

Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria

Charles J. Dorman, Jay C.D. Hinton and Andrew Free

The H-NS protein of *Escherichia coli* is a major component of the bacterial nucleoid and is a pleiotropic regulator of genome stability, recombination and gene expression^{1,2}. The StpA protein is 58% identical to H-NS at the level of amino acid sequence. Initially, StpA was identified as a multicopy suppressor of a splicing-defective *td* intron of bacteriophage T4 (Refs 3,4) and of *adi* gene expression in an *hns* mutant⁵. These highly related proteins, or paralogues (i.e. intraspecies homologues), have both common and distinct properties. Each can constrain DNA supercoils *in vitro* and they can repress transcription *in vivo* and *in vitro* from a synthetic *gal* promoter containing an upstream curved sequence⁶. However, StpA is tenfold more efficient than H-NS in promoting RNA annealing and strand exchange during splicing of a mutant *td* intron *in vivo* or the wild-type intron *in vitro*^{4,6}.

Each protein can inhibit both its own gene promoter and that of the other (negative cross and autoregulation)⁶. However, the patterns of expression differ (see Table 1). The production of *hns* mRNA is confined mainly to the logarithmic phase of growth⁷⁻⁹, although certain fusion constructs give conflicting results^{10,11}. Overall, the cell seems to attempt to maintain a constant ratio of H-NS to DNA (Ref. 8). Transcription of the *stpA* gene is repressed by H-NS in rich growth media; it is confined to a brief period in mid-logarithmic growth and can be induced strongly by osmotic stress and, to a lesser extent, by an increase in growth temperature¹². A more sustained induction of *stpA* is observed in minimal medium and this is dependent on the leucine-responsive regulatory protein (LRP)¹³. This stronger expression is abolished in an H-NS-independent manner if the culture undergoes carbon starvation¹². Expression of the *hns* gene is not affected by

The bacterial nucleoid-associated proteins H-NS and StpA can form homomeric or heteromeric complexes, a parallel with protein HU. Thus, functional modulation of H-NS and StpA by one another and by other proteins with appropriate interaction domains is possible. This has implications for bacterial pathogenesis and adaptation to environmental stress.

C.J. Dorman* is in the Dept of Microbiology, Moyné Institute of Preventive Medicine, University of Dublin, Trinity College, Dublin 2, Ireland; J.C.D. Hinton is in the Nuffield Dept of Clinical Biochemistry, Institute of Molecular Medicine, University of Oxford, Oxford, UK OX3 9DS; A. Free is in the Institute of Cell and Molecular Biology, University of Edinburgh, Darwin Building, King's Buildings, Mayfield Road, Edinburgh, UK EH9 3JR.
*tel: +353 1 608 2013,
fax: +353 1 679 9294,
e-mail: cjdorman@tcd.ie

growth in minimal medium and requires a different activator [the factor for inversion stimulation (FIS)] for optimal expression both *in vivo* and *in vitro*⁷. DsrA, an antisense RNA, blocks translation of *hns* mRNA by hybridization. Although DsrA RNA does not interact with *stpA* directly, the reduction in the intracellular level of H-NS protein causes StpA levels to increase¹⁴.

Two-domain structure of H-NS and StpA

Initial indications that H-NS consists of two functional domains came from genetic studies. Mutations within the carboxy-terminal domain, between amino acid residues 90 and 120, reduce DNA-binding activity, whereas mutations in the amino-terminal domain, between residues 12 and 65, retain DNA-binding activity but remove repressor function, probably because of a loss of protein-protein interactions¹⁵⁻¹⁸. Mutations at the extreme amino terminus have been reported to affect repression without altering DNA-binding or oligomerization properties¹⁹.

Confirmation that the carboxy-terminal domain of H-NS contains the DNA-binding region comes from studies with deletion derivatives¹⁹ and biochemical evidence that the environment of the W109 residue is altered upon DNA binding²⁰. A core DNA-binding motif (TWYG-GR-P) lies between residues 108 and 116 and is highly conserved among the proteobacterial clade of Gram-negative bacteria (Fig. 1). Furthermore, the RNA-annealing activity and the ability to promote trans-splicing *in vitro* are also functions of the carboxy-terminal domain of StpA (Ref. 21). Consequently, we refer to the carboxy-terminal region of H-NS as the nucleic-acid-binding domain.

The ability of the amino-terminal domain of H-NS to mediate oligomerization has been confirmed by protein cross-linking and gel filtration^{17,18}. The

oligomerization domain lies within a region predicted to be predominantly α -helical, which has broad similarity to eukaryotic structural coiled-coil proteins such as spectrin, myosin and dystrophin¹. We used the 'Coils' program from the Swiss Institute for Experimental Cancer Research (ISREC) (http://www.isrec.isbsib.ch/software/COILS_form.html) to analyse the sequences of mutant H-NS proteins defective in oligomerization. The program predicts that the region from amino acids 15 to 70 of H-NS will form a coiled-coil with high probability, but that a mutant protein (L30P)¹⁷ that no longer oligomerizes will not have a coiled-coil structure in the first half of this region. Similarly, mutating L30 to D affects both oligomerization ability¹⁷ and the coiled-coil prediction, whereas replacing this residue with A or K has a much smaller effect on oligomerization and results in mutant proteins that still have a high coiled-coil prediction for the whole of the oligomerization domain. Thus, there is a very good correlation between the oligomerization ability and the potential of this region to form a coiled-coil.

The boundary between the amino- and carboxy-terminal domains of H-NS and StpA is formed by a protease-sensitive linker extending between residues 76 and 88 (Fig. 1)²¹. This is the most divergent part of the proteins and is composed of amino acids that occur commonly in other linkers; no mutations affecting H-NS or StpA function have been isolated in this sequence^{17-19,21}.

MdbA, a microcin DNA-binding protein from uropathogenic *E. coli* (UPEC), was suggested originally to have limited amino-terminal homology to H-NS (Ref. 21). However, the insertion of a single base pair at nucleotide 405 of the *mdbA* gene sequence corrects a frame shift (possibly resulting from a sequencing error) to produce an open reading frame with full-length homology to H-NS (Fig. 1). The search for other H-NS homologues has revealed interesting patterns of domain conservation. A single example of a protein that consists simply of a tandemly duplicated carboxy-terminal H-NS domain without additional amino acid sequences has been reported (the KorB protein of plasmid pKM101)²². More commonly, proteins have been identified that have the carboxy-terminal nucleic-acid-binding domain fused to unrelated amino-terminal regions. Seven proteins that contain the nucleic-acid-binding domain of H-NS alone are shown in Fig. 1. The amino-terminal regions are predicted to form coiled-coils for all seven proteins, suggesting conservation at the level of function rather than sequence. For instance, the Orf4 protein encoded by the IncM plasmid R446, involved in the regulation of conjugation

Table 1. Differential expression of the *hns* and *stpA* genes^a

Regulator/condition ^b	<i>hns</i> expression	<i>stpA</i> expression
H-NS	Represses ^{6,10,11}	Represses ^{6,12,13}
StpA	Represses ⁶	Represses ⁶
LRP	Not determined	Activates ^{12,13}
FIS	Activates ⁷	No effect ¹²
DsrA	Represses ¹⁴	No effect ¹⁴
Logarithmic growth	Activates ⁷⁻⁹	Transiently activates ¹²
Minimal medium	No effect ^c	Activates ^{12,13}
Temperature increase	No effect ^c	Activates ^{12,13}
Cold shock	Activates ⁴⁰	Not determined
Osmotic stress	No effect ^{c,9}	Activates ¹²
Carbon starvation	No effect ^c	Represses ¹²
DNA relaxation	No effect ⁴¹	Represses ¹²

^aTrans-acting factors and environmental conditions affecting expression of the *Escherichia coli* *hns* and *stpA* genes.

^bAbbreviations: LRP, leucine-responsive regulatory protein; FIS, factor for inversion stimulation.

^cA. Free and C.J. Dorman, unpublished.

pilus expression, possesses a carboxy-terminal H-NS domain²³. Similarly, the LEE (locus of enterocyte effacement) pathogenicity island of enteropathogenic *E. coli* (EPEC) includes a gene coding for a protein that contains the nucleic-acid-binding domain of H-NS (Ref. 24). The existence of such hybrid proteins, which contain amino-terminal coiled-coil regions linked to the homologue of the carboxy-terminal domain of H-NS, shows that the H-NS nucleic-acid-binding domain has been used to generate novel regulatory DNA-binding proteins that have properties distinct from those of H-NS. The various amino-terminal coiled-coil regions are thought to mediate oligomerization of the H-NS-related DNA-binding proteins in a fashion that is analogous to heteromer formation between the *fos* and *jun* transcription factors. Heteromeric interactions among these eukaryotic proteins have been reviewed elsewhere (see Ref. 25).

Homomeric and heteromeric complexes

Protein cross-linking and genetic studies have shown that H-NS and StpA form homomeric and heteromeric complexes *in vivo* and that multimer formation is mediated by the amino-terminal domains of the proteins^{15,18,21}. The precise number of subunits in a functional H-NS or StpA complex is unclear, but it is probably at least tetrameric¹⁶. Whether or not H-NS forms even larger oligomeric structures remains an open question, and this topic has been covered in a recent review². Evidence that the oligomeric state of H-NS influences the ability of the protein to recognize curved DNA and to bend DNA has been presented, together with data showing that deletion of Pro116 or substituting it with alanine, appears to disrupt oligomerization¹⁶. Because Pro116 lies within the DNA-binding and not the oligomerization domain of H-NS (Fig. 1), its effect on oligomerization might be indirect. Further work is required to resolve this matter.

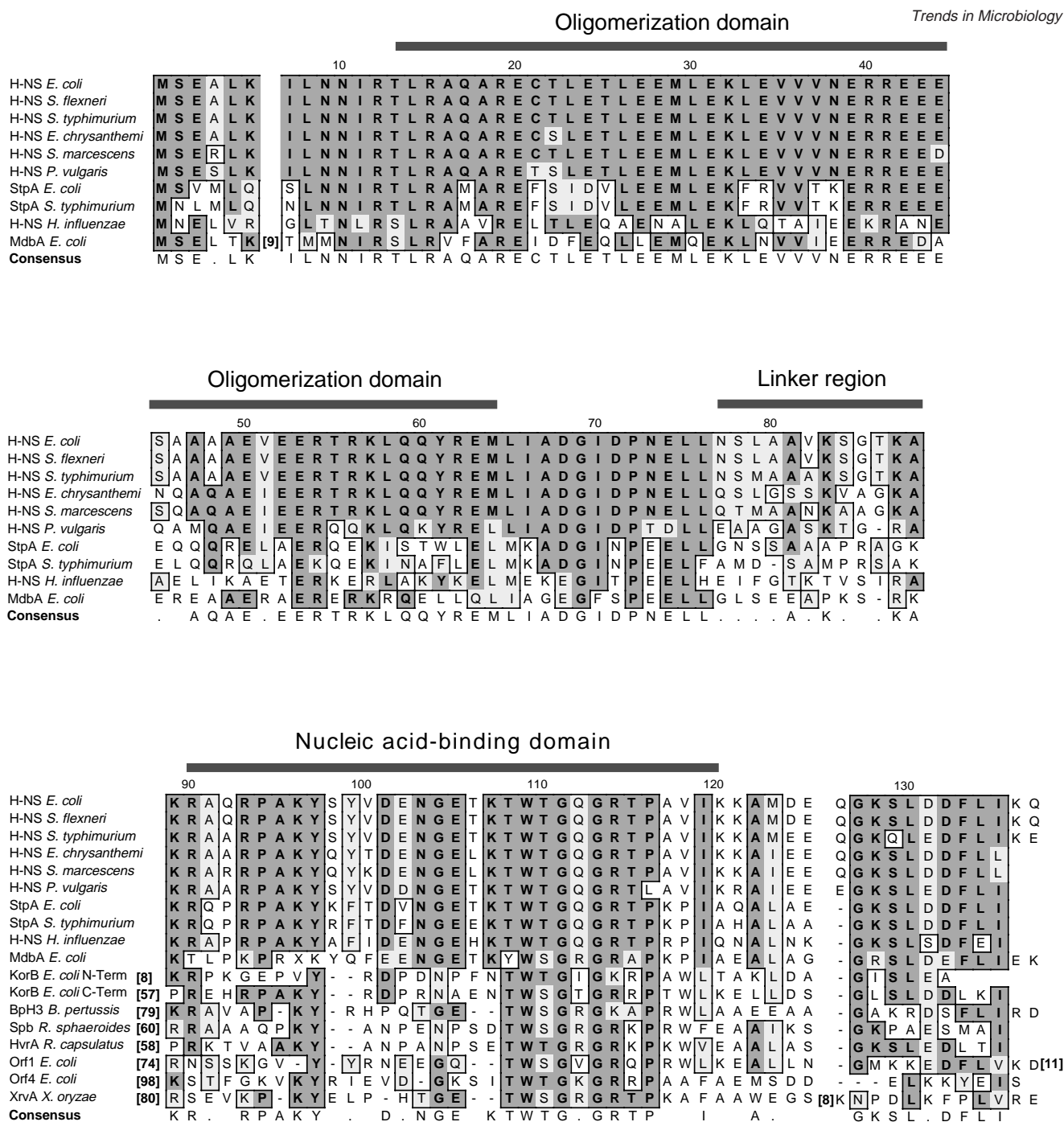


Fig. 1. H-NS-related bacterial proteins. Multiple sequence alignment of H-NS homologues. Bold residues with dark shading indicate sequence identity in >50% of all sequences, and boxed residues with light shading show conservative changes of residues in >50% of all sequences. The consensus sequence is indicated below the aligned sequences, and the numbers refer to the residues from the *Escherichia coli* H-NS amino acid sequence. The accession numbers for the sequences are: H-NS *E. coli*, Sw:P08936; H-NS *Shigella flexneri*, Sw:P09120; H-NS *Salmonella typhimurium*, Sw:P17428; H-NS *Erwinia chrysanthemi*, EMBL:X89444; H-NS *Serratia marcescens*, Sw:P18955; H-NS *Proteus vulgaris*, Sw:P18818; StpA *E. coli*, Sw:P30017; StpA *S. typhimurium*, GB:AF009363; H-NS *Haemophilus influenzae*, Sw:P43831; MdbA *E. coli*, GB:U47048 (the frame-shift at nucleotide 405 has been corrected, and residue 105 of MdbA has been designated 'X'); KorB *E. coli*, PIR:I79262 [KorB N-term and KorB C-term represent the amino-terminal (residues 9–49) and carboxy-terminal (residues 58–101) regions of the KorB protein, respectively]; BpH3 *Bordetella pertussis*, GB:U82566; Spb *Rhodobacter sphaeroides*, PIR:S82916; HvrA *Rhodobacter capsulatus*, Sw:P42505; Orf1 *E. coli*, GB:AF022236 [a protein virtually identical to Orf1 (L0054) has recently been found in *E. coli* O157:H7; GB:AF071034]; Orf4 *E. coli*, PIR:S34257; XrvA *Xanthomonas oryzae*, EMBL:X97866. The bracketed numbers in the Orf1, XrvA and MdbA sequences indicate amino acids that are not shown. The KorB, Orf1, Orf4 and XrvA proteins have no similarity to H-NS before residue 89. Bracketed numbers before these incomplete sequences indicate the position within the whole sequence. See Ref. 21 for further references. EMBL, GB, PIR and Sw refer to the EMBL, Genbank, Protein Information Resource and Swissprot databases, respectively. The oligomerization domain, the linker region and the nucleic-acid-binding domain are indicated.

The ability of H-NS and StpA to form homomeric and heteromeric complexes raises the issue of the physiological roles of the different complexes. Data from two-dimensional gel analysis of total protein in *E. coli* show that certain genes require both H-NS and StpA for normal regulation. Although these proteins might act independently in each case, the observation is consistent with a specific role for the heteromeric complexes⁶. Recent work has shown that the *bgl* operon of *E. coli*, a classic H-NS-regulated system, can be repressed by the amino-terminal domain of H-NS targeted to the promoter by the DNA-binding domain of StpA (Ref. 15). By contrast, the *proU* operon, another well-established example of an H-NS-repressed promoter, is indifferent to physiological levels of StpA. Thus, although many bacterial genes are subject to repression by H-NS, only a subset is also influenced by StpA. Currently, there are no clear examples of promoters regulated by StpA alone.

The differential expression of H-NS and StpA is likely to influence the abundance of the homomeric and heteromeric complexes in the cell at any time. H-NS seems to be present constantly at a level of ~100 000 molecules per genome (J.C.D. Hinton, unpublished), whereas StpA is expressed only transiently in rich growth medium¹², and the intracellular level is thought to be much lower (J.C.D. Hinton, unpublished). However, StpA expression is enhanced in minimal medium and in response to environmental changes, such as temperature or osmotic shock^{12,13}. Therefore, StpA might become an important cellular protein under environmental conditions encountered by the bacterium outside the laboratory, such as during interaction with an animal or plant host.

HU α and HU β

A direct analogy can be drawn with protein HU in *E. coli*. This is another highly abundant nucleoid-associated protein, being present at ~30 000 dimers per cell in exponentially growing cultures²⁶. Like eukaryotic histones, HU is a basic protein and, in *E. coli*, it is a heterodimer of α and β subunits of 9.5 kDa with an amino acid sequence similarity of ~70%. Thus, as in the case of H-NS and StpA, HU α and HU β can be regarded as paralogous. The subunits are encoded by the genes *hupA* and *hupB*, respectively. Analogous to H-NS and StpA (Refs 21,27), HU shows little sequence specificity in its binding sites for DNA, although it displays a preference for bent or kinked DNA (Ref. 28). HU contributes to transposition, DNA replication initiation at *oriC* and modulation of the binding specificity of other DNA-binding proteins and also affects the organization of the nucleoid and the transcriptional control of several genes²⁶.

Like *hns* and *stpA*, the *hup* genes in *E. coli* are subject to differential regulation (see Table 2). The *hupA*

Table 2. Differential expression of the *hupA* and *hupB* genes^a

Regulator/condition ^b	<i>hupA</i> expression	<i>hupB</i> expression
HU α	Represses ²⁹	Represses ²⁹
HU β	Represses ²⁹	Represses ²⁹
CRP	Activates ^{30,31,c}	Activates ³⁰
FIS	Activates ³⁰	Represses ³⁰
Logarithmic growth	Active earlier ³¹	Active later ³¹
Amino acid starvation	No effect ³⁰	Represses ³⁰

^aTrans-acting factors and environmental conditions affecting expression of the *Escherichia coli* *hupA* and *hupB* genes.

^bAbbreviations: CRP, cyclic AMP receptor protein; FIS, factor for inversion stimulation.

^cCRP activates both *hupA* and *hupB* at the transcriptional level, but its positive effect on the production of HU α protein is much less than that on HU β (Ref. 31).

and *hupB* promoters are both negatively autoregulated²⁹ but have differential responses to growth phase, amino acid starvation and the regulatory proteins FIS and CRP (the factor for inversion stimulation and the cyclic-AMP receptor protein, respectively)^{30,31}. As a result of this regulation, the composition of HU varies throughout the growth phase, with α 2 homodimers predominating in early log phase and the α β heterodimer predominating in stationary phase. Therefore, it seems that *E. coli* has the ability to vary the relative amounts of the HU subunits in response to growth (via FIS and autoregulation), carbon source (via CRP) and amino acid availability (via the stringent response). Indications of the significance of homo- and heterodimer formation by HU subunits come from genetic studies. Mutants deficient in *hupA* expression are less fit than *hupB* mutants under laboratory growth conditions, which is consistent with a predominance of α 2 homodimers in early log phase. However, the α β form is required for long-term survival in stationary phase, indicating that it has functions that cannot be provided by the α 2 homodimer. Moreover, although the α 2 and α β forms of HU can constrain DNA supercoils, the β 2 form, present in small amounts in stationary phase, cannot do this^{31,32}. Therefore, changes in the composition of HU seem to have consequences for its ability to organize bacterial chromatin and to influence gene expression.

Niche adaptation and homomer and heteromer formation

The differential expression of StpA and H-NS, and of HU α and HU β , provides the cell with the capacity to vary the composition of its nucleoid-associated proteins as a function of the environment and of growth phase. Bacteria other than *E. coli* possess similar or identical sets of such proteins. *Shigella flexneri*, *Salmonella typhimurium* and *Erwinia chrysanthemi* have a similar H-NS and StpA complement to *E. coli*, but *Haemophilus influenzae* has only one *hns/stpA*-like gene, and Gram-positive bacteria seem to have neither H-NS nor StpA (Refs 1,33; A. Free, unpublished). In *E. coli*, *E. chrysanthemi*, *S. flexneri* and *S. typhimurium*, HU is encoded by two *hup* genes³⁴. By contrast, *H. influenzae* has just a single *hupA* gene,

Questions for future research

- The ability of StpA and H-NS to form heteromers and their patterns of expression might suggest that pure homomers of either protein do not exist *in vivo*. By analogy with the α and β subunits of HU, should H-NS and StpA be considered as functional parts of a heteromer, rather than in isolation?
- StpA and H-NS homomers are significantly different in their abilities to bind RNA and DNA and to repress transcription. Will an understanding of the biochemical properties of H-NS–StpA heteromers give us functional data that will relate to the situation *in vivo*?
- Can any of the proteins bearing homology to H-NS at the carboxyl terminus interact with H-NS–StpA complexes to modify further their properties when they are co-expressed in the same cell?
- By which mechanism(s) do H-NS and StpA influence transcription? Is direct contact with RNA polymerase involved?

whereas *Pseudomonas aeruginosa* has just an *hupB*-like gene³⁴. The Gram-positive bacterium *Bacillus subtilis* has a homodimeric HU-like protein, but, in phylogenetic analysis, the gene that codes for it groups separately from *hupA* and *hupB* (Ref. 34).

This capacity for homomer and heteromer formation in the H-NS/StpA and HU proteins, which is characteristic of *E. coli* and its close relatives, might reflect a need for greater flexibility of response to environmental change than is the case for other bacteria, or perhaps these bacteria achieve such flexibility by other means. One intriguing possibility relates to the MdbA protein, which is encoded by plasmid p24-2 from UPEC (Ref. 35). MdbA possesses both the oligomerization and nucleic-acid-binding domains of H-NS and, thus, has the potential to interact with H-NS–StpA complexes and provide a further level of modulation of their properties when expressed in *E. coli*. This could contribute to the ability of these pathogenic strains to adapt gene regulation to the variety of conditions encountered during the infection process.

Conclusions

The host–pathogen interactions engaged in by bacteria are dynamic and require the pathogen to adapt to a series of environments, many of which are hostile. To succeed, the bacteria bring into play a sophisticated network of gene-regulatory functions, most of which are aimed at modulating the transcriptional profile of the cell. H-NS is a negative regulator of invasion-gene expression in *S. flexneri* and of adhesin expression in pathogenic strains of *E. coli* and contributes to the ability of *S. typhimurium* to infect mice^{36–39}. With their pleiotropic effects on transcription (and other DNA-based transactions), and their ability to form alternative protein–protein relationships with different properties, the H-NS/StpA nucleoid-associated proteins are well placed to assist the pathogen in the subtleties of host adaptation. Further studies on how the properties of these complexes vary as a function of protein subunit composition should establish to what extent this potential is utilized.

Acknowledgements

We thank members of the Dorman and Hinton laboratories for many useful discussions and comments on the manuscript. A.F. was supported by a Wellcome Trust Prize Fellowship (Ref. 044711/Z/95/Z) and currently holds a Wellcome Trust Advanced Training Fellowship. J.C.D.H. is supported by Wellcome Trust Programme Grant Ref. 045490. C.J.D. and J.C.D.H. are members of an EU TMR network (Contract ERBFMRXCT98-0164).

References

- 1 Atlung, T. and Ingmer, H. (1997) *Mol. Microbiol.* 24, 7–17
- 2 Williams, R.M. and Rimsky, S. (1997) *FEMS Microbiol. Lett.* 156, 175–185
- 3 Zhang, A. and Belfort, M. (1992) *Nucleic Acids Res.* 20, 6735
- 4 Zhang, A. *et al.* (1995) *RNA* 1, 783–793
- 5 Shi, X. and Bennett, G.N. (1994) *J. Bacteriol.* 176, 6769–6775
- 6 Zhang, A. *et al.* (1996) *EMBO J.* 15, 1340–1349
- 7 Falconi, M. *et al.* (1996) *Mol. Microbiol.* 19, 965–975
- 8 Free, A. and Dorman, C.J. (1995) *Mol. Microbiol.* 18, 101–113
- 9 Hinton, J.C.D. *et al.* (1992) *Mol. Microbiol.* 6, 2327–2337
- 10 Dersch, P. *et al.* (1993) *Mol. Microbiol.* 8, 875–889
- 11 Ueguchi, C. *et al.* (1993) *Mol. Gen. Genet.* 236, 171–178
- 12 Free, A. and Dorman, C.J. (1997) *J. Bacteriol.* 179, 909–918
- 13 Sondén, B. and Uhlin, B.E. (1996) *EMBO J.* 15, 4970–4980
- 14 Lease, R.A. *et al.* (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 12456–12461
- 15 Free, A. *et al.* (1998) *J. Bacteriol.* 180, 994–997
- 16 Spurio, R. *et al.* (1997) *EMBO J.* 16, 1795–1805
- 17 Ueguchi, C. (1997) *J. Mol. Biol.* 274, 145–151
- 18 Williams, R.M. *et al.* (1996) *J. Bacteriol.* 178, 4335–4343
- 19 Ueguchi, C. *et al.* (1996) *J. Mol. Biol.* 263, 149–162
- 20 Tippner, D. and Wagner, R. (1995) *J. Biol. Chem.* 270, 22243–22247
- 21 Cusick, M.E. and Belfort, M. (1998) *Mol. Microbiol.* 28, 847–857
- 22 Moré, M.I. *et al.* (1996) *J. Bacteriol.* 178, 4392–4399
- 23 Tietze, E. and Tschäpe, H. (1994) *J. Basic Microbiol.* 34, 105–116
- 24 Elliott, S.J. *et al.* (1998) *Mol. Microbiol.* 28, 1–4
- 25 Hodges, R. (1992) *Curr. Biol.* 2, 122–124
- 26 Drlica, K. and Rouvière-Yaniv, J. (1987) *Microbiol. Rev.* 51, 301–319
- 27 Jordi, B.J.A.M. *et al.* (1997) *J. Biol. Chem.* 272, 12083–12090
- 28 Pontiggia, A. *et al.* (1993) *Mol. Microbiol.* 7, 343–350
- 29 Kohno, K. *et al.* (1990) *J. Mol. Biol.* 213, 27–36
- 30 Claret, L. and Rouvière-Yaniv, J. (1996) *J. Mol. Biol.* 263, 126–139
- 31 Claret, L. and Rouvière-Yaniv, J. (1997) *J. Mol. Biol.* 273, 93–104
- 32 Tanaka, H. *et al.* (1995) *Mol. Gen. Genet.* 248, 518–526
- 33 Fleischmann, R.D. *et al.* (1995) *Science* 269, 496–512
- 34 Oberro, J. and Rouvière-Yaniv, J. (1996) *J. Bacteriol.* 178, 293–297
- 35 O'Brien, G.J. and Mahanty, H.K. (1994) *Plasmid* 31, 288–296
- 36 Dorman, C.J. and Porter, M.E. (1998) *Mol. Microbiol.* 29, 677–684
- 37 Göransson, M. *et al.* (1990) *Nature* 344, 682–685
- 38 Harrison, J. *et al.* (1994) *Mol. Microbiol.* 13, 133–140
- 39 White-Ziegler, C.A. *et al.* (1998) *Mol. Microbiol.* 28, 1121–1137
- 40 La Teana, A. *et al.* (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 10907–10911
- 41 Free, A. and Dorman, C.J. (1994) *Mol. Microbiol.* 14, 151–161