Volume 16 Number 5 1988 Nucleic Acids Research

The interaction of *Escherichia coli* integration host factor with the cohesive end sites of phages λ and 21

Winnie Xin and Michael Feiss

Department of Microbiology, College of Medicine, and Interdepartmental Genetics Ph.D. Program, University of Iowa, Iowa City, IA 52242, USA

Received November 30, 1987; Revised and Accepted January 25, 1988

ABSTRACT

The interaction of *E. coli* integration host factor (IHF) with the cohesive end sites (*cos*'s) of phages λ and 21 has been studied by the DNAase I footprinting technique. Six potential sites in *cos* λ differ from the consensus IHF binding sequence by 1 to 3 base pairs. Of the six, one site, I1, binds IHF strongly. The I1 segment protected by IHF contains two sequences that closely match the IHF consensus binding sequence. Another site, I2, binds IHF moderately well, and three sites: IO', I3 and I4 bind IHF very weakly. The IO site does not bind IHF under the conditions used here. In phage 21 the DNA segment extending to the right from the cohesive ends, which contains three potential IHF binding sites, was examined. Two sites bind IHF well; I1, the 21 analogue of one of the λ I1 sites, and IO, a site not analogous to a λ site. The third 21 site, I2, binds IHF moderately well, as does the analogous I2 site in λ. The significance of the results for λ DNA packaging is discussed.

INTRODUCTION

A virus depends on its host for many processes, and some cases involve specific interactions between virus and host elements. An example of a specific interaction is the interaction of phage λ DNA with *E. coli* integration host factor (IHF). IHF is a small, heterodimeric, site-specific, DNA binding protein. The alpha subunit (MW=11,224) is encoded by the *himA* gene (1), and the beta subunit (MW=10,581) is the product of the *hip* gene (2). IHF is required for site-specific integration of λ DNA into the host chromosome. IHF binds specifically to several sites within λ *attP* (3). More recent studies indicate IHF acts at *attP* to assist in the formation of a nucleosome-like structure involving IHF, λ integrase, and *attP* (4, 5).

IHF also acts in the formation of a multiprotein complex at the origin of replication of pSC101 (6). In addition to acting in the formation of complex DNA:protein structures, IHF functions in gene regulation. For example, IHF binds to a site at the beginning of the λ *cII* gene, and enhances

cII expression at a posttranscriptional level, perhaps by influencing mRNA dissociation from the DNA (7). Study of a number of IHF binding sites yielded the following consensus sequence: 5'-PyNPyAANNNNTTGAT(A/T)-3' (8).

IHF has been implicated in the DNA packaging of the evolutionarily-related phages λ and 21. The enzyme terminase plays a central role in DNA packaging in the lambdoid phages (9). The λ terminase is a heterooligomer of two subunits: the small subunit gpNu1 and the large subunit gpA, the gene products of the *Nu1* and *A* genes, respectively. Terminase is a multifunctional protein that binds λ DNA at *cos*, introduces the staggered nicks to generate the cohesive ends, and binds the prohead to generate the DNA:terminase:prohead complex that leads to packaging of the DNA (10). The cohesive end site, *cos*, is bipartite, consisting of *cosN*, the nicking site, and *cosB*, the site where terminase initially binds λ DNA. The structure of *cos* is shown in Fig. 1; the *Nu1* and *A* genes are just to the right of *cos*.

Phage 21 produces a terminase that shares about 60% sequence identity with λ terminase (11); the genes *1* and *2* produce the small subunit, gp1, and the large subunit, gp2, respectively. IHF is essential for an early, terminase-dependent step in 21 DNA packaging (12). The IHF requirement can be bypassed, because 21 *her* mutants can be found that are able to grow in cells lacking IHF. The *her* mutations affect the amino-terminal part of gp1, the part of gp1 that is known from other studies to be involved in DNA binding. These results suggest IHF helps 21 terminase to bind at *cos*.

IHF plays an accessory role in λ DNA packaging. Bear et al. (13) described a mutant, λ *cos*154, that is dependent on IHF for DNA packaging. The mutation occurs in a sequence (called R) that is repeated several times in *cos* and which was postulated to be the terminase binding site. Because λ *cos*154 is IHF-dependent, Bear et al., suggested that λ terminase normally is able to bind to *cos* well enough for growth, but that when one binding site is inactive terminase requires IHF to form a functional complex with *cos*. With this model in mind, Bear et al. searched *cos* for the IHF consensus binding sequence, and found several (Fig. 1). We have identified two additional sites, I0 and I0' (Fig. 1). In 21, a similar search (11) found several potential IHF binding sequences (Fig. 2).

We have done DNA footprinting experiments to examine these putative binding sites as a first step in understanding the role of IHF in viral DNA packaging.

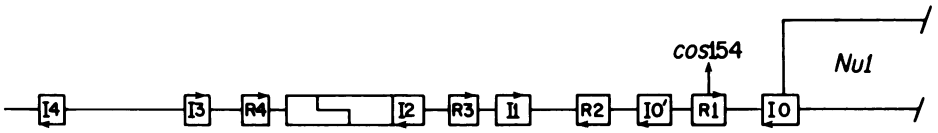

Figure 1: Schematic diagram of presumptive terminase binding sites and IHF binding sites (I) at λ *cos*. The center structure represents the staggered nicks that are introduced by terminase at λ *cosN*. Each R site contains a 16 bp sequence postulated to be the terminase binding site. The I sites resemble the IHF binding consensus sequence. *cos*154 is a single base transition mutation in one of the R sites. The IO site partially overlaps the leftmost λ gene, *Nul*.

MATERIALS AND METHODS

Plasmids

Plasmid pLW110 has a λ DNA segment extending from the *Bcl*I site at 47942 through *cosN* to an *Eco*RI site at 194 cloned into pBR322 digested at the *Bam*HI and *Eco*RI sites (14). Base pairs of λ and 21 are numbered as in Daniels et al. (15); the base at the 5' end of the left cohesive end is designated 1 and numbering proceeds rightward. The 5' to 3' strand extending to the right from position 1 is called the top strand, and the 3' to 5' strand is the bottom strand. Plasmid pSF1 has a λ segment extending from the *Hind*III site at 44141 through *cosN* to the *Bam*HI site at 5505 cloned into pBR322 (16).

Plasmid pBA2 is a derivative of pACYC184 (17) containing the *cos*/terminase segment from λ-21 hybrid 51 (18). λ-21 hybrid 51 is a recombinant from a λ vs. 21 cross, such that the segment from position 1, at the left cohesive end, to 490, a site within the *I* gene, is derived from 21 and the remainder of the chromosome is derived from λ. pBA2 was constructed by ligating the λ-21 hybrid 51 segment extending from the *Hind*III site at 44141 through *cos* to the *Bam*HI

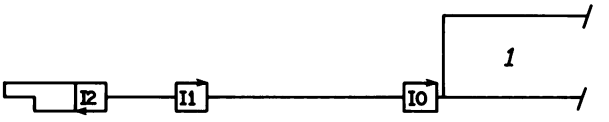

Figure 2: Schematic diagram of the phage 21 *cos* segment to the right of 21 *cosN*. The left cohesive end is shown at the left of the diagram. The I sites are the possible IHF binding sites at phage 21 *cos*. The leftmost gene on the 21 chromosome, gene *1*, follows the IO site. The postulated terminase binding sites (R sequences) are not shown in this diagram. The 21 R sequences are at the same positions as the λ R sites (Fig. 1), but the DNA sequence of the 21 repeats is different from the sequence of the λ R sites.

site at 5505 into *Hin*dIII, *Bam*HI-digested pACYC184. pBA4, a shortened derivative of pBA2, was made by deleting the segment from the λ *Bgl*I site at 2666 to the *Bgl*I site of pACYC184. Plasmids were prepared in large quantity by the method of Bestwick et al. (19).

3'-labeling

The procedure for DNA-labeling at a 3' end by filling in with DNA polymerase I (Klenow fragment, Boehringer-Mannheim) and an appropriate α-^{32}P-dNTP or α-^{32}P-ddNTP (Amersham) was as described in Maniatis et al. (20). Restriction enzymes were obtained from New England Biolabs, Bethesda Research Laboratories or Boehringer-Mannheim. The labeled fragment was purified from a 6% polyacrylamide gel by electroelution and ethanol precipitations.

5'-labeling

Plasmid DNA was digested and treated with bacterial alkaline phosphatase to remove 5' terminal phosphates. Then the digest mix was extracted with phenol and ethanol precipitated. The sample was then labeled with T4 DNA polynucleotide kinase (New England Biolabs) and γ-^{32}P-ATP (Amersham). A secondary restriction enzyme digestion was performed after the labeling. The labeled fragment was purified from a 6% polyacrylamide gel by electroelution and ethanol precipitations.

DNA sequencing

The Maxam and Gilbert DNA sequencing method (21) was used for sequencing end-labeled DNA fragments.

Footprinting experiments

The method of Galas and Schmitz (22), as modified for IHF by Craig and Nash (3) was used. Purified end-labeled DNA fragment was incubated with various amounts of IHF in 100 µl binding buffer. Binding buffer contained 52 mM Tris-HCl (pH 7.4), 70 mM KCl, 10% (v/v) glycerol, 1.1 mM EDTA, 1.0 mM β-mercaptoethanol, 7.0 mM $MgCl_2$, 3.0 mM $CaCl_2$, and 200 µg/ml bovine serum albumin. After 20 min at 25°C, 5.0 µl of 5 µg/ml pancreatic DNAase I (Worthington) in the binding reaction buffer was added and incubation continued for 30 sec. The reactions were then stopped by addition of 100 µl of stop solution containing 0.6 M NH_4Ac, 0.1 M EDTA, and 20 µg/ml sonicated calf thymus DNA. The sample was ethanol precipitated and washed and loaded onto an 8% sequencing gel.

Purified IHF is a gift from Dr. Howard Nash from National Institute of Mental Health. Prior to use, IHF was diluted in 50 mM Tris-HCl (pH 7.4), 10% (v/v) glycerol, 800 mM KCl, and 2 mg/ml bovine serum albumin.

IHF binding consensus	1 2 3 4 5 6 7 8 $9^{10}11^{12}13^{14}$ 15 5'-Py N Py A A N N N N T T G A T (A/T)-3'	mis-matches	location
*cos*λ			
IO	C t T C C a t t g T T C A T T	3/10	197-183*
IO'	C a A A A a g c c T C G C T T	3/10	146-132
I1A	C a T A A c t t a A T G T T T	2/10	72-86
I1B	C t T A A t g t t T T T A T T	1/10	77-91
I2	C g G A A a a t t T T C A T A	2/10	44-30
I3	T g T G A t a t g T A G A T G	3/10	48439-48453
I4	T c G T A t t a a T T G A T C	3/10	48386-48372
*cos*ϕ21			
IO	G a T A A t g g g T T G T T T	2/10	172-186
I1	T a T A A c t t t T T G T T T	1/10	72-86
I2	T g A A A a a t t T T C A T A	2/10	44-30

Figure 3: Consensus sequence of IHF binding sites and the possible binding sites at λ *cos* and phage 21 *cos*. The IHF consensus is that of Gardner and Nash (8). Base pairs are numbered as according to Daniels et al. (15).

RESULTS

Six sequences that resemble the IHF consensus binding sequence are located in or near *cos*λ. Two of these sequences, I3 and I4, are to the left of *cosN* and four, IO, IO', I1 and I2, are to the right of *cosN* (Fig. 1). The I1 sequence actually contains two overlapping sequences, I1A and I1B, that closely match the IHF consensus binding sequence. In phage 21, between *cosN* and the *I* gene, there are three potential binding sites for IHF; designated IO, I1 and I2 (Fig. 2). Figure 3 is a comparison of these nine sequences with the IHF consensus binding sequence. We have examined the ability of IHF to protect these potential IHF binding sites against DNAase I attack.

Protection by IHF at the λ I1 and I2 sites

Plasmid pLW110 was cleaved at the *Ban*II site of pBR322 and the *Eco*RI site at 194 of λ, generating a 850 bp fragment. This fragment was 3'-labeled by filling in the *Eco*RI end, using Klenow fragment and α-^{32}P-ddATP, and then footprinted. Our footprinting data (Figure 4A) shows that IHF strongly protects the λ I1 site at all concentrations of IHF studied. At IHF concentration of 45 nM in the binding reaction, the same concentration used to study the λ *att* site (3) a 46-bp segment from 65 to 110 is protected. At 5-fold and 10-fold increased IHF concentrations, the pattern of protection is very similar to the protection at the 1X IHF concentration except that the nucleotides C at 72 and A at 94 show enhancement of DNAase I attack.

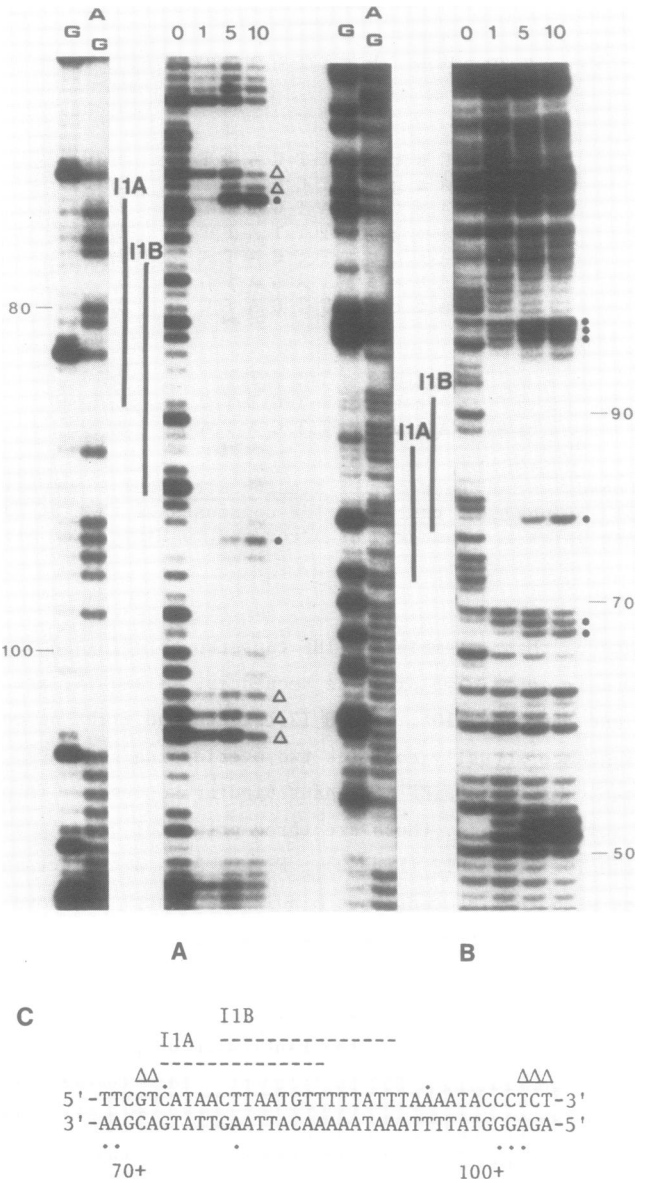

Figure 4: Footprint of IHF at the I1 site of λ. DNAase I digestion pattern of A) the top strand, and B) the bottom strand, of the DNA fragment containing the λ I1 site. For the top strand (A), an 850 bp fragment extending from the *Ban*II site of pBR322 through *cos* to the λ *Eco*RI site at 194 was used. The *Eco*RI end was 3'-labeled. For the bottom strand (B), a 223 bp fragment extending from the λ *Ppu*MI site at 48473 to the λ *Eco*RI site at 194 was used. The fragment was 3'-labeled at the *Ppu*MI site. Base pair positions are given

at the sides of the figure. The first two lanes are Maxam and Gilbert G and AG sequencing reactions. Lanes 0, 1, 5 and 10 show the DNAase I digestion pattern of the I1-containing fragment after incubation of the fragment with 0 nM, 45 nM, 225 nM and 450 nM IHF protein. C) Diagram of the two overlapping possible IHF binding sequences at the λ I1 site. 46 bases (from 65 to 110) are protected on the top strand and 30 bases (from 70 to 99) are protected on the bottom strand. Bases marked with triangles are bases in the protected region which are not well protected by IHF even at the high IHF concentrations. Bases with dot show enhanced DNAase I cleavage at high IHF concentrations.

To examine the protection of the opposite strand, a 223 bp fragment extending from the λ *Ppu*MI site at 48473 to the *Eco*RI site at 194 was isolated from a digest of pLW110. This fragment was 3'-labeled by filling in the *Ppu*MI end, using Klenow fragment and α-^{32}P-dGTP, and then footprinted. The data is shown in Figure 4B. At 45 nM IHF concentration, strong protection from 70 to 99 is detected. Enhancement of DNAase I attack of the A at 78 within the protected region and several bases on the margins of the protected region can be seen at 5-fold and 10-fold increased IHF concentrations.

The region protected by IHF at I1 site is relatively wide compared to the 15-bp IHF-binding consensus. By careful examination of DNA sequence at the I1 site, we found that the λ I1 site consists of two overlapping IHF-binding consensus sequences, named I1A and I1B, respectively (Figure 4C). The base pairs covered by both consensus sequences are from 72 to 91. The I1B is more centered than the I1A in the wide protected I1 region. The enhancement at the nucleotide C at 72 falls at the first nucleotide of the I1A consensus sequence. At this point, we do not know which of the two IHF-binding sequences or whether both consensus sequences are recognized and protected by IHF.

The I2 site was also examined in the gel shown in Fig. 5. The protection at I2 is rather weak. At higher IHF concentrations, the 30 bp segment extending from 26 to 55 on the bottom strand shows diminished cleavage and several enhancements. The 5'-AAAATTTT-3' sequence of I2 is not cleaved by DNAase I, so the extent of binding of this sequence by IHF cannot be discerned.

Results on protection of the top strand at I2 were obtained as follows. Plasmid pLW110 was cut with *Eco*RI at the λ site at 194 and with *Ban*II at the site in pBR322, generating a 850 bp fragment. The fragment was 3'-labeled at the *Eco*RI site and used for footprinting. At I2, a segment from 21 to 53 was protected at the higher IHF concentrations, as shown in Fig. 6.

There is no protection by IHF at I0 and weak protection of I0' of λ

A 444 bp fragment extending from the λ *Ppu*MI site at 48473 to the λ *Bgl*II site at 415 was isolated, 3'-labeled at the *Ppu*MI end, and footprinted.

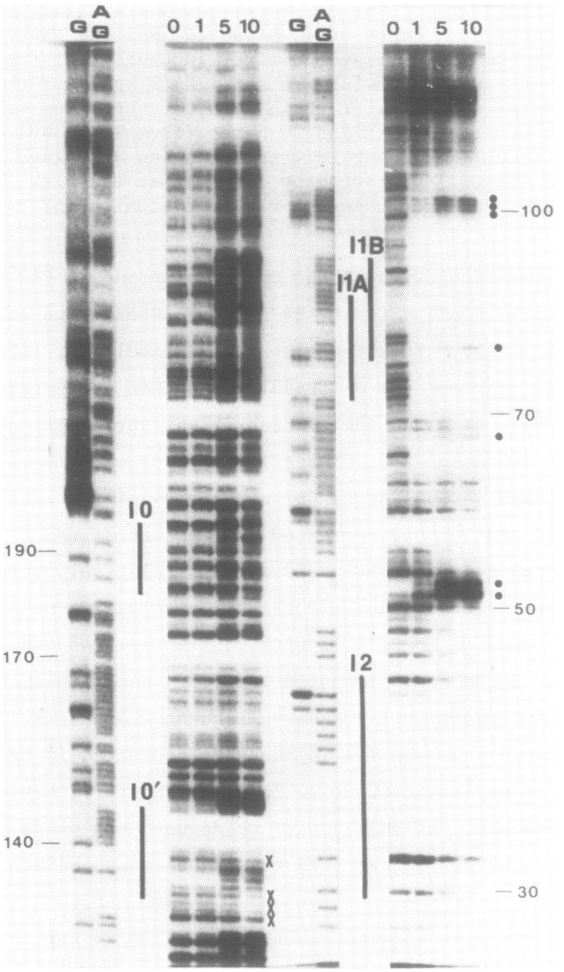

Figure 5: DNAase I footprints of λ I1, I2, I0 and I0' sites. The G and AG lanes are Maxam and Gilbert sequencing reactions of the DNA fragment containing the λ I1, I2, I0 and I0' sites. The 444 bp fragment used extends from the λ *Ppu*MI site at 48473 to the λ *Bgl*II site at 415. The fragment was 3'-labeled at the *Ppu*MI end. Base pair positions are given at the sides of the figure. Lanes 0, 1, 5 and 10 are the DNAase I digestion pattern of the same DNA fragment after incubation with 0 nM, 45 nM, 225 nM and 450 nM IHF, respectively. The two overlapping IHF binding sequences are also indicated. Symbols as in Fig. 4. Bases with x show reduced DNAase I cleavage at high IHF concentration.

The results, shown in Figure 5, show no significant protection of the bottom strand at the I0 site. There is weak diminution of cleavage at some base pairs in I0' at the 10-fold increased concentration of IHF. The top strand was also examined by using a fragment extending from the λ *Dra*I site at 90

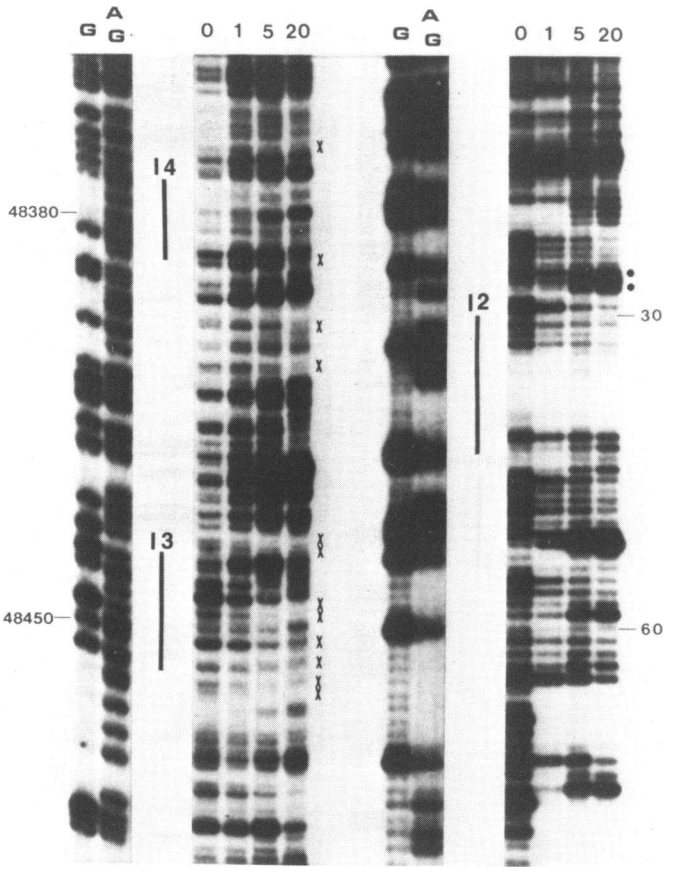

Figure 6: DNAase I footprints of the λ I2, I3 and I4 sites. The G and AG lanes are Maxam and Gilbert sequencing reactions. The fragment used was the 850 bp segment from the pBR322 *Ban*II site to the λ *Eco*RI site at 194. The fragment was 3'-labeled at the *Eco*RI end. Base pair positions are given at the sides of the figure. Lanes 0, 1, 5 and 20 are the DNAase I digestion pattern of the DNA fragment containing the I2, I3 and I4 sites after the incubation of the DNA with 0 nM, 45 nM, 225 nM and 900 nM IHF, respectively. Symbols as in Figures 4 and 5.

to the *Bgl*II site at 415. This fragment was 5'-labeled at the *Dra*I end. Again, there was no significant protection of the I0 site and weak protection of some bp at the I0' site was found (data not shown).

The I3 and I4 sites of λ are weakly protected by IHF at high concentrations

I3 and I4 are present on the footprint of the top strand shown in Figure 6. The results show that at low IHF concentration (45 nM), neither I3 nor I4 is protected against DNAase I attack. At 5 and 10-fold increased

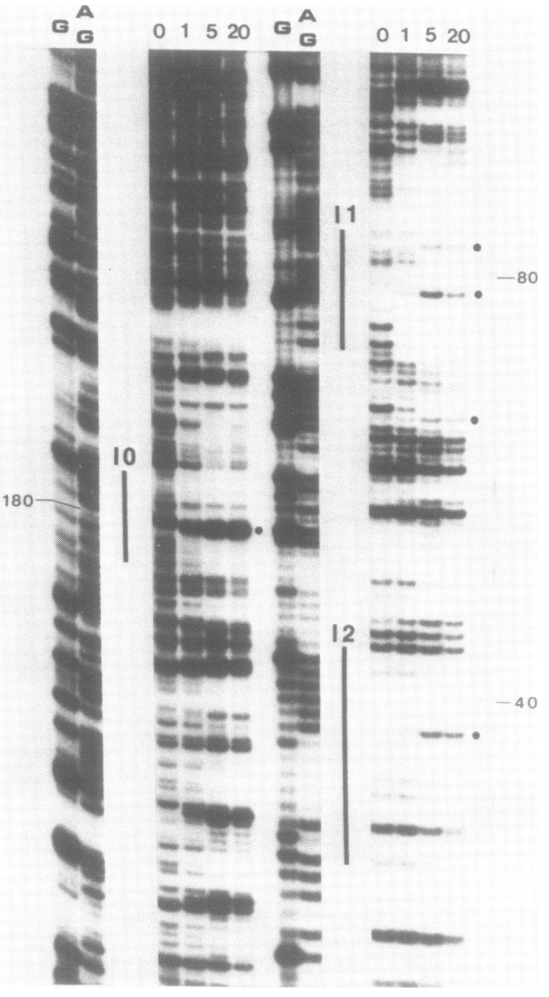

Figure 7: DNAase I footprints of phage 21 *cos* fragment containing the 21 I0, I1 and I2 sites. The 348 bp fragment extends from the λ *Ppu*MI site at 48473 to the 21 *Nco*I site at 319. The DNA was 3'-labeled at the *Ppu*MI end. Base pair positions are given at the sides of the figure. G and AG lanes are Maxam and Gilbert sequencing reactions. Lanes 0, 1, 5, and 20, show the DNAase I digestion patterns of 21 *cos* fragment after incubation with 0 nM, 45 nM, 225 nM and 900 nM IHF, respectively. Symbols as in Fig. 4.

IHF concentrations, both I3 and I4 show some diminution of DNAase cleavage at some base pairs, suggesting weak binding by IHF. To approach the I3 and I4 sites more closely, the 693 bp fragment extending from the λ *Xmn*I site at

39 to the *Ban*II site of pBR322 was 5'-labeled at the *Xmn*I end and footprinted. Again, weak binding of I3 and I4 by IHF, this time on the bottom strand, was indicated by diminution of DNAase cleavage at a few base pair positions (data not shown). It may be noted in Figure 6 that bp outside I3 and I4 show enhanced cleavage by DNAase I. These enhancements may indicate the formation of a complex structure at high IHF concentrations. Diminution of cleavage at bp outside the I sites is not seen, indicating that the diminutions found within the I sequences represent specific interactions with IHF.

IHF protects three sites at the left end of phage 21 DNA

Lambdoid phage 21 has a *cosN* that is functionally identical to *cosN* of λ. In *cosB* of phage 21 there are also repeat sequences analogous to the R sequences at λ *cosB* (not shown in Figure 2), but the sequence of the 21 repeats is different from the λ R sequence. This divergence of the DNA sequence in the repeats between the two phages presumably reflects the different DNA packaging specificities between the two phages. DNA sequencing at the left end of phage 21 has revealed three potential IHF binding sites (11), I2, I1 and I0 (Figure 2). These sites are at roughly the same locations as the I2, I1 and I0 sites of λ, respectively, but differ in details. The 21 I2 site differs from the λ I2 by only two base pairs. The 21 I1 site contains only one IHF binding consensus sequence; it is equivalent to λ I1A. The 21 I0 site is oriented in the opposite direction and several base pair upstream from the λ I0 site; it includes the putative Shine-Dalgarno sequence of the 21 *I* gene, which encodes the smaller subunit of 21 terminase. Plasmid pBA4, which contains the 21 DNA sequence from *cosN* to 490, was used to study the binding of IHF to 21 *cos*. The plasmid was digested by *Ppu*MI at 48473 of the λ sequence and *Nco*I at 319 of the phage 21 sequence. The 348 bp fragment was isolated and 3'-labeled at the *Ppu*MI end. At 45 nM IHF, positive footprinting results were observed at the 21 I1 and I0 sites (Figure 7). The segments protected on the bottom strand are from 63 to 99 for I1 and from 163 to 206 for I0. When the IHF concentration was increased 5 and 20-fold, the 21 I2 site also showed diminished cleavage by DNAase I and the protection at the 21 I1 and I0 sites became stronger. The I2 segment protected extends from 26 to 49. There are some sites of enhancement of DNAase I cleavage at the higher IHF concentrations. The DNAase I cleavage of the A at 42 in the 21 I2 region, the A at 78 in the 21 I1 region and the C at 179 in the 21 I0 region are most enhanced at the high IHF concentration.

A summary of the IHF protection results is shown in Table 1.

Table 1. Relative strength of IHF protection

*cos*λ:	I4	I3	I2	I1	I0'	I0
	+/-	+/-	+	+++	+/-	-
*cos*ϕ21:			I2	I1		I0
			+	++		++

+++: strong protection at all the IHF concentrations tested.
++: positive protection at 1X IHF concentration with increasing protection at higher IHF concentrations
+: no detectable protection at 1X IHF concentration and weak protection at higher IHF concentrations.
+/-: weak protection of some base pairs at high IHF concentrations.
-: no detectable protection at all the IHF concentrations tested.

DISCUSSION

IHF binds to specific sites at λ *cos*

Our work shows that IHF binds to specific sites at *cos*λ and *cos*21. The strength of protection by IHF among the sites at λ *cos* varies. None of the six potential IHF binding sites at λ *cos* has a perfect match to the IHF binding consensus sequence (Fig. 3). Inspection of the nucleotide sequence and DNAase I protection patterns of the IHF binding sites at λ *cos* reveals several common features. First, the more closely a site matches the consensus sequence, the better the protection. The two binding sequences at the I1 site have the best matches to the consensus sequence. I1A has two mismatches and I1B has only one mismatch and these sites are most strongly protected by IHF at all concentrations. On the other hand, the I0, I0', I3 and I4 sites all have three mismatches and they show very poor or no protection by IHF.

Secondly, each of the poor binding sites differs from the consensus sequence at a position known to be very important for IHF binding. For example, the A at position 4 of the consensus sequence is important because an A-to-C transversion mutation in the IHF site in the left end of IS1 abolishes IHF binding (23). I0, I3 and I4 have a base other than A at position 4. The T at position 11 of the IHF consensus sequence is also important for the interaction with IHF, as follows. Gardner and Nash (8) found that changing both T_{10} and T_{11} to Gs results in loss of IHF binding. T_{10} is apparently not crucial for IHF binding, because pBR322 I site contains an A at position 10 and is able to bind IHF (24). The above two studies indicate that T_{11} is important for IHF binding. The I0' and I3 sites contain a base other than T at position 11. Thus each of the weak or non-binding

sequences in *cos*λ is changed for a base that is important for IHF binding.

The regions protected by IHF seem very wide compared to the 15-bp IHF binding consensus. The region protected by IHF at H1 of λ *att* covers 24 bps; the protection of IHF at H2 of λ *att* is 34 bps wide on one strand and 42 bps on the other (3). At λ *cos* similar results are obtained. In the case of the λ I1 site, the protected region spans about 46-bps on the top strand and 30-bps on the bottom strand (Figure 4). The λ I1 site consists of two possible binding sequences and the protected region covers both sequences. We do not know which of the two binding sequences or whether both are recognized by IHF at this point. A further genetic study is needed to answer this question. Comparison of the pattern of protection between λ I1 and 21 I1 suggests that only one protein molecule binds to the λ I1 site, as follows. The 21 I1 contains only one IHF consensus and it is structurally analogous to the I1A of λ, but the width of the region protected at the 21 I1 site is very similar to that of λ I1 site.

Comparison of IHF binding sites at λ *cos* and phage 21 *cos*

Lambdoid phage 21 requires IHF as a host factor for terminase action in DNA packaging while λ DNA packaging does not absolutely require IHF as a host factor *in vivo*. For λ, cleavage of *cosN* can be carried out *in vitro* in a reaction requiring terminase and a host factor, either IHF or THF (terminase host factor). The latter has been identified biochemically; it is a basic protein with a molecular weight of 22 kd (25). The gene encoding THF has not been identified.

Comparing the protection by IHF to the right of the λ and 21 *cosN*s, we found that certain differences exist between them. The 21 IO site runs in an opposite direction to the λ IO site and overlaps the Shine-Dalgarno sequence of the 21 *I* gene, which codes for the smaller subunit of 21 terminase. IHF protects the 21 IO site but it does not bind to the λ IO site. IHF has been found to promote the translation initiation of the λ *cII* gene transcript by binding to the site 3-bp upstream of the *cII* Shine-Dalgarno sequence (7). We have not determined whether IHF plays any role at the 21 IO site in 21 *I* gene translation initiation. Also, the DNA sequence divergence between the two phages at their I1 sites results in only one possible IHF binding sequence at the 21 I1 site while there are two possible IHF binding sequences at the λ I1. We have not examined the *cos* fragment to the left of 21 *cosN*.

From the data we have obtained, we can conclude that λ *cos* and 21 *cos* respond differently to IHF. 21 *cos* contains three clear IHF binding sites

to the right of its *cosN* and two of them are well protected by IHF in footprinting experiments. On the other hand, λ *cos* has more sites and each site seems to have its own characteristics. Also, *in vivo* and *in vitro* data have shown that 21 DNA packaging absolutely requires IHF as a host factor (12) while λ DNA packaging can proceed in cells lacking IHF.

Weak IHF binding sites.

We have found that the I0', I3 and I4 sites of λ, which differ from the consensus sequence at 3 positions, bind IHF weakly. These sites are unusual in that most of the IHF binding sites have been described to date are strong sites that match the consensus well. The significance of the weak sites in *cos*λ is unclear. They may function *in vivo*, or they may be sequences that fortuitously match the consensus sequence. The latter is to be expected in AT-rich segments of DNA such as *cos*λ. On the other hand, genetic studies indicate that IHF sites that have several mutations retain residual activity *in vivo* (8, 26). These studies suggest that IHF binding sites that bind weakly or not at all in footprint experiments may still be able to be bound by IHF with sufficient affinity to function *in vivo*. Further work on this point is needed.

Possible roles of IHF in λ DNA packaging

IHF has a surprisingly varied functional repertoire, acting in gene expression and in structure formation. IHF acts post-transcriptionally in the regulation of λ *cII* expression (7). IHF acts positively and negatively at the transcriptional level in phage Mu (27).

IHF is involved in structure formation in a number of cases. At the replication origin of plasmid pSC101, IHF assists in the formation of a nucleoprotein structure that includes the *ori* DNA, the *dnaA* protein, and the plasmid initiator protein. In pSC101, IHF binds at a position that contains a static bend, and the bending of the DNA is increased by IHF (6). The pSC101 results indicate that IHF may play a role in structure formation by bending the DNA.

In λ site-specific recombination, IHF is involved in formation of complex nucleoprotein structures at *attP* and *attR* (4, 5, 28, 29). Structure formation involves cooperative interactions between IHF and integrase (3, 5). IHF may also be involved in bending the DNA at the *att* sites, because *att* DNA, like the pSC101 origin, appears to be bent (30). The role of IHF at *cos* seems likely to be analogous to its role at *attP*, *attR* and at the pSC101 *ori*, i.e., in structure formation, though there is little direct evidence to date and the argument is based on the similarity of *cos* to the *att* sites and the pSC101 *ori*. The structure of *cos* is complex and is probably an interspersed array of binding

sites for IHF and terminase, in a manner analogous to the *att* and *ori* sites. The DNA of *cos* is also bent (Yeo, personal communication; Kosturko, personal communication). A third similarity to the *att* and *ori* sites is functional. Echols (31) has argued that complex nucleoprotein structures are formed for DNA sites that function in processes, like site-specific recombination and initiation of replication, in which great specificity is required. The introduction of nicks at *cosN* and the initiation of packaging would also seem to require great specificity. Therefore, we speculate that the role of IHF at *cos* is to assist in structure formation, perhaps by bending the DNA and/or cooperative interactions with terminase.

ACKNOWLEDGEMENTS

We thank Howard Nash who generously gave us purified IHF for this study. We appreciate helpful discussions with Linda Kosturko and our coworkers here at Iowa. We thank Marcia Reeve for preparation of the manuscript. Supported by NIH grant AI-12581.

REFERENCES

1. Miller, H.I., and Friedman, D.I. (1980). Cell 20, 711-718.
2. Flamm, E.L., and Weisberg, R.A. (1985). J. Mol. Biol. 183, 117-128.
3. Craig, N.L., and Nash, H. A. (1983). Cell 35, 795-803.
4. Richet, E., Abcarian, P., and Nash, H.A. (1986). Cell 46, 1011-1021.
5. Thompson, J.F., Moitoso de Vargas, L., Skinner, S.E., and Landy, A. (1987). J. Mol. Biol. 195, 481-493.
6. Stenzel, T.T., Patel, P., and Bastia, D. (1987). Cell 49, 709-717.
7. Mahajna, J., Oppenheim, A., Rattray, A., and Gottesman, M. (1986). J. Bact. 165, 167-174.
8. Gardner, J., and Nash, H.A. (1986). J. Mol. Biol. 191, 181-189.
9. Feiss, M. (1986). Trends in Genetics 2, 100-104.
10. Becker, A., Marko, M., and Gold, M. (1977). Virology 78, 291-305.
11. Miller, G., and Feiss, M. (1985). J. Mol. Biol. 185, 246-249.
12. Feiss, M., Frackman, S., and Sippy, J. (1985). J. Mol. Biol. 183, 239-240.
13. Bear, S., Court, D., and Friedman, D.I. (1984). J. Virol. 52, 966-972.
14. Feiss, M., Widner, W., Miller, G., Johnson, G., and Christiansen, S. (1983). Gene 24, 207-218.
15. Daniels, D., Schroeder, J., Szybalski, W., Sanger, F., Coulson, A., Hong, G., Hill, D., Peterson, G., and Blattner, F. (1983). In Lambda II, Cold Spring Harbor Laboratory, ed. R.W. Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg. pp 519-676.
16. Feiss, M., Siegele, D.A., Rudolph, C.A., and Frackman, S. (1982). Gene 17, 123-130.
17. Chang, A.C.Y., and Cohen, S.N. (1978). J. Bact. 134, 1141-1156.
18. Frackman, S., Siegele, D.A., and Feiss, M. (1985). J. Mol. Biol. 183, 225-238.
19. Bestwick et al. (1983). Anal. Biochem. 133, 79-84.8
20. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1984). Molecular Cloning, A Laboratory Manual. New York: Cold Spring Harbor Laboratory.

21. Maxam, A., and Gilbert, W. (1980). Methods in Enzymology 65, 499-559.
22. Galas, D., and Schmitz, A. (1978). Nuc. Acids Res. 5, 3157-3170.
23. Prentki, P., Chandler, M., and Galas, D. J. (1987). EMBO J. 6, 2479-2489.
24. Gamas, P., Chandler, M.G., Prentki, P., and Galas, D.J. (1987). J. Mol. Biol. 195, 261-272.
25. Gold, M., and Parris, W. (1986). Nucl. Acids Res. 14, 9797-9809.
26. Thompson, J.F., Waechter-Brulla, D., Gumport, R.I., Gardner, J.F. de Vargas, L.M., and Landy, A. (1986). J. Bact. 168, 1343-1351.
27. Krause, H.M., and Higgins, N.P. (1986). J. Biol. Chem. 261, 3744-3752.
28. Pollock, T.J., and Nash, H.A. (1983). J. Mol. Biol. 170, 1-18.
29. Better, M., Lu, C., Williams, R.C., and Echols, H. (1982). Proc. Natl. Acad Sci., USA 79, 5837-5841.
30. Ross, W., and Landy, A. (1982). J. Mol. Biol. 156, 523-529.
31. Echols, H. (1986). Science 233, 1050-1056.