JOURNAL OF VIROLOGY, Aug. 1985, p. 458–465
0022-538X/85/080458-08$02.00/0

Vol. 55, No. 2

DNA Packaging Initiation of *Salmonella* Bacteriophage P22: Determination of Cut Sites Within the DNA Sequence Coding for Gene *3*

HORST BACKHAUS†

Institut für Genetik und Mikrobiologie der Universität, D8000 Munich 19, Federal Republic of Germany

Received 17 October 1984/Accepted 29 March 1985

DNA packaging of *Salmonella* phage P22 starts at a defined site on a concatemer of P22 genomes. The molecular ends formed at the packaging initiation site (*pac*) map within a region of ca. 120 base pairs and may contain any of the four nucleotides at their 5′ end. The determination of the positions of the cuts within the sequence demonstrates a characteristic distribution of cut sites which apparently cannot be attributed to the sequence organization of the involved regions. Symmetric elements of the sequence might serve as signals for a recognition event(s) at *pac* in a separate process preceding the cutting reaction. The region of packaging initiation is located within the sequence coding for gene *3*. The 3 protein is responsible for the site specificity of this process. We find no significant homology to Nu1 protein, which appears to have an analogous or similar function in the DNA maturation of *Escherichia coli* phage lambda.

Phage-controlled DNA replication synthesizes polymers of tandemly reiterated phage genomes, concatemers. These are the substrates for DNA encapsidation into preformed proheads. During this process, double-stranded breaks are introduced which generate the mature form of phage DNA present in phage particles (12). The enzymatic activity responsible for this DNA maturation in lambdoid *Escherichia coli* phages was called "ter" nuclease. In phage lambda, it forms the unique cohesive ends, 12 nucleotides protruding at the 5′ end, at a specific map position (30, 48).

Phage P22 belongs to the same group of phages as inferred from the regulation of its life cycle and the analogous organization and partial sequence homology of its genome (43). However, it employs a different mechanism of end formation during DNA maturation. P22 DNA molecules contain a terminal redundancy of ca. 3% which is permuted with respect to its position in the genome (34). The distinct localization of ends in populations of DNA molecules from wild-type P22 and deletion and insertion variants implied that the "headful" DNA packaging model of Streisinger and co-workers (40) applies to DNA maturation of phage P22 in a precisely defined manner (46); i.e., the packaging of P22 DNA initiates at a unique site on the genome and proceeds in one direction, sequentially packaging headfuls from the concatemer. The first headful cut generates a new left end for the second headful, and so forth. The position of these headful cuts is not restricted to P22-specific DNA sequences (45, 46). In excision-defective lysogens, a packaging series initiated at prophage sequences will proceed into bacterial DNA sequences and form transducing particles (23, 24, 38, 47).

Other experiments localized the position of cuts and the signal sequence for packaging initiation on the physical and near gene *3* on the genetic map of P22 (21, 23, 47). The site was called *pac* by Jackson and co-workers (21). Molecular ends at *pac* are formed at six different positions within a region of ca. 120 base pairs (bp), as demonstrated by Casjens and Huang (4).

The terminase activity of P22 resides in (or is controlled by) the products of genes *3* and *2* located near *pac*. Amber mutations in these genes abolish DNA maturation (1, 3, 26). The 3 protein determines the site specificity of packaging initiation, as inferred from the mapping of mutations in mutants selected for their altered transduction abilities, and the phenotype of suppressed amber and transduction mutants demonstrating a reduced, changed, or abolished specificity of packaging initiation (6, 22, 33, 36, 43).

In phage lambda, the products of genes *Nul* and *A* are subunits of the terminase. By using partially purified proteins, site-specific nucleolytic activity (Ter activity) is observed in vitro by the formation of molecular ends (2, 19, 41, 42). The *Nul* and *A* genes map at a place in the lambda genome which is analogous to that of gene *3* and *2* of P22, and the molecular weights of these proteins are of the same order of magnitude (49). The determinants for *cos* cutting and the specific substrate recognition among these proteins are not known in phage lambda.

Lambda DNA packaging proceeds in one direction and in a sequential manner along concatemeric DNA similarly to P22 packaging (13). Substrate recognition of phage lambda—e.g., mediating the packaging specificity of lambda phage versus phage 21, both of which contain identical cohesive ends—depends on a signal(s) located in the right arm of *cos* (14). The second cut (and possibly all subsequent cuts in a series) does not depend on the specificity signal and seems to be triggered by a headful signal as in P22 (14, 15). Thus, the specificity signal of phage lambda might serve a similar function in DNA packaging initiation as does *pac* in P22. Both processes seem to have much in common, and in a simplified general model, the main difference would be the site specificity of the nuclease activity of lambda terminase.

MATERIALS AND METHODS

Bacteria, phages, and plasmids. *Salmonella typhimurium* DB 21 was the prototrophic su$^-$ host; 3002 *cys* am su$^+$ carries an amber suppressor and was isolated in our lab (33); H174

† Present address: Gesellschaft für Strahlen und Umweltforschung, Institut für Biologie, D-8042 Oberschleissheim, Federal Republic of Germany.

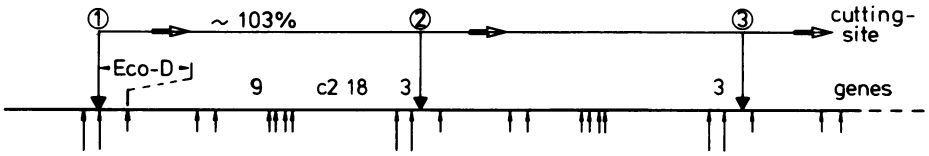

FIG. 1. Sequential packaging of P22 DNA. Cutting positions along a concatemer of P22 DNA are shown. Packaging and cutting is started at site 1, which was called *pac* and also defines the zero position in published restriction maps (7, 10, 21). Cuts 2 and 3 and all subsequent cuts are defined by headful packaging (ca. 103% of P22 genome fitting into phage capsid). They are less precise than the first cut and do not generate a defined band after *Eco*RI restriction. Positions of genes *3* (packaging specificity), *9* (tail fiber), *c2* (repressor), and *18* (replication) are indicated. *Eco*RI sites are indicated with small arrows, and the longer arrows mark the two *Sma*I sites. The distance between two restriction sites at identical positions constitutes the P22 genome length of 42.5 kilobases.

recA$^-$ *leu*$^-$ S^r (Yamamoto) and CC4412 *pro*C90 *hsd*LT (r^-m^+) *hsd*SA (r^-m^+) S^r (C. Colson) were used as hosts for plasmid pGM1000 in biological tests; and strains DB 5064 and DB 5065 carry prophages with deletions as described by Chan and Botstein (5). In DB 5064, the deletion extends from the right side of the prophage map and includes gene *2*; in DB 5065, the deletion includes gene *3* also. *E. coli* strain 490A with the relevant genotype *hsd*R$^-$ *hsd*M$^-$ *rec*A$^-$ was used as a host for plasmids in plasmid DNA preparations.

Phages used were P22 *c2*-5 (H5) (the source of phage DNA); P22 *3*$^-$amN6 (D. Botstein); P22 *3*$^-$amH24 and P22 *3*$^-$amH30 (Soska; obtained from H. Prell); and P22 *c2*-5 *3*$^-$amH24 *erf2*. New *3*$^-$am mutants have been obtained by hydroxylamine mutagenesis and defined by their supressor sensitivity and the different plating on strains DB 5064 and DB 5065.

Plasmid pGM1000 has been previously described (41). The end sites formed by packaging initiation (4; see below) map within the *Hin*dIII-*Xho*I fragment of P22 DNA replacing the 622-bp *Hin*dIII-*Sal*I fragment of pBR322 in this plasmid.

Chemicals and enzymes. Most restriction enzymes, calf intestine phosphatase, DNA polymerase (large fragment), and polynucleotide kinase were from Boehringer GmbH, Mannheim, Federal Republic of Germany. *Hin*fI and *Asu*I (*Sau*96I) were purchased from New England BioLabs, Schwalbach, Federal Republic of Germany. *Xho*I, *Dde*I, and *Hae*III were from Bethesda Research Laboratories, GmbH, Neu Isenberg, Federal Republic of Germany.

Determination of 5'-terminal nucleotides. Fragments labeled at their 5' end were incubated with 1 μg of DNaseI and 1 μg of snake venom diesterase in 10 mM Tris-hydrochloride (pH 8.5)–10 mM $MgCl_2$ in a volume of 30 μl at 37°C for 1 h to obtain complete digestion of the 5' mononucleotides. After the addition of four deoxynucleotides (0.5 mg/ml), the mixture was applied onto Whatman 3MM paper and electrophoresed at 3×10^3 V in 0.5% pyridine–5% acetate (wt/vol). After the marker dye bromophenol blue had migrated 10 cm, the paper was dried and sprayed two times with 0.1% ammonia. The monophosphate dots visualized in UV light were cut out, and the radioactivity was counted in toluene with PPO (2,5-diphenyloxazole; 5 g/liter) and dimeth-yl-POPOP [1,4-bis(5-phenyloxazolyl)benzene; 0.3 g/liter].

DNA preparations, end labeling, and sequencing. Phage particles were purified by a discontinuous (1.3 and 1.7 g/cm^3) and a continuous CsCl gradient (32,000 rpm, 24 h, Beckman Ti50 rotor). DNA was prepared by two phenol extractions and extensive dialysis in 10 mM Tris-hydrochloride–0.1 mM EDTA (pH 8.0). Plasmids were amplified by the procedure of Clewell and Helinski (9). Cells were lysed with lysozyme, EDTA, and sodium dodecyl sulfate as described by a Godson and Vapnek (18), and plasmid DNA was purified in a cesium chloride–ethidium bromide gradient (45,000 rpm, 48 h, Ti 50 rotor). Fragments of a restriction digest were isolated from agarose gels (after cutting out the stained bands) by electroeluting the DNA into electrophoresis buffer in a sealed dialysis bag.

For end-labeled fragments from polyacrylamide gels, the diffusion method into extraction buffer described by Maxam and Gilbert (27) was used. *Hae*III-digested lambda dv1 DNA (a gift from G. Hobom) or *Hin*fI- and *Eco*RI-digested pBR322 DNA were used as molecular weight standards. End labeling and sequencing essentially followed the protocol of Maxam and Gilbert (27). The 5'-end labeling of flush ends (*Hae*III, *Sma*I, or especially the *pac* ends of *Eco*RI-D) was performed in 45 mM glycine-NaOH (pH 9.5)–5% (vol/vol) glycerol–5 mM dithiothreitol–10 mM $MgCl_2$ without a denaturing step. Usually *Asu*I (*Sau*96I), *Hin*fI, and *Dde*I fragments were used for 5'-end labeling. In few cases, the 3' end was end labeled with Klenow fragment of polymerase I or terminal transferase. Either isolated fragments or, in most cases, the fragments of a restriction digest were labeled and separated on a polyacrylamide gel directly or after a secondary restriction. The products of the base-specific sequencing reactions were separated on 8, 15, or 20% gels with 8 M urea (60 by 25 by 0.02 cm) and run with constant power at a temperature of 50 to 60°C. The 8% gels were dried after 10% acetic acid treatment to a thin film. For autoradiography, Kodak XAR-5 X-ray film was used without an intensifying screen.

Evaluation of sequence data. Overlapping sequences, restriction sites, and protein reading frames were found by the computer program of Larson and Messing (25). Homologies and inverted repeats were evaluated by matrix plotting programs of K. May. Programs were run on an Apple IIe microcomputer.

RESULTS

Mapping of molecular ends generated at *pac*. In P22 DNA, the cuts formed by packaging initiation at *pac* and digestion with *Eco*RI generate a fragment (*Eco*RI-D) of ca. 4 kilobases which is present in a molar ratio of 0.3 (Fig. 1; 21). Apparently, one-third of all mature P22 chromosomes are formed as first headfuls. *Eco*RI-D was isolated from preparative agarose gels. Recutting with *Sma*I and *Xho*I affected the migration of *Eco*RI-D without additional smaller fragments being visible. A 10-fold increase in DNA concentration made a set of DNA fragments of 400 to 550 bp demonstrable after recutting with *Sma*I. A double digest with *Sma*I and *Xho*I showed a sharp, clear visible band of 185 bp and a set of fragments forming at least five bands that were heterogeneous with respect to their intensity and sharpness (data not shown). For subsequent restriction analysis, end-labeled fragments were used as described above.

Figure 2 shows an autoradiogram of a 7.5% polyacrylamide gel with fragments resulting from recutting of end-

labeled *Eco*RI-D with *Sma*I, *Xho*I, and *Asu*I. The approximate position of the *Hin*dIII, *Sma*I, and *Xho*I sites was known (7, 10, 21). Additional data that had been published while this work was in progress (4) were helpful in establishing additional restriction maps. For a determination of restriction sites to the left of the *pac* cut sites, a fragment of P22 DNA formed by *Hin*dIII and *Xho*I (sites 1 and 9 in Fig. 3) was used (due to terminal redundancy, this fragment is present at the right end of mature DNA in all particles). Most of this analysis was performed by recutting of 5'-end-labeled fragments and analysis of cleavage products on autoradiographs of polyacrylamide gels.

Figure 3 shows the restriction map of the *Hin*dIII-*Xho*I fragment, including the cutting region. The mapping results for the end sites agree well with those published by Casjens and Huang (4). Following a proposal of these authors, I shall differentiate between the position of the endonucleolytic cuts (*end* sites) and the possibly nonidentical recognition (*pac*) signals. End region 6 was not analyzed in this work. It is very rarely attacked and is seen as a faint band in Fig. 2. These ends could not be reisolated for the type of analysis to be described. In the sequence (see Fig. 5), the approximate position of this region is given by a broken line.

Determination of 5'-terminal nucleotides. The clustering of cutting in specific end regions can be explained by at least two alternate possibilities. (i) The cutting activity has a specificity for a certain sequence which occurs several times at *pac*. Variations in these sequences result in different affinities (identity of *pac* and *end* signals). (ii) Proteins engaged in packaging initiation recognize one specific sequence at *pac*, and the cutting of DNA occurs in a separate process which is less precise but depends on recognition.

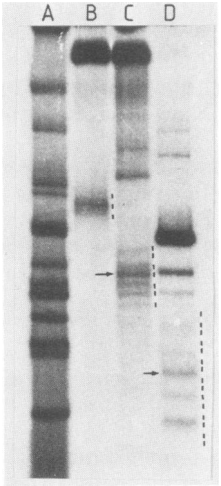

FIG. 2. Analysis of the *pac* ends of *Eco*RI-D. *Eco*RI-D was end labeled as described in the text. Restriction enzyme digests were run on a 7.5% polyacrylamide gel, and the bands were visualized by autoradiography. The arrows point to the band generated by the most frequently attacked cutting region (D3). The dashed line to the right of a lane indicates the approximate region of the *pac* end distribution. In addition to the *pac* ends and the prominent *Eco*RI end (the protruding 5' ends of this site are more efficiently labeled than the blunt *pac* ends), some contaminating bands of undefined origin are visible above the set of *pac* ends in lanes C and D. Lane A, λ dv1. *Hae*III was the marker dye. Fragment lengths (in bp) are as follows: 1,490, 1,310, 880, 534, 460, 340, 290, 270, 215, 210 (two bands), 175, 137, 123, and 82. Lane B, *Eco*RI-D-*Sma*I. Lane C, *Eco*RI-D-*Xho*I. Lane D, *Eco*RI-D-*Asu*I (*Sau*96I).

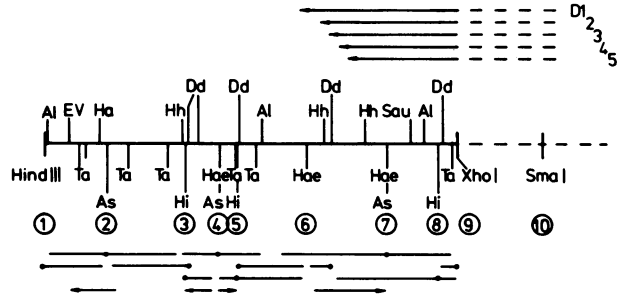

FIG. 3. Restriction map and sequencing strategy. The 5' to 3' direction of sequence readings is indicated from the site of end labeling marked by a dot. Arrows indicate length and direction of sequence readings after 3'-end labeling. The position of *end* sites analyzed by mapping the *pac* ends of *Eco*RI-D are shown above the map (arrowheads). Abbreviations used for restriction enzymes are as follows: Al, *Alu*I; As, *Asu*I (*Sau*96I); Dd, *Dde*I; EV, *Eco*V; Ha, *Hpa*II; Hae, *Hae*III; Hh, *Hha*I; Hi, *Hin*fI; Sau, *Sau*3a; and Ta, *Taq*I.

For a further analysis of cutting specificity, the 5'-terminal nucleotides of *Eco*RI-D were determined. Approximately 5 pmol of *Eco*RI-D were cut with *Sma*I and end labeled with kinase as described above. The fragments were recut with *Xho*I and separated on a 7.5% polyacrylamide gel. The bands containing the *Sma*I end (185 bp) and five bands with the in vivo-generated ends at *pac* (D1 to D5 in Fig. 3 and Table 1) were cut out after ethidium bromide staining. Fragment isolation by elution and the quantitative evaluation of 5'-end nucleotides are described above.

Table 1 shows the distribution of end nucleotides in the different bands. The *Sma*I-6 end serves as an internal control. In contrast, in bands D1 to D5, each end nucleotide is represented, and no band contains only one of them. The bands contain various frequencies of end nucleotides of which the quantitative relation might be influenced by the labeling specificity of kinase. Thus, the initiating cutting reaction does not occur at precisely defined positions.

From the rate of end labeling and other observations, it may be concluded that cutting generates blunt ends as follows. (i) The sum of Čerenkov counts in all ends generated in vivo is of the same order of magnitude as in the blunt *Sma*I-6 end, which is inefficiently labeled because of its sequence. Due to the procedures of labeling, recutting, and isolation of bands, this end is present in the analysis in an amount equimolar to the sum of *pac* ends. (ii) The amount of label in the different bands corresponds to their intensity by ethidium bromide staining and is comparable to the results of Casjens and Huang (4), which were obtained by a different method.

Sequence of the *pac* region and the position and frequency of cutting within this sequence. It was hoped that sequence data and the exact cutting positions within this sequence might suggest possible recognition sites for packaging initiation, cutting, or both. The sequence of the *Hin*dIII-*Xho*I fragment was established as described above. The sequencing strategy is shown in Fig. 3. Sequence information is from both strands or at least two gels giving unequivocal readings. The sequence is presented below (see Fig. 5). Numbering of bases was chosen to avoid discrepancies with publishing restriction maps (4, 10). The central G of the *Dde*I site, which is located at the most frequently attacked cutting region for packaging initiation, was given the number 1. Statistically significant repeats which might indicate poten-

TABLE 1. End nucleotides of *pac* ends

Base of 5′-end nucleotide	% Distribution in bands analyzed						
	D1	D2	D3	D4	D5	Sum	*Sma*I-6
C	29.2	9.3	14.6	2.7	23.4	14.7	2.9
A	39.0	66.6	38.2	45.7	46.1	45.3	2.2
G	9.4	8.9	4.6	37.3	19.4	13.5	91.7
T	22.4	15.2	42.6	14.3	11.1	26.5	3.3
Čerencov counts of isolated bands (% distribution)	23,200 (4.8)	85,800 (17.8)	201,500 (41.7)	78,800 (16.3)	93,400 (19.4)	482,700	330,000

tial elements of (recognition) signal structures or regulatory secondary structures are indicated.

The exact position of cutting sites was determined as follows. *Eco*RI-D was digested with *Asu*I, and the fragments were 3′ end labeled at the *Asu*I sites with [^{32}P]dGTP and the large fragment of DNA polymerase I. (A minor fraction of the *pac* ends might be labeled under these conditions when recessed 3′ ends [with a cytosine in the opposite strand] are present or formed by the "proofreading" 3′- to 5′-exonuclease function of polymerase.) The fragments were separated on a 7.5% polyacrylamide gel from which the ladder of heterogeneous fragments containing the *pac* ends was isolated. This mixture of fragments was loaded on a sequencing gel together with the four base-specific reactions of a fragment identically labeled at the same *Asu*I site (7) (*Asu*I fragment 4 to 7, recut with *Hin*fI [Fig. 3]). The position of a band containing an in vitro-generated cut at its 5′ end will then appear at the position of the base in the sequence ladder preceding the cut, due to the chemical cleavage method inherent in the Maxam and Gilbert sequencing technique.

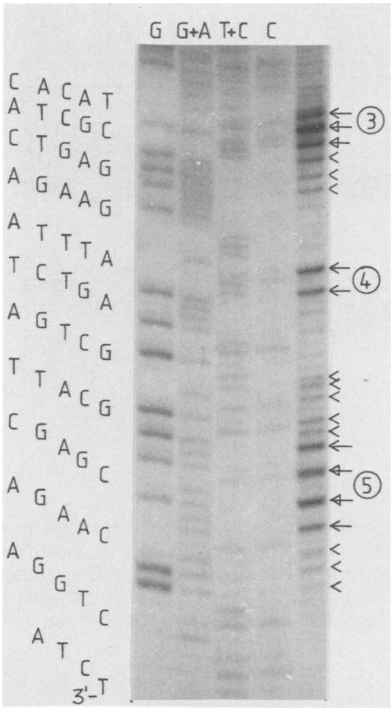

FIG. 4. *pac* ends analyzed on a sequencing gel. The sequence is read from the 3′ to 5′ direction in the upper strand. The cutting positions of regions 5, 4, and 3 can be analyzed on this gel. *Asu*I site 7 in Fig. 3 was 3′ end labeled for this analysis, as described in the text.

Figure 4 shows a sequencing gel together with the *pac* fragments containing the clustered cut positions. The position and relative frequency of cuts are indicated in Fig. 5 above the upper strand. In cutting region 1, the exact positions could not be determined accurately, and in region 6, the ends were not analyzed in this study.

The most striking result of this analysis is the clustering of cut positions around central cut positions. The central and most prominent cuts are always 2 bp apart from each other. This seems not to be an artifact of the experimental procedure, as *pac* fragments labeled at their 5′ end and also recut with *Asu*I show the same distribution pattern (at a slightly different position in the gel due to the labeling of the opposite strand [data not shown]).

Localization of gene *3* on the *Hin*dIII-*Xho*I fragment. During the course of this work, the *Hin*dIII-*Xho*I fragment was cloned in pBR322 (pGM1000; see above). For all *3*$^-$ amber mutants, including a set of newly isolated mutants (see above), marker rescue with pGM1000 could be demonstrated. In addition, complementation of the gene *3* product after infection of a *rec*$^-$ host with an *erf*$^-$ *3*$^-$ double mutant was observed (data not shown). The distribution of initiator and terminator codons with an open reading frame of suitable length allows the tentative assignment of gene *3* in the sequence with ambiguity concerning the initiating codon.

The putative gene *3* reading frame has two possible initiator codons (positions −263 and −269 [Fig. 5]). Overlapping segments are homologous to the 5′ end of 16S RNA (Shine-Dalgarno tracts [37]) in this region. A comparison to known ribosome-binding sites (39) does not allow us to unambigously assign the initiating codon (from this analysis, a transcription initiation at the second AUG might be predicted). The reading frame is terminated by two terminator codons following each other: position 224, opal; and position 227, ochre. From the first initiator codon, the reading frame codes for a 164-amino-acid protein of 18,871 daltons. This value agrees well with data from sodium dodecyl sulfate protein gels for the molecular size of the gene *3* product (17 kilodaltons [50]).

The sequence proximal to gene *3* contains terminator codons in all reading frames, the most proximal to the *Hin*dIII site being at positions −580 (I), −510 (II), and −462 (III, gene *3* reading frame). Thus, at least 200 noncoding bp will separate gene *3* from the preceding one, which may be gene *19*, the gene coding for P22 endolysin. Whether the intercistronic region might serve any function, e.g., in gene expression, remains unknown. No statement is possible on reading frames distal to gene *3*. The end sites for packaging initiation map within the gene. Also, the *pac* site which is

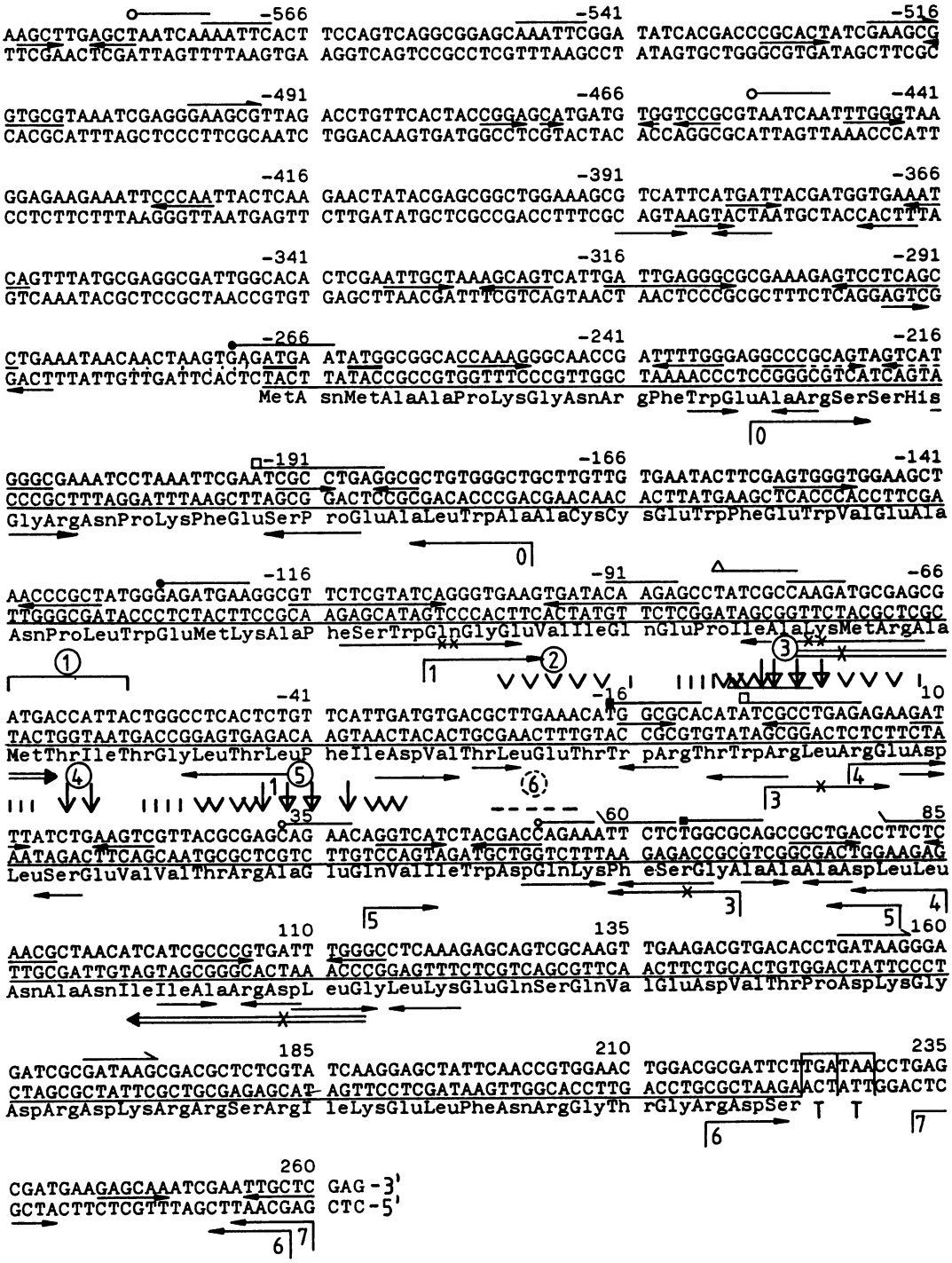

FIG. 5. Sequence of the *Hin*dIII-*Xho*I fragment of P22 DNA (Fig. 3). Direct repeats (given different symbols) are indicated above the sequence; inverted repeats (between and below the sequence) point to their symmetry center. The indication of repeats is based solely on their statistical significance. In general, the probability for a random occurrence of the marked repeats is below 0.02. Various symbols indicating the (increasing) relative frequency of cutting at the marked positions included the following: (|), (⋁), (↓), and (⏚). Numbers in circles refer to the mapping data of *pac* ends in Fig. 3. The amino acid sequence of 3 protein is given below its coding sequence.

recognized by gene *3* product for packaging initiation is located within this gene (23; the recognition site can now be localized between position 1 and 114 [unpublished data]). A comparison of gene *3* with the analogous lambda gene *Nul* will be discussed below.

DISCUSSION

Following earlier interpretations, the cut sites observed in my experiments could be the product of a site-specific nuclease with a limited precision, if at least five similar *pac*

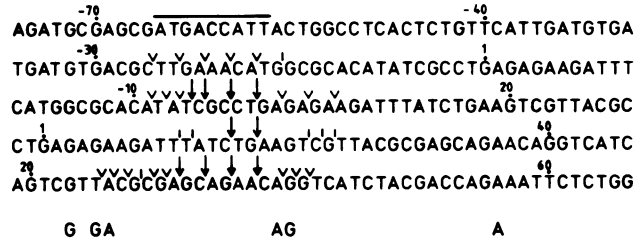

FIG. 6. Alignment of the end regions 1 (first line) to 5. End region 1 was aligned to give a maximum number of matching positions with the other end regions. For the alignment of end regions 2 to 5, the main cutting positions were used. For symbols and numbers, see the legend to Fig. 5.

sites are present with sequence variations affecting recognition, cleavage, or both. The five main end regions are aligned in Fig. 6 with respect to the main cutting positions. Although some similarities exist, especially between end regions 4 and 5, it cannot be concluded from this comparison that the special distribution of cut sites is an inherent feature of the sequence itself, as predicted by the hypothesis. The "consensus positions" given at the bottom of Fig. 6 are most certainly accidentally matching bases (this region is rich in purine bases), suggesting that the *end* sites are not likely to play a role in the step of "substrate recognition" (12).

Alternatively, recognition of a singular *pac* site might be a prerequisite for a cleavage reaction. This recognition could act as an allosteric effector for the nuclease activity of the protein(s) engaged in the process of substrate recognition at *pac* and the DNA cleavage. By analogy to lambda terminology, the proteins involved should be termed terminase. The distribution of *end* sites is the consequence of the susceptibility of phosphodiester bonds to the nucleolytic sites of activated terminase. The main cut sites are separated by 2 bp, and the end regions are 18 to 20 bp apart from each other (40 bp in the case of regions 1 and 2). This peculiar distribution of cut positions is not easily compatible with movement of free terminase complex along the DNA. A fixed position at one site of the DNA (possibly, the *pac* site), however, could impose stringent steric limitations to nucleolytic activity of a second domain at another position on the DNA. A possible translocation of DNA with respect to the nucleolytic domain may result in the spacing of cutting regions of which the specific sequence possibly influences cutting probability.

There is a formal similarity of these observations to the mechanism of restriction nucleases of type I. These enzymes apparently recognize a site and induce DNA cleaving at a distant region after DNA translocation while remaining bound to the recognition site (50; the property of acting as a DNA-dependent ATPase in DNA translocation is certainly another feature that is attractive as a model for terminase action in DNA packaging). In addition, this interpretation seems to be in accordance with observations on the properties of lambda terminase (see below).

In other bacteriophages, the heterogeneous ends formed by DNA maturation demonstrate that the cleavage reaction is not a sequence-specific event (11, 17). Most clearly in the case of phage Mu 1, DNA recognition depends on a signal at the left end of the Mu genome. Starting at a distance of 56 bp from the end of the Mu genome within the linked heterogeneous bacterial sequences, the cleavage sites are scattered in a regular spacing of ca. 11 bp (17; J. Teifel, unpublished data).

Of all bacteriophages relying on headful mechanism for the length determination of mature DNA, another allosteric effector for a second cleavage reaction must be employed that is a signal from packaging density of the prohead. In sequential packaging of P22, the second and all subsequent cleavages must be triggered by such a headful signal, possibly affecting the nucleolytic activity of the same protein that made the first cut. The headful cuts, however, are scattered over a longer region (unpublished data). Alternatively, both cuts—the initiating first and all subsequent headful cuts—could occur in a concerted way after head filling. In the case of P22, however, it has been demonstrated that cutting in the *pac* region may occur without concomitant DNA packaging in mutants defective in prohead assembly (26). It is not known whether this prohead-independent *pac* cleavage generates the same molecular ends as does normal packaging initiation. The argument of biological expediency demands a strict regulation of such a process in P22 DNA maturation to prevent the frequent production of DNA molecules cut between their *pac* sites. These will not contain the terminal redundancy necessary for recombinative ring formation after infection, which is a prerequisite for lysogenization or lytic growth. Which parameters of DNA substrate or terminase might be critical for imposing a limitation on initiating cleavages can only be speculated upon at the moment.

In P22, the products of genes *2* and *3* are components of terminase (26). Their specific interaction has been demonstrated in vitro, and the specificity for *pac* recognition can be assigned to 3 protein (22, 31, 33, 44). A nucleolytic activity of this complex and a specific interaction with components of the prohead has not been demonstrated directly, although it may be concluded from the in vitro packaging properties of the complex (32, 41). The property of being stabilized by ATP could be indicative of an ATPase function of the complex (31).

In the amino acid sequence of protein 3, the hydrophilic carboxy terminus with its predominating basic amino acids may be noted (from amino acid position 140 to the carboxy terminus of protein 3). Basic amino acids also predominate in the first 25 positions. These basic sections are compensated by acidic regions more central in the primary sequence. Our results demonstrate that this gene is disrupted by packaging initiation. The recognition site *pac* is also part of this sequence. The relevant nucleotide sequence and the question of whether it may be separated completely from the cut sites remains to be elucidated.

From data cited above, which point to a functional analogy of both proteins, a comparison to Nu1 protein of phage lambda makes sense. The complete nucleotide sequence of phage lambda is known, and from this the protein sequence of *Nu1* can be deduced (35). The comparison of the amino acid sequences shows only little homology. In Nu1 protein, the most hydrophilic regions are in the center of the primary sequence around amino acids 60 and 90. A search for alignments yields no more than 12% of homologous positions. Sections of the sequences show better homology (20% for regions of 50 amino acids). These homologies generally do not coincide with structural alignments that were tested by computer-assisted secondary structure predictions with the rules of Chou and Fasman (8; data not shown). For a statement on the significance of such homologies, more information about structure and functional domains of these proteins must be available.

Recently, the sequences necessary for packaging in vitro and in vivo, *cos* cutting, and packaging specificity of phage

lambda have been minimalized in experiments with cosmids (plasmids with *cos* sites) as substrates (16, 20, 28, 29). The results imply that the *cos* region contains two separate sites, one, *cosN*, in which the staggered nicks are introduced and a terminase-binding site, *cosB*. An interesting feature concerning the P22 model was the observation that the lambda terminase-binding site seems to stimulate in vitro *cos* cutting and packaging only when located at the correct distance to the *cosN* site (29).

A comparison of the DNA sequence around *cos* to the *pac* region of P22 does not reveal any significant sequence homologies or similarities in sequence organization with respect to direct and inverted repeats. Also, *cos* maps in a large intercistronic region preceding gene *Nu1*, whereas the initiating cuts in P22 disrupt gene *3*. In contrast to general similarities in map location, molecular weight, and their function as a subunit of terminase, this comparison stresses the evolutionary and functional divergence of the proteins and the process of DNA packaging in both phages.

ACKNOWLEDGMENTS

I thank F. Sanger and S. Casjens for communicating results before publication and U. Hamker for help in preparation of this manuscript.

This work was supported by the Deutsche Forschungsgemeinschaft.

LITERATURE CITED

1. **Adams, M. B., M. Hayden, and S. Casjens.** 1983. On the sequential packaging of bacteriophage P22 DNA. J. Virol. **46**:673–677.
2. **Becker, A., and M. Gold.** 1978. Enzymatic breakage of the cohesive end site of phage lambda DNA: terminase (ter) reaction. Proc. Natl. Acad. Sci. U.S.A. **75**:4199–4203.
3. **Botstein, D., C. H. Waddell, and J. King.** 1973. Mechanism of head assembly and DNA encapsulation in *Salmonella* phage P22. I. Genes, proteins, structures and DNA maturation. J. Mol. Biol. **80**:669–695.
4. **Casjens, S., and W. H. Huang.** 1982. Initiation of sequential packaging of bacteriophage P22 DNA. J. Mol. Biol. **157**: 287–298.
5. **Chan, R. K., and D. Botstein.** 1972. Genetics of bacteriophage P22. I. Isolation of prophage deletions which affect immunity to superinfection. Virology **49**:257–267.
6. **Chelala, C. A., and P. Margolin.** 1974. Effects of deletions on cotransduction linkage in *Salmonella typhimurium*: evidence that bacterial chromosome deletions affect the formation of transducing DNA fragments. Mol. Gen. Genet. **131**:97–112.
7. **Chisholm, R. L., R. J. Deans, E. N. Jackson, D. A. Jackson, and J. E. Rutila.** 1980. A physical gene map of the bacteriophage P22 late region: genetic analysis of cloned fragments of P22 DNA. Virology **102**:172–189.
8. **Chou, P. Y., and G. D. Fasman.** 1974. Conformational parameters for aminoacids in helical, β-sheet and random coil regions calculated from proteins. Biochemistry **13**:211–222.
9. **Clewell, D. B., and D. R. Helinski.** 1972. Effects of growth conditions on the formation of the relaxation complex of supercoiled ColE1 deoxyribonucleic acid and protein in *Escherichia coli*. J. Bacteriol. **110**:1135–1142.
10. **Deans, R. J., and E. N. Jackson.** 1979. Restriction endonuclease *Hin*dIII cleavage site map of bacteriophage P22. Virology **95**:359–372.
11. **Deichelbohrer, J., W. Messer, and T. A. Trautner.** 1982. Genome of *Bacillus subtilis* bacteriophage SPP1: structure and nucleotide sequence of *pac*, the origin of DNA packaging. J. Virol. **42**:83–90.
12. **Earnshaw, W. C., and S. R. Casjens.** 1980. DNA packaging by the double stranded DNA bacteriophages. Cell **21**:319–331.
13. **Feiss, M., and A. Bublitz.** 1975. Polarized packaging of bacteriophage lambda chromosomes. J. Mol. Biol. **94**:583–594.
14. **Feiss, M., R. A. Fisher, D. A. Siegele, B. P. Nichols, and J. E. Donelson.** 1979. Packaging of the bacteriophage lambda chromosome: a role for base sequences outside *cos*. Virology **92**:56–67.
15. **Feiss, M., and D. A. Siegele.** 1979. Packaging of the bacteriophage lambda chromosome: dependence of *cos* cleavage on chromosome length. Virology **92**:190–200.
16. **Feiss, M., J. Kobayashi, and W. Widner.** 1983. Separate sites for binding of bacteriophage λ DNA by terminase. Proc. Natl. Acad. Sci. U.S.A. **80**:955–959.
17. **George, M., and A. I. Bukhari.** 1981. Heterogenous host DNA attached to the left end of mature bacteriophage Mu DNA. Nature (London) **292**:175–176.
18. **Godson, G. M., and D. Vapnek.** 1973. A simple method of preparing large amounts of ϕX174 RFI supercoiled DNA. Biochim. Biophys. Acta **299**:516–520.
19. **Gold, M., and A. Becker.** 1983. The bacteriophage λ terminase. J. Biol. Chem. **258**:14619–14625.
20. **Hohn, B.** 1983. DNA sequences necessary for packaging of bacteriophage λ DNA. Proc. Natl. Acad. Sci. U.S.A. **80**:7456–7460.
21. **Jackson, E. N., D. A. Jackson, and R. J. Deans.** 1978. *Eco* RI analysis of bacteriophage P22 DNA packaging. J. Mol. Biol. **118**:365–388.
22. **Jackson, E. N., F. Laski, and C. Andres.** 1982. Bacteriophage P22 mutants that alter the specificity of DNA packaging. J. Mol. Biol. **154**:551–563.
23. **Kufer, B., H. Backhaus, and H. Schmieger.** 1982. The packaging initiation site of phage P22. Analysis of packaging events by transduction. Mol. Gen. Genet. **187**:510–515.
24. **Kwoh, D. Y., and J. Kemper.** 1978. Bacteriophage P22-mediated specialized transduction in *Salmonella typhimurium*: identification of different types of specialized transducing particles. J. Virol. **27**:535–550.
25. **Larson, R., and J. Messing.** 1982. Apple II software for M13 shotgun DNA sequencing. Nucleic Acids Res. **10**:39–49.
26. **Laski, F., and E. N. Jackson.** 1982. Maturation cleavage of P22 DNA concatemers in absence of DNA packaging. J. Mol. Biol. **154**:565–579.
27. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
28. **Miwa, T., and K. Matsubara.** 1982. Identification of sequences necessary for packaging DNA into lambda phage heads. Gene **20**:267–279.
29. **Miwa, T., and K. Matsubara.** 1983. Lambda phage DNA sequences affecting the packaging process. Gene **24**:199–206.
30. **Mousset, S., and R. Thomas.** 1969. *Ter*, a function which generates the ends of the mature lambda chromosome. Nature (London) **221**:242–244.
31. **Poteete, A. R., and D. Botstein.** 1979. Purification and properties of proteins essential to DNA encapsulation by phage P22. Virology **95**:565–573.
32. **Poteete, A. R., V. Jarvik, and D. Botstein.** 1979. Encapsulation of phage P22 DNA in vitro. Virology **95**:550–564.
33. **Raj, A. S., A. Y. Raj, and H. Schmieger.** 1974. Phage genes involved in the formation of generalized transducing particles in *Salmonella* phage P22. Mol. Gen. Genet. **135**:175–184.
34. **Rhoades, M., L. A. MacHattie, and C. A. Thomas, Jr.** 1968. The P22 bacteriophage molecule I. The mature form. J. Mol. Biol. **37**:21–40.
35. **Sanger, F., A. R. Coulson, G. F. Hong, D. F. Hill, and G. B. Petersen.** 1982. Nucleotide sequence of bacteriophage lambda DNA. J. Mol. Biol. **162**:729–773.
36. **Schmieger, H., and H. Backhaus.** 1976. Altered cotransduction frequencies exhibited by *HT*-mutants of *Salmonella*-phage P22. Mol. Gen. Genet. **143**:307–309.
37. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. U.S.A. **71**:1342–1346.
38. **Smith, H. O.** 1968. Defective phage formation by lysogens in integration deficient phage P22 mutants. Virology **34**:203–223.
39. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Charac-

terization of translational initiation sites in *E. coli*. Nucleic Acids Res. **10**:2971–2995.
40. **Streisinger, G., J. Emrich, and M. M. Stahl.** 1967. Chromosome structure in phage T4. III. Terminal redundancy and length determination. Proc. Natl. Acad. Sci. U.S.A. **57**:292–295.
41. **Strobel, E., W. Behnisch, and H. Schmieger.** 1984. *In vitro* packaging of mature phage DNA by *Salmonella* phage P22. Virology **133**:158–165.
42. **Sumner-Smith, M., A. Becker, and M. Gold.** 1981. DNA packaging in the lambdoid phages: the role of lambda genes *Nul* and *A*. Virology **111**:642–646.
43. **Susskind, M. M., and D. Botstein.** 1978. Molecular genetics of bacteriophage P22. Microbiol. Rev. **42**:385–413.
44. **Tye, B. K.** 1976. A mutant of phage P22 with randomly permuted DNA. J. Mol. Biol. **100**:421–426.
45. **Type, B. K., R. K. Chan, and D. Botstein.** 1974. Packaging of an oversize transducing genome by *Salmonella* phage P22. J. Mol. Biol. **85**:485–500.
46. **Tye, B. K., J. A. Hubermann, and D. Botstein.** 1974. Non random circular permutation of phage P22 DNA. J. Mol. Biol. **85**:501–527.
47. **Weaver, S., and M. Levine.** 1978. Replication in situ and DNA encapsulation following induction of an excision-defective lysogen of *Salmonella* phage P22. J. Mol. Biol. **118**:389–411.
48. **Wu, R., and E. Taylor.** 1971. Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage lambda DNA. J. Mol. Biol. **57**:441–511.
49. **Youderian, P., and M. M. Susskind.** 1980. Identification of the products of bacteriophage P22 genes, including a new late gene. Virology **107**:258–269.
50. **Yuan, R.** 1981. Structure and mechanism of multifunctional restriction endonucleases. Annu. Rev. Biochem. **50**:285–315.