Characterization of the Small Subunit of the Terminase Enzyme of theBacillus subtilisBacteriophage SPP1

ARTICLE *in* VIROLOGY · MARCH 1998

Impact Factor: 3.32 · DOI: 10.1006/viro.1997.9017 · Source: PubMed

2 AUTHORS, INCLUDING:

Juan C Alonso
Spanish National Research Council (C…

217 PUBLICATIONS **5,298** CITATIONS

SEE PROFILE

VIROLOGY 242, 279–287 (1998)
ARTICLE NO. VY979017

Characterization of the Small Subunit of the Terminase Enzyme of the *Bacillus subtilis* Bacteriophage SPP1

Aranzazu Gual and Juan C. Alonso[1]

Centro Nacional de Biotecnología, C.S.I.C., Campus Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

Received September 3, 1997; accepted December 24, 1997

The small subunit of bacteriophages SPP1 and SF6 terminase, G*1*P, share 71% identity clustered in three conserved segments (I, II, and III). Within segment I the helix–turn–helix DNA-binding domain was mapped, whereas segment III was found to be nonessential. For terminase activity, chimeric G*1*Ps, obtained by domain swapping between gene *1* of SPP1 and the SF6 origin (Chi1 to Chi4), were purified. The chimeric proteins behave in all respects similarly to the G*1*P of SPP1 or SF6. The major determinant for G*1*P:G*1*P interactions was found to lie within segment II. We showed that a G*1*P derivative (G*1*P*) lacking the 62 N-terminal residues (segment I), and Chi1 lacking the 45 C-terminal residues (segment III) interact with G*1*P. The N-terminal domain of G*1*P is necessary for terminase subunit assembly, because the large subunit of the terminase (G*2*P) interacts only with G*1*P and Chi1, but fails to do so with G*1*P*. These results suggest that segment III and the extended C-terminal part of SPP1 G*1*P do not play a major role in DNA recognition and that G*1*P recognizes an extended nucleotide sequence and DNA structure.

INTRODUCTION

Initiation of packaging of double-stranded bacteriophage DNA concatemers involves the specific interaction of the prohead with virus DNA in a process mediated by a phage-encoded DNA recognition and cleavage (terminase) protein (reviewed in Black, 1989; Murialdo, 1991; Catalano *et al.,* 1995). The terminase enzymes described so far are heterooligomers composed of a small and a large subunit, and with few exceptions, the terminases do not display a significant level of sequence similarity among them (reviewed in Black, 1989). It is thought that the small terminase subunit forms a nucleoprotein structure that helps to position the terminase large subunit at its target site termed *cos* or *pac* (reviewed in Black, 1989).

The terminase enzyme of *Bacillus subtilis* bacteriophages SPP1 and the highly related SF6 are composed of a small (G*1*P) and a large (G*2*P) subunit which are the products of genes *1* and *2,* respectively (Chai *et al.,* 1994). The SPP1 and SF6 G*1*P (native molecular masses of 190 to 210 kDa and 160–180 kDa, respectively) share 71% identity clustered in three discrete regions (segments I, II, and III) (Chai *et al.,* 1994). Within segment I (residues 11 to 49) lies the putative NTP-binding motif and the information to specifically recognize the packaging initiation subsites (termed *pac*L *and pac*R) by a helix–turn–helix motif (located between residues 23 and 44) (see Fig. 1A). This is consistent with the fact that the major groove binding dye methyl green, which does not distort sequence-directed bending of DNA, competes with G*1*P for binding at *pac*L DNA (Chai and Alonso, 1996) and that a G*1*P derivative (G*1*P*) lacking the first 62 residues interacts with wild-type G*1*P but fails to bind DNA (Chai *et al.,* 1994). It is thought that the G*1*P segment II (residues 65 to 120) or III (residues 137 to 157) or the concerted action of both are involved in the G*1*P:G*1*P interaction. The putative phosphate-binding loop (AXXXXGK$^{L}/_{A}$) was predicted within segment II (Chai *et al.,* 1994). Purified G*1*P, however, fails to hydrolyze ATP or dATP (Chai *et al.,* 1994). No apparent biological role can be assigned to the carboxy terminal region of the G*1*P of the SPP1 origin (Chai *et al.,* 1995; Fig 1A). In spite of the differences between the SPP1 and SF6 phages at the nucleotide and amino acid level of G*1*P, the G*2*Ps of the two phages are identical (our unpublished results).

The SPP1 *pac* region can be subdivided into three discrete sites (*pac*L, *pac*C, and *pac*R) (Chai *et al.,* 1995). The nuclease activity associated with the terminase large subunit (G*2*P) cleaves at *pac*C, located between *pac*L and *pac*R sites (Deichelbohrer *et al.,* 1982; Chai *et al.,* 1992, 1995). G*1*P binds cooperatively to the encapsidated (*pac*R) and nonencapsidated (*pac*L) DNA ends and holds the two binding sites together in a DNA loop (Chai *et al.,* 1995). The G*1*P recognition site at *pac*L is embedded in a sequence-directed DNA bend and G*1*P binding to *pac*L DNA enhances DNA bending (Chai *et al.,* 1995). Distamycin, a minor groove binder that induces local distortions of the DNA, inhibits G*1*P–*pac*L complex formation, but other minor groove binding compounds that

[1] To whom correspondence and reprint requests should be addressed.

0042-6822/98 $25.00

do not distort DNA failed to compete with G*1*P for *pac*L DNA binding (Chai and Alonso, 1996). The competition of G*1*P with distamycin for DNA binding at the *pac*L site was also shown to be independent of the order of addition of the reactants. Furthermore, cationic metals, which generate a repertoire of DNA structures different from the one caused by the minor groove binder distamycin, can partially reverse the distamycin-induced inhibition of G*1*P binding to *pac*L DNA (Chai and Alonso, 1996). DNase I footprinting experiments suggest that each G*1*P target site contains two discrete binding domains, termed Box a in *pac*L and Box c in *pac*R (Chai *et al.,* 1995). The interaction of G*1*P with *pac*L DNA was observed only on one side of the double helix. Furthermore, G*1*P bound to *pac*L and G*1*P bound to *pac*R give rise to a loop of 204 bp (or about 20 turns of the DNA helix). Additional evidence for DNA looping is provided by the alternating DNase I-hypersensitive and DNase I-resistant sites in the complexed DNA, which appear with an approximate periodicity of 10 bp (Chai *et al.,* 1995). On the basis of published data it has been hypothesized that the natural sequence-directed bend that exists within the *pac*L site is the architectural element that facilitates the assembly of a nucleoprotein complex and hence the initiation of DNA encapsidation by bacteriophage SPP1 (Chai *et al.,* 1995; Chai and Alonso, 1996).

In this work, we have purified and characterized different chimeric SPP1:SF6 G*1*Ps and the SPP1 G*2*P. The purified G*1*Ps bind cooperatively to the nonadjacent *pac*L and *pac*R sites. The interaction of the different G*1*Ps with the intrinsically bent *pac*L DNA facilitates the formation of a higher-order nucleoprotein structure. The binding of the G*1*Ps within *pac*L DNA, which occurs on only one face of the double helix, was studied. G*1*P specifically binds to a discrete DNA region and the binding is stabilized by a DNA structure imposed by the protein. We have confirmed that the G*1*P:G*1*P interacting domain lies within the central portion (residues 63 to 138) of the protein and have determined by different biochemical approaches that the major determinant for G*1*P:G*2*P interaction lies in segment I. The nonessential segment III and the extended C-terminal part of SPP1 G*1*P (see Chai *et al.,* 1994) are dispensable in the G*1*P:G*1*P and G*1*P:G*2*P interactions. The results presented here suggest that G*1*P recognize an extended nucleotide sequence and a DNA structure.

RESULTS AND DISCUSSION

The sequences of G*1*P of SPP1 and the SF6 origin are 71% identical and have three segments (I, II, and III) in common (Chai *et al.,* 1994; Fig. 1A). In a previous study, we showed that certain chimeric G*1*Ps, obtained by domain swapping between gene *1* of SPP1 and SF6, complement the defect of SPP1 conditional lethal mutants in gene *1* (Chai *et al.,* 1994; Fig. 1B).

In the case of bacteriophage λ, the subunit assembly domain lies within the C-terminal domain of the terminase small subunit (Frackman *et al.,* 1985). To map the protein residues involved in G*1*P–DNA, G*1*P–G*1*P, and G*1*P–G*2*P interactions we have purified and characterized SPP1 wild-type G*1*P, a SPP1 G*1*P derivative lacking the first 62 residues (G*1*P*) and four different SPP1 and SF6 G*1*P chimeric proteins. In chimeras Chi1 and Chi2, segment III and the C-terminal end are absent and in chimeras Chi3 and Chi4, segment III has been exchanged between SPP1 and SF6 G*1*P (Chai et al., 1994, Fig. 1B)].

Chimeric G1Ps lacking part of the C-terminal end bind cooperatively to SPP1 *pac* DNA

A 320-bp DNA fragment (0.3 nM) containing the SPP1 242-bp *Xho*II–*Hin*fI (part of *pac*L, *pac*C, and *pac*R sites) was radiolabeled and used as probe in a filter-binding assay using the highly purified G*1*P and chimeric G*1*Ps and G*1*P DNA. Complex formation was determined as a function of G*1*P concentration (Fig. 2). SPP1 G*1*P (predicted molecular mass of 20.7 kDa) and SF6 G*1*P (16.7 kDa) exist in solution as oligomer at about picomolar concentration (the native protein is a decamer in solution) (Chai *et al.,* 1995). The dependence of DNA retention on SPP1 or SF6 G*1*P concentration appeared to follow a sigmoidal curve, indicating cooperativety of binding. The apparent equilibrium constant (K_{app}) of SPP1 G*1*P and *pac*-DNA was estimated to be approximately 10 nM at pH 7.5 and 37°C (Fig. 2). The same K_{app} is observed when SF6 G*1*P was used (Chai *et al.,* 1994, 1995). At the protein concentration midpoint, about 20 SPP1 G*1*P (Fig. 2) or SF6 G*1*P protomers bind to each of the 320-bp DNA fragments in a cooperative manner (Chai *et al.,* 1995).

Under identical experimental conditions, the K_{app} for *pac*-DNA of chimeric SPP1–SF6 proteins was determined. Chi1 (predicted molecular mass 15.5 kDa) and Chi4 (20.3 kDa) have a K_{app} for *pac*-DNA of about 32 and 23 nM, respectively. These are slightly lower than wild-type proteins, whereas the K_{app} of Chi2 (16.0 kDa) and Chi3 (17.0 kDa) were similar (9 and 8 nM, respectively) to that of SPP1 or SF6 G*1*P. It is likely, therefore, that the absence of segment III and the C-terminal domain neither affect the affinity of G*1*P for the SPP1 *pac*-containing (Chi2, see Fig. 1B) nor the cooperativity of the G*1*P–DNA interaction. The latter parameter is used as an indirect measurement of G*1*P oligomerization. This is consistent with the fact that a G*1*P derivative (G*1*P*) lacking the first 62 residues interacts with wild-type G*1*P but fails to bind DNA (Chai *et al.,* 1994). Furthermore, modification of residues at position 123 to 138 and the addition of 31 residues on the C-terminal part of SF6-G*1*P (Chi4) only marginally affected the DNA-binding activity of the protein (2-fold), yet did not affect the cooperativity of binding

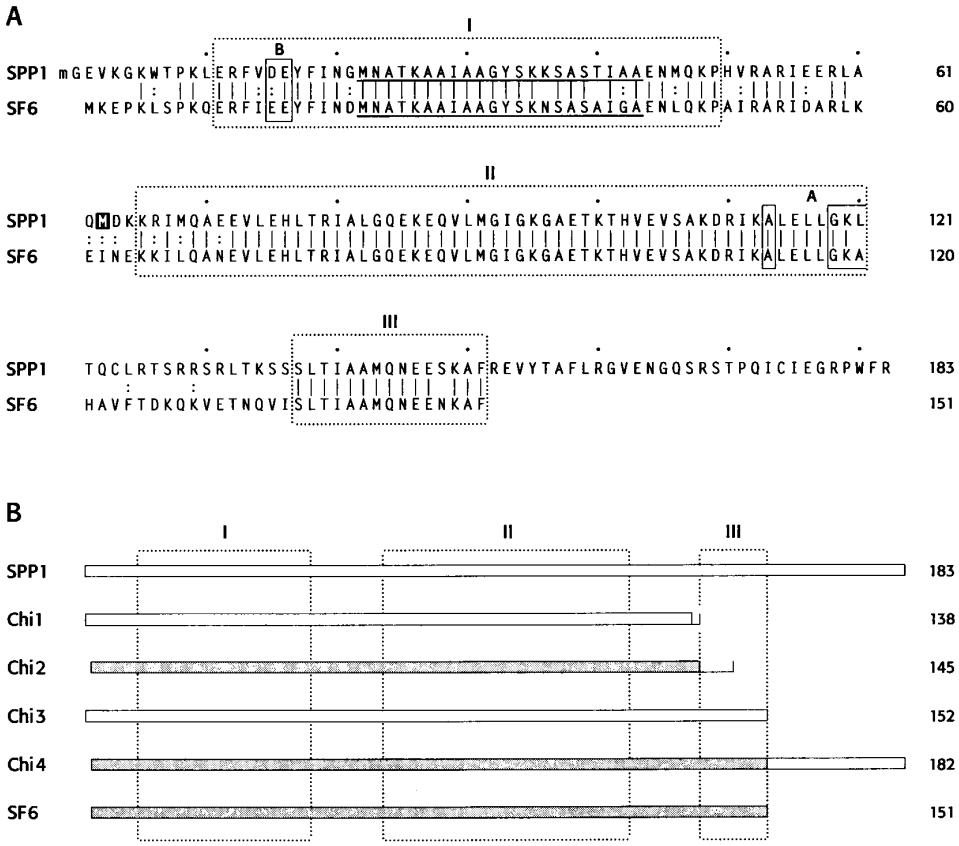

FIG. 1. Amino acid sequence comparison of phage SPP1 and SF6 G*1*P. (A) The deduced amino acid sequences (one letter code) of G*1*P and G*1*P* of SPP1 and SF6 are aligned. The vertical bars denote identical residues, dots conserved residues, and the absence of a bar or dot indicates nonconserved residues. The putative NTP-binding and NTP-hydrolysis motifs (A and B respectively of Walker *et al.,* 1982) are boxed. The putative HTH motif is underlined. The N-terminal methionine absent in the purified SPP1 G*1*P is shown in lowercase and the N-terminal methionine of G1P* is framed (Chai *et al.,* 1994). Segments I, II, and III are indicated as boxed (broken line). (B) Open and filled bars denote the G*1*P of SPP1 and SF6 origin, respectively. The thin lines denote new residues. Chi1 and Chi3 contain the first 136 and 148 residues of SPP1 G*1*P, respectively. Chi2 and Chi4 contain the first 136 and 151 residues of SF6 G*1*P. Chi1 and Chi2 contain at their C-terminal end 2 and 9 residues of vector origin. Chi3 and Chi4 contain 4 and 31 residues at their C-terminal end of SF6 and SPP1 G*1*P origin, respectively. The polypeptide length is indicated. Segments I, II, and III are boxed (broken line).

when compared to the G*1*P of SPP1 or SF6 origins (see Fig. 2). This is in good agreement with our previous results showing that plasmid-borne chimeric proteins (Chi1, Chi2, Chi3, and Chi4) complement the defect of conditional lethal mutants in SPP1 gene *1* (Chai *et al.,* 1994). A plasmid-borne Chi1 and Chi 4 fully complemented the G*1*P defect while in cells bearing a plasmid-borne Chi2 and Chi3, a 17- to 2-fold lower plating efficiency was detected (Chai *et al.,* 1994).

Chimeric G*1*Ps interaction with *pac*L DNA occurs on one face of the DNA double helix and forms a specialized nucleoprotein complex

The G*1*P-binding region on *pac*L contains two direct repeated segments, termed Box a (Chai *et al.,* 1995). The presence of Box a within the intrinsically bent DNA and the G*1*P-induced bending suggests that the length of the DNA segment protected by G*1*P is probably determined by the ability of the protein to direct the DNA to follow a certain conformation imposed by the G*1*P decamers.

A periodic pattern of DNase I hypersensitive sites is observed when wild-type G*1*P (SPP1 or SF6 origin) is incubated with a *pac*L DNA-containing fragment (Chai *et al.,* 1995). To test whether the chimeric G*1*Ps interact with only one face of the DNA double helix and direct the DNA to follow a certain conformation imposed by G*1*P, the 271-bp *Hpa*II–*Bsm*I *pac*L–DNA fragment (coordinates 1 to 271) was used as a probe in DNase I protection experiments using the wild-type or chimeric G*1*Ps (see Fig. 3).

As previously reported, the top strand in the G*1*P–*pac*L complex displayed nine domains of protection interspaced by phosphodiester bonds hypersensitive to DNase I cleavage (Chai *et al.,* 1995; Fig 3). The hypersensitive sites are separated by 10 $\pm$ 1 nucleotides, which is about one helical turn (assuming 10.5-bp per turn) in double-stranded B-form DNA (Fig. 3). The bottom strand exhibited seven protected stretches, two of which were remarkably long (16 $\pm$ 1 nucleotides), with the protected domains interspaced by sites hypersensitive

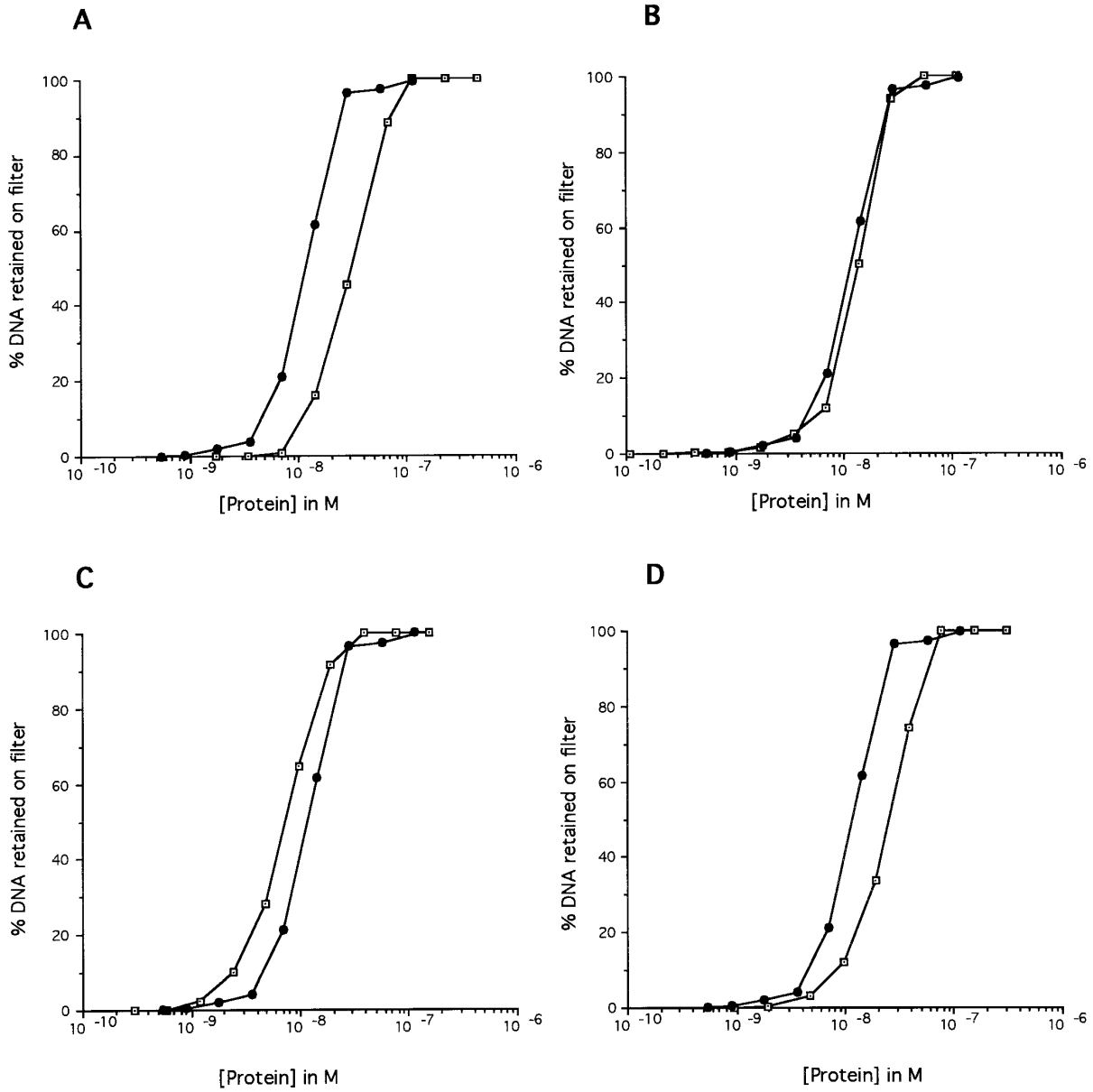

FIG. 2. *pac* region DNA binding by wild-type or chimeric G*1*Ps. The ^{32}P-labeled 320-bp DNA (0.3 nM) (containing the 242-bp *Xho*II–*Hin*fI *pac*-cleavage site) in buffer A (25 mM Tris–HCl, pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$) was brought to 37°C; increasing amounts of G*1*P or chimeric G*1*Ps were added (final volume 20 μl) and the incubation was continued for 15 min. The DNA retained on the filter after washing was corrected for the retention of ^{32}P-labeled *pac*–DNA in the absence of G*1*P (about 2 to 3% of total input). Filled circle, wild-type G*1*P; open square, chimeric G*1*P. In A, Chi1; in B, Chi2; in C, Chi3; and in D, Chi4.

to DNase I cleavage (data not shown). These periodic anomalies in the DNase I cleavage pattern imply that the DNA segment is highly flexible (reviewed by Hagerman, 1990).

In the presence of 120 nM of the G*1*P chimeric proteins (Chi1, Chi2, Chi3, and Chi4) and 5 nM of ^{32}P-labeled *pac*L DNA, the same pattern of protection and interspaced hypersensitivity to DNase I cleavage as with wild-type G*1*P were observed (Chai et al., 1995; Fig. 3).

DNase I footprinting with similar characteristics has been observed for curved DNA that is wrapped around a protein and in looped DNA (Kirkegaard and Wang, 1981; Morrison and Cozzarelli, 1981; Drew and Travers, 1985; Hochschild and Ptashne, 1986). Drew and Travers (1985) explained this phenomenon by proposing that helical grooves located on the inner surface of the bend/loop are sterically occluded and that grooves on the outer face are the sites of enhanced DNase I cleavage. It is likely, therefore, that the chimeric G*1*Ps assemble in a ring-like structure as reported for the SF6 and SPP1 G*1*P (Chai *et al.,* 1995). The *pac*L DNA wraps around a wild-type G*1*P, which has a ring-like structure (see Chai *et al.,* 1995). Since similar DNaseI protected and hypersensitive sites are observed when wild-type G*1*P is replaced

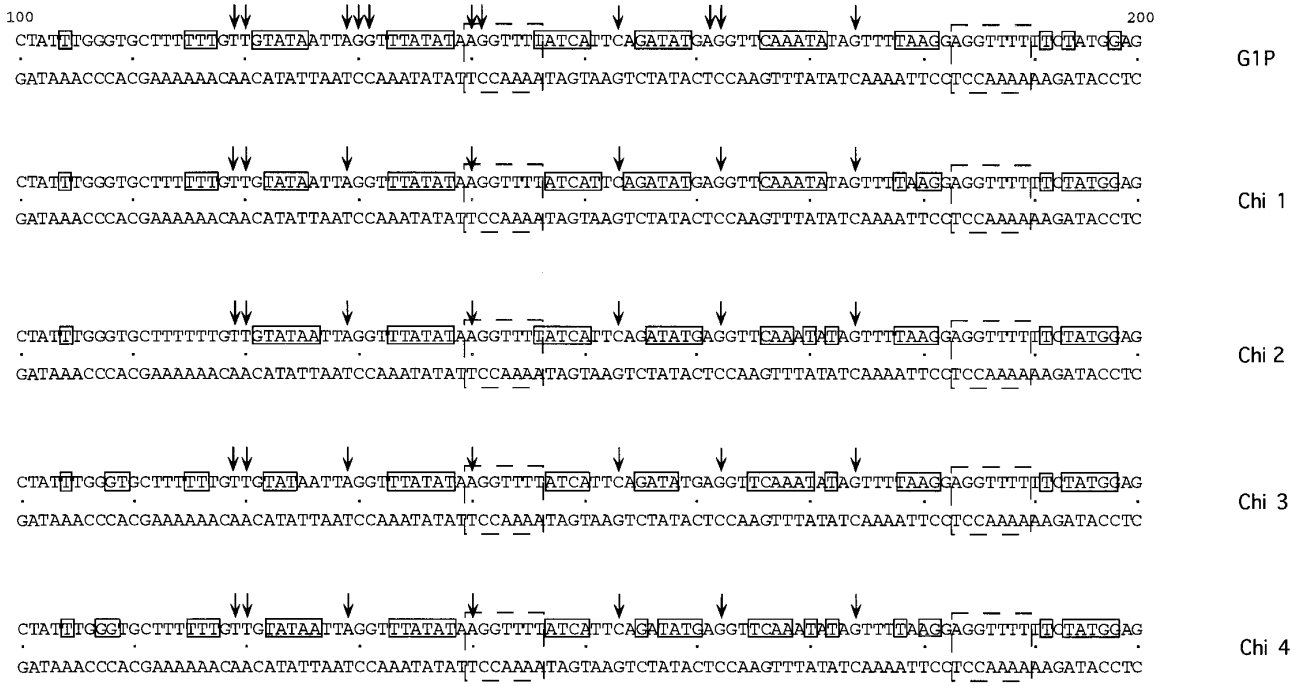

FIG. 3. Schematic representation of the top strand sequences of DNase I protection and hypersensitive sites of *pac*L DNA by wild-type and chimeric G*1*Ps. The 271-bp *Hpa*II–*Bsm*I *pac*L–DNA fragment was end-labeled, incubated with different concentrations of wild-type G*1*P or chimeric G*1*Ps, and partially digested with DNase I. In the top strand (coordinates 100 to 200) the regions protected from DNaseI cleavage are boxed, whereas the hypersensitive sites are denoted by filled arrows. Both strands of the directly repeated box a are boxed (broken lines). The proteins used to detect *pac*L protected and hypersensitive sequences are indicated at the right side.

by the chimeric proteins, we can infer that segment III (absent in Chi1 and 2) is not involved in the G*1*P:G*1*P interaction. It is likely, therefore, that chimeric G*1*Ps also have a ring-like structure.

Wild-type or chimeric G*1*Ps interact with Box a and impose a defined conformation to the *pac*L DNA

The DNase I cleavage pattern from G*1*P–DNA complexes revealed that at low protein concentration a clear contact with two nonadjacent sites [one at the left hand side (*pac*L site) and another one at the right hand side (*pac*R site) of the *pac*C site] was observed. No apparent nucleotide sequence identity could be detected between the *pac*R and *pac*L sites. However, two sets of 7-bp direct repeats, termed Box a (within *pac*L) and Box c (within *pac*R), were observed within the protected region.

To test the relevance of these boxes in the interaction of G*1*P with *pac*-containing DNA (*pac*L–*pac*C–*pac*R), we used methylation interference experiments and the missing-nucleoside assay. The former experiments would allow us to identify those adenine or guanine residues that, when methylated, interfere with G*1*P binding, whereas the latter identify the nucleosides that are important for the binding of G*1*P to DNA. Both types of experiments failed to identify any specific residues whose methylation or absence impaired the binding of G*1*P or chimeric proteins to *pac*-containing DNA (data not shown). It is likely therefore, that the G*1*P binding sites (Boxes a and c) are equally important and G*1*P interacts with two different set of sequences.

We have previously shown that the specific DNase I pattern of protection and hypersensitivity were observed when the left hand-side Box a (between nucleotides 140 to 146, left) of the SPP1 *pac*L segment (Δ*pac*L) was replaced by an unrelated nucleotide sequence (Chai *et al.,* 1995; see Fig. 3). Furthermore, the presence of Boxes a within the intrinsically bent DNA and the G*1*P-induced bending suggests that the precise position at which G*1*P associates with the *pac*L DNA segment and the face of the double helix that is in contact with (or exposed away from) G*1*P determine the extent that the protein causes the DNA to follow a defined conformation (positioning) (Chai *et al.,* 1995; Chai and Alonso, 1996). To test this hypothesis, we constructed *pac*L substrates (Δ*pac*L) in which we systematically deleted DNA sequences from the right-hand side of the segment depicted in Fig. 4. In the presence of 120 nM of wild-type G*1*P and 5 nM of ^{32}P-labeled *pac*L DNA (coordinates 1 to 271), the typical pattern of protection and interspaced hypersensitivity to DNase I cleavage was observed (Figs. 3 and 4). The replacement of coordinates 189 and 271 (rightmost Box a between nucleotides 183 and 189) with nonspecific sequences had no effect on the pattern of protection and hypersensitivity to DNase I (Fig. 4). To obtain the same

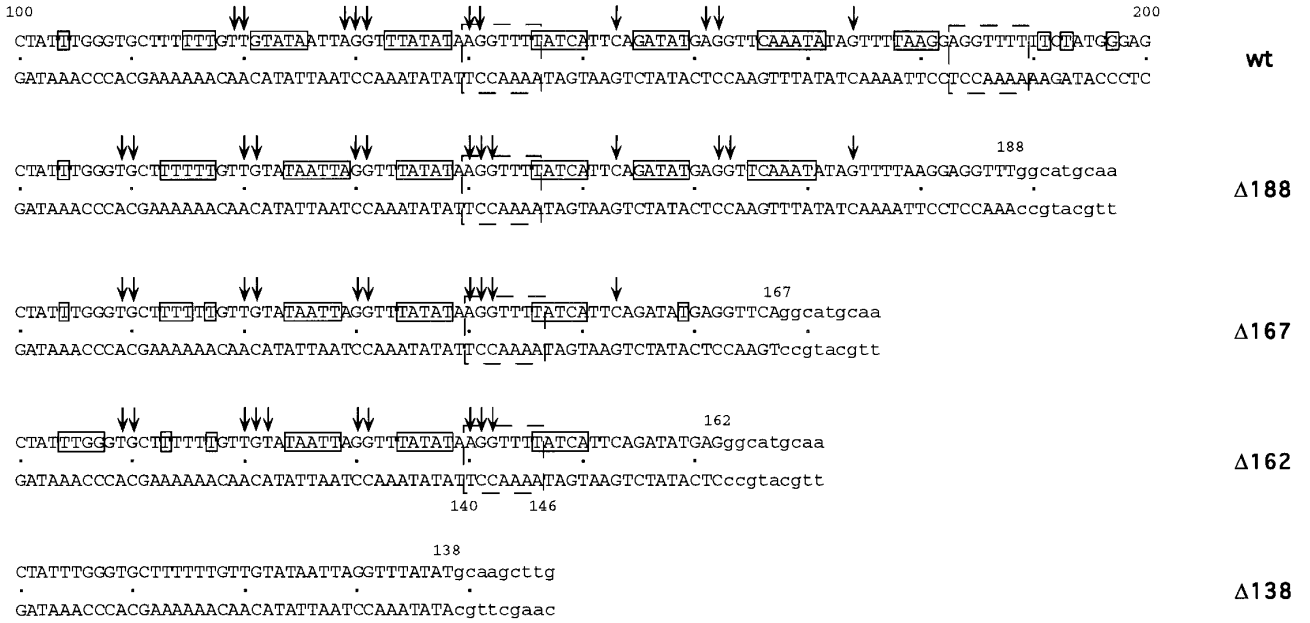

FIG. 4. Alignment of DNase I protection and hypersensitive sites to partial or full *pac*L sequences by wild-type G*1*P. The 271-bp *Hpa*II–*Bsm*I *pac*L–DNA fragment was end-labeled, incubated with different concentrations of wild-type G*1*P or chimeric G*1*Ps, and partially digested with DNase I. In the top strand the regions protected from DNaseI cleavage are boxed, whereas the hypersensitive sites are denoted by filled arrows. Both strands of the directly repeated box a are boxed (broken lines). Only the relevant SPP1 wild-type (wt) *pac*L region (coordinates 100 to 200) is indicated. In *pac*L Δ188 (100 to 188), in *pac*L Δ167 (100 to 167), in *pac*L Δ162 (100 to 162), and in *pac*L Δ138 DNA fragment (100 to 138), the uppercase letters show the SPP1 *pac*L nucleotide sequence, whereas the lowercase letters denote non-SPP1 sequences.

pattern of protection and hypersensitivity to DNase I in the replacements between coordinates 168 to 271 and 163 to 271, however, we needed twice the amount of G*1*P (240 nM). When both Boxes a (coordinates 183 to 189 and 141 to 147) were replaced by nonspecific sequences, the pattern of protection and hypersensitivity to DNase I cleavage was not detected even in the presence of 480 nM of G*1*P. The same results were obtained with the chimeric proteins (data not shown). It is likely, therefore, that the presence of only one Box a, within the intrinsically bent DNA, is sufficient to direct the DNA to follow a certain conformation imposed by the G*1*P ring-like decamer, but that the amount of protein required is higher. This is consistent with the fact that the selective loss of any single nucleotide does not affect the interaction of G*1*P with *pac*-containing DNA (see above) and that *in vivo* DNA packaging is not markedly affected by the deletion of either the *pac*L or the *pac*R DNA region (see Bravo *et al.,* 1990).

G*1*P segments I and II are the major determinants of G*1*P:G*1*P and G*1*P:G2P interaction

To study the G*1*P interacting domain, polyclonal antibodies raised against purified G*1*P or G*1*P* were inmobilized in a protein A–Sepharose column (Fig. 5). G*1*P (540 nM) and G*1*P* (770 nM) (lacking the first 62 residues) were loaded onto the column. G*1*P* was able to retain G*1*P and Chi1 in the column, whereas BSA, a nonspecific control, was present in the flow-through volume (data not shown). These results suggest that G*1*P*, which lacks the first 62 residues (segment I), is able to interact with G*1*P and with the Chi1 chimera that lacks the last 45 residues. From these data we inferred that the major G*1*P:G*1*P interaction domain was located within coordinates 63 and 137. The same type of interaction was observed when the SPP1 and SF6 G*1*P or Chi2 chimera were analyzed. These proteins share no identity between residues 122 and 138. It is likely, therefore, that the major G*1*P:G*1*P interaction domain is between residues 63 and 121 (see Fig. 1A). These results are in good agreement with those obtained using gel filtration experiments (see Chai *et al.,* 1994, 1995).

To characterize a possible interaction between G*1*P and G*2*P, purified G*1*P, G1P*, or the Chi1 protein were immobilized in a protein A–Sepharose column. G*1*P (540 nM), G*1*P* (770 nM), or Chi1 protein (900 nM) was immobilized in the column. G*2*P (210 nM) was then loaded. G*1*P retained G*2*P in the column at 0.5 M NaCl concentration, whereas BSA, a nonspecific control, was present in the flow-through volume. At 1 M NaCl, about 80 to 90% of the input G*2*P was present in the elution volume. The same results were obtained when Chi1 was immobilized on the protein A–Sepharose column (data not shown). When G*1*P* was immobilized in the column, about 60 to 70% of the input G*2*P loaded onto the column was present in the flow-through volume and 20 to 25% in the 0.3 M NaCl wash. It is likely, therefore, that the major determinant for G*1*P:G*2*P interaction is within segment I,

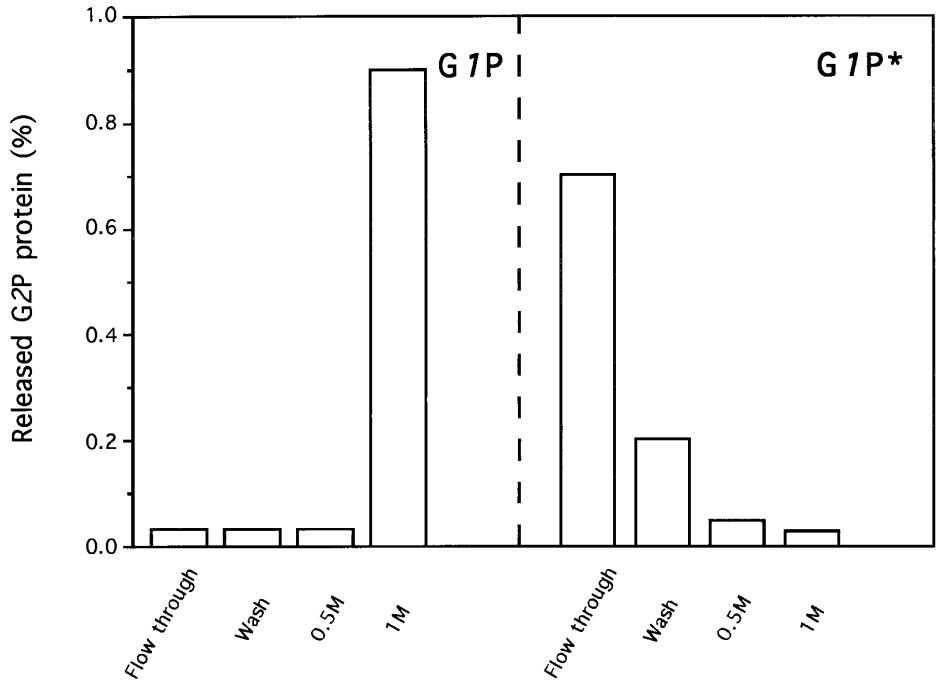

FIG. 5. Protein–protein interaction measured by polyclonal antibodies immobilized on a protein A–Sepharose column. G1P and G1P* (shown at the top) were individually immobilized by specific antibodies and protein A–Sepharose in buffer B (50 mM Tris–HCl, pH 7.5, 5% glycerol) containing 300 mM NaCl, and then G2P was loaded. The flow-through, the wash in buffer B containing 300 mM NaCl, and the elution fractions in buffer B containing 500 mM or 1 M NaCl were concentrated. Data are shown as the percentage of released to input G2P, which was determined by densitometric analysis of the protein bands observed in SDS–PAGE.

because the information provided by the N-terminal 62 residues seems to be an important component (see Fig. 1A). At the present time, we cannot assay the role of segment II in the G1P:G2P interaction. Since DNA was absent in the G1P and G2P assembly reaction, it is likely that terminase is assembled before its interaction with the *pac* site.

Within the N-terminal 62 residues, including segment I (residues 11 to 49), lies the putative NTP- and DNA-binding motifs and the major determinant for the G1P:G2P interaction. Within segment II lies the putative phosphate-binding loop ($AXXXXGK^{L}/_{A}$) and the G1P:G1P interacting domain. Segment III and the C-terminal end seem to be dispensable (Chai *et al.,* 1994; this report).

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli strains used were JM103, for propagating plasmids (Yanisch-Perron *et al.,* 1985), and BL21(DE3), for protein overexpression (Studier, 1991).

The plasmid-borne gene *1* chimeric genes [pBT419 (Chi1), pBT420 (Chi2), pBT421 (Chi3), pBT422 (Chi4)], and SPP1 gene *1* (pBT115) (Chai *et al.,* 1994), pBT363, pBT397, and pBT398 (Chai *et al.,* 1995), pT712 (GIBCO-BRL), pQE11, and pREP4 (Quiagen), pUC18 (Yanisch-Perron *et al.,* 1985), and pLysE (Studier, 1991) have been described previously. Plasmids pCB193 (Δ188), pCB194 (Δ167), pCB167 (Δ162), and pCB192 (Δ138) were obtained from the plasmid-borne *pac* site (pBT398) by making nested deletions with *Exo*III enzyme. The deletion end-points were determined by nucleotide sequencing. The plasmid-borne gene *2* (pCB191) was obtained by cloning gene *2* into *Hin*dII–*Hin*dII-linearized pQE11 and its overexpression was carried out in BL21(DE3) cells bearing plasmid pREP4.

Enzymes and reagents

Dithioerythreitol, isopropyl-β-D-thiogalactopyranoside, lysozyme, and the ExoIII kit were from Merck, Calbiochem, Boehringer, and Pharmacia, respectively. Cellulose phosphate was from Whatman, Heparin Sepharose LC-6B, Hitrap-Sepharose, protein A–Sepharose, and Q-Sepharose were from Pharmacia. [α^{32}P]dATP was from Amersham Corp (F.R.G.). Ultrapure acrylamide was from Serva.

DNA manipulations

Covalently closed circular plasmid DNA was purified by using the SDS lysis method (Sambrook *et al.,* 1989) followed by purification on a cesium chloride–ethidium bromide gradient. Gel-purified DNA fragments were end-labeled by filling in the restriction site using the large fragment of DNA Pol I in the presence of [α^{-32}P]dATP and dTTP, dCTP, and dGTP. Analytical and preparative gel

electrophoreses of plasmid DNAs and restriction fragments were carried out either in 0.8 % (w/v) agarose/Tris–acetate–EDTA horizontal slab gels or with 4% (w/v) polyacrylamide/Tris–borate gels. The concentration of DNA was determined using the molar extinction coefficients of 6500 $M^{-1} \times cm^{-1}$ at 260 nm. The amount of DNA is expressed in moles.

Protein purification

Wild-type SPP1 G*1*P and SPP1 G*1*P* were purified as previously described (Chai *et al.,* 1994). The chimeric G*1*Ps were purified using the same protocol. G*2*P was purified using a Hitrap nickel-chelating column according to the manufacters instructions (Pharmacia). The concentrations of G*1*P, G*1*P*, the chimeric G*1*Ps, and G*2*P are expressed as moles of protein protomers.

Filter binding assay

The formation of DNA complexes was measured by using alkali-treated filters (Millipore, 0.45 μm type HAWP) as previously described (Chai *et al.,* 1995). The standard reaction (20 μl) was carried out in a mixture of a solution containing 0.3 nM (5 fmol) of ^{32}P-labeled 242-bp *Xho*II–*Hin*fI DNA (obtained as a 320-bp *Hin*dIII–*Eco*RI DNA fragment from plasmid pBT363) and increasing concentrations of SPP1 G*1*P or chimeric G*1*Ps (0.45, 0.9, 1.8, 3.5, 7, 14, 28, 56, and 112 nM) in buffer A (25 mM Tris–HCl, pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$) and incubated for 15 min at 37°C. Ice-cold buffer A (1 ml) was added to the mixture to stop the reaction, which was then filtered through KOH-treated filters. Filters were dried and the amount of radioactivity bound to the filter was determined by scintillation counting. The DNA retained on the filter was corrected for the retention of radioactively labeled DNA in the absence of G*1*P. The specific activity of the labeled DNA was measured as TCA-precipitable material.

DNA footprinting assay

DNase I cleavage reactions were performed essentially as described by Galas and Schmitz (1978). Reaction mixtures (20 μl) consisting of 5 nM of ^{32}P-labeled DNA (the SPP1 271-bp *Hpa*II–*Bsm*I DNA segment was taken as a 308-bp fragment from plasmid pBT397) were equilibrated at 37°C for 20 min with different concentrations of G*1*P or chimeric G*1*Ps (26, 42, 84, 169, 338, 676, and 1352 nM) in buffer A. DNase I was added to obtain one cut per DNA molecule on average and the digestion was for 2 min at 37°C. The reactions were terminated by the addition of 25 mM EDTA. The DNA was precipitated and then resuspended in denaturing formamide loading buffer and separated in 6% denaturing PAGE as described (Sambrook *et al,.* 1989). Autoradiographs of the dried gels were subsequently made.

Methylation interference assays

The end-labeled DNA fragments used for methylation interference experiments (Siebenlist and Gilbert, 1980) were the same as those indicated for the DNase I footprinting (pBT397) and filter-binding (pBT363) assays. End-labeled DNA was partially methylated with 50 mM dimethyl sulfate in 50 mM sodium cacodylate (pH 8), 25 mM $MgCl_2$, 0.1 mM EDTA. After 10 min at room temperature, the reaction was stopped by addition of β-mercaptoethanol (0.2 M final concentration), and the DNA was precipitated twice with ethanol. Methylated DNA was incubated with wild-type G*1*P (670 nM) in buffer A and 3 μg of poly(dI–dC) in a total volume of 20 μl. Free and protein-bound DNA were resolved by electrophoretic mobility shift assays (EMSA), eluted from the gel, and precipitated with ethanol. DNA was resuspended in 100 μl of 1 M piperidine and incubated for 30 min at 90°C to generate breaks at methylated positions. Piperidine was evaporated under vacuum with 1 vol of ethanol, and the DNA analyzed using 6% urea–polyacrylamide gels.

Missing-nucleoside assays

The end-labeled *pac* DNA fragments used were the same as those indicated for methylation interference assay. DNA (20 μl) was treated by the addition of 3 μl of a freshly prepared solution containing 4 mM EDTA, 2 mM ammonium iron(II)–sulfate hexahydrate, 16 mM sodium ascorbate, and 1.5% H_2O_2 to obtain DNA fragments which contained on average no more than one gap per molecule (Tullius and Bombrosky, 1983). After 4 min, the reaction was stopped by addition of 2 μl of 100 mM thiourea and 2 μl of 0.5 M EDTA. Samples were diluted 1:1 with water, the DNA was precipitated and then resuspended in 20 μl of 25 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, and 2.5 μg of poly[d(I–C)]. The wild-type G*1*P (670 nM) was added and the mixture incubated for 15 min at 37°C. Bound and unbound DNA were separated by EMSA, excised from the gel, purified by electroelution, precipitated with ethanol, and analyzed with 8 % urea–polyacrylamide gels.

Affinity chromatography analysis

Polyclonal rabbit antibodies raised against SPP1 G*1*P or G*1*P* were coupled to a protein A–Sepharose column as recommended by the supplier (Pharmacia). G*1*P (540 nM), G*1*P* (770 nM), Chi1 (900 nM), Chi2 (700 nM), Chi3 (610 nM), Chi4 (490 nM), and G*2*P (210 nM) were incubated together or separately at room temperature for 60 min in buffer B (50 mM Tris–HCl, pH 7.5, 5% glycerol) containing 300 mM NaCl and then loaded onto the AntiG*1*P– or AntiG*1*P*–protein A–Sepharose column (20 μl column) equilibrated with the same buffer. The columns were then washed with 15 column volumes of buffer B containing 300 mM NaCl and eluted with 5

column volumes of buffer B containing 500 and 1 M NaCl. Fractions were analyzed by a 12.5% SDS–PAGE.

ACKNOWLEDGMENTS

This research was partially supported by Grants PB 96-0817 from DGCICYT and 06G/004/96 from the Consejería de Educación y Cultura de la Comunidad de Madrid to J.C.A. We are grateful to F. W. Studier for providing the BL21(DE3) strain. We thank S. Chai for her help in the early stages of this work and L. Yuste for excellent technical assistance.

REFERENCES

Black, L. W. (1989). DNA packaging in dsDNA bacteriophages. *Annu. Rev. Microbiol.* 43, 267–292.

Bravo, A., Alonso, J. C., and Trautner, T. A. (1990). Functional analysis of the *Bacillus subtilis* bacteriophage SPP1 *pac* site. *Nucleic Acids Res.* 18, 2881–2886.

Catalano, C. E., Cue, D., and Feiss, M. (1995) Virus DNA packaging: The strategy used by phage λ. *Mol. Microbiol.* 16, 1075–1088.

Chai, S., Bravo, A., Lüder, G., Nedlin, A., Trautner, T. A., and Alonso, J. C. (1992). Molecular analysis of the *Bacillus subtilis* bacteriophage SPP1 region encompassing gene *1* to *6*. The products of gene *1* and gene *2* are required for *pac* cleavage. *J. Mol. Biol.* 124, 87–102.

Chai, S., Kruft, V., and Alonso, J. C. (1994). Analysis of the *Bacillus subtilis* bacteriophage SPP1 and SF6 gene *1* product: a protein involved in the initiation of headful packaging. *Virology* 202, 930–939

Chai, S., Lurz, R., and Alonso, J. C. (1995). The small subunit of the terminase enzyme of *Bacillus subtilis* bacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. *J. Mol. Biol.* 252, 386–398.

Chai, S., and Alonso, J. C. (1996). Distamycin-induced inhibition of formation of a nucleoprotein complex between the terminase small subunit G1P and the non-encapsidated end (*pac*L site) of *Bacillus subtilis* bacteriophage SPP1. *Nucleic Acids Res.* 24, 282–288.

Deichelbohrer, I., Messer, W., and Trautner, T. A. (1982). Genome of *Bacillus subtilis* bacteriophage SPP1: Structure and nucleotide sequence of *pac,* the origin of DNA packaging. *J. Virol.* 42, 83–90.

Drew, H. R., and Travers, A. A. (1985). DNA Bending and its relation to nucleosome positioning *J. Mol. Biol.* 186, 773–790.

Frackman, S., Siegele, D. A., and Feiss, M. (1985). The terminase of bacteriophage λ: Functional domains for cosB binding and multimer assembly. *J. Mol. Biol.* 183, 225–238.

Galas, D. J., and Schmitz, A. (1978). DNase footprinting: A simple method for detection of protein–DNA binding specificity. *Nucleic Acids Res.* 5, 3157–3170.

Hagerman, P. J. (1990). Sequence-directed curvature of DNA. *Annu. Rev. Biochem.* 59, 755–781.

Hochschild, A., and Ptashne, M. (1986). Cooperative binding of repressors to sites separated by integral turns of the DNA helix. *Cell* 44, 681–687.

Kirkegaard, K., and Wang, J.C. (1981). Mapping the topography of DNA wrapped around gyrase by nucleolytic and chemical probing of complexes of unique DNA sequences. *Cell* 23, 721–729.

Morrison, A., and Cozzarelli, N.R. (1981). Contacts between DNA gyrase and its binding site on DNA: Features of symmetry and asymmetry revealed by protection from nucleases. *Proc. Natl. Acad. Sci. USA* 78, 1416–1420.

Murialdo, H. (1991). Bacteriophage DNA maturation and DNA packaging. *Annu. Rev. Biochem.* 60, 125–153

Sambrook, J., Maniatis, T., and Fritsch, E.F. (1989). "Molecular Cloning. A Laboratory Manual." 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Siebenlist, U., and Gilbert, W. (1980). Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *Proc. Natl. Acad. Sci. USA* 77, 122–126.

Studier, F. W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* 219, 37–44.

Tullius, T. D., and Dombrosky, B. A. (1986). Hydroxyl radical footprinting: High-resolution information about DNA–protein contacts and application to λ repressor and cro protein. *Proc. Natl. Acad. Sci. USA* 83, 5469–5473.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). Distantly related sequences in the α- and β-subunits of ATP syntase myosin, kinase and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945–951.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103–119.