure 7**). On the other hand, although the divergence between B35SSB and Φ29SSB and is bigger, the predicted secondary structure of the C-terminal region is very similar. Altogether, we can differentiate two main regions in B35SSB, the C-terminal domain, a highly conserved novel fold that probably plays an essential role in the ssDNA-binding of Bam35-Φ29 SSBs proteins and an N-terminal domain, less conserved and that might also contribute to DNA binding by enhancing the cooperativity, which is compatible with the DNA binding properties of B35SSB variants F48A and S49A.

In some cases, such as GA-1 and T4 SSBs, the N-terminal domain has a role in cooperativity through self-association mediated by electrostatic interactions (Casas-Finet et al., 1992; Gascón et al., 2002). Nevertheless, within the C-terminal region of B35SSB, we have identified two conserved positively charged residues, K130 and K156, whose non-conservative mutation severely affects the ssDNA-binding ability. EMSA and cross-linking assays suggest a putative role in cooperative binding. However, their predicted location at the limits of the binding groove in the structural models (**Figure 7B**) would also be compatible with a role in the interactions with the sugar-phosphate backbone. That notwithstanding, it should be emphasized that dissection of a unique role in protein–DNA or protein–protein interactions is challenging, as cooperative binding requires both features. Indeed, high salt concentrations disrupt ssDNA binding with no detected effect on cooperativity, suggesting that electrostatic interactions can be required for both protein–protein or protein–DNA contacts. Another residue located in this region, V124, may also contribute to B35SSB cooperativity. This residue is well conserved among Bam35-related SSBs, its ssDNA-binding ability is highly unstable and could not be stabilized in cross-linking assays where no oligomers were observed.

SSBs-ssDNA binding is mainly driven by base-stacking binding to aromatic residues, cation-stacking interactions, and, additionally, hydrophobic and hydrogen- bonding interactions, salt bridges and H bonds (reviewed in ref. Dickey et al., 2013). The generation of B35SSB variants according to this criterion and the conservation level among either the Bam35-Φ29 SSBs or the group of betatectiviral SSBs helped us to propose some residues essential for cooperativity (discussed above) and ssDNA binding. As expected, the level of conservation is in agreement with the role of the different residues as it is shown in the case of the barely affected variants of the poorly conserved residue Y150 (**Figure 5** and **Supplementary Figure 3**).

The ssDNA binding groove of T4 SSB contains three aromatic residues (Y115, Y137 and Y186), being Y115 and Y186 involved in ssDNA binding, whereas Y137 seems dispensable (Shamoo et al., 1989). Thus, the protein–DNA interactions within the putative ssDNA binding groove of B35SSB would be different, as Y117 would correspond to Y137. According to our structural models, Y117 would be within the putative ssDNA binding groove (**Figure 7B**), but contrary to Y137 it is essential for base stacking interactions. Indeed, B35SSB residue Y117 would correspond to the residue Y76 in Φ29SSB (**Supplementary Figure 7**). Intrinsic tyrosine fluorescence quenching assays indicated that Φ29SSB Y76 would be involved in the ssDNA binding (Soengas et al., 1994, 1995). The phenotype of the conservative replacement of Φ29SSB Y76 to phenylalanine is similar to the obtained with the variant Y117F of B35SSB and the non-conservative replacement to alanine resulted in a loss of solubility in both proteins. Additionally, the Φ29SSB variant Y76S which conserves the OH group but not the aromatic ring resulted in a loss of binding ability (Soengas, 1996; de la Torre et al., 2019). Altogether, these results point out this amino acid as an essential residue in ssDNA binding through base-stacking interactions, which is in agreement with its high conservation among the whole clade of viral SSBs.

On the other hand, the B35SSB Y63 variants showed a similar phenotype as Y117 (**Figure 5**) suggesting a similar role in ssDNA binding. This residue is also well conserved among this cluster as an aromatic residue that can be tyrosine or phenylalanine (**Supplementary Figure 3**) which indicates the relevance of the aromatic ring similarly to Y117. However, the instability of the Y63A variant (**Supplementary Figure 2**) and its position outside the predicted DNA binding groove (**Figure 7**), suggests that this residue may be also essential for proper protein folding and/or stability in solution, similar to tyrosines 73 and 92 in the case of T4 SSB (Shamoo et al., 1989). Lastly, among the residues only conserved in betatectiviral SSBs, F48, S49, D85, and V87, our results suggest that F48 could also be necessary for ssDNA binding while D85 and V87 would be required the stability of the protein as their substitution leads to insoluble proteins (**Supplementary Figure 2**).

Overall, in the absence of a more detailed structural characterization of the protein, we can conclude that we have identified several aromatic and positively charged conserved residues essential for ssDNA-binding capacity (F48, Y63, Y117, K130, and K156), being aromatic residues essential for protein–DNA interaction.

Besides its specific binding ability to ssDNA, we have characterized the role of B35SSB enhancing processive DNA replication and its DNA unwinding capacity. The stimulation of M13 ssDNA processive replication, a situation that resembles the replicative intermediates generated during Φ29-like DNA replication of linear genomes, has been described for other SSBs such as Φ29, GA-1, and Nf SSBs. In those cases, the SSB prevents the formation of non-productive binding of the DNA polymerase

to ssDNA (Gascón et al., 2000b). On the other hand, the helix-destabilizing capacity of T4 and Φ29-like SSBs has been proposed to reduce the energy required for the unwinding process in DNA replication (Soengas et al., 1995; Pant et al., 2004), which would contribute to the replication stimulation by B35SSB. The unwinding capacity has also been linked to a role in the initiation step of replication (Matsumoto and Ishimi, 1994). However, we did not detect an effect of B35SSB in the early steps of TP-DNA replication suggesting that its function in replication would occur later during processive replication. On the other hand, it should be noted that the SSBs role in TP-initiation has been described for adenovirus and PRD1 SSBs and has been suggested to be related to, among others, their capacity to bind dsDNA, absent in B35SSB (Pakula et al., 1990; Tucker et al., 1994).

The role of multiple SSBs in replication can be linked to their interaction with other replication proteins, normally polymerases, as illustrated by *E. coli,* T7 and T4 SSBs (Pestryakov and Lavrik, 2008). Particularly, the increase of the elongation rate and the dissociation of the T7 SSB is associated with an interaction with its cognate DNA polymerase (Cerrón et al., 2019). In the case of B35SSB, *in vitro* stimulation of its cognate DNA polymerase seems unspecific. Together with the absence of detected interactions between the B35SSB and other viral proteins, these results point to a generalized role of B35SSB in replication that does not need specific interactions with other viral factors (Berjón-Otero et al., 2017). This is also the case for Φ29SSB, whose interaction with its cognate polymerase has not been detected and indeed it might be dispensable for viral replication (Tone et al., 2012; de la Torre et al., 2019).

A Highly Divergent OB-Fold-Like SSB in Protein-Primed Viral Genomes

As explained above, oligonucleotide/Oligosaccharide (OB)-fold domains are present in most of single-stranded DNA binding proteins (SSBs) and other DNA binding proteins implicated in genome maintenance and safeguarding in all domains of life. In dsDNA viruses, the OB-fold based SSBs are very common, and they have been classified into five groups, each one represented by the structures of the SSBs of *E. coli*, phages T7 and T4, herpesvirus I3 protein and the archaeo-eukaryotic Replication Protein A (Kazlauskas et al., 2016). However, the dsDNA viruses with a protein-primed DNA replication mechanism could be one of the very few exceptions to the universal presence OB-fold in SSBs (Kazlauskas and Venclovas, 2012). In this work, we have characterized the product of ORF 2 of Bam35 (B35SSB) as a proficient SSB that would be involved in the protein-primed replication of the viral genome. Phylogenetic analysis and structural predictions indicate that B35SSB shares a common fold with a diverse group of sequences which include unrelated podoviruses with a protein-primed mechanism, such as *Bacillus* viruses Φ29 or GA-1. This fold, covering the 91 C-terminal residues of B35SSB is made up of five β-strands and two α-helixes and thus it may resemble an OB-fold. However, cluster analysis showed that similarity with known OB-fold containing proteins is very remote. Structural comparisons showed a possible similarity with T4 SSB, although it could not be detected with some of the models, underlining the distant relationship of this conserved fold with known OB-fold containing SSBs, as predicted from the clustering results.

Therefore, we can conclude that Bam35-Φ29 SSBs contain a novel and conserved fold, only distantly related to the canonical OB-fold that may have evolved independently. Within this group, B35SSB and betatectiviral orthologs form a highly compact clade within a very diverse group of sequences including Φ29SSB and related viruses from genus *Salasvirus* and related *Picovirinae,* along with predicted proteins from genomic and metagenomic projects that were annotated as bacterial but might as well be encoded by unforeseen viruses. Although other tectiviruses are described or suggested to replicate their genome by a protein-primed mechanism (Caldentey et al., 1992; Philippe et al., 2018; Caruso et al., 2019), B35SSB would be specific for betatectiviruses and, strikingly, clearly related with highly distant podoviruses. In line with this, viruses from *Betatectivirus* and *Salasvirus* genera infect related hosts from *Bacillus* genus (Alcaraz et al., 2010), which supports a possible ancient common ancestor for Φ29 and Bam35 SSBs, providing a possible scenario for a gene transfer event.

In conclusion, we have identified and thoroughly characterized the single-stranded protein of *Bacillus* virus Bam35, highly conserved among betatectiviruses. This protein belongs to a novel family of SSBs from protein-primed replicating genomes whose conserved fold may be a diverged OB-fold, only distantly related to the T4 group of OB-fold containing SSBs. Thus, Bam35-Φ29 SSBs would represent the first occurrence of an OB-fold-like domain in a protein-primed virus. Within this group, B35SSB stands out by its high cooperativity and ssDNA-binding capacity and may play an important role in DNA stabilization during protein-primed replication.

MATERIALS AND METHODS

Proteins

The Bam35 TP (B35TP), Bam35 DNA polymerase (B35DNAP), Φ29 DNA polymerase (Φ29DNAP), Φ29 terminal protein (Φ29TP), and Φ29 P5 (Φ29SSB) were from the laboratory stocks and purified as described (Martin and Salas, 1988; Lázaro et al., 1995; Mencía et al., 2011; Berjón-Otero et al., 2015). *Escherichia coli* SSB (EcoSSB) was purchased from Qiagen.

The Bam35 genomic DNA was used to amplify the gene 2 flanked by *Kpn*I and *Bam*HI sites by PCR with B35SSB_FW_*Kpn*I and B35SSB_RV_*Bam*HI primers (**Supplementary Table 1**) and Vent DNA Polymerase (New England Biolabs). The digested PCR product was cloned into a pET-52b(+) plasmid (Novagen) in frame with an N-terminal Strep-tag II obtaining the B35SSB expression vector pET52b:B35SSB. B35SSB mutants were obtained by direct mutagenesis using different oligonucleotides (**Supplementary Table 1**) and Pfu DNA Polymerase (Agilent). The expression vectors were used to transform *E. coli* XL-1 and verified by sequencing the complete ORF2 using T7_FW primer. B35SSB and its variants were expressed in *E. coli* BL21(DE3) harboring the pET52b:B35SSB plasmid. Cultures were grown in 200 ml of ZYM-5052 autoinduction medium (Studier, 2005)

with 150 mg/l ampicillin for 16 h at 26 or 30°C. Cells were harvested and disrupted by grinding with alumina, suspended in 6 ml of buffer A [50 mM Tris-HCl pH 8, 1 mM EDTA, 0.04% (v/v) β-mercaptoethanol, 500 mM NaCl, complete^{TM} protease inhibitor cocktail (Roche), 5% (v/v) glycerol] per gram of bacteria and sonicated to reduce the viscosity of the extract by fragmentation of the genomic DNA. Samples were centrifuged for 5 min at 480 × *g* at 4°C to remove cell debris and alumina. To extract the soluble protein fraction, samples were centrifuged for 20 min at 17,200 × *g* at 4°C. Soluble fractions were applied to 1 ml of Strep-tactin column (IBA) per 3 g of bacterial pellet, according to the manufacturer's instructions. The resin was washed with 10 column bed volumes (CV) of buffer A and 15 CV of Strep-tactin washing buffer (IBA). Proteins were eluted in 5 CV of Strep-tactin elution buffer (IBA) and dialyzed overnight against 50 mM Tris-HCl pH 7.5, 50% (v/v) glycerol, 1 mM EDTA, 7 mM β-mercaptoethanol and 0.1 M NaCl. Final protein concentration and purity were analyzed by SDS/PAGE followed by Coomassie blue staining.

Although different growing temperatures were assayed for the expression of D85A, Y117A, and Y63A (**Supplementary Figure 2A**), only small amounts of soluble proteins were obtained precluding purification and analysis of D85A and Y117A and resulting in low concentration and purity of Y63A.

The Strep-tag II from the B35SSB fusion protein was removed by digestion using HRV 3C Protease. The digestion reaction of 15 μg of B35SSB was carried out in a final volume of 100 μl of digestion buffer [50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.05% (v/v) Tween 20, 1 mM DTT] with 3 units of HRV 3C Protease (Novagen) for 16 h at 4°C (**Supplementary Figure 2B**, lanes 1–3). To remove the protease, the reaction was incubated with 100 μl of Glutathione Sepharose Fast Flow resin (Sigma-Aldrich) for 1.5 h at 4°C on a rotating wheel. After centrifugation for 1 min at 1,000 × *g* at 4°C, the supernatant containing the purified untagged protein was collected (lane 6). To evaluate the yield of the purification procedure, the resin was further washed with 50 μl of digestion buffer (lane 4), and the remaining bound protein was eluted by incubation with 20 μl of 2X Laemmli SDS-PAGE sample buffer [8% (v/v) 2-mercaptoethanol, 3.6% (w/v) SDS, 24% (v/v) glycerol, 240 mM Tris HCl pH 6.8, 0.010% (w/v) bromophenol blue] for 3 min at 95°C (lane 5). We then verified that the N- terminal Strep-tag fusion does not affect either B35SSB oligomeric state (**Supplementary Figure 2C**) or its capacity to stimulate processive DNA replication (**Supplementary Figure 2D**). These results indicate that the presence of the N-terminal Strep-tag does not have an important impact on B35SSB characteristics.

Analysis of Nucleic Acid Binding Capacity

Nucleic acid binding capacity was analyzed by gel mobility shift assays (EMSA). The ssDNA and RNA substrates (**Supplementary Table 1**) were 5′ labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The dsDNA substrate was obtained by hybridization of the labeled 50-mer oligonucleotide with an excess of 50-mer complementary (50-mer_c) unlabeled oligonucleotide in hybridization buffer (0.2 M NaCl, 60 mM Tris-HCl pH 7.5).

Binding between the indicated substrate and protein was carried out in a final volume of 20 μl in binding buffer (50 mM Tris-HCl pH 7.5, 1 mM DTT, 4% (v/v) glycerol, 0.1 mg/ml BSA, 0.05% (v/v) Tween 20, 1 mM EDTA, 50 mM NaCl) for 10 min at 4°C. Samples were analyzed by electrophoresis in 4% (w/v) polyacrylamide gels (acrylamide/bis-acrylamide 80:1, w/w) containing 12 mM Tris-acetate, pH 7.5 and 1 mM EDTA, and run at 4°C in the same buffer. Protein–DNA complexes were detected by autoradiography as a shift in the migrating position of the labeled DNA. Quantification of protein–DNA complexes was performed using ImageJ software and plots were obtained using the ggplot package for R software (Wickham, 2009).

Multimerization and Cooperativity Assays

Analysis of oligomerization state by analytical ultracentrifugation was carried out by glycerol gradients (15–30%) with 50 mM Tris-HCl, 1 mM EDTA, 7 mM 2-mercaptoethanol and 200 mM NaCl were formed in 5-ml Beckman polyallomer centrifuge tubes (13 × 51 mm). On top of the glycerol gradient, 200 μl of a solution containing 8 μg of B35SSB, and 5 μg of BSA (Sigma-Aldrich) and Φ29TP as molecular weight markers, were loaded. Samples were centrifuged at 58,000 rpm in a Beckman TST 60.4 rotor for 24 h at 4°C. Gradients were fractionated from the bottom of the tube and analyzed by SDS– PAGE followed by protein staining with Coomassie Brilliant Blue.

Glutaraldehyde cross-linking assays of B35SSB in presence of different ssDNA substrates (**Supplementary Table 1**) were carried out as described (Jaenicke and Rudolph, 1986). Briefly, B35SSB was incubated with the indicated ssDNA oligonucleotides for 10 min at room temperature (RT) in 100 μl final volume of cross-linking buffer [50 mM sodium phosphate buffer pH 7.5, 50 mM NaCl, 0.05% (v/v) Tween 20]. Although different substrates and protein:DNA ratio were tested, the standard assay contained 2 μg of B35SSB (1 μM) incubated with 165 ng of the indicated oligonucleotide (0.1 μM for a 50-mer substrate).

Cross-links were formed by the addition of 1% (v/v) glutaraldehyde and incubation for 4 min at 20°C. The reaction was quenched by adding 5 μl of 2 M $NaBH_4$ (freshly prepared in 0.1 M NaOH) for 20 min at 20°C. Total protein was precipitated by adding 0.03% (v/v) sodium deoxycholate and 3% (v/v) TCA. Samples were incubated in ice for 15 min and centrifugated for 10 min at 14,000 × *g* at 4°C. The supernatant was discarded, and the pellet was resuspended in 20 μl of 1X Laemmli SDS-PAGE sample buffer [4% (v/v) 2-mercaptoethanol, 1.8% (w/v) SDS, 12% (v/v) glycerol, 120 mM Tris HCl pH 6.8, 0.005% (w/v) bromophenol blue] and heated to 95°C for 5 min. Samples were analyzed by electrophoresis in 8–16% polyacrylamide (w/v) gradient SDS gels followed by staining with Coomassie Brilliant Blue.

Helix-Destabilizing Assay

Unwinding assays substrate was obtained by hybridization of the genomic M13mp18 single-stranded circular DNA with the labeled primed M13 universal primer (M13UP: 5′ CAGGAAACAGCTATGAC 3′) in hybridization buffer (0.2 M NaCl, 60 mM Tris-HCl). M13UP was labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase (Sambrook and Russell, 2001). The assay was carried out in 12.5 μl of unwinding buffer [50 mM Tris-HCl pH 7.5, 1 mM DTT, 4% (v/v) glycerol, 0.1 mg/ml BSA, 0.05% (v/v) Tween 20, 1 mM EDTA, 50 mM NaCl] with 4 nM labeled primed M13 ssDNA and the indicated concentration of B35SSB or Φ29SSB. Reactions were incubated for 1 h at 37°C and stopped with 1.25 μl of 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 30% (v/v) glycerol and 0.5% (w/v) SDS. Samples were analyzed by electrophoresis in 0.1% (w/v) SDS 8-16% (w/v) gradient polyacrylamide gels. Gels were dried and the helix-destabilizing activity was detected by autoradiography as a change in the mobility of the labeled primer due to its displacement from the M13 DNA molecule.

Protein-Primed Initiation and Elongation Assays

Reactions of TP-initiation and elongation were carried out, as described in Berjón-Otero et al. (2016), in a final volume of 25 μl of reaction buffer [50 mM Tris-HCl pH 8.5, 1 mM DTT, 4% (v/v) glycerol, 0.1 mg/ml BSA, 0.05% (v/v) Tween 20] with 5 μM of dTTP (initiation assays) or all 4 dNTPs (elongation assays) supplemented with 2 μCi of [α-^{32}P]dTTP, a preincubated mix (10 min) with 2.5 nM of B35DNAP and 34 nM of B35TP, 10 mM $MgCl_2$, 315 nM of the 29 or 60-mer single-stranded oligonucleotide with Bam35 genome left end sequence (**Supplementary Table 1**) as the template and the corresponding concentration of SSB. The reactions also contained a basal concentration (27.6 mM NaCl) contributed by the reaction components.

Processive Replication of Singly Primed M13 DNA

The primed M13 ssDNA was obtained by hybridization of the genomic M13mp18 single-stranded circular DNA with the M13UP primer in hybridization buffer (0.2 M NaCl, 60 mM Tris-HCl). Reactions of replication were carried out in a final volume of 25 μl of reaction buffer [50 mM Tris-HCl, 1 mM DTT, 4% (v/v) glycerol, 0.1 mg/ml BSA, 0.5% Tween 20] with 50 nM B35DNAP, 1.4 nM primed M13 ssDNA, 40 μM dNTPs, 10 mM MgCl2 and 0.5 μCi [α-^{32}P]dATP, and the indicated concentrations of SSB. The reactions contained 16 mM NaCl (final concentration), contributed by the reaction components. To increase the efficiency of replication, the polymerase and the template were incubated for 10 min at RT prior to the addition of the SSB. Reactions were incubated at 37°C for the indicated times and stopped by adding 30 mM EDTA and 0.5% (v/v) SDS. Non-incorporated radiolabeled dATP was removed by Sephadex G-50 spin filtration. The lambda DNA ladder was labeled by filling in with Klenow fragment (New England Biolabs) in the presence of [α-^{32}P]dATP. Results of replication were analyzed by electrophoresis in alkaline 0.7% agarose gels, and the gels were subsequently dried and autoradiographed.

B35SSB Sequence Analysis and Phylogeny

Bam35 ORF2 product sequence (B35SSB, UniProt database accession Q6 × 3W7) was used as the protein query to find related proteins using the Jackhmmer tool from the HMMER web server (Potter et al., 2018). The analysis was performed using Uniref90 as database and 10–5 as *E*-value. Search converged after 5 iterations and 112 significant query matches were obtained. The retrieved protein sequences were used to obtain a multiple sequence alignment (MSA) using MAFFT (parameters: –maxiterate 1000 –retree 1) (Katoh and Standley, 2013). The MSA was then trimmed with trimAl (parameter -gappyout) (Capella-Gutierrez et al., 2009) and used to generate a best-fit maximum-likelihood phylogeny using IQTREE2 version 2.0.7 (parameters -B 5000 -alrt 5000 -redo) (Kalyaanamoorthy et al., 2017; Minh et al., 2020). According to the Bayesian information criterion, the selected best-fit model was LG + F + R5. Lastly, the tree was visualized with the ggtree package for R software (Yu, 2020). Secondary structure prediction of the MSA was obtained with the JPred4 server through the JalView web server version 2.11 (Drozdetskiy et al., 2015; Procter et al., 2021).

To assess the possible relationship of the Bam35-Φ29 SSBs group with previously known OB-fold containing SSBs, we used an alignment-independent clustering approach. First, a comprehensive dataset aiming to span all known diversity was generated by merging previously used sets of sequences (Szczepankowska et al., 2011; Kazlauskas and Venclovas, 2012), updated with our own searches with TopSearch (Wiederstein et al., 2014)[1]. The final dataset contained 2,750 non-redundant sequences. Then, HMM profiles for each sequence were generated and the profiles were compared all-against-all with HHsearch to cluster them with CLANS (Frickey and Lupas, 2004; Soding et al., 2005).

The B35SSB structural models were obtained with five different servers (January 2021), based either in fold-recognition or distant-based approaches, i-Tasser (Roy et al., 2010), Phyre2 (Kelley et al., 2015), RaptorX (Källberg et al., 2012), trRossetta (Yang et al., 2020), and RosettaCM (Kim et al., 2004; Song et al., 2013). Besides the evaluation provided by each method, all models were assessed by VoroMQA and ModFold8 protein structure quality assessment independent methods (Maghrabi and McGuffin, 2017; Olechnovič and Venclovas, 2017). Models were then visualized with Open-Source Pymol (Schrödinger, 2021).

Structural searches in PDB database were performed using Dali Protein Structural Comparison Server (Holm, 2020).

DATA AVAILABILITY STATEMENT

Uniref90 database was downloaded from Uniprot site (https://www.uniprot.org/downloads) and T4 SSB structure was retrieved

[1]https://topsearch.services.came.sbg.ac.at

from Protein Data Bank (https://www.rcsb.org/) under the accession number 1GPC.

AUTHOR CONTRIBUTIONS

MS and MR-R: conceptualization, funding acquisition, and supervision. AL, DK, and MR-R: formal analysis, investigation, and writing – review and edit. AL and MR-R: writing – original draft preparation. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by grants from the Spanish Ministry of Science, Innovation and Universities (PGC2018-093726-B-100 AEI/FEDER UE to MS and PGC2018-093723-A-100 to MR-R) and Fundación Ramón Areces (VirHostOmics). AL was holder of a Ph.D. fellowship (FPU15/05797) from the Spanish Ministry of Science, Innovation and Universities. Fundación Ramón Areces Grant VirHostOmics to MS and MR-R.

ACKNOWLEDGMENTS

An institutional grant from Fundación Ramón Areces and Banco Santander to the Centro de Biología Molecular Severo Ochoa is acknowledged. We are also grateful to all the members of the MR-R lab for discussions and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.699140/full#supplementary-material

REFERENCES

Ackermann, H. W., Roy, R., Martin, M., Murthy, M. R., and Smirnoff, W. A. (1978). Partial characterization of a cubic *Bacillus* phage. *Can. J. Microbiol.* 24, 986–993.

Alcaraz, L. D., Moreno-Hagelsieb, G., Eguiarte, L. E., Souza, V., Herrera-Estrella, L., and Olmedo, G. (2010). Understanding the evolutionary relationships and major traits of Bacillus through comparative genomics. *BMC Genomics* 11:332. doi: 10.1186/1471-2164-11-332

Antony, E., and Lohman, T. M. (2019). Dynamics of *E. coli* single stranded DNA binding (SSB) protein-DNA complexes. *Semin. Cell Dev. Biol.* 86, 102–111. doi: 10.1016/j.semcdb.2018.03.017

Berjón-Otero, M., Lechuga, A., Mehla, J., Uetz, P., Salas, M., and Redrejo-Rodríguez, M. (2017). Bam35 tectivirus intraviral interaction map unveils new function and localization of phage ORFan proteins. *J. Virol.* 91:e00870-17. doi: 10.1128/JVI.00870-17

Berjón-Otero, M., Villar, L., de Vega, M., Salas, M., and Redrejo-Rodríguez, M. (2015). DNA polymerase from temperate phage Bam35 is endowed with processive polymerization and abasic sites translesion synthesis capacity. *Proc. Natl. Acad. Sci. U.S.A.* 112, E3476–E3484. doi: 10.1073/pnas.1510280112

Berjón-Otero, M., Villar, L., Salas, M., and Redrejo-Rodríguez, M. (2016). Disclosing early steps of protein-primed genome replication of the Gram-positive tectivirus Bam35. *Nucleic Acids Res.* 44, 9733–9744. doi: 10.1093/nar/gkw673

Blanco, L., Lazaro, J. M., de Vega, M., Bonnin, A., and Salas, M. (1994). Terminal protein-primed DNA amplification. *Proc. Natl. Acad. Sci. U.S.A.* 91, 12198–12202.

Boon, M., De Zitter, E., De Smet, J., Wagemans, J., Voet, M., Pennemann, F. L., et al. (2020). "Drc", a structurally novel ssDNA-binding transcription regulator of N4-related bacterial viruses. *Nucleic Acids Res.* 48, 445–459. doi: 10.1093/nar/gkz1048

Byrne, B. M., and Oakley, G. G. (2019). Replication protein A, the laxative that keeps DNA regular: the importance of RPA phosphorylation in maintaining genome stability. *Semin. Cell Dev. Biol.* 86, 112–120. doi: 10.1016/j.semcdb.2018.04.005

Caldentey, J., Blanco, L., Savilahti, H., Bamford, D. H., and Salas, M. (1992). In vitro replication of bacteriophage PRD1 DNA. Metal activation of protein-primed initiation and DNA elongation. *Nucleic Acids Res.* 20, 3971–3976.

Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Caruso, S. M., deCarvalho, T. N., Huynh, A., Morcos, G., Kuo, N., Parsa, S., et al. (2019). A novel genus of actinobacterial tectiviridae. *Viruses* 11:1134. doi: 10.3390/v11121134

Casas-Finet, J. R., Fischer, K. R., and Karpel, R. L. (1992). Structural basis for the nucleic acid binding cooperativity of bacteriophage T4 gene 32 protein: the (Lys/Arg)3(Ser/Thr)2 (LAST) motif. *Proc. Natl. Acad. Sci. U.S.A.* 89, 1050–1054. doi: 10.1073/pnas.89.3.1050

Cernooka, E., Rumnieks, J., Tars, K., and Kazaks, A. (2017). Structural Basis for DNA recognition of a single-stranded DNA-binding protein from *Enterobacter* phage Enc34. *Sci. Rep.* 7:15529. doi: 10.1038/s41598-017-15774-y

Cerrón, F., de Lorenzo, S., Lemishko, K. M., Ciesielski, G. L., Kaguni, L. S., Cao, F. J., et al. (2019). Replicative DNA polymerases promote active displacement of SSB proteins during lagging strand synthesis. *Nucleic Acids Res.* 47, 5723–5734. doi: 10.1093/nar/gkz249

Chisty, L. T., Quaglia, D., and Webb, M. R. (2018). Fluorescent single-stranded DNA-binding protein from *Plasmodium falciparum* as a biosensor for single-stranded DNA. *PLoS One* 13:e0193272. doi: 10.1371/journal.pone.0193272

de la Torre, I., Quiñones, V., Salas, M., and Del Prado, A. (2019). Tyrosines involved in the activity of phi29 single-stranded DNA binding protein. *PLoS One* 14:e0217248. doi: 10.1371/journal.pone.0217248

Dickey, T. H., Altschuler, S. E., and Wuttke, D. S. (2013). Single-stranded DNA-binding proteins: multiple domains for multiple functions. *Structure* 21, 1074–1084. doi: 10.1016/j.str.2013.05.013

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* 43, W389–W394. doi: 10.1093/nar/gkv332

Dubiel, K., Myers, A. R., Kozlov, A. G., Yang, O., Zhang, J., Ha, T., et al. (2019). Structural mechanisms of cooperative DNA binding by bacterial single-stranded DNA-binding proteins. *J. Mol. Biol.* 431, 178–195. doi: 10.1016/j.jmb.2018.11.019

Frickey, T., and Lupas, A. (2004). CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* 20, 3702–3704. doi: 10.1093/bioinformatics/bth444

Gamsjaeger, R., Kariawasam, R., Gimenez, A. X., Touma, C., McIlwain, E., Bernardo, R. E., et al. (2015). The structural basis of DNA binding by the single-stranded DNA-binding protein from *Sulfolobus solfataricus*. *Biochem. J.* 465, 337–346. doi: 10.1042/BJ20141140

Gascón, I., Carrascosa, J. L., Villar, L., Lázaro, J. M., and Salas, M. (2002). Importance of the N-terminal region of the phage GA-1 single-stranded DNA-binding protein for its self-interaction ability and functionality. *J. Biol. Chem.* 277, 22534–22540. doi: 10.1074/jbc.M202430200

Gascón, I., Gutiérrez, C., and Salas, M. (2000a). Structural and functional comparative study of the complexes formed by viral Φ29, Nf and GA-1 SSB proteins with DNA. *J. Mol. Biol.* 296, 989–999. doi: 10.1006/jmbi.2000.3521

Gascón, I., Lázaro, J. M., and Salas, M. (2000b). Differential functional behavior of viral Φ29, Nf and GA-1 SSB proteins. *Nucleic Acids Res.* 28, 2034–2042.

Gillis, A., Fayad, N., Makart, L., Bolotin, A., Sorokin, A., Kallassy, M., et al. (2018). Role of plasmid plasticity and mobile genetic elements in the entomopathogen *Bacillus thuringiensis* serovar israelensis. *FEMS Microbiol. Rev.* 42, 829–856. doi: 10.1093/femsre/fuy034

Hernandez, A. J., and Richardson, C. C. (2019). Gp2.5, the multifunctional bacteriophage T7 single-stranded DNA binding protein. *Semin. Cell Dev. Biol.* 86, 92–101. doi: 10.1016/j.semcdb.2018.03.018

Holm, L. (2020). DALI and the persistence of protein shape. *Protein Sci.* 29, 128–140. doi: 10.1002/pro.3749

Hyland, E. M., Rezende, L. F., and Richardson, C. C. (2003). The DNA binding domain of the gene 2.5 single-stranded DNA-binding protein of bacteriophage T7. *J. Biol. Chem.* 278, 7247–7256. doi: 10.1074/jbc.M210605200

Jaenicke, R., and Rudolph, R. (1986). Refolding and association of oligomeric proteins. *Methods Enzymol.* 131, 218–250. doi: 10.1016/0076-6879(86)31043-7

Jalasvuori, M., Palmu, S., Gillis, A., Kokko, H., Mahillon, J., Bamford, J. K., et al. (2013). Identification of five novel tectiviruses in *Bacillus* strains: analysis of a highly variable region generating genetic diversity. *Res. Microbiol.* 164, 118–126. doi: 10.1016/j.resmic.2012.10.011

Jose, D., Weitzel, S. E., Baase, W. A., and von Hippel, P. H. (2015). Mapping the interactions of the single-stranded DNA binding protein of bacteriophage T4 (gp32) with DNA lattices at single nucleotide resolution: gp32 monomer binding. *Nucleic Acids Res.* 43, 9276–9290. doi: 10.1093/nar/gkv817

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., et al. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7, 1511–1522. doi: 10.1038/nprot.2012.085

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kazlauskas, D., Krupovic, M., and Venclovas, C. (2016). The logic of DNA replication in double-stranded DNA viruses: insights from global analysis of viral genomes. *Nucleic Acids Res.* 44, 4551–4564. doi: 10.1093/nar/gkw322

Kazlauskas, D., and Venclovas, C. (2012). Two distinct SSB protein families in nucleo-cytoplasmic large DNA viruses. *Bioinformatics* 28, 3186–3190. doi: 10.1093/bioinformatics/bts626

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kim, D. E., Chivian, D., and Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 32, W526–W531. doi: 10.1093/nar/gkh468

Kim, Y. T., Tabor, S., Bortner, C., Griffith, J. D., and Richardson, C. C. (1992). Purification and characterization of the bacteriophage T7 gene 2.5 protein. A single-stranded DNA-binding protein. *J. Biol. Chem.* 267, 15022–15031.

Koonin, E. V., Krupovic, M., and Yutin, N. (2015). Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. *Ann. N.Y. Acad. Sci.* 1341, 10–24. doi: 10.1111/nyas.12728

Kozlov, A. G., and Lohman, T. M. (1999). Adenine base unstacking dominates the observed enthalpy and heat capacity changes for the *Escherichia coli* SSB tetramer binding to single-stranded oligoadenylates. *Biochemistry* 38, 7388–7397. doi: 10.1021/bi990309z

Krupovic, M., and Koonin, E. V. (2015). Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. *Nat. Rev. Microbiol.* 13, 105–115. doi: 10.1038/nrmicro3389

Kumaran, S., Kozlov, A. G., and Lohman, T. M. (2006). Saccharomyces cerevisiae replication protein A binds to single-stranded DNA in multiple salt-dependent modes. *Biochemistry* 45, 11958–11973. doi: 10.1021/bi060994r

Kur, J., Olszewski, M., Dlugolecka, A., and Filipkowski, P. (2005). Single-stranded DNA-binding proteins (SSBs) – sources and applications in molecular biology. *Acta Biochim. Pol.* 52, 569–574.

Lázaro, J. M., Blanco, L., and Salas, M. (1995). Purification of bacteriophage Φ29 DNA polymerase. *Methods Enzymol.* 262, 42–49.

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. doi: 10.1146/annurev.bi.63.070194.002523

Lohman, T. M., Overman, L. B., and Datta, S. (1986). Salt-dependent changes in the DNA binding co-operativity of *Escherichia coli* single strand binding protein. *J. Mol. Biol.* 187, 603–615. doi: 10.1016/0022-2836(86)90338-4

Maghrabi, A. H. A., and McGuffin, L. J. (2017). ModFOLD6: an accurate web server for the global and local quality estimation of 3D protein models. *Nucleic Acids Res.* 45, W416–W421. doi: 10.1093/nar/gkx332

Martin, G., and Salas, M. (1988). Characterization and cloning of gene 5 of *Bacillus subtilis* phage phi 29. *Gene* 67, 193–201.

Matsumoto, K., and Ishimi, Y. (1994). Single-stranded-DNA-binding protein-dependent DNA unwinding of the yeast ARS1 region. *Mol. Cell. Biol.* 14, 4624–4632. doi: 10.1128/MCB.14.7.4624

Mencía, M., Gella, P., Camacho, A., de Vega, M., and Salas, M. (2011). Terminal protein-primed amplification of heterologous DNA with a minimal replication system based on phage Φ29. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18655–18660. doi: 10.1073/pnas.1114397108

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Murzin, A. G. O. B. (1993). (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* 12, 861–867.

Naue, N., Beerbaum, M., Bogutzki, A., Schmieder, P., and Curth, U. (2013). The helicase-binding domain of *Escherichia coli* DnaG primase interacts with the highly conserved C-terminal region of single-stranded DNA-binding protein. *Nucleic Acids Res.* 41, 4507–4517. doi: 10.1093/nar/gkt107

Olechnovič, K., and Venclovas, Č (2017). VoroMQA: assessment of protein structure quality using interatomic contact areas. *Proteins* 85, 1131–1145. doi: 10.1002/prot.25278

Pakula, T. M., Caldentey, J., Gutiérrez, C., Olkkonen, V. M., Salas, M., and Bamford, D. H. (1993). Overproduction, purification, and characterization of DNA-binding protein P19 of bacteriophage PRD1. *Gene* 126, 99–104.

Pakula, T. M., Caldentey, J., Serrano, M., Gutiérrez, C., Hermoso, J. M., Salas, M., et al. (1990). Characterization of a DNA binding protein of bacteriophage PRD1 involved in DNA replication. *Nucleic Acids Res.* 18, 6553–6557.

Pant, K., Anderson, B., Perdana, H., Malinowski, M. A., Win, A. T., Pabst, C., et al. (2018). The role of the C-domain of bacteriophage T4 gene 32 protein in ssDNA binding and dsDNA helix-destabilization: kinetic, single-molecule, and cross-linking studies. *PLoS One* 13:e0194357. doi: 10.1371/journal.pone.0194357

Pant, K., Karpel, R. L., Rouzina, I., and Williams, M. C. (2004). Mechanical measurement of single-molecule binding rates: kinetics of DNA Helix-destabilization by T4 gene 32 protein. *J. Mol. Biol.* 336, 851–870. doi: 10.1016/j.jmb.2003.12.025

Paytubi, S., McMahon, S. A., Graham, S., Liu, H., Botting, C. H., Makarova, K. S., et al. (2012). Displacement of the canonical single-stranded DNA-binding protein in the thermoproteales. *Proc. Natl. Acad. Sci. U.S.A.* 109, E398–E405. doi: 10.1073/pnas.1113277108

Perales, C., Cava, F., Meijer, W. J., and Berenguer, J. (2003). Enhancement of DNA, cDNA synthesis and fidelity at high temperatures by a dimeric single-stranded DNA-binding protein. *Nucleic Acids Res.* 31, 6473–6480.

Pestryakov, P. E., and Lavrik, O. I. (2008). Mechanisms of single-stranded DNA-binding protein functioning in cellular DNA metabolism. *Biochem. Mosc.* 73, 1388–1404. doi: 10.1134/s0006297908130026

Philippe, C., Krupovic, M., Jaomanjaka, F., Claisse, O., Petrel, M., and le Marrec, C. (2018). Bacteriophage GC1, a novel tectivirus infecting gluconobacter cerinus, an acetic acid bacterium associated with wine-making. *Viruses* 10:39. doi: 10.3390/v10010039

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204. doi: 10.1093/nar/gky448

Procter, J. B., Carstairs, G. M., Soares, B., Mourão, K., Ofoegbu, T. C., Barton, D., et al. (2021). Alignment of biological sequences with jalview. *Methods Mol. Biol.* 2231, 203–224. doi: 10.1007/978-1-0716-1036-7_13

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-A resolution. *Proc. Natl. Acad. Sci. U.S.A.* 94, 6652–6657. doi: 10.1073/pnas.94.13.6652

Ravantti, J. J., Gaidelyte, A., Bamford, D. H., and Bamford, J. K. (2003). Comparative analysis of bacterial viruses Bam35, infecting a gram-positive host, and PRD1, infecting gram-negative hosts, demonstrates a viral lineage. *Virology* 313, 401–414.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5

Salas, M. (1991). Protein-priming of DNA replication. *Annu. Rev. Biochem.* 60, 39–71. doi: 10.1146/annurev.bi.60.070191.000351

Salas, M., and de Vega, M. (2016). "Protein-primed replication of bacteriophage Φ29 DNA," in *The Enzymes*, eds S. K. Laurie and O. Marcos Túlio (Cambridge, MA: Academic Press), 137–167.

Salas, M., Holguera, I., Redrejo-Rodriguez, M., and de Vega, M. (2016). DNA-binding proteins essential for protein-primed bacteriophage Φ29 DNA replication. *Front. Mol. Biosci.* 3:37. doi: 10.3389/fmolb.2016.00037

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 4th Edn. New York, NY: Cold Spring Harbor Laboratory Press.

Savilahti, H., Caldentey, J., Lundstrom, K., Syvaoja, J. E., and Bamford, D. H. (1991). Overexpression, purification, and characterization of *Escherichia coli* bacteriophage PRD1 DNA polymerase. In vitro synthesis of full-length PRD1 DNA with purified proteins. *J. Biol. Chem.* 266, 18737–18744.

Schrödinger, L. L. C. (2021). *The PyMOL Molecular Graphics System, Version 2.4.*

Shamoo, Y., Friedman, A. M., Parsons, M. R., Konigsberg, W. H., and Steitz, T. A. (1995). Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature* 376, 362–366. doi: 10.1038/376362a0

Shamoo, Y., Ghosaini, L. R., Keating, K. M., Williams, K. R., Sturtevant, J. M., and Konigsberg, W. H. (1989). Site-specific mutagenesis of T4 gene 32: the role of tyrosine residues in protein-nucleic acid interactions. *Biochemistry* 28, 7409–7417. doi: 10.1021/bi00444a039

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. doi: 10.1080/10409230802341296

Shi, H., Zhang, Y., Zhang, G., Guo, J., Zhang, X., Song, H., et al. (2013). Systematic functional comparative analysis of four single-stranded DNA-binding proteins and their affection on viral RNA metabolism. *PLoS One* 8:e55076. doi: 10.1371/journal.pone.0055076

Soding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–W248. doi: 10.1093/nar/gki408

Soengas, M. S. (1996). *Caracterización Estructural y Funcional de la Proteína de Unión a DNA de Cadena Sencilla Del Bacteriófago Φ29*. Ph.D. Thesis. Madrid: Universidad Autónoma de Madrid.

Soengas, M. S., Esteban, J. A., Salas, M., and Gutiérrez, C. (1994). Complex formation between phage Φ29 single-stranded DNA binding protein and DNA. *J. Mol. Biol.* 239, 213–226.

Soengas, M. S., Gutiérrez, C., and Salas, M. (1995). Helix-destabilizing activity of Φ29 single-stranded DNA binding protein: effect on the elongation rate during strand displacement DNA replication. *J. Mol. Biol.* 253, 517–529. doi: 10.1006/jmbi.1995.0570

Song, Y., DiMaio, F., Wang, R. Y.-R., Kim, D., Miles, C., Brunette, T., et al. (2013). High-resolution comparative modeling with RosettaCM. *Structure* 21, 1735–1742. doi: 10.1016/j.str.2013.08.005

Stabell, F. B., Egge-Jacobsen, W., Risoen, P. A., Kolsto, A. B., and Okstad, O. A. (2009). ORF 2 from the *Bacillus cereus* linear plasmid pBClin15 encodes a DNA binding protein. *Lett. Appl. Microbiol.* 48, 51–57. doi: 10.1111/j.1472-765X.2008.02483.x

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Exp. Purif.* 41, 207–234.

Stuiver, M. H., and van der Vliet, P. C. (1990). Adenovirus DNA-binding protein forms a multimeric protein complex with double-stranded DNA and enhances binding of nuclear factor I. *J. Virol.* 64, 379–386.

Szczepankowska, A. K., Prestel, E., Mariadassou, M., Bardowski, J. K., and Bidnenko, E. (2011). Phylogenetic and complementation analysis of a single-stranded DNA binding protein family from lactococcal phages indicates a non-bacterial origin. *PLoS One* 6:e26942. doi: 10.1371/journal.pone.0026942

Tone, T., Takeuchi, A., and Makino, O. (2012). Functional linkages between replication proteins of genes 1, 3 and 5 of Bacillus subtilis phage Φ29. *Genes Genet. Syst.* 87, 347–356.

Touma, C., Kariawasam, R., Gimenez, A. X., Bernardo, R. E., Ashton, N. W., Adams, M. N., et al. (2016). A structural analysis of DNA binding by hSSB1 (NABP2/OBFC2B) in solution. *Nucleic Acids Res.* 44, 7963–7973. doi: 10.1093/nar/gkw617

Tucker, P. A., Tsernoglou, D., Tucker, A. D., Coenjaerts, F. E., Leenders, H., and van der Vliet, P. C. (1994). Crystal structure of the adenovirus DNA binding protein reveals a hook-on model for cooperative DNA binding. *EMBO J.* 13, 2994–3002.

Wadsworth, R. I., and White, M. F. (2001). Identification and properties of the crenarchaeal single-stranded DNA binding protein from Sulfolobus solfataricus. *Nucleic Acids Res.* 29, 914–920. doi: 10.1093/nar/29.4.914

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.

Wiederstein, M., Gruber, M., Frank, K., Melo, F., and Sippl, M. J. (2014). Structure-based characterization of multiprotein complexes. *Structure* 22, 1063–1070. doi: 10.1016/j.str.2014.05.005

Witte, G., Fedorov, R., and Curth, U. (2008). Biophysical analysis of *Thermus aquaticus* single-stranded DNA binding protein. *Biophys. J.* 94, 2269–2279. doi: 10.1529/biophysj.107.121533

Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., and Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. *PNAS* 117, 1496–1503. doi: 10.1073/pnas.1914677117

Yu, G. (2020). Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinform.* 69:e96. doi: 10.1002/cpbi.96

Yu, G., Lam, T. T.-Y., Zhu, H., and Guan, Y. (2018). Two methods for mapping and visualizing associated data on phylogeny using Ggtree. *Mol. Biol. Evol.* 35, 3041–3043. doi: 10.1093/molbev/msy194

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Lechuga, Kazlauskas, Salas and Redrejo-Rodríguez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.