Analysis of the IHF binding site in the regulatory region of bacteriophage Mu

Piet A.van Rijn, Pieter van de Putte and Nora Goosen*
Laboratory of Molecular Genetics, Department of Biochemistry, Gorlaeus Laboratories, Leiden University, Einsteinweg 5, 2300 RA Leiden, The Netherlands

Received April 9, 1991; Revised and Accepted May 7, 1991

ABSTRACT

In bacteriophage Mu the converging early and repressor transcriptions are both stimulated by binding of IHF to the same region, which is located just upstream of the early promoter (Pe) and 100 base pairs downstream of the repressor promoter (Pc). Within this region two sequences are present (*ihfa* and *ihfb*) that match the consensus sequence for IHF binding. These sequences are partially overlapping and in inverted orientation. In this paper we describe the effect of mutations in the non-overlapping part of *ihfa* and *ihfb* on the binding of IHF. We show that IHF has a very strong preference to bind to *ihfb* even when a mutated *ihfa* has a better match with the consensus. A stretch of A residues located nine base pairs from the *ihfb* sequence appears to play an important role in the stability of the DNA-IHF complex, but not in the discrimination between the two putative binding sites. In addition we describe the effect of the mutations on the stimulation of early and repressor transcription. We show that for activation of the Pc promoter a stable complex between IHF and the DNA is required, whereas for normal Pe stimulation a much weaker DNA-IHF interaction is sufficient.

INTRODUCTION

The Integration Host Factor (IHF) of *Escherichia coli* has originally been identified as a protein that is essential for integration of the lambda genome into the *E. coli* chromosome (reviewed in 1). IHF is a histone-like protein consisting of two basic subunits, encoded by the *himA* and *himD* (*hip*) genes (2,3). It shows a strong homology with HU which is an other member of the family of histone-like proteins (4). Unlike HU, IHF is a DNA binding protein which recognizes a specific sequence. Binding sites of IHF (*ihf*) contain the consensus sequence WATCAANNNNTTR (5,6,7). Some sequences however that show a perfect match to the consensus sequence are not bound by IHF (8,9,10). This indicates that also the DNA flanking the consensus is involved in the binding. Yang and Nash have shown that upon IHF binding to the *ihf* sites in the *attP* of lambda a stretch of A's is contacted that is located nine base pairs from the 5'-end of the consensus sequence (7). It has been suggested that this contact is important for the stability of IHF binding. However not all known *ihf* sites contain such A stretches at that position. The functional *ihf* sites are frequently found in A/T-rich regions (11), suggesting that also DNA bending might play a role in IHF binding. Binding of IHF has been shown to involve minor groove recognition (7) and the binding results in a strong bending of the DNA by about 140° (12).

Besides in lambda integration, IHF is involved in many other processes viz. replication, transposition and regulation of both viral and bacterial genes (for a review see 11). In bacteriophage Mu transcriptions from the early promoter Pe and the repressor promoter Pc are stimulated by IHF (13,14,15). Although binding of IHF to the same region just upstream Pe is involved in both processes, the mechanism by which Pe and Pc are activated must be different. The transcriptions from Pe and Pc are in opposite directions and overlapping by about 36 base pairs. As a consequence the IHF binding region is located 100 bp downstream the transcription start of Pc. The distance between Pc and this IHF binding region can be enlarged by at least another 100 base pairs without the loss of Pc activation (16). The stimulation of Pe-directed transcription however is abolished by the same 100 base pairs insertion. Furthermore the transcription stimulation of Pe requires a helix dependent orientation of the IHF and RNA polymerase binding sites on the DNA, whereas for Pc stimulation no such helix dependency was found (16). Within the IHF binding region two consensus sequences can be indicated which are partially overlapping and in opposite orientation (*ihfa* and *ihfb*), see Fig. 1. It has been postulated that one consensus sequence might be involved in Pe activation and that the other consensus sequence would be required for the stimulation of the repressor transcription by IHF (13).

In order to investigate the affinity of IHF for both putative binding sequences, we introduced mutations in the non-overlapping part of these sequences and studied the IHF binding to the mutated DNA. Here we show that IHF has a very strong preference to bind to one consensus sequence. Furthermore we show that mutations that reduce the stability of the IHF-DNA complex abolish the activation of the Pc promoter without affecting the stimulation of the early transcription.

* To whom correspondence should be addressed

MATERIALS AND METHODS

Bacterial strains and plasmids

Strain PP1674 is BW313 (*dut*, *ung*) (17) harbouring pGP655 (F^+ Tc^R) (18) and was used to isolate ssDNA of M13mp11 derivatives for site-directed mutagenesis. All plasmids were transformed to JM101 (19), except the plasmid containing the regulatory region of Mu in tandem which was maintained in strain DH5α (*recA*) (20).

The *E. coli* strains used in the galactokinase assay were AB1157 (*galK*, *himD*$^+$) and the isogenic strain PP1954 (*galK*, *himD*) (15).

The *galK* fusion plasmids are schematically represented in Fig. 2. The *galK*expression plasmids pCA95 (21) and pGP182 (16) and the *galK* fusion plasmids pGP133, pGP134, pGP139 (13), pGP185 (16) and pGP188 (15) have been described. Plasmid pGP185-*BL* which carries a *Bam*HI linker in the region flanking the *ihf* site (Fig. 1) has been described (16). Plasmid pGP714 containing the permuted Mu fragment, corresponding to position 818 to 1242, was constructed by ligation of the *Eco*RI-*Cla*I fragment of pGP133 to the *Eco*RI and the *Acc*I site of plasmid pSB118 (22). Next the *Eco*RI fragment (containing the Mu sequence) was cloned in tandem resulting in pGP714. In a similar way pGP713 (tandem sequences of mutant 1+2) was constructed. All cloning procedures were carried out essentially as described (23).

Oligo-directed mutagenesis

Phage M13mp11 containing the *Eco*RI-*Sma*I fragment of pGP134 (Fig. 2) was used to introduce mutations in the *ihf* sites. Site-directed mutagenesis was carried out according to Kunkel (17). The oligonucleotides 5′-TGATTACTGCCTAACGCGTT-3′ and 5′-AACGCGTTGGGAAATAAGGA-3′ were used to isolate mutations in *ihfa* (mutant 1), in *ihfb* (mutant 2) and in both sites (mutant 1+2), see Fig. 1. The oligonucleotides used to construct mutant 3+4 are: 5′-CGCGTTAGGAAAT-3′ and 5′-ACTGA-TCAACGCG-3′ (Fig. 1). For the introduction of the *ihf* mutations in the Pe-*galK* fusion plasmids the *Eco*RI-*Cla*I fragments of the different M13mp11 mutants were inserted into pGP182 leading to pGP185 derivatives. Insertion of the *Hin*dIII fragment of the M13mp11 mutants in pGP139 resulted in pGP133 derivatives (with the *galK* gene under control of Pc).

DNaseI protection experiments

*Cla*I-*Hin*dIII fragments corresponding to the Mu sequence from position 818 to 1004 were labeled uniquely at the 3′-end of either the *Cla*I or the *Hin*dIII site by Klenow fragment of DNA polymerase I and α^{32}P-CTP. The labeled fragments were purified from 5% polyacrylamide gels. Subsequently 22 μl reactions containing labeled fragment (less than 5 nM) and various amounts of IHF were incubated for 20 minutes at 37°C in IHF binding buffer (50 mM Tris-HCl pH 7.5, 70 mM KCl, 7 mM $MgCl_2$, 3 mM $CaCl_2$, 1 mM βmercapto-ethanol, 200 μg/ml bovine serum albumin and 10% glycerol). Next 0.25 ng DNaseI was added and the mixture was incubated at 25°C for 10 minutes exactly. For the footprinting with fragments containing mutant 3+4, 1.25 ng DNaseI was used and the mixture was incubated at 14°C for 20 minutes. The reaction was quenched by addition of an equal volume of stop solution (0.5 μg/ml calf thymus DNA, 600 mM Na-Acetate pH 7.5, 80 mM EDTA). The samples were extracted twicewith phenol, twice with chloroform, followed by precipitation with 2 volumes of ethanol and washing with 70% ethanol. The air-dried samples were resuspended in sequence loading mix and electrophoresed on denaturing 6 or 8% polyacrylamide sequence gels (24).

Retardation assay

Derivatives of pGP188 or pGP185 were digested with *Eco*RI, *Hin*dIII and *Cla*I and labeled with Klenow Fragment of DNA polymerase I and [α^{32}-P]-CTP. After labelling, the DNA fragments were purified by phenol extraction and ethanol precipitation. Next, the fragments were resuspended in IHF binding buffer (50 mM Tris-HCl pH 7.5, 70 mM KCl, 7 mM $MgCl_2$, 3 mM $CaCl_2$, 1 mM βmercapto-ethanol, 200 μg/ml bovine serum albumin and 10% glycerol). About 50 ng DNA was incubated with varying amounts of purified IHF at 37°C for 20 minutes and electrophoresed on a 5% polyacrylamide gels at 4 or 40°C, as indicated. Purified IHF was a generous gift from H.A.Nash (the National Institute of Mental Health).

Galactokinase assay

Galactokinase (GalK) activities of cells containing *galK* fusion plasmids were determined essentially as described (25). GalK activities are expressed as nanomoles phosphorylated galactose per min per 10^8 cells at 32°C.

Determination of the sequence-induced bending of the DNA

Plasmids pGP713 and pGP714 (containing the *Eco*RI fragment in tandem) were digested with restriction enzymes which cut the *Eco*RI fragment only once, resulting in fragments of the same length but different composition. After phenol extraction and ethanol precipitation, the samples were resuspended in 10 mM Tris-HCl, 1 mM EDTA pH 7.6 and electrophoresed on 10% polyacrylamide gels at 4 or 60°C, as indicated.

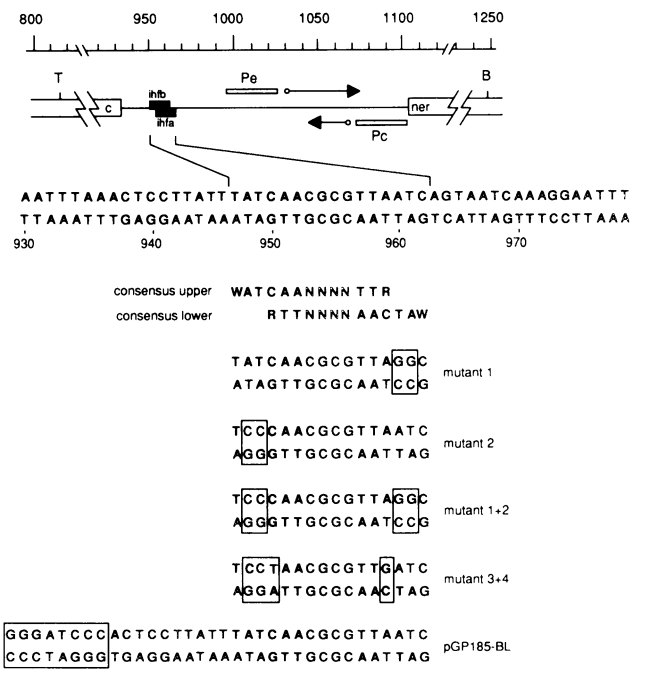

Figure 1. Nucleotide sequence of the mutations in the IHF binding region. The regulatory region of Mu is schematically represented, showing the relative positions of the early promoter (Pe) and the repressor promoter (Pc). Relevant restriction sites are T=*Taq*I and B=*Bal*I. Bases in *ihfa* and *ihfb* that match the consensus are indicated by bold symbols. The stretch of A residues at the 5′ site of *ihfb* is also in bold symbols. The base substitutions in the different mutants are boxed.

RESULTS

IHF binding to mutant ihf sites

Within the IHF binding region of phage Mu there are two possible consensus sequences *ihfa* and *ihfb* (Fig. 1). Comparison of the DNaseI protection pattern of this region by IHF with the non-symmetric footprint pattern with respect to the consensus of the *ihf* sites in the *attP* of lambda suggested that IHF binds primarily by recognition of the *ihfb* consensus sequence (14). However it is not ruled out that *ihfa* can also berecognized by IHF with a lower efficiency.

In order to study the affinity of IHF for both putative *ihf* sites, we introduced mutations in *ihfa* and *ihfb* by site-directed mutagenesis. The mutations were introduced in that part of the consensus sequence which is not shared by the consensus sequence in the other strand. Mutant 1 contains two base substitutions in *ihfa* at positions 960 (T→C) and 961 (A→C) of the Mu sequence and mutant 2 contains the similar mutations in *ihfb* at positions 948 and 949. Mutant 1+2 carries all four subtitutions (Fig. 1). First we compared the binding of IHF to the wild type sequence and to the different mutant *ihf* sites using the DNaseI footprinting technique (Fig. 3 and 4). The footprints of the mutant 1, mutant 2 and mutant 1+2 DNA fragments show the same pattern of protection and enhanced cleavage as the footprint of the wild type DNA. The DNA in the 'lower' strand (Fig. 3) is protected from about position 940 to 969 with an enhanced cleavage by DNaseI at the right border of the protected region. This enhanced cleavage is characteristic for the 5'-border of the DNaseI footprint of IHF showing that it recognizes the *ihfb* sequence (5,14). The DNA of the 'upper' strand is protected from position 927 to 963 and from 967 to 970. In this strand there is also a cleavage at the right border within the protected region, which is also characteristic for the recognition of the *ihfb* consensus (5). In the 'upper' strand we observed also an enhanced cleavage site at position 955.

Apparently in all three mutants IHF still primarily recognizes the *ihfb* consensus. This is surprising, since in mutant 2 the *ihfa* sequence matches the consensus better than the mutated *ihfb* (Fig. 1). Although in mutant 2 (with a mutated *ihfb*) the affinity for IHF is somewhat reduced, it is clear that although two out of nine base pairs of the recognized consensus in this mutant have been changed, the binding of IHF is not severely affected. A more striking reduction of IHF binding affinity is observed with mutant 1+2. This suggests that also bases adjacent to the *ihfb* consensus sequence are important for the recognition of this consensus by IHF.

To test whether the affinity of IHF for *ihfa* will increase when the differences between *ihfa* and *ihfb* with respect to the consensus are more extreme, we introduced two extra mutations in mutant 2. The mutation T→C at position 959 leads to a better match of *ihfa* to the consensus and the mutation C→T at position 950 makes *ihfb* deviate even more from the consensus (Fig. 1). With the resulting mutant 3+4 it was very difficult to obtain protection against DNaseI by IHF. Only after preincubation with large amounts of IHF and partial DNaseI digestion at lower temperature, a very weak protection pattern can be observed (Fig. 5). However with these IHF concentrations we also observed some non-specific binding. Apparently the binding of IHF to mutant 3+4is very weak. A closer examination of the footprint pattern of mutant 3+4 reveals that it seems to be different from the wild type. In the 'upper' strand the characteristic enhanced cleavage site at the right border of the protected region (position 964) is absent, whereas now an enhanced cleavage site at the left border seems to be present.

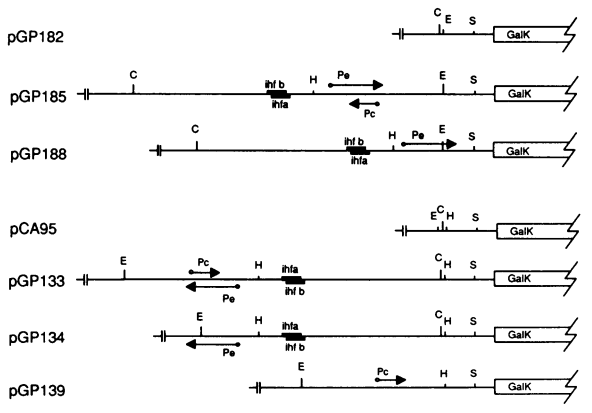

Figure 2. Schematic representation of the *galK* fusion plasmids. The arrows indicate the directions of the transcription from Pe and Pc. Relevant sites are E=*Eco*RI, C=*Cla*I, H=*Hin*dIII and S=*Sma*I. Plasmids pGP185 and pGP133 contain the Mu sequence from position 818 to 1242 in pGP182 and pCA95 respectively. Plasmids pGP188 and pGP134 contain the Mu sequence from position 818 to 1063 in pGP182 and pCA95. Plasmid pGP139 is pCA95 containing the Mu sequence from 1004 to 1242.

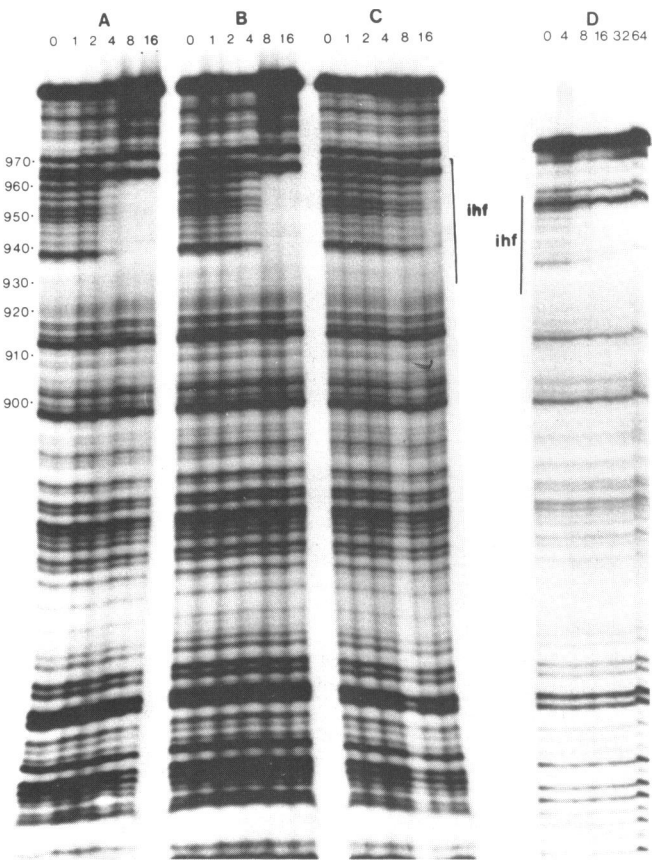

Figure 3. DNaseI protection pattern of the 'lower' strand by IHF. The *Cla*I-*Hin*dIII fragment was labeled at the *Cla*I site. Before DNaseI digestion, incubation was carried out with different amounts of IHF as indicated at the top of each lane (expressed in nanogram). The numbering of the sequence is with respect to the left end of the Mu sequence. Panel **A**: wild type *ihf*, panel **B**: mutant 1, panel **C**: mutant 2 and panel **D**: mutant 1+2.

In the 'lower' strand an enhanced cleavage site can be observed at position 948 which is absent in the wild type footprint. These differences might indicate that in mutant 3+4 IHF now recognizes *ihfa* but it is clear that although in this mutant *ihfa* matches the consensus sequence almost perfectly, the affinity for IHF is nevertheless extremely low.

IHF bends the DNA by about 140°, which can be shown by the highly decreased electrophoretic mobility of IHF-DNA complexes (10,12). Since the retardation of DNA fragments by IHF not only depends on the degree of bending, but also on the position of the IHF binding site we also tested the discrimination between *ihfa* and *ihfb* in gel retardation experiments. Fig. 6A shows a retardation gel of the *Cla*I-*Hin*dIII fragment of the wild type, mutant 1 and mutant 2 DNA (performed at 40°C). In all three cases the IHF-DNA complex (in which as the footprints showed IHF is bound to *ihfb*) migrates at the same position indicating that the degree of bending is not influenced by the base substitutions. The retarded fragment using DNA of mutant 1+2 also migrates at the same position as the wild type fragment (Fig. 7). The stability of the IHF-DNA complex with mutant 1+2 however is much lower compared to mutant 1 and mutant 2. A specific retarded band can only be observed when the experiment is performed at 4°C (Fig. 7), but not at 40°C (Fig. 6B). Also when DNA of mutant 3+4 is used a specific IHF-DNA complex can only be detected at low temperature and much higher IHF concentrations are needed (Fig. 7). The position of this retarded mutant fragment however clearly differs from that of the wild type fragment. This different mobility could be the result of IHF binding to the same sequences as in the wild type DNA, but inducing a different degree of DNA bending as a result of the mutations. Considering the DNaseI footprint results however, we regard it as more likely that the different mobility of the IHF-DNA complex of mutant 3+4 reflects the binding of IHF to different sequences, by recognition of *ihfa* instead of *ihfb*.

Role of the flanking sequences in IHF binding

From the experiments described above it is clear that the preference of IHF to bind *ihfb* in the wild type Mu sequence can not be ascribed to the differences in the sequence between *ihfa* and *ihfb*. Therefore the DNA flanking the consensus should play an important role in the recognition of the IHF binding site. Yang & Nash (7) haveshown that binding of IHF to the target sequences in *attP* of phage lambda involves a close contact with a stretch of A residues located about nine base pairs from the 5'-site of the consensus. Such a track of A's is present at a comparable position with respect to *ihfb* (positions 935−937) but not with respect to *ihfa* (Fig. 1).

To test whether the presence of this A track is responsible for the preferred IHF binding to *ihfb*, we isolated a mutant in which two of the three A residues have been deleted. The sequence of the IHF binding region of the resulting plasmid (pGP185-*BL*) is indicated in Fig. 1 and the footprint of the 'upper' strand of

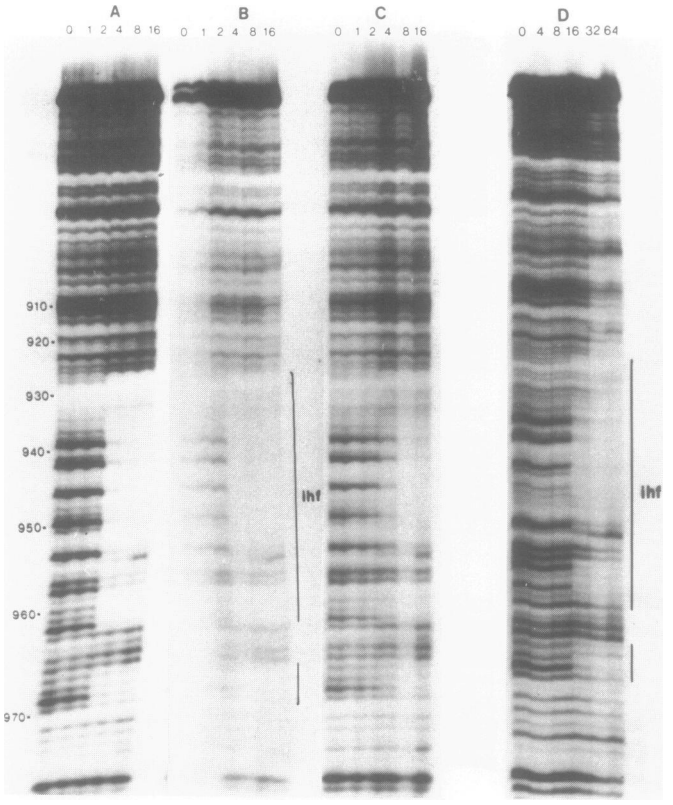

Figure 4. DNaseI protection pattern of the 'upper' strand by IHF. The *Cla*I-*Hin*dIII fragment was labeled at the *Hin*dIII site. Before DNaseI digestion, incubation was carried out with different amounts of IHF as indicated at the top of each lane (expressed in nanogram). The numbering of the sequence is with respect to the left end of the Mu sequence. Panel **A**: wild type *ihf*, panel **B**: mutant 1, panel **C**: mutant 2 and panel **D**: mutant 1+2.

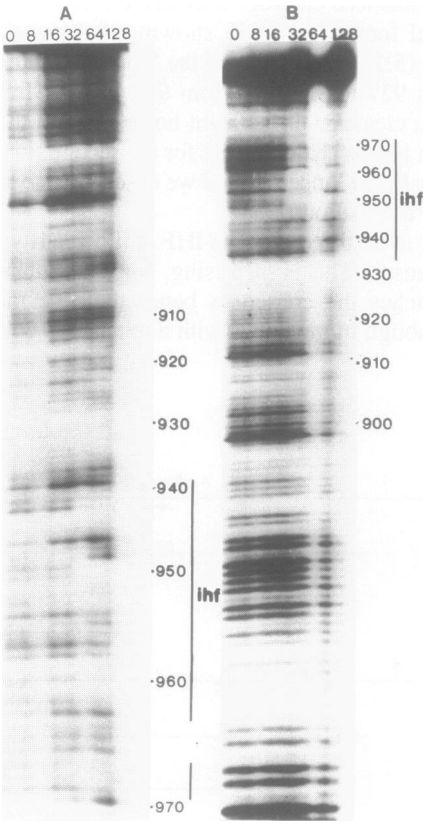

Figure 5. DNaseI protection pattern of mutant 3+4 DNA by IHF. The *Cla*I-*Hin*dIII fragment was either labeled at the *Hin*dIII site ('upper' strand) (panel **A**) or at the *Cla*I site ('lower' strand) (panel **B**). Before DNaseI digestion, incubation was carried out with different amounts of IHF as indicated at the top of each lane (expressed in nanogram). The numbering of the sequence is with respect to the left end of Mu.

the *Hin*dIII-*Cla*I fragment of pGP185-*BL* is in Fig. 8. From the characteristic enhanced cleavage sites at the right border (at position 964) and in the center of the consensus (position 955) it is clear that in the deletion mutant IHF still binds to *ihfb*. The protection pattern at the left border however seems to be less extended (about seven base pairs smaller) in comparison to the pattern of the wild type. This indicates that due to the absence of the two A residues the interaction of IHF with this part of the DNA is absent or very weak. A gel retardation assay with pGP185-*BL* DNA (Fig. 9) shows that the disturbed interaction with the 5′ flanking region of the consensus does not result in a different degree of DNA bending in the IHF-DNA complex since the retarded fragment migrates at the same position as in the wild type situation. It also shows that the affinity of IHF for this mutant DNA is severely reduced. The IHF-DNA complex can be detected when the retardation is performed at low temperature with relatively high IHF concentrations (Fig. 9A) but is hardly observed at 40°C (Fig. 9B). Apparently the stretch of A residues in the 5′ flanking region are very important for the stability of the IHF-DNA complex but there are no indications that they play a decisive role in the discrimination between *ihfa* and *ihfb*.

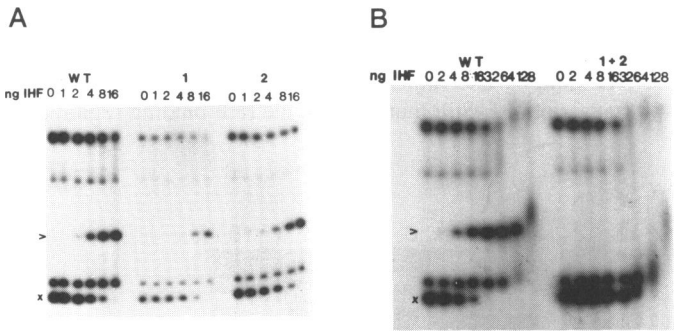

Figure 6. Retardation of DNA fragments by IHF at 40°C. DNA fragments of wild type DNA (WT), mutant 1 DNA (1), mutant 2 DNA (2) and mutant 1+2 DNA (1+2) were incubated with different amounts of IHF as indicated in nanogram. The temperature during electrophoresis was 40°C. The x indicates the position of the unbound *Cla*I-*Hin*dIII fragments (180 bp) with the *ihf* site. The IHF bound *Cla*I-*Hin*dIII fragments are indicated by arrow heads. The 250 bp *Hin*dIII-*Eco*RI fragments that do not contain an *ihf* site serve as an internal control for non-specific binding.

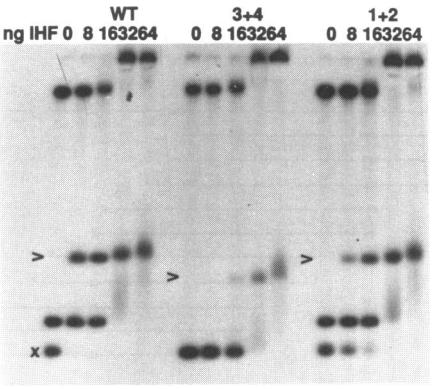

Figure 7. Retardation of DNA fragments by IHF at 4°C. DNA fragments of wild type DNA (WT), mutant 3+4 DNA (3+4) and mutant 1+2 DNA (1+2) were incubated with different amounts of IHF as indicated in nanogram. The temperature during electrophoresis was 4°C. The × indicates the position of the unbound *Cla*I-*Hin*dIII fragments (180 bp) with the *ihf* site. The IHF bound *Cla*I-*Hin*dIII fragments are indicated by arrow heads. Note that the mutant 3+4 plasmid used does not contain the 250 bp *Hin*dIII-*Eco*RI fragment.

The DNA conformation of the Mu regulatory region

The regulatory region of Mu is very A/T-rich (about 75%) and it might be intrinsically bent, as tracks of A residues can induce a bend in the DNA (26). When the flanking sequences of the IHF binding region are part of such a bend the *ihfa* and *ihfb* sequences might not be situated in the same relative position with respect to the DNA-directed bend. Therefore the preferential binding of IHF to *ihfb* might be a consequence of the intrinsic bending of the DNA which could be in a favourable direction for *ihfb* and in an unfavourable direction for *ihfa*.

We have determined the intrinsic bending by electrophoresis of linear permuted fragments containing the IHF binding region corresponding to the Mu sequence from positions 818 to 1242 (pGP714). Digestion of pGP714 with restriction enzymes that cutonly once within this sequence result in fragments of the same size but with a different composition (Fig. 10B). Electrophoresis on a native polyacrylamide gel at low temperature (4°C) shows that the fragments differ in mobility (Fig. 10A), indicating that indeed the Mu fragment contains a sequence-directed bend. By plotting the relative mobility of the different fragments (Fig. 11) the center of the bend can be determined which is located near the *Nhe*I site (position 1032 of the Mu sequence) which is close

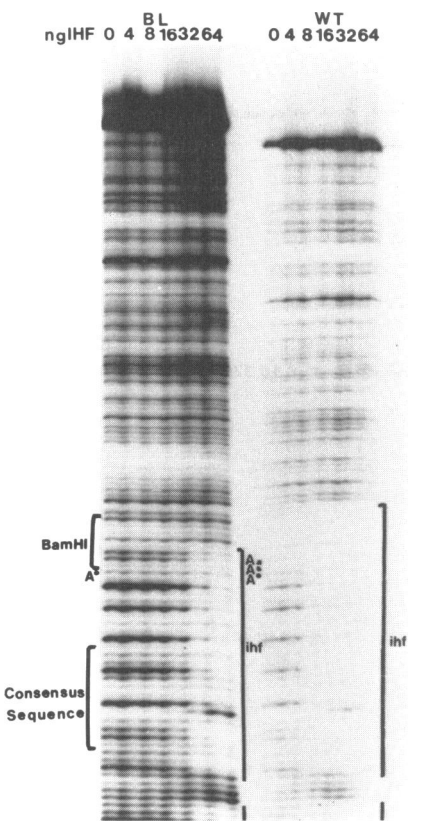

Figure 8. DNaseI protection pattern of the wild type *ihf* site (WT) and the *ihf-BL* mutant (BL). The *Cla*I-*Hin*dIII fragments were labeled at the *Hin*dIII site ('upper' strand). Before DNaseI digestion, incubation was carried out with different amounts of IHF as indicated at the top of each lane (expressed in nanogram). The positions of the *ihfb* consensus sequence, the stretch of A's and the *Bam*HI site are indicated.

to the start of the early transcription. Since the sequence of the entire regulatory region (from position 860 to 1115) is very A/T-rich with many repetitive sequences of four or five A residues it is very likely that this regulatory region, including the *ihf* site has an overall curved structure. Indeed when we studied the migration of the *Cla*I-*Hin*dIII fragment (positions 818 to 1004) which contains the *ihf* site but not the early transcription start also a deviant mobility was observed (not shown) indicating that also the DNA to the left of the *Hin*dIII site is bent. Due to the small size of this fragment however the center of this bend could not be determined exactly. The curved structure of this IHF binding region might be an important factor in the preferential binding to *ihfb*.

The effect of the mutant ihf sites on the transcription stimulation

To determine the effect of the *ihf* mutants on the stimulation of early and repressor transcriptions we introduced these into *galK* fusion plasmids which have the structural *galK* gene under control of either the early or the repressor promoter (as described in materials and methods). The *galK* fusion plasmids that we used for this purpose carry a mutated early promoter (δPe), in which one base pair between the −10 and the −35 region is deleted, resulting in a reduced early transcription. This δPe promoter was used for several reasons. Firstly the stimulation of Pc transcription is hard to determine in the presence of the relatively strong wild type Pe promoter, due to the interference of the early transcription with the opposing Pc transcription (15). With the weaker δPe promoter this interference is nearly absent. Secondly the stimulation of transcription by IHF from a weak promoter (δPe) is more explicit than the stimulation from the stronger wild type promoter. Alterations in the stimulation of early transcription will therefore be more readily observed when δPe is used.

The first conclusion that can be drawn from the transcription results (Table 1) is that in mutant 3+4, which presumably binds IHF to *ihfa* with a low affinity, both the stimulation of Pe and that of Pc are completely abolished. Apparently the *ihfa* consensus can not substitute for the *ihfb* consensus in Pe and Pc stimulation. This is either due to the very low affinity of IHF for *ihfa* in mutant 3+4 or to the invertedorientation of *ihfa* compared to *ihfb*. Evidently, also in the unmutated IHF binding region there is only one IHF binding site (*ihfb*) that is active in the stimulation of both transcriptions. Still mutations in the IHF binding region can

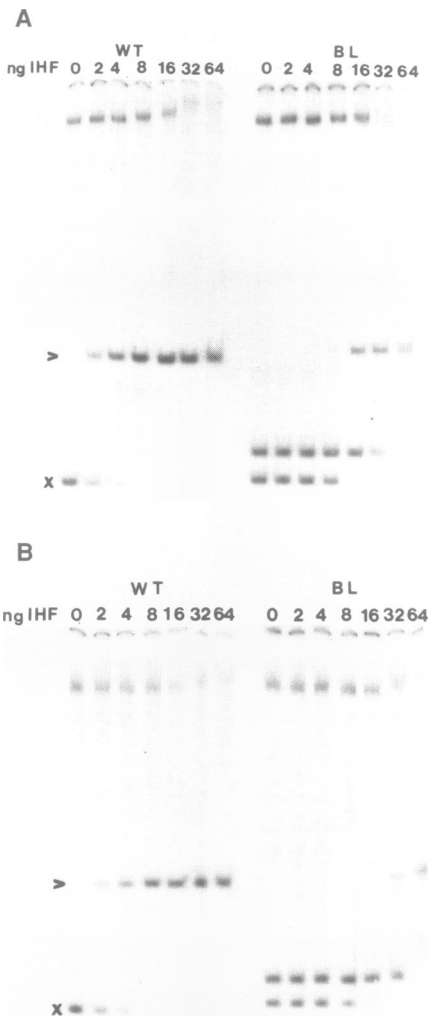

Figure 9. Retardation of wild type and *ihf-BL* mutant DNA fragments by IHF. DNA fragments of wild type DNA (WT) and mutant *ihf-BL* (BL) were incubated with different amounts of IHF as indicated in nanogram. The temperature during electrophoresis was either 4°C (A) or 40°C (B). The × indicates the position of the unbound *Cla*I-*Hin*dIII fragments (180 bp) with the *ihf* site. The IHF bound *Cla*I-*Hin*dIII fragments are indicated by arrow heads. Note that the wild type plasmid used in this experiment does not contain the 250 bp *Hin*dIII-*Eco*RI fragment.

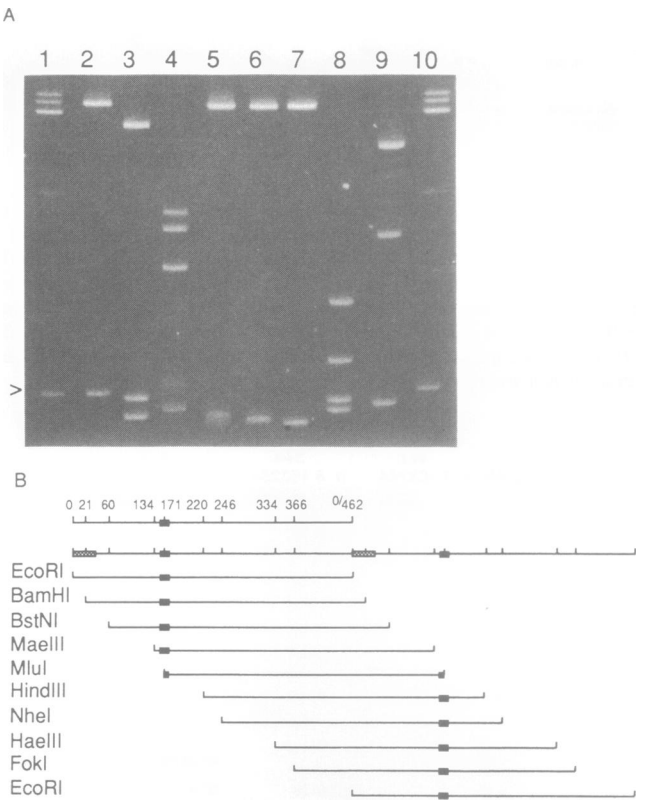

Figure 10. Mobility of permuted DNA fragments. **A**: An ethidiumbromide-stained polyacrylamide gel (10%) after electrophoresis at 4°C of plasmid pGP714 (containing the Mu fragment in tandem) digested with different restriction enzymes. Lane 1 and 10: *Eco*RI, 2: *Bam*HI, 3: *Bst*NI, 4: *Mae*III, 5: *Mlu*I, 6: *Hin*dIII, 7: *Nhe*I, 8: *Hae*III and lane 9: *Fok*I. The arrow head indicates the position of the Mu fragment. **B**: Diagram of the permuted DNA fragments obtained after digestion with the different restriction enzymes as indicated. The position of the *ihf* site is shown by a filledbox.

have a differential effect on the two transcriptions. The most striking example of this is observed with the mutant that lacks two of the three A residues in the 5′ flanking region of *ihfb* (pGP185-*BL*). Transcription stimulation of Pe is unaltered in this mutant. (Note that the results of pGP185-*BL* should be compared with the parental plasmid pGP185, which was shown to have a higher level of Pe expression (15). Pc activation however is completely abolished by the same mutations (pGP133-*BL*). Gel retardation showed that the stability of the IHF-DNA complex of pGP185-*BL* is severely reduced. Apparently such a weakened IHF-DNA complex can still function normally in Pe activation, but for Pc stimulation a more stable complex seems to be required.

A difference in stimulation of Pe and Pc is also observed with mutant 2 (Table 1). The 2 base substitutions in *ihfb* of mutant 2 significantly reduce the IHF stimulation of Pe, whereas they have no or very little effect on Pc (Table 1). Footprint and retardation experiments with mutant 2 DNA showed that DNaseI protection and band shifts were obtained at approximately the same protein concentrations as for the wild type DNA. The reduced stimulation of Pe transcription can therefore not be explained by a reduced affinity of IHF for the mutant DNA. Apparently the mutations do somehow affect the IHF-DNA complex in such a way that it is less active in Pe stimulation while it is still normally active in Pc activation.

The results with mutant 2 become even more striking when we compare them with those of mutant 1+2. DNaseI footprinting showed that the affinity of IHF for mutant 1+2 DNA is severely reduced compared to wild type or mutant 2 DNA (Fig. 3 and 4). Also gel retardation showed that the stability of the IHF-DNA complex is seriously affected since the complex could only be detected as a retarded band at 4°C (Fig 7) but not at 40°C (Fig 6B). In correlation with the reduced stability of the IHF-DNA complex the stimulation of Pc transcription is also reduced (Table 1). The stimulation of Pe however is less affected and it is significantly higher than that of mutant 2. So apparently, comparable to what was observed for mutant pGP185-*BL*, a

Table 1a. GalK activities of Pe-galK fusion plasmids. Each GalK value is the average of at least 4 measurements. The deviation in the GalK activities varied from 15–30%

plasmid	AB1157 IHF+	PP1954 IHF−	factor of stimulation
pGP188	80	4	20.0
pGP188.1	73	4	18.3
pGP188.2	29	4	7.3
pGP188.1+2	56	4	14.0
pGP188.3+4	4	3	1.3
pGP185	346	45	7.6
pGP185.*BL*	316	44	7.2

Table1b. GalK activities of Pc-*galK* fusion plasmids.

plasmid	AB1157 IHF+	PP1954 IHF−	factor of stimulation
pGP133	35	6	5.8
pGP133.1	54	8	6.8
pGP133.2	30	6	5.0
pGP133.1+2	18	7	2.6
pGP133.3+4	5	5	1.0
pGP133.*BL*	5	6	0.8

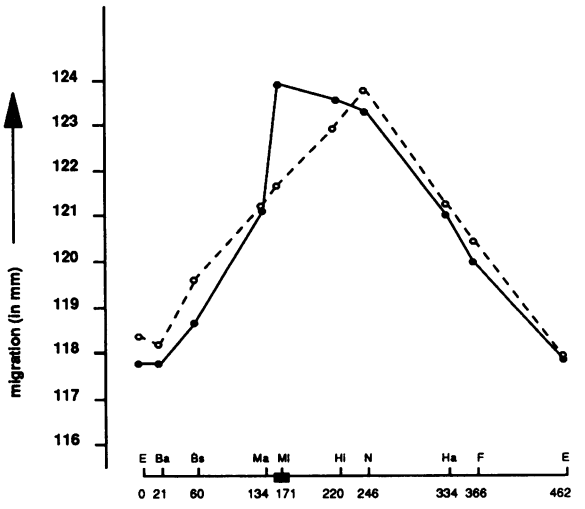

Figure 11. Relative mobilities of permuted fragments. The different permuted fragments as indicated in Fig 10 were electrophoresed on a native 10% polyacrylamide gel at 4°C. The migration distances of the fragments are plotted against the positions of the different restriction sites. The wild type DNA is indicated by open symbols and mutant 1+2 DNA by filled symbols. The *ihf* site is indicated by a filled box and the relevant restriction sites are: E=*Eco*RI, Ba=*Bam*HI, Bs=*Bst*NI, Ma=*Mae*III, Ml=*Mlu*I, Hi=*Hin*dIII, N=*Nhe*I, Ha=*Hae*III, F=*Fok*I.

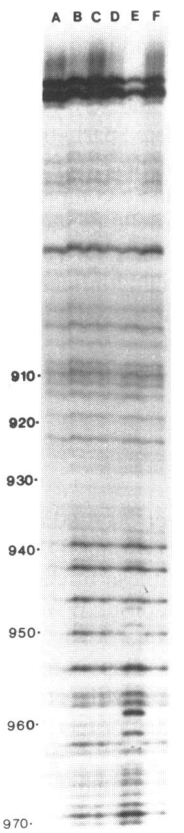

Figure 12. Pattern of partial DNaseI digestion of the 'upper' strand in the absence of IHF. The *Cla*I-*Hin*dIII fragment was labeled at the *Hin*dIII site. Positions are indicated with respect to the left end of Mu. Lane **A**, **B** and **F**; wild type *ihf*, lane **C**; mutant 1, lane **D**; mutant 2 and lane **E**; mutant 1+2.

weakened interaction between IHF and its binding site does not hamper Pe activation but has a large effect on Pc.

How can we explain the observed difference between mutant 2 and mutant 1+2 with regard to the Pe stimulation? The affinity of IHF for mutant 1+2 was shown to be much lower than for mutant 2, but still Pe is more activated in mutant 1+2.

The DNaseI digestion pattern (Fig. 12) of the DNA of mutant 1+2 (in the absence ofproteins) is clearly different from the pattern of mutant 1, mutant 2 or wild type DNA. Mutant 1+2 DNA shows a reduction of the DNaseI attack at positions 950 and 962, whereas cleavages at positions 947, 949, 959 and 961 are enhanced. Although these enhanced cleavages are located at the positions of the introduced mutations it is clear that it is the combination of the four substitutions that is responsible for this phenomenon, since the DNaseI patterns of mutant 1 and mutant 2 are the same as for wild type DNA. Apparently the four base substitutions induce a conformational change in the DNA which in some way may be involved in the unexpected high level of transcription stimulation in mutant 1+2. To study the conformational change in more detail we compared the sequence-directed bend that is present in the wild type DNA with the sequence-directed bend of the mutant 1+2 DNA by the permutation assay (Fig. 11). In the wild type DNA the center of the sequence-directed bend had already been shown to be located near the *Nhe*I site, corresponding to position 1032 of the Mu sequence. In the mutant 1+2 DNA the center of the bend seems to be shifted in the direction of the *Mlu*I site which is in the *ihf* site at position 955. Since the entire regulatory region of Mu (from position 860 to 1115) seems to be curved the shift in center of the bend indicates that the mutations in mutant 1+2 create a relatively sharp bend in the DNA near the *ihf* site. Such a local conformational change of the DNA might facilitate the interaction between IHF and RNA polymerase at Pe, thereby enhancing the IHF-stimulated early transcription in mutant 1+2.

In summary we can conclude from the transcription data that for stimulation of Pc a stable complex of IHF and DNA seems to be required. In contrast Pe can also be stimulated when the interaction between IHF and DNA is much weaker, but this stimulation seems to be highly influenced by changes in the local DNA conformation, possibly by changing the interaction between IHF and RNA polymerase.

DISCUSSION

In the past five years over 30 binding sites for IHF in *E. coli* phages, plasmids or chromosomal DNA have been reported (for a review see 11). All of these binding sites contain sequences that closely match the consensus WATCAANNNNTTR (5,6,7). On the other hand also sequences have been reported that show a perfect match to the consensus, but that bind IHF only poorly or not at all (8,10). Apparently for efficient binding of IHF there is more involved than just the primary sequence of the consensus. The IHF binding site that plays a role in the regulation of early and repressortranscription of phage Mu is a good tool to study the DNA requirements for IHF binding. It contains two inverted partially overlapping consensus sequences (*ihfa* and *ihfb*) for one of which (*ihfb*) IHF shows a strong preference. The experiments reported in this paper show that even an *ihfb* containing two mutations is preferred over an *ihfa* that almost perfectly matches the consensus.

How can IHF discriminate between both consensus sequences? The bases in the center of the consensus can not be the cause for the preferential binding to *ihfb* since they are exactly the same for both sequences. Apparently the flanking sequences must play a crucial role in the recognition of an IHF binding site. Lee et al (31) reported that mutations in the A-rich region that flanks the *ihf* consensus of the H' site in lambda have a drastic effect on IHF binding. This same region has been reported to make a close contact with IHF (5). In full agreement we observed that deletion of two of the three adenines that flank the *ihfb* site reduce the affinity of IHF for this site. It is not likely however that the stretch of A's that flanks *ihfb* and that is absent in the comparable flanking region of *ihfa* is the bases for the discrimination between both sites since also in the deletion mutant only *ihfb* is bound by IHF. Furthermore several functional IHF binding sites have been reported that contain no A residues at all in the comparable flanking region, e.g. the *IIa* and *IIb* sites of pBR322 (9), the *ihf* site in the *ori* of pSC101 (27) and the *H'* site in *attP* of phage P22 (8). Therefore we suggest that whether a consensus sequence is an IHF binding site or not, is not determined primarily by the sequence of the flanking DNA but by its secondary structure.

The experiments with the base substitution mutants described in this paper confirm this hypothesis. The affinity of IHF for mutant 1 (which contains two base substitutions just adjacent to the 5'-site of *ihfb*) or mutant 2 (containing two substitutions in *ihfb*) as determined by DNaseI footprint and gel retardation is hardly affected, whereas the binding of IHF to mutant 1+2 (containing all four base pair substitutions) is severely reduced. This difference in binding of IHF to mutant 1 and mutant 2 DNA on one hand and to mutant 1+2 DNA on the other hand can not solely be explained by the difference in the primary sequence of the DNA's but is probably due to a combined effect of the mutations in mutant 1+2 on the secondary structure. That the introduction of four GC base pairs in a DNA region that is extremely A/T-rich (75%) induces a conformational change in the DNA is clear from the DNaseI cleavage pattern of the mutant 1+2 DNA (in the absence of IHF). Gel retardation essays showed that this conformational change also affects the local bending of the DNA. We therefore argue that the bending of the DNA might be an important factor in the recognition of an *ihf* site and the stability of an IHF-DNA complex. Recently it has been shown (29) bycomparing the binding of IHF to different DNA sequences that do not contain the IHF consensus, that there is a preference for binding to curved DNA. It is conceivable that such a preference also exists for curved sequences that do contain the consensus. The DNA of the Mu regulatory region shows an overall bent structure. The sequence directed bend present in the IHF binding region could promote the IHF-induced bending resulting in a more stable complex. A requisite for such a promotion however is that both the sequence- and IHF-induced bend should be in the same direction. The local strong bending that seems to be present near the IHF binding site in mutant 1+2 might be in an unfavourable direction, thereby reducing the affinity for IHF. The sequence directed bend in the wild type DNA might also be the cause for the preferential binding of *ihfb*.

The gel-retardation experiments showed that the presumed IHF-*ihfa* complex (mutant 3+4) is less retardet than the IHF-*ihfb* complex (wild type). This might suggest that the center of the IHF induced bending in the IHF-*ihfa* complex is located closer to the *Hin*dIII end of the fragment used than in the IHF-*ihfb* complex. In other words, the center of the IHF-induced bending is expected to be located asymmetrically with respect to the WAT-

CAANNNNTTR consensus, around the left part of this sequence. The WATC part of *ihfa* is located 12 base pairs from the WATC part of *ihfb* in the complementary strand and thus at the opposite site of the DNA helix. If the center of the IHF induced bend would be within this WATC sequence, the IHF-induced bend in the case of recognition of the *ihfa* site would be in opposite direction of the bend induced by binding of IHF to *ihfb*. A sequence directed bend that favours binding to *ihfb* will in that case antagonize the binding to *ihfa*.

The DNA bending induced by IHF is believed to be important for the different processes in which the host factor is involved. In the *attP* site of lambda the central IHF binding site (*H2*) can be replaced by another source of DNA bending like the CRP binding site or 'A-tract' DNA (28). This does not necessarily mean that the bending of the DNA is the only function of IHF in all IHF-regulated processes. Mutants of IHF have been isolated that are inactive in one IHF-regulated process and active in another one (D.Roberts, pers.comm.), which suggests that IHF can act in more than one way. Also the mechanisms by which transcriptions from Pe and Pc in phage Mu are stimulated by IHF are distinct, since mutations in the Mu DNA can be isolated that abolish the stimulation from one promoter without affecting the stimulation of the other promoter (16 and this paper). What is most striking in the case of Mu however is that IHF stimulates both transcriptions by binding to only one site in the regulatory region.

From the results presented in this paper it is clear that mutations that affect the stability of the IHF-DNA complex also affect the stimulation of Pc transcription. For Pe activation however a relatively weak complex seems to be sufficient. The most striking example of this difference is observed with the mutant that lacks two of the three A residues in the region flanking the 5′ site of the consensus. Due to the reduced affinity of IHF for this region the stability of the IHF DNA complex is strongly reduced and the Pc stimulation is completely lost. In contrast the stimulation of Pe transcription is perfectly normal, indicating that a stable interaction of IHF with the 5′ flanking region is not essential in this process. Previous experiments (16) already suggested that the activation of Pe directed transcription is mediated by protein-protein interactions between IHF and RNA polymerase. Apparently also in weaker IHF-DNA complexes such an interaction with RNA polymerase can still occur and it is unlikely that the bending of the DNA by IHF plays a crucial role in this process. The positioning of both proteins on the DNA does play an important role in Pe activation as was shown by the helix dependency of the distance between Pe and *ihf* (16). We therefore predict that the mutations in mutant 2 slightly change the relative positioning of IHF and RNA polymerase, since this mutant shows no reduced binding of IHF, but does show a reduced stimulation of Pe. The more drastic conformational change in the DNA caused by the mutations in mutant 1+2 which on one hand result in a reduced affinity of IHF for this DNA might on the other hand improve the relative positioning of IHF and RNA polymerase. It has already been shown that also the insertion of 1 bp between the *ihf* site and Pe seems to enhance the contact between IHF and RNA polymerase (16).

For Pc activation the necessity for a stable IHF-DNA complex and the fact that such a complex can also be active at a distance of over 200 bp (16) suggest that a conformational change of the DNA is the basis of this process. The IHF induced bending of the DNA could be very important in this. Since no extensive studies on the conformation of the Mu regulatory region in the presence or absence of IHF have been done, one can only speculate about the nature of the proposed IHF-induced conformational change.

A possible mechanism by which changing of the DNA conformation could modulate Pc transcription has already been suggested by Higgins et al (30). They postulated that the IHF bending in a supercoiled molecule is stably localized to the apex of a negative supercoiled loop thereby positioning the Pe and Pc promoters. In their model the positioning of the Pc promoter by IHF is in such a way that transcription is inhibited. This model was based on their results obtained *in vitro* which showed that on a supercoiled molecule IHF inhibits Pc transcription. This paper and previous work (13,15,16) however show that *in vivo* Pc transcription is activated by IHF. Thediscrepancy between the *in vivo* and *in vitro* results is most probably due to the difference in strength of the Pe promoters used in these studies. Van Rijn et al showed that the effect of IHF on Pc transcription is highly dependent on the strength of the Pe promoter (15). In the presence of the strong wild type promoter (as used by Higgins in the *in vitro* studies) Pc transcription is inhibited by IHF, whereas in the presence of a weaker mutated Pe (as used by us in *in vivo* studies) Pc is stimulated. Therefore the inhibition of Pc transcription by IHF observed *in vitro* is most probably not a direct effect of IHF but the indirect result of the stimulation of the opposing Pe transcription.

To explain the stimulation of Pc transcription by IHF we would like therefore to propose an adapted model. The bending that is present in the Mu regulatory region as a result of the many tracks of A residues might (in the absence of IHF) also result in a more or less specific topological orientation of this region in a supercoiled molecule. If the position of the Pc promoter within such a molecule is unfavourable for RNA polymerase binding (e.g. in the tip of a branch), repressor transcription would consequently be relatively low. The binding of IHF might result in a different topological orientation, now placing the *ihf* site in the tip of the branch. As a result the Pc promoter might come in a more favourable position resulting in an elevated transcription. For such a topological change of the Pc region, the *ihf* site has to be in the vicinity (within a few hundred base pairs) of the promoter, but no absolute distance or helical orientation will be required.

REFERENCES

1. Nash,H.A. (1981) *Ann.Rev.Genet.* **15**, 143–167
2. Miller,H.I. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 691–698
3. Flamm,E.L. and Weisberg,R.A. (1985) *J. Mol. Biol.* **183**, 117–128
4. Drlica,K. and Rouviere-Yaniv,J. (1987) *Microbiol. Rev.* **51**, 301–319
5. Craig,N. and Nash,H.A. (1984) *Cell* **39**, 707–716
6. Goodrich,J.A., Schwartz,M.L. and McClure,W.R. (1990) *Nucleic Acid Res.* **18**, 4993–5000
7. Yang,C.-C. and Nash,H.A. (1989) *Cell* **57**, 869–880
8. Leong,J.M., Nunes-Düby,S., Lesser,C.F., Youderian,P., Susskind,M.M. and Landy,A.(1985) *J. Biol. Chem.* **260**, 4468–4477
9. Gamas,P., Chandler,M.G., Prentki,P. and Galas,D.J. (1987) *J. Mol. Biol.* **195**, 261–272
10. Prentki,P., Chandler,M.G. and Galas,D.J. (1987) *EMBO J.* **6**, 2479–2487
11. Friedman,D.I. (1988) *Cell* **55**, 545–554
12. Thompson,J.F. and Landy,A. (1988) *Nucleic Acids Res.* **16**, 9687–9705
13. Goosen,N., van Heuvel,M., Moolenaar,G.F. and van de Putte,P. (1984) *Gene* **32**, 419–426
14. Krause,H.M. and Higgins,N.P. (1986) *J. Biol. Chem.* **261**, 3744–3752
15. van Rijn,P.A., Turk,S.C.H.J., Goosen,N. and van de Putte,P. (1989) *Nucleic Acids Res.* **17**, 10203–10212

16. van Rijn,P.A., Goosen,N. and van de Putte,P. (1988) *Nucleic Acids Res.* **16**, 4595–4605
17. Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 488–492
18. Groenen, M.A.M., Kokke,M. and van de Putte,P. (1986) *EMBO J.* **5**, 3687–3690
19. Messing,J. and Vieira,J. (1982) *Gene* **19**, 269–276
20. Hanahan,D. (1985) in *DNA Cloning. A Practical Approach. Techniques for transformation of E.coli.* (Glover,D.M.,ed) vol. 1 IRL Press, Oxford, pp. 109–135
21. van Sluis,C.A., Moolenaar,G.F. and Backendorf,C. (1984) *EMBO J.* **2**, 2313–2318
22. Norrelander,J., Kempe,T. and Messing,J. (1983) *Gene* **26**, 101–106
23. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, NY
24. Sanger,F., Nicklen,S. and Toulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467
25. McKenney,K., Shimitake,H., Court,D., Schmeissner,U., Brady,C. and Rosenberg,M. (1981) in *Gene Amplification and Analysis* (Chririkan,J.G., and Papas,T.S., eds) pp. 383–415
26. Wu, H.M. and Crothers,D.M. *Nature* **308**, 509–513 (1984)
27. Stenzel,T.T., Patel,P. and Bastia,D. (1987) *Cell* **49**, 709–717
28. Goodman,S.D. and Nash,H.A. (1989) *Nature* **341**, 251–254
29. Bonnefoy,E. and Rouviere-Yaniv,J. (1991) *EMBO J.* **10**, 687–696
30. Higgins,N.P., Collier,D.A., Kilpatrick,M.W. and Krause,H.M. (1989) *J.Biol.Chem.* **264**, 3035–3042
31. Lee,E.C., MacWilliams,M.P., Gumport,R.I. and Gardner,J.F. (1991) *J.Bacteriol.* **173**, 609–617