IL NUCLEOIDE

- la molecola di DNA cromosomico di E.coli è lunga circa 1.6mm

- è contenuta in una cellula di 2 μm di lunghezza e 1μm di larghezza

- Un compattamento casuale della molecola determinerebbe un volume di circa 200mm^3 circa 400 volte superiore al volume del nucleoide

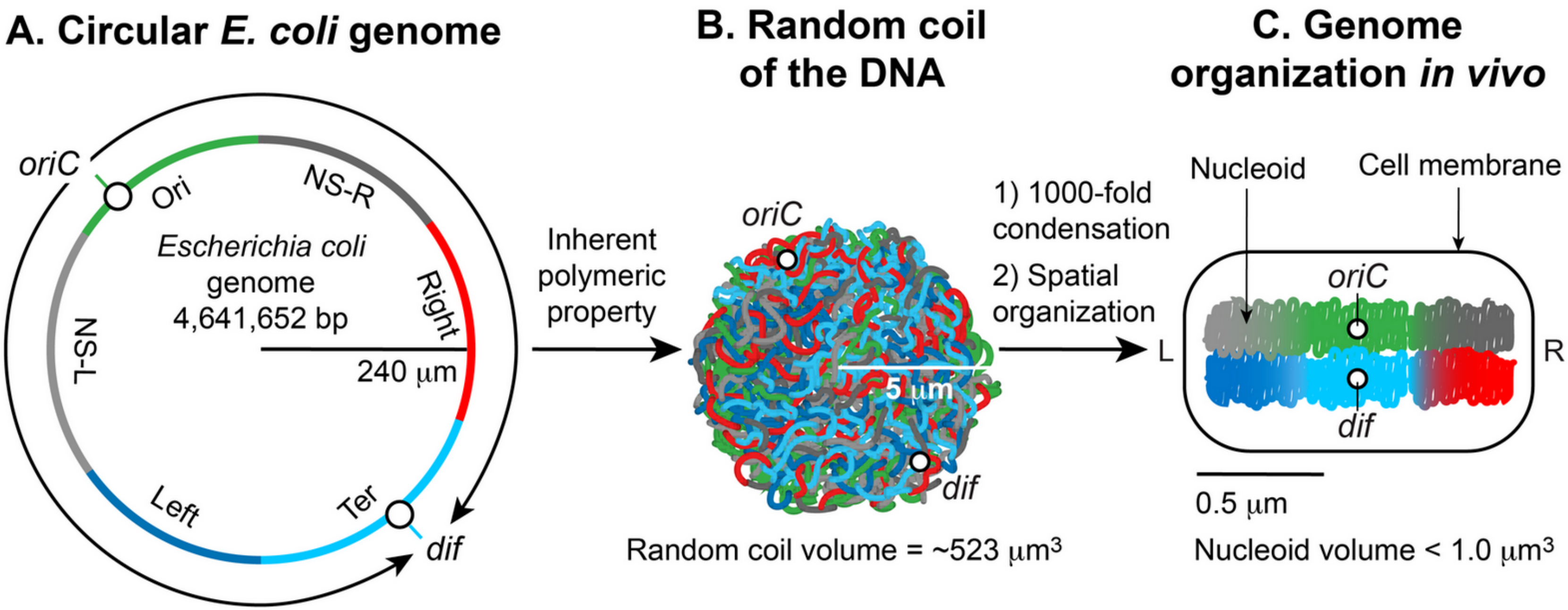
- Il volume del nucleoide è di circa di E.coli 0.5 μm

- Il cromosoma è quindi estremamente organizzato in anse topologicamente indipendenti circa 100 da 50 kb

L’organizzazione del genoma batterico è caratterizzata dalla presenza di macrodomini funzionali , ampie regioni di DNA nel qule ogni singolo gene ha un corretto livello di espressione che dipende dal suo orientamento dalla sua posizione rispetto all’origine.

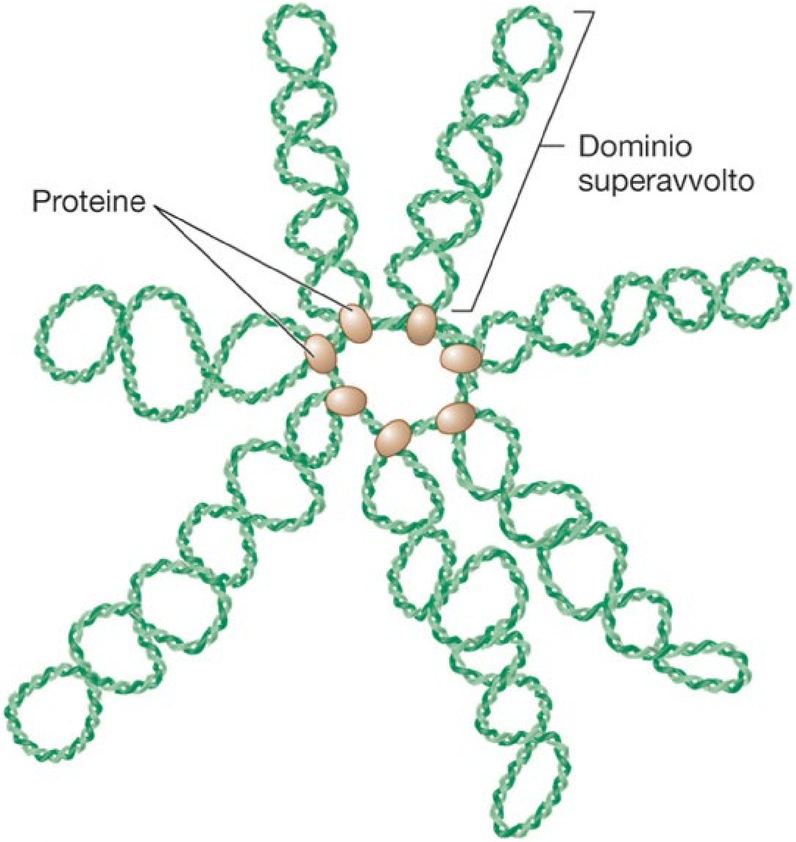
Se alcune regioni vengono invertite come orientamento sul genoma (pur rimanendo presenti i geni) si può avere la non espressione o espressione molto ridotta: queste grandi inversioni possono indurre anche morte nella cellula .

Elemento importante è anche il posizionamento rispetto all’origine: geni posizionati in posizione prossimale all’origine sono espressi di più in modo statisticamente significativo rispetto a quelli in posizione distale


A. Circular *E. coli* genome
oriC
Ori
NS-R
Escherichia coli
genome
4,641,652 bp
Right
NS-L
240 µm
Left
Ter
dif
Inherent
polymeric
property
B. Random coil
of the DNA
oriC
5 µm
dif
Random coil volume = ~523 µm³
1) 1000-fold
condensation
2) Spatial
organization
C. Genome
organization *in vivo*
Nucleoid
Cell membrane
oriC
L
R
dif
0.5 µm
Nucleoid volume < 1.0 µm³

Il cromosoma batterico è organizzato in numerosi domini superavvolti stabilizzati dal legame con proteine specifiche alla base dell'ansa.

In Escherichia coli si calcolano circa 100 domini



(d) Cromosoma con domini superavvolti

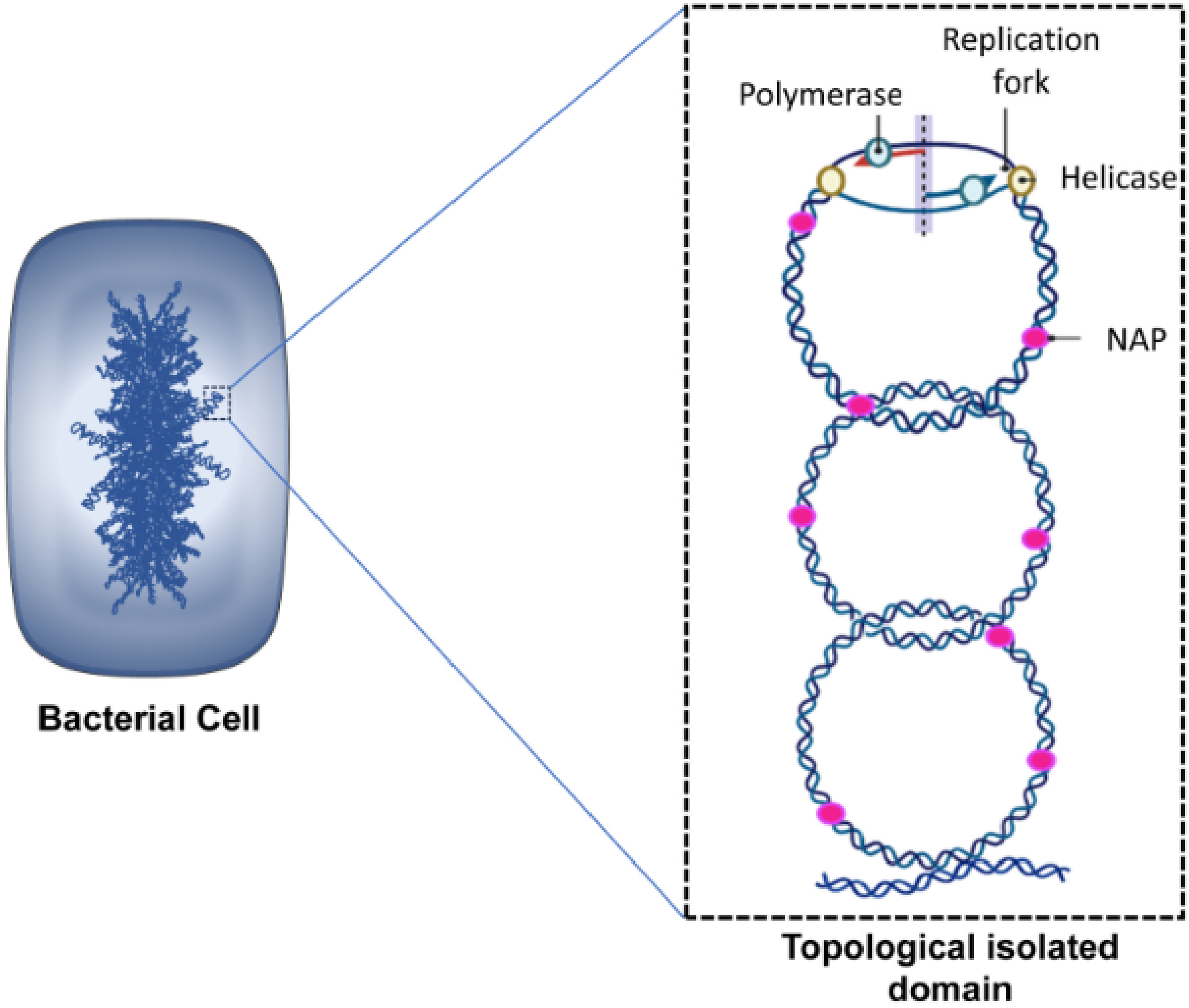
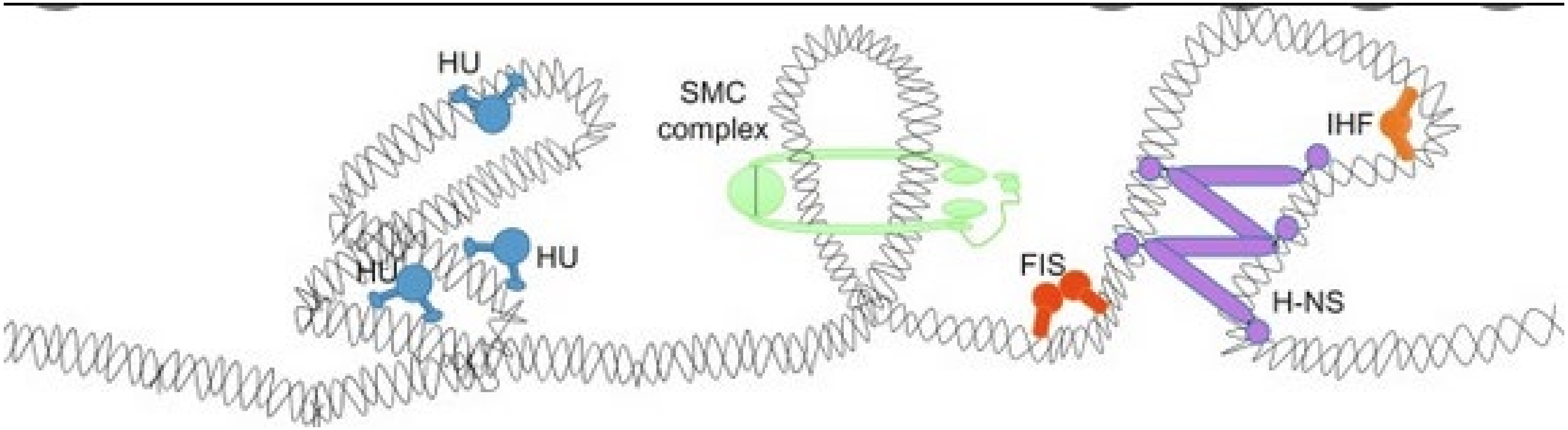
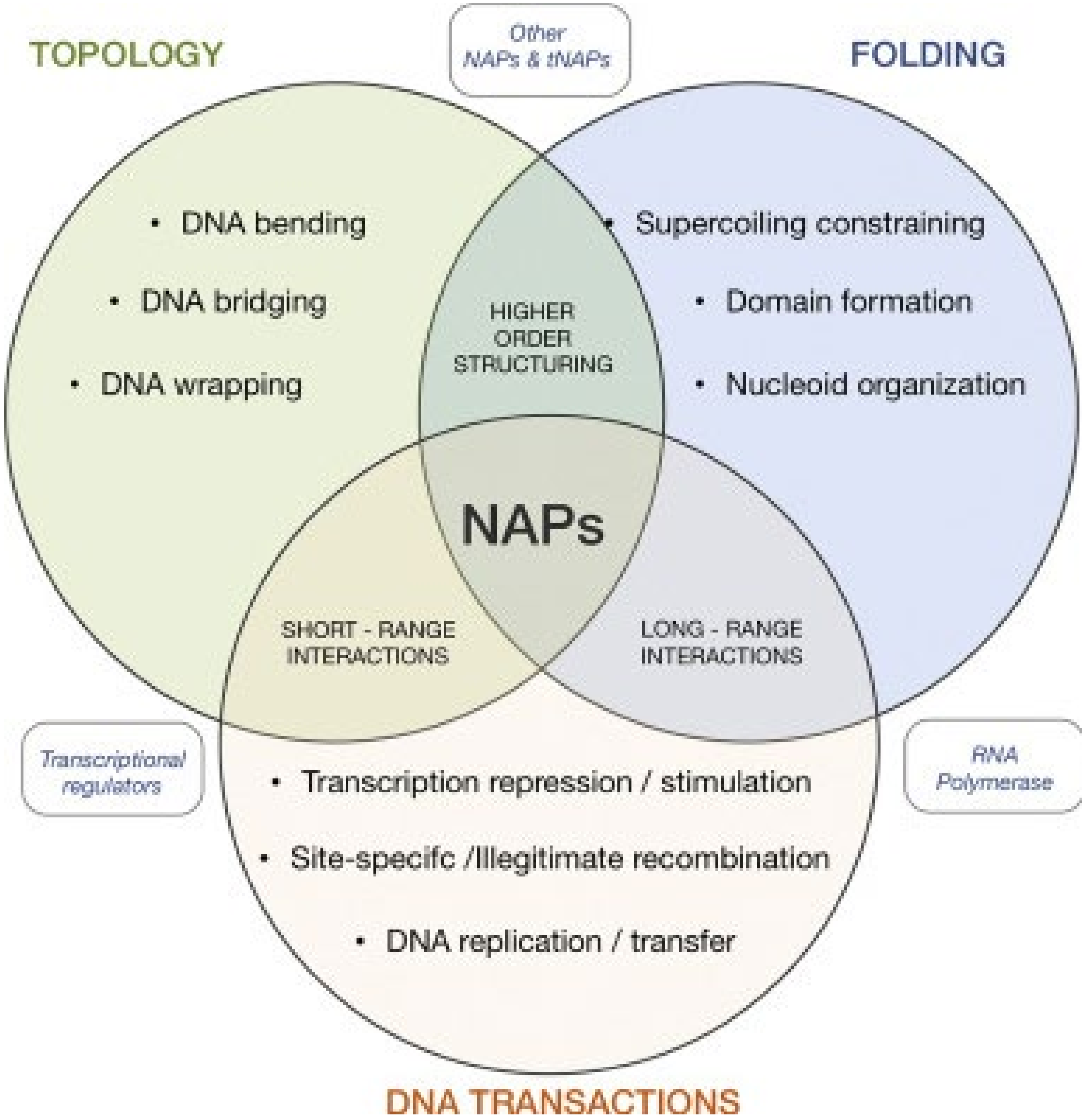


Fig. 1. :Illustration of bottlebrush nucleoid model: The figure displays topological domains or loops (bristles) emerging from the central nucleoid core. These loops are spatially organized and can vary in size and position in response to cellular and metabolic states via DNA-binding proteins. DNA binding proteins, such as NAPs, serve important functions by bridging various DNA segments, forming loops, and assisting in DNA replication, transcription, and chromosome segregation, thereby dynamically altering the nucleoid structure (Wang et al., 2013; Verma et al., 2019).



Quali sono le proteine associate al nucleoide?
Come riescono a compattare il cromosoma batterico?

SMC complex corrisponde alle proteine MUK


TOPOLOGY
Other NAPs & tNAPs
FOLDING
• DNA bending
• DNA bridging
• DNA wrapping
HIGHER ORDER STRUCTURING
• Supercoiling constraining
• Domain formation
• Nucleoid organization
NAPs
SHORT - RANGE INTERACTIONS
LONG - RANGE INTERACTIONS
Transcriptional regulators
RNA Polymerase
• Transcription repression / stimulation
• Site-specifc /Illegitimate recombination
• DNA replication / transfer
DNA TRANSACTIONS

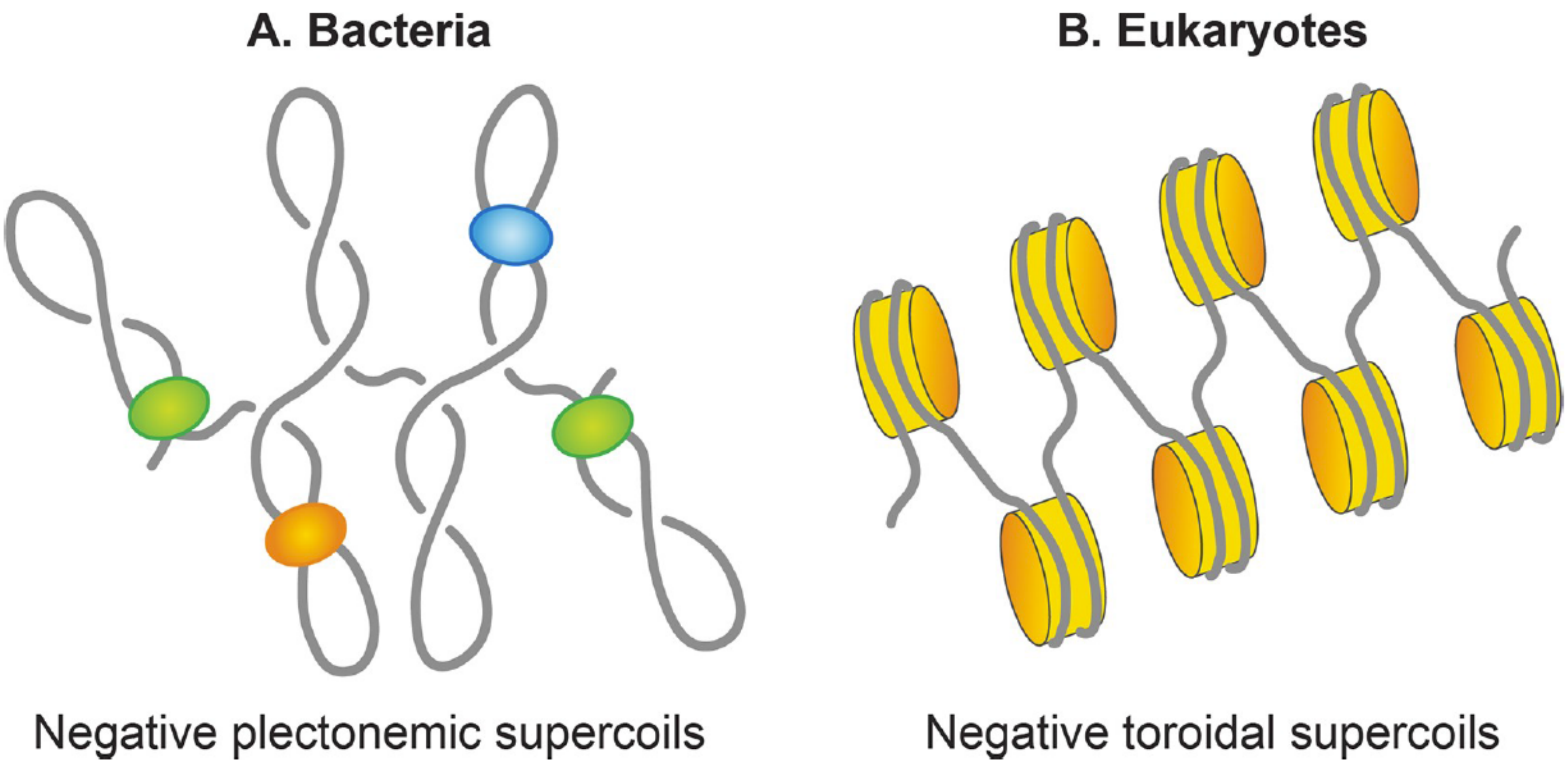


Fig 5. Basic units of genomic organization in bacteria and eukaryotes. A. A bacterial genome organizes as plectonemic supercoils. Half of the supercoils are present in free form, and nucleoid-associated proteins (NAPs), shown as colored spheres, restrain the remaining half. **B.** In contrast, a eukaryotic genome organizes as toroidal supercoils, induced by the wrapping of DNA around histone proteins (orange color). An octamer of histones with 146 wrapped DNA refers to as nucleosome, and the genome organizes into a repeating array of nucleosomes.

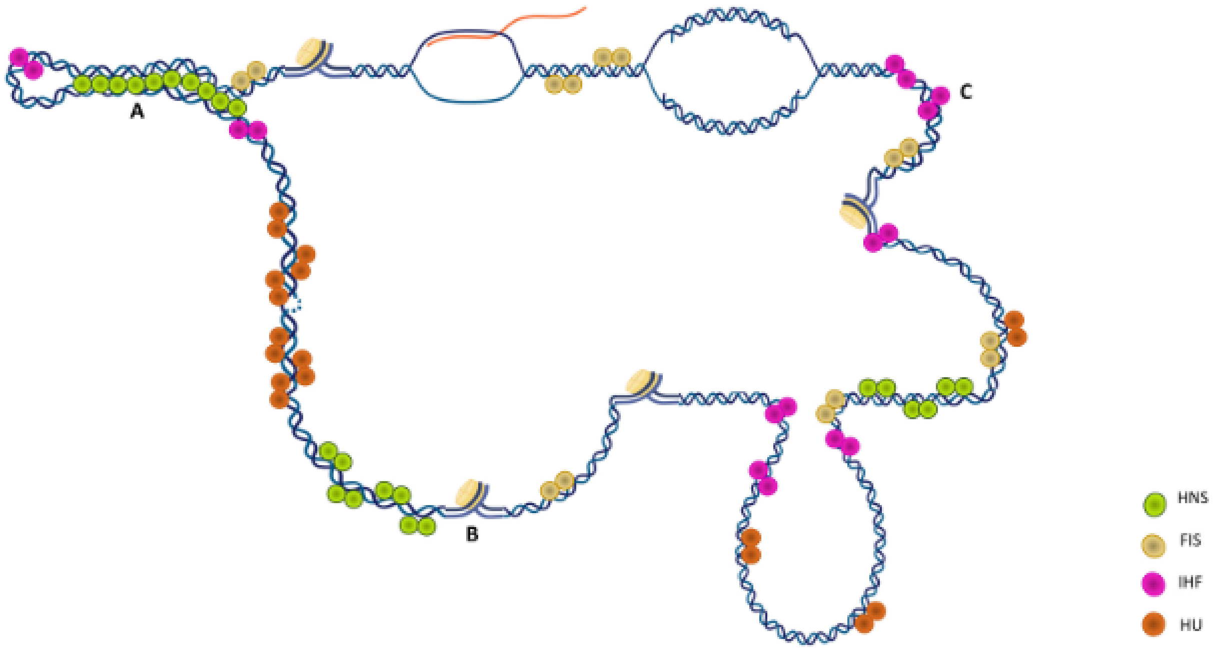
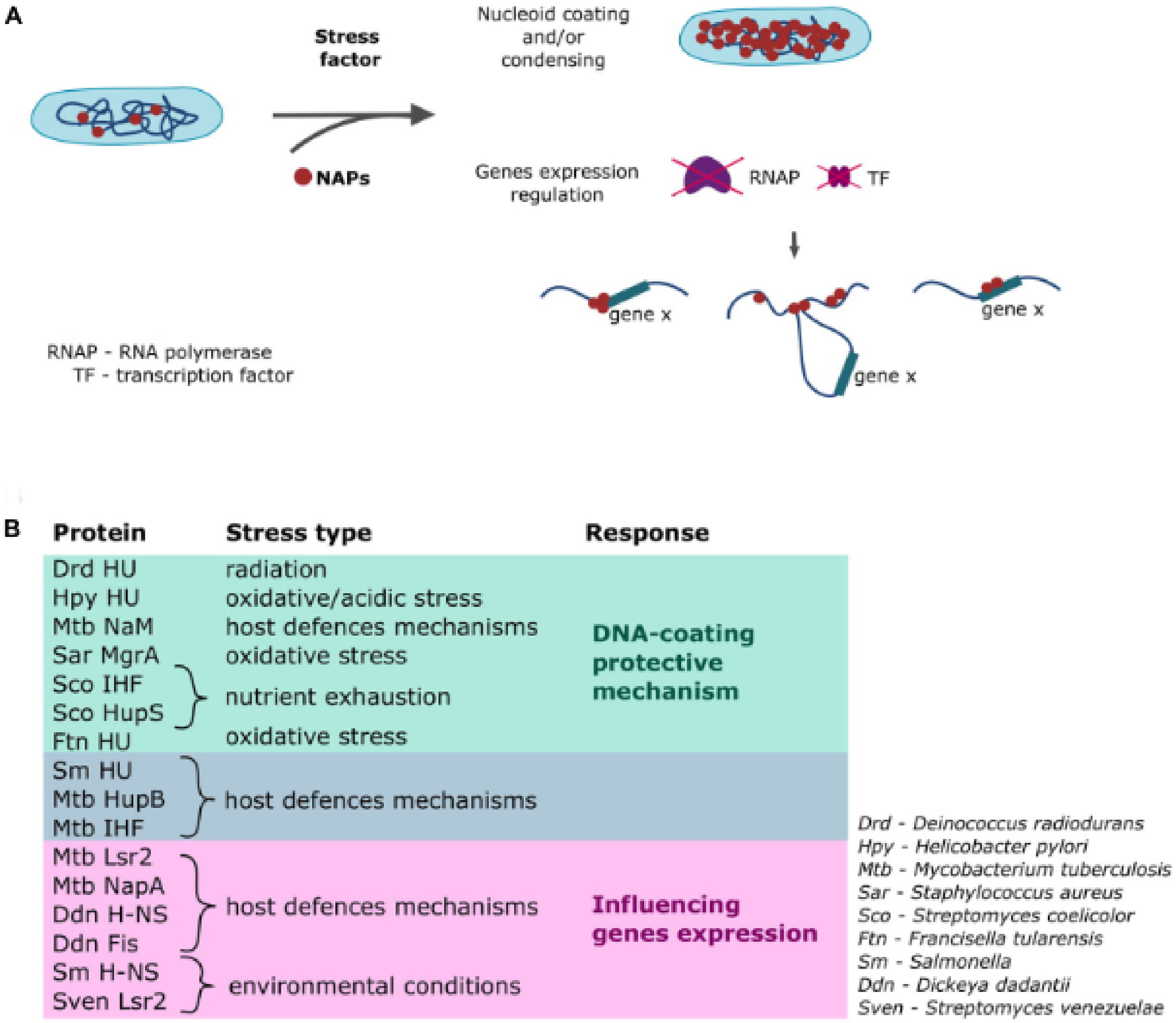


Fig. 2. : Schematic representation of DNA supercoiling aided replication and transcription processes by diverse set of NAPs (A) bridgers (B) wrappers (C) benders: DNA bridging and stiffening activity are shown by H-NS dimers (green), bending by 50° –90° and wrapping shown by FIS dimers (yellow), bending nearly by 160° shown by IHF dimers (pink), and binding nicks, gaps shown by HU (orange) (Dillon and Dorman, 2010).



Protein	Stress type	Response
Drd HU	radiation	**DNA-coating protective mechanism**
Hpy HU	oxidative/acidic stress	
Mtb NaM	host defences mechanisms	
Sar MgrA	oxidative stress	
Sco IHF	nutrient exhaustion	
Sco HupS		
Ftn HU	oxidative stress	
Sm HU	host defences mechanisms	
Mtb HupB		
Mtb IHF		
Mtb Lsr2	host defences mechanisms	**Influencing genes expression**
Mtb NapA		
Ddn H-NS		
Ddn Fis		
Sm H-NS	environmental conditions	
Sven Lsr2		

FIGURE 2 | Involvement of NAPs in stress responses. **(A)** General mechanisms through which NAPs act in response to a stress factor (Dillon and Dorman, 2010; Meyer and Grainger, 2013; Kriel et al., 2018; Trojanowski et al., 2019). **(B)** Examples of the homologs of the canonical *E. coli* NAPs involved in the cellular response triggered upon detection of stress conditions.

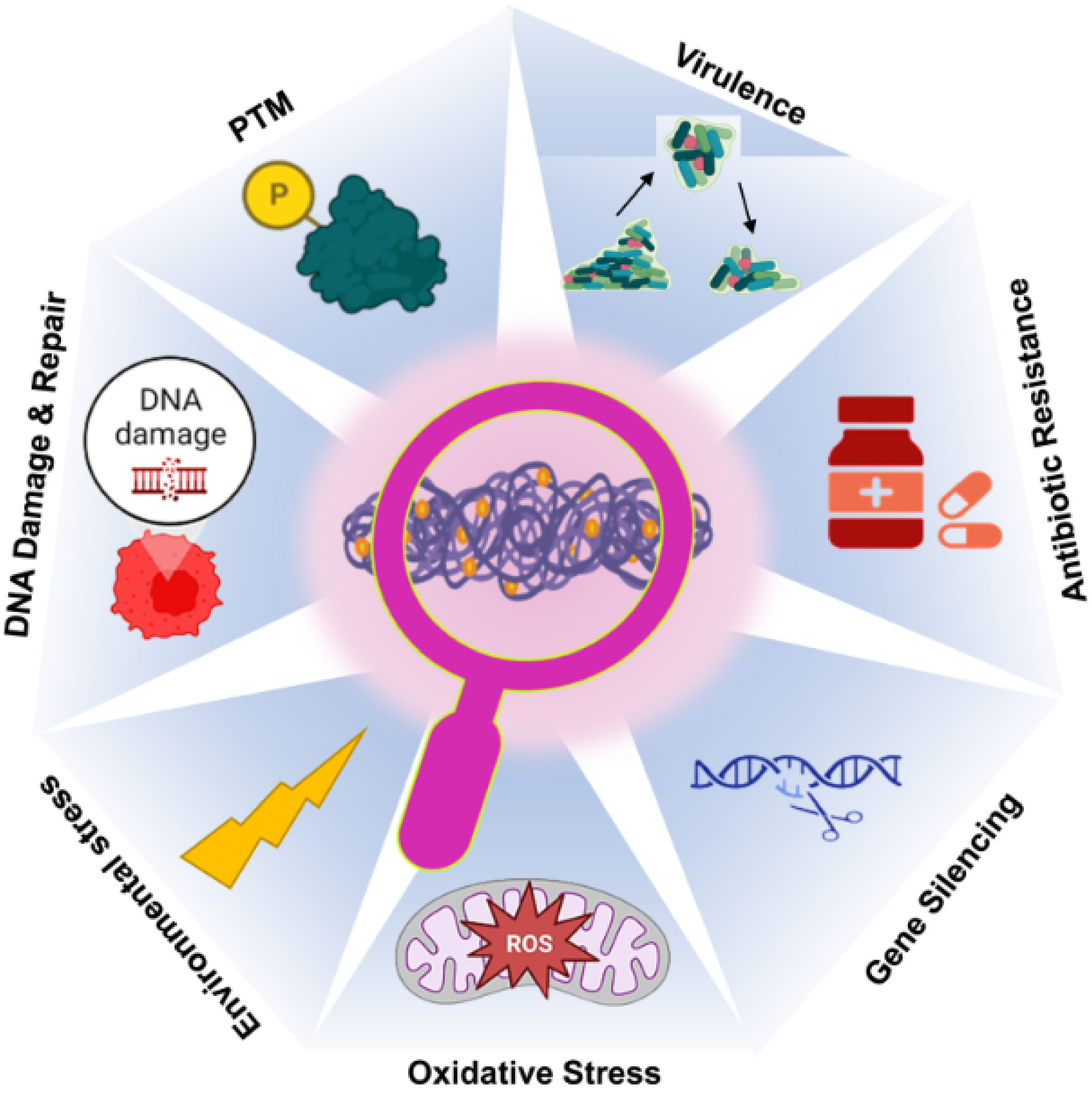


Fig. 4. :A graphical illustration of the various roles of NAPs: Nucleoid-associated proteins (NAPs) serve diverse functions in the ESKAPE group, from enhancing virulence and antibiotic resistance to regulating gene expression, managing oxidative and environmental stress, and aiding DNA repair and post-translational modification

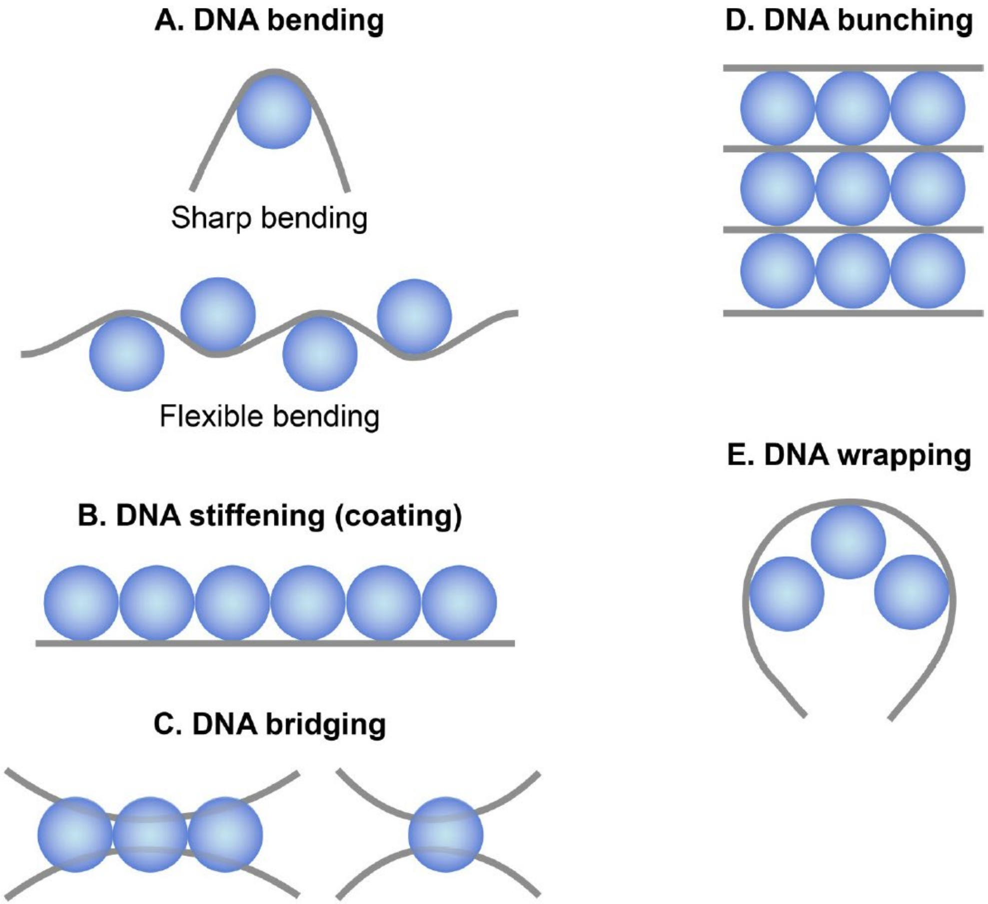
Le principali proteine associate al nucleoide batterico (NAP nucleoid associated proteins)

NAP	**Properties and abundance**	**Structure**	**Genes**	**Main functional activities**
HU (Heat Unstable nucleoid protein)	• basic • abundant in exponential phase (15000-30000 dimers/cell)	heterodimer; HUα , 9.2 kDa HUβ, 9.5 kDa	*hup*A (90.4 min) *hup*B (9.9 min)	• compacts DNA into nucleosome-like structures • induces DNA curvature • recognizes curved DNA, gapped regions, and 3-/4way junctions • involved in DNA replication and recombination
IHF (Integration Host Factor)	• basic • abundant in stationary phase (25000-3000 dimers/cell) • high amino acid identity between IHF and HU subunits • DNA binding preference: WATCAANNNNTTR	heterodimer; IHFα, 11.2 kDa IHFβ, 10.7 kDa	*him*A (38.6 min) *him*D (25 min)	• induces very strong DNA curvature (up to 140°) • participates in site-specific recombination, transposition, and DNA replication
FIS (Factor for Inversion Stimulation)	• basic • abundant in exponential phase (20000-40000 dimers/cell) • DNA binding preference: GNYAWWWTRNC	homodimer, 2x11.2 kDa	*fis* (73.4 min)	• induces strong DNA curvature (up to 90°) • alters DNA topology • participates in site-specific recombination, transposition, and DNA replication
H-NS (Histone-like Nucleoid Structuring protein)	• non basic • 20000-40000 dimers/cell • binding form may be tetramer or higher oligomer • induced during cold-shock	homodimer, 2 x 15.4 kDa	*hns* (27.8 min)	• recognizes curved DNA • alters DNA topology • induces DNA curvature • influences recombination

Table 1. Properties and the abundance of major nucleoid-associated proteins of *E. coli*.

Protein	Molecular mass (kDa)	Native functional unit	Abundance[1] in growth phase	Abundance[1] in stationary phase
HUα and HUβ	~ 9	Homo- and hetero-dimer	55,000 (23)	30,000 (12.5)
IHFα and IHFβ	~ 11	Heterodimer	12,000 (5)	55,000 (23)
H-NS	~ 15	Homodimer	20,000 (8)	15,000 (6)
Fis	~ 11	Homodimer	60,000 (25)	Undetectable
Dps	~ 19	Dodecamer	6,000 (0.4)	180,000(12.5)

[1] Abundance (molecules/cell) data were taken from [16]. The number in the parenthesis is micromolar concentration calculated using the following formula: (number of native functional units/Avogadro number) x (1/cell volume in liter) x 10^3. Cell volume in liter (2 x 10^{-15}) was determined by assuming volume of the *E. coli* cell to be 2 μm^3.



Non solo curvatura… le diverse funzioni delle proteine associate al nucleoide sul DNA

A. Bend curvare
B. Stiffen irrigidire,
Coat rivestire
C. Bridge creare ponti
B. Bunch raggruppare
E. Wrap avvolgere

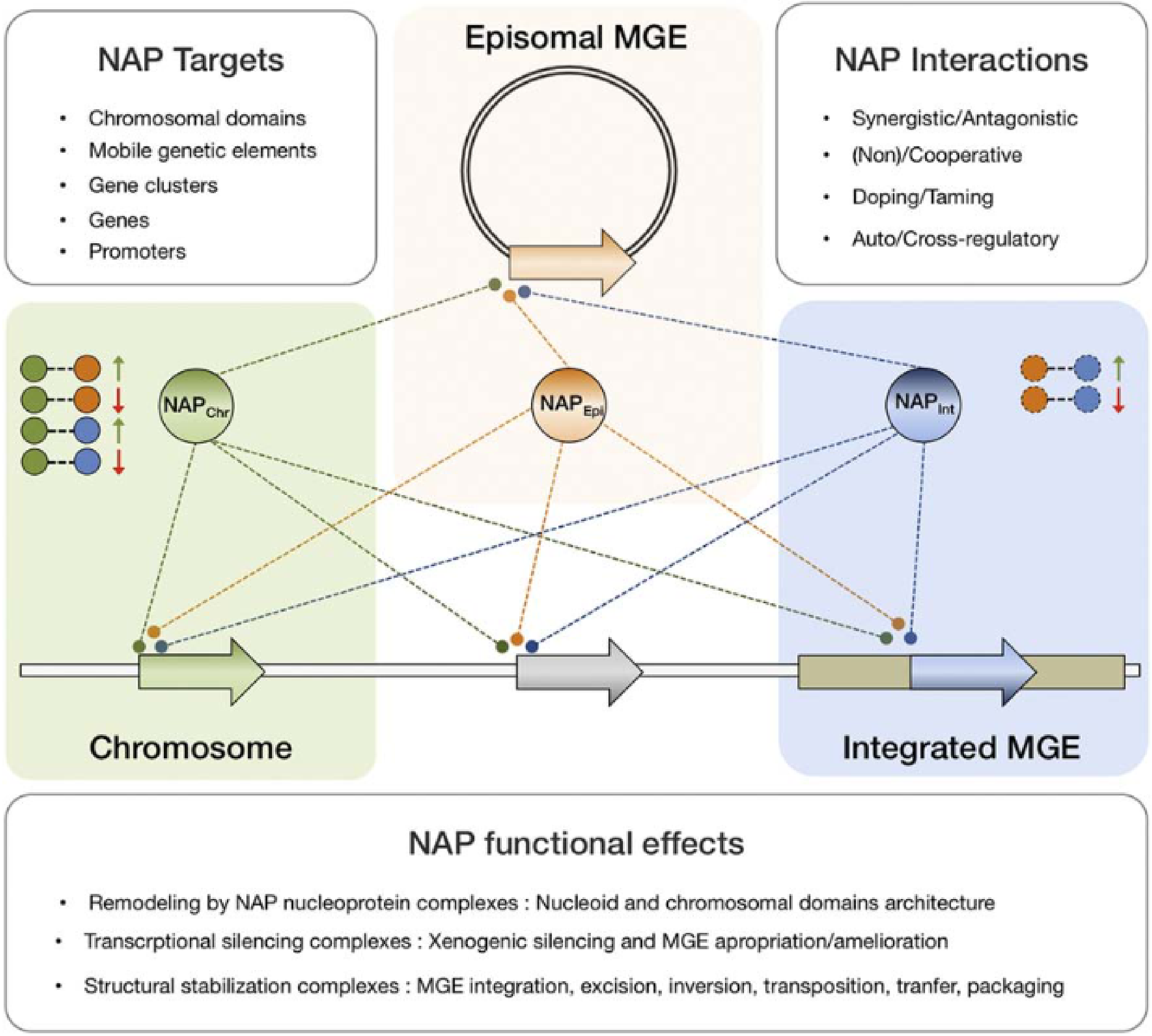


Fig. 2. Network of acknowledged interactions between endogenous and foreign NAPs occurring in bacteria. Endogenous NAPs are encoded chromosomally (NAP_{Chr}), whereas foreign NAPs (NAP_{Epi}; NAP_{Int}) are encoded in episomal (double line circle) or integrated (beige box) mobile genetic elements. The genes are represented as filled arrows and their cognate protein products as circles. The gene-protein pairs are colored according to their origin: green for endogenous; orange for episomal and blue for integrated. Protein-DNA interactions are represented by connecting lines in the main scheme, and protein-protein interactions are represented as connected circles in the upper corner of the green and blue boxes. The nature of the interaction is represented by positve (synergistic) or negative (antagonistic) symbols colored green or red, respectively. NAP proteins main targets, the types of interactions they establish and the functional outputs of those interactions are indicated in the accopmpaigning text boxes.

Characteristics of the main nucleoid-associated proteins in bacteria.

NAPs	Size (kDa)	Accession ID	Representative organism	Presence in MGEs	Acknowledged function
H-NS family proteins					
H-NS	15	P0ACF8	*Escherichia coli*	Y	Xenogeneic silencing. Nucleoid structuring
StpA	15	P0ACG1	*Escherichia coli*	ND	Functional analog of H-NS
MvaT	14	Q9HW86	*Pseudomonas aeruginosa*	Y	Functional analog of H-NS
Ler	14	A0A0H0PFT0	*Escherichia coli*	Y	Homologue and antagonist of H-NS. Activator of LEE
Hfp			*Escherichia coli*	Y	Functional analog of H-NS
BpH3	14	O07507	*Bordetella pertussis*	ND	Functional analog of H-NS. Essential for *B. pertussis*
Bv3F	13	A4JS72	*Burkholderia vietnamiensis*	ND	
HvrA	11	P42505	*Rhodobacter capsulatus*	ND	
Lsr2	12	P9WIP7	*Mycobacterium tuberculosis*	ND	Functional analog of H-NS
XrvA	15	Q56835	*Xanthomonas oryzae*	ND	
Rok	22	O34857	*Bacillus subtilis*	ND	Functional analog of H-NS. Repressor of ComK
HU/IHF family proteins					
HU	9	P0ACF0	*Escherichia coli*	Y	DNA replication, repair, recombination packaging
IHF	11	P0A6X7/P0A6Y1	*Escherichia coli*	Y	DNA transposition, recombination, plasmid replication
Fis family proteins					
Fis	11	P0A6R3	*Escherichia coli*	ND	Gene regulation, nucleoid architecture, DNA remodeling
Other					
Lrp	19	P0ACJ0	*Escherichia coli*	Y	Gene regulation
EbfC	11	O51418	*Borrelia burgdorferi*	ND	Gene regulation
NdpA	37	A0A024L1K9	*Escherichia coli*	Y	

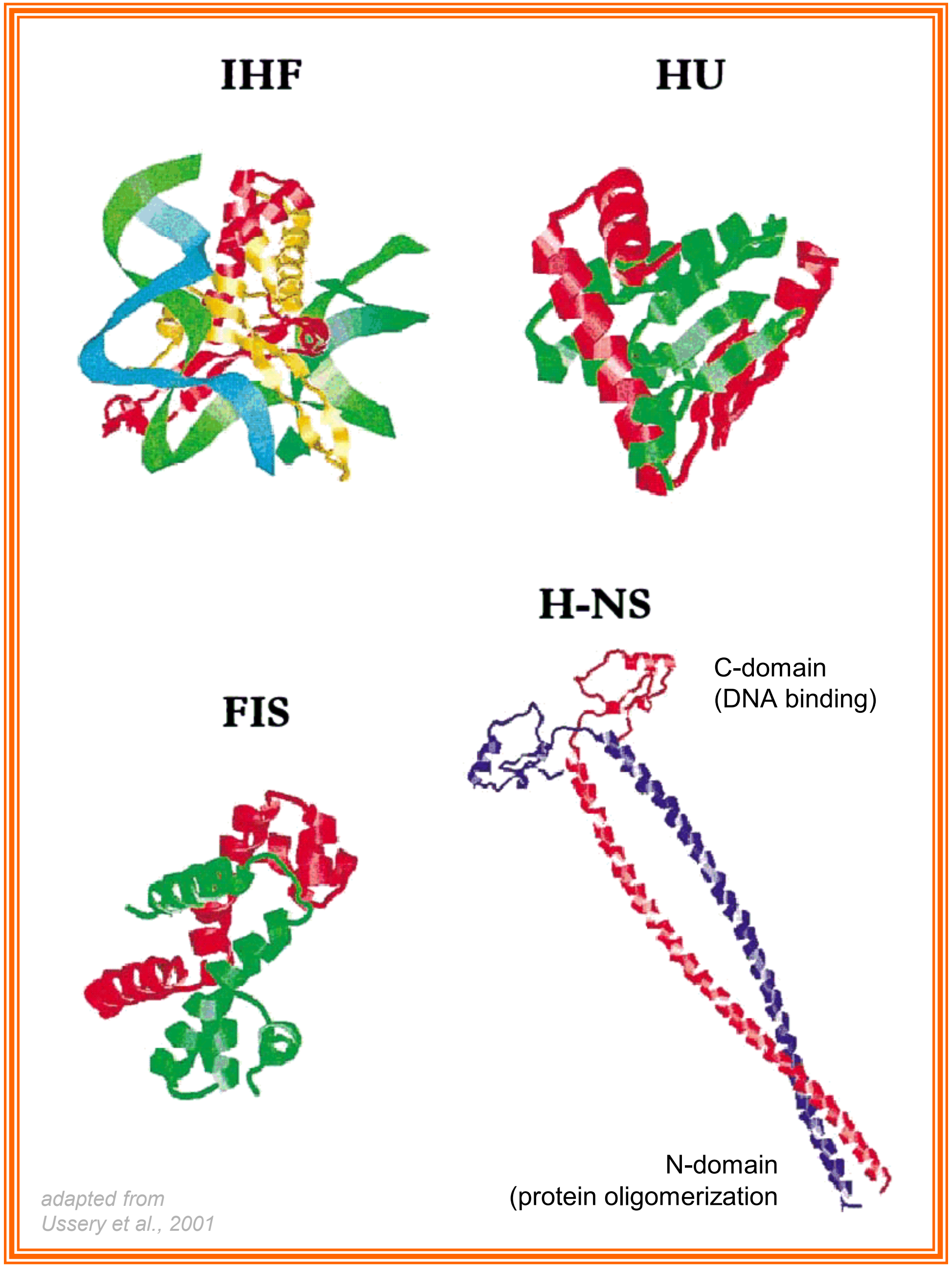
La proteina HU

Caratteristiche

- proteina basica
- molto abbondante 30000 copie/cellula
- la più abbondante tra le proteine del nucleoide
- nessuna sequenza consenso di legame al DNA

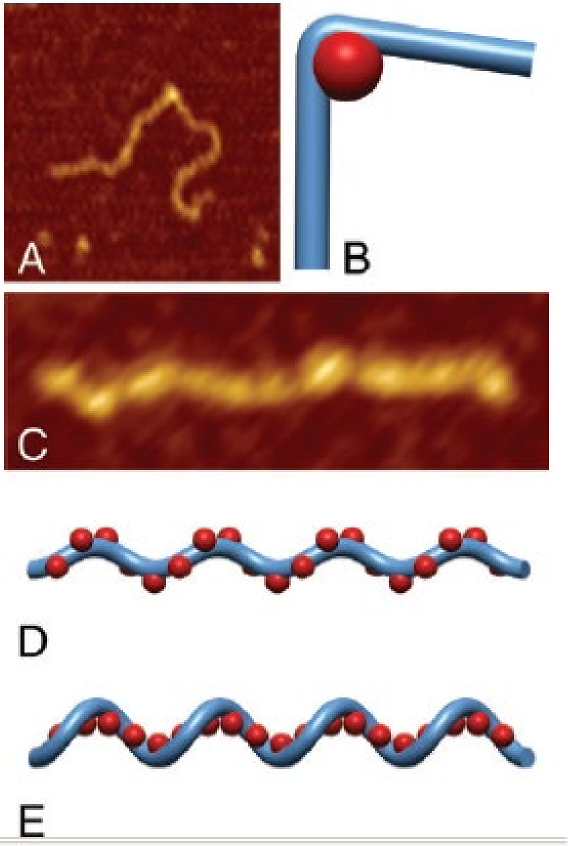
Struttura	pM	geni
eterodimero		
Hua	9.2 kDa	hupA (90.4 min)
Hub	9.5 kDa	hupB (9.9 min)

Nucleoid proteins are homo- or heterodimers

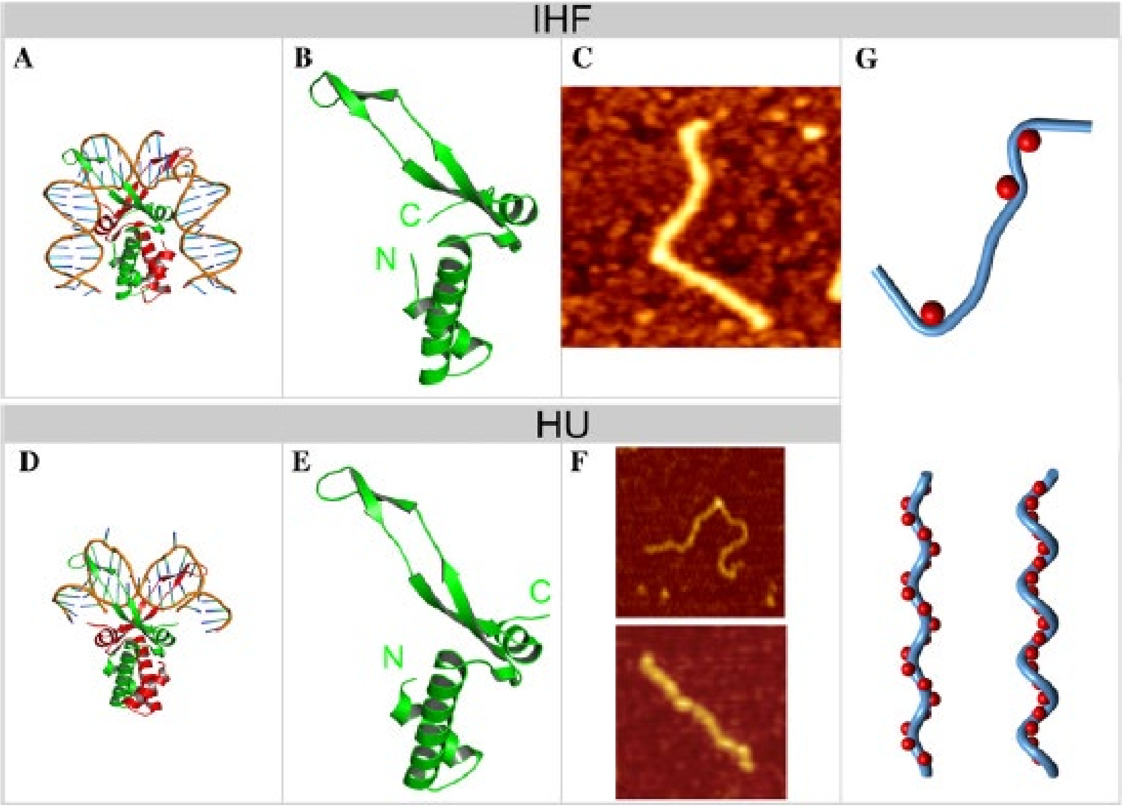


adapted from
Ussery et al., 2001

HU lega il DNA e lo ripiega



Le proteine che ripiegano il DNA

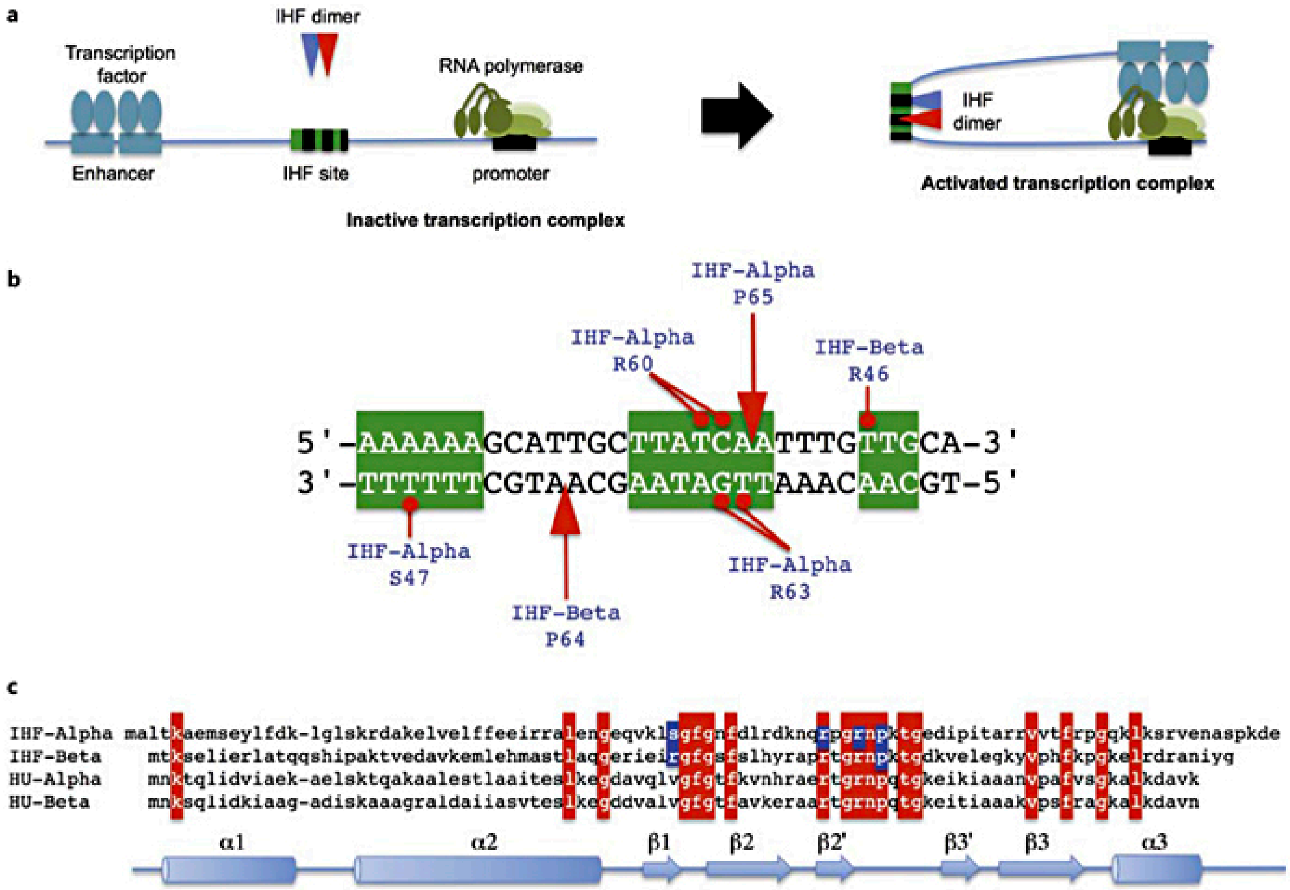


La proteina IHF   Integration Host factor

Caratteristiche
- proteina basica
- 5-10 volte meno abbondante di HU
- abbondante in fase stazionaria
- debole specificità di sequenza per il legame al DNA (YAANNNNTTGATW)

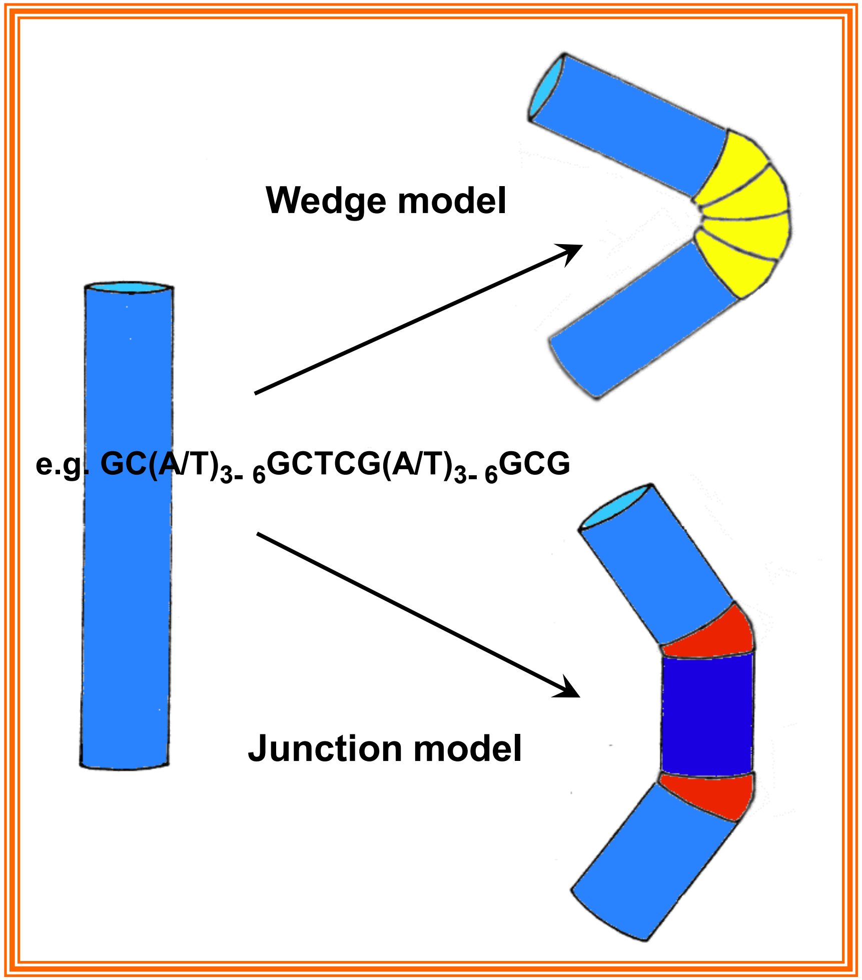
Struttura	pM	geni
Eterodimero		
IHFα	11.2 kDa	*himA* (38.6 min)
IHFβ	10.5 kDa	*himD* (25 min)

From: **Function of Nucleoid-Associated Proteins in Chromosome Structuring and Transcriptional Regulation**
Journal of Molecular Microbiology and Biotechnology. 2015;24(5-6):316-331. doi:10.1159/000368850



IHF and its paralogue HU. **a** The interaction of IHF with its target site and the consequences for the pathway of the DNA are shown. Here, RNA polymerase (containing σ54) has formed an inactive complex with a promoter, and the bending of the DNA by IHF causes a transcription factor, bound as two dimers to two copies of the upstream-located enhancer sequence, to make physical contact with RNA polymerase, activating transcription. Not to scale. **b** The details of the IHF site sequence are shown, with the highlighted residues being the conserved members of the IHF binding site consensus. The amino acids in the α- and the β-subunits of IHF that interact with the DNA sequence are shown. In the cases of proline residues P65 (α-subunit) and P64 (β-subunit), the protein makes an insertion into the minor groove of the DNA duplex, bending it by up to 180° . **c** An alignment of the α- and β-subunits of the paralogous IHF and HU proteins from *E. coli* strain W3110 is shown together with a summary of the main structural features of each monomer. Amino acids that are completely conserved in all four proteins are highlighted..

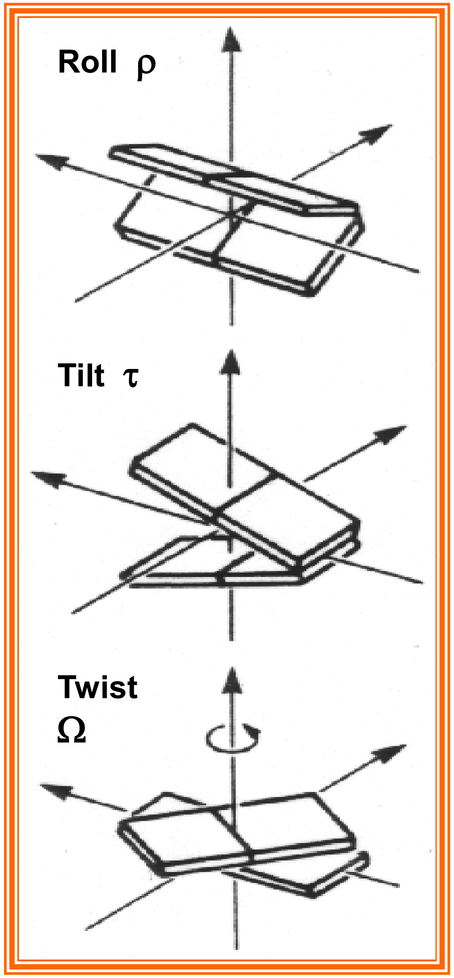
Intrinsically curved DNA



Wedge model

e.g. GC(A/T)$_{3-6}$GCTCG(A/T)$_{3-6}$GCG

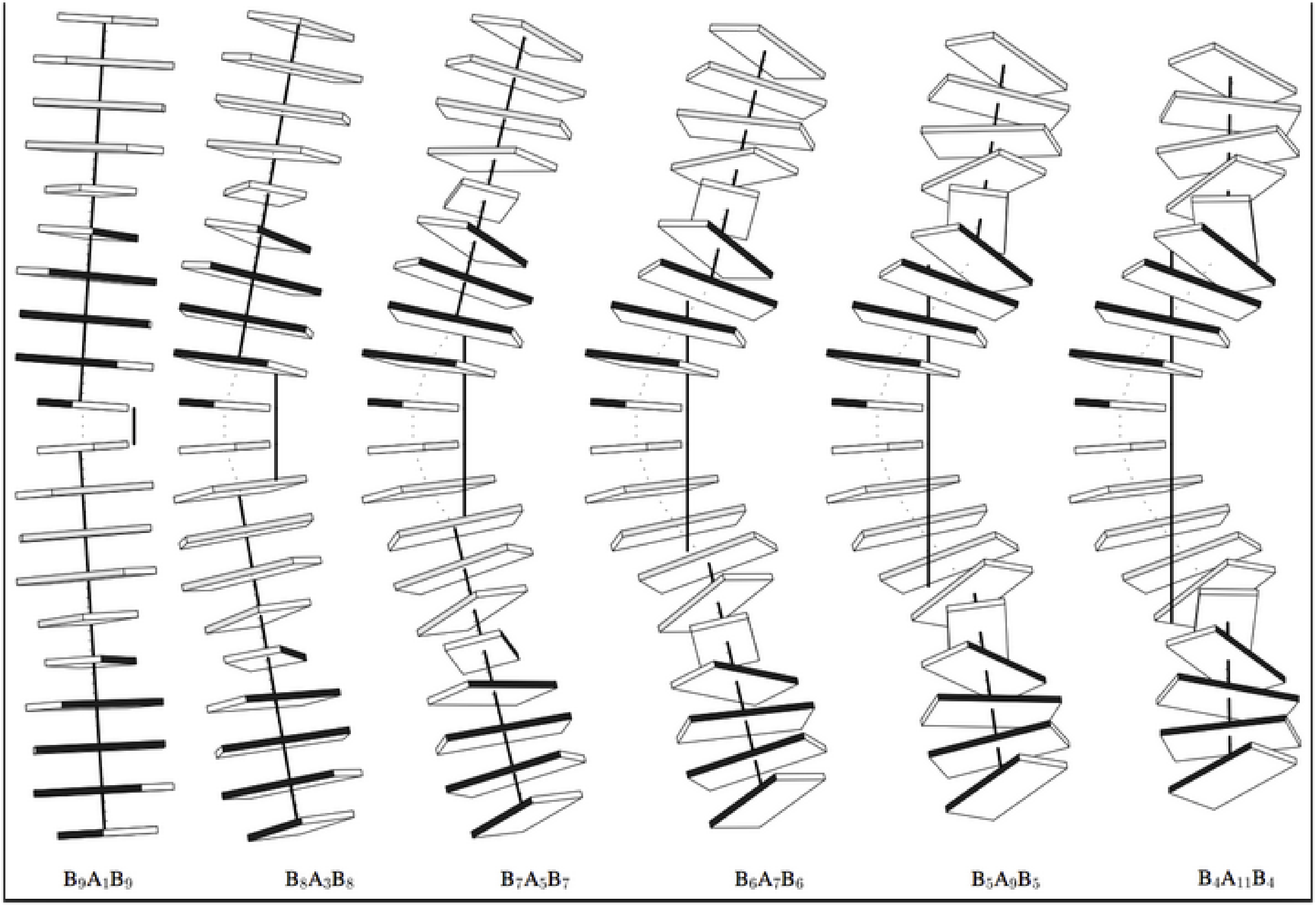
Junction model



Roll ρ

Tilt τ

Twist Ω


$B_9A_1B_9$
$B_8A_3B_8$
$B_7A_5B_7$
$B_6A_7B_6$
$B_5A_9B_5$
$B_4A_{11}B_4$

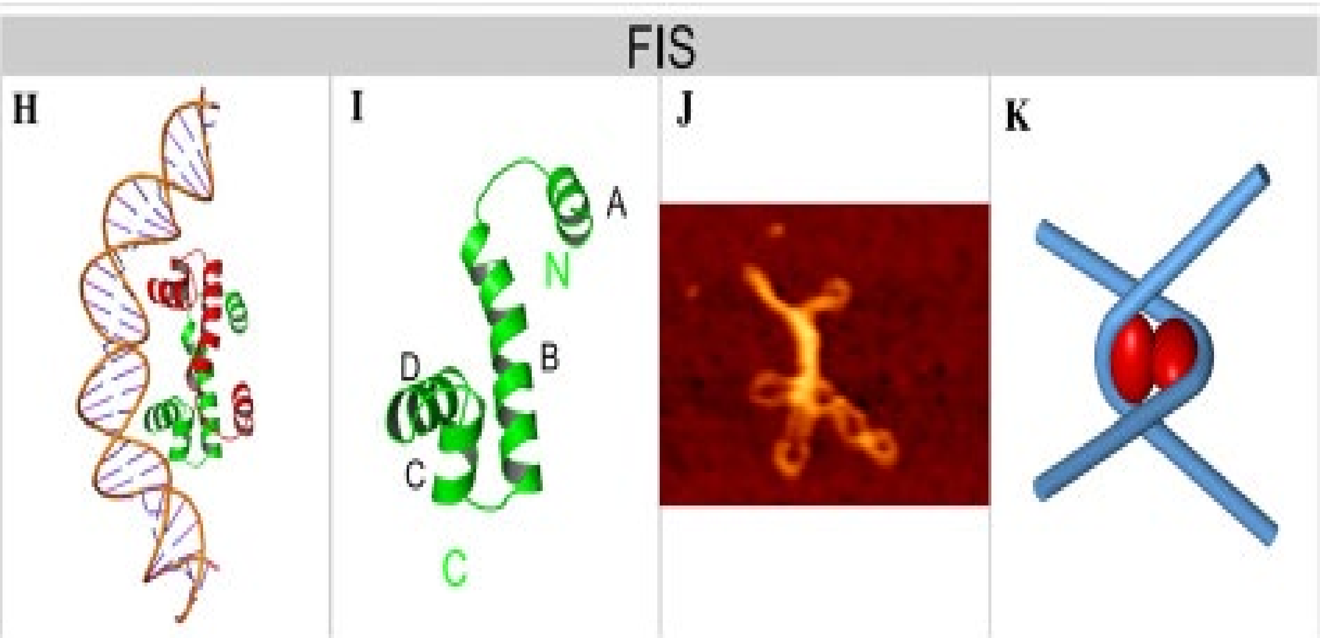
FIS Factor for Inversion Stimulation

Caratteristiche

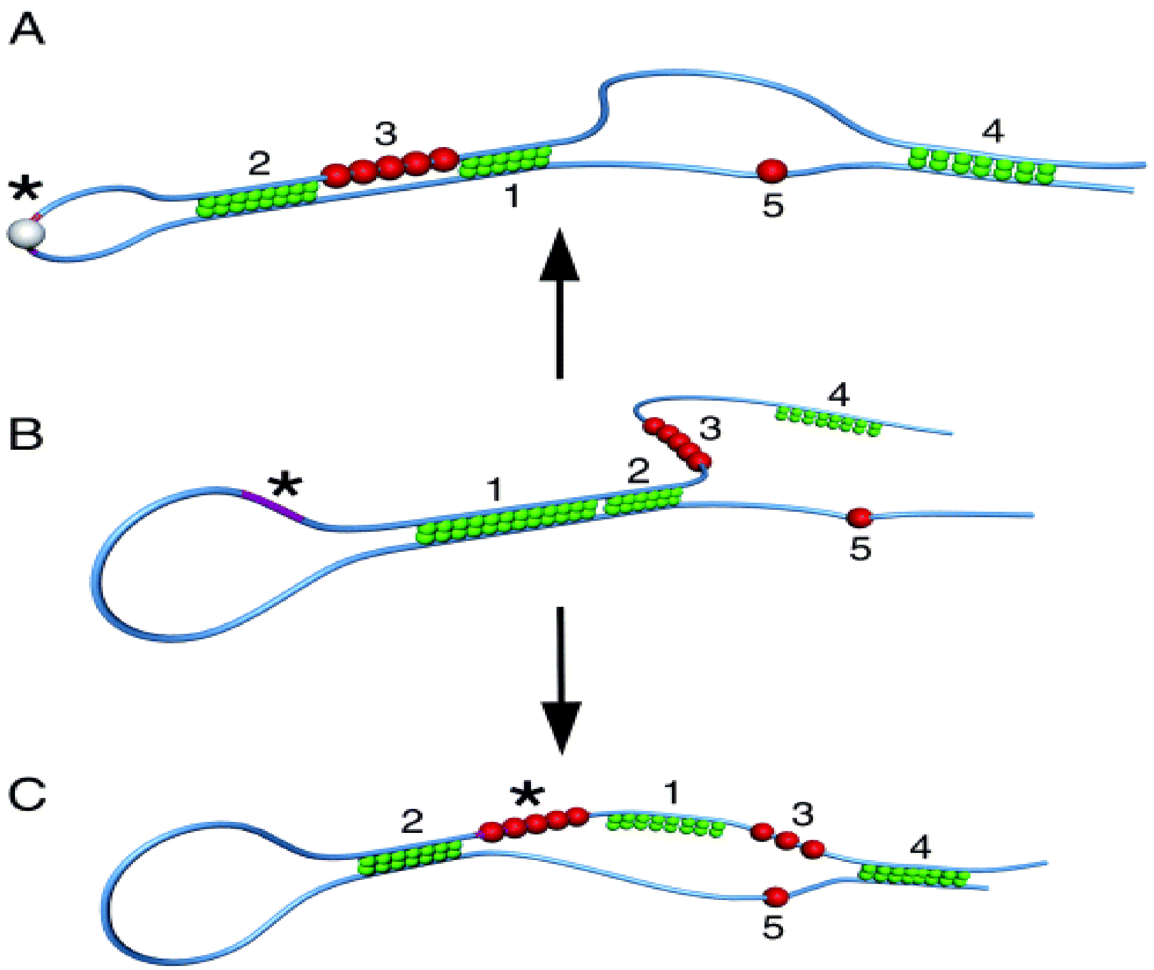
- Proteina basica
- abbondante in fase esponenziale
- 10.000-60.000 copie
- scarsa specificità di sequenza (KNNYRNNWNNYRNNM)

W	**TA**
R	**GA**
K	**GT**
Y	**CT**

Struttura	pM	geni
OMODIMERO		
FIS	2x 11.5 kDa	fis (27.4 min)


FIS
H
I
A
N
B
D
C
C
J
K

The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin



Molecular Microbiology, Volume: 56, Issue: 4, Pages: 858-870, First published: 07 April 2005, DOI: (10.1111/j.1365-2958.2005.04598.x)

The nucleoid is very dynamic in nature. The organization of the supercoiled loops and the relative orientation of sites within are affected by the binding of nucleoid-associated proteins such as HU, H-NS, IHF and Fis. Each of these proteins has different functional interactions with DNA and the proteins of this family can be roughly fitted into two categories: DNA-bridgers and DNA-benders (see main text). Within the context of a supercoiled loop, the configuration is determined by the proteins that are simultaneously bound within the loop. The organization within loops is important for DNA compaction and can also play a role in the modulation of transcription. DNA is depicted in cyan. The green spheres correspond to H-NS (following the mechanism of binding as shown in Fig. 1B). The red spheres correspond to HU (either in its bending or its rigidification mode). The grey sphere corresponds to a DNA bending protein, such as IHF, Fis or HU. Although the described type of reorganization can take place only within a supercoiled loop, DNA topology has been omitted in the visualization of these loops for the purpose of simplicity.

The (arbitrary) starting configuration of a loop is shown in B. H-NS bridges hold together DNA tracts within the loop, some HU molecules are shown rigidifying the DNA, or bending it at another site and one tract of DNA is covered with H-NS but *not* bridged. A binding site is present within the tip of the loop (in purple, and indicated with an asterisk). The binding of proteins at this site results in reconfiguration of the loop. In the first example (A), this site is bound by a protein that bends DNA (e.g. IHF), which imposes relocation of the bound protein to the tip of the loop. In parallel, other bound proteins within the loop will become relocated. In this example, one of the H-NS tracts (2) and the adjacent region with HU bound (3) follow the left-ward directed movement of the tip of the loop. A part of a second H-NS tract (1) is found back bound on the right side of the region with HU. Finally, the spatial vicinity of the loose H-NS tract (4) has changed such that it has become bound to the DNA alongside the DNA bending HU molecule (5). In the second example (C), HU becomes co-operatively bound to the binding site (or rather region). It is energetically unfavourable for a rigid tract of DNA to find itself located within the tip of a loop and therefore it moves away into the stem. The H-NS tract directly aside of the bound HU (1) moves along with it to the right (where it remains loose), whereas the second H-NS tract (2) is found back bridging an area on the left of the region with HU bound (3). Finally, the spatial vicinity of an initially loose H-NS tract (4) has also changed such that it has become bound to the DNA alongside the DNA bending HU molecule (5).

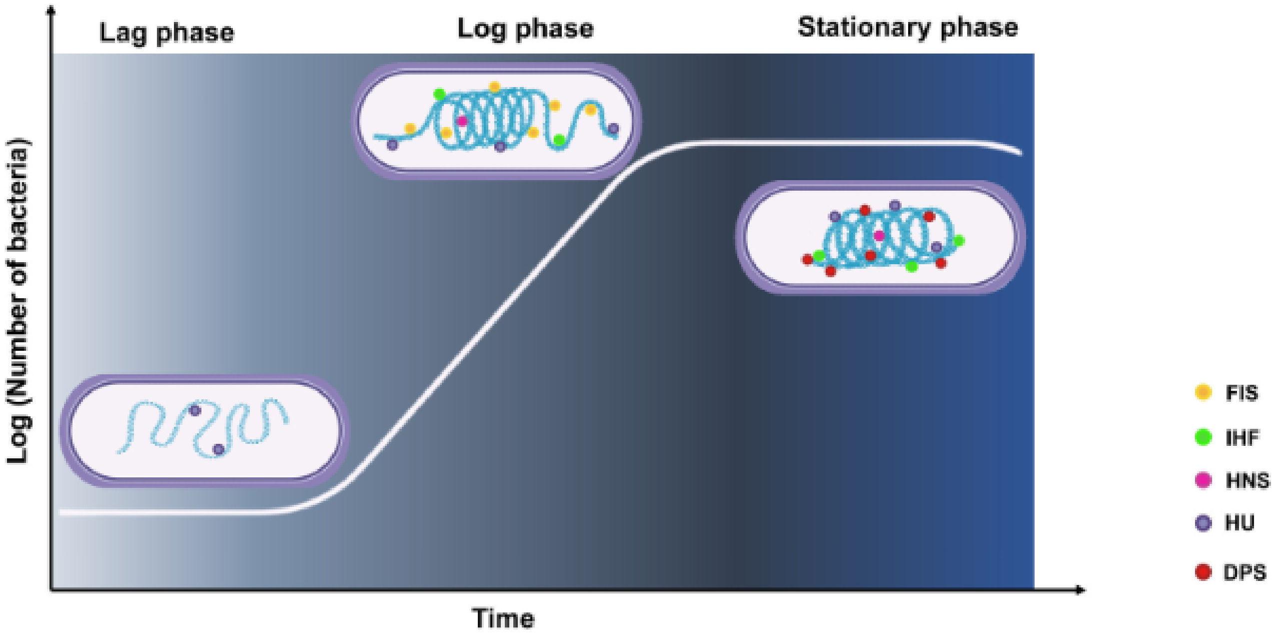
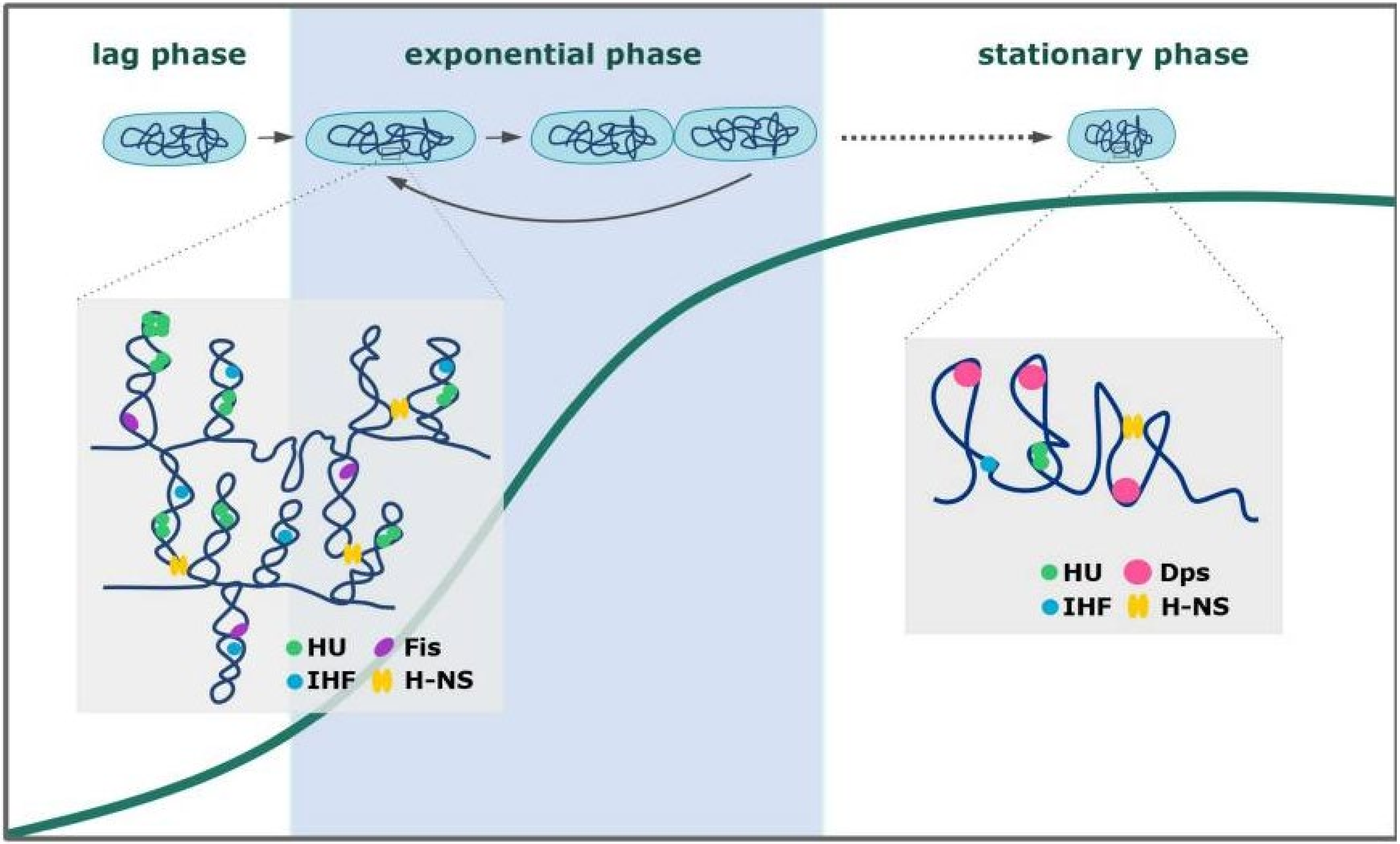


Fig. 3. : Association of bacterial growth with NAPs: As the cell goes from nutrient-sufficient lag to nutrient-deficient stationary growth phase different NAPs bind DNA at different phases. FIS (yellow) is the most abundant protein during the early log phase and is later joined by the rest of the NAPs during the log phase, with DPS (red) being the most dominant protein during the stationary phase (Ali Azam et al., 1999).

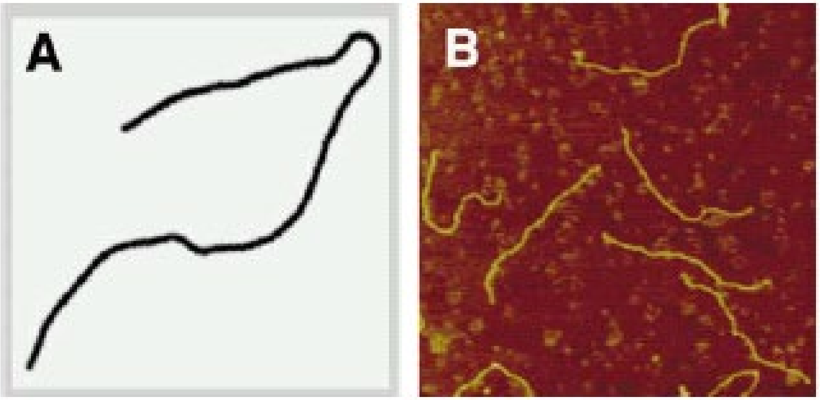


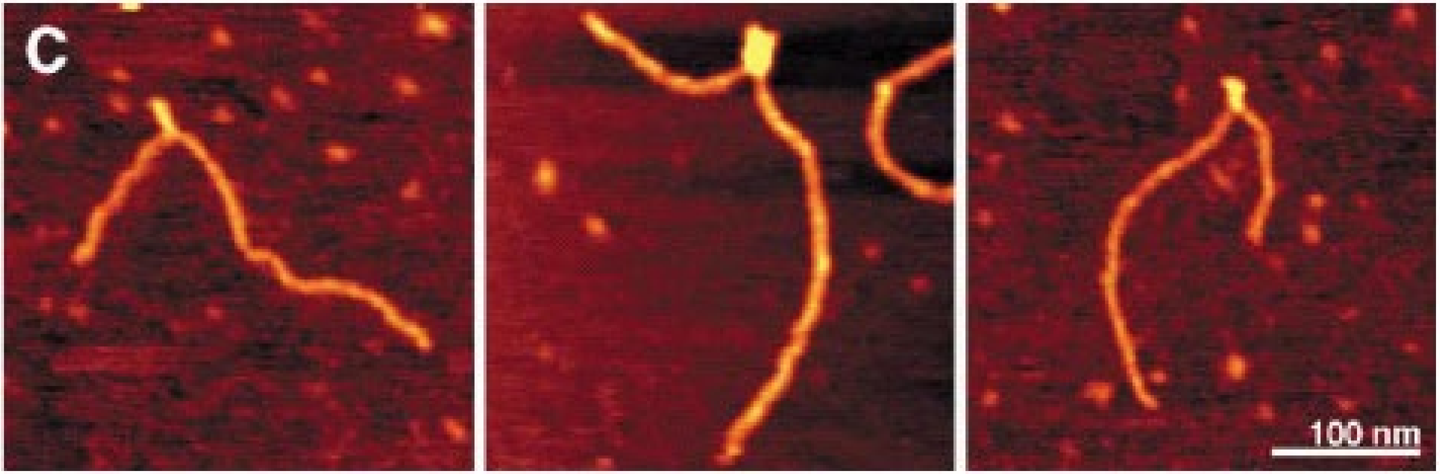
Chromosome organization during the growth of *Escherichia coli*. The expression patterns of *E. coli* NAPs reflect the chromosome compaction level (higher in the stationary than in the exponential phase) and cellular processes that involve certain NAPs

H-NS
(Histone-like Nucleoid Structuring protein)

- Abundant peptide (~ 20000 copies/cell)
- Small (136 aminoacids, 15.5 kD), non-basic peptide
- Mainly acts as a homodimer or tetramer
- Able to form heterodimers with StpA or HhA
- *In vitro* binding to DNA is non-specific; induces high DNA compaction
- Higher affinity for intrinsically curved DNA; able to bend DNA *in vitro*
- Global regulator: controls 5 % of the whole *E.coli* protein coding sequences
- Generally acts as a transcriptional repressor of virulence genes outside the host

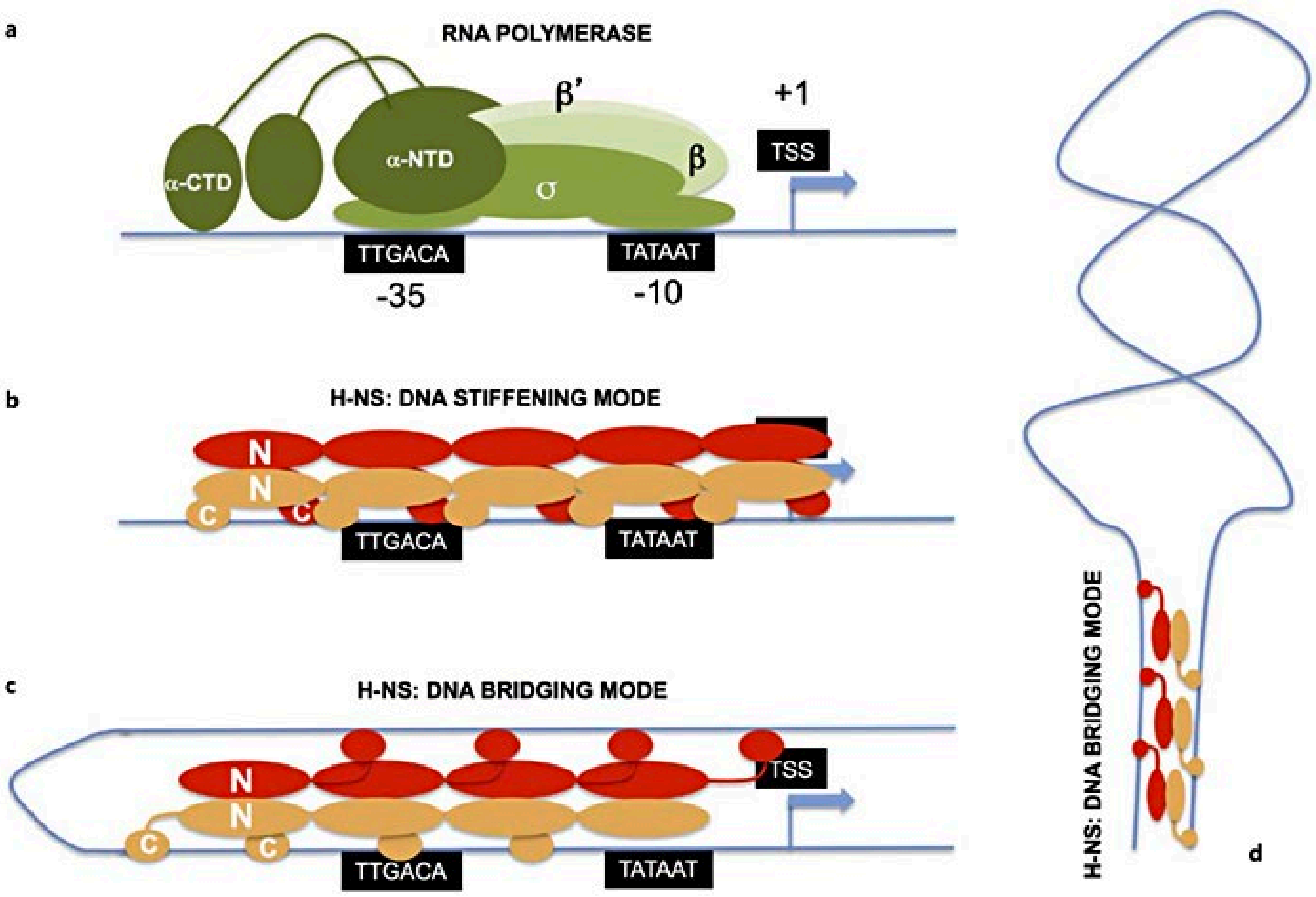
La proteina H-NS riconosce sequenze di DNA curvo ed è in grado di indurre curvatura nel DNA





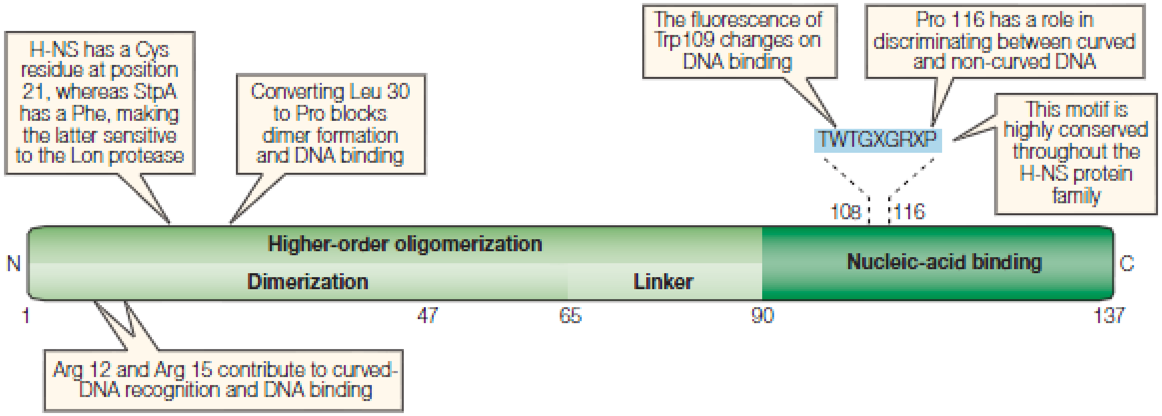
From: **Function of Nucleoid-Associated Proteins in Chromosome Structuring and Transcriptional Regulation**

Journal of Molecular Microbiology and Biotechnology. 2015;24(5-6):316-331. doi:10.1159/000368850



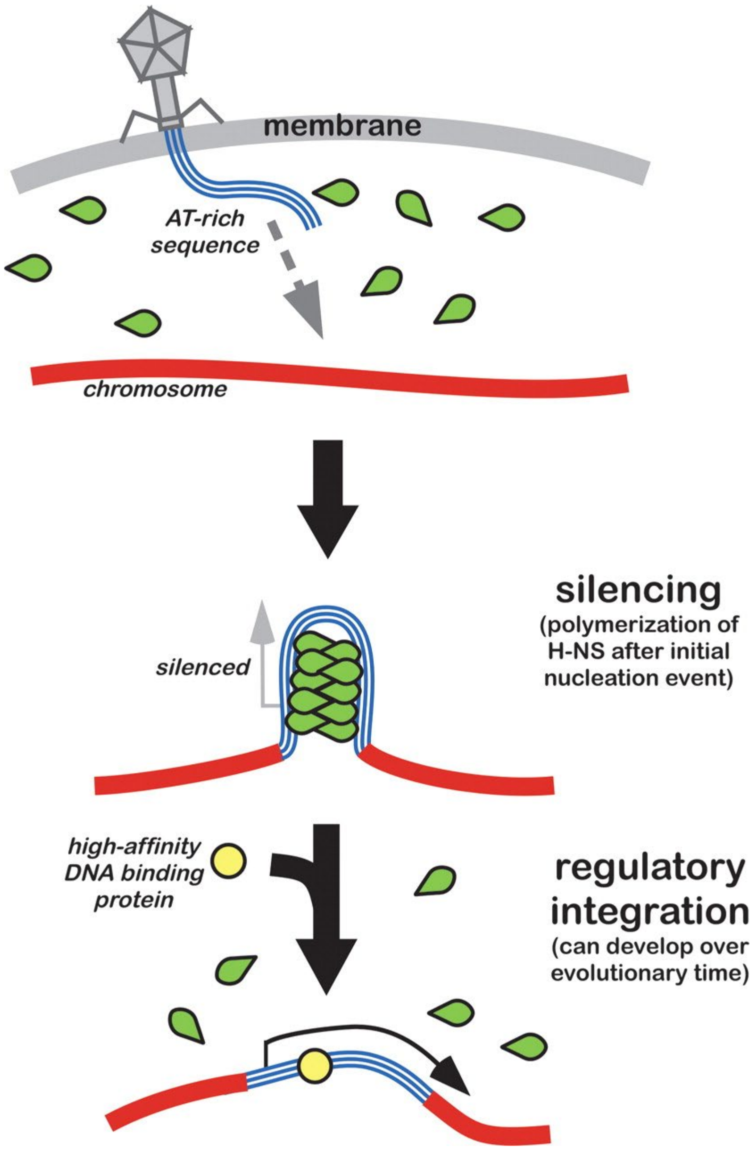
H-NS, transcription silencing and chromosome microdomain formation. **a** A standard representation of RNA polymerase bound to a transcription promoter is shown consisting of the principal components of the holoenzyme: the α-subunit, in two copies with their carboxyl-terminal domains (CTD) and amino terminal domains (NTD) shown connected by flexible linkers. The β-, β′- and σ-subunits are also illustrated. The locations of the transcription start site (TSS, +1), the -10 and -35 elements are also shown together with the consensus DNA sequences for the -10 and -35 motifs of promoters that are bound by the RpoD sigma factor of RNA polymerase. **b** The same promoter sequence is shown decorated by the H-NS protein in its DNA stiffening mode, excluding RNA polymerase and silencing transcription. Here H-NS polymerizes along the DNA duplex, and the two DNA-binding motifs of each H-NS dimer bind to the same DNA molecule in *cis*. The H-NS monomers are arranged in an antiparallel orientation within each dimer. **c** H-NS is shown bound to the same promoter element in its bridging mode. Here, the DNA-binding domains of each H-NS dimer (shown in antiparallel configuration) bind to spatially widely separated segments of the same DNA molecule, creating a DNA-protein-DNA bridge that excludes RNA polymerase from the promoter. **d** The bridging function of H-NS can also form loops in DNA, including the 10- to 15-kb microdomain loops that contribute to the higher-order structure of the bacterial nucleoid. The drawings are not to scale.

Struttura della proteina H-NS



La proteina H-NS è costituita da soli 137 AA
Il dominio di oligomerizzazione è localizzato al N terminale ed è costituito da brevi sequenze di AA (1-8, 12-19, e 23-47) capaci di formare 3 strutture ad alfa elica. I linker flesibili che separano le 3 a eliche permettono alle eliche 1 e 2 di ripiegarsi facilitando la formazione di oligomeri tra i diversi dimeri

H-NS ed il silenziamento di regione geniche acquisite per HGT



Il legame di H-NS a regioni di DNA esogeno ricche in AT silenzia l'espressione genica.

L'eventuale presenza di una proteina regolatrice sequenza specifica con un elevata affinità per il DNA può competere con H-NS per eliminare il silenziamento e permettere nuovamente l'espressione genica in condizioni specifiche.

In questo modo, la cellula ospite può tollerare la presenza di sequenze di DNA estraneo e in seguito inserire la sequenza in un network di regolazione preesistente

H-NS e il silenziamento dei geni acquisiti tramite HGT

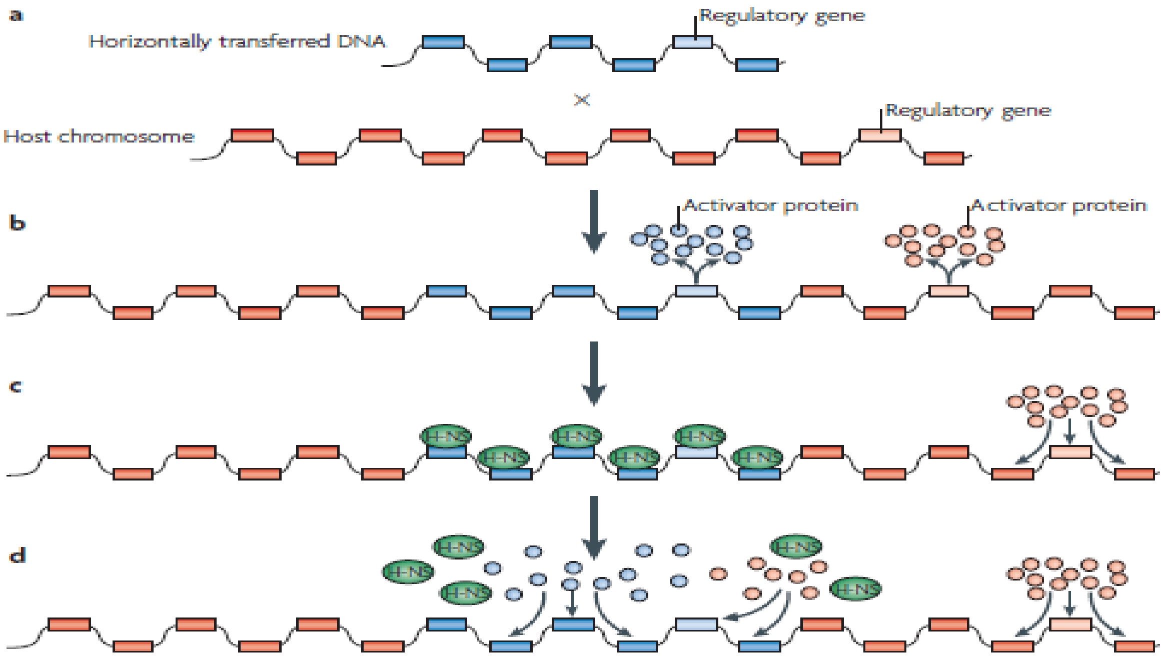
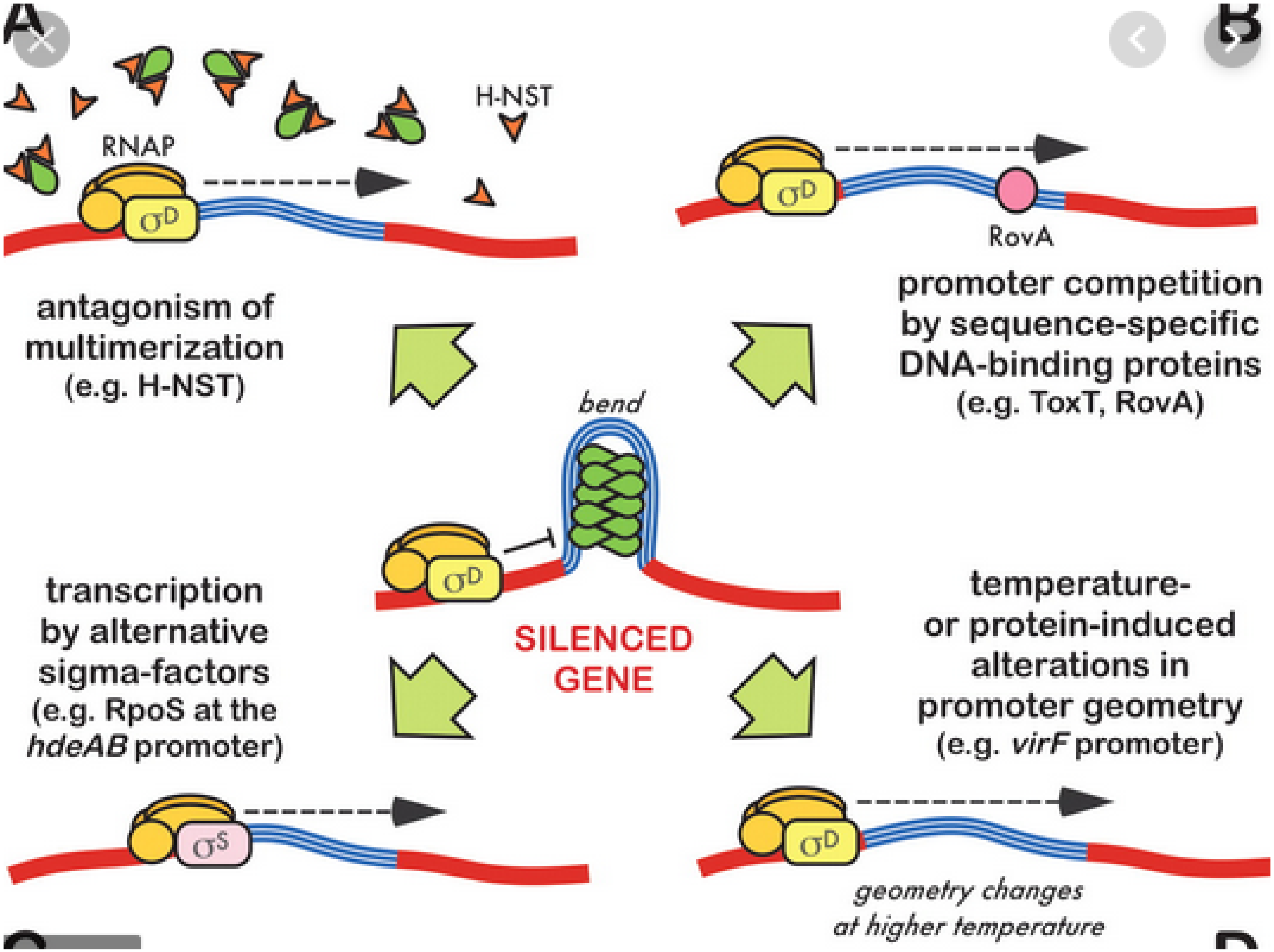
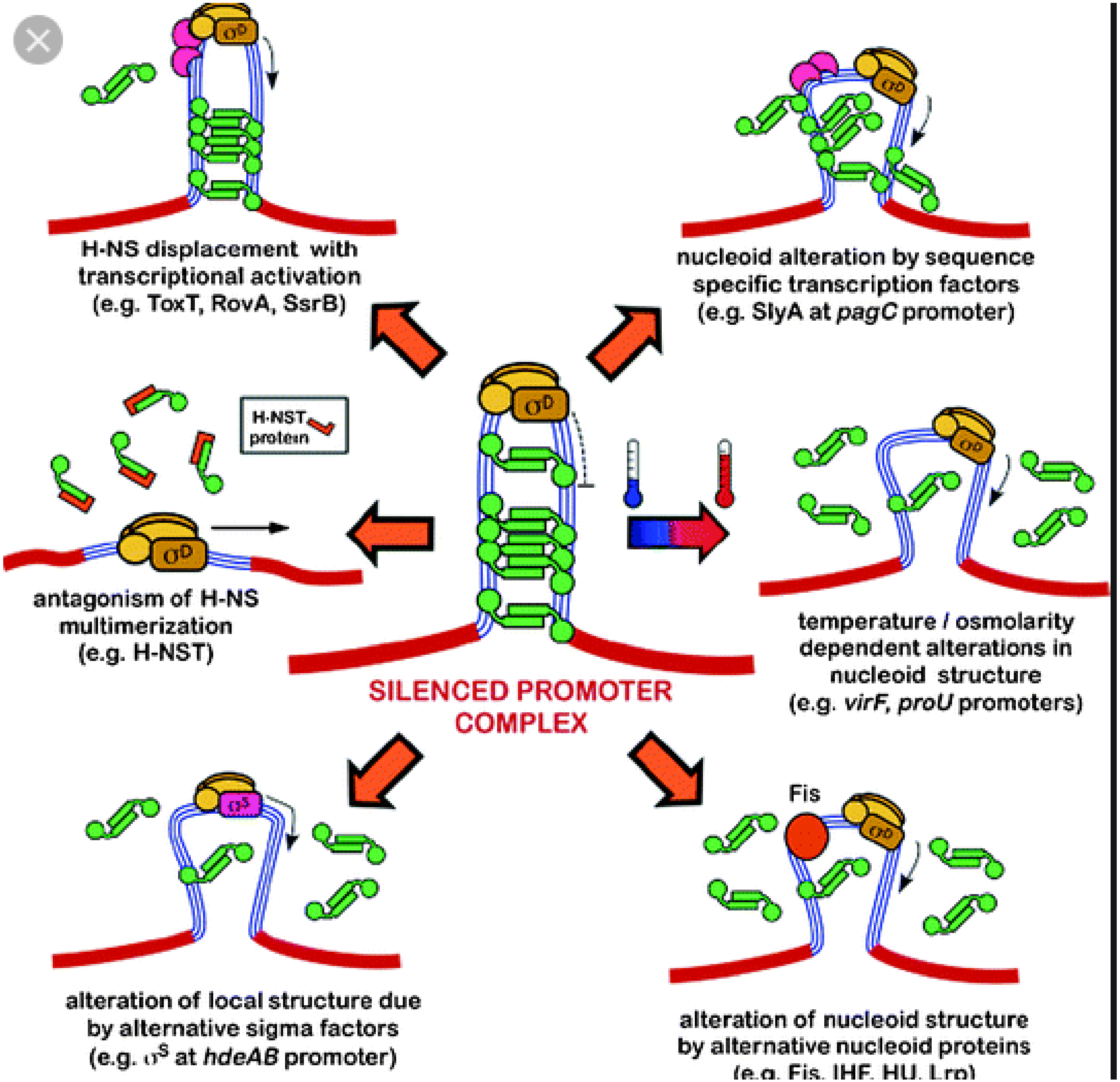


Figure 1 | **H-NS as a gene silencer.** The figure shows a model for the involvement of H-NS in the repression of horizontally transferred genes, and how this repression can be relieved. **a** | A segment of DNA carrying six genes enters the cell as a result of horizontal transfer. One of these is a regulatory gene that encodes a transcription activator that is specific for its own gene and the other five genes in the cluster. The newly arrived genes integrate into the chromosome of the host bacterium (shown in red). **b** | Once inserted, the horizontally transferred genes and the ancestral genes coexist as a contiguous DNA sequence that is distinguished by the higher AT content of the insertion (blue). **c** | The H-NS protein quickly targets and downregulates the promoters of the genes with high AT sequences. **d** | This transcription repression can be relieved in numerous ways. Changes to DNA structure, particularly the planar curvature, induced by environmental signals, such as an increase in temperature, might dislodge H-NS. The activator protein encoded by the horizontally transferred regulatory gene (blue) might displace H-NS by remodelling the DNA near its binding sites through structural perturbations, such as DNA bending. An activator encoded by a regulatory gene in the ancestral chromosome (red) might displace H-NS by the same mechanism. A regulatory relationship between the ancestral activator and the new DNA sequences could arise by different routes: (i) suitably positioned sites for activator binding might fortuitously already exist in the horizontally transferred genes; (ii) the activator protein might evolve to bind to appropriately positioned sites; (iii) sites might evolve in the horizontally transferred DNA to which the ancestral activator can now bind; (iv) or some combination of these scenarios might apply.


A
H-NST
RNAP
σD
antagonism of multimerization (e.g. H-NST)
B
σD
RovA
promoter competition by sequence-specific DNA-binding proteins (e.g. ToxT, RovA)
bend
σD
SILENCED GENE
transcription by alternative sigma-factors (e.g. RpoS at the *hdeAB* promoter)
σS
temperature- or protein-induced alterations in promoter geometry (e.g. *virF* promoter)
σD
geometry changes at higher temperature



H-NS displacement with transcriptional activation (e.g. ToxT, RovA, SsrB)

nucleoid alteration by sequence specific transcription factors (e.g. SlyA at *pagC* promoter)

antagonism of H-NS multimerization (e.g. H-NST)

SILENCED PROMOTER COMPLEX

temperature / osmolarity dependent alterations in nucleoid structure (e.g. *virF*, *proU* promoters)

alteration of local structure due by alternative sigma factors (e.g. σ^S at *hdeAB* promoter)

alteration of nucleoid structure by alternative nucleoid proteins (e.g. Fis, IHF, HU, Lrp)

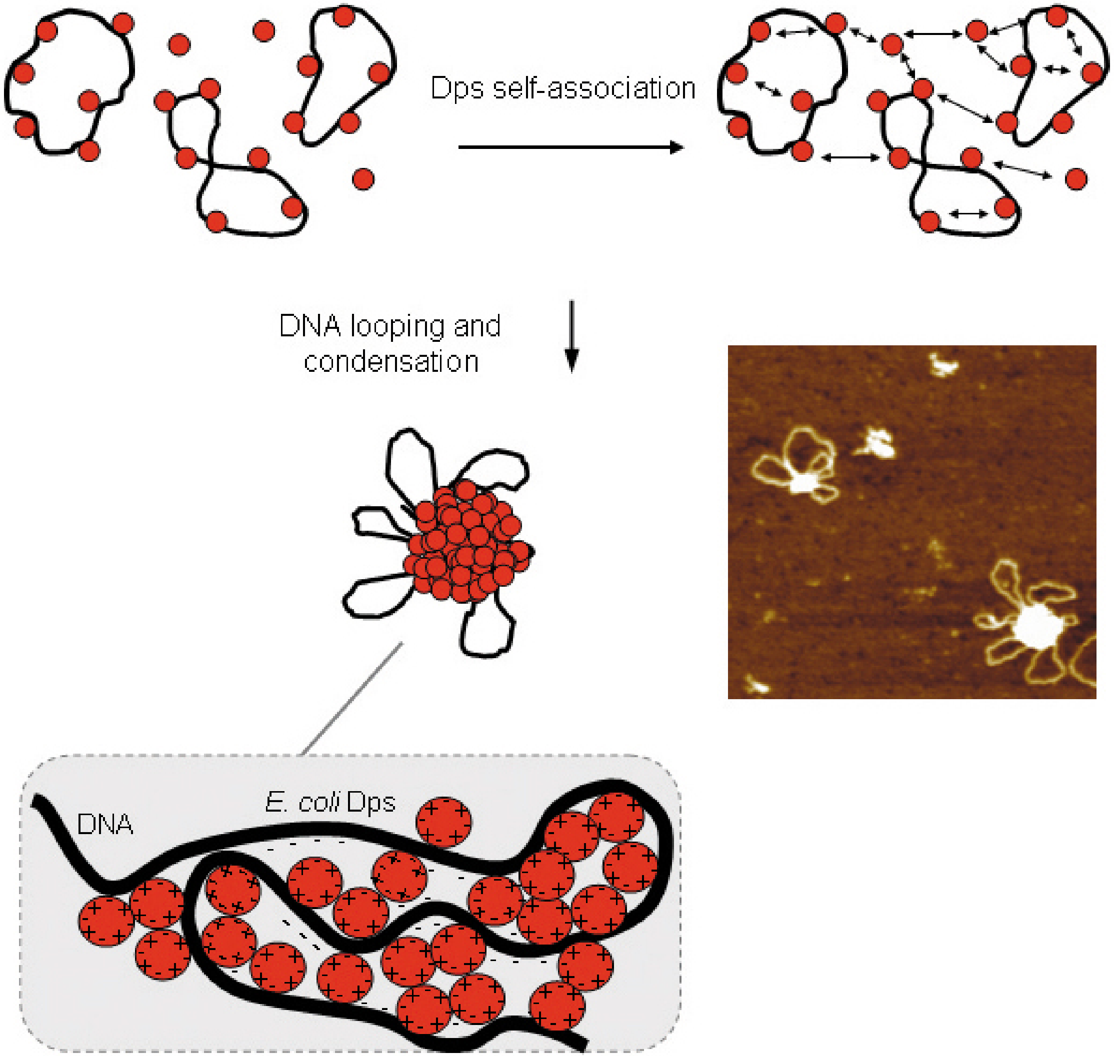
La proteina Dps (Dna Binding protein from starved cells)

forma un complesso costituito da 12 monomeri di 19 KDa

Il complesso Dps contiene uno ione Fe e rende il DNA resistente allo stress ossidativo

E' presente in alto numero di copie circa 20.000

Si lega al DNA a livello di sequenze non specifiche



I residui di lisina localizzati all’estremità N terminale di Dps carichi positivamente promuovono la condensazione del DNA in quanto interagiscono sia con il DNA che con le regioni cariche negativamente delle molecole adiacenti di Dps

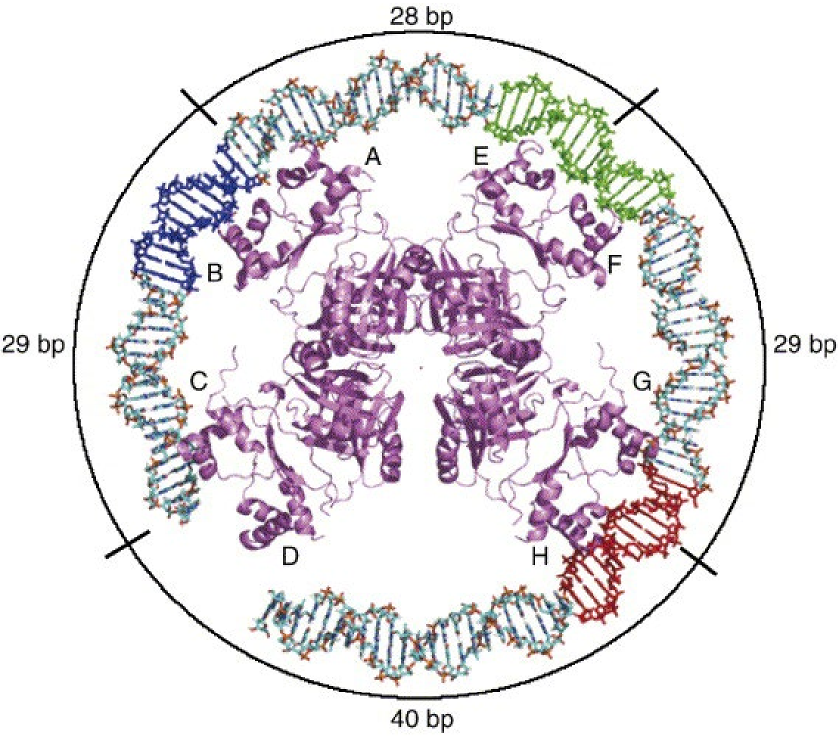
Lrp Leucine responsive regulatory protein

influenza ala trascrizione del 10% dei geni di E.coli e a seconda del target il suo effetto può essere potenziato o meno dalla presenza di leucina.

I geni regolati comprendono geni coinvolti nell'acquisizione e metablismo degli AA oltre a geni di virulenza quali quelli coinvolti nella sintesi di alcuni pili.

Lrp riconosce una sequenza consenso degenerata sul DNA e modifica la struttura del DNA con il suo legame.

Esiste in diversi stati oligomerici , dimero, ottamero o esadecamero.



Ottamero di LRP

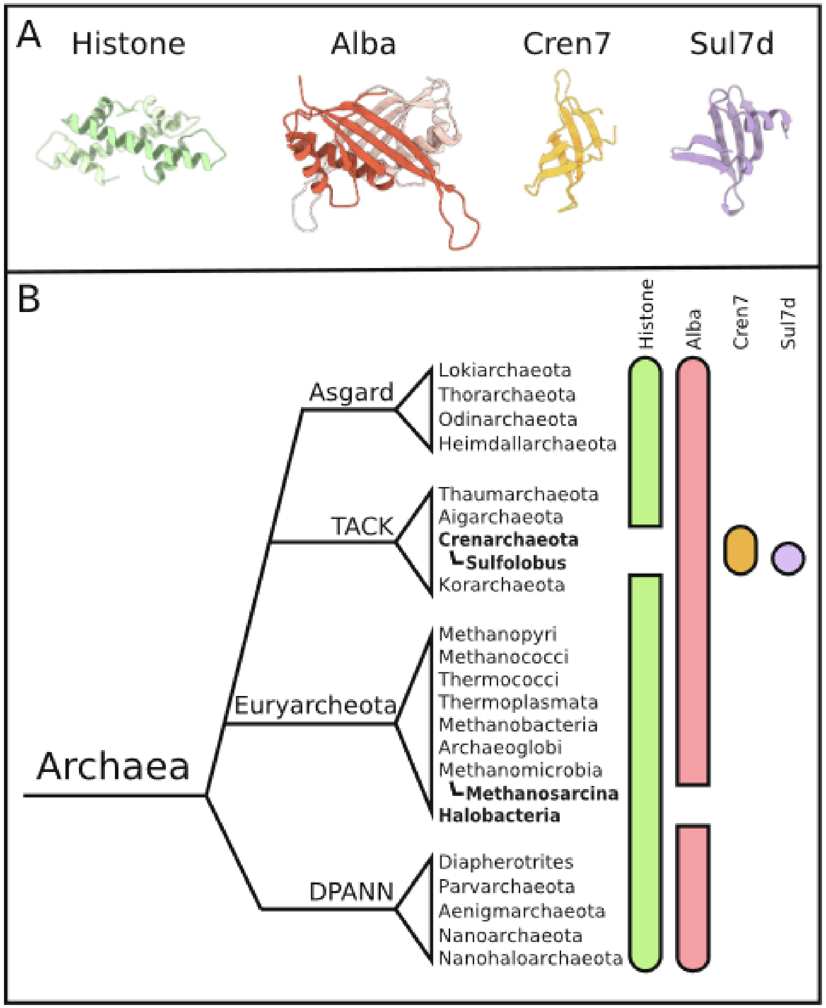
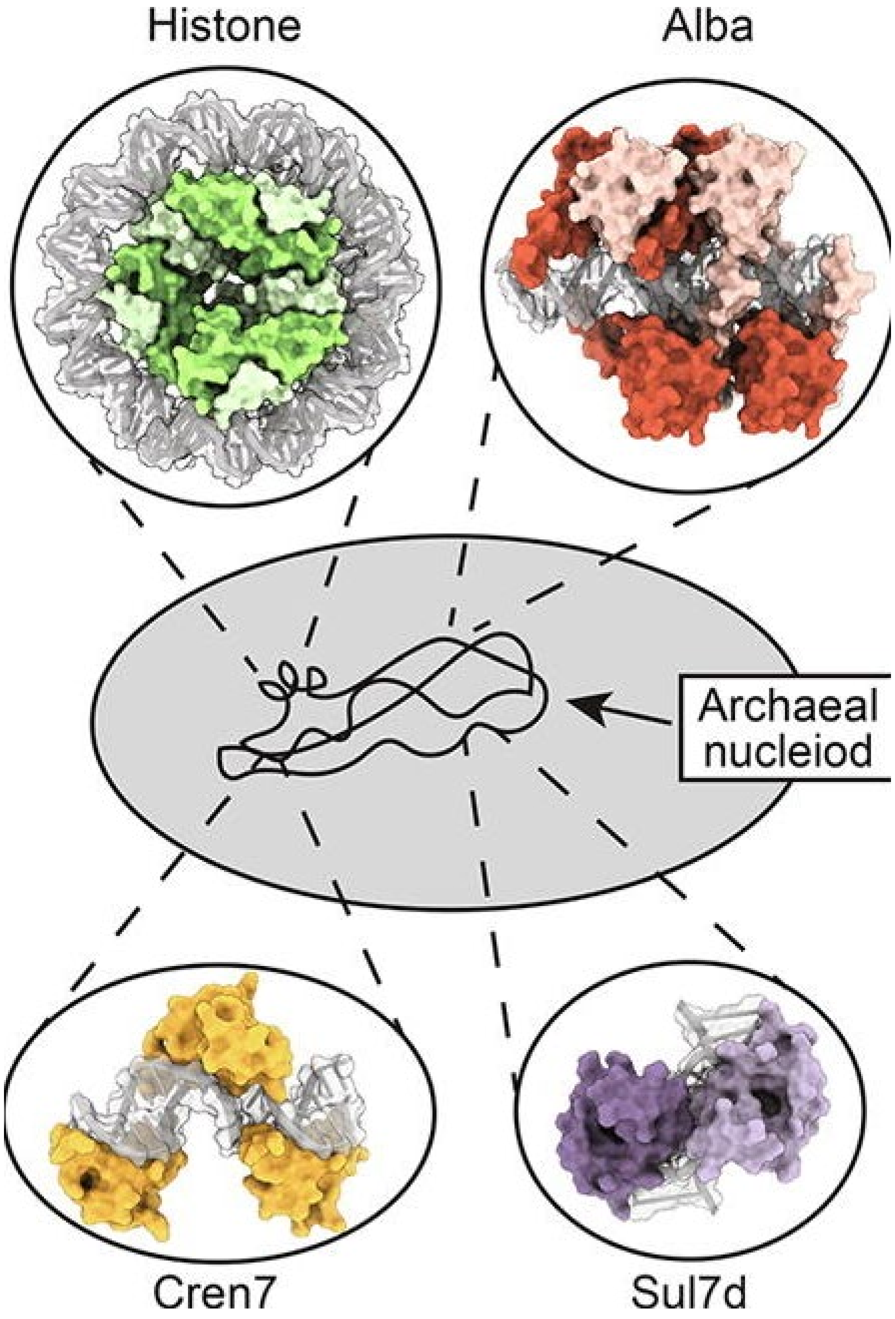
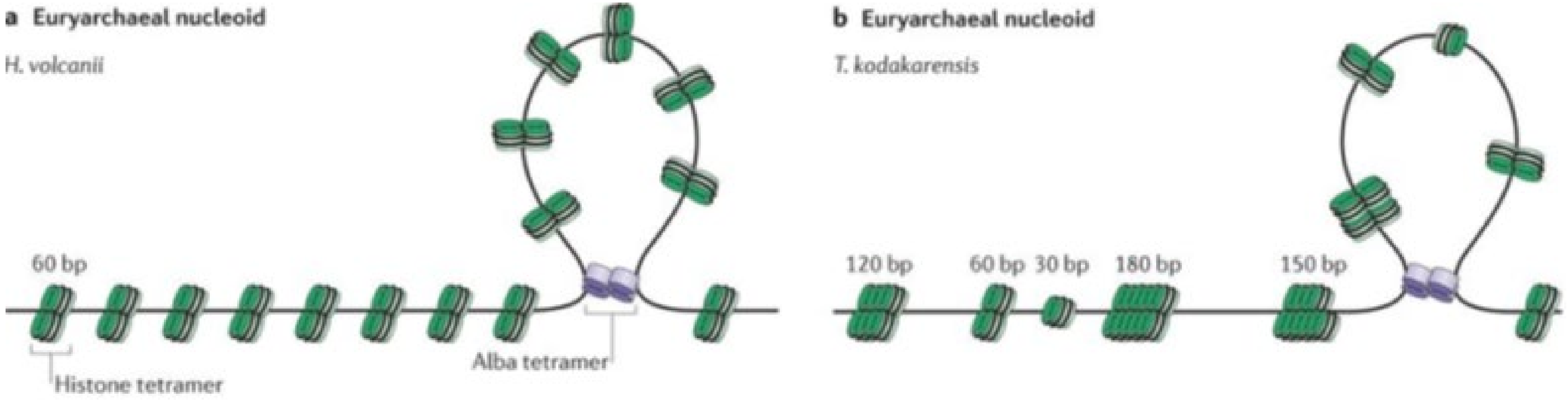


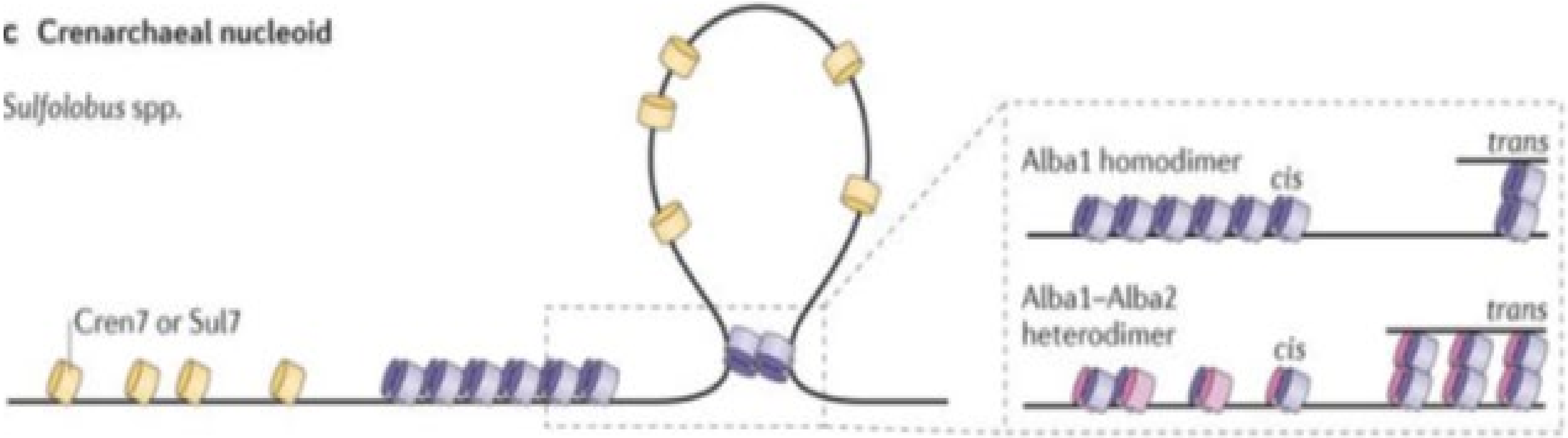
Figure 1. Groups of archaea utilize different proteins to structure their genomes. (A) Crystal structures of the four archaeal DNA-structuring proteins summarized in this review. (B) A simplified phylogeny, based on Williams et al.[9] showing the distribution of chromatin structure-related proteins among archaea. Although debate over the true taxonomy of Archaea is ongoing, we have placed four widely discussed branches on this tree: Asgard, TACK, Euryarcheota, and DPANN. Colored bars beside each phylum denotes the type of chromatin-organizing protein present in those species. Many archaea encode both histones and Alba, but bold text highlights taxa without identified histone or Alba sequences, as shown by gaps in the green and red bars, respectively. Cren7 and Sul7d exist only in Crenarcheota, which do not typically c.ontain histones.


Histone
Alba
Archaeal nucleiod
Cren7
Sul7d

The structure of the archaeal nucleoid varies among different archaeal species depending on the chromatin proteins they express



a,b | The euryarchaeal nucleoid is mainly organized by histone proteins that bend or wrap DNA, as well as by Alba that binds to DNA as a homodimer or a heterodimer and that forms looped structures by bridging two DNA duplexes. In *Haloferax volcanii*, histone proteins form tetrameric nucleoprotein structures that wrap about 60 bp of DNA around their surface (part **a**). These nucleosomes form a regular 'beads-on-a-string' structure similar to eukaryotic chromatin. In *Thermococcus kodakarensis*, histone proteins assemble into multimeric forms that cover variable sizes of DNA ranging from 30 bp (indicative of a dimer binding) to 450 bp (part **b**).



c | The crenarchaeal nucleoid is organized by proteins that bend DNA (for example, Cren7 and Sul7 in *Sulfolobus* spp.), as well as by Alba that either forms looped structures by bridging two DNA duplexes or forms stiff filaments by binding cooperatively side by side. The best-studied chromatin proteins belong to the Alba superfamily, which is widely distributed and almost universally present in archaea[15]. Alba seems to have an ancient evolutionary history and considerable functional plasticity[16]. Most Alba proteins interact with RNA in addition to binding to double-stranded DNA (dsDNA)and have been suggested to function in RNA metabolism· In euryarchaeal methanogenic archaea, Alba proteins are low-abundance, sequence-specific dsDNA-binding proteins[19], whereas in **crenarchaeal** organisms, it was shown that **Alba is a highly abundant cellular** protein that binds to dsDNA without apparent sequence specificity. Alba assembles into dimers, which are homodimeric or heterodimeric depending on whether paralogues are encoded and on their relative amounts.

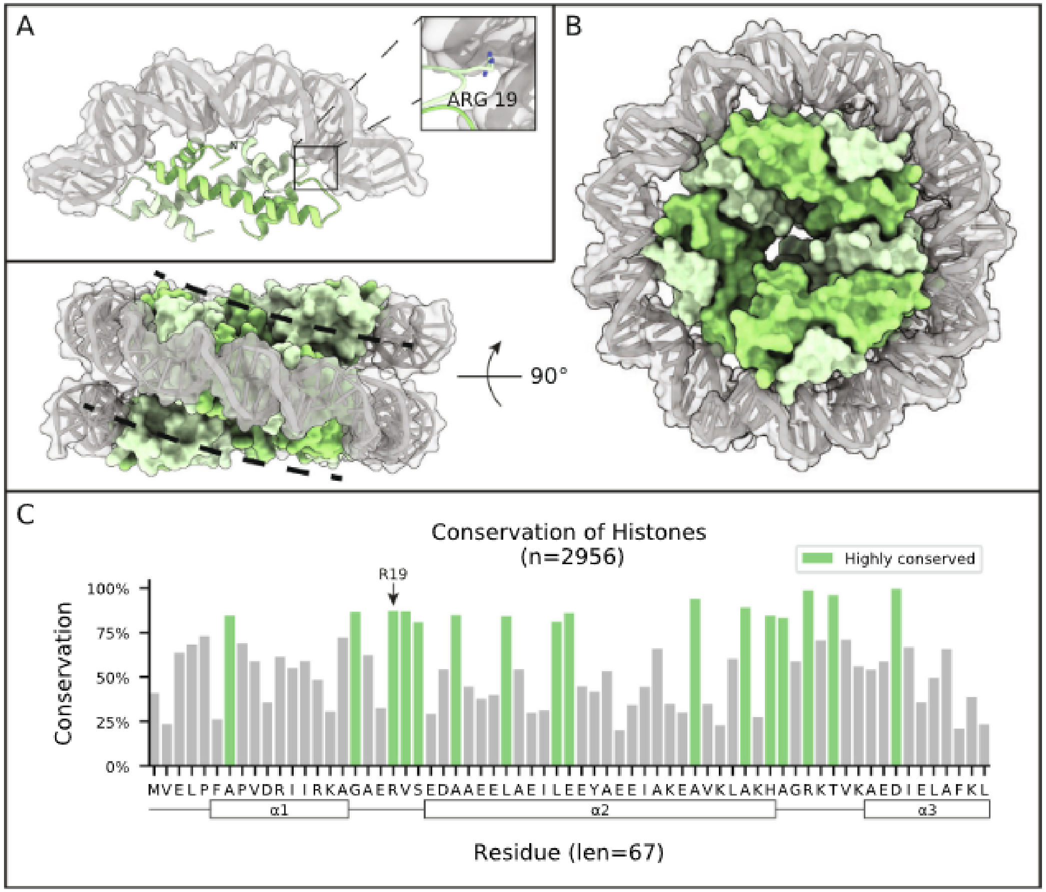


Figure 2. Archaeal histones wrap DNA into a superhelix. (A) The crystal structure of HMfB dimers bound to 90 bp of DNA (only 30 bp of DNA is shown with one dimer bound) shows the arching of DNA induced upon binding by histones (PDB PDB5T5T5K). The RMSD between this histone arrangement and DNA-bound eukaryotic dimers is only 1.7 Å.[10] (B) Crystal contacts in the 5T5K structure suggests an organization for higher-order chromatin compaction, where repeated stacking of dimers wraps DNA into a nucleosome-like helical ramp. Shown from both face and side views, is a model of 4 sets of HMfB dimers wrapping 120 bps of DNA. (C) Conservation of putative histone homologs found in Archaea. Highlighted regions are residues conserved at least one standard deviation more than the mean conservation across the alignment. Black frames in the sequence identify predicted α-helical structures, based on homology with HMfB. The short loops connecting the α-helices are involved in DNA .binding (see A).

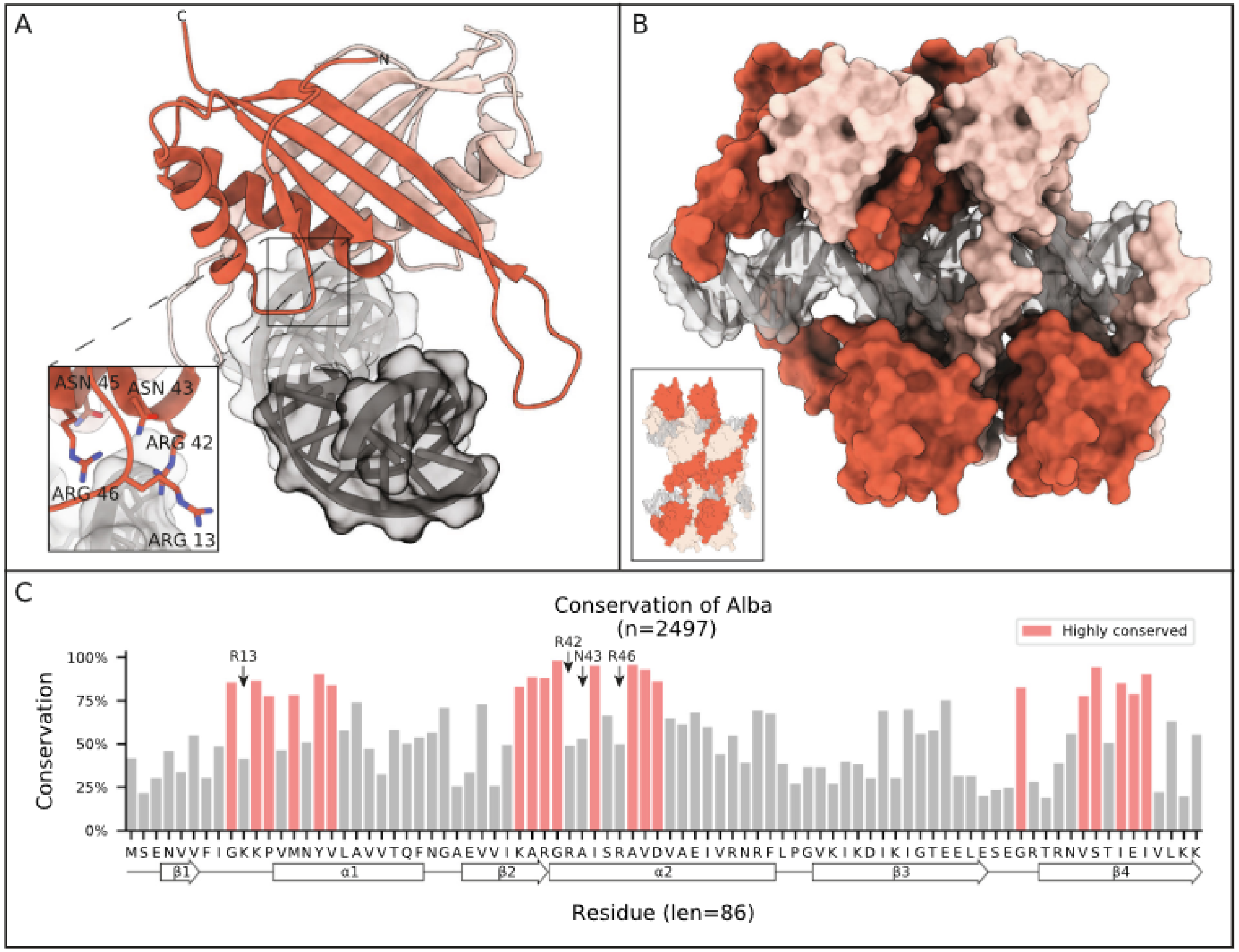
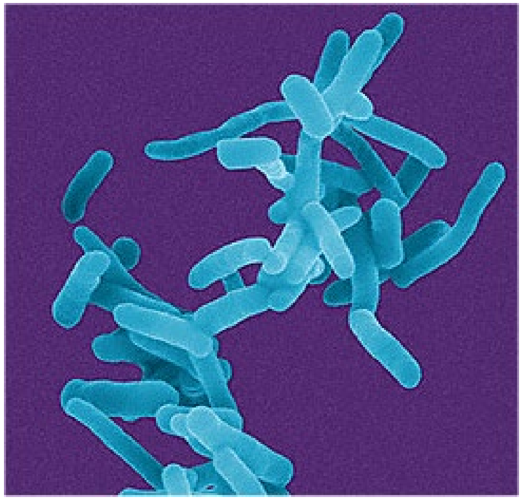


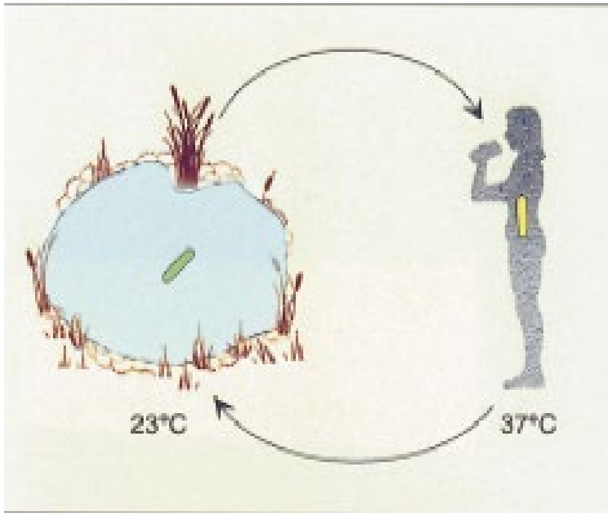
Figure 3. Alba dimers encase DNA. (A) Crystal structure of an Alba dimer from *Aeropyrum pernix* K1 bound to 16 bp of DNA (PDB PDB3U6U6Y). Only the first 4 bp, shown in dark grey, were resolved in the asymmetric unit, the rest are modeled in based on adjacent asymmetric units. Inset shows important DNA binding residues. (B) A model suggested by Tanaka et al. of higher order chromatin filament induced by continuous Alba dimer binding (based on PDB PDB3U6U6Y). The inset shows model of adjacent antiparallel Alba-DNA filaments. (C) Conservation of Alba family proteins found in Archaea. Highlighted regions are residues conserved at least one standard deviation more than the mean conservation across the alignment. Secondary structural elements are projected along the residue consensus sequence on the bottom of the plot.

Shigella



- is a Gram negative, facultative anaerobe
- is an intracellular pathogen
- is the etiological agent of bacillary dysentery, an acute diarrheal disease
- causes 160 million of episodes, determining 1.1 million deaths/year in children and infants in developing countries.

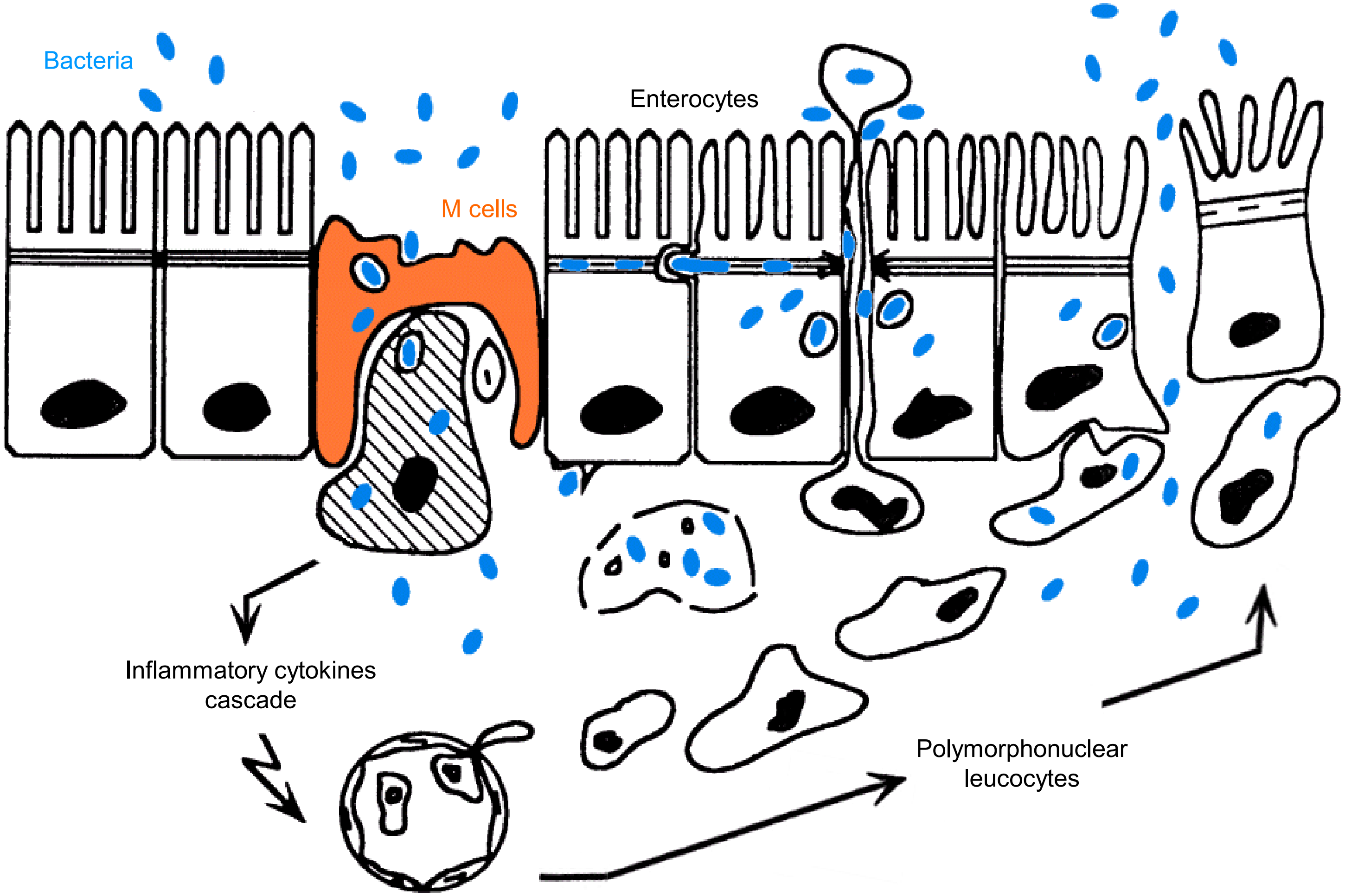
Infection is spread via fecal-oral route

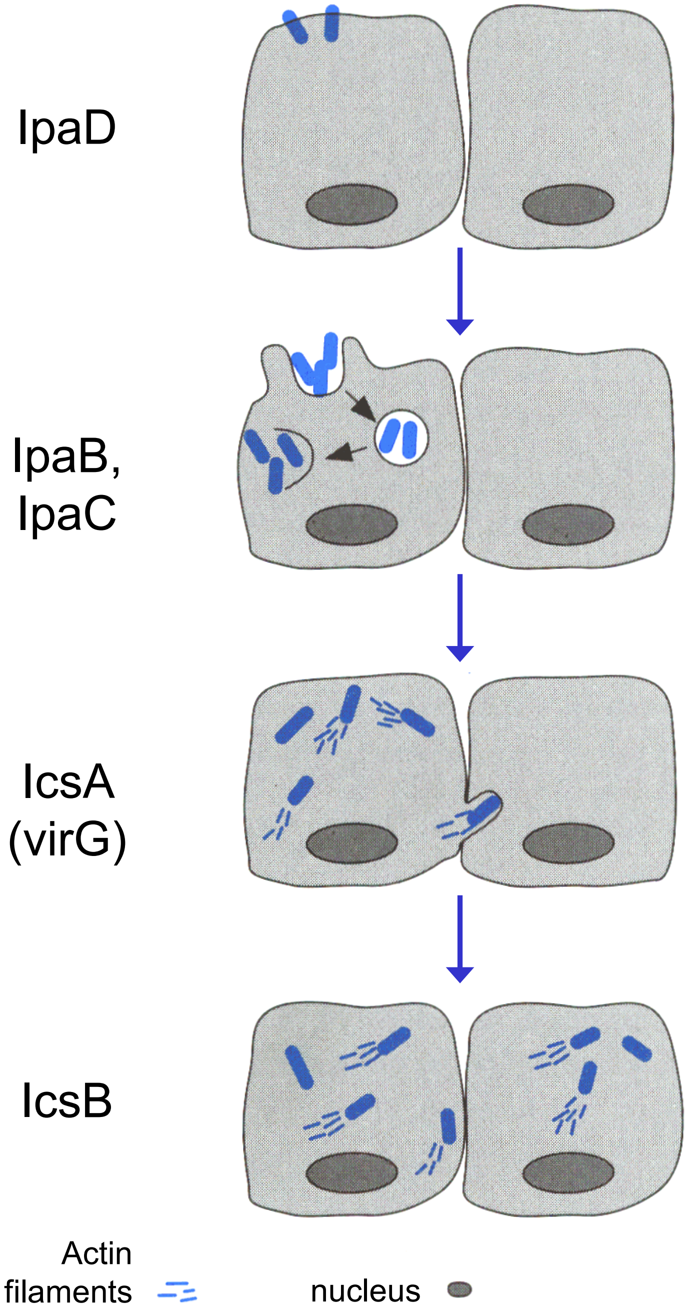


Subgrouped in four "species":

- *Shigella flexneri*
- *Shigella dysenteriae*
- *Shigella boydii*
- *Shigella sonnei*

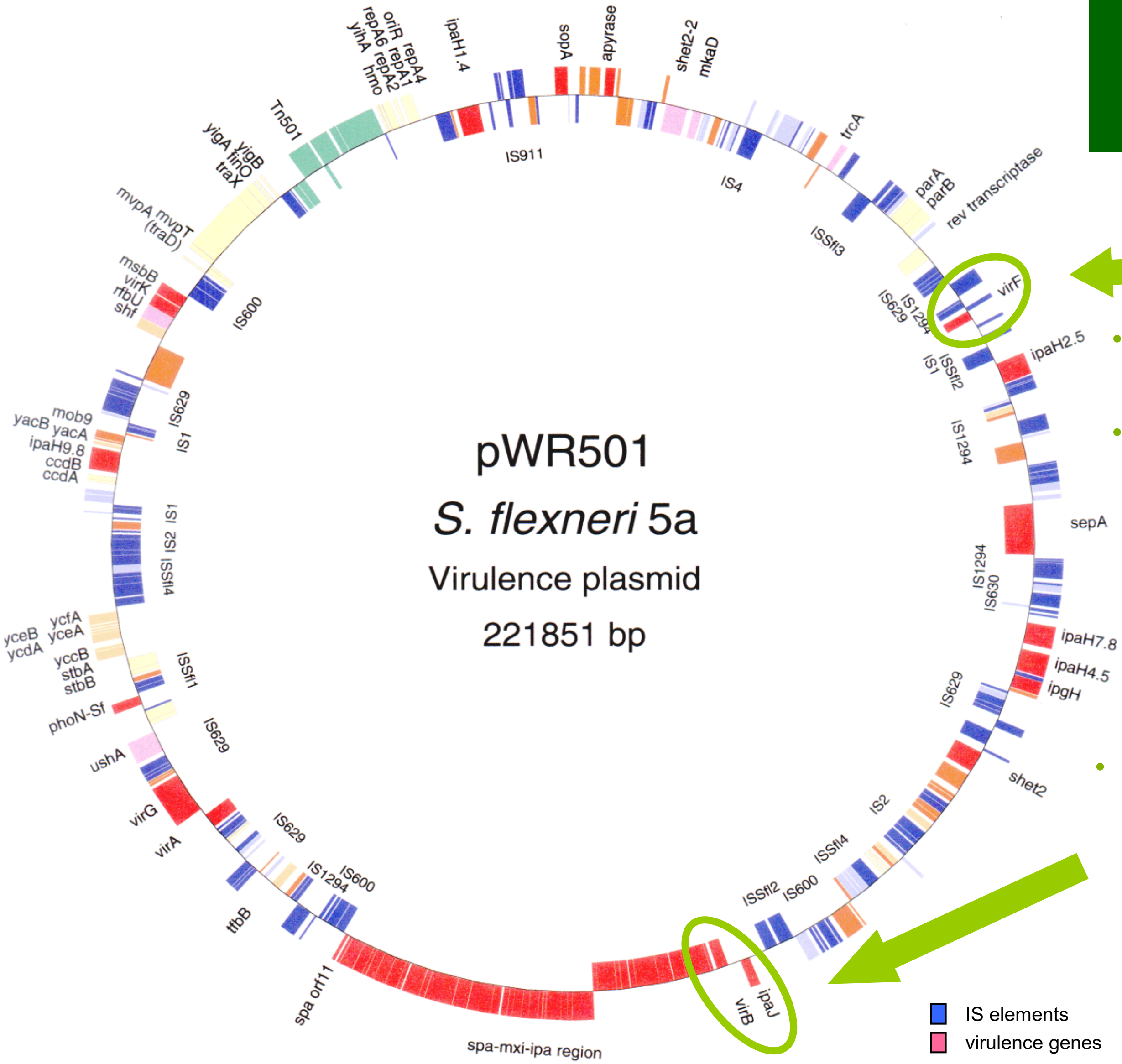
Model for *Shigella* invasion of the colonic mucosa





Proteins involved
in the invasion
process are
encoded by
a virulence
plasmid (pINV)

Genetic organization of pINV



🟦 IS elements

🟥 virulence genes

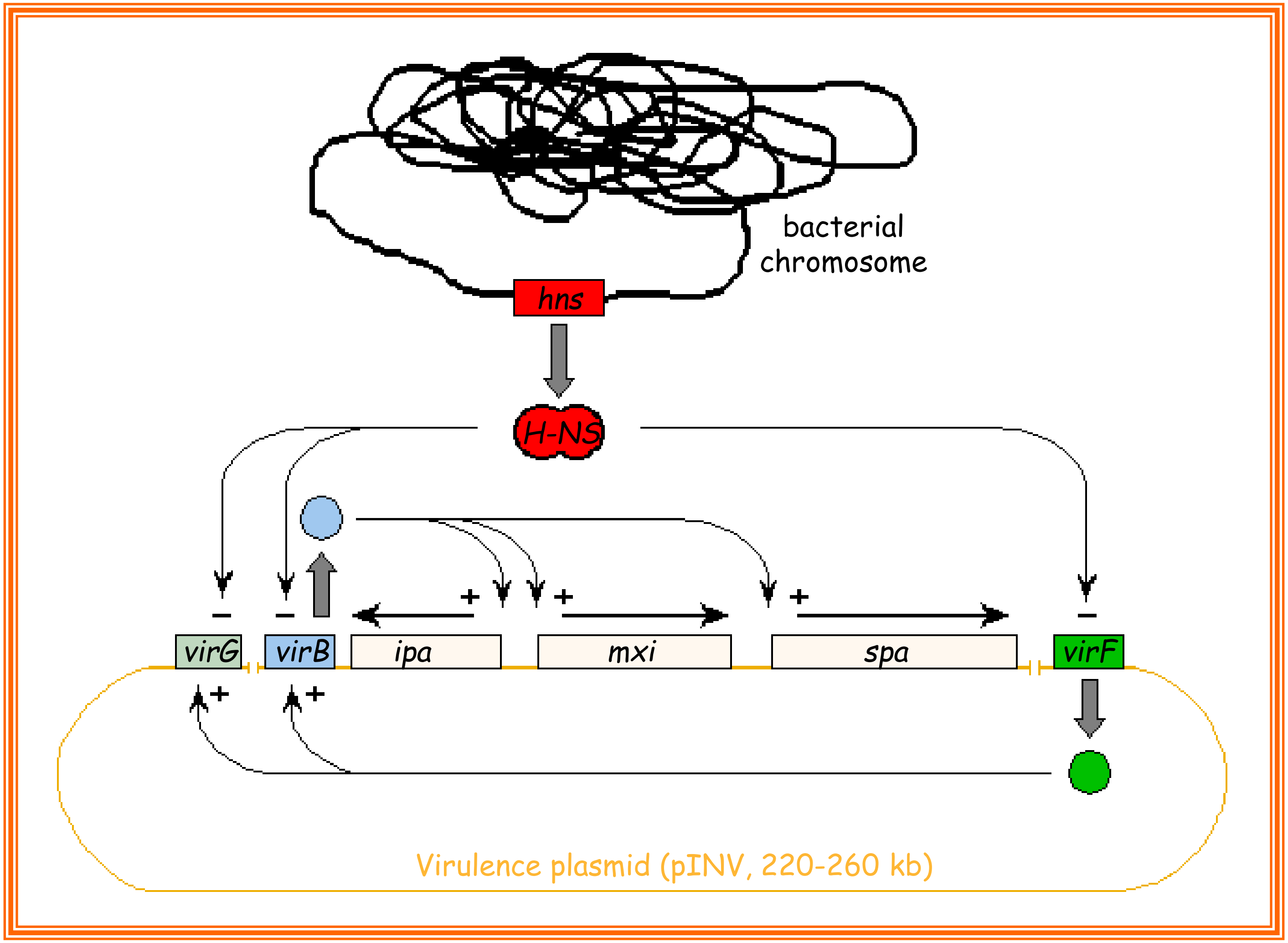
virF is …

- … located on a “desert island”
- … the first positive activator of pINV virulence genes

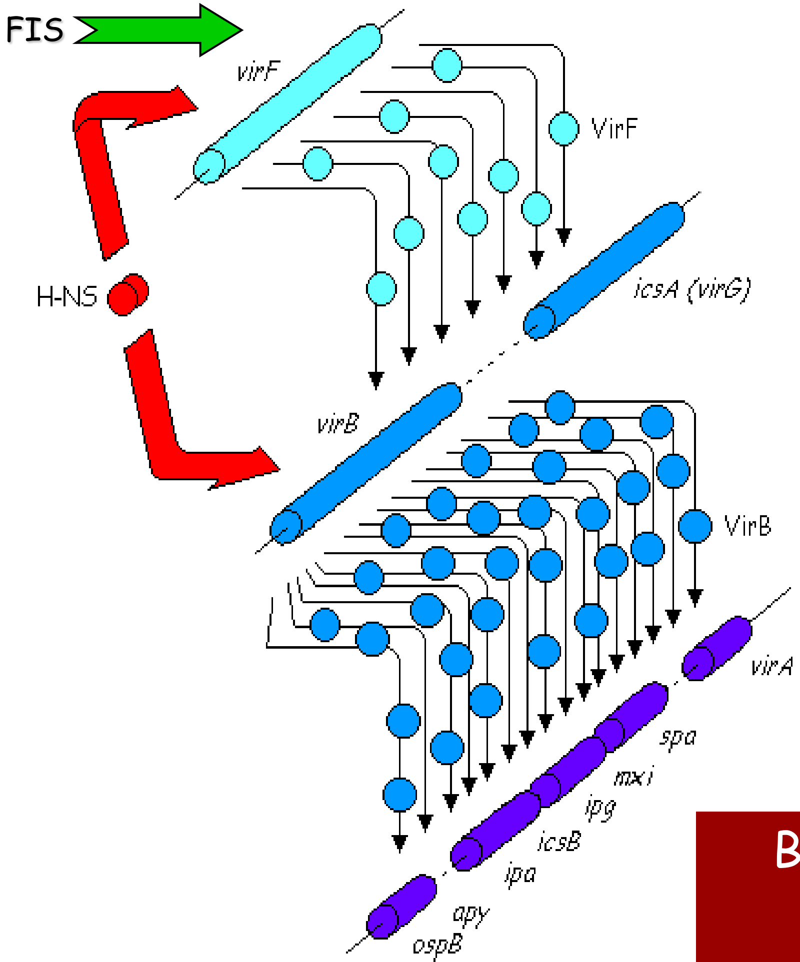
virB is …

- … located within the main Pathogenicity Island
  … the second positive regulator of the plasmid virulence regulons

H-NS controls the virulence regulon in *Shigella* and in *E.coli* EIEC



The expression “cascade” of virulence genes

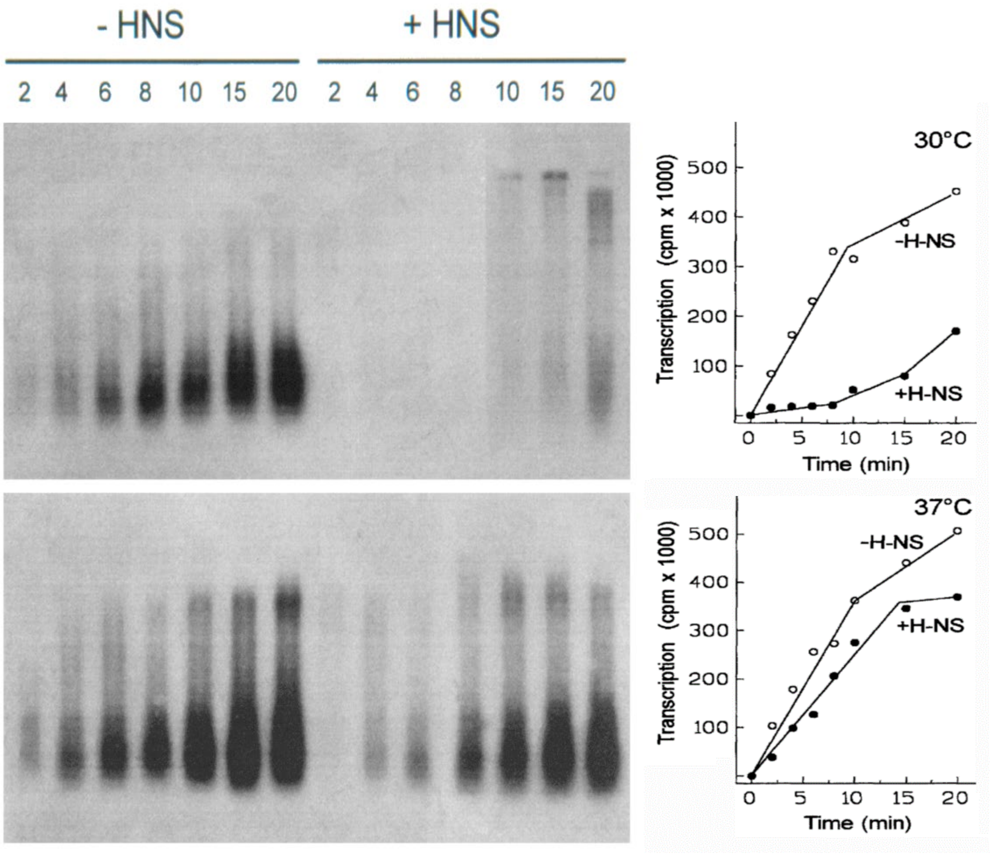


VirF is …

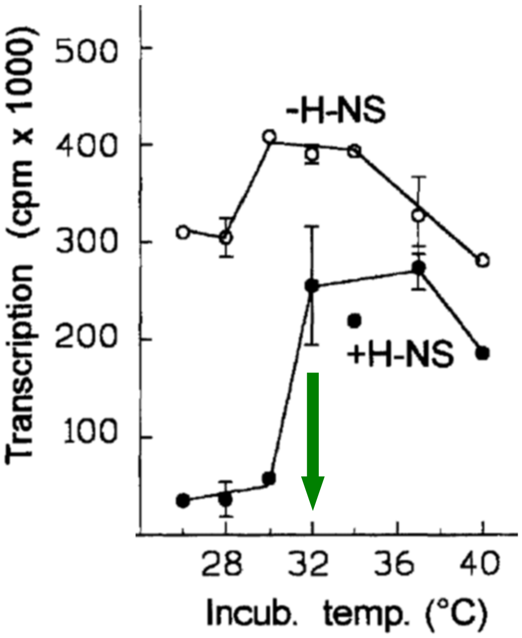
- … expressed at 37°C
- … is controlled antagonistically by two nucleoid proteins H-NS (repressor) and FIS (activator)

By which mechanism is the *virF* gene allowed to be expressed only at the host temperature?

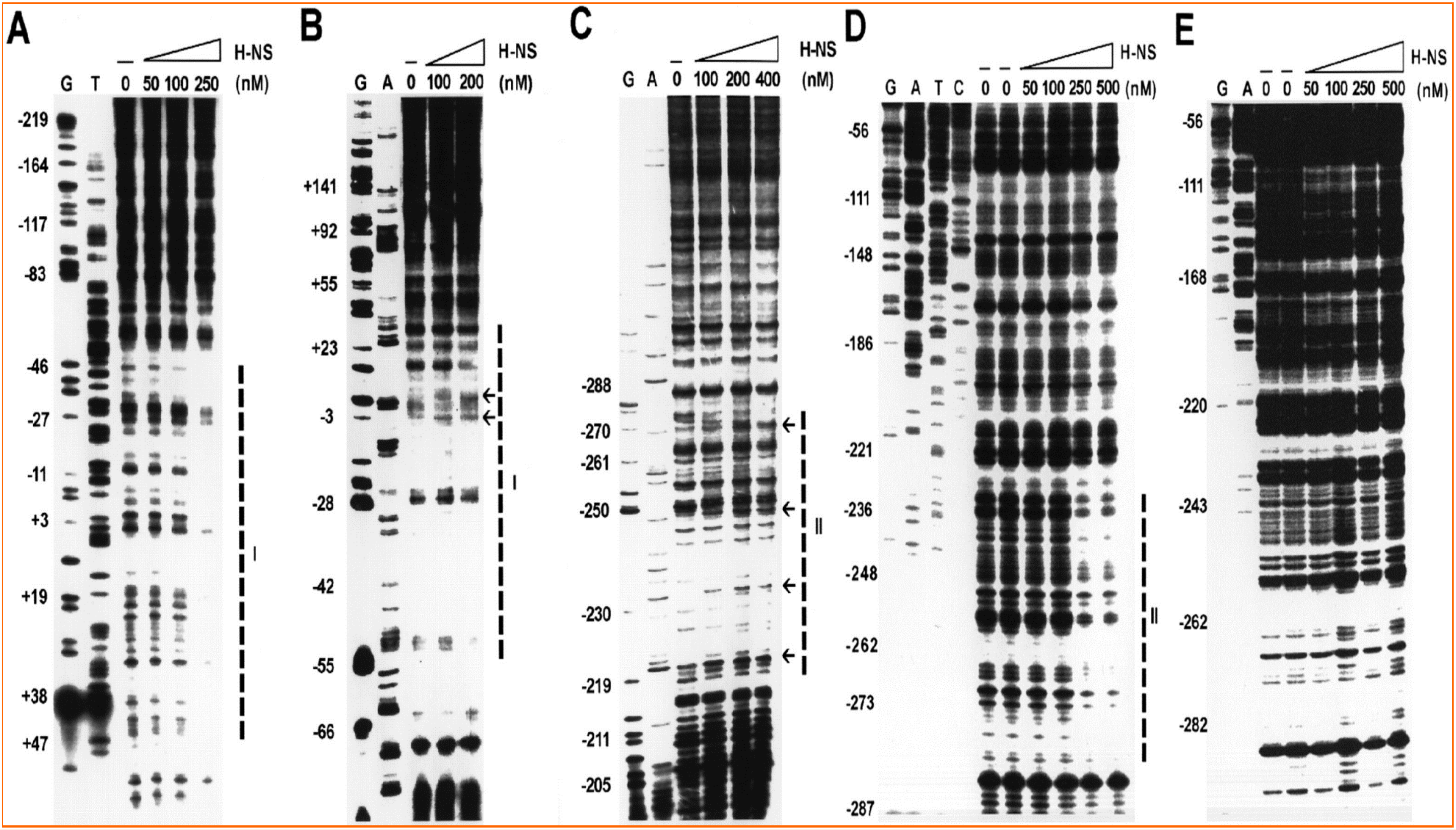
H-NS is able to bind to and repress *virF* only at low temperature



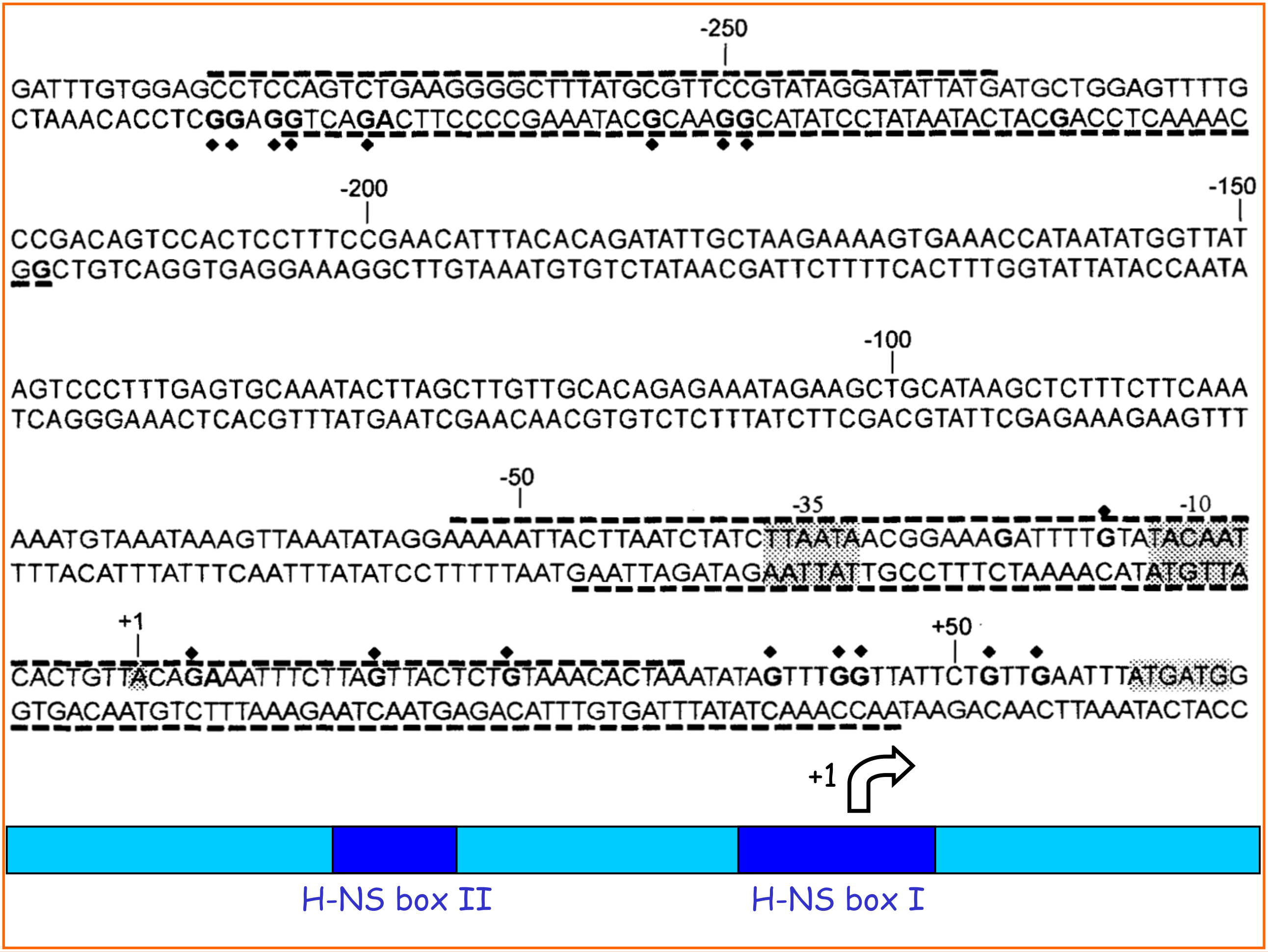
…more precisely:
virF transcription is inhibited by H-NS only below 32°C



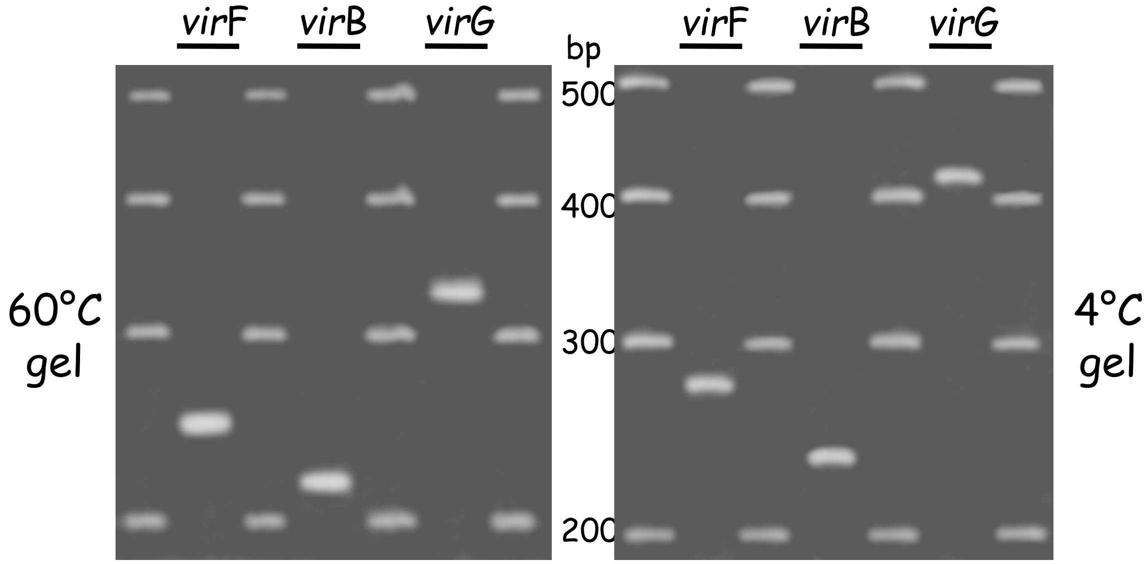
DNaseI footprinting of the *virF* promoter region by H-NS



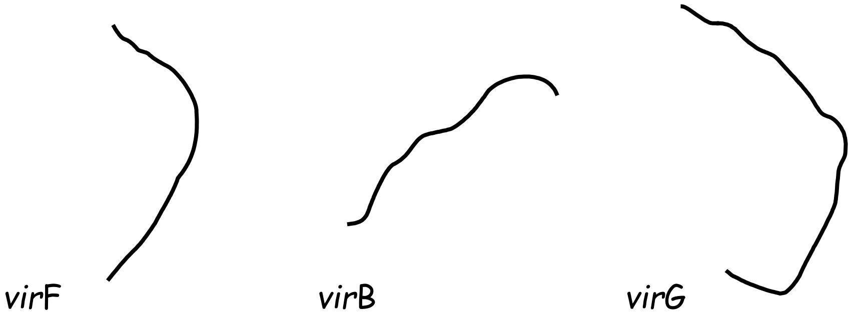
The *virF* promoter region



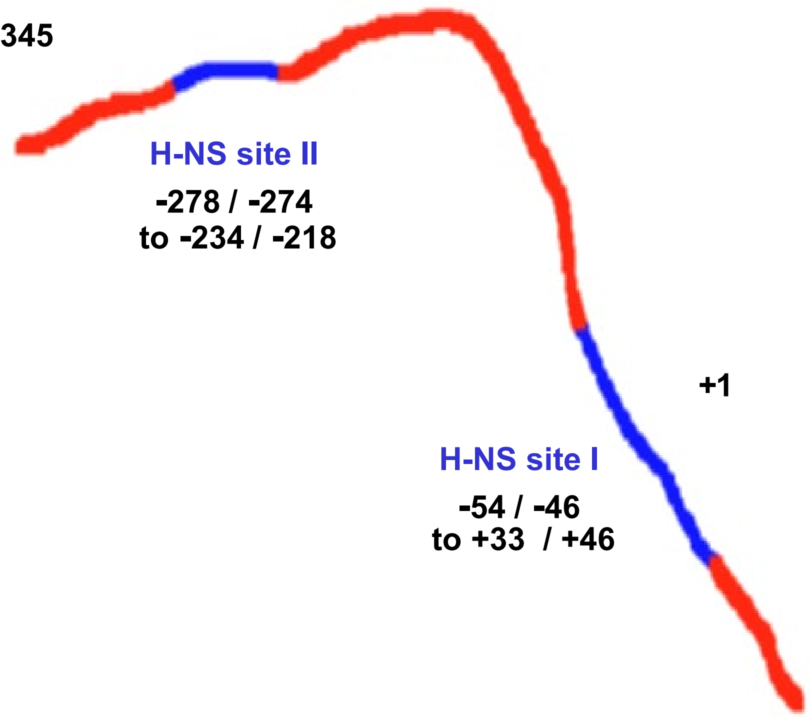
Curvature in H-NS regulated vir promoters



Computer-generated models



Within the *virF* promoter region H-NS recognizes two sites separated by a region endowed with signficant intrinsic curvature



-140

-345

H-NS site II

-278 / -274
to -234 / -218

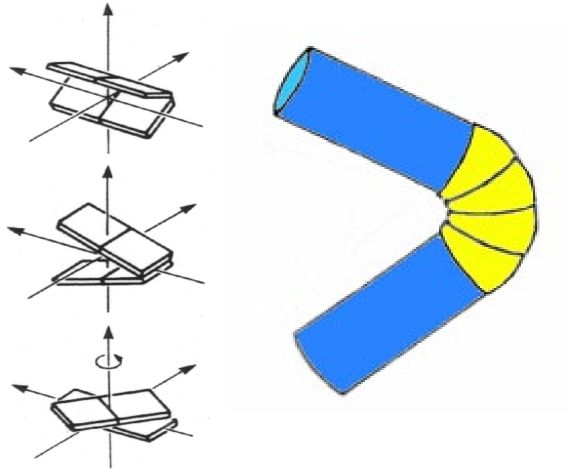
+1

H-NS site I

-54 / -46
to +33 / +46

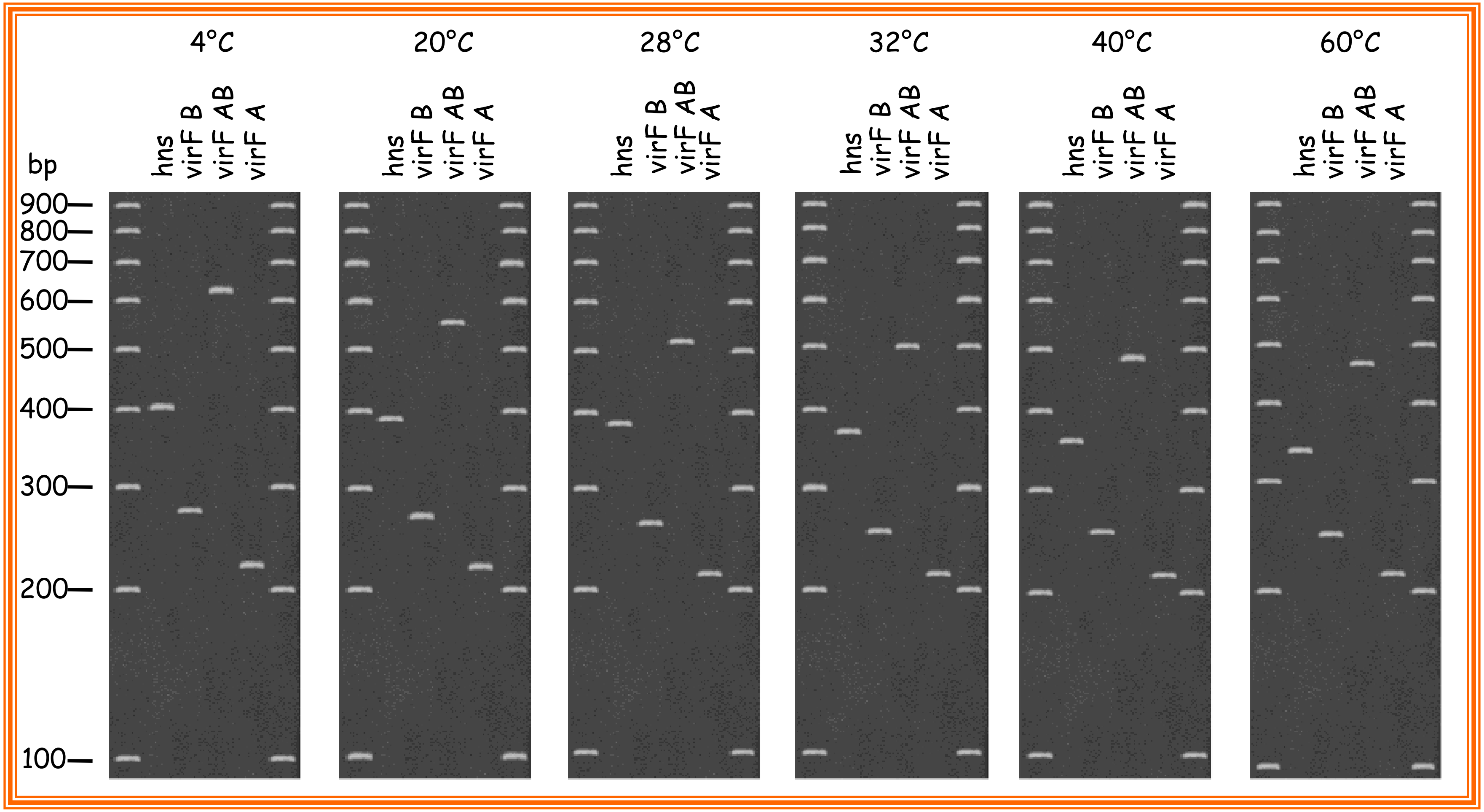
+104

DNA bending : the **wedge model**
for sequence mediated curvature

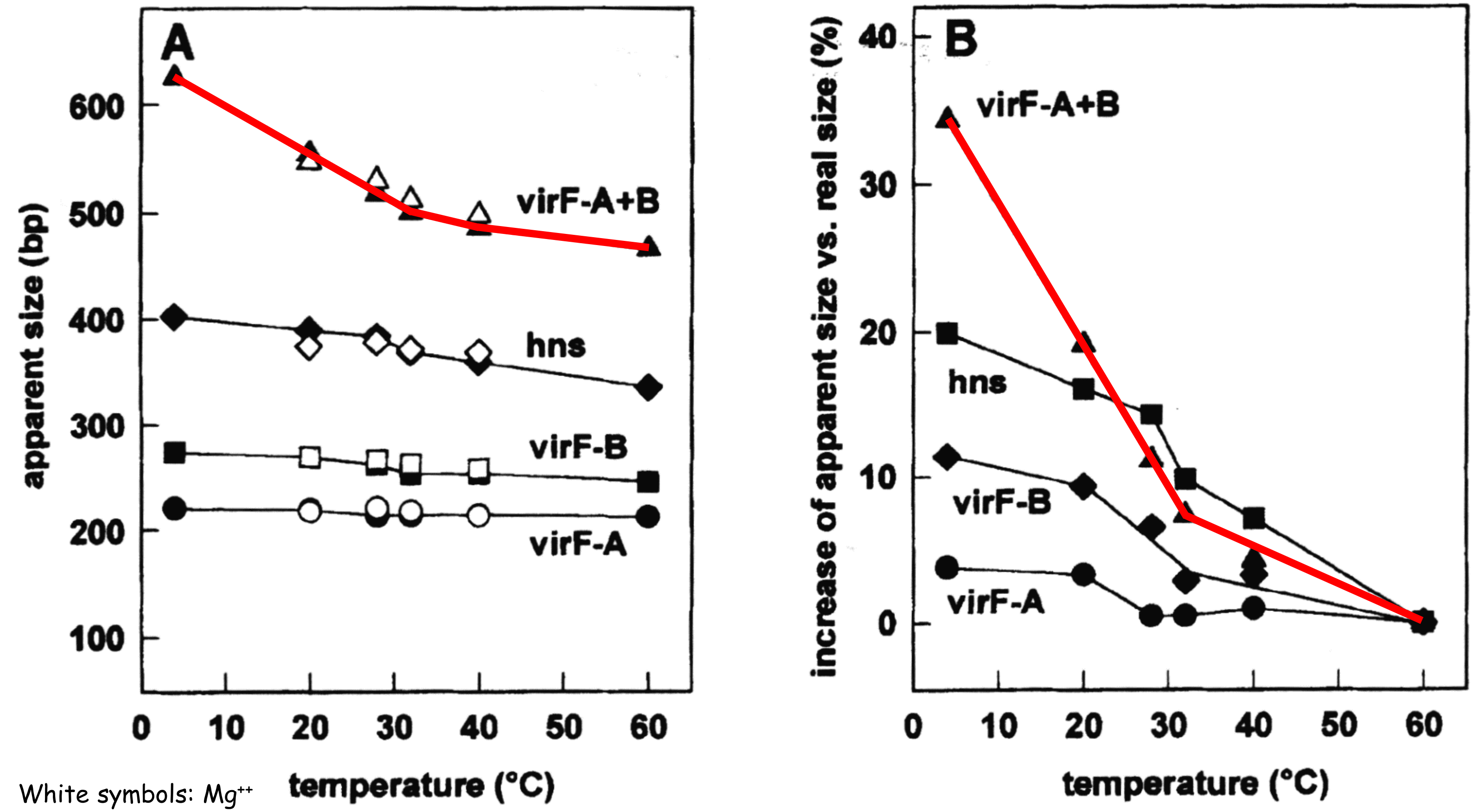


Temperature-dependent curvature of the *virF* promoter

Acrilamide separations



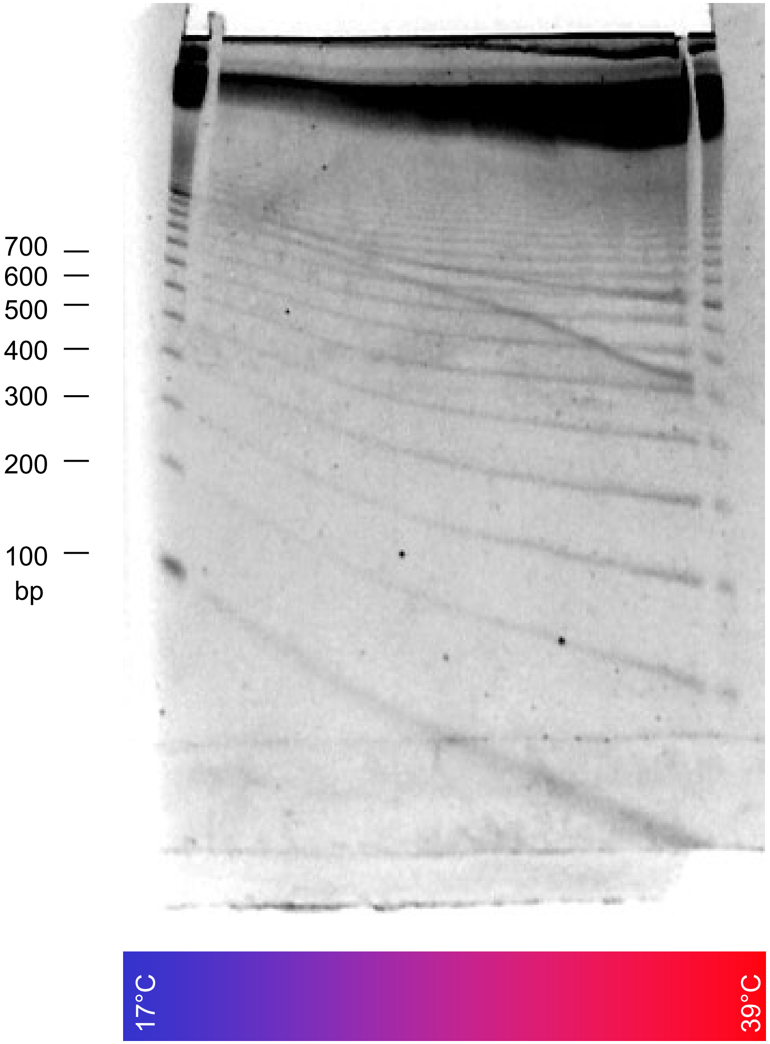
The curvature of the *virF* promoter is strongly temperature-dependent

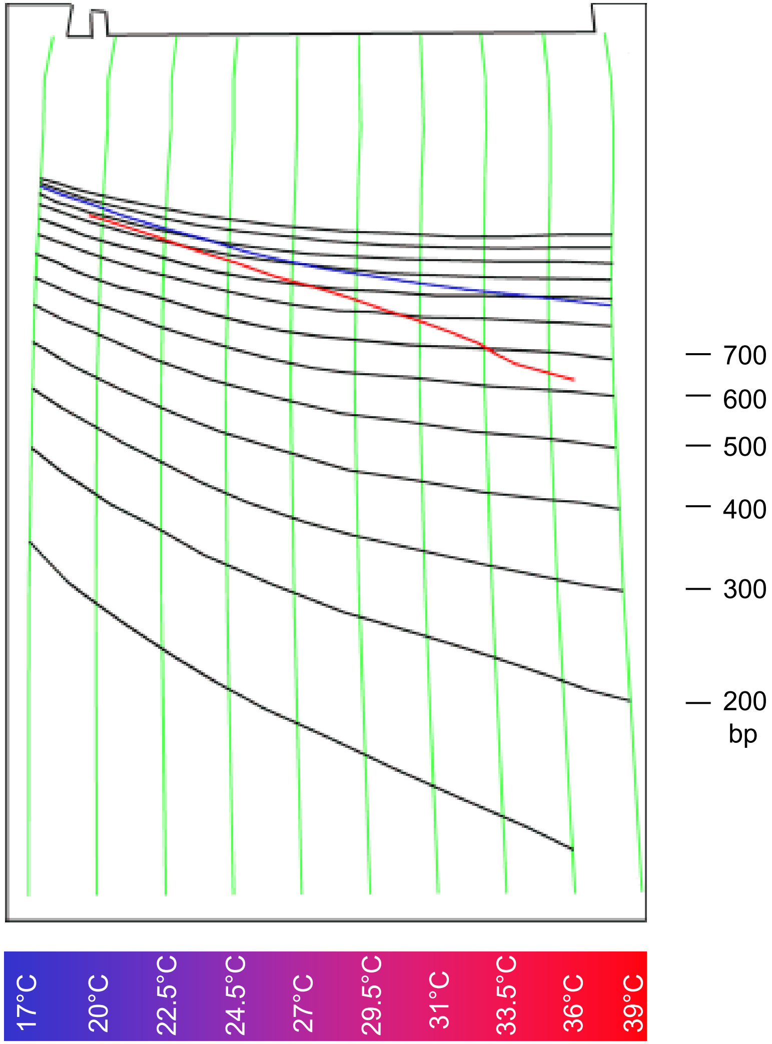


White symbols: Mg^{++}
Black symbols: no Mg^{++}

Temperature-dependent curvature of the *virF* promoter

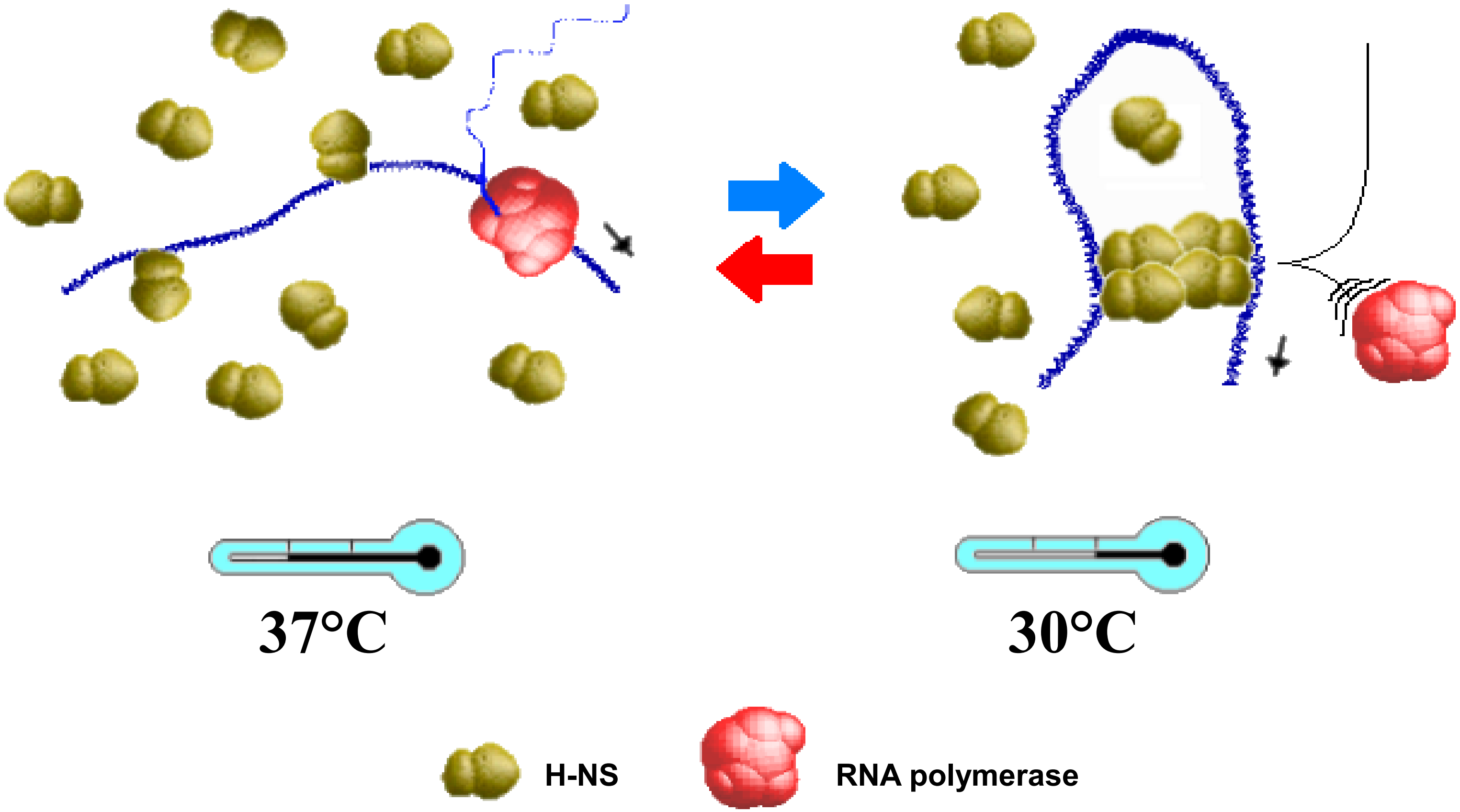
TGGE separation





Temperature-dependent *vir*F expression

Working model

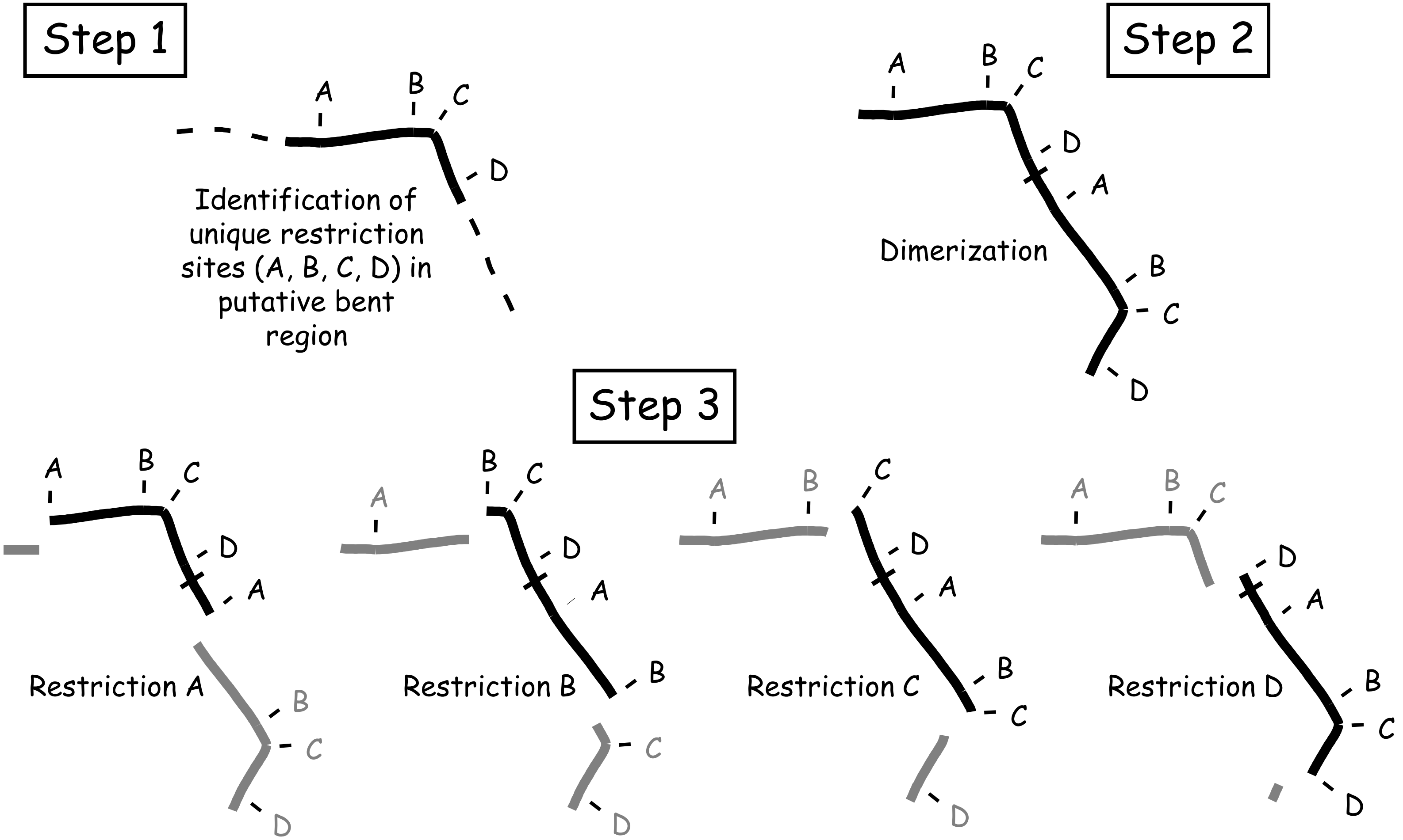


By which mechanism is the *virF* gene
allowed to be expressed
only at the host temperature?

Small RNAs are emerging as key regulators
of virulence gene expression in bacteria.
Is this true also in Shigella?

The circular permutation assay

Rationale



Step 1

Identification of unique restriction sites (A, B, C, D) in putative bent region

Step 2

Dimerization

Step 3

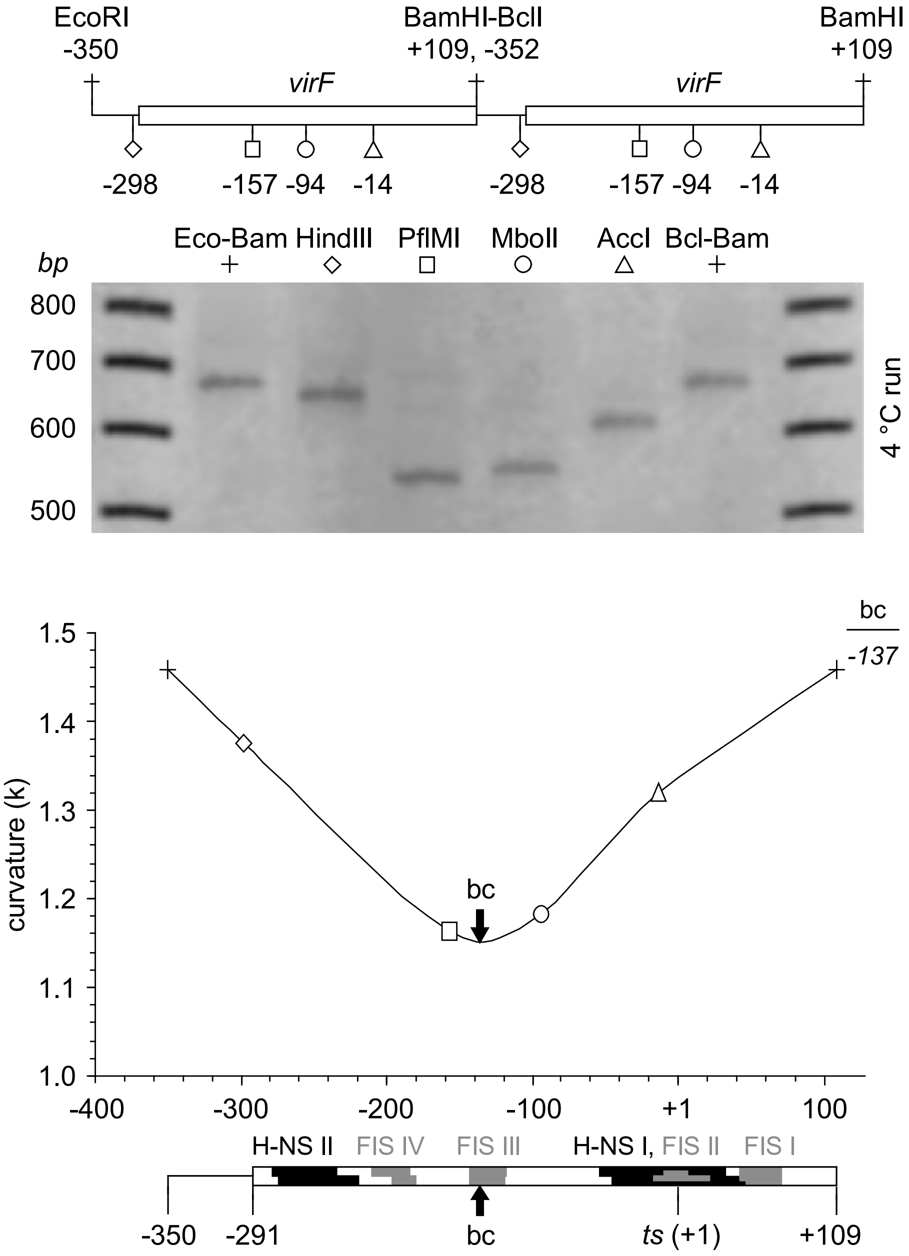
Restriction A

Restriction B

Restriction C

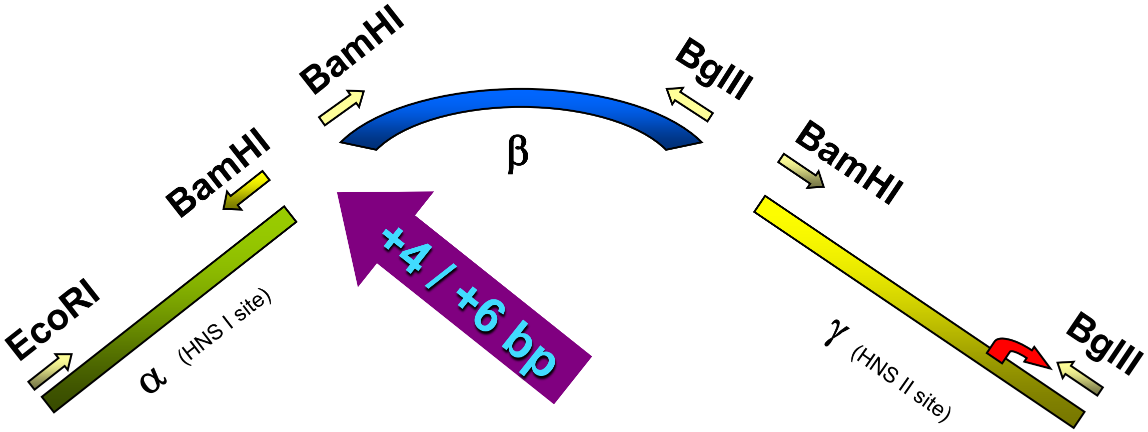
Restriction D

Circular permutation assay on the *virF* promoter region

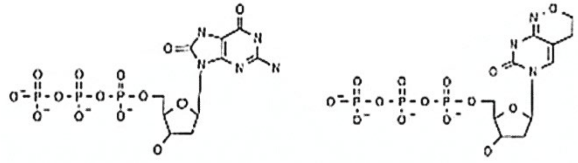


The bending centre maps halfway between the H-NS boxes and is located ~140 bp upstream the transcription start site (+1)

Molecular dissection of the *virF* promoter: mutagenesis of the β-region and shifting the β-region by ~½helix turn



8-oxo-dGTP **dPTP**



A=T ➡ T=A, C≡G

A=T ➡ G≡C

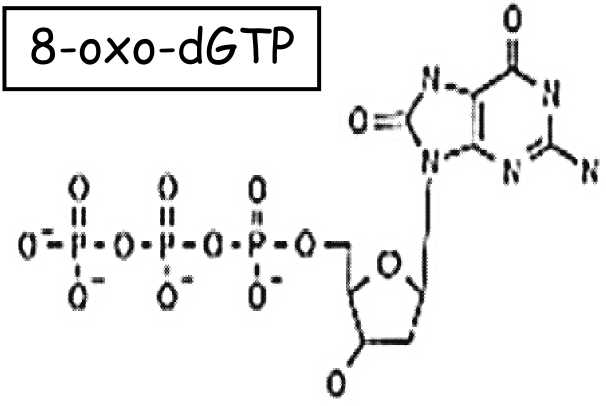
GCGCGCAAAAAGCGCGTAAAAGCGCGATAAAGC

GCGCGCAGAGAGCGCGTAGAAGCGCGACAAAGC

Mutagenesis of the bent region

Mutagens used

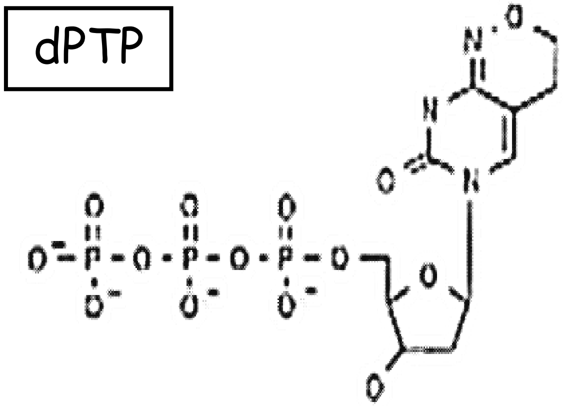
8-oxo-dGTP



A=T → T=A

C≡G

dPTP



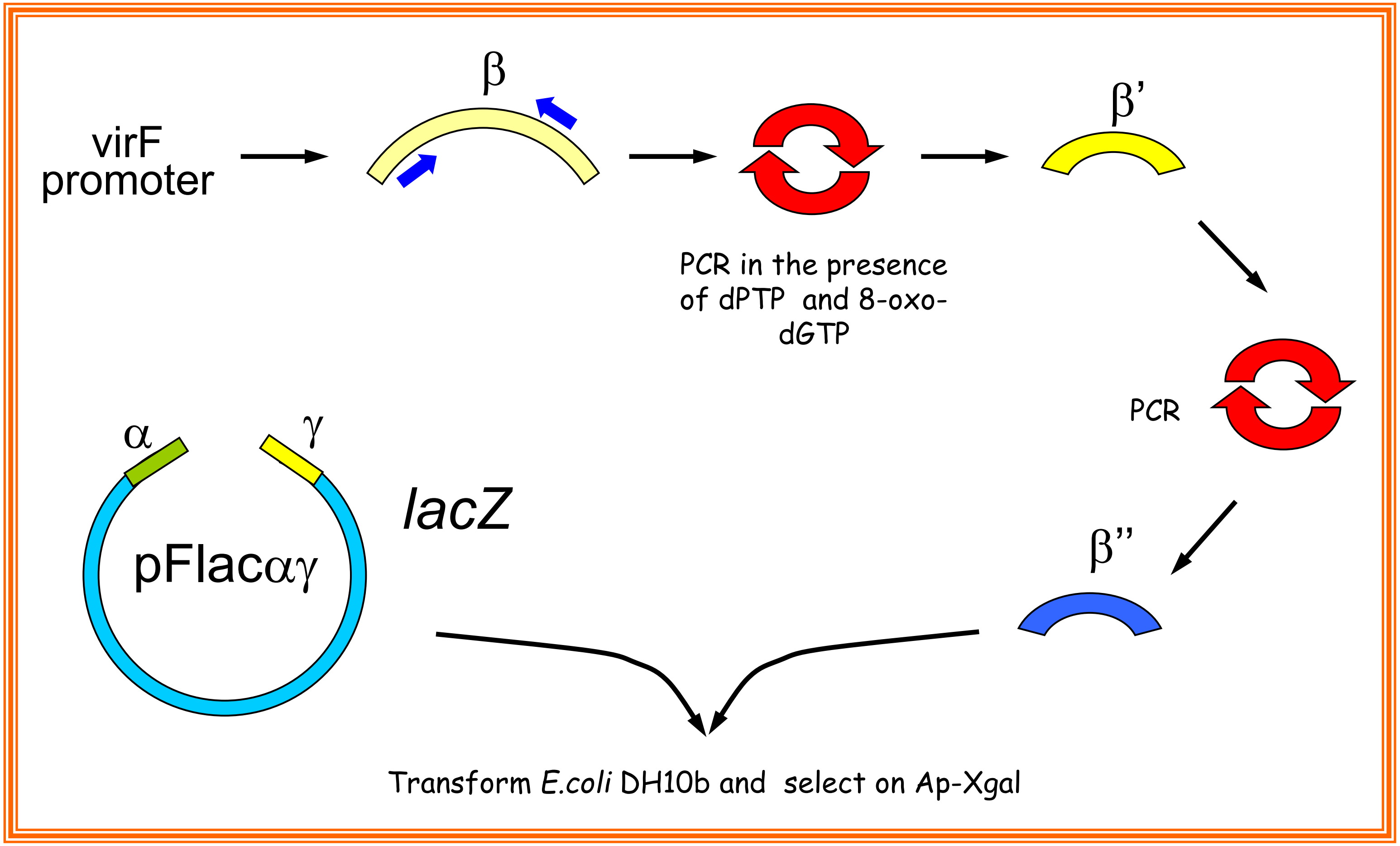
A=T → G≡C

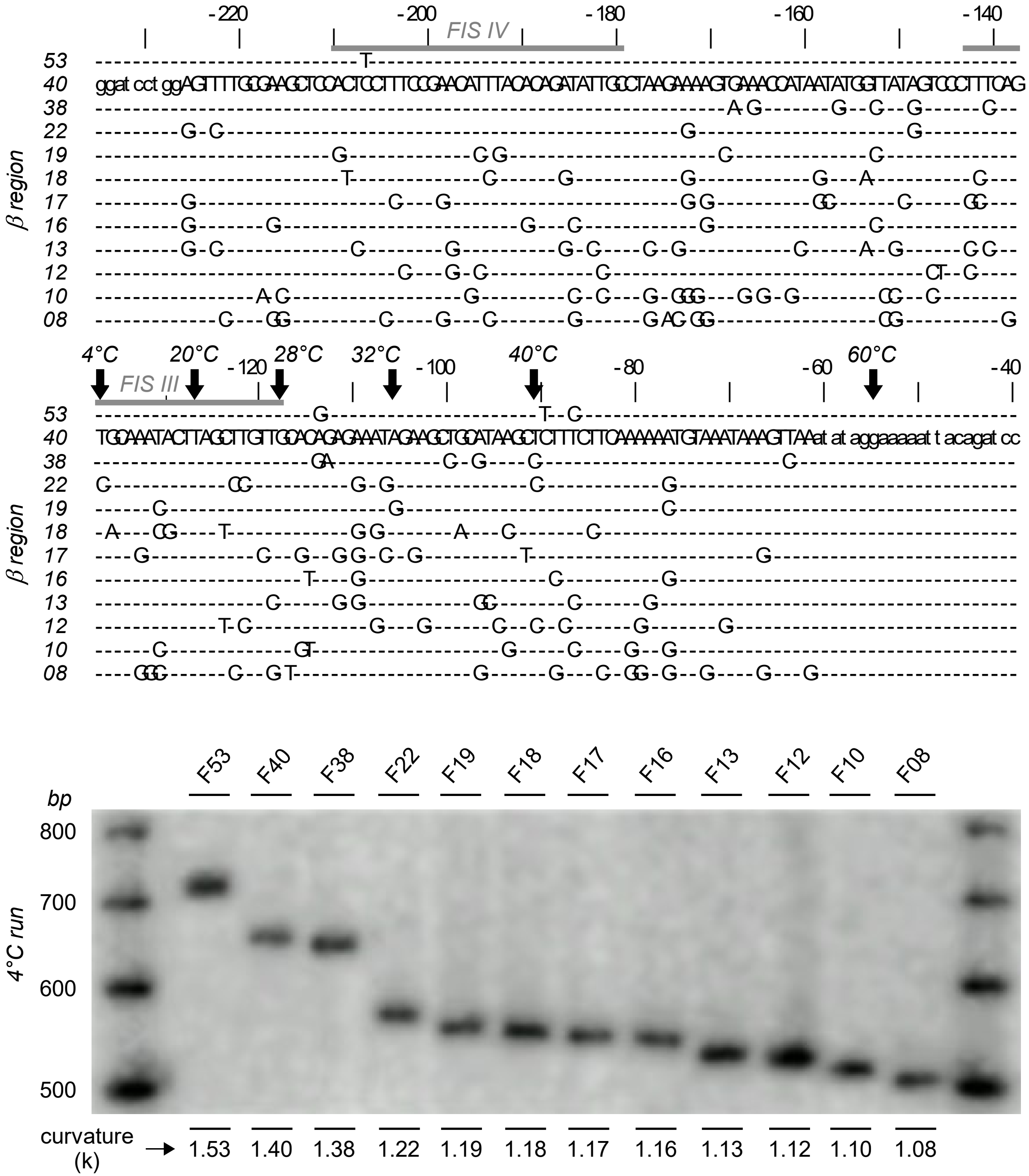
GCGCGCAAAAAGCGCGTAAAAGCGCGATAAAGC

GCGCGCAGAGAGCGCGTAGAAGCGCGACAAAGC

Mutagenesis of the bent region

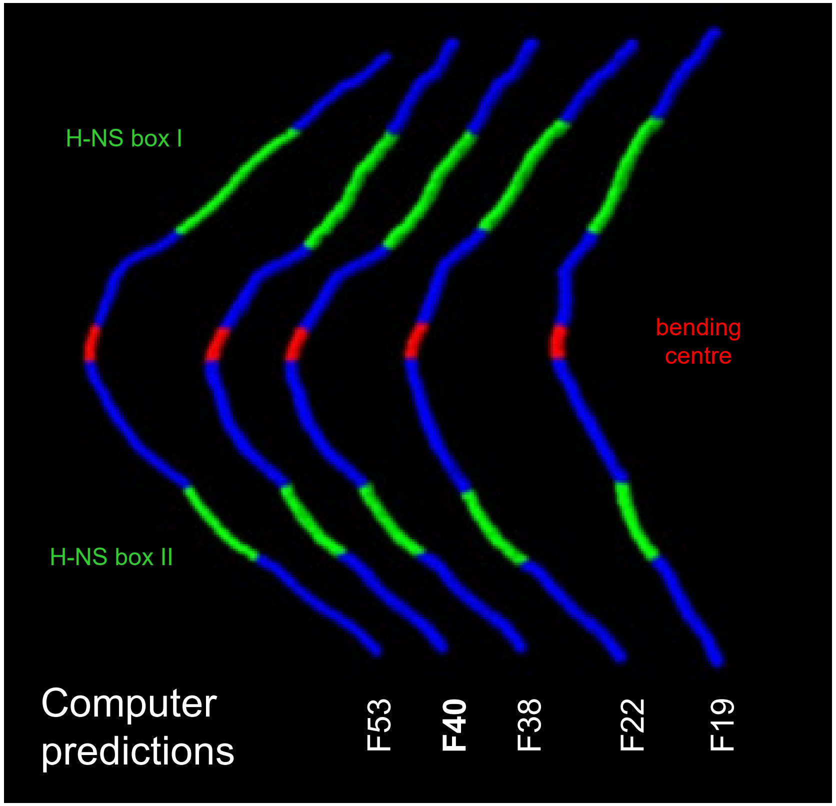
Strategy for the construction of β-region mutants

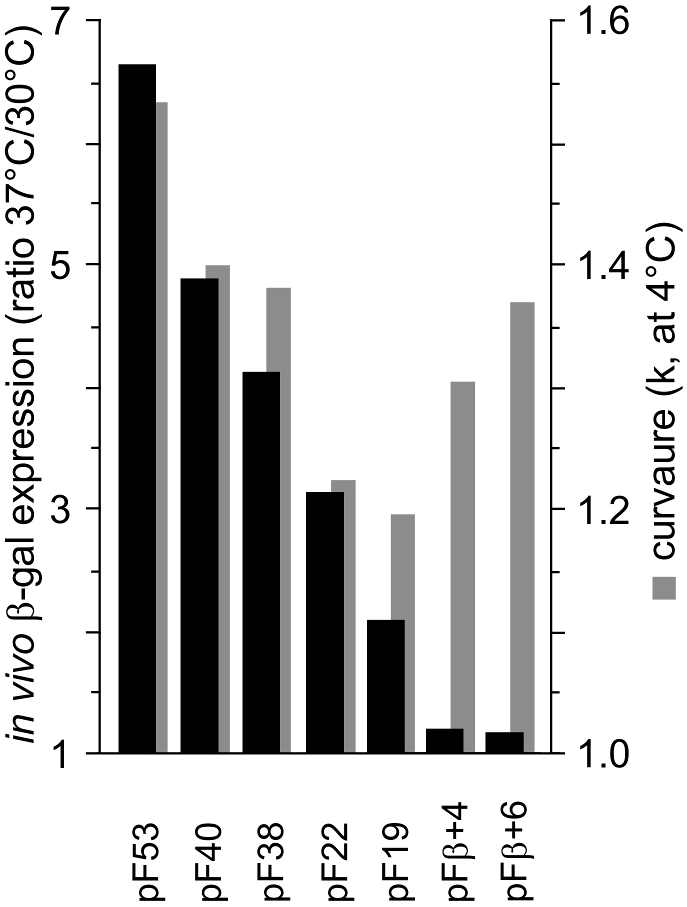




Acrilamide separation (4°C) of β-region mutants

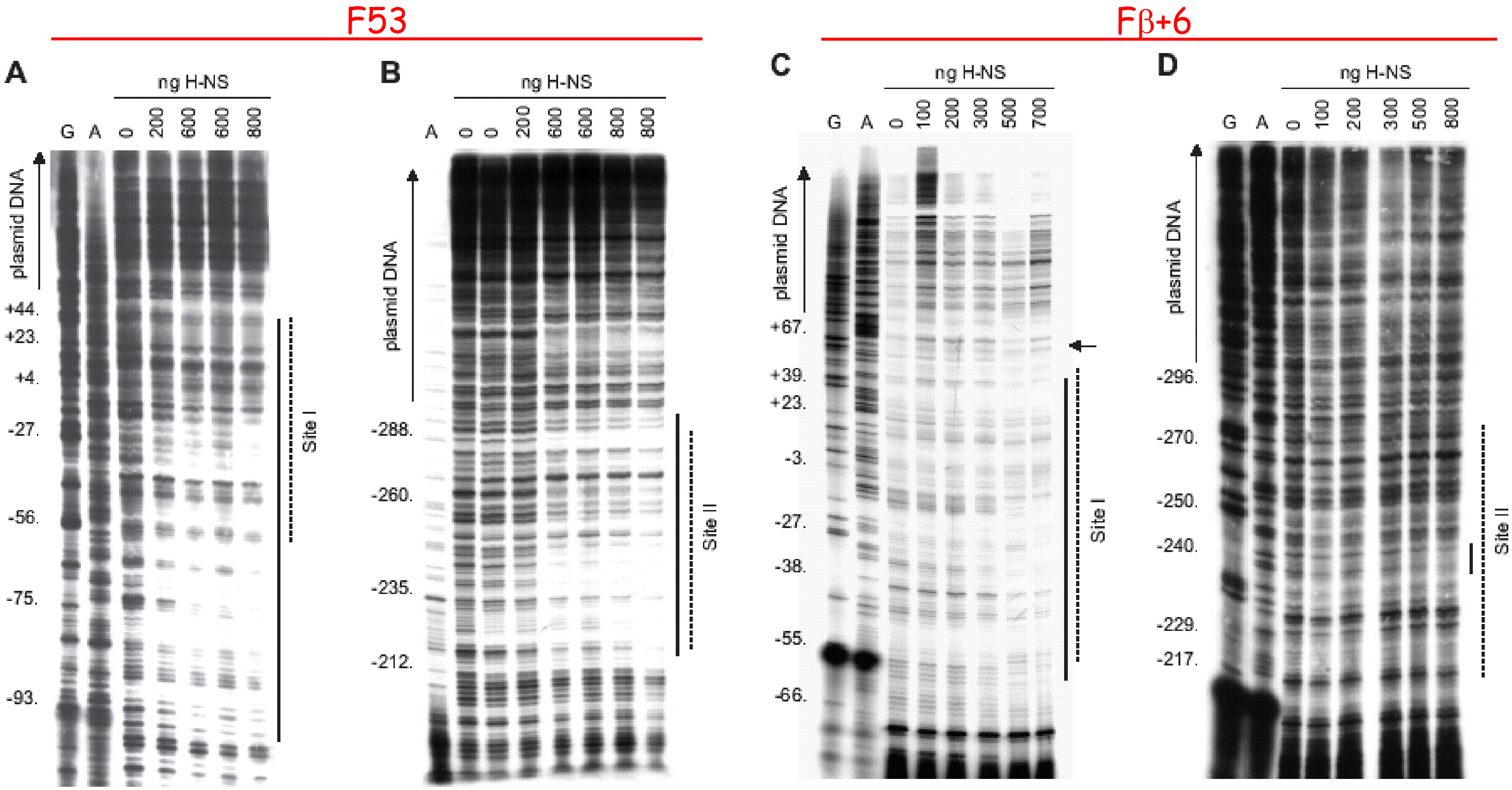
Correlation between intrinsic curvature and thermoregulated expression of the *virF* promoter



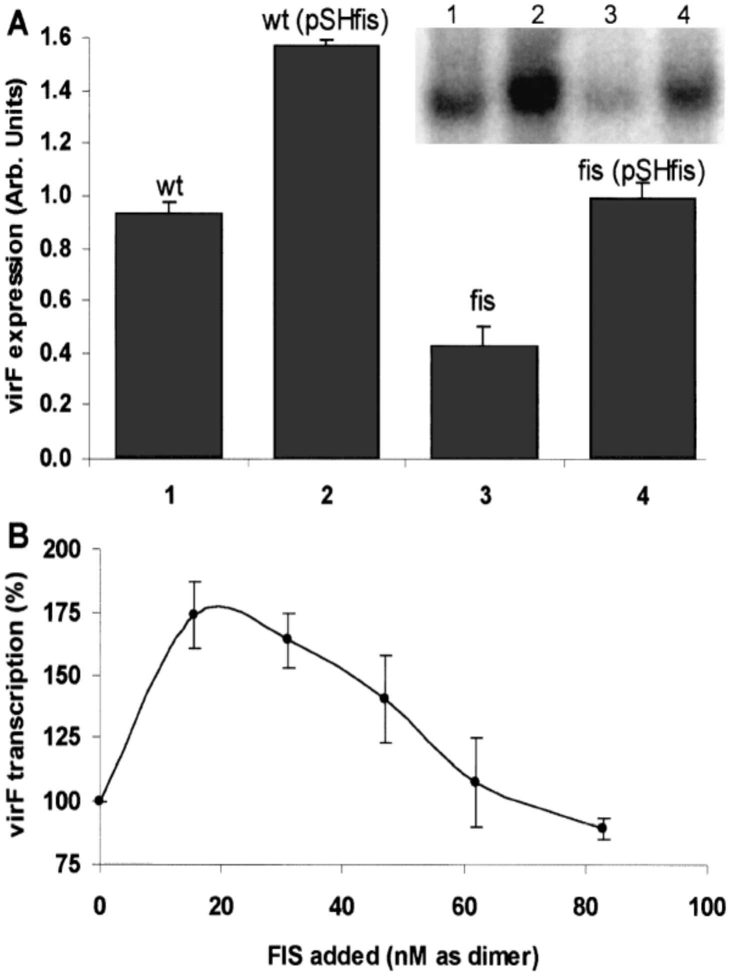


DNAseI footprints reveal that H-NS:

- recognizes a wider site I in the strongly bent mutant (F53)
- is unable to recognize site II in the Fβ+6 mutant



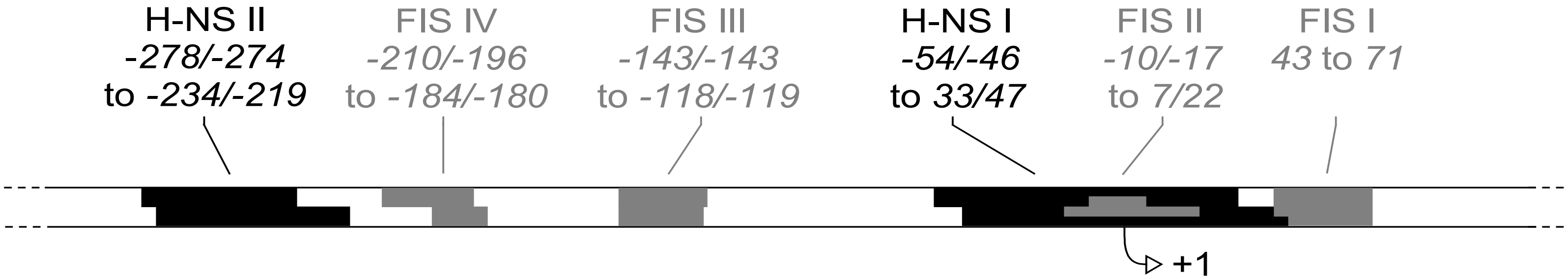
The *virF* gene is positively regulated by **FIS**, another nucleoid protein



FIS (Factor for Inversion Stimulation)

- Basic
- homodimer (2 x 11.5 kDa)
- abundant in exponential phase
- very weak DNA sequence specificity
- transcriptional regulator
- participates in site-specific recombination and transposition

FIS has four binding sites within the *virF* promoter …



H-NS II
-278/-274
to *-234/-219*

FIS IV
-210/-196
to *-184/-180*

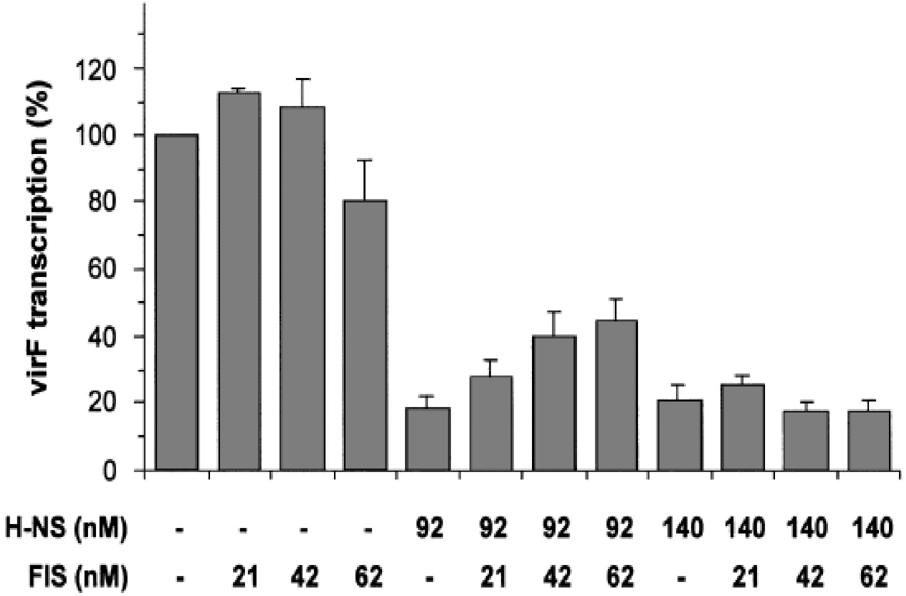
FIS III
-143/-143
to *-118/-119*

H-NS I
-54/-46
to *33/47*

FIS II
-10/-17
to *7/22*

FIS I
43 to *71*

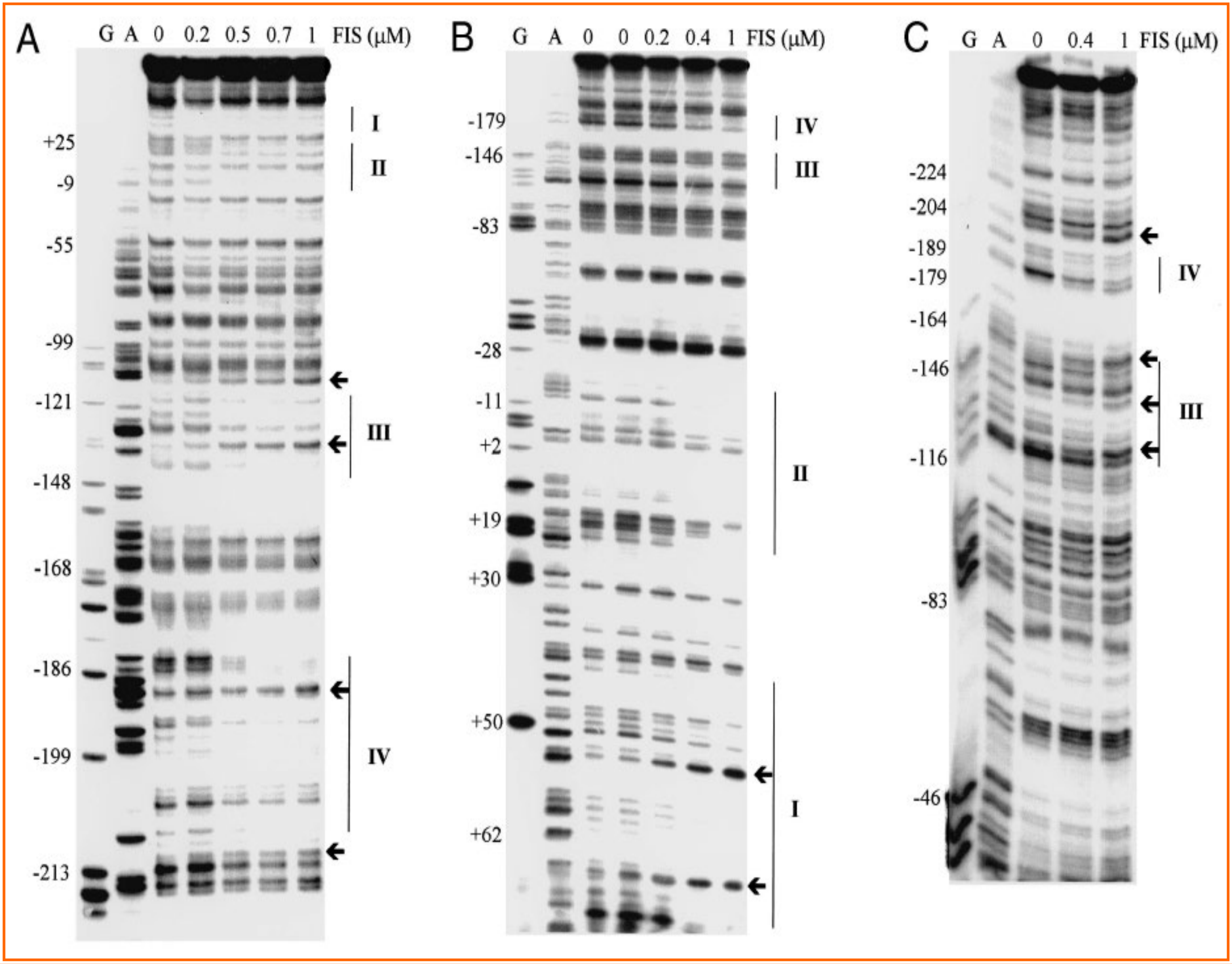
+1



In vitro transcription in the presence of both, H-NS and FIS

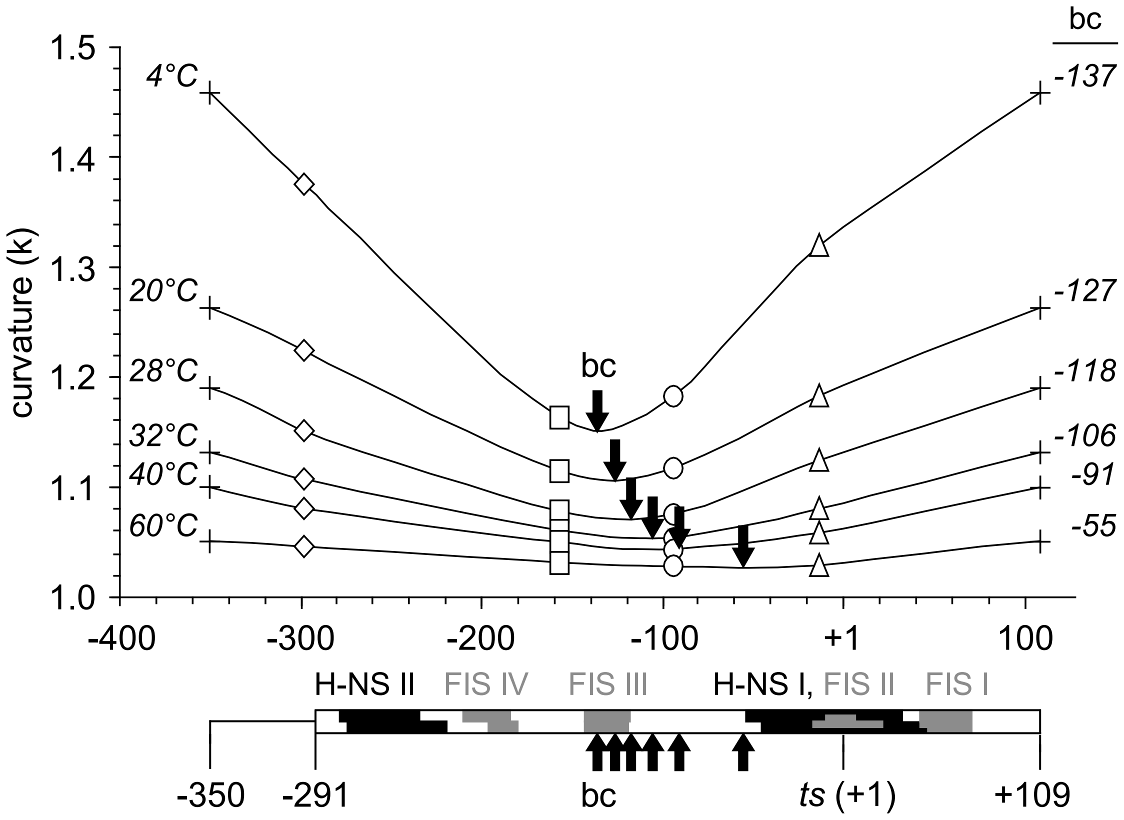
… and alleviates
H-NS-mediated
repression
of the *virF* promoter
at 31°C

Identification of FIS binding sites on the *virF* promoter region



Circular permutation assay on the *virF* promoter region

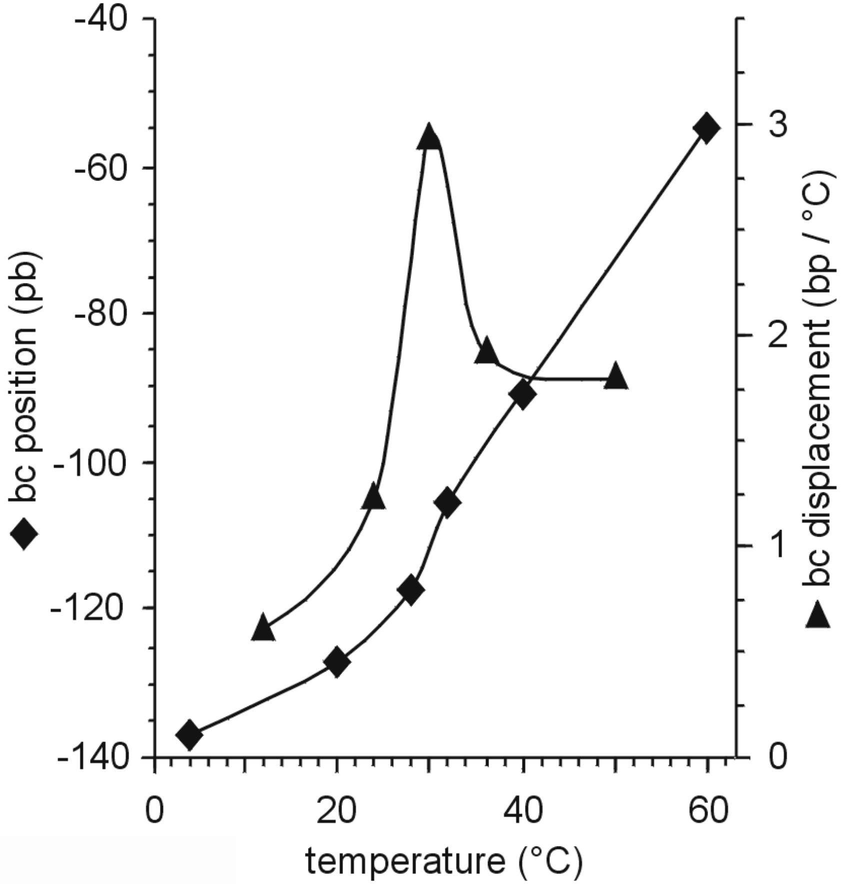
Effect of temperature



Curvature is reduced as temperature increases

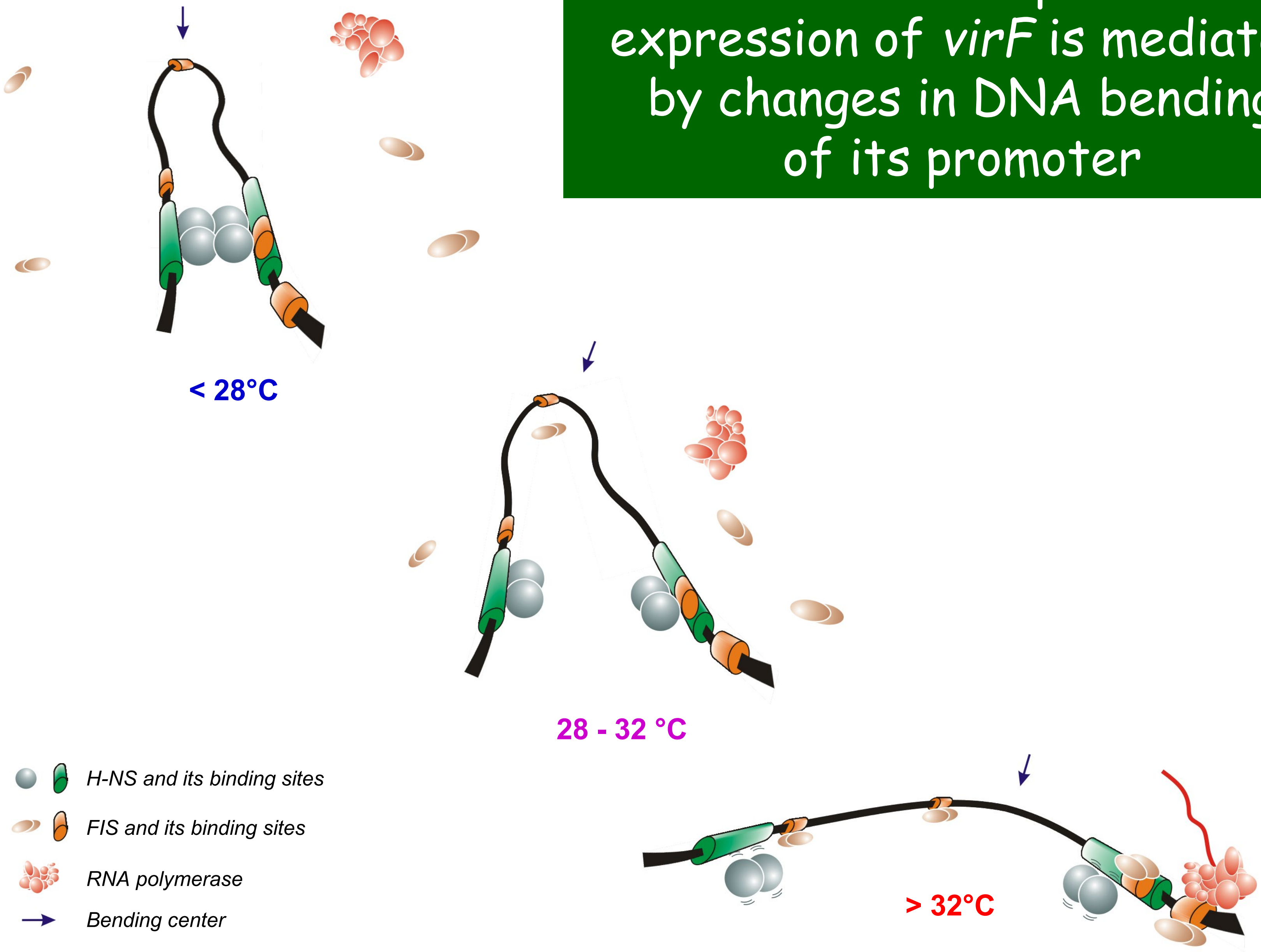
Do temperature changes alter the position of the bending centre?

… yes, the bending centre of the *virF* promoter shifts considerably with temperature

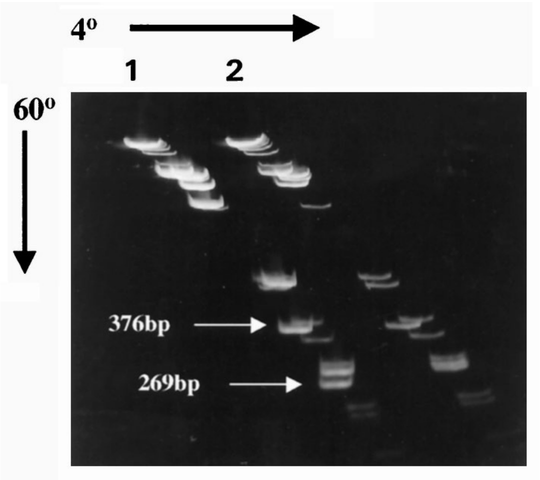


- The displacement of the bending centre is not a linear function of temperature.
- The maximum displacement occurs between 28°C and 32°C.

The thermodependent expression of *virF* is mediated by changes in DNA bending of its promoter

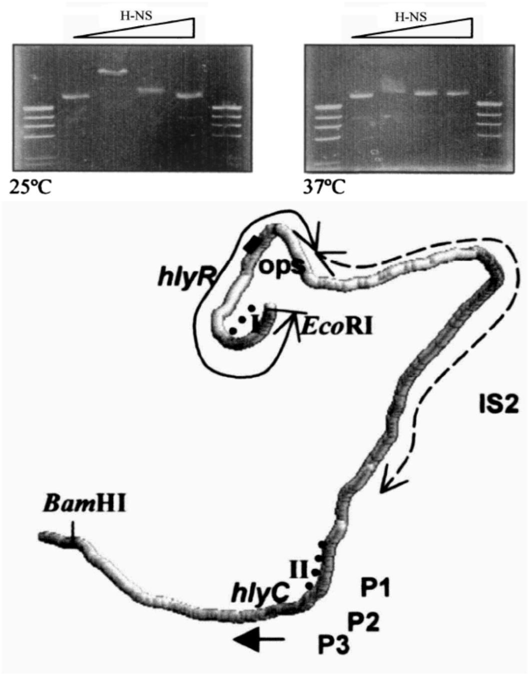


Does DNA curvature regulate virulence genes as a function of host temperature also in other bacterial pathogens?



The *Yersinia enterocolitica* virulence plasmid contains DNA bends which melt at 37°C

(Rohde et al.)



In pathogenic *E. coli* the plasmid hemolysin operon is regulated by temperature-dependent binding of H-NS to curved DNA

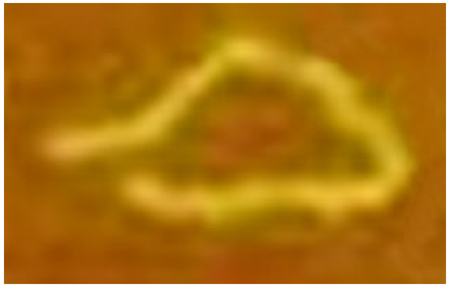
(Madrid et al.)

From pathogenic bacteria to *E. coli* :

How far does intrinsic DNA curvature sustain bacterial transcription?

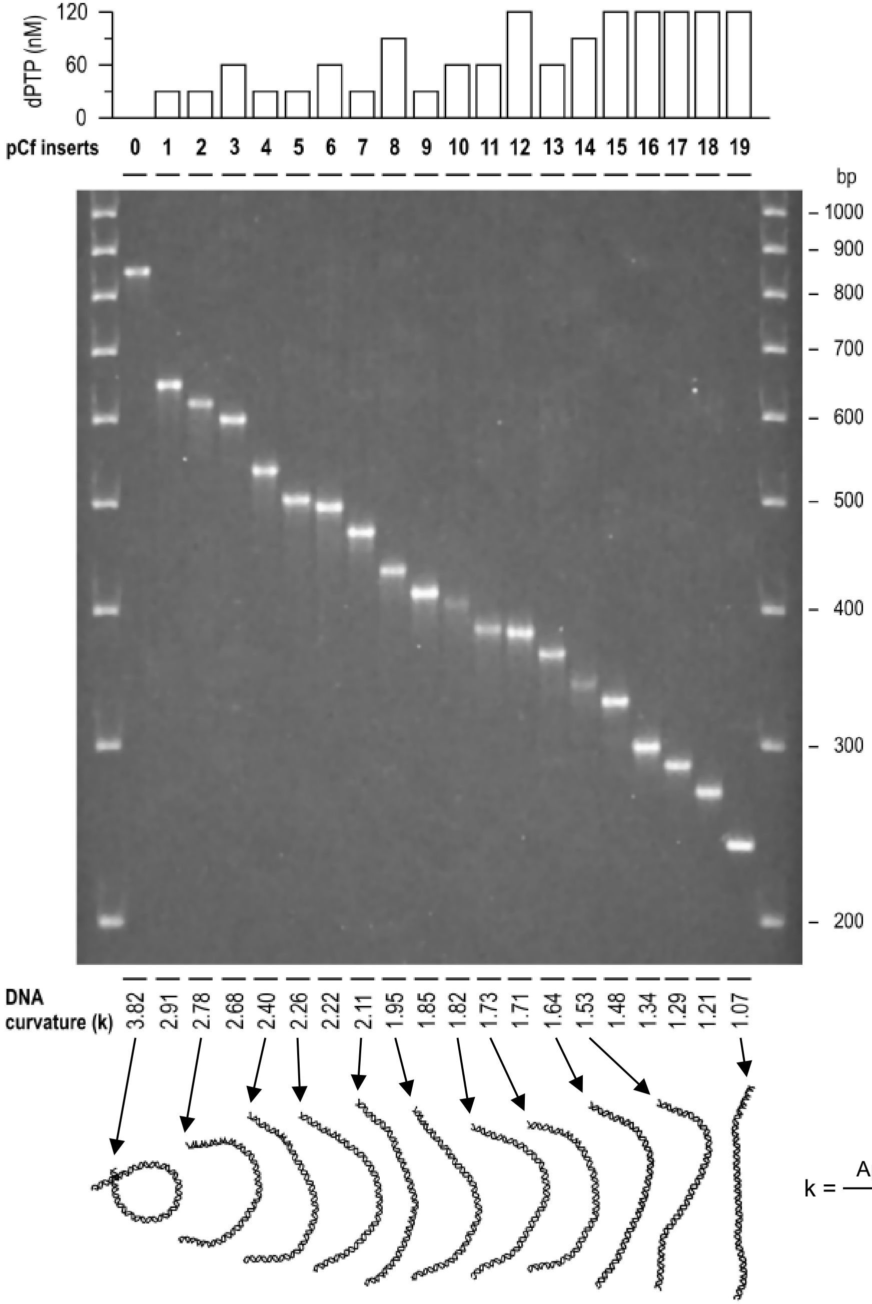
The experimental approach

The 211 bp fragment from the kinetoplast DNA of Crithidia fasciculata, known to be endowed with strong curvature,



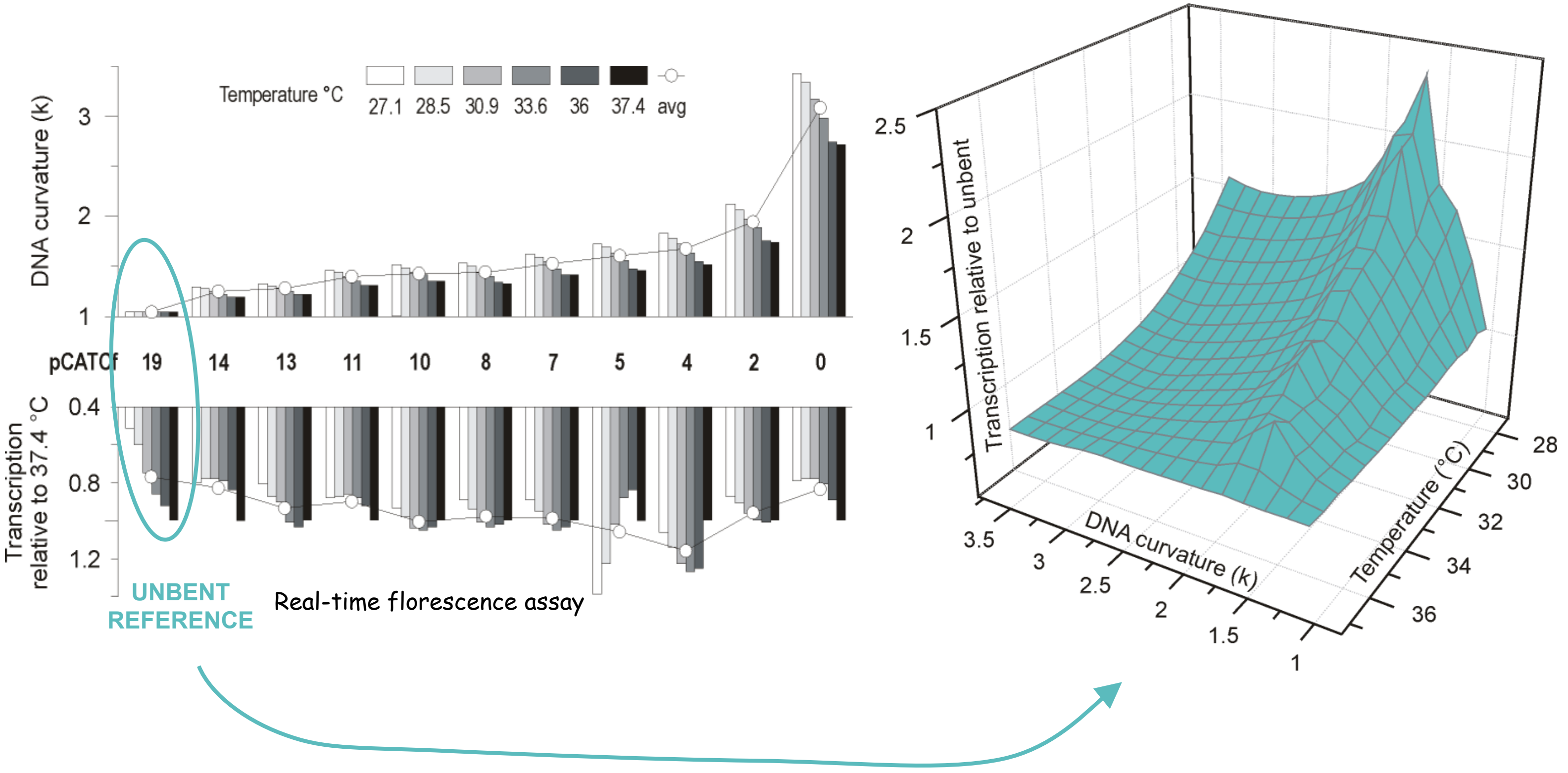
has been randomly mutagenized in order to obtain a spectrum of fragments covering a wide curvature range.

Then, mutagenized fragments were cloned upstream (-45) a reporter gene and …



$$k = \frac{\text{Apparent size}}{\text{Real size}}$$

… their temperature-dependent transcription profiles were analyzed in vitro …



UNBENT REFERENCE

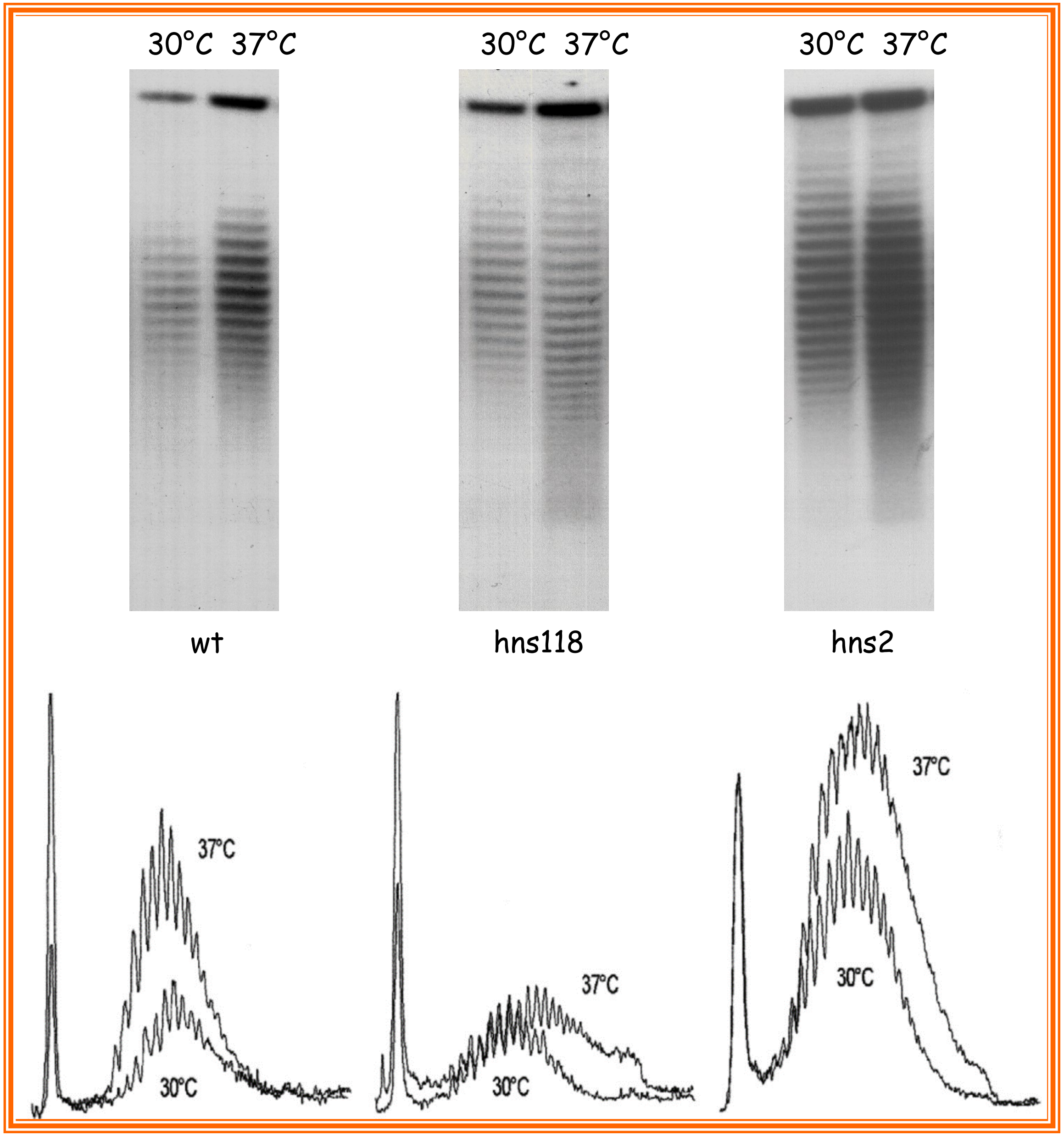
Real-time florescence assay

In short:

Curved DNA regions are frequently located upstream bacterial promoters.

Their marked temperature-sensitivity makes them excellent candidates as transcriptional modulators responding to environmental stimuli.

A narrow range of curvature is able to sustain bacterial transcription in vitro.



Distribution of pMYSH6504 topoisomers in *hns*$^{+}$ and *hns*$^{-}$ strains at 30°C and at 37°C

Influence of FIS on the H-NS mediated repression of the *virF* promoter at 31°C