The EMBO Journal vol.12 no.13 pp.5019-5027, 1993

Specific interaction of IHF with RIBs, a class of bacterial repetitive DNA elements located at the 3′ end of transcription units*

Frédéric Boccard and Pierre Prentki

Department of Molecular Biology, University of Geneva, 1211 Geneva 4, Switzerland

Communicated by M.Hofnung

*In memory of Pierre Prentki (deceased July 24, 1993)

The prokaryotic REP (repetitive extragenic palindromes) or PU (palindromic units) sequences are often associated with other repetitive elements, forming arrangements which have been called 'BIMEs' (bacterial interspersed mosaic elements). It is estimated that the *Escherichia coli* chromosome carries ~300-500 BIMEs, whose biological role is at present unknown. We have identified a family of BIMEs consisting of two converging REP sequences flanking a 35 bp conserved segment which carries a static DNA bend and a binding site for IHF, the integration host factor of *E.coli*. We estimate that the *E.coli* genome harbors ~100 copies of this module, which we name 'RIB' (reiterative *ihf* BIME). We have analyzed by gel retardation and by footprinting the *in vitro* interaction of IHF with individual RIBs, and shown that the protein binds strongly and specifically to their center. We have also demonstrated binding of IHF to the chromosomal population of RIBs, using a new approach which combines two-dimensional bandshift and Southern blotting. RIB elements are at the end of transcription units, and thus define a new class of *ihf* sites. Possible implications for genome structure and DNA topology are discussed.

Key words: bacterial interspersed mosaic elements (BIMEs)/ DNA gyrase/genome structure/REP sequences/supercoiling

Introduction

The integration host factor (IHF) is a small, heterodimeric 'histone-like' protein which was discovered through its role in the site-specific integration of bacteriophage λ DNA into the *E.coli* chromosome (Miller and Friedman, 1977; Williams *et al.*, 1977; Kikuchi and Nash, 1978). The DNA-binding properties of IHF are remarkable: it binds to specific sites in DNA, which approximate the consensus YAANNNNTTGATW (Craig and Nash, 1984; Gamas *et al.*, 1987) in a favorable context (Goodrich *et al.*, 1990); it makes most of its sequence-specific contacts in the minor groove of DNA (Yang and Nash, 1989); it bends the DNA sharply upon binding (Prentki *et al.*, 1987a; Stenzel *et al.*, 1987), with an estimated angle of ~140° (Thompson and Landy, 1988; Kosturko *et al.*, 1989).

Most of the IHF-binding sites (*ihf*) originally identified are present on mobile genetic elements such as phage, plasmids and transposons (for a review, see Friedman, 1988), where they usually play a role in the assembly of specialized nucleoprotein structures (Echols, 1986) involved in site-specific recombination, replication, transposition or transcription. As shown by the elegant work of Goodman and Nash (1989), the main role of IHF is an architectural one: in *att*P, swapping of *ihf* sites by DNA bends can be effective in supporting the λ integration reaction.

Much less is known about the number and the role of *ihf* sites in the *E.coli* genome. Using a two-dimensional bandshift approach, we showed that the *E.coli* chromosome carries a minimum of 70 strong *ihf* sites (Boffini and Prentki, 1991). A number of *ihf* sites have recently been characterized in the *E.coli* genome, usually in promoter regions, where binding of the protein modulates the expression of several genes at the transcription initiation level (for a review, see Freundlich *et al.*, 1992).

A computer analysis of the available *E.coli* sequences has led us to detect other potential *ihf* sites at the 3′ end of a significant number of genes. Upon further inspection, we noticed that many of these sites were flanked by sequences called 'REP' (repetitive extragenic palindromes) or 'PU' (palindromic units) (Higgins *et al.*, 1982; Gilson *et al.*, 1984). These are small (~30-40 bp), highly conserved imperfect inverted repeats, found downstream of many *E.coli* genes. They have been reported to bind DNA gyrase with the help of HU (Yang and Ames, 1988, 1990), DNA polymerase I and another uncharacterized protein (Gilson *et al.*, 1990), and to be involved in mRNA stabilization (Newbury *et al.*, 1987; for a review, see Lupski and Weinstock, 1992). REP sequences are often arranged in clusters of several alternating copies, which were called 'BIMEs' (bacterial interspersed mosaic elements) by Gilson *et al.* (1991a) or 'REP elements' by Dimri *et al.* (1992). Gilson *et al.* (1991a) have shown that BIMEs contain additional repeated sequence motifs. Although present in 300-500 copies in the *E.coli* genome, their function is largely unknown. BIMEs have also been found in other bacteria related to *E.coli* and can be classified in three species-specific groups (*E.coli*, *Salmonella* and *Klebsiella*) (Gilson *et al.*, 1991b; Bachellier *et al.*, 1993).

In this communication, we identify a class of BIMEs, which we call 'reiterative *ihf* BIMEs' (RIBs), consisting of two converging REP sequences flanking a conserved region, ~35 bp in length, which carries an *ihf* consensus site. We show, by gel retardation and by footprinting, that IHF binds strongly and specifically to the center of this element. We propose that one of the functions of RIBs could be to modulate the binding of DNA gyrase to certain inter-operonic regions, thereby insulating domains from changes in supercoiling state generated during processes involving DNA.

Results

Detection and structure of RIBs

BIMEs (Gilson *et al.*, 1991a) consist of one to several alternating copies of REP sequences linked to other repetitive

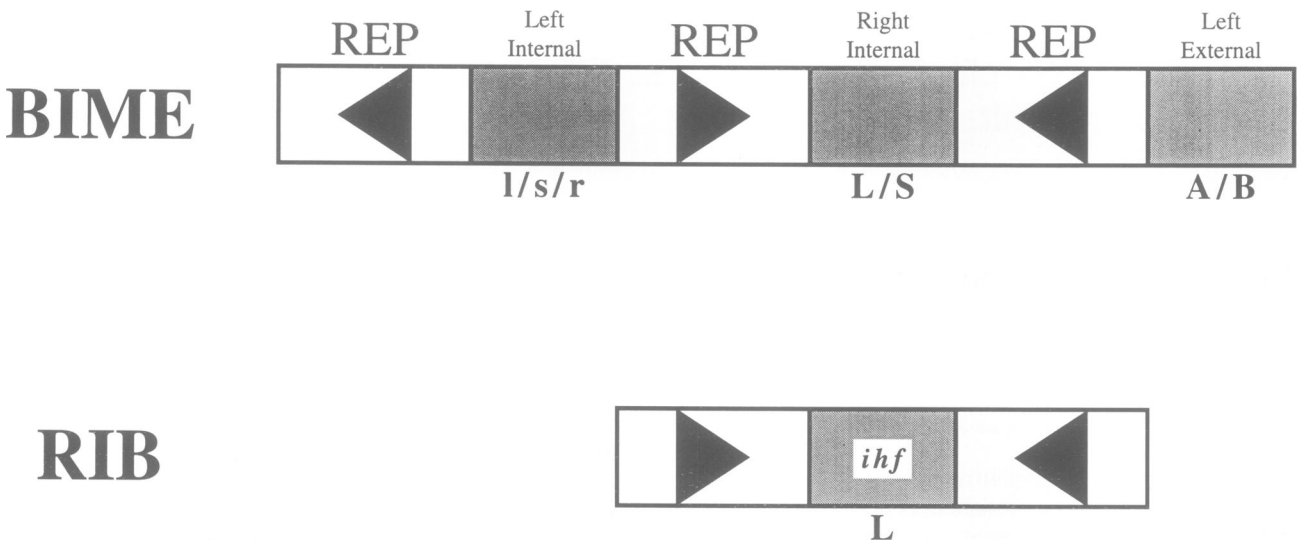

Fig. 1. Structure of a typical BIME (top). Most of the symbols and nomenclature of Gilson *et al.* (1991a) have been retained. REP (PU) sequences, represented by white rectangles with an internal triangle, alternate in orientation. They are linked to other moderately conserved motifs, which can be located between two divergent (l, s or r) or two convergent (L or S) REPs, or attached to the left end of a REP, at the border of the element (A or B). 'Left' and 'right' refer to the end of REP which abuts the interval. What we define as 'RIB' (bottom) consists of two convergent REPs flanking an L motif with a functional *ihf* site. RIBs are not necessarily flanked by the motifs shown in the top diagram.

elements (Figure 1). These other motifs, which can be either internal (S, L, s, l and r), or external (A and B), show various degrees of conservation. The most homogeneous, both in length and in sequence, is the L motif. Our computer analysis (see Materials and methods) revealed the presence of a good match to the *ihf* consensus within L elements.

A compilation of all known L elements with their flanking REP sequences is presented in Table I. This arrangement has several salient features, which are listed below.

The ihf site. The central L segments (31–36 bp, with >90% of them at 35 ± 1 bp) were aligned according to the *ihf* consensus. Out of the 41 sequences listed, three have a perfect match and 32 differ by one base only, at positions in which variations can be tolerated (Goodrich *et al.*, 1990). These sequences are listed at the top of the table. Only six sequences, listed at the bottom, have a poorer match and would not be expected to bind IHF efficiently. Overall, this alignment improves the apparent homogeneity of the population and stresses that the bases belonging to the *ihf* consensus show stronger conservation than average (see L consensus in Table I).

The flanking REP sequences. The REP sequences listed to the left vary in size between 28 and 37 bp, with two-thirds of them at 29 bp. The sequence conservation among these REPs is well above that in the average population of REPs. This homogeneity is even more striking on the right hand side REPs. Their sizes are between 32 and 34 bp, with >80% at 34 bp, and there is very little variation in their sequence. In other words, it appears that the central L segment carrying the *ihf* site tends to be associated with particular sub-classes of REPs, previously called Z1 and Y by Gilson *et al.* (1991b) (left and right hand REP sequences respectively in Table I). Thus, the BIMEs listed at the top of Table I form a well conserved structure, which we call 'RIB' (reiterative *ihf* BIME). A few atypical RIBs ('pseudo-RIBs') are listed at the bottom of the table.

The position of RIBs. In bacteria, transcription units are often composed of multiple genes. Individual REP sequences or many different classes of BIMEs can be found between genes belonging to the same operon (see Higgins *et al.*, 1988). However, in all the sequence files listed in Table I which include transcription units characterized as complete, the corresponding RIB always appears downstream of the last gene, and in either orientation with respect to the direction of transcription. Therefore, with the only caveat that some of these sequence files may later be shown to belong to longer transcription units, RIBs appear to be excluded from intra-operonic positions, a feature which may give an important clue to their possible function (see Discussion). It is interesting to note that at least half of the RIBs listed in Table I are at the end of transcription units encoding a membrane-associated protein. Finally, since ~40% of the *E.coli* sequences are represented in the current database, it can be estimated by extrapolation that the total number of RIB elements in the *E.coli* genome could amount to ~100 copies.

Binding of IHF to RIBs

We used the gel retardation assay (Fried and Crothers, 1981; Garner and Revzin, 1981) to demonstrate that IHF is able to bind to RIBs. Four representative RIBs were chosen for analysis: those situated at the end of the *uvrD*, *gyrB*, *hag* and *dppA* genes. In addition, we included two control BIMEs: the *gdhA* pseudo-RIB, which has an L element with a poorer match to the *ihf* consensus (Table I), and the *sodB* BIME, which carries a completely unrelated motif (S) between the two REP sequences, and is devoid of an *ihf* site. The corresponding fragments (180–220 bp) were generated by PCR amplification from *E.coli* genomic DNA, using specific primers annealing convergently next to the respective RIB, and cloned into the Bluescript SK^- plasmid vector (see Materials and methods). The binding probes were then generated by end-labeling.

Table I. Compilation of *E.coli* RIBs from the EMBL database

	REP	L (ihf) motif	REP	3′ of gene	Map location (min)	MB protein
KATE/OP	GCCTGATGGCTTCGCTTATCAGGCCTAC	GTGTTTCCTGCAATATATTGAATTTGCACAGTTTT	GTAGGCCGGATAAGGCGTTCACGCGCATCCGGC	katE	37.90	
SNGLPD	GCCTGATGCGTACGCTTATCAGGCCTAC	GTGGTTTATGCAATATATTGAATTTGCATGGTCTT	GTAGGCCAGATAAGACGTTCACGTCGCATCCGGC	glpD[a]	75.53	+
SECDF-TSX/OP	GCCTGATCGCTACGCTTATCAGGCCTAC	GTTATTTCAGCAATATATTGAATTTACGTGCTTTT	GTTGGCCGGACAAAGCGTTTACGCCGCATCCGGC	tsx-orf6[a]	9.18	+
AVT/OP	GCCTGATGCGCTACGCTTATCAGGCCTAC	ATGATTTCTGCAATATATTGAATTTGGAAGAATTT	GTAGGCCGGATAAGGCGCTTTACGGATCTGGC	----	80.30	
MTLD	GCCTGATGCGCTACGCTTATCAGGCCTAC	AGGATGCATCACAATTTGTTGAATTTGCACGTTCTT	GTAGGCCGCATAAGGCGCTTACGCGCATCCGC	mtlA[b]	80.62	+
FPP	GCCTGATGCGCTCCGCTTATCAGGCCTAC	GAAAATTCTGCAATGTATTGAATTTGCACGATTTT	GTAGGCCGGATAAGGCGTTAACGCGCATCCGGC	dppA	79.21	+
SRM	GCCTGATGCGCTACGCTTATCAGGCCTAC	GTGGCTCATGCAATATATTGAATTTGCACGATCTT	GTAGGCCGGATAAGCGTTCACGCCGCATCCGGC	srmB	-----	
CYS	GCCTGATGCGCTACGCTTATCAGGCCTAC	AAGGTTTCTGCAATATATTGAATTAGCACGATTTT	GTAGGCCGGATAAGGCGTTTACGCCGCATCCGGC	cysM[a]	52.30	+
FDAPGK	GCCTGATGCGCTACGCTTATCAGGCCTAC	ACGTATCCTGCAATATATTGAATTTGCAGGTTTTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	fda [a]	63.40	
GYRBF	GCCTGATGCGCTACGCTTATCAGGCCTAC	GAAAATCTTGCAATGTATTGAATTTTCATGATTTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	gyrB[a]	83.43	+
HAG	GCCTGATGCGCTGCGCTTATCAGGCCTAC	AAGTTGAATTGCAATTTATTGAATTTGCACATTTTT	GTAGGCCGGATAAGGCGTTTACGCCGCATCCGGC	hag [a]	42.60	+
PFKA	GCCTGATGCGCTTCGCTTATCAGGCCTAC	ATGAATTCTGCAATTTATTGAATTTGCAAACTTTT	GTAGGCCGGATAAGGCGTTCGCGCCGCATCCGGC	pfkA[a]	88.26	
UGP	GCCTGATGCGCTGCGCTTATCAGGCCTAC	TGGGTGAGTGGCAATATGTTGAATTTGCACGATCTT	GTAGGCCTGATAAGGCGTTTACGCCGCATCCGGC	livP[a]	75.82	+
NAGACD	GCCTGATGCGCTACGCTTATCAGGCCTAC	GCAGCTCCTGCAATATATTGAATTTTCATGCTTTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	nagB[a]	15.16	+
METJA	GCCTGATGCGCTACGCTTATCAGGCCTAC	GTCATATTGCAATTTATTGAATTTGCACGAACTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	metJ[a]	89.00	
EXBBD A-B	GCCTGATGCGCTACGCTCATCAGGCCTAC	AAAATCTATTGCAACATGTTGAATCTTCGTGCGTTT	GTAGGCCGGATAAGGCGTTCACGACGCATCCGGC	exbD[a]	65.00	+
UVRB2	GCCTGATGCGCTACGCTTATCAGGCCTAC	ATTTTCTCCGCAATATATTGAATTTGCGCGGTTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	uvrB	17.49	
RHSDRS	GCCTGATGCGCTACGCTTATCAGGCCTAC	GAGGATCGTGCAATATATTGATTTTGCAAGATTGT	GTAGGCCGGATAAGGCGTTTACGCCGCATCTGGC	----	11.65	
DACB/OP	CCCTGATGCGCTACGCTTATCAGGCCTAC	AAGAGAATTGCAATATGTTGAATTTGCACGATTTT	GTAGGCCGGATAAGGCGTTTACGCCGCATCCGGC	dacB	69.12	
GLDE/OP	GCCTGATGCGCTACGCTTATCAGGCCTAC	GCTAGCTGTGCAATCCATTGATTTTGCACAATTTT	GTAGGCTGGATAAGGCGTTCACGCCGCATCCGGC	gcd	03.10	+
GLYA/OP	GCCTGATGCGCTACGCTTATCAGGCCTAC	ACGGAGATTGCAATATATTGAATTTGCAAAGTTTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	glyA[a]	54.90	
ENTD/OP	GCCTGATGCGCTACGCTTATCAGGCCTAC	ATGGTCCTGCAATATATTGAATTGGCAAGATTTTT	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGC	fepA[a,c]	13.41	+
MSYB/OP	GCCTGATGCGCTACGCTTATCAGGCCTAC	ATCGTTCCTGCAATTTATTGATTTTGCAAGACTTT	GTGGGTCGGATAAGGCGTTCACGCCGCATCCGGC	msyB[a]	23.10	+
PMSR/OP	GCCTGATGCGCTATGCCAATCAGGCCTAC	GCAGTCCTTGCAATATATTGAATTTGACGGATTTT	GTAGGCCTGGTAAGGCGTTTACGCCGCATCTGGC	pmsR	-----	
RNPBW/OP	GCCTGATGCGCTACGCTTATCAGGCCTAC	GCCATCTCTGCAATATATTGAATTTGCGTGCTTTT	GTAGGCAGGATAAGGCGTTCACGCCGCATCCGGC	orfX	68.19	
PLSB/OP	GCCTGATGCGCGGCGCTTCTCAGGCCTAC	ACAACACATCGCAATTTATTGAATTTGCAGATTATG	GAAGGCCGGATAAGGCGTTTTCGCCGCATCCGGC	plsB[a]	91.86	+
CYD/OP	ACCTGAATGCGCTACGCTTGCCAGTGTCTAC	ATCATCTCTGCAATATATTGAATTTGCCTGCTTTT	GTAGACCGAATAAAGCGTTCATGCCGTATCCGGC	----	16.60	
UVRD/OP	GCCGGATGCGGCGTAAACGCCCTATCCGGCCTAC	ATGACGTTGCAATTTATTGAATCTGAATGATTTTT	TAGGCCAAATAAGGTGCGCAGCACCGCATCCGGC	uvrD	85.90	
NTRLA/OP	ACACCGTCTGACGTGCTACGCCTGTCAGGCCTAT	TCGACTCCTGCAATGTATTGAATTTGCATAGTTTT	ATAGGTCGAATAAGGCGTTCACGCGCATCCGGC	ntrlA	37.95	
RCSBC	GCCAGATCGACGCTGACGCGTCTTATCTGGCCTAC	TTTAATGTAATTGTTGAAATAATGGGAATC	GTAGGCCGGATAAGGCGTTTACGCCGCATCCGGC	rcsB rcsC[a]	48.10	+
ASLAB/OP	GCCGGATGCGATGCTGACGCATCTTATCCAGCCTAC	AGAACGCTGCAATTTTAATGAATTTGCACGATCAT	GTAGGCCGGATAAGGCGTTTACGCCGCATCCGGC	aslA aslB[a]	85.92	
MUTT	GCCTGATGCGACGCTGGCGCGTCTTATCAGGCCTAA	GGGATTTCTAACTCATTGATAAATTTGTTTTT	GTAGGTCGGATAAGGCGTTCACGCCGCATCCGAC	mutT[a]	02.52	+
THR	GCCTGATGCGACGCTGGCGCGTCTTATCAGGCCTAC	GTTAATTCTGCAATATATTGAATCTGCATGCTTTT	GTAGGCAGGATAAGGCGTTCACGCCGCATCCGGC	thrC[a]	00.05	
MGLABC	GCCGGATGCAACGCTGGCGCGTTTTATCCGGCCTAC	AATAAGCTCAACATACTGATTTGTATGGTTTT	GTAGGCCGGATAAGGCGTTCAAGCCGCATCCGGC	mglC[a]	46.32	+
GACAR	GCCTGATGCGACGCTTGCCGCGTCTTATCAGGCCTAC	GCCAGACAGCGCAATAGCCTGATTTAGCGTGATTTT	GTAGGTCGGATAAGGCGTTTATGCCGCATCCGAC	gltP	-----	+
consensus L		RRHDDDYHYTGCAATWTATTGAATTTGCRBGHTTTTT				
GLNHPQ	GCCTGATCGTACGCTTATCAGGCCTAC	AGGATATCTGGCAACTTATTAAAATTGCATGAACTT	GTAGGACGGATAAGGCGTTCACGCGCATCCGGC	glnQ	18.00	+
GLNS	GCCTGATGCGCTACGCTTATCAGGCCTAC	ATGATCTCTGCAACACACCGAGCCCGCGTGCTTTT	GCAGGCCGGATAAGGCGTTCGCGCCGCATCCGGC	glnS[a]	15.21	
LIVHMG	GCCGGATGCGGCGTAAACGCCTTATCCGGCCTAC	AAGATCCAAAGAAATCAGTAAATTGCAACACACATT	GTAGGCCTGATAAGCGTAGCGCATCAGGC	livJ[a]	75.91	+
TRPR	GCCTGATGCGACGCTGCCGCGTCTTATCATGCCTAC	CAAACATATTGAAATTACGGGTATTT	GTAGGACGGATAAGGCCTTCACGCCGCATCCGGC	trpR[a]	99.68	

All BIMEs with converging REP sequences separated by 20−40 bp are listed. The name of the corresponding EMBL file is given in the first column, ending with '/OP' when the complementary strand is shown. The internal segments have been aligned to the *ihf* consensus, YAAnnnnTTGATW. Bases matching the consensus are underlined. The vast majority differ by no more than 1 bp from the *ihf* consensus (top of the table), and are called 'RIBs' (see text). The consensus for the central 'L' segment, with a 75% chance of occurrence at each position, is also given. Sequences with a poorer match to the *ihf* consensus or with atypical L motifs are listed at the bottom ('pseudo-RIBs' in the text). The table also indicates the gene at the 3′ end of which the RIB element is found, its chromosomal map location, and the presence of a protein associated with the membrane (indicated by + in the 'MB protein' column) in the same transcription unit. A few RIB or pseudo-RIB elements have been omitted from this list: those situated 3′ of *aspC*, *melB*, *priC*, because the sequence of one of their flanking REP is incompletely known; *metBL*, which has a 50 bp center with an L-like sequence, and a poor *ihf* site; *polA*, which carries a truncated L (29 bp), and is followed by a completely atypical REP. Our original sequence data (HAG) have been used in place of the sequence of the HAG48 file.

[a]Indicates RIBs for which there is evidence that they are located at the end of transcription units. Conclusive transcription data are not available for the others.

[b]A revision of the original sequence of the intercistronic region between *mtlA* and *mtlD* (file ECMTLD1) revealed the presence of a ρ-independent transcription terminator downstream of RIB *mtl*A.

[c]In the case of RIB *Ent*D, a minor transcript for genes on both sides of that RIB might be present (Armstrong *et al.*, 1989).

The results are presented in Figure 2A. They show clearly that the *uvrD*, *gyrB*, *hag* and *dppA* RIBs are efficiently retarded by low concentrations of IHF (5 nM, lanes b) and that they form stable and specific complexes. The *gdhA* module binds poorly to the protein, and the *sodB* BIME shows only weak and non-specific binding at the highest protein concentration. To strengthen our confidence in these results, we have also analyzed simultaneously the binding of IHF to two elements which had responded either positively (*dppA*) or negatively (*sodB*) in Figure 2A. In this mixture, only the largest of the two (*dppA*) is retarded by IHF (Figure 2B), as expected. The strength of IHF binding to RIBs, despite their imperfect match to the consensus, suggests a contribution from the context. We note, at the right end of the L segment (Table I), a short run of Ts: this sequence is also present in the strongest *ihf* site of λ *att*P (H′) where it has been shown to be contacted by the protein (Yang and Nash, 1989) and to be important for binding efficiency (Lee *et al.*, 1991).

Another observation can be made from Figure 2. Because of particular cloning configurations, all the DNA fragment probes are of slightly different absolute length. But a

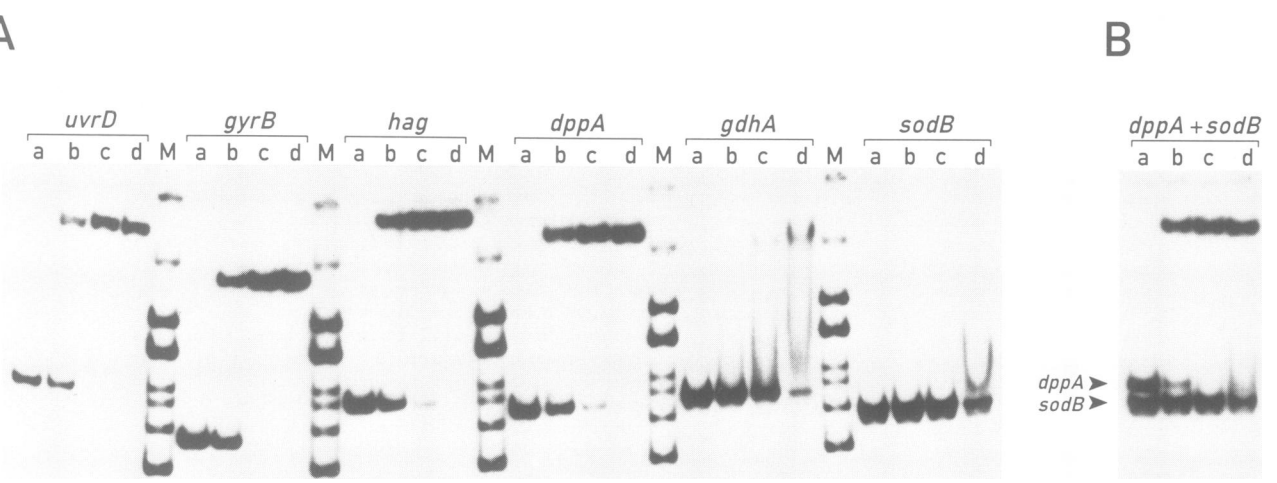

Fig. 2. Binding of IHF to RIBs. The DNA binding probes were generated by end-labeling of *Hin*dIII digests of the plasmid clones (except *gyrB*: *Bam*HI digest), followed by a second digestion with *Bam*HI (except *gyrB*: *Xho*I). The DNA (0.25 nM) was incubated with purified IHF at concentrations of 0 nM (lanes a), 5 nM (lanes b), 10 nM (lanes c) or 20 nM (lanes d). Lanes M contain a size standard (pBR::IS*1* # 11, digested with *Hin*fI and end-labeled). In panel **A**, individual DNA probes were used; in panel **B**, equimolar amounts of the *dppA* and *sodB* probes were mixed before incubating with IHF.

comparison of the known length of each naked DNA fragment probe with its apparent size, measured using size standards present on the gel, shows that the four fragments that bind IHF well also have a slightly aberrant (slower) electrophoretic mobility. The R_L value (relative mobility), measured on several gels, is 0.92 (*uvrD*, *gyrB*, *hag*) or 0.93 (*dppA*), whereas that for the non-binding *sodB* is completely normal (1.00). Interestingly, the poorly binding element (*gdhA*), displays an almost normal value (0.98). As aberrant gel mobilities are often associated with structure distortions, such as DNA bending (for a review, see Lane *et al.*, 1992), we determined whether this was the case with RIBs.

Circular permutation

A simple way to demonstrate the presence of a static bend in a DNA sequence is the circular permutation method devised by Wu and Crothers (1984). It is based on the observation that DNA molecules of identical length have different electrophoretic mobilities depending on the position of the bend relative to the ends. To prove that the anomalous mobility of RIB elements was due to the presence of a bend in their center, we cloned the *uvrD* RIB and the *gdh*A pseudo-RIB in the 'bending vector' pCY4 (Prentki *et al.*, 1987b), a pBR322 derivative with a duplication of the 375 bp *Eco*RI–*Bam*HI interval separated by a polylinker cloning site.

The electrophoretic pattern of the *uvrD* clone (data not shown) is diagnostic of DNA bending: fragments in which the cloned RIB element is near an end had a higher mobility than fragments in which it is near the middle. Such a periodic behavior was barely detectable with the *gdh*A pseudo-RIB, indicating the presence of a very minor bend. The presence of a stronger bend in *uvrD* RIB was also revealed on different gels by the slower mobility of fragments containing *uvrD* RIB. By plotting the R_L values from different gels (data not shown), the bend center was mapped to within a few base pairs of the *ihf* site present in the L element. The generality of the presence of a static bend at the center of RIB elements is apparent from other experiments, presented below.

DNase I footprinting

Footprinting experiments (Galas and Schmitz, 1978) were performed to verify that the interaction between IHF and RIB takes place at the site predicted at the center of the element. The pattern obtained with the element distal to *gyrB* is shown in Figure 3A (bottom strand). Plasmid pFBG30 was cut with *Xho*I, 3′ end-labeled and redigested with *Bam*HI.

The DNase I footprint left by IHF is of a typical size for this protein [35–40 bp (Gamas *et al.*, 1987)] and coincides almost exactly with the central L segment. There is little protection of the flanking REP sequences. Footprinting of the other strand was also performed (data not shown); the combined protection results are presented on the diagram in Figure 3B.

It should be stressed that, even in the absence of IHF, the region is poorly sensitive to DNase I, except for two phosphodiester bonds on each strand which react strongly. This seems to be a 'signature' of L elements, as it is also observed with the *uvrD* and *dppA* RIBs (data not shown), but not with the vast majority of *ihf* sites unassociated with REPs analyzed by us and by others.

Binding of IHF to a RIB element in the presence of a genome equivalent

To add more significance to the *in vitro* binding studies of IHF–RIB interactions, we decided to examine IHF binding to the *gyrB* site in the presence of a complete equivalent of the *E.coli* genome, under conditions where the IHF:DNA ratio is comparable to that found in the cell (Materials and methods). To do this, we developed an approach which combines gel retardation and Southern blotting. The experiment is shown in Figure 4.

E.coli genomic DNA was digested with *Hin*cII + *Hin*fI, incubated with 0, 50, 100 and 150 nM IHF, and electrophoresed on a non-denaturing polyacrylamide gel under conditions used for gel retardation. An ethidium bromide stain of the gel is shown in Figure 4A. The presence

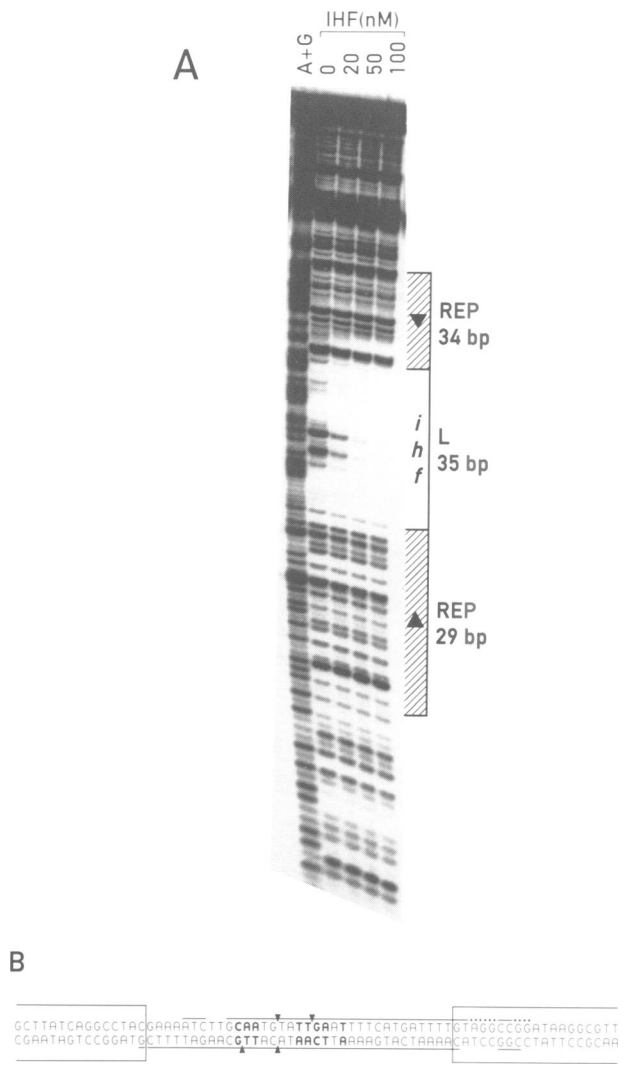

Fig. 3. IHF protection of RIB at the end of *gyrB*. (**A**) DNase I footprinting of a DNA fragment 3′ end-labeled at the *Xho*I site. The concentrations of IHF in the binding reactions are indicated at the top of each lane. The first lane is an A+G sequencing ladder. The position of the RIB is diagrammed at the right of the figure. (**B**) Combined IHF protection patterns on the L element of *gyrB*. Base pairs corresponding to those in the *ihf* consensus are in bold. The regions protected in footprinting experiments are indicated by horizontal lines. The dotted line indicates uncertainty in the extent of protection. The phosphodiester bonds which show strong reactivity in the absence of IHF are indicated by arrowheads.

of the protein has little effect on the general electrophoretic behavior of the chromosomal digest.

The same gel was then blotted onto a nylon membrane and probed with nick-translated *gyrB* RIB DNA. The results (Figure 4B) show that only one band, of the size expected for *gyrB* (441 bp), is detected in the absence of protein (lane a) in conditions of high stringency. This band is efficiently retarded by IHF (lanes b and c). The specificity of this interaction was demonstrated by stripping the labeled DNA from the filter and probing it again with an end-labeled *sodB* BIME (Figure 4C). Again, one band only lights up in the no protein lane, but as expected, this band is not affected by the presence of IHF. The same result was observed by probing with the *gdhA* pseudo-RIB (data not shown), indicating that this method revealed only specific binding of IHF to chromosomal fragments.

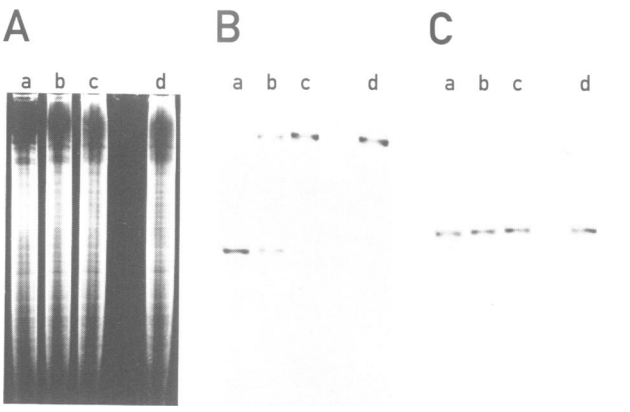

Fig. 4. Southern-blotted bandshift of *E.coli* genomic DNA. (**A**) Ethidium bromide stained polyacrylamide gel (5%). A *Hinc*II + *Hin*fI digest of *E.coli* genomic DNA was incubated with 0 (lane a), 50 (lane b) 100 (lane c) or 150 nM (lane d) IHF before non-denaturing electrophoresis. (**B**) Autoradiogram of a Southern blot of the same gel, using a nick-translated *gyrB* RIB probe, labeled with ^{32}P. (**C**) Southern blot of the same gel, after stripping of the *gyrB* probe and rehybridization with a *sodB* BIME probe.

Genome scanning for interactions between IHF and RIBs

The few members of the RIB family that we have analyzed individually all bound IHF efficiently. To demonstrate that IHF binding is a general property of RIB elements, we have taken advantage of the sequence conservation of the L motif to design a specific probe, and used a two-dimensional bandshift approach (Boffini and Prentki, 1991), which permits visualization of protein binding sites in genomic DNA.

The principle of the two-dimensional bandshift is simple: it consists of two successive gel electrophoresis steps performed perpendicularly at two different temperatures. The DNA fragments which interact with a protein are retarded in the first dimension, performed in the cold, and detached from the protein during the second dimension, performed at a high temperature. Hence, they migrate ahead of a diagonal formed by all the fragments that have not been bound by the protein, and which thus behave similarly in both dimensions.

E.coli genomic DNA was digested with *Alu*I + *Hin*fI, and incubated with 0, 100 or 150 nM IHF, loaded onto a 5% polyacrylamide gel, and subjected to the two-dimensional bandshift procedure. These two restriction enzymes were chosen because they have cleavage sites in only a few of the RIB elements listed in Table I. The results are presented in Figure 5. An ethidium bromide stain of the gel shows little general effect of IHF on the chromosomal pattern, except for the appearance of a small 'cloud' at the top of the gel, ahead of the diagonal (lanes b and c). This must correspond to a significant number of chromosomal fragments with *ihf* sites, which were strongly retarded.

A Southern blot of the same gel probed with a labeled degenerate oligonucleotide (OL-RIB) corresponding to the L segment is shown in Figure 5B and C. The degeneracy of this OL-RIB probe was limited by introducing variable nucleotides only at positions corresponding to the least conserved bases in L. This implies that OL-RIB will anneal completely to 25% of the RIBs, it will present a 1 bp mismatch in 50% of the cases, and in the remaining 25% it will have two or more differences. Approximately 30

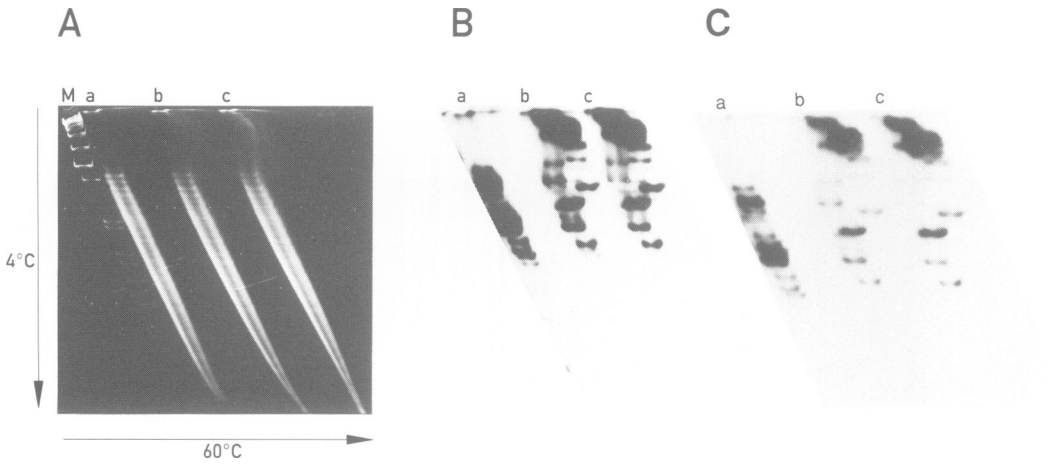

Fig. 5. Binding of IHF to a population of RIBs detected by a two-dimensional Southern bandshift. The first dimension (vertical arrow) was at 4°C, and the second at 60°C (see Materials and methods for details). (**A**) Ethidium bromide stained gel. A *Hin*fI + *Alu*I digest of *E.coli* genomic DNA was incubated with 0 (lane a), 100 (lane b) or 150 nM (lane c) IHF, and subjected to two-dimensional electrophoresis. Lane M is a size marker. (**B**) Autoradiogram of a Southern blot of the same gel, using a ^{32}P end-labeled oligonucleotide probe (OL-RIB), specific for the L-motif. (**C**) Shorter exposure of the autoradiogram shown in panel B.

signals are detected, with varying intensities (lane a). The same number was observed with a *Sau*3AI digest of the genome (data not shown).

Strikingly, in the presence of IHF (lane b, 100 nM; lane c, 150 nM) all of these bands are completely or almost completely displaced from the diagonal. By vertical projection onto the diagonal, these bands correspond to signals observed in lane a. Hence, all members of the sub-population of RIBs visualized in this experiment are specifically bound by IHF, albeit with some variability in the efficiency.

Similar experiments using as a probe OL-REP34 specific for the 34 bp REP sequence (right REP in Table I) were performed. The same signals displaced from the diagonal detected with OL-RIB were seen but the overall pattern of the diagonal was only slightly changed in presence of IHF (data not shown), indicating that the non-RIB BIMEs were not bound by IHF.

The data presented in Figure 5 also confirm that a large fraction of RIB elements carry a DNA bend. Whereas the general pattern of *E.coli* DNA gives a perfect diagonal (Figure 5A, lane a), it is apparent that, even in the absence of IHF, a large fraction of the fragments detected by the probe are found slightly ahead of the diagonal (Figure 5B, lane a). This is because the anomalous mobility conferred by DNA bending is maximal at low temperature (Diekmann, 1987) and is suppressed during electrophoresis at high temperature (Mizuno, 1987). The irregularity of the diagonal probably reflects the heterogeneity in the position of the bend in each restriction fragment.

Discussion

IHF–DNA interactions have been studied extensively with many mobile genetic elements but the role of chromosomal *ihf* sites has only recently been documented (see Freundlich *et al.*, 1992). These characterized *ihf* sites (~25) were located 5′ of genes, in promoter regions, and IHF binding modulated the transcription level. We have shown here that IHF interacts with a new class of sites ('RIBs', ~100 copies), which are located at the 3′ end of transcription units. This high number of *ihf* sites is consistent with the pleiotropic *in vivo* properties of IHF.

RIB elements have been characterized structurally as consisting of two converging REP sequences flanking a conserved segment (L motif) with an *ihf* consensus. It should be noted that the high degree of conservation of the L motif suggests a possible interaction with more than one protein (see Herman and Schneider, 1992). Despite their modular structure, several arguments indicate that RIBs are to be considered as discrete units of 100–110 bp. First, L elements are never found without being associated with REP sequences. Second, these particular REPs are strongly conserved in size and in sequence, being usually 29 and 34 bp for left and right hand sequences respectively. Third, although frequently flanked by additional repetitive elements (see Gilson *et al.*, 1991a), the immediate sequence context of RIBs is variable. RIB elements thus constitute a well defined and distinctive sub-class of BIMEs.

RIBs differ from other BIME elements in their ability to bind IHF. This was shown by classical methods, such as gel retardation (Figure 2) and footprinting (Figure 3) on representative cloned sequences, and was confirmed and extended by a new approach, which combines the resolution provided by two-dimensional bandshifts with the power of detection of Southern blotting (Figure 5), hence permitting the study of specific protein–DNA interactions involving a collection of chromosomal sites. This approach could easily be applied in a wide variety of other situations, in which it is desirable to detect, characterize and clone DNA fragments carrying protein-binding sites through their interaction with a regulatory protein.

We also found that RIB DNA has some intrinsic curvature at or near the *ihf* site. This observation has two implications: the function of this bend may simply be to put the DNA in a path which facilitates IHF binding, as has been found with some, but not all, *ihf* sites; but it also means that RIB elements may sometimes fulfil their role in the cell in the absence of IHF, albeit with a reduced efficiency. This is consistent with the fact that under most physiological situations, the presence of IHF is not essential to the host (see Freundlich *et al.*, 1992).

The map location of the known RIB elements does not seem to follow a particular bias (Table I), and it is likely that the hundred or so estimated copies of RIBs are distributed more or less evenly around the chromosome. Laundon and Griffith (1988) have shown that curved DNA segments, by making an apex on plectonemic supertwisted DNA, define 'ends' of supercoiled DNA domains. Interestingly, the number of RIBs coincides roughly with the number of independent domains of supercoiling in the nucleoid, estimated to be 50–100 (Pettijohn, 1982). REP sequences are able to interact with a chromoid-associated protein (see Gilson *et al.*, 1991b) and it is tempting to speculate that IHF-mediated DNA bending at RIBs may facilitate the formation of a higher-order DNA nucleoprotein structure involved in the organization of the nucleoid. A detailed comparison of isolated nucleoids from IHF^+ and IHF^- strains would therefore be of great interest.

Another function of RIBs is suggested by the strong bias observed in their *local* position: in contrast to the general population of BIMEs, which can be found at intra-operonic locations (i.e. between genes of the same transcription unit), RIB elements are found at the very end of transcription units (Table I). The conservation of position of RIBs at the 3′ end of transcription units has to be put in parallel with the ability of REP sequences to bind DNA gyrase, among other proteins. Such an interaction may lead RIB elements to play a key role in the structural definition of transcription units beside additional functions which may be common to other BIMEs. Liu and Wang (1987) have shown that movement of RNA polymerase along a DNA molecule can result in the formation of negative supercoils behind the transcription complex, and of positive supercoils ahead, if the ends of the DNA molecule or the transcription apparatus itself are not free to rotate.

In the model we propose (Figure 6), IHF plays an architectural role in the formation of nucleoprotein complexes involving at least RIB DNA and gyrase, reminiscent of the interaction of IHF with the Int topoisomerase and *att*P in the λ intasome. Such complexes may thus act as topological boundaries which would prevent the propagation of supercoiling changes into adjacent transcription units, and RIB elements are ideally located to bind DNA gyrase and function as swivels for the removal of transcription-generated positive supercoils (Figure 6). The size of RIB elements puts them within the range of the amount of DNA wrapped in a complex with a gyrase dimer [~115 bp as judged by electron microscopy, ~140 bp by DNase I protection (Kirchhausen *et al.*, 1985)]. This model does not exclude the possibility that other BIMEs can also remove positive supercoils, but it is conceivable that by bending RIB elements, IHF forms a loop which modulates the interaction of gyrase with the flanking REP sequences or its topoisomerase activity (Figure 6). This would also influence transcription of the operon bordered by RIBs, implying that IHF may also exert a retro-regulatory role by binding at the 3′ end of genes. In this light, it may be significant that one of the RIBs studied here occurs at the end of the *gyrB* transcription unit, encoding the B subunit of DNA gyrase itself, whose promoter has been shown to be regulated by supercoiling (Menzel and Gellert, 1983). Earlier observations of synergistic effects of mutations affecting IHF and GyrB, have already suggested an interaction between IHF and gyrase (Friedman *et al.*, 1984). Of particular importance would be

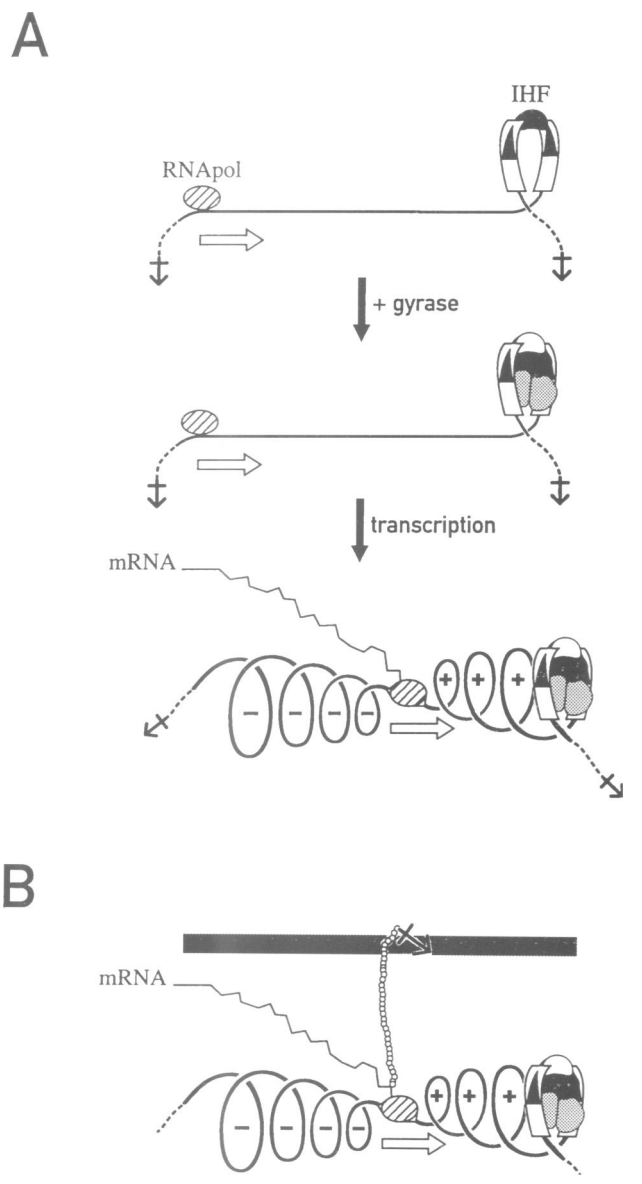

Fig. 6. A model for the role of IHF binding at RIBs. (**A**) The first line shows a transcription unit which is part of a chromosomal DNA domain. Stabilization of chromosomal loops at their bases (symbolized by anchors) can occur by protein–protein interactions or by attachment to a matrix, a well documented situation in eukaryotes in which topoisomerase II is involved (see Gasser and Laemmli, 1987). Upon binding to the RIB element, IHF (represented by a black semi-circle) induces a strong bend between the two REP sequences. The looped structure thus formed would have the right size and geometry to accommodate a gyrase tetramer (two A and two B subunits), which would interact symmetrically with the flanking REP sequences. The enzyme thus bound at the RIB would then be positioned to remove the positive supercoils generated ahead of RNA polymerase during transcription (Liu and Wang, 1987); IHF may or may not remain bound to the RIB–gyrase complex at this stage (indicated by an empty semi-circle). This model thus puts IHF in a position to modulate gyrase binding and activity. (**B**) In *E.coli*, the formation of such 'twin supercoiled domains' is favored if the nascent RNA codes for a protein carrying a transmembrane domain (anchor) which is cotranslationally inserted in the membrane, thereby further restricting the rotation of the transcriptional machinery (Pruss and Drlica, 1986; Lynch and Wang, 1993). It is noteworthy that at least half of the RIBs listed in Table I are at the end of transcription units encoding a membrane-associated protein.

to determine whether IHF binding is affected by positive supercoiling and whether IHF remains attached to its DNA site after gyrase binding.

The roles we have proposed for RIBs, in the definition of chromosomal domains and as topological boundaries, are not mutually exclusive, and may be intimately linked. It is striking to note that despite their abundance in the *E.coli* genome, neither RIB elements nor even REP sequences have been found in mobile genetic elements. The reason for this absence is at present unclear but is consistent with a topological or structural role in complex DNA genomes.

The biological significance of RIB elements, and of their ability to interact with IHF, recently received further support from the discovery of a BIME element in *Klebsiella aerogenes* (at the 3′ end of *pulA*), which fulfils the criteria for RIBs: it consists of two convergent REP-like sequences flanking an L-like motif (Bachellier *et al.*, 1993). The REP sequences and the L motif overall are divergent from those found in *E.coli*, but an *ihf* site is still present, stressing the conservation of the ability of RIBs to bind IHF.

Materials and methods

Computer search for ihf sites

A compilation of *ihf* sites led Goodrich *et al.* (1990) to propose a 48 bp matrix which includes the previously defined consensus (YAAnnnnTTGATW) and assigns a similarity score relying on the occurrences of each base at each position. This matrix was used to scan the *E.coli* sequences present in the EMBL database (release 33) for potential *ihf* sites, demanding only that the TG dinucleotide be present, since it is fully conserved in the compilation and since mutations at any of these two bases are deleterious for binding (Murtin and Prentki, unpublished results). Sequences scoring higher than 45 were retained for further analysis.

Visual inspection of these sequences showed that ~25 of them were placed in a context which displayed strong resemblances extending over a total length of ~100 bp. The sequences were then aligned around the *ihf* consensus. This revealed that the putative *ihf* sites were present on a conserved ~35 bp segment, bracketed by inverted repeats consisting of REP sequences. Scanning for potential REP or L sequences in the database was performed by using the FASTA program (Pearson and Lipman, 1988), and alignments were done with the PILEUP algorithm (Needleman and Wunsch, 1970), both from the GCG package (University of Wisconsin).

Bacterial strains and plasmids

The standard *E.coli* strain used for transformation and plasmid DNA preparations was DH5α. Genomic DNA was prepared from DF101 (C600/λ). The pBR322 derivatives pBR322::IS*1* # 11 (carrying an IS*1* insertion at coordinate 3291; Prentki *et al.*, 1987c) and pCY4 (Prentki *et al.*, 1987b) have been described previously. Plasmid pOP24 carries a 6.9 kb *Eco*RI fragment with *gyrB* on a pBR325 backbone. Bluescript SK⁻ (Stratagene) was used routinely as cloning vector. The construction of the recombinant plasmids is described below. Their names and the respective BIMEs are: pFBG1, *uvrD*; pFBG30, *gyrB;* pFBG23, *hag*; pFBG14, *dppA*; pFBG25, *gdhA*; pFBG18, *sodB*.

DNA manipulations, transformation and sequencing

For preparation of plasmid DNA, tailing of DNA with terminal deoxynucleotidyl transferase, end-labeling of DNA fragments, digestion with restriction enzymes, ligation and transformation of *E.coli*, well established procedures were used (Ausubel *et al.*, 1989; Sambrook *et al.*, 1989). Sequencing of double-stranded DNA was as described by Murphy and Kavanagh (1988). Linear fragments were extracted and purified from agarose gels with Geneclean II (Bio 101).

The following pairs of oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer and used as primers for PCR amplification (Saiki *et al.*, 1988) from lysed bacterial colonies: 5′-TGGCTGGTGG-CGGCATACGC-3′ and 5′-CCATGATTTCCGTCGAGCGG-3′ (*uvrD*); 5′-CTTCGAAAACGTCTCTATCG-3′ and 5′-CCACCGGGTATACAG-ATTG-3′ (*dppA*); 5′-CTGCTGCAGGGTTAATCGTT-3′ and 5′-TTAC-AGGTAAATTCCAGGCA-3′ (*hag*); 5′-GTGAAGGTTGCCGATGCG-3′ and 5′-TGGCCCATAAGAAATGTTCC-3′ (*gdhA*); 5′-GGAGCACTTC-TGGGCGCT-3′ and 5′-ACAGTTGCTGACAACGTGC-3′ (*sodB*). The colonies were resuspended in 100 μl of lysis buffer (1% Triton, 20 mM Tris−HCl pH 8.3, 2 mM EDTA) and incubated at 96°C for 10 min in a DNA Thermal Cycler (Perkin Elmer Cetus). 5 μl were used as template in a 50 μl reaction volume. Successive incubations at 94°C (30 s), 50°C (30 s) and 72°C (1 min) were repeated 25−30 times in an automated thermal cycler, followed by a final incubation at 72°C (10 min).

For cloning, the PCR products were tailed with dCTP, under conditions which add ~20 bases to each 3′ end (Maniatis *et al.*, 1982), inserted into similarly dGTP tailed-pSK⁻ and transformed into *E.coli* (Sambrook *et al.*, 1989). A 221 bp *Cla*I−*Hin*cII fragment containing *gyrB* RIB was cloned in pSK⁻ digested with *Sma*I and *Cla*I. Recombinants were selected on LA plates supplemented with ampicillin (100 μg/ml), X-Gal (40 μg/ml) and IPTG (20 μg/ml).

In vitro binding of IHF

The DNA binding reactions for gel retardation have been described previously (Prentki *et al.*, 1987a). Briefly, IHF was incubated for 30 min at 25°C with DNA in binding buffer (50 mM Tris−HCl pH 7.5, 70 mM KCl, 7 mM $MgCl_2$, 3 mM $CaCl_2$, 1.1 mM EDTA, 10% glycerol, 200 μg/ml BSA, 1 mM β-mercaptoethanol). The samples were loaded on polyacrylamide gels at 4°C without loading dye, under voltage (3 V/cm). Electrophoresis was performed in a cold room at 14 V/cm for ~2 h. Running buffer was 1 × TBE (Sambrook *et al.*, 1989). The gels had been prerun for several hours to stabilize the conductivity. For two-dimensional bandshift, conditions were as described before (Boffini and Prentki, 1991).

For chromosomal DNA binding reactions, chromosomal DNA was extracted as described by Ausubel *et al.* (1989). After digestion with restriction enzyme(s), DNA was phenol-treated, ethanol-precipitated and resuspended in water. IHF (100 or 150 nM) was incubated with 3−5 μg of DNA in 20 μl of binding buffer. The number of IHF protomers per cell is estimated to be at least 20 000 (M.Ditto and R.Weisberg, personal communication cited in Freundlich *et al.*, 1992). In our conditions, the ratio of IHF to DNA approximates that found in the cell: 2000 IHF molecules are present per genome equivalent.

Conditions for footprinting with purified IHF were exactly as described by Gamas *et al.* (1987), except that the last phenol−chloroform extraction of the samples was omitted.

Southern-blotted one- and two-dimensional bandshifts

After electrophoresis, one-dimensional or two-dimensional gels were stained with ethidium bromide, photographed, soaked twice for 30 min in 99% formamide at 42°C and finally for 15 min in electrotransfer buffer (12 mM Tris base, 6 mM sodium acetate, 0.3 mM EDTA pH 7.4). The DNA was then electrotransferred onto a Hybond-N membrane (Amersham) using the same buffer. Electrotransfer was carried out for 2 h at 600 mA with cooled buffer. After transfer, the membrane was allowed to dry and the DNA was crosslinked with UV. Prehybridizations, hybridizations, washing and stripping of the probe were performed as recommended by the supplier of the membrane. DNA fragments were labeled by nick translation or by random priming (Sambrook *et al.*, 1989).

Two oligonucleotide probes were synthesized: OL-RIB (5′-AAANYN-TGCAAATTCAATAWATTGCA-3′) and OL-REP34 (5′-GCCGGATG-CGGCGTRAACGCCTTATCCGGCCTAC-3′), specific for the L element and the 34 bp REP, respectively. They were labeled using [γ-^{32}P]ATP from Amersham and T4 polynucleotide kinase from NEB (Sambrook *et al.*, 1989). Southern hybridizations using OL-RIB and OL-REP34 probes were finally washed twice for 20 min at 48°C in 1 × SSPE (180 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA pH 7.7), 0.1% SDS and at 48°C in 0.1 × SSPE, 0.1% SDS, respectively.

Acknowledgements

We thank Cynthia Steinberger for computer analysis, Sundiatu Dixon-Fyle for the gift of IHF, Mick Chandler, Nicolas Roggli and Eduardo Gonzalez for help with the figures, Dany Rifat for the synthesis of oligonucleotides, and Sophie Bachellier, Andrea Boffini, Lucien Caro, Mick Chandler, Michel Clerget, Sundiatu Dixon-Fyle, Hans Geiselmann, Emmanuel Käs and Christine Murtin for discussions and for comments on the manuscript. This work was supported by grants 31-9073.87 and 31-32692.91 from the Swiss National Science Foundation.

References

Armstrong,S.K., Petis,G.S., Forrester,L.J. and McIntosh,M.A. (1989) *Mol. Microbiol.*, **3**, 757−766.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G.,

Smith,J.A. and Struhl,K. (1989) *Current Protocols in Molecular Biology*. John Wiley & Sons, New York.
Bachellier,S., Perrin,D., Hofnung,M. and Gilson,E. (1993) *Mol. Microbiol.*, **7**, 537–544.
Boffini,A. and Prentki,P. (1991) *Nucleic Acids Res.*, **19**, 1369–1374.
Craig,N.L. and Nash,H.N. (1984) *Cell*, **39**, 707–716.
Diekmann,S. (1987) *Nucleic Acids Res.*, **15**, 247–265.
Dimri,G.P., Rudd,K.E., Morgan,M.K., Bayat,H. and Ames,G.F.-L. (1992) *J. Bacteriol.*, **174**, 4583–4593.
Echols,H. (1986) *Science*, **233**, 1050–1056.
Freundlich,M., Ramani,N., Mathew,E., Sirko,A. and Tsui,P. (1992) *Mol. Microbiol.*, **6**, 2557–2563.
Fried,M.G. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6524.
Friedman,D.I. (1988) *Cell*, **55**, 545–554.
Friedman,D.I., Olson,E.J., Carver,D. and Gellert,M. (1984) *J. Bacteriol.*, **157**, 484–489.
Galas,D.J. and Schmitz,A. (1978) *Nucleic Acids Res.*, **5**, 3157–3170.
Gamas,P., Chandler,M.G., Prentki,P. and Galas,D.J. (1987) *J. Mol. Biol.*, **195**, 261–272.
Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047–3060.
Gasser,S.M. and Laemmli,U.K. (1987) *Trends Genet.*, **3**, 16–22.
Gilson,E., Clément,J.-M., Brutlag,D. and Hofnung,M. (1984) *EMBO J.*, **3**, 1417–1421.
Gilson,E., Perrin,D. and Hofnung,M. (1990) *Nucleic Acids Res.*, **18**, 3941–3952.
Gilson,E., Saurin,W., Perrin,D., Bachellier,S. and Hofnung,M. (1991a) *Nucleic Acids Res.*, **19**, 1375–1383.
Gilson,E., Saurin,W., Perrin,D., Bachellier,S. and Hofnung,M. (1991b) *Res. Microbiol.*, **142**, 217–222.
Goodman,S.D. and Nash,H.N. (1989) *Nature*, **341**, 251–254.
Goodrich,J.A., Schwartz,M.L. and McClure,W.R. (1990) *Nucleic Acids Res.*, **18**, 4993–5000.
Herman,N.D. and Schneider,T.D. (1992) *J. Bacteriol.*, **174**, 3558–3560.
Higgins,C.F., Ames,G.F.-L., Barnes,M., Clément,J.-M. and Hofnung,M. (1982) *Nature*, **298**, 760–762.
Higgins,C.F., McLaren,R.S. and Newbury,S.F. (1988) *Gene*, **72**, 3–14.
Kikuchi,A. and Nash,H.A. (1978) *J. Biol. Chem.*, **253**, 7149–7157.
Kirchhausen,T., Wang,J.C. and Harrison,S.C. (1985) *Cell*, **41**, 933–943.
Kosturko,L.D., Daub,E. and Murialdo,H. (1989) *Nucleic Acids Res.*, **17**, 317–334.
Lane,D., Prentki,P. and Chandler,M. (1992) *Microbiol. Rev.*, **56**, 509–528.
Laundon,C.H. and Griffith,J.D. (1988) *Cell*, **52**, 545–549.
Lee,E.C., MacWilliams,M.P., Gumport,R.I. and Gardner,J.F. (1991) *J. Bacteriol.*, **173**, 609–617.
Liu,L.F. and Wang,J.C. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 7024–7027.
Lupski,J.R. and Weinstock,G.M. (1992) *J. Bacteriol.*, **174**, 4525–4529.
Lynch,A.S. and Wang,J.W. (1993) *J. Bacteriol.*, **175**, 1645–1655.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Menzel,R. and Gellert,M. (1983) *Cell*, **34**, 105–113.
Miller,H.I. and Friedman,D.I. (1977) In Bukhari,A., Shapiro,J. and Adhya,S. (eds), *DNA Insertion Elements, Plasmids and Episomes*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 349–356.
Mizuno,T. (1987) *Nucleic Acids Res.*, **17**, 6827–6841.
Murphy,G. and Kavanagh,T. (1988) *Nucleic Acids Res.*, **16**, 5198.
Needleman,S.B. and Wunsch,C.D. (1970) *J. Mol. Biol.*, **48**, 443–453.
Newbury,S.F., Smith,N.H., Robinson,E.C., Hiles,I.D. and Higgins,C.F. (1987) *Cell*, **48**, 297–310.
Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl Acad. Sci.* USA, **85**, 2444–2448.
Pettijohn,D.E. (1982) *Cell*, **30**, 667–669.
Prentki,P., Chandler,M. and Galas,D.J. (1987a) *EMBO J.*, **6**, 2479–2487.
Prentki,P., Pham,M.-H. and Galas,D.J. (1987b) *Nucleic Acids Res.*, **15**, 10060.
Prentki,P., Pham,M.-H., Gamas,P., Chandler,M. and Galas,D.J. (1987c) *Gene*, **61**, 91–101.
Pruss,G.J. and Drlica,K. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 8952–8956.
Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) *Science*, **239**, 487–491.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Stenzel,T.T., Patel,P. and Bastia,D. (1987) *Cell*, **49**, 709–717.
Thompson,J.F. and Landy,A. (1988) *Nucleic Acids Res.*, **16**, 9687–9705.
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, Cold Spring Harbor, NY, pp. 357–361.
Wu,H.-M. and Crothers,D.M. (1984) *Nature*, **308**, 509–513.
Yang,C.-C. and Nash,H.A. (1989) *Cell*, **57**, 869–880.
Yang,Y. and Ames,G.F.-L. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 8850–8854.
Yang,Y. and Ames,G.F.-L. (1990) In Drlica,K. and Riley,M. (eds), *The Bacterial Chromosome*. American Society for Microbiology, Washington, DC, pp. 211–225.

Received on June 16, 1993; revised on September 9, 1993