The EMBO Journal vol.6 no.8 pp.2479–2487, 1987

Escherichia coli integration host factor bends the DNA at the ends of IS*1* and in an insertion hotspot with multiple IHF binding sites

Pierre Prentki, Michael Chandler[1] and David J.Galas

Molecular Biology, University of Southern California, Los Angeles, CA 90089-1481, USA and [1]Centre de Recherche de Biochimie et Génétique Cellulaires du CNRS, 118 route de Narbonne, 31062 Toulouse Cedex, France

Communicated by L.Caro

The integration host factor of *Escherichia coli* (IHF) is a small, histone-like protein which participates in the integration of bacteriophage λ into the *E. coli* chromosome and in a number of regulatory processes. Our recent footprinting analysis has shown that IHF binds specifically to the ends of the transposable element IS*1*, as well as to several sites within a short segment of the plasmid pBR322. We have extended our studies of the binding of the IHF molecule to these sites *in vitro* using a gel retardation assay. We report here that IHF bends the DNA upon binding, as judged from the strong cyclic dependence of the protein-induced mobility shift on the position of the binding site. Using cloned, synthetic ends of IS*1* as substrates, we have found that some mutations within the conserved bases of the IHF consensus binding sequence abolish binding, and that alterations of the flanking sequences can greatly reduce IHF binding. The presence of multiple IHF sites on a single DNA fragment increases binding very little, indicating that IHF does not bind cooperatively in this complex. We discuss the possibility that DNA bending is related to the role IHF plays in forming and stabilizing nucleoprotein complexes, and suggest that bending at the IHF sites may be important to its diverse effects in the cell.

Key words: IHF/transposition/DNA bending/mobility of protein–DNA complexes/protein–DNA sliding

Introduction

DNA is subject to a number of structural variations that may differ sharply from B form. Indeed, B-form DNA itself has been shown to exhibit sequence-dependent variations in its helix parameters. One of the best-studied phenomena of this type is the bending of DNA. Bending can be intrinsic to a DNA sequence (Marini *et al.*, 1982; Bossi and Smith, 1984; Wu and Crothers, 1984; Koo *et al.*, 1986) or it can be induced by the binding of proteins such as CRP (Gronnenborn *et al.*, 1984; Wu and Crothers, 1984), restriction enzyme *Eco*RI (Frederick *et al.*, 1984), Cro repressor (Ohlendorf *et al.*, 1982) or initiators of plasmid (Mukherjee *et al.*, 1985) or phage (Zahn and Blattner, 1985) replication. Some proteins such as the cI repressors from phage λ (Griffith *et al.*, 1986) and 434 (Anderson *et al.*, 1985) or the lactose repressor (Wu and Crothers, 1984), however, do not bend the DNA detectably. Although a direct demonstration of bending requires spectroscopic techniques or electron microscopy, it has been shown by Crothers and co-workers that bent DNA can be detected by the anomalous mobilities of DNA fragments in polyacrylamide gels. We have used this gel retardation technique to study the specific interaction of the integration host factor (IHF) protein with DNA.

The IHF of *Escherichia coli* is a heterodimeric protein consisting of the products of the *him*A gene (11.2 kd; Miller, 1984) and the *him*D or *hip* gene (10.5 kd; Flamm and Weisberg, 1985). It is required for integration of bacteriophage λ both *in vivo* (Miller and Friedman, 1977; Williams *et al.*, 1977) and *in vitro* (Kikuchi and Nash, 1978); for a review, see Weisberg and Landy (1983). IHF has been found to participate in a variety of other processes in the cell, including gene expression (Hoyt *et al.*, 1982; Friedman *et al.*, 1984; Goosen *et al.*, 1984; Mahajna *et al.*, 1986; Gamas *et al.*, 1987a) and DNA replication (Gamas *et al.*, 1986).

Previous studies of phage attachment sites (Craig and Nash, 1984; Leong *et al.*, 1985) established that IHF binds in a sequence-specific manner to a set of sequences with the consensus YAA . . . RTTGAT(A,T). The exact mode of IHF binding still remains to be elucidated. Our recent finding that IHF binds specifically to a single site at each end of the transposable element IS*1* (Gamas *et al.*, 1987b) led us to investigate the nature of the interaction of IHF with DNA other than phage attachment sites. We found that in plasmid pBR322, there is a small segment of ~300 bp, derived from transposable element Tn*3*, that contains four binding sites for IHF. This region constitutes an insertion hotspot for IS*1* (Zerbib *et al.*, 1985). It is this region and the ends of IS*1*, separately and together, that have been used to study the binding of IHF to DNA.

Results

Binding of IHF to DNA fragments carrying a single site

We have identified by footprinting several IHF binding sites in pBR322 and one on either end of the transposable element IS*1* (Gamas *et al.*, 1987b). We have also constructed synthetic IS*1* ends, which include the IHF-binding sites (Prentki *et al.*, 1987). To evaluate the strength of binding of IHF to these sites, we first examined the equilibrium distribution of bound and unbound DNA on isolated fragments, using the gel retardation technique (Fried and Crothers, 1981; Garner and Revzin, 1981).

The test plasmids used here are pBR322 and derivatives carrying synthetic IS*1* ends. Cleavage of pBR322 with *Eco*RI and *Ssp*I yields a fragment of 194 bp which carries two IHF-binding sites, while cleavage with *Eco*RI and *Hin*PI yields a fragment of 105 bp with one IHF site (Figure 1). The structure of plasmids of the pMP series is shown in Figure 2. The ends of IS*1* flank a selectable DNA fragment (Ω; Prentki and Krisch, 1984) inserted into the *Eco*RI site of pBR322. Plasmid pMP20 carries the terminal 28 bp of the left end of IS*1* (Figure 2). Cleavage of pMP20 with *Hin*dIII, *Ssp*I and *Eco*RV yields fragments of 156 bp (*Hin*dIII/*Eco*RV), 86 bp (*Hin*dIII) and 248 bp (*Hin*dIII/*Ssp*I), which carry zero, one and three IHF sites respectively. In this section, we first examine binding of DNA fragments carrying one site only.

The results of gel retardation assays using end-labelled pBR322 and pMP20 DNA fragments are shown in Figures 3 and 4. The 105-bp fragment of pBR322 (Figure 1, lanes e–g) or the 86-bp fragment of pMP20 (Figure 2A), each of which carry one IHF site, are progressively chased into unique retarded DNA–protein

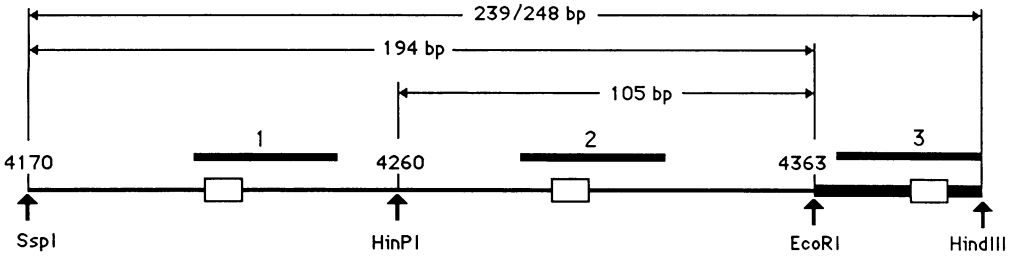

Fig. 1. IHF-binding sites and footprints within the DNA fragments used in this study. The *Ssp*I–*Eco*RI segment originates from pBR322; the coordinates are those of Sutcliffe (1978) as revised by Peden (1983). IHF sites 1 and 2 (boxes) and footprints (thick horizontal line) correspond to sites I and IIa of pBR322 (Gamas *et al.*, 1987b). Site 3 belongs to synthetic DNA segments whose sequences are found in Figure 2.

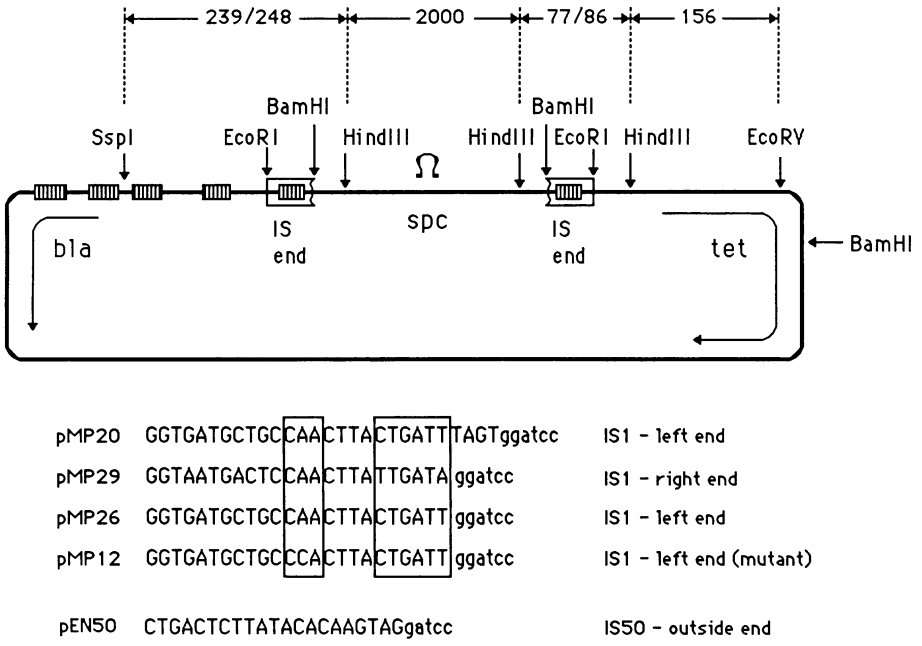

Fig. 2. Plasmid structures. All plasmids are derivatives of pBR322. The sites of IHF binding (hatched boxes) in plasmid pBR322 are in a region derived from the transposon Tn*3* between the genes for resolvase (*tnp*R) and beta-lactamase (*bla*). The role of this cluster of sites is unknown. The synthetic IS ends are indicated by boxes. Each plasmid carries the same sequence (shown at the bottom of the figure) in an inverted repeat configuration at both ends of the Ω fragment. The sizes of the end-labelled fragments are shown at the top (arrows). Fragments carrying synthetic ends may differ in size depending on the length of the IS end.

complexes as the concentration of IHF increases. Since the protein was in large molar excess over the DNA, we estimate the binding constants for the single site carried by the 86-bp pMP20 and 105-bp pBR322 fragments (see Materials and methods) as 10^8 M^{-1} and 4×10^7 M^{-1} respectively. The absence of binding to fragments of similar size carrying no IHF site indicates that the binding represented by these constants is specific in character.

We have also investigated binding to fragments in which the synthetic IS*1* end has been altered. The 82-bp fragments from pMP29 and pMP26, which carry shorter (24 bp) variants of the wild-type right or left end of IS*1* respectively (Figure 2) exhibit binding constants of 3×10^7 M^{-1} (pMP29) and 10^7 M^{-1} (pMP26). The more efficient binding to the pMP29 fragment may be a reflection of the higher match to the IHF consensus (note the T at position 8). However, the significant reduction of binding of pMP26, compared with pMP20, by changes that did not alter the nucleotides specified by the consensus binding site, suggests that sequences flanking this consensus influence the interaction with the protein. The 82-bp fragment from pMP12, in which a highly conserved A at position 2 of the consensus was replaced with a C (Figure 2) showed no detectable binding (Figure 4D). The same was true for the 77-bp fragment of plasmid pEN50 (Figure 4E), carrying the first 21 bp of IS50, which do not include an IHF site.

Correlation between the number of sites and the number of retarded bands

In the presence of IHF the 194-bp fragment of pBR322 exhibits two bands of lower mobility whereas the 105-bp fragment exhibits only one (Figure 3). Similarly, the 248-, 86- and 156-bp fragments of pMP20 exhibit respectively, three, one and zero retarded bands (Figure 4A). There is thus a strict correlation between the number of specific binding sites and the number of retarded bands. This is also apparent in the behaviour of the 244-bp fragments of pMP29 (Figure 4B) and pMP26 (Figure 4C), which also display three retarded bands. When the number of binding sites on the fragment is reduced to two by a mutation at the IS*1* end as in pMP12 (Figure 4D), or by substituting it with an IS50 end as in pEN50 (Figure 4E), only two retarded bands are observed. The same correlation is apparent in recent binding studies with IHF (Thompson *et al.*, 1986) or with the

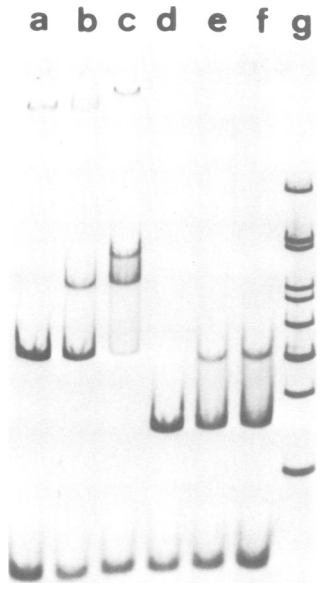

Fig. 3. IHF-induced gel retardation. DNA of plasmid pBR322 was end-labelled at the unique *Eco*RI site, recut either with *Ssp*I + *Hin*dIII **(lanes a–c)** or *Hin*PI + *Hin*dIII **(d–f)**, and electrophoresed onto a 5% polyacrylamide gel (1.5 mm thick), after prior incubation with purified IHF protein in the following estimated active amounts: 0 nM **(a** and **d)**, 8 nM **(b** and **e)** or 16 nM **(c** and **f)**. An end-labelled digest (*Eco*RI + *Hin*fI) of pBR322 **(g)** served as a size standard. Following electrophoresis the gel was dried and autoradiographed.

resolvase of γδ (Hatfull and Grindley, 1986). At higher IHF concentrations, we also see lower mobility complexes in a weak, diffuse group above the third band (data not shown). These are probably due to non-specific binding.

It is thus reasonable and consistent with the observations on other proteins to assume that the third band from the three-site fragment, and the second band from the two-site fragment, represent complexes in which all sites are occupied. Indeed, footprinting of this region of pBR322 at similar protein concentrations has demonstrated that all the sites are occupied (Gamas *et al.*, 1987b). The intermediate bands (I and II in Figure 4A), on the other hand, could represent either (i) complexes in which one (band I) or two (band II) sites are stably occupied; or (ii) complexes in which one (band I) or two (band II) protein molecules are distributed over all available sites but occupy a given site as a function of the relative binding affinity of that site. Results presented later favour the second hypothesis.

IHF induces DNA bending

It has been suggested that the role of IHF in the formation of the condensed nucleoprotein structure at λ attP is to introduce bends in the DNA (Craig and Nash, 1984). It was therefore of interest to determine whether IHF alone is able to induce changes in DNA conformation. To test this possibility, we constructed a plasmid that could be used to generate DNA fragments of identical length and sequence composition with a single site for IHF binding located at different positions. Using this set of fragments we could then determine the mobility of the DNA–protein complex as a function of the position of the protein on the DNA. The scheme is based on that devised by Wu and Crothers (1984) to examine sequence-directed bending of kinetoplast DNA and protein-induced bending of DNA by the CRP protein in the lactose operon control region.

A 410-bp *Bam*HI fragment of pMP20 (Figure 2) was inserted into the *Bam*HI site of pBR322, in an orientation which restores resistance to tetracycline. The resulting plasmid, pMP40 (Figure 5a) carries an IHF site (the synthetic IS*1* end) flanked by a tandem duplication of pBR322 between the unique *Eco*RI and *Bam*HI sites (coordinates 1–375). This segment of pBR322 carries no detectable IHF sites (data not shown). IHF gel retardation assays (Figure 5b) using the fragments shown in Figure 5a demonstrate a striking variation in mobility of the various complexes as a function of the position of the site. The relative mobilities, as represented by the m_r value of the complexes (see Materials and methods), are plotted in Figure 5c as a function of the position of the end of the fragment with respect to the central *Bam*HI site. Note that the curves reach the point of minimum mobility precisely where the IHF site would be at the centre of the fragment. This result is diagnostic of bending of the DNA (Wu and Crothers, 1984). The data indicate that the bend centre lies very close to the IHF site, between the *Bam*HI and *Eco*RI sites, but closer to the *Bam*HI site.

We have measured the m_rs for these fragments at two different gel concentrations. Retardation of the protein–DNA complexes is enhanced at higher concentrations (see below), although the curves have their maxima at the same position and retain the same shape (Figure 5c).

Mobilities of multi-site fragments

Since the mobility of IHF–DNA complexes is strongly dependent on the position of the site due to bending of the DNA, a fragment carrying multiple sites would be expected to experience different successive shifts in mobility as proteins fill the available sites. Thus we should be able to gain information about the occupation of different sites on such a fragment from the succession of mobility shifts.

To make such inferences, however, we must make several assumptions about the behaviour of the complexes in the gel. As shown by Fried and Crothers (1983) for the lactose repressor, the ratio of protein to DNA molecules in a retarded band remains constant (and an integer), and the transfer of proteins between molecules is slow, presumably due to the trapping of complexes in the 'cage' of the gel. On the other hand, we would expect transfer of a protein molecule from site to site on the same DNA molecule (Winter *et al.*, 1981; Berg and von Hippel, 1985) to be as rapid as in solution. If this process were slow the band would spread out because there would be a fluctuating range of mobilities, corresponding to the positions of the bound protein. Thus when we say that binding to one particular site is primarily responsible for the mobility we assume a dynamic equilibrium in which the protein 'sees' all the sites but spends most of the time on that particular site. An additional argument against the view that during electrophoresis the proteins remain bound to a particular site arises from the observation that IHF bends the DNA: each complex with a different arrangement of proteins on the sites would have a different mobility. Thus for a three-site fragment we would expect to see a total of seven retarded bands, rather than three.

To investigate the contributions of the individual sites to the mobilities of multi-site complexes, we can examine the effects of alterations in the binding properties of one site (the synthetic IS*1* end at the extremity of the *Hin*dIII–*Ssp*I fragments; see Figure 1). We know the binding constants of each of these sites from the behaviour of the short *Hin*dIII fragments (see above).

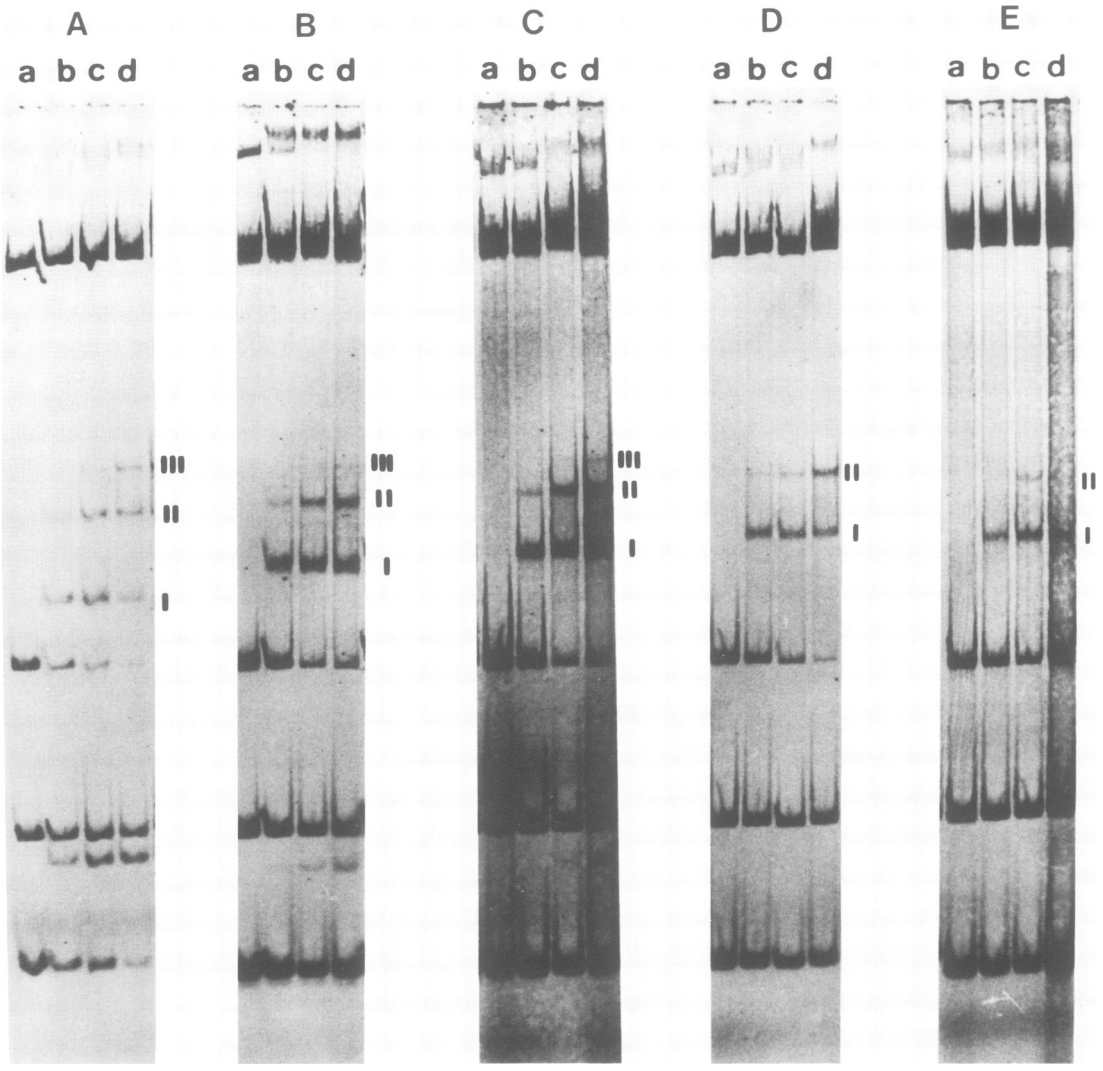

Fig. 4. Binding of IHF to DNA fragments carrying synthetic binding sites. DNA of plasmids pMP20 **(panel A)**, pMP29 **(panel B)**, pMP26 **(panel C)**, pMP12 **(panel D)** and pEN50 **(panel E)** were end-labelled at the *Hin*dIII sites, recut with *Ssp*I + *Eco*RV and incubated with purified IHF protein: 0 nM **(lanes a)**, 2.5 nm **(lanes b)**, 5 nM **(lanes c)** and 10 nM **(lanes d)**. Electrophoresis was performed on a 0.4 mm thick native 'sequencing' -size gel. Complexes of the *Hin*dIII–*Ssp*I fragments are indicated (I, II and III).

We pointed out earlier the correspondence between the number of IHF binding sites and the number of retarded bands. Close inspection of the position of each of these retarded bands (Figure 4) clearly shows that for fragments with three active sites, only band III (all sites occupied) displays identical mobilities. Bands I (one protein bound) and II (two proteins bound) are strikingly different in each plasmid. Moreover, the mobilities of bands I and II are progressively reduced as the affinity of site 3 (Figure 1) is reduced. It appears that by reducing the affinity of site 3, we caused the binding to shift towards the center. As reported above, the binding constant for site 3 in pMP20 is greater than for site 2. In pMP29, the binding constants of sites 2 and 3 are comparable. The shift toward the centre that we would expect if the sites are occupied preferentially according to these binding constants is thus borne out.

Assessment of cooperativity

Cooperative protein–protein interactions (involving IHF) have recently been observed at the attP locus of phage λ (Bushman *et al.*, 1985). The experiments presented above can also be used to ask whether a weak site (e.g. site 3 of pMP29) is made stronger by adjacent binding of IHF molecules on the same DNA molecule. It is apparent (Figure 4B) that all three sites on the 244-bp fragment are occupied at protein concentrations which are comparable with those necessary to obtain a similar level of binding of site 3 alone (82-bp fragment). This suggests that there is little or no cooperative binding at site 3. We can determine the limits of the effect by estimating the effective binding constant of site 3 in the presence of two bound proteins. This quantity is determined by measuring the fragments bound in bands II and III and taking the ratio. Several independent experiments indicate that this factor is not greater than 2. For this particular arrangement of binding sites, it appears that there is no appreciable cooperativity between IHF molecules. The same is true for pMP26 (Figure 4C).

A similar lack of cooperativity was observed on linear fragments when one of the three IHF sites within attP was mutated (Gardner and Nash, 1986). However, the observation that some attP regions with mutated IHF sites still support site-specific recombination (Gardner and Nash, 1986; Thompson *et al.*, 1986),

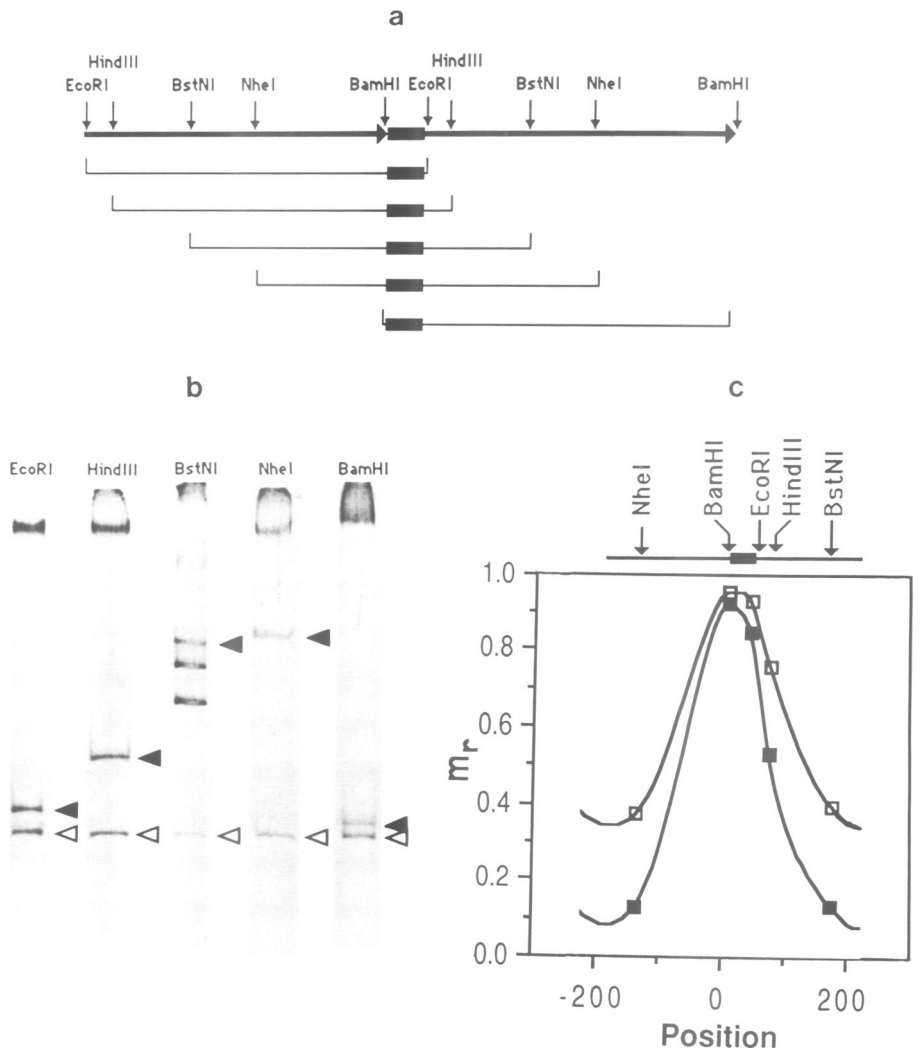

Fig. 5. Circular permutation of the IHF binding site to test for bending. **(a)** Structure of the relevant region of plasmid pMP40, a derivative of pBR322 carrying a tandem duplication of the 375-bp *Eco*RI−*Bam*HI fragment bracketing the terminal 28 bp at the left end of IS*1* (black rectangle). Digestion of pMP40 with the set of restriction enzymes indicated by arrows is formally equivalent to cutting at various positions a circle of DNA consisting of the IHF site joined to the ends of the 375-bp fragment. **(b)** Electrophoretic analysis. End-labelled fragments of pMP40 DNA cut with the indicated enzyme were incubated with purified IHF protein (10 nM) and electrophoresed on a 5% polyacrylamide gel. Bound and unbound DNA fragments are indicated with filled or empty triangles respectively. Additional fragments in the *Bst*NI lane result from the multiple cuts in the pBR322 backbone. **(c)** Mapping the bend. Relative mobilities (m_r) of protein−DNA complexes are plotted against the map position of the restriction site, such that each position corresponds to a fragment obtained by digestion with that restriction enzyme, as shown in **(a)**. Coordinate zero is arbitrarily assigned to the centre of the *Bam*HI site. Each point represents the average of several measurements, performed on 5% (□) or 7.5% (■) polyacrylamide gels.

as well as binding studies on supercoiled substrates (Richet *et al.*, 1986), suggest that cooperative protein−protein interactions involving IHF may require supercoiling.

Dependence of electrophoretic mobilities on gel concentration

Various studies on the behaviour of DNA fragments have shown that anomalous mobility is emphasized as the gel concentration is increased (Marini *et al.*, 1982; Ross *et al.*, 1982; Brown and Smith, 1977). The physical explanation for this effect is not known, but it is reasonable to assume that as the average pore size of the gel decreases, the more compact structure of bent DNA migrates more slowly because of an increase in effective cross-section of the molecule (Marini *et al.*, 1982).

We have examined the mobilities of fragments carrying IHF sites, and of IHF−DNA complexes as a function of acrylamide concentration. Typical results are shown in Figure 6. Two features of the data plotted here are striking. The dependence of mobility on gel concentration is linear over a wide range of conditions: this is the behaviour some previous work leads us to expect (Stellwagen, 1983; and references therein). An unexpected feature of the plots is that when extrapolated to zero gel concentration the mobilities (m_rs) converge very close to a value of one. Examination of other protein−DNA complexes (to be published elsewhere) suggests that the extrapolation to approximately one may be a peculiarity of IHF. This difference in electrophoretic behaviour may reflect more profound differences in the structure of DNA−protein complexes. Once understood, it is conceivable that the slopes and intercepts of these gel concentration plots may provide a simple method to characterize protein−DNA complexes.

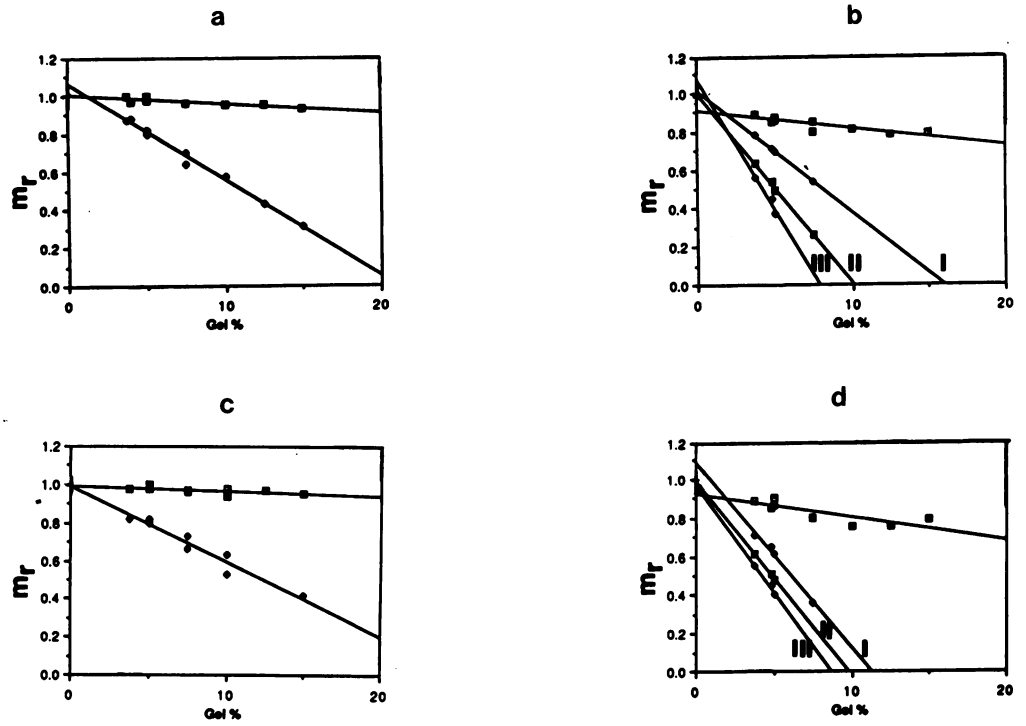

Fig. 6. Mobilities of DNA fragments and IHF–DNA complexes as a function of gel concentration. The diagrams summarize measurements performed on the 86-bp **(a)** and 248-bp **(b)** fragments of pMP20 and the 82-bp **(c)** and 244-bp **(d)** fragments of pMP29. Relative mobilities (see Materials and methods) of free DNA fragments are indicated by empty squares. The mobilities of the first, second and third retarded bands on multi-site fragments are indicated by I, II and III respectively (see Figure 4).

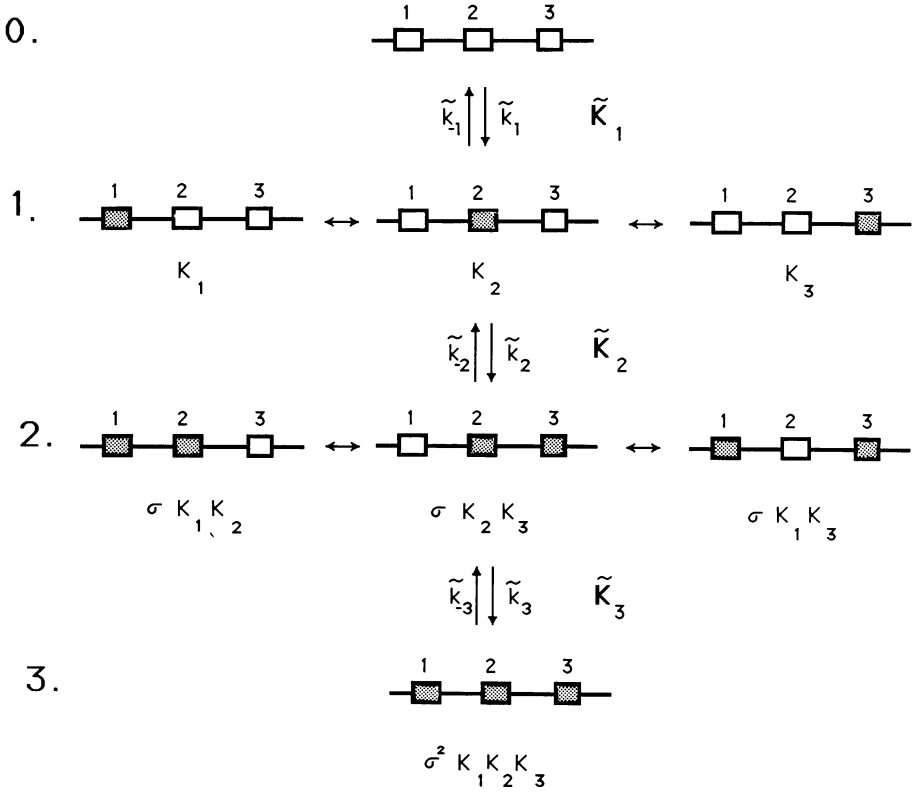

Fig. 7. Possible configurations of fragment molecules on a DNA fragment with three binding sites. The diagram represents the unoccupied (open rectangles) and protein occupied (hatched rectangles) binding sites and all possible transitions between the configurations. Transitions in the vertical direction represent position changes of proteins. Lower case ks represent kinetic rate constants, and σ cooperativity indexes.

Discussion

The principal result reported here is the demonstration that specific binding of IHF bends the DNA molecule. Previous studies of the λ integration reaction by several laboratories have shown that the nucleoprotein complex at attP is compact and somewhat similar to a nucleosome. This 'intasome' (Better *et al.*, 1982; Pollock and Nash, 1983; Bushman *et al.*, 1985; Echols, 1986) involves five binding sites for the Int protein, three sites for IHF and 230 bp of DNA. Its compact nature, as well as topological evidence, imply that the DNA is wrapped around the proteins. We have shown here that occupation of a single IHF binding site can induce a bend in the DNA. Furthermore, the centre of the bend extrapolates to the IHF recognition sequence, implying that the bend occurs at (or close to) a position where the protein is in close contact with the DNA. It is possible that the capacity of IHF to bend DNA is central to its role in the formation and stabilization of the 'intasome' complex.

Our results demonstrate that the effect of IHF-induced bending on mobility is large, perhaps a reflection of the degree of bending. Although the permutation experiment was carried out only for the IS*1*-end IHF site, the effects on mobility of IHF in the multiple-site experiments indicate that bending occurs at these sites also. IHF-induced bending has also recently been observed at the H1 site of λ attP (H.A.Nash, personal communication). Among the sites which bind IHF (Craig and Nash, 1984; Leong *et al.*, 1985; Gardner and Nash, 1986; Thompson *et al.*, 1986; Gamas *et al.*, 1987b) there are known differences in affinity. It is striking that the sequence differences between the IS*1* ends in pMP20 and pMP26 are outside the region identified as the consensus (Figure 2). Sequences flanking the usual consensus can thus have a significant effect on the affinity of the site for IHF. The nature of the role played by this sequence context is unknown, and may be important to the understanding of the function of IHF. The gel retardation system used here will be a powerful tool in the study of the sequence specificity of IHF.

A closely related question is whether sequence-specified bends in the DNA affect IHF binding. Such a relationship is possible because energy is required to bend the DNA, and binding to a molecule that contains a bend similar in its direction to the induced bend would require a smaller energy change. This should be reflected in an increase in the binding constant. Likewise, an incompatible bend could reduce the binding constant. The fragment of pBR322 that contains the IHF sites shows distinctly anomalous mobility in the absence of protein (Stellwagen, 1983; our unpublished results; and see Figure 6) suggesting sequence-directed bending. Note that attP DNA fragments, which also carry several IHF sites, have anomalous mobilities (Ross *et al.*, 1982). However, the permutation experiment, in which all unbound fragments exhibit the same mobility (Figure 5b), implies that sequence-directed bending does not occur at the IS*1* site. Although an influence on binding is not ruled out, prior bending of the DNA is apparently not necessary for IHF binding.

Experiments in which the binding constant of one site on a multi-site fragment was varied led us to conclude that, under the gel electrophoresis conditions used here, a protein bound on a DNA fragment 'sees' all the sites present on that DNA molecule. Since binding of CRP is known to bend the DNA (Wu and Crothers, 1984), we would predict that a similar effect should be observable for this protein. In examining the data reported by Fried and Crothers (1983) we found a situtation in which this occurred. A fragment containing two CRP sites was examined in a wild-type and mutant version of the principal site. It was observed that the first principal retarded band (one protein bound) of the mutant has a shifted mobility, but the second band is the same, suggesting that in this DNA−protein system also the protein 'sees' both sites.

Circular permutation experiments using variant IS*1* ends show that for most fragments the mobility of the IHF−DNA complex is increased as the binding constant is reduced (P.Prentki *et al.*, unpublished). The opposite relation is observed when the binding site is positioned at the extremity of the fragment. Similar results have recently been reported with two CRP binding sites, and it was concluded that a stronger binding is accompanied by a greater bending of the DNA in the complex (Liu-Johnson *et al.*, 1986). Our observation that IHF is capable of sliding in the complex suggests another possible explanation. The complex may exist in two states: *bent* when the protein is bound at its specific site, and *undistorted* when it is bound nonspecifically (sliding). Decreasing the binding constant to the specific site would result in the protein spending a larger fraction of its time bound nonspecifically, implying a less severe effect on the gel mobility. Crosslinking experiments to try to distinguish between these hypotheses are in progress.

Among the biological effects of IHF is its requirement for the replication/maintenance of plasmid pSC101. The putative IHF site in the origin of replication of pSC101 (Gamas *et al.*, 1986) has recently been shown to bind IHF and to undergo bending (unpublished work with G.Churchward). In the presence of IHF, copies of IS*1* in the cell will be bent at both ends. The role of such a protein-induced conformation change is not known, but segregating a transposable element into a segment of DNA with bends located at both ends is likely to serve some function. In a similar fashion, the multiple IHF-induced bends in the segment of pBR322 may be responsible for its behaviour as an insertional hotspot for IS*1*.

The presence of IHF sites in many positions where a bend would be advantageous may thus reflect a major role for this protein in the cell. IHF may also affect some cell systems because of the occurrence of a binding site in a region where a bend modifies a function indirectly, perhaps analogous to the effect of supercoiling. In summary, we suppose that IHF manages the conformation of the DNA and interacts with other proteins in nucleoprotein complexes: the pleiotropic effects of IHF mutations may be a reflection of the involvement of IHF in forming such structures *in vivo*.

Materials and methods

Plasmids and strains

The plasmid pBR322 has been described (Sutcliffe, 1978; Peden, 1983). The construction of plasmids pMP20, pMP26, pMP29 and pMP12 and their use to study IS*1* transposition will be described in detail elsewhere (P.Prentki *et al.*, in preparation). Briefly, oligonucleotides carrying sequences at the ends of IS*1* were synthesized on a Systec Microsyn 1450A or on an Applied Biosystems 381A DNA synthesizer, and cloned in an inverted repeat configuration at the extremities of the Ω fragment (Prentki and Krisch, 1984), using a procedure inspired by the 'linker tailing' technique (Lathe *et al.*, 1984). The sequence of the cloned oligonucleotides was verified by dideoxy DNA sequencing (Sanger *et al.*, 1977), as described for double-stranded DNA templates (Zagursky *et al.*, 1985). The standard strain used for plasmid constructions and DNA preparations was DH1 (*gyr*A96 *rec*A1 *end*A1 *thi*1 *hsd*R17 *sup*E44; Hanahan, 1983).

DNA preparations

Plasmid DNA was prepared according to the method of Ish-Horowicz and Burke (1981) and purified by centrifugation in CsCl/EtBr gradients. Enzymes were purchased from New England Biolabs, Boehringer Mannheim or Bethesda Research Laboratories, and used according to the manufacturers' recommendations. End labelling at the 3′ end of DNA fragments was performed using the Klenow fragment of *E. coli* DNA polymerase I and [α-^{35}S]dATP (sp. act. 1200 Ci/mmol; New England Nuclear). DNA polymerase was inactivated by a 20-min incubation at 70°C. The DNA was subjected to further digestion with the appropriate en-

zymes, extracted with phenol/chloroform (1:1 v/v) and ether, and precipitated in ethanol in the presence of ammonium sulphate.

Electrophoresis of protein–DNA complexes

The protein–DNA complexes were studied using modifications of the techniques of Fried and Crothers (1981) and Garner and Revzin (1981). Polyacrylamide gels (38:1 acrylamide to bis-acrylamide ratio) were prepared in a Tris (89 mM)–borate (89 mM)–EDTA (0.5 mM) buffer (pH 8.3). The gels were pre-run for several hours at running voltage (~ 10 V/cm), to stabilize conductivity. All electrophoresis was performed at 4°C following our observation that weaker IHF complexes are best detected in the cold. About 25 ng of end-labelled plasmid DNA was incubated with varying amounts of purified IHF protein (a generous gift of H.Nash), in 25 μl binding buffer (50 mM Tris–HCl, pH 7.4, 70 mM KCl, 1 mM EDTA, 1 mM β-mercaptoethanol, 7 mM $MgCl_2$, 3 mM $CaCl_2$, 10% glycerol, 200 μg/ml BSA (BRL) for 30 min at 25°C. Five μl of loading buffer (glycerol, 50%; BSA, 1 mg/ml; xylene cyanol, 0.05% w/v) was added, and the mixture loaded under low voltage. The purified IHF was stored as described by Nash and Robertson (1981). Care was taken to mix the protein before dispensing for the binding reactions to avoid aggregation of IHF.

To represent the degree of retardation in the migration of DNA and complexes we have calculated, for each band, the ratio of the actual mol. wt of the fragment involved to the apparent mol. wt. This latter value was determined from the calibration curve which related mol. wt to migration distance. This ratio, the relative mobility which we designate m_r, is then a measure of the anomalous migration of the fragment (with and without bound protein). Diekmann and Wang (1985) have used the same ratio, which they call the k factor. The quantity called R_f is often used for this purpose but is distinctly different (e.g. see Stellwagen, 1983).

Calculation of binding constants

The binding constants for IHF are sufficiently low that binding experiments can be done at protein–DNA ratios much greater than one. In this regime the binding constants can be estimated rather simply. For a single binding site on the DNA we need only know the concentration of the protein and the fractions of bound and unbound DNA. These fractions were obtained from densitometric analysis of the autoradiograms.

The fraction of the DNA bound at a given protein concentration, P, can be calculated directly from the definition of the equilibrium constant for the bimolecular binding reaction, K_B. It is given by

$$f_{bound} = \frac{K_B P}{1 + K_B P}$$

The concentration of free protein, P, is simply the total concentration of protein put into the reaction, because it exceeds the DNA concentration by more than a factor of 10 in most cases. Thus, we can estimate the binding constant by calculating $K_B = f_{bound}/[P\,(1-f_{bound})]$.

The concentration of protein in this expression is the concentration of active IHF capable of binding DNA. The purified IHF has been used as if the protein was 100% active for binding. This is a potential source of error in our measurements of binding constants. We estimate that our absolute values are accurate to within a factor of 2. Relative values are more accurate.

The analysis of binding to fragments containing multiple binding sites is more complex (Figure 7). In calculating the equilibrium distributions of proteins on DNA we must take into account the binding and dissociation of protein molecules from DNA fragments, as well as the binding and dissociation from one binding site and rebinding at another on the same fragment.

We assume here that the dissociation of protein molecules from the DNA is effectively prevented within the gel. This may be accomplished by the (albeit poorly understood) 'caging' effect, but must be helped substantially by the low ionic strength of the gel buffer. The overall effect is one of freezing the number distribution of protein molecules per DNA molecule when the complexes are run into the gel. Thus transitions between vertical levels shown in Figure 7 are not permitted. On the other hand, the 'sliding' or transfer of proteins from one site to another should continue.

The distribution of complexes at equilibrium in solution can be calculated from the partition function [following the analysis of Crothers (1968) and Fried and Crothers (1981)]. Consider a fragment carrying three sites to which the protein can bind, as shown in Figure 7. The equilibrium constants defining the number distribution of proteins on DNA (vertical transitions in Figure 7), in terms of the binding constants for the specific sites, are

$$\tilde{K} = K_1 + K_2 + K_3$$
$$\tilde{K}_2 = (K_1K_2 + K_1K_3 + K_2K_3)/(K_1 + K_2 + K_3)$$
$$\tilde{K}_3 = K_1K_2K_3/(K_1K_2 + K_1K_3 + K_2K_3)$$

A simple representation of the effect of cooperativity is provided by the σ in the figure. A more detailed analysis of the multi-site equilibrium will be provided elsewhere.

Acknowledgements

We are grateful to Minh-Ha Pham and Didier Sanchez for expert technical assistance, to Howard Nash and Carol Robertson for the generous gift of purified IHF protein and to Ed Nolan for constructing plasmid pEN50. We thank Howard Nash and Pascal Gamas for useful discussions. This work was supported by a CNRS/NSF grant to M.C. and D.C., and by NIH grant AI1936 to D.G.

References

Anderson,J., Ptashne,M. and Harrison,S. (1985) *Nature*, **316**, 596–601.
Berg,O.G. and von Hipple,P.H. (1985) *Annu. Rev. Biophys. Chem.*, **14**, 131–160.
Better,M., Lu,C., Williams,R.C. and Echols,H. (1982) *Proc. Natl. Acad. Sci. USA.*, **79**, 5837–5841.
Bossi,L. and Smith,D.M. (1984) *Cell*, **39**, 643–652.
Brown,N.L. and Smith,M. (1977) *J. Mol. Biol.*, **116**, 1–30.
Bushman,W., Thompson,J.F., Vargas,L. and Landy,A. (1985) *Science*, **230**, 906–911.
Craig,N. and Nash,H. (1984) *Cell*, **39**, 707–716.
Crothers,D.M. (1968) *Biopolymers*, **6**, 575–584.
Diekmann,S. and Wang,J.C. (1985) *J. Mol. Biol.*, **186**, 1–11.
Echols,H. (1986) *Science*, **233**, 1050–1056.
Flamm,E.L. and Weisberg,R.A. (1985) *J. Mol. Biol.*, **185**, 117–128.
Frederick,C.A., Grable,J., Melia,M., Samudzi,C., Jen-Jacobson,L., Wang,B.-C., Greene,P., Boyer,H. and Rosenburg,J.M. (1984) *Nature*, **309**, 327–331.
Fried,M. and Crothers,D. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
Fried,M. and Crothers,D. (1983) *Nucleic Acids Res.*, **11**, 141–158.
Fried,M. and Crothers,D. (1984) *J. Mol. Biol.*, **172**, 241–262.
Friedman,D.I., Olson,E.R., Carver,D. and Gellert,M. (1984) *J. Bacteriol.*, **157**, 484–489.
Gamas,P., Burger,A.C., Linder,P., Churchward,G., Caro,L., Galas,D. and Chandler,M. (1986) *Mol. Gen. Genet.*, **204**, 85–89.
Gamas,P., Caro,L., Calas,D. and Chandler,M. (1987a) *Mol. Gen. Genet.*, **207**, 302–305.
Gamas,P., Chandler,M., Prentki,P. and Galas,D. (1987b) *J. Mol. Biol.*, **195**, 261–273.
Gardner,J.F. and Nash,H.A. (1986) *J. Mol. Biol.*, **191**, 181–189.
Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047–3060.
Goosen,N., van Heuvel,M., Modenaar,G. and van de Putte,P. (1984) *Gene*, **32**, 419–426.
Griffith,J., Hochschild,A. and Ptashne,M. (1986) *Nature*, **322**, 750–752.
Gronnenborn,A.M., Nermut,M.V., Eason,P. and Clore,G.M. (1984) *J. Mol. Biol.*, **179**, 751–757.
Hanahan,D. (1983) *J. Mol. Biol.*, **166**, 557–580.
Hatfull,G.F. and Grindley,N. (1986) *Proc. Natl. Acad. Sci. USA.*, **83**, 5429–5433.
Hoyt,A., Knight,D., Das,A., Miller,H. and Echols,H. (1982) *Cell*, **31**, 565–573.
Ish-Horowicz,D. and Burke,J. (1981) *Nucleic Acids Res.*, **9**, 2989–2999.
Kikuchi,A. and Nash,H.A. (1978) *J. Biol. Chem.*, **253**, 7149–7157.
Koo,H.-S., Wu,H. and Crothers,D.M. (1986) *Nature*, **320**, 501–506.
Lathe,R., Kieny,M.P., Skory,S. and Lecocq,J.P. (1984) *DNA*, **3**, 173–182.
Leon,J., Nunes-Duby,S., Lessere,C., Youderian,P., Susskind,M. and Landy,A. (1985) *J. Biol. Chem.*, **260**, 4468–4477.
Liu-Johnson,H.-N., Gartenberg,M.R. and Crothers,D.M. (1986) *Cell*, **47**, 995–1005.
Mahajna,J., Oppenheim,A.B., Rattray,A. and Gottesman,M. (1986) *J. Bacteriol.*, **165**, 167–174.
Marini,J.C., Levene,S.D., Crothers,D.M. and Englund,P.T. (1982) *Proc. Natl. Acad. Sci. USA.*, **79**, 7664–7668.
Miller,H.I. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 691–698.
Miller,H.I. and Friedman,D. (1977) In Bukhari,A., Shapiro,J. and Adhya,S. (eds), *DNA Insertion Elements, Plasmids, and Episomes.* Cold Spring Harbor Laboratory Press, New York, pp. 349–356.
Mukherjee,S., Patel,I. and Bastia,D. (1985) *Cell*, **43**, 189–197.
Nash,H. and Robertson,C. (1981) *J. Biol. Chem.*, **256**, 9246–9253.
Ohlendorf,D.S., Anderson,W.F., Fisher,R.G., Takeda,Y. and Matthews,B.W. (1982) *Nature*, **298**, 718–723.
Peden,K.W.C. (1983) *Gene*, **22**, 275–280.
Pollock,T. and Nash,H.A. (1983) *J. Mol. Biol.*, **170**, 1–18.
Prentki,P. and Krisch,H.M. (1984) *Gene*, **29**, 303–313.
Prentki,P., Gamas,P., Galas,D.J. and Chandler,M. (1987) In Kelly,T. and McMacken,R. (eds), *UCLA Symposia on Molecular and Cellular Biology: Mechanisms of DNA Replication and Recombination.* Allan R.Liss, New York, pp. 719–734.
Richet,E., Abcarian,P. and Nash,H.A. (1986) *Cell*, **46**, 1011–1021.
Ross,W., Shulman,M. and Landy,A. (1982) *J. Mol. Biol.*, **156**, 505–529.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA.*,

74, 5463–5467.
Stellwagen,N.C. (1983) *Biochemistry*, **22**, 6186–6193.
Sutcliffe,J.G. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 77–90.
Thompson,J.F., Waechter-Brulla,D., Gumport,R.I., Gardner,J.F., Moitoso de Vargas,L. and Landy,A. (1986) *J. Bacteriol.*, **168**, 1343–1351.
Weisberg,R. and Landy,A. (1983) In Hendrix,R. *et al.* (eds), *Lambda II.* Cold Spring Harbor Laboratory Press, New York, pp. 211–250.
Williams,J.G.K., Wulff,D.L. and Nash,H.A. (1977) In Bukhari,A., Shapiro,J. and Adhya,S. (eds), *DNA Insertion Elements, Plasmids, and Episomes.* Cold Spring Harbor Laboratory Press, New York, pp. 357–361.
Winter,R.B., Berg,O.G. and von Hippel,P.H. (1981) *Biochemistry*, **20**, 6961–6968.
Wu,H.-M. and Crothers,D.M. (1984) *Nature*, **308**, 509–513.
Zagursky,R.J., Baumeister,K., Lomax,N. and Berman,M. (1985) *Gene Anal. Techn.*, **2**, 89–94.
Zahn,K. and Blattner,F.R. (1985) *EMBO J.*, **4**, 3605–3616.
Zerbib,D., Gamas,P., Chandler,M., Prentki,P., Bass,S. and Galas,D. (1985) *J. Mol. Biol.*, **185**, 517–524.

Received on March 20, 1987; revised on May 7, 1987