CHAPTER 11

Gene Regulation by HMGA and HMGB Chromosomal Proteins and Related Architectural DNA-Binding Proteins

Andrew A. Travers

Abstract

The eukaryotic abundant high mobility group HMGA and HMGB proteins can act as architectural transcription factors by promoting the assembly of higher-order protein-DNA complexes which can either activate or repress gene expression. The structural organisation of both classes of protein is similar with either a single or repeated DNA binding domain preceding a short negatively charged C-terminal tail. In the HMGB class of proteins the HMG DNA-binding domain binds non-specifically and introduces a sharp bend into DNA whereas the AT-hook in the HMGA protein binds preferentially to A/T rich regions of DNA and stabilises a B-DNA structure. The acidic tails are hypothesised to facilitate the interaction of the proteins with nucleosomes by binding to the positively charged histone tails. Both classes of protein also interact with a large number of transcription factors that bind to specific DNA sequences.

Introduction

The eukaryotic nucleus contains three classes of abundant chromatin associated proteins – the High Mobility Group proteins of the HMGA, HMGB and HMGN classes (originally termed HMGI/Y, HMG1/2 and the HMG14/17; for recent nomenclature changes see ref. 1), so called because they were initially identified on the basis of their rapid migration through starch gels.[2] The HMGA and HMGB classes of chromosomal proteins in general share some common characteristics, notably a conserved acidic region and the ability to interact with several different transcription factors. A major role is to organise the structure of DNA-protein complexes in the context of chromatin.

Architectural DNA Binding Proteins

In both the eukaryotic nucleus and the bacterial nucleoid the trajectory of the DNA double helix is normally tightly constrained so that not only can the DNA be compacted without entanglement but also to provide an appropriate environment for the enzymatic machinery involved in DNA transcription, replication and recombination. This organisation is normally effected by abundant DNA binding proteins, termed architectural DNA-binding proteins,[3] that either induce DNA bending or facilitate the formation of multicomponent DNA-protein complexes. The term 'architectural' in this context implies that the protein is required for organising DNA but the proteins that fall within this definition are often otherwise functionally distinct and would include, for example, the histone octamer, abundant eukaryotic

DNA Conformation and Transcription, edited by Takashi Ohyama.

chromosomal proteins, such as the HMGA and HMGB proteins, abundant proteins associated with the prokaryotic nucleoid, such as FIS, H-NS, IHF and HU, as well as bona fide transcription factors exemplified by the TATA-binding protein (TBP). Some of these proteins have more than one architectural function. For example FIS can stabilise particular configurations of supercoiled DNA plasmids and also act to promote the assembly and activity of transcription, replication and recombination complexes. Many of the more generalised architectural proteins may be regarded as facilitators and are often not essential for viability while some more 'specialised' proteins, such as the histone octamer and TBP, are clearly essential.

Principles of Transcription Factor Induced Bending

The bending of DNA by transcription factors and by other protein complexes is a major component in the establishment of the overall morphology of protein-DNA complexes. This bending is usually a consequence of indirect readout, a mechanism by which the selectivity of binding is dependent not on making direct contacts between the aminoacids and bases, i.e., direct readout, but instead on the physicochemical properties of the DNA molecule itself.

Recognition of DNA by transcription factors often involves both direct and indirect readout. However, the principles of indirect readout are well illustrated by the histone octamer which, although not a transcription factor itself, completely lacks direct contacts between the aminoacid side chains and the bases of the bound DNA. The octamer binds 147 bp of DNA which are wrapped in a left-handed superhelix with a total curvature of approximate 10 radians.[4] This curvature contrasts with the stiffness of DNA in solution where the average persistence length (P), defined as the length over which the average deflection of the polymer axis caused by thermal agitation is one radian, is 140-150 bp,[5] i.e., the same length as that bound by the histone octamer. For DNA molecules that are not anisotropically curved the affinity of the DNA for the octamer is directly proportional to the flexibility (the inverse of the stiffness).[6] However the dependence of the binding energy on P is some 10-fold lower than the dependence of the bending energy in solution on P. This implies that the histone octamer increases the apparent flexibility substantially to compensate for the average increase in DNA curvature on binding.

How might this change in flexibility be effected? The histone core provides a DNA binding surface in the form of a positively charged ramp. On binding to this ramp the negative charges on one side of the DNA are neutralised. This asymmetric neutralisation, which can be mimicked in free DNA,[7] creates an imbalance in charge distribution on opposite sides of the double helix so that repulsion between the opposing sugar-phosphate backbones on the unneutralised side facilitates bending by increasing the width of the grooves. Concomitantly, the reduction in this repulsion on the inside of the bend permits greater freedom in the motions of the base-pairs, with a corresponding reduction in the width of the grooves. The greater flexibility of the motions between base-pairs is reflected in the periodic variation of twist and roll with groove width such that the ranges of values assumed for both are substantially larger than the corresponding ranges observed for DNA molecules free in solution.[8]

The correlation between flexibility and affinity for the histone octamer only applies strictly when a DNA molecule does not possess intrinsic anisotropic curvature. When it does the affinity may be relatively higher or lower. For example, the intrinsically curved TATA DNA sequence whose curvature is compatible with the surface of the histone octamer binds with an affinity that would normally be characteristic of a substantially more flexible isotropic binding site.[6,9] In this case binding is favoured by the lower entropic penalty on binding relative to an isotropically flexible molecule.[8] However if the intrinsic bend is too great and therefore less compatible with the protein binding surface the affinity is reduced relative to an isotropically flexible molecule.[10]

An extension of this principle of asymmetric alteration of the ionic environment of DNA is provided by the transcription factor TBP and the HMG-domain, found in HMGB proteins, a class of abundant chromosomal proteins and certain transcription factors such as SRY and LEF-1.[11] The HMGB proteins consist essentially of a small L-shaped protein domain with a

cluster of hydrophobic residues on its inner surface and an extended unstructured basic region. When these proteins bind to DNA they produce a bend of 95-120° over about six base-pairs and decrease both the axial and torsional stiffness.[12] On the outer surface of the bend the hydrophobic 'wedge' towards the apex of the L binds in and widens the minor groove, concomitantly untwisting the DNA. This effect is believed to be facilitated by a local reduction in the dielectric constant which increases the repulsion between opposing sugar-phosphate backbones on the approach of the protein to DNA.[13,14] At the same time the basic region neutralises the phosphates bounding the major groove on the inside of the bend thus decreasing the repulsive forces and permitting the narrowing of the groove. Additionally the protein inserts, or intercalates, hydrophobic aminoacids into either a single base-step or into two base-steps that are themselves separated by a single base-step. The extent to which this intercalation increases or simply stabilises the induced bend is unclear. The bend induced by the intercalation contrasts with the smooth DNA bending induced by the histone octamer since the intercalation effectively introduces a kink in the DNA such that the stacking interactions between adjacent base-pairs are very substantially reduced.

In the TATA-binding protein this same principle of hydrophobic interactions predominates. Here two pairs of phenylalanine residues are intercalated at steps separated by 6 bases, kinking the DNA by ~45° at each intercalation site.[15,16] Between these pairs of phenylalanine residues a hydrophobic surface rests snugly within the minor groove. Again the minor groove is widened and untwisted. However, unlike the HMGB proteins there is no charge neutralisation on the opposing major groove face of the bent DNA and indeed the sharpness of the induced curvature is less than that for the HMGB proteins.

In other transcription factors there is substantial variation in the degree of induced bending. The *Escherichia coli* CAP (aka CRP) factor is a good example of mixed direct and indirect readout. This dimeric protein induces a bend of ~45° per monomer.[17] In this case the major bend occurs where the recognition helix of the helix-turn-helix motif binds in the major groove on the inside of the bend, concomitantly making direct contacts with the DNA bases and neutralising the sugar-phosphate backbone in the immediate vicinity.[18] Flanking the central recognition palindrome is a basic ramp which binds DNA and increases the overall DNA bend by indirect readout in a manner analogous to the histone octamer.

Biological Functions of DNA Bending

Although one of the principal roles of DNA bending in the living cell is to maintain the compaction of DNA, it also has important functions in transcriptional control and, in particular, in the assembly of regulatory complexes. A major consequence of introducing a tight bend into DNA is to bring DNA sequences which are far apart on a linear representation of a DNA molecule into close spatial proximity. This effect, which is also characteristic of plectonemically supercoiled DNA, is mediated in chromatin by the HMG-domain transcription factors, such as TCF-1, LEF-1 and SRY. In the case of TCF-1 acting at the enhancer of the *TCR* promoter, the bend induced by the factor brings together a normally unstable complex of the Ets-1 and PEBP2α DNA-binding proteins and ATF/CREB activator proteins to form a stable complex.[19] This example is probably a particular case of the more general phenomenon in which the DNA between a transcription factor and its target protein partner must be bent for protein-protein contacts to occur. The ease of bending will depend critically on the distance and the helical phase difference between the binding sites of the factor and its target. Normally unless one or both of the partner proteins are flexible contact will be facilitated when the binding sites are in helical phase, primarily because of the constraints on the torsional flexibility of DNA. However, at least in vitro, the constraints imposed by both torsional and axial rigidity can be overridden by the abundant DNA-bending proteins of the HMGB class. In the presence one of these proteins a requirement for an integral number of helical turns between binding sites is no longer crucial.[20] Furthermore the involvement of the HMGB protein in the formation of the complex need only be transient.

To what extent are variations in DNA flexibility reflected in genomic organisation? An excellent example of the dependence of biological function on DNA structure is provided by genome of the enteric bacterium *E. coli*. In this organism the strongest promoters for DNA transcription, often those directing the synthesis of rRNA and tRNA, are almost invariably associated with A/T rich, and hence flexible, DNA sequences extending upstream for 100-300 base-pairs from the transcription startpoint.[21] The activity of many of these promoters is strongly dependent on a high negative superhelical density stored in the DNA. This would in principle favour both DNA untwisting at sequences such as TATAAT close to startpoint[22] and also left-handed DNA wrapping around the protein complex responsible for initiating transcription. In many of these highly active promoters the DNA sequence also imparts curvature to the region, a feature that correlates both with the presence of multiple activating binding sites for the abundant DNA bending protein FIS (Factor for Inversion Stimulation).[21,23] These sites are often organised in helical phase such that the binding of FIS could constrain a negative superhelical loop. Indeed in the *rrnA* P1 regulatory region the presence of a far upstream FIS site centred at position -222 from the transcription startpoint results in the constraint of an additional supercoil in the initiation complex.[24] A primary function of this FIS-induced DNA looping is to promote the wrapping of DNA around the RNA polymerase prior to the initiation of transcription,[25] a phenomenon that has also been proposed for the activation of the *lac* promoter by the CAP DNA-bending transcription factor,[26] and consequently to facilitate the extended wrapping characteristic of the open complex.[27] In turn the FIS-induced constraint of negative superhelicity buffers this type of promoter against changes in the unconstrained superhelical density.[24,25,28] In other such promoters an alternative model proposes that the upstream activating sequence contains regions that are highly susceptible to DNA untwisting.[29] For both DNA wrapping and untwisting in the upstream region both models predict that the topological unwinding is transmitted to the 'TATA' sequence and promotes its untwisting.

DNA bending may also be required for the establishment of repressive regulatory complexes. Here, a DNA loop is often formed by the binding of an oligomeric repressor to two sites that are distant from each other along the DNA sequence. This loop, which can be as tightly bent as nucleosomal DNA, prevents the binding of RNA polymerase to the regulated promoter. Examples of this mode of regulation include repression by the AraC, LacI and GalR proteins.[30-33]

HMGA Proteins

The vertebrate HMGA proteins are small proteins of ~100-110 aminoacids and contain tandem copies, usually three, of a characteristic DNA-binding domain, the AT-hook, together with a C-terminal acidic region (Fig. 1).[34] The AT-hook is not restricted to the HMGA proteins as such as it is also found in a related *Drosophila* chromosomal protein, D1, which contains multiple copies of the motif,[35,36] in the motor subunits of various ATP-dependent chromatin remodelling complexes and in certain transcription factors that also contain a primary sequence-specific DNA-binding domain where it is assumed to act as an auxiliary DNA-binding element.[37]

The AT-hook is an unstructured short motif with the consensus sequence Arg/Lys-Pro-Arg-Gly-Arg-Gly-Pro-Arg/Lys[37,38] and selectively binds in the minor groove of A/T-rich regions of DNA.[39,40] The central Arg-Gly-Arg core adopts an extended conformation deep within the groove with the arginine side chains making extensive hydrophobic contacts along the base of the groove.[41] The proline residues change the trajectory of the backbone allowing the basic residues flanking the core mediate electrostatic and hydrophobic contacts with the DNA backbone. When bound to DNA the surface of the core motif contacting the DNA is concave and resembles that of the DNA binding drug netropsin, which has a similar selectivity for A/T-rich sequences.[41] For both netropsin and an AT-hook one consequence of DNA binding is a modest widening of the minor groove with a concomitant stabilisation of B-DNA structure. In some cases this widening results in a change in the direction of bending of DNA, particularly when that bending is dependent on a narrow minor groove width in A/T-rich sequences.[41,42]

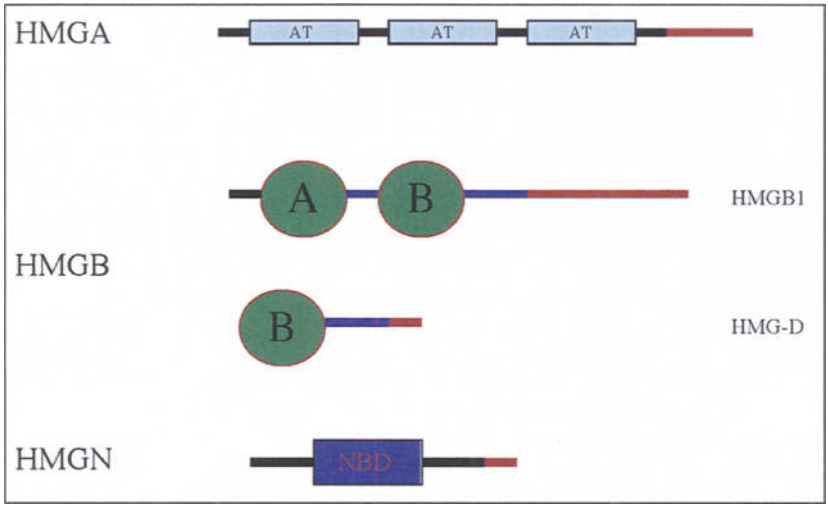

Figure 1. Comparison of the structural organisation of the HMG proteins. Regions of net negative charge outside the principal DNA binding or nucleosome binding domains are indicated in red and those of net positive charge in blue. The AT-hook DNA binding domain (AT) in HMGA proteins and the A- and B-type HMG domains in HMGB proteins are indicated. The nucleosome binding domain (NBD) in HMGN proteins is also shown.

Thus a small intrinsic bend of ~20° towards the minor groove in the IFN-β enhancer is reversed on binding HMGA1.[43] This could then facilitate recognition of the opposing major groove by transcription factors binding to specific sequences. Nevertheless although the HMGA proteins induce only small changes in DNA structure they bind tightly to DNA ligands with distorted or unusual features. These include supercoiled DNA, four-way junctions and base-unpaired regions of AT-rich DNA (reviewed in ref. 34). Strikingly in vitro these proteins can also introduce supercoils into relaxed DNA, possibly by stabilising cross-overs and thereby stabilising DNA loops.[44] This ability has been suggested as an explanation of the observation that in vitro HMGA1 represses the chick globin β^A gene promoter in the absence of the 3' enhancer but strongly activates transcription in its presence, regardless of whether or not the substrate is free DNA or is assembled into nucleosomes.[45] This stabilisation of loops could be mediated by the presence of multiple AT-hooks on each protein.

In mammals the HMGA proteins are encoded by two functional genes, *HMGA1* and *HMGA2*.[34] Alternative splicing of the transcripts of these genes increases the variety of protein products, of which the most abundant are HMGA1a (HMG-I) and HMGA1b (HMG-Y). These proteins appear to perform a variety of functions, of which the most studied are related to chromatin structure and to the facilitation or inhibition of transcription factor binding. In vitro HMGA proteins bind to nucleosomes, notably at the exit and entry points to the nucleosome core particle where they are in close proximity to histones H2A, H2B and H3.[46,47] The proteins can also bind to internal sites where they can induce local changes in the rotational setting of the wrapped DNA. The binding to the nucleosomal DNA is mediated by the AT-hooks but (by analogy to the HMGB class of proteins) it is conceivable that the acidic tail may also be involved in contacting the histone octamer and could perform a similar role to that of the HMGB proteins.

In mitotic chromosomes the HMGA proteins are associated with particular bands and these proteins are localised to the base of large chromatin loops in close proximity to scaffold-attachment regions (SARs).[48] As a consequence it has been suggested that the HMGA proteins are involved in the maintenance of the condensed mitotic chromosome structure in these regions.

Evidence supporting this view was adduced from the observation that synthetic 'MATH' proteins containing AT-hooks interfere with chromosome condensation during mitosis.[49] In contrast another model suggests that co-operative binding of HMGA molecules to a looped chromatin domain in interphase nuclei will facilitate the formation of an 'open' chromatin structure that is competent for transcription by competing with histone H1.[50] Although some experiments demonstrate that the MATH proteins can counteract the spreading of heterochromatin, as shown in particular by suppressing position-effect variegation in flies.[51] The mechanism by which this is accomplished remains to be established. However the expression of the HMGA proteins is strongly correlated with cell growth and is characteristically high in neoplastically transformed cells.[52]

HMGA1 has been shown to interact directly with a large variety of transcription factors including AT-1, ATF-3, NF-Y, IRF-1, SRF, NF-κB, p50, Tst-1/Oct-6 and c-Jun.[34] In some cases the protein regulates the formation of an enhanceosome. Thus at the virus-inducible β-interferon enhancer a complex containing both HMGA proteins and transcription factors forms and then acts to recruit RNA polymerase II and its associated general transcription factors.[53,54] In other cases HMGA proteins can block enhanceosome formation.[55] This modulation of transcription factor binding may be integrated with the regulation of chromatin organisation. Thus HMGA1a enhances the binding of the ATF-3 to a site at the edge of a nucleosome positioned on the HIV-1 promoter.[56] This combination of bound proteins can then recruit the remodelling complex hSWI/SNF. The interactions of HMGA with these factors can be modulated by covalent modifications including phosphorylation and methylation.[34]

HMGB Proteins

HMGB proteins are characterised by the HMG-box, a DNA-binding domain specific to eukaryotes. A major characteristic of this domain is to introduce a sharp bend into DNA (Fig. 2). Accordingly the domain also binds preferentially to a variety of distorted DNA structures, especially those in which the distortion itself induces a bend. These include negatively supercoiled DNA, small DNA circles, cruciforms, DNA bulges and cisplatin modified DNA.[10] The HMG-box domain is also found in several related types of protein, for example transcription factors such as SRY and LEF-1, and subunits of many chromatin remodelling complexes. All these proteins are predominantly nuclear and appear to act primarily as architectural facilitators in the manipulation of nucleoprotein complexes; for example, in the assembly of complexes involved in recombination and the initiation of transcription, as well as in the assembly and organisation of chromatin.

The archetypal HMGB proteins are highly abundant (~10-20 copies per nucleosome in the mammalian nucleus[57]) and often occur in two major forms, HMGB1 and HMGB2, originally termed HMG1 and HMG2, in vertebrates.[1] The two distinguishing features of these highly homologous proteins are two similar, but distinct, tandem HMG-box domains (A and B), and a long acidic C-terminal 'tail', consisting of ~30 (HMG1) or 20 (HMG2) acidic (aspartic and glutamic acid) residues, linked to the boxes by a short, predominantly basic linker (Fig. 1). However the most abundant HMG-box domain proteins in *Saccharomyces cerevisiae*, Nhp6ap and Nhp6bp (non-histone proteins 6A and 6B respectively), contain only a single HMG box, and lack an acidic tail. Likewise the two major HMG-box domain proteins in *Drosophila melanogaster*, HMG-D and HMG-Z, have only a single HMG box but, unlike the yeast proteins, contain a short C-terminal acidic tail in addition to a basic region (Fig. 1). These abundant proteins in yeast and *Drosophila* may be the general functional counterparts of HMGB1 and 2 in vertebrates.

The precise functions of the chromosomal HMGB proteins in vivo for a long time remained obscure. However there is now substantial evidence that they interact directly with both transcription factors and with the histone octamer. These interactions can affect transcription factor access to chromatin either directly or by promoting chromatin remodelling. In the latter case the proteins may facilitate repression or activation.

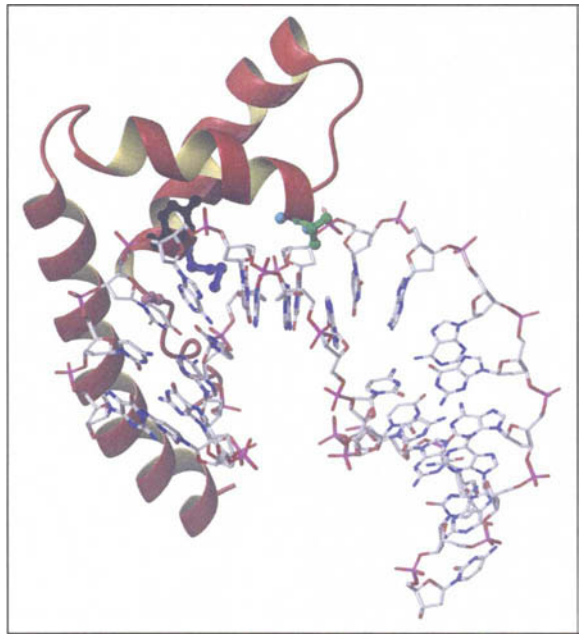

Figure 2. DNA bending by an HMG domain. The figure shows the DNA binding domain of the *Drosophila* High Mobility Group protein HMG-D binding to a short DNA fragment. The protein binds in the minor groove widening the groove and concomitantly stabilising a bend of ~100° in the DNA. This is achieved by inserting the sidechains of hydrophobic aminoacid residues (space-filling representation) between adjacent base-pairs at two locations separated by one base-step. The α-helices of the protein are depicted in red and yellow. The DNA structure shown contains a 'bulge' in which two adjacent bases on one of the strands are unpaired. Reproduced with permission from ref. 88.

There are two established cases in which the assembly of nucleoprotein complexes containing sequence-specific DNA-binding proteins is promoted by the DNA-bending properties of HMGB1 and 2, i.e., the proteins have a classical architectural role. First, in V(D)J recombination the lymphocyte-specific proteins RAG1 and RAG2 (human recombination activating genes 1 and 2) appear to recruit HMGB1 and 2 to the appropriate sites in chromatin[58-62] presumably by protein-protein contacts with the RAG1 homeodomain. Here they ensure the "12/23 rule". This requires that V(D)J recombination occurs only between specific recombination signal sequences (RSS). Each RSS is made up of a conserved heptamer and nonamer sequence separated by a non-conserved spacer of either 12 or 23 base pairs. HMGB1 (in concert with RAG1,2) facilitates recombination probably by bending the DNA between the two conserved sequences spaced by 23 bp and stabilising a nucleoprotein complex. The HMGB protein plays the dual role of bringing critical elements of the 23-RSS heptamer into the same phase as the 12-RSS to promote RAG binding and of assisting in the catalysis of 23-RSS cleavage. Recent footprinting experiments indicate that the HMGB1 (or HMGB2) protein is positioned 5' of the nonamer in 23-RSS complexes, interacting largely with the side of the duplex opposite the one contacting the RAG proteins.[63] A second instance in which an abundant HMGB protein may facilitate nucleoprotein complex assembly is in the formation of an enhanceosome containing the Epstein-Barr virus replication activator protein ZEBRA and HMGB1,[64] the two proteins bind cooperatively, HMGB1 binding to, and presumably bending, a specific DNA sequence between two ZEBRA recognition sites. Bending of DNA by HMGB1 and 2 has also been invoked to explain the essential role of these proteins in initiating DNA replication by loop formation at the MVM (minute virus of mice) parvovirus origin of replication.[65]

In vitro HMGB proteins can enhance the binding of various transcription factors (e.g. adenovirus MLTF, Oct-1 and 2, HoxD9, p53, steroid hormone receptors, Rel proteins, p73, Dof2 and the Epstein-Barr activator Rta) to their cognate DNA binding sites (reviewed in ref. 66). Similarly rat SSRP1 has been shown to facilitate the DNA binding of serum response factor[67] and human SSRP1 is associated with the γ isoform of p63 in vivo at the endogenous *MDM2* and *p21*$^{waf1/cip1}$ promoters.[68] In most of these cases, the interaction of the HMG protein with the transcription factor has been detected in vitro and could, in principle, serve as the mechanism for recruitment of HMGB1 or 2 to particular DNA sites. In some cases transfection experiments indicate functional interactions in vivo. Direct interactions between Nhp6p and the Gal4p and Tup1p transcription factors have also been inferred in vivo by a split-ubiquitin screen and confirmed by a pull-down assay.[69] Although the demonstrated interactions in vitro so far involve an HMGB protein and a single transcription factor, it is entirely possible that in vivo, in a natural regulatory context, the bending of DNA by HMGB1 and 2 could potentially allow the recruitment of a second transcription factor to the complex, in an analogous manner to the action of sequence-specific HMG-box transcription factors[1] such as LEF-1 in the enhanceosome at the T cell receptor alpha (TCRα).[70] HMGB1 may play a catalytic, chaperone role, since it does not appear to be stably incorporated into the final complex. Although a role for HMGB1 and HMGB2 induced DNA bending in the facilitated binding of transcription factors, while being entirely plausible, has not been directly established, it is strongly suggested by the ability of HU to substitute for the HMGB1-stimulated binding of the Epstein Barr virus transactivator Rta to its cognate binding sites.[71] A possible role for an HMGB1-induced change in DNA conformation in facilitation of transcription factor binding is also suggested by the observation that HMGB1 promotes binding of p53 to linear DNA but not to 66 bp DNA circles.[72] However, in this case the data do not distinguish between possible effects of DNA bending or untwisting.

The biological roles of the HMGB proteins have been studied using gene knock-outs. In mice the loss of HMGB1 but not of HMGB2 is lethal although in the former knock-out there are pleiotropic effects on glucose metabolism while in the latter spermatogenesis is impaired.[73,74] This suggests a functional redundancy between members of the HMGB1 and 2 family. A similar situation occurs with Nhp6ap and Nhp6bp in yeast.[75] However, the different phenotypes of the HMGB1 and HMGB2 null mice probably reflect specific roles for the two proteins in different tissues.[73,74] In *S. cerevisiae* the transcriptional effects of *NHP6* are not general but gene-specific. At the *CHA1* locus, loss of *NHP6* results both in an increase in the basal level of transcription and in a substantial decrease in the induced level.[76] This suggests an effect at the level of chromatin. The *CHA1* regulatory region contains a positioned nucleosome which occludes the TATA box under non-inducing conditions. On induction the TATA region becomes accessible.[77] However in the mutant strain, consistent with the increased basal level transcription, the chromatin structure of the TATA region in the uninduced state is similar to that in the induced wild-type strain. *NHP6* thus appears to be required for establishment of the organised chromatin structure characteristic of the uninduced state. The RSC remodelling complex is also required for this process[78] suggesting that RSC and Nhp6p may cooperate to remodel chromatin.

Further insights into how Nhp6ap and Nhp6bp function were provided by studies on the *HO* gene.[79] Loss of *NHP6* function can be suppressed by mutations that increase nucleosome accessibility and mobility, and enhanced by those with the opposite effect. Mutations both in the *SIN3* and *RPD3* genes, encoding components of a histone deacetylase complex, and in *SIN4*, partially restore wild-type function in cells lacking both Nhp6ap and Nhp6bp, while loss of the histone acetylase Gcn5p (also a component of the SAGA histone acetylase complex) in the same cells results in a more severe phenotype. Rpd3p and Gcn5p contribute to the dynamic balance between histone acetylation and deacetylation.[80] Both histone acetylation and the *sin* (SWI/SNF independence) phenotype are correlated with chromatin unfolding[81,82] and/or enhanced nucleosome accessibility[83] while histone deacetylation would be expected to

favour folding. On this argument one role of the Nhp6p proteins would be to antagonise folding and possibly promote nucleosome accessibility.

The ability of the HMGB proteins to promote both transcription factor binding to their cognate sites and also chromatin remodelling implies that these activities could be coordinated to alter chromatin structure in the vicinity of a factor binding site. Like the HMGA proteins the abundant HMGB proteins bind to nucleosomes at sites close to the DNA exit and entry points. An insight into how HMGB proteins might alter the accessibility of nucleosomal DNA was provided by the observation that HMGB1 could facilitate the binding and subsequent remodelling function of the ACF remodelling complex in vitro.[84] Further observations showed that HMG-D, a *Drosophila* HMGB protein, when bound to nucleosome core particles increased the accessibility of nucleosomal DNA to restriction endonucleases at particular sites.[85] These sites were asymmetrically distributed, one site being located at one end of the bound DNA and the other in the vicinity of the nucleosome dyad. This effect required the acidic tail of the HMGB protein: without it the HMG-D reduced accessibility at all sites tested on the nucleosome. This result argues that certain HMGB proteins can alter the structure of nucleosomes and to do so presumably by interacting with an available basic region of the histone octamer. From the distribution of the sites with increased accessibility a prime candidate would be one (but not both) of the N-terminal tails of histone H3 or one of the C-terminal tails of histone H2A. It is important to note that the yeast Nhp6 proteins lack an acidic region and so could not interact with histones directly in this way. However they can associate with two other proteins, Pob3p and Spt16p, to form a complex, SPN, involved in chromatin remodelling.[86,87] Both these proteins contain extensive acidic regions and so, in principle, could substitute for the lack of an acidic region in Nhp6p.

Concluding Remarks

The abundant HMGA and HMGB chromosomal proteins share several common features. Both interact with nucleosomes, both can also bind a set of transcription factors, both are involved in enhanceosome formation and both can facilitate the recruitment of chromatin remodelling complexes. Interestingly the HMGN class of HMG proteins shares with the HMGA and HMGB classes the ability to interact with nucleosomes and also possesses a C-terminal region with a net negative charge.

References

1. Bustin M. Revised nomenclature for high mobility group (HMG) chromosomal proteins. Trends Biochem Sci 2001; 26:152-153.
2. Johns EW, Forrester S. Studies on nuclear proteins. The binding of extra acidic proteins to deoxyribonucleoprotein during the preparation of nuclear proteins. Eur J Biochem 1969; 8:547-551.
3. Grosschedl R, Giese K, Pagel J. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. Trends Genet 1994;10:94-100.
4. Luger K, Mäder AW, Richmond RK et al. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 1997; 389:251-260.
5. Hagerman PJ. Flexibility of DNA. Annu Rev Biophys Biophys Chem 1988; 17:265-286.
6. Travers A. The structural basis of DNA flexibility. Phil Trans R Soc Lond A 2004; in press.
7. Strauss JK, Maher LJ 3rd. DNA bending by asymmetric phosphate neutralization. Science 1994; 266:1829-1834.
8. Richmond TJ, Davey CA. The structure of DNA in the nucleosome core. Nature 2003; 423:145-150.
9. Widlund HR, Cao H, Simonsson S et al. Identification and characterization of genomic nucleosome-positioning sequences. J Mol Biol 1997; 267:807-817.
10. Anselmi C, Bocchinfuso G, De Santis P et al. Dual role of DNA intrinsic curvature and flexibility in determining nucleosome stability. J Mol Biol 1999; 286:1293-1301.
11. Scipioni A, Anselmi C, Zuccheri G et al. Sequence-dependent DNA curvature and flexibility from scanning force microscopy images. Biophys J 2002; 83:2408-2418.
12. Thomas JO, Travers AA. HMG1 and 2, and related architectural DNA-binding proteins. Trends Biochem Sci 2001; 26:167-174.

13. Travers AA. Reading the minor groove. Nat Struct Biol 1995; 2:615-618.
14. Elcock AH, McCammon JA. The low dielectric interior of proteins is sufficient to cause major structural changes in DNA on association. J Am Chem Soc 1996; 118:3787-3788.
15. Kim JL, Nikolov DB, Burley SK. Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 1993; 365:520-527.
16. Kim Y, Geiger JH, Hahn S et al. Crystal structure of a yeast TBP/TATA-box complex. Nature 1993; 365:512-520.
17. Zinkel SS, Crothers DM. Comparative gel electrophoresis measurement of the DNA bend angle induced by the catabolite activator protein. Biopolymers 1990; 29:178-181.
18. Shultz SC, Shields GC, Steitz TA. Crystal structure of a CAP-DNA complex: the DNA is bent by 90°. Science 1991; 253:1001-1007.
19. Giese K, Cox J, Grosschedl R. The HMG domain of the lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. Cell 1992; 69:185-195.
20. Ross ED, Hardwidge PR, Maher LJ, 3rd. HMG proteins and DNA flexibility in transcription activation. Mol Cell Biol 2001; 21:6598-6605.
21. Pedersen AG, Jensen LJ, Brunak S et al. A DNA structural atlas for *Escherichia coli*. J Mol Biol 2000; 299;907-930.
22. Drew HR, Weeks JR, Travers AA. Negative supercoiling induces spontaneous unwinding of a bacterial promoter. EMBO J 1985; 4:1025-1032.
23. Travers A, Muskhelishvili G. DNA microloops and microdomains: a general mechanism for transcription activation by torsional transmission. J Mol Biol 1998; 279:1027-1043.
24. Rochman M, Aviv M, Glaser G et al. Promoter protection by a transcription factor acting as a local topological homeostat. EMBO Rep 2002; 3:335-360.
25. Pemberton IK, Muskhelishvili G, Travers AA et al. FIS modulates the kinetics of successive interactions of RNA polymerase with the core and upstream regions of the *E. coli tyrT* promoter. J Mol Biol 2002; 318:651-663.
26. Buc H. Mechanism of activation of transcription by the complex formed between cyclic AMP and its receptor in *Escherichia coli*. Biochem Soc Trans 1986; 14:196-199.
27. Rivetti C, Guthold M., Bustamante C. Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex. EMBO J 1999; 18:4464-4475.
28. Auner H, Buckle M, Deufel A et al. Mechanism of transcriptional activation by FIS: role of core promoter structure and DNA topology. J Mol Biol 2003; 331:331-344.
29. Hatfield GW, Benham CJ. DNA topology-mediated control of global gene expression in *Escherichia coli*. Annu Rev Genet 2002; 36:175-203.
30. Dunn TM, Hahn S, Ogden S et al. An operator at -280 base pairs that is required for repression of the *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc Natl Acad Sci USA 1984; 81:5017-5020.
31. Lyubchenko YL, Shlyakhtenko LS, Aki T et al. Atomic force microscopic demonstration of DNA looping by GalR and HU. Nucleic Acids Res 1997; 25:873-876.
32. Krämer H, Niemolle M, Amouyal M et al. *lac* repressor forms loops with linear DNA carrying two suitably placed *lac* operators. EMBO J 1987; 6:1481-1491.
33. Law SM, Bellomy GR, Schlax PJ et al. In vivo thermodynamic analysis of repression with and without looping in *lac* constructs. Estimates of free and local *lac* repressor concentrations and of physical properties of a region of supercoiled plasmid DNA in vivo. J Mol Biol 1993; 230:161-173.
34. Reeves R. Molecular biology of HMGA proteins: hubs of nuclear function. Gene 2001; 277:63-81.
35. Levinger L, Varshavsky A. Protein D1 preferentially binds A + T-rich DNA in vitro and is a component of *Drosophila melanogaster* nucleosomes containing A + T-rich satellite DNA. Proc Natl Acad Sci USA 1982; 79:7152-7156.
36. Ashley CT, Pendleton CG, Jennings WW et al. Isolation and sequencing of cDNA clones encoding *Drosophila* chromosomal protein D1. A repeating motif in proteins which recognize at DNA. J Biol Chem 1989; 264:8394-8401.
37. Aravind L, Landsman D. AT-hook motifs identified in a wide variety of DNA-binding proteins. Nucleic Acids Res 1998; 26:4413-4421.
38. Reeves R, Nissen MS. The AT-DNA-binding domain of mammalian high mobility group I chromosomal proteins: a novel peptide motif for recognizing DNA structure. J Biol Chem 1990; 265:8573-8582.
39. Churchill MEA, Travers AA. Protein motifs that recognize structural features of DNA. Trends Biochem Sci 1991; 16:92-97.
40. Solomon M, Strauss F, Varshavsky A. A mammalian high mobility group protein recognizes any stretch of six A-T base pairs in duplex DNA. Proc Natl Acad Sci USA 1986; 83:1276-1280.

41. Huth JR, Bewley CA, Nissen MS et al. The solution structure of an HMG-I(Y) DNA complex defines a new architectural minor groove binding motif. Nat Struct Biol 1997; 4:657-665.
42. Goodsell DS, Kopka ML, Dickerson RE. Refinement of netropsin bound to DNA: bias and feedback in electron density map interpretation. Biochemistry 1995; 34:4983-4993.
43. Falvo JV, Thanos D, Maniatis T. Reversal of intrinsic DNA bends in the IFN β gene enhancer by transcription factors and the architectural protein HMG I(Y). Cell 1995; 83:1101-1111.
44. Nissen MS, Reeves R. Changes in superhelicity are introduced into closed circular DNA by binding of high mobility group protein I/Y. J Biol Chem 1995; 270:4355-4360.
45. Bagga R, Michalowski S, Sabnis R et al. HMG I/Y regulates long-range enhancer-dependent transcription on DNA and chromatin by changes in DNA topology. Nucleic Acids Res 2000; 28:2541-2550.
46. Reeves R, Nissen MS. Interaction of high mobility group-I(Y) nonhistone proteins with nucleosome core particles. J Biol Chem 1993; 268:21137-21146.
47. Reeves R, Wolffe AP. Substrate structure influences binding of the non-histone protein HMG-I(Y) to free and nucleosomal DNA. Biochemistry 1996; 35:5063-5074.
48. Saitoh Y, Laemmli UK. Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. Cell 1994; 76:609-622.
49. Strick R, Laemmli UK. SARs are *cis* DNA elements of chromosome dynamics: synthesis of a SAR repressor protein. Cell 1995; 83:1137-1148.
50. Zhao K, Kas E, Gonzalez E et al. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. EMBO J 1993; 12:3237-3247.
51. Girard F, Bello B, Laemmli UK et al. In vivo analysis of scaffold-associated regions in *Drosophila*: a synthetic high-affinity SAR binding protein suppresses position effect variegation. EMBO J 1998; 17:2079-2085.
52. Tallini G, Dal Cin P. HMGI(Y) and HMGI-C dysregulation: a common occurrence in human tumors. Adv Anat Pathol 1999; 6:237-246.
53. Agalioti T, Lomvardas S, Parekh B et al. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-β promoter. Cell 2000; 103:667-678.
54. Kim TK, Maniatis T. The mechanism of transcriptional synergy of an in vitro assembled interferon-β enhanceosome. Mol Cell 1997; 1:119-129.
55. Klein-Hessling S, Schneider G, Heinfling A et al. HMG I(Y) interferes with the DNA binding of NF-AT factors and the induction of the interleukin 4 promoter in T cells. Proc Natl Acad Sci USA 1996; 93:15311-15316.
56. Henderson A, Holloway A, Reeves R et al. Recruitment of SWI/SNF to the human immunodeficiency virus type I promoter. Mol Cell Biol 2004; 24:389-397.
57. Duguet M, de Recondo AM. A deoxyribonucleic acid unwinding protein isolated from regenerating rat liver. Physical and functional properties. J Biol Chem 1978; 253:1660-1666.
58. Sawchuk DJ, Weis-Garcia F, Malik S et al. V(D)J recombination: modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA-bending proteins. J Exp Med 1997; 185:2025-2032.
59. van Gent DC, Hiom K, Paull TT et al. Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J 1997; 16:2665-2670.
60. Kwon J, Imbalzano AN, Matthews A et al. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. Mol Cell 1998; 2:829-839.
61. West RB, Lieber MR. The RAG-HMG1 complex enforces the 12/23 rule of V(D)J recombination specifically at the double-hairpin formation step. Mol Cell Biol 1998; 18:6408-6415.
62. Aidinis V, Bonaldi T, Beltrame M et al. The RAG1 homeodomain recruits HMGB1 and HMGB2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. Mol Cell Biol 1999; 19:6532-6542.
63. Swanson PC. Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. Mol Cell Biol 2002; 22:1340-1351.
64. Ellwood KB, Yen YM, Johnson RC et al. Mechanism for specificity by HMG-1 in enhanceosome assembly. Mol Cell Biol 2000; 20:4359-4370.
65. Cotmore SF, Christensen J, Tattersall P. Two widely spaced initiator binding sites create an HMG1-dependent parvovirus rolling-hairpin replication origin. J Virol 2000; 74:1332-1341.
66. Travers AA, Thomas JO. Chromosomal HMG-box proteins. In: Zlatanova J, Leuba SH, eds. Chromatin structure and dynamics: state-of-the-art. New Comprehensive Biochemistry. Amsterdam: Elsevier Science, 2004:103-134.
67. Spencer JA, Baron MH, Olson EN. Cooperative transcriptional activation by serum response factor and the high mobility group protein SSRP1. J Biol Chem 1999; 274:15686-15693.

68. Zeng SX, Dai MS, Keller DM et al. SSRP1 functions as a co-activator of the transcriptional activator p63. EMBO J 2002; 21:5487-5497.
69. Laser H, Bongards C, Schüller J et al. A new screen for protein interactions reveals that the *Saccharomyces cerevisiae* high mobility group proteins Nhp6A/B are involved in the regulation of the GAL1 promoter. Proc Natl Acad Sci USA 2000; 97:13732-13737.
70. Giese K, Kingsley C, Kirshner JR et al. Assembly and function of a TCRα enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. Genes Dev 1995; 9:995-1008.
71. Mitsouras M, Wong B, Arayata C et al. The DNA architectural protein HMGB1 displays two distinct modes of action that promote enhanceosome assembly. Mol Cell Biol 2002; 22:4390-4401.
72. McKinney K, Prives C. Efficient specfic DNA binding by p53 requires both its central and C-terminal domains as revealed by studies with High-Mobility Group 1 protein. Mol Cell Biol 2002; 22:6797-6808.
73. Calogero S, Grassi F, Aguzzi A et al. The lack of chromosomal protein HMGB1 does not disrupt cell growth but causes hypoglycaemia in newborn mice. Nat Genet 1999; 22:276-279.
74. Ronfani L, Ferraguti M, Croci L et al. Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2. Development 2001; 128:1265-1273.
75. Costigan C, Kolodrubetz D, Snyder M. NHP6A and NHP6B, which encode HMG1-like proteins, are candidates for downstream components of the yeast SLT2 mitogen-activated protein kinase pathway. Mol Cell Biol 1994; 14:2391-2403.
76. Moreira JMA, Holmberg S. Chromatin-mediated transcriptional regulation by the yeast architectural factors NHP6A and NHP6B. EMBO J 2000; 19:6804-6813.
77. Moreira JM, Holmberg S. Nucleosome structure of the yeast *CHA1* promoter: analysis of activation-dependent chromatin remodeling of an RNA-polymerase-II-transcribed gene in TBP and RNA pol II mutants defective in vivo in response to acidic activators. EMBO J 1998; 17:6028-6038.
78. Moreira JM, Holmberg S. Transcriptional repression of the yeast *CHA1* gene requires the chromatin-remodeling complex RSC. EMBO J 1999; 18:2836-2844.
79. Yu Y, Eriksson P, Stillman DJ. Architectural factors and the SAGA complex function in parallel pathways to activate transcription. Mol Cell Biol 2000; 20:2350-2357.
80. Verdone L, Wu J, van Riper K et al. Hyperacetylation of chromatin at the *ADH2* promoter allows Adr1 to bind in repressed conditions. EMBO J 2002; 21:1101-1111.
81. Tse C, Sera T, Wolffe AP et al. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol 1998; 18:4629-4638.
82. Horn PJ, Crowley KA, Carruthers LM et al. The SIN domain of the histone octamer is essential for intramolecular folding of nucleosomal arrays. Nat Struct Biol 2002; 9:167-171.
83. Anderson JD, Lowary PT, Widom J. Effects of histone acetylation on the equilibrium accessibility of nucleosomal DNA target sites. J Mol Biol 2001; 307:977-985.
84. Bonaldi T, Längst G, Strohner R et al. The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. EMBO J 2002; 21:6865-6873.
85. Ragab A, Travers A. HMG-D and histone H1 alter the local accessibility of nucleosomal DNA. Nucleic Acids Res 2003 31:7083-7089.
86. Formosa T, Eriksson P, Wittmeyer J et al. Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. EMBO J 2001; 20:3506-3517.
87. Brewster NK, Johnston GC, Singer RA. A bipartite yeast SSRP1 analog comprised of Pob3 and Nhp6 proteins modulates transcription. Mol Cell Biol 2001; 21:3491-3502.
88. Cerdan R, Payet D, Yang JC, Travers AA, Neuhaus D. HMG-D complexed to a bulge DNA: an NMR model. Protein Sci 2001;10:504-518.