Structural Analysis of Bacillus subtilis SPP1 Phage Helicase Loader Protein G39P*

Received for publication, September 11, 2002, and in revised form, January 21, 2003
Published, JBC Papers in Press, February 13, 2003, DOI 10.1074/jbc.M209300200

Scott Bailey‡§, Svetlana E. Sedelnikova‡, Pablo Mesa¶, Sylvia Ayora¶, Jon P. Waltho††††, Alison E. Ashcroft‡‡, Andrew J. Baron‡‡, Juan C. Alonso¶¶, and John B. Rafferty†††††

From the ‡Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom; *Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CSIC, Campus Universitario Autónomo de Madrid, Cantoblanco, 28049 Madrid, Spain; ¶¶Astbury Centre for Structural Molecular Biology, School of Biology, University of Leeds, Leeds LS2 9JT, United Kingdom

The Bacillus subtilis SPP1 phage-encoded protein G39P is a loader and inhibitor of the phage G40P replicative helicase involved in the initiation of DNA replication. We have carried out a full x-ray crystallographic and preliminary NMR analysis of G39P and functional studies of the protein, including assays for helicase binding by a number of truncated mutant forms, in an effort to improve our understanding of how it interacts with the helicase and with the phage replisome organizer, G39P. Our structural analyses reveal that G39P has a completely unexpected bipartite structure comprising a folded N-terminal domain and an essentially unfolded C-terminal domain. Although G39P has been shown to bind its G40P target with a 6:6 stoichiometry, our crystal structure and other biophysical characterization data reveal that the protein probably exists predominantly as a monomer in solution. The G39P protein is proteolytically sensitive, and our binding assays show that the C-terminal domain is essential for helicase interaction and that removal of just the 14 C-terminal residues abolishes interaction with the helicase in vitro. We propose a number of possible scenarios in which the flexibility of the C-terminal domain of G39P and its proteolytic sensitivity may have important roles for the function of G39P in vivo that are consistent with other data on SPP1 phage DNA replication.

The initiation of DNA replication is a key step in the life cycle of all cells and as such its careful and precise control is essential. Studies of prokaryotic systems centered mainly on Escherichia coli, and its extrachromosomal elements have identified the following three key stages involved in DNA replication initiation: first, the recognition of the DNA replication origin and initial melting of the DNA strands; second, the recruitment of the replication machinery to the origin; and third, the remodeling of a replication complex to trigger the transition from a stable origin-bound complex to a mobile replication machine (1).

Initiation of DNA replication in E. coli proceeds via a sequence of events involving a replisome-DNA replication on a Gram-positive bacterial environment is that of Bacillus subtilis, the most well characterized example of initiation of DNA replication. The initiation of SPP1 replication requires the phage-encoded products of genes 38, 39, and 40 (G39P, G39P, and G40P) in addition to the host DNA polymerase III and DnaG primase (4). G38P (a monomer with a predicted molecular mass of 29,997 Da) acts as a close functional equivalent to DnaA (although the proteins share no sequence similarity) and is the replisome organizer of the SPP1 system. The G38P protein specifically interacts with its cognate site present in multiple copies at the phage replication origin (oriC). This interaction occurs in the absence of ATP and is thought to induce the local unwinding of the adjacent A + T-rich region present within oriC to initiate θ-type DNA replication (5–7). G40P is a DnaB-like helicase and as such is a ring-shaped hexamer, capable of unwinding duplex DNA with a 5′ to 3′ polarity in a reaction fueled by nucleotide 5′-triphosphate hydrolysis (6, 8). G39P is predominantly a monomer (molecular mass 14,610 Da) when free in solution and forms a specific interaction with ATP-activated G40P that inactivates the ssDNA binding, ATPase, and unwinding activities of the helicase. Targeting of G40P by G39P to the G38P-bound oriL then functions to activate G40P upon delivery (6). It is believed that G40P, in the form of the G39P-G40P-ATP complex, is delivered to G38P-bound oriL via the specific protein-protein interaction of the helicase-bound G39P with the G38P bound at oriL. These interactions result in the formation of an unstable nucleoprotein oriL-G38P-G39P-G40P-ATP intermediate, with subsequent release of G38P/G39P heterodimers that leaves the ATP-activated G40P complex to bind the melted DNA.

* This work was supported in part by European Commission Grants B104-CT98-0106 and QLK2-CT-2000-00634 and an EMBL grant under the “Human Capital and Mobility” Programme. The Krebs Institute is a designated Biotechnology and Biological Sciences Research Council Biomolecular Sciences Centre and a member of the North of England Structural Biology Centre. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1NO1) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Present address: Dept. of Molecular Biophysics and Biochemistry, Yale University, Bass Center, Rm. 415, 266 Whitney Ave., New Haven, CT 06520-8114.

† Royal Society Olga Kennard Fellow. To whom correspondence should be addressed. Tel.: 44-114-2222809; Fax: 44-114-2726697; E-mail: j.rafferty@sheffield.ac.uk.

The abbreviations used are: G38P, gene X product; a.u., asymmetric unit; ss, single-stranded; wt, wild type; BstT1A, 2-[bis(2-hydroxyethyl)]-amino]-2-hydroxymethylpropene-1,3-diol; Ni-NTA, nickel-nitritotriacetic acid; ATP, adenosine 5′-O-(thiotriphosphate).
Structure of SPP1 Phage G39P Helicase Loader

The values in parentheses refer to the outer shell.

| MAD phasing | Inflection | Peak | Remote |
|-------------|------------|------|--------|
| Wavelength (Å) | 0.9795 | 0.9793 | 0.9717 |
| Resolution (Å) | 20–2.4 | 20–2.4 | 20–2.4 |
| Outer shell Å | (2.46–2.40) | (2.46–2.40) | (2.46–2.40) |
| No. observations | 161,039 | 162,010 | 159,504 |
| Unique reflections | 28,352 | 28,373 | 28,331 |
| Data completeness (%) | 93 (81) | 97.6 (93.9) | 93 (81) |
| Rmerge (%) | 77.5 (49.5) | 76.6 (49.7) | 75.7 (49.9) |
| Phasing powera | 5.4 (26.2) | 5.4 (26.3) | 4.4 (30.2) |

a Phasing power = <Fobs/Fcalc> – 1, where Fcalc is the calculated heavy atom structure factor amplitude.

b Rmerge = <Iobs/Icalc> – 1. Phasing power and Rmerge statistics are given for phase set used in calculation of first map at 3-Å resolution.

c FoM = mean figure of merit.

d Defined in PROCHECK (40).

e After translation, librations, and screw tensor refinement.

EXPERIMENTAL PROCEDURES

Crystal Structure Determination—The details of wt G39P and G39P112 variant constructs, their overexpression, purification, and crystallization are described elsewhere. Following purification, the wt protein was crystallized from ammonium phosphate, and preliminary x-ray diffraction data were collected at room temperature on home laboratory x-ray sources. These data suggested that the crystals belong to space group P61212 (or P6522) with cell dimensions a = b = 105.3 Å, c = 47.4 Å, and a = b = 65.3 Å, and diffractions to 3.5 Å. However, the reproducibility of these crystals was very poor, and an analysis by mass spectrometry of freshly purified protein and the crystals themselves revealed multiple fragments derived from proteolytic cleavage at the C terminus of the protein. The proteolysis terminated markedly after the final 14 residues of G39P had been removed, and thus gene 39 was engineered to produce a mutant construct that coded for a truncated protein comprising only the first 112 of 126 residues of G39P. The G39P112 protein and selenomethionine incorporated form were prepared in a similar manner to wild type protein but crystallized from ammonium sulfate in space group P212121 with cell dimensions a = 85.6 Å, b = 89.7 Å, c = 47.6 Å. The crystals have three monomers in the asymmetric unit and an assumed solvent content of 47% based on a V_M (Matthews coefficient) of 2.3 Å³ Da⁻¹. The SeMet G39P112 crystals were used in a subsequent multiwavelength anomalous dispersion experiment in which data were collected using a Mar Research 345 imaging plate scanner at the European Synchrotron Radiation Facility on station BM30 using inverse beam geometry to collect Friedel pairs. The data for each wavelength were processed individually and scaled in such a way as to preserve anomalous signal using the HKL Suite of programs (13). The data processing statistics are shown in Table I. The positions of six selenium atoms were found using the program SOLVE (14). These positions were then refined, and initial phases were calculated in the program MLPHARE (15) following the pseudo-MIR procedure (16). Phasing statistics are shown in Table I. An electron density map was calculated at 3.0 Å and subsequently improved by solvent flattening and histogram matching with the program DM (17). This map was of good quality with readily identifiable regions of secondary structure. Following a preliminary trace of the secondary structure, non-crystallographic symmetry operators were determined for the three monomers found in the asymmetric unit, and the map was averaged and phase-extended to 2.4 Å using DM. The model fitted to the resultant map was submitted to refinement using the program REFMAC (18). Iterative cycles of phase combination of the partial structure phases and those from the multiwavelength anoma-

TABLE I

Data collection, phasing, and refinement statistics

| Source | Data completeness (%) | Rmerge (%) | Phasing power |
|--------|-----------------------|------------|--------------|
| MAD phasing | 93 (81) | 77.5 (49.5) | 5.4 (26.2) |
| Defined in PROCHECK (40). |
| FoM | 0.74 | | |

2 S. Bailey, S. E. Sedelnikova, P. Mesa, S. Ayora, J. C. Alonso, and J. B. Rafferty, submitted for publication.
### Structure of SPP1 Phage G39P Helicase Loader

The G39P protein crystallized with three independent monomers in the a.u., but they adopted an arrangement around a non-crystallographic 6-fold axis parallel to the crystallographic c axis (Fig. 2). When viewed along the direction of the c axis, the disordered C-terminal domains are located on the exterior of the helical arrays of monomers formed by the screw axes. The cavities in the crystal lattice are clearly sufficient to accommodate the C-terminal domains of the G39P112 variant as supported by mass spectrometric analysis of the crystal used in the structure determination that confirmed the presence of intact variant (data not shown). However, it would seem that the extra bulk provided by the 14 C-terminal residues of the intact wt protein necessitates an alternative crystal packing arrangement that appears to be less stable as judged by the noticeably poorer crystal reproducibility and diffraction quality. The inter-monomer contacts are made predominantly between residues immediately following helices $\alpha$A and $\alpha$C and those preceding helices $\beta$B and D. The residues involved are both polar and hydropheric, and the helices include two completely buried water molecules. Pairwise superposition of the $\alpha$-carbon positions of each of the monomers gives root mean square deviation values of 0.2 to 0.3 Å. Approximately 1200 Å$^2$ or 30% of the surface area of each monomer is buried in the interfaces with other monomers in the non-crystallographic 6-fold helix and a further ~10% on average in contacts with other subunits in the crystal lattice. The missing polypeptide chain in the electron density map, extending from residue Lys-67, is directed toward a large cavity in the crystalline lattice (Fig. 2) where it adopts a flexible, mainly unfolded state (see NMR analysis below). Thus at least in the crystal lattice, the G39P112 protein appears to exist as a monomer.

Calculations of the electrostatic surface potential of the folded N-terminal domain of G39P reveals a somewhat negatively charged surface overall. However, there is a notable localized, highly negative patch on the surface formed by residues at the N terminus of helix $\alpha$A and the N terminus and loop preceding helix $\alpha$C that also lies adjacent to the last observed residue in the map, Lys-67. The distribution of charge is even more striking when one examines the helical packing of the

[1] Nielsen, P.D., Klenow, D.L., Rosewood, S., and Rossmann, M.G. (1972) J. Mol. Biol. 62, 155–166.

[2] Blobel, G. and Potter, H. (1965) Proc. Natl. Acad. Sci. U.S.A. 54, 246–252.

[3] Frangione, B., Bernheim, D., and Farkas, K. (1972) J. Mol. Biol. 66, 457–481.

[4] Blobel, G. and Potter, H. (1965) Proc. Natl. Acad. Sci. U.S.A. 54, 246–252.

[5] Blobel, G. and Potter, H. (1965) Proc. Natl. Acad. Sci. U.S.A. 54, 246–252.

[6] Blobel, G. and Potter, H. (1965) Proc. Natl. Acad. Sci. U.S.A. 54, 246–252.

[7] Blobel, G. and Potter, H. (1965) Proc. Natl. Acad. Sci. U.S.A. 54, 246–252.
G39P monomers in the crystal that reveals the helical array to have a very predominantly negatively charged outer surface with the uncharged or positive surface mostly buried in intermonomer contacts or close to the helical axis (Fig. 2). The unobserved, flexible C-terminal domain may modify the apparent surface charge, but the calculated pI is 4.9 for the C-
Structure of SPP1 Phage G39P Helicase Loader

Fig. 2. The arrangement of monomers of G39P112 in the crystal lattice and their electrostatic surface potential. A, ribbon trace of the three monomers in the a.u. labeled 1–3 and shown in red, green, and blue. B, G39P112 monomers shown as ribbon trace and equivalent surface colored for electrostatic potential (red ≤−10 kcal (mol)/H18528 and blue ≥10 kcal (mol)/H11002). View is orthogonal to the non-crystallographic 6, screw axis and shows two copies of the a.u. (monomers labeled 1–3 and 1′, 2′, 3′). C, the packing of adjacent helical repeats of G39P112 monomers in the crystal lattice as viewed along the non-crystallographic 6, screw axes, with the final residues of the N-terminal domains marked C, indicating the locations of the C-terminal domains. Figure was produced using the programs GRASP (39), MOLSRCRIPT (37), and Raster3D (38).

terminal 59 residues of G39P (the calculated pl for the N-terminal 67 residues is 5.2) and thus might suggest a generally negatively charged surface. Apart from the slight imbalance in positively and negatively charged residues that leads to the acidic pl, the C-terminal domain of G39P does not show a particularly abnormal distribution of residue type.

Molecular replacement attempts to determine the structure of the wt G39P protein using the lower resolution data collected to 3.5 Å from the seemingly related P6122/P6522 crystal form are ongoing but have so far been unsuccessful.

Analysis of Internal Mobility—In order to determine to what extent the disorder observed for the C-terminal half of the molecule reflected conformational heterogeneity in solution rather than disorder within the crystal, the 1H-NMR behavior of G39P was investigated. One- and two-dimensional 1H-H-TOCSY experiments were recorded on samples of both wt G39P and G39P112 mutant protein forms (Fig. 3). Spectra recorded at room temperature for both forms of the protein revealed two domains with very different degrees of motion, as revealed by differential NMR relaxation rates. For the C-terminal ~50 residues, i.e. about half the size of G39P, NMR relaxation rates are slow and are thus dominated by internal mobility far in excess of the overall rotation of the protein. Resonances from these residues show little chemical shift dispersion away from their random coil values, indicative of conformational averaging, and high intensity cross-peaks in TOCSY spectra, as illustrated by the correlations between the aromatic ring protons of Phe-76 and Tyr-80 in Fig. 3B. The intensity of the primary amide cross-peaks of glutamines 84, 90, 104, and 107 and asparagines 99 and 110 is also apparent in Fig. 3B. Backbone amide proton resonances from the mobile C-terminal domain, upon heating the sample, are severely attenuated in intensity by solvent exchange, following saturation of the water resonance (Fig. 3A). This demonstrates that there is weak or no hydrogen bonding involving this part of the protein backbone other than to solvent molecules.

The resonances from the N-terminal residues, at room temperature, have far lower intensity than would be expected for a protein of around 13–15 kDa, indicative of a self-association process under the conditions of the NMR experiments. Upon heating, resonances from this region sharpen markedly, indicating the thermal dissociation of the aggregate (Fig. 3A). In contrast to the resonances corresponding to the mobile region, these resonances display the chemical shift dispersion of a normally folded domain and include, for example, the easily identifiable indole NH of Trp-33 located at the N terminus of helix oB and found within the hydrophobic core of the crystal structure of the N-terminal domain.

The NMR experiments indicate that in solution the protein behaves as a two-domain entity. One domain, corresponding to approximately the 67 N-terminal residues observed in the electron density map from the x-ray experiment, exists in the fold determined above, although at room temperature it is involved in a thermally labile, self-association process. The second domain, corresponding to the remaining 60 C-terminal residues that are not observed in the electron density map, has rapid internal motion and no well defined and stable fold involving immobilized side chains. A detailed comparison of the spectra from the intact wt and truncated forms of G39P revealed no major difference between the two forms. Our findings strongly
the relevant number in a concentration range. Molecular weight standards are shown, and the

G

P112 (P(Afiltration column chromatograms are shown of wtG

P as observed in the crystal and that reported previously in artifact but reflected an underlying flexibility that may be

structure for the C-terminal region was not a result of the support the idea that the disorder observed in the crystal structure for the C-terminal region was not a result of the truncation of the protein nor was it merely some form of crystal artifact but reflected an underlying flexibility that may be closely related to the function of the protein.

Analysis of G39P Oligomeric State—In order to examine further the apparent difference between the oligomeric state of G39P as observed in the crystal and that reported previously in solution (6), the wt protein and the G39P112 mutant were subjected both to gel filtration and analytical ultracentrifugation analyses.

The gel filtration studies suggested that under the conditions tested (40 mM Tris-HCl, pH 8.0, containing 100 mM NaCl at 4 and 25 °C) both the wt G39P and the G39P112 mutant apparently exist largely as a dimer when in the micromolar concentration range (Fig. 4) and as an equilibrium between dimer and monomer when in the nanomolar concentration range (data not shown). However, the NMR data above reveals a rapid equilibrium in solution between dissociated and aggregated states of the protein and emphasizes the need for a more cautious interpretation of the results of gel filtration experiments that are necessarily carried out over much longer time scales. The observation of a species approximating to the size of a dimer might actually arise from the rapid interchange between the monomeric and aggregated forms of the protein and is further complicated by an increase in the hydrodynamic radius arising from the flexible C-terminal domain. Gel filtration of G39P samples subjected to protein cross-linking in the micromolar concentration range apparently revealed monomers, dimers, trimers, and higher order oligomers, but interpretation of these results carries the same caveats as described for the non-cross-linked gel filtration analysis with respect to the rapid equilibration between the aggregated states of the protein and its hydrodynamic radius.

In the analytical ultracentrifugation experiments, samples of the wt G39P and G39P112 mutant proteins were subjected to sedimentation velocity measurements under similar solvent conditions to those used in the gel filtration experiment but at both acidic and basic pH values (10 mM Bis-Tris-HCl, pH 6.0, or Tris-HCl, pH 8.0, 160 mM KCl) and a protein concentration in the micromolar range (Fig. 5). The average mass calculated from both absorbance and interference scans of the solute front was −12.0 kDa for wt G39P and 12.4 kDa for G39P112, which fits well with a monomer form. However, the plots of the rate of change of the concentration (dc/dt) versus the sedimentation coefficient S look slightly irregular for the wt G39P sample at pH 8.0, and a better fit to a mixture of monomer and dimer can be made, and it is reasonable to conclude that there is an equilibrium between monomer and higher order oligomer forms favoring the monomeric species under these conditions.

Limited Proteolysis Study of G39P—An investigation of the general susceptibility of G39P to proteolytic degradation was performed using proteinase K and both wtG39P and an N-terminal His-tagged variant. The analysis revealed a much more pronounced sensitivity to proteolytic cleavage in the C-terminal half of the protein that resulted in fragments corresponding to residues 1–79, 1–87, 1–90, 1–94, and 1–106 (Fig. 6). The identity of the fragments was confirmed by the correspondence of the molecular weights determined by mass spectrometry and by retention on a Ni-NTA column of equivalent fragments (as assessed by SDS-PAGE) from the N-terminal His-tagged variant.

Genetic and Biochemical Analysis of Deletion Mutants—A series of SPP1 gene 39 mutants was studied in complementation assays, and their expressed protein products corresponding to fragments of G39P were examined in vitro to test their ability to interact with G40P-ATPγS and inhibit its helicase, ATPase, or ssDNA binding activities (6). The results of these studies can be analyzed in the light of the domain structure of G39P revealed by this study.

The gene 39 mutant, G39P112, encoding the truncated protein corresponding to residues 1–112, was established in a plasmid-borne system, and this was used to test its ability to complement the defect of the SPP1sus53 conditional lethal mutant. The SPP1sus53 mutant allele has a suppressible mutation at the eighth codon of gene 39, and although a plasmid-borne wt 39 gene fully complements the defect of SPP1sus53, leading to a phage yield indistinguishable from the titer of wt SPP1, the plasmid-borne gene 39-112 mutant is unable to do so. This is consistent with the fact that an SPP1 conditional lethal mutant with a suppressible mutation at codon 103 of gene 39 (SPP1sus22) has been isolated previously (4) and suggests that G39P112 is inactive as a loader and/or inhibitor of the G40P helicase in vivo.

The full-length wt G39P protein is able to interact with G40P-ATPγS and to inhibit all three associated activities (ssDNA binding, ATPase, and helicase activity (6)). Assays have been performed on fragments of G39P and have shown that the G39P112 variant can neither exert a negative effect on G40P activities nor compete out the wt protein from the G39P-G40P-ATP complex (data not shown). A series of G39P truncated variants with N termini deleted up to residue 73 still show interaction with G40P-ATP as assessed by affinity chromatography assay, but a variant consisting of the N-terminal residues 1–68 does not (Fig. 7).
DISCUSSION

Our structural analysis has shown that the G39P protein has two domains: a stably folded 67-residue N-terminal domain and a highly flexible and largely unfolded 59-residue C-terminal domain. This correlates well with our biochemical observations that suggest a bipartite nature for G39P in which the C-terminal domain has been implicated as the fragment of the protein responsible for the interaction with the G40P helicase. The difference in the folding behavior of the two domains of the protein is striking and unexpected and may also indicate some functional significance.

Multifunctional, multidomain proteins are common in biology, but examples as small as the G39P protein are more rare. There is an increasing body of evidence to suggest that “natively unfolded” proteins are quite common in vitro and possibly in vivo (24) and that some can adopt more structured forms only in the presence of partner or target molecules or other ligands (25). Many of these “natively unfolded” proteins have been implicated in disease states such as various forms of cancer, Alzheimer’s and Parkinson’s diseases, and myotonic dystrophy (24), and their unfolded nature has been linked often to their pathological effects. The presence of ordered domains coupled to other largely unfolded domains as observed for the N- and C-terminal domains in G39P also has a precedent in other structural studies, and in many of these cases a functional significance has been assigned to the flexibility of the domains (26–28). Studies on the Salmonella typhimurium regulatory protein FlgM suggest that this protein is intrinsically unstructured when in dilute solution in vitro (26). However, the C-terminal domain of FlgM is observed to adopt a more structured form either when its partner molecule, the RNA polymerase α-factor σ28, is added (26), when it is in vivo, or when in vitro conditions are adjusted to match more closely those found in living cells (29). Thus, it is possible that the C-terminal domain of G39P may possess more structure under in vivo conditions, even in the absence of partner molecules, than we have observed in our experiments. However, it might also be the case that the C-terminal domain of G39P is inherently flexible and has little structure when not in a complex to enable the optimal interaction of G39P with its G40P helicase partner. This interaction has an apparent 6:6 G39P:G40P stoichiometry (6), and the flexibility may be essential for the correct formation of a hexameric arrangement of G39P on the surface of G40P that is still accessible for interaction with origin bound G38P. Indeed G39P may need to bind potentially to a number of monomer forms of the G40P within the hexamer as these may vary depending upon either the relative conformations of the helicase subunits or the state of loading of ATP nucleotide or its hydrolyzed products as observed for the T7 gene 4 helicase (30). Use of a largely unfolded state to bind a variety of targets has been observed previously (31), and the cyclin-dependent kinase inhibitor, p21, has a completely unfolded native state that is suggested to enhance its ability to bind multiple protein targets. Another possible reason for
maintaining a flexible C-terminal domain in G39P could be to enable inactivation of the protein by rapid proteolytic degradation (25). We have observed that removal of just the C-terminal 14 residues impairs G40P binding. Within the cell, random unregulated DNA binding and unwinding by G40P-ATP would be deleterious, and hence some control and targeting of its function is required through the combined action of G39P and G39P to ensure loading at the origin of replication. However, once replication has commenced and the replication machinery moved on from the origin, problems can arise from DNA damage, and the replication fork can stall with release of the replicative machinery and the requirement for subsequent reloading of these components after damage repair. Recently, it has been shown that the loading of G40P at any stalled replication fork by the SPP1 phage-encoded G35P protein can lead to replication fork reactivation (32). At this point, binding of G40P by G39P could be harmful, and thus the levels of G39P might need to be kept low either through its interaction with other factors such as G35P or by its degradation. Indeed, G39P accumulates very fast after phage infection and reaches a plateau at minute 5, remains constant up to minute 18, and goes to initial basal levels after minute 20, whereas levels of G40P accumulate with similar kinetics to those of G39P but remain constant until phage lysis (33). The reason for the apparently abrupt stop in the proteolysis of G39P after the removal of the 14 C-terminal residues under the experimental conditions used prior to crystallization is still under investigation as there is no obvious protease target site at this point. Our proteolytic degradation experiments reveal that the extent of protease sensitivity corresponds well with a more structured N-terminal domain and a less structured C-terminal domain, although the presence of substantial amounts of discrete fragments during the initial stages of proteolysis may argue for some limited structural flexibility might be intrinsic to the ability of the protein to bind G38P and act as a linker in the transfer of the G40P helicase onto its ssDNA target. Upon transfer of the G40P-ATP to the DNA, the G39P dissociates as a heterodimer with G38P (6). The predicted pI of G38P is 9.0, and hence it is likely to have a positive electrostatic surface consistent with its DNA binding function, but this feature might also be important in the interaction with G39P given its overall negative surface charge distribution (see Fig. 2).

Our analyses of the oligomeric state of G39P when free in
solution suggest that it is most likely in a monomeric form at the sub-micromolar concentrations found in vivo but that it can form higher order species or aggregates as the local concentration increases. The crystal structure implies that the oligomerization of the monomers probably does not proceed via the initial formation of 2-fold rotationally symmetric dimers but rather the gradual building up of larger species via growing chains of monomers that have 6-fold symmetry potential. Indeed the primary function of the N-terminal domain of G39P may be oligomerization for presentation of the C-terminal domain to partner proteins, but we are currently investigating further possible roles in the interaction with G38P.

Thus, this first crystal structure for a helicase loader inhibitor protein involved in σ-type DNA replication has revealed an unexpected, highly plastic, bipartite structure that has developed to fulfill multiple interaction functions and to ensure the critical loading of the replicative DNA helicase on the DNA origin of replication.

Acknowledgments—We thank Michel Roth and the station staff on BM30 at the European Synchrotron Radiation Facility and the staff at the Central Laboratory of the Research Councils Daresbury Synchrotron Radiation Source laboratory.

REFERENCES
1. Baker, T. A., and Bell, S. P. (1998) Cell 92, 295–305
2. Marians, K. J. (1992) Annu. Rev. Biochem. 61, 673–719
3. Lee, D. G., and Bell, S. P. (2000) Curr. Opin. Cell Biol. 12, 280–285
4. Pedre, X., Weise, F., Chai, S., Luder, G., and Alonso, J. C. (1994) J. Mol. Biol. 236, 1324–1340
5. Ayora, S., Missich, R., Mesa, P., Weise, F., Chai, S., Lu, R., and Alonso, J. C. (1999) J. Mol. Biol. 291, 50–64
6. Ayora, S., Stasiak, A., and Alonso, J. C. (1999) J. Mol. Biol. 288, 71–85
7. Ayora, S., Mesa, P., Weise, F., Stasiak, A., and Alonso, J. C. (2002) Nucleic Acids Res. 30, 2280–2289
8. Bárcena, M., San Martin, C., Weise, F., Ayora, S., Alonso, J. C., and Carazo, J. M. (1998) J. Mol. Biol. 283, 809–819
9. Ayora, S., Langer, U., and Alonso, J. C. (1998) FEBS Lett. 439, 59–62
10. Martínez-Jiménez, M. I., Mesa, P., and Alonso, J. C. (2002) Nucleic Acids Res. 30, 5056–5064
11. Jones, C. E., Mueser, T. C., Dudax, K. C., Kreuzer, K. N., and Nossal, N. G. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 8312–8318
12. Ismael, F. T., Alley, S. C., and Benkovic, S. J. (2002) J. Biol. Chem. 277, 20555–20562
13. Otwonowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
14. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 849–861
15. Otwonowski, Z. (1991) in Proceedings of the CCP4 Study Weekend, Warrington, January 25–26, 1991 (Wolf, W., Evans, P. R., and Leslie, A. G. W., eds) pp. 80–88, SERC Daresbury Laboratory, Warrington, UK
16. Ramakrishnan, V., and Biou, Y. (1997) Methods Enzymol. 276, 538–557
17. Cowtan, K. (1994) Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31, 34–38
18. Marshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D Biol. Crystallogr. 53, 240–255
19. Winn, M. D., Isupov, M., and Marshudov, G. N. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr. 57, 122–133
20. Martin, J. R., Jerala, R., Kroonzitko, L., Zerovnik, E., Turk, V., and Walther, J. P. (1994) Eur. J. Biochem. 225, 1181–1194
21. Stafford, W. F. (1992) Anal. Biochem. 203, 295–301
22. Philo, J. S. (2000) Anal. Biochem. 279, 151–163
23. Chai, S., Stepan, U., Luder, G., Trautner, T. A., and Alonso, J. C. (1993) Gene (Amst.) 129, 41–49
24. Uversky, V. N. (2002) Eur. J. Biochem. 269, 2–12
25. Dyson, H. J., and Wright, P. E. (2002) Curr. Opin. Struct. Biol. 12, 54–60
26. Daughdrill, G. W., Chadey, M. S., Karlinsky, J. E., Hughes, K. T., and Dahlquist, F. W. (1997) Nat. Struct. Biol. 4, 285–291
27. Wright, P. E., and Dyson, H. J. (1999) J. Mol. Biol. 293, 321–331
28. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Biochemistry 41, 6573–6582
29. Dedmon, M. M., Patel, C. N., Young, G. B., and Pielak, G. J. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 12681–12684
30. Singleton, M. R., Sawaya, T. E., Ellenberger, T., and Wigley, D. B. (2000) Cell 101, 589–600
31. Kiwinski, R. W., Hengst, T. L., Reed, S. I., and Wright, P. E. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 11504–11509
32. Ayora, S., Missich, R., Mesa, P., Lu, R., Yang, X., Egelman, E. H., and Alonso, J. C. (2002) J. Biol. Chem. 277, 35989–35997
33. Weise, F. (1997) Biochemische Charakterisation der Initiation der DNA-Replikation des Bacillus subtilis-Bacteriophagen SPP1. Ph.D. thesis, Freie Universität, Berlin, Germany
34. Bárcena, M., Ruiz, T., Donato, L. E., Brown, S. E., Dixon, N. E., Radermacher, M., and Carazo, J. M. (2001) EMBO J. 20, 1462–1468
35. Barton, G. L. (1993) Protein Eng. 6, 37–40
36. Eneaouf, R. M. (1999) Acta Crystallogr. Sect D Biol. Crystallogr. 55, 938–940
37. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
38. Merritt, E. A., and Woolfson, N. M. (1990) Acta Crystallogr. Sect D Biol. Crystallogr. 46, 213–220
39. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Proteins, Structure, Function, and Genetics 11, 281–296
40. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
Structural Analysis of *Bacillus subtilis* SPP1 Phage Helicase Loader Protein G39P
Scott Bailey, Svetlana E. Sedelnikova, Pablo Mesa, Sylvia Ayora, Jon P. Waltho, Alison E. Ashcroft, Andrew J. Baron, Juan C. Alonso and John B. Rafferty

*J. Biol. Chem.* 2003, 278:15304-15312.
doi: 10.1074/jbc.M209300200 originally published online February 13, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M209300200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 38 references, 6 of which can be accessed free at http://www.jbc.org/content/278/17/15304.full.html#ref-list-1