Journal of Cellular Biochemistry 50:111–123 (1992)

PROSPECT

Homeotic Protein Binding Sites, Origins of Replication, and Nuclear Matrix Anchorage Sites Share the ATTA and ATTTA Motifs

Teni Boulikas

Linus Pauling Institute of Science and Medicine, Palo Alto, and Institute of Molecular Medical Sciences, Palo Alto, California 94306

Abstract Nuclear matrix organizes the mammalian chromatin into loops. This is achieved by binding of nuclear matrix proteins to characteristic DNA landmarks in introns as well as proximal and distal sites flanking the 5′ and 3′ ends of genes. Matrix anchorage sites (MARs), origins of replication (ORIs), and homeotic protein binding sites share common DNA sequence motifs. In particular, the ATTA and ATTTA motifs, which constitute the core elements recognized by the homeobox domain from species as divergent as flies and humans, are frequently occurring in the matrix attachment sites of several genes. The human apolipoprotein B 3′ MAR and a stretch of the Chinese hamster DHFR gene intron and human HPRT gene intron shown to anchor these genes to the nuclear matrix are mosaics of ATTA and ATTTA motifs. Several origins of replication also share these elements. This observation suggests that homeotic proteins which control the expression level of many genes and pattern formation during development are components of the nuclear matrix. Thus, the nuclear matrix, known as the site of DNA replication, might sculpture the crossroads of the differential activation of origins during development and S-phase and the control of gene expression and pattern formation in embryogenesis. © 1992 Wiley-Liss, Inc.

Key words: nuclear matrix, origins of replication, homeodomain proteins, transcription factors, development, chromatin, ATTA motif

NUCLEAR MATRIX, TRANSCRIPTIONAL ENHANCERS, AND THE DOMAIN CHROMATIN STRUCTURE

The 25 million or so nucleosomes in a single mammalian nucleus are organized into 60,000 chromatin loops by the interaction of specific sequences of DNA with the nuclear matrix proteins [Paulson and Laemmli, 1977; Mirkovitch et al., 1984; Gasser and Laemmli, 1986]. Such landmarks occur within introns of the various genes [Cockerill and Garrard, 1986; Käs and Chasin, 1987; Sykes et al., 1988], as well as within proximal and distal sites flanking the 5′ and 3′ ends [Bode and Maass, 1988; Phi-Van and Strätling, 1988; Levy-Wilson and Fortier, 1989]. Table I summarizes the AT-rich sequences which are abundant in the matrix attachment region of several genes that have been studied. A striking finding is that the MAR of the human apolipoprotein B gene is a mosaic of ATTA, ATTTA, and other related motifs repeated over a 555 bp stretch. Similar motifs are present in virtually all MAR sequences determined (Table I), with the exception of the *Xenopus* ribosomal RNA gene repeat MAR [Marilley and Gassend-Bonnet, 1989] and that of human Ha-*ras* oncogene [Boulikas et al., 1992]. These data reinforce the idea that some noncoding sequences, such as MARs, may represent repetitive mass binding sites for regulatory nonhistone proteins; the presence of a large domain with a number of weak binding sites can lead to the formation on a strong DNA-protein complex [Zuckerkandl and Villet, 1988], especially when one single high-affinity binding site is present in the complex [Scheuermann, 1991]. Such high-affinity protein-DNA complexes are presumably formed when total nuclear matrix proteins interact with specific fragments of DNA [e.g., Cockerill and Garrard, 1986].

MAR sequences have been demonstrated to have transcriptional enhancer activity [Xu et al., 1989; Blasquez et al., 1989a; Stief et al., 1989;

Received June 3, 1992; accepted June 16, 1992.

Address reprint requests to Teni Boulikas, Institute of Molecular Medical Sciences, 460 Page Mill Rd., Palo Alto, CA 94306.

© 1992 Wiley-Liss, Inc.

TABLE I. Nuclear Matrix Anchorage Regions (MARs)

DNA sequence	Gene and location	Reference
TAATTAAATATTTTATAATTAAAATATTTA The entire MAR is composed of this 30-nucleotide motif, repeated 18.5 times	Human apolipoprotein B gene 3′ proximal MAR	Levy-Wilson and Fortier [1989]
187 TAAATTATATAACTAATTTTAATTATAAAATTAA CACTAATTATAATAGCATTATTAATGAAATTAACATT GATTATAATAAAACAACATTAATAATAAAATTAAC	Chinese Hamster dihydrofolate reductase gene MAR (within an intron)	Käs and Chasin [1987]
CTTAATGTATATAATCTTTTAGAGGTAAAATC-20nt- GTAAATATTCTTTGA-21nt-CTAAATTATATGTCATAT TAACTGGTTAAATTAATATAAATTTGT-26nt- ATATTTTT-8nt-AAAAATATGACTAATAATAATTTA GCACAAAAAATATT-8nt-CTTTAATTC-38nt- CTATAATCCCATAATTTTGAAAACTATTTATTAG-33nt- CTATTTA (topoisomerase II sites are underlined; a sequence common to several MARs is in grey)	DNA sequence motifs with resemblance to homeotic protein binding sites within a 364 bp stretch of the JC intron of the mouse κ immunoglobulin gene	Blasquez et al. [1989b]
129 CTATTAGCTTTAATTATTTTTACATTTAG-19nt- CATAATTTTG-24nt-CAAATTATC-43nt-GAATTTTAA 6nt-TGAATATAAAA-50nt-TATTACTCATTACTAGGAAT TAAATAAAAT-40nt-TCATTTTTATCTTTTATTTTTG TTTGTTTTTTTTG (shaded areas represent yeast ARS)	DNA sequence motifs from the human hypoxanthine phosphoribosyl transferase gene MAR (within the first intron)	Sykes et al. [1988]
CCCATTATTGG CCTAAATG GGTTTTAAGTAATGTATTCTCTATTA GTGAATAATGATGTT CAATTAAATGTCCAC (topoisomerase II sites are underlined)	Motifs from 3′ MAR of chicken α-globin gene	Farache et al. [1990]
−471 AATAAATAAA	*Drosophila* Sgs-4 gene MAR	Gasser and Laemmli [1986]
−583 AATAAATAAT −712 AATAAATAAA −957 TTATTTTATA −942 TTTTATTATT	*Drosophila* histone H1–H3 intergenic region MAR	Gasser and Laemmli [1986]
−914 AATAAATAAA −346 TTATATTATT	*Drosophila* adult alcohol dehydrogenase gene MAR	Gasser and Laemmli [1986]
−4159 AATAAATAAT −4190 AATAATATATA −3763 TTTTATTATT	*Drosophila* larval alcohol dehydrogenase gene MAR	Gasser and Laemmli [1986]
+260 TATAAATAAA −516 TAATTTTTTT	*Drosophila* Hsp70 87A7 gene MAR	Gasser and Laemmli [1986]

Phi-Van et al., 1990; Klehr et al., 1991]. This, together with earlier findings showing that transcription takes place on the nuclear matrix [e.g., Ciejek et al., 1983; see Boulikas, 1987], seems to suggest that understanding the nature of DNA sequences that are attached to the nuclear matrix is important for a comprehensive view of the regulation of gene expression in eukaryotic cells.

We may view transcriptional enhancers as elements, permanently attached to the nuclear matrix. The promoter region of their genes might be anchored to the nuclear matrix via hydrophobic interactions between nonhistone proteins bound to the enhancer and promoter elements and looping out of the DNA, first proposed by Mirkovich and co-workers [1984; see Fig. 2 in Boulikas, 1992]. This idea divides nuclear ma-

trix attachment sites into permanent (enhancers, origins of replication) and temporal (related to the transcriptional and repair activity of the genes as well as developmental stage). The constitutive or permanent class of MARs is present in all cell types irrespective of the transcriptional activity of the genes involved and seems to include origins of replication [e.g., Razin et al., 1986]. Usually one such MAR is found in the 5′ far upstream site and a second MAR in the 3′ downstream site of a gene, delineating the boundaries of its functional chromatin domain. Examples are the 19 kb locus of the chicken lysozyme gene [Phi-Van and Strätling, 1988] and the 49 kb locus of the human apolipoprotein B gene [Levy-Wilson and Fortier, 1989].

The facultative (or functional or transient) class of MARs seems to depend on the transcriptional and repair activity of the gene. Indeed, repair activities are also restricted to the nuclear matrix [McCready and Cook, 1984; Mullenders et al., 1988]. The facultative type of MAR is therefore detected only in cell types harboring the particular gene in an active chromatin structure. This type of MAR is located between nucleotides −2.7 and −1.8 kb to the 5′ side of the human apolipoprotein B gene [Levy-Wilson and Fortier, 1989].

Similarly, two types of nuclear matrix structures have been described: 1) An external part, lining the interior of the nuclear envelope, with lamins A, B, and C as the major components, apparently mediating the topological compartmentalization of the peripheral chromatin [see Hancock and Boulikas, 1982; Boulikas, 1987; Berezney, 1991]. Direct contacts between lamins and DNA, as determined by crosslinking with UV [Boulikas, 1986b], along with lamin-nucleosome interactions [Yuan et al., 1991], are the major determinants of this rather unspecific chromatin binding to the peripheral nuclear matrix. 2) An internal nuclear matrix component composed of thick, highly branched polymorphic fibers which might serve as the core structure for the internal matrix architecture [e.g., He et al., 1990].

Among the thousands of nuclear matrix proteins, few have been identified and well characterized to date. These include topoisomerase II [Berrios et al., 1985], the chicken ARBP protein, which recognizes the consensus sequence ATTTCA$^{C}_{G}$TTGTAAAA in the MAR of the chicken lysozyme locus [von Kries et al., 1991], the yeast ACBP protein which interacts with the ARS element [Hofmann and Gasser, 1991], matrin 3, an acidic protein of the rat internal nuclear matrix network displaying a 96% identity with the human matrin 3 [Belgrader et al., 1991], matrin F/G [Hakes and Berezney, 1991], lamins A, B, and C, which form the peripheral nuclear matrix [Fisher et al., 1986; McKeon et al., 1986; Osman et al., 1990], and several others.

Transcription factors have been described that possess a highly negatively charged amino acid domain. Such acidic factors include HMG 1 and 2 which contain a contiguous stretch of 41 negatively charged amino acid residues [Walker et al., 1980] and which act as general class transcription factors [Tremethick and Molloy, 1986, 1988], the human and *Xenopus* UBF transcription factors that contain a contiguous stretch of 55 negatively charged amino acids [Jantzen et al., 1990; Bachvarov and Moss, 1991] and the SPT5 and SPT6 gene products in yeast that are essential transcription factors and display an extremely high density of negative charges in their N-terminal segments [Swanson et al., 1990, 1991].

Matrin 3, a component of the internal nuclear matrix network, similarly possesses an acidic domain [Belgrader et al., 1991]. This finding supports our proposal that a major class of transcription factors is represented by components of the nuclear matrix that might be involved in the removal of histones from DNA via their negatively charged domain [Boulikas, 1992]. A similar function has been attributed to poly(ADP-ribosyl)ated molecules in nuclei which are highly negatively charged and appear at specific sites on the nuclear matrix whenever strand breaks occur on the DNA [see Boulikas, 1992]. Nucleosome dissolution into histones and DNA is a well-established process found to occur at the promoter, but not coding, region of genes during transcription activation, thus making the DNA accessible to RNA polymerase II and accessory proteins of the transcription initiation complex [e.g. Richard-Foy and Hager, 1987; Archer et al., 1991]. This process was proposed to occur on facultative MARs [Boulikas, 1992]. However, most matrix proteins, via their hydrophobic interactions, might simply cause the looping out of DNA between two matrix protein binding sites [von Kries et al., 1991].

The nature of the DNA that anchors the various genes to the nuclear matrix has been the subject of vigorous studies. Its size varies from

about 360-bp [Blasquez et al., 1989b] to 1 kb [Jackson et al., 1990] and 7 kb [Mielke et al., 1990]. Such figures show the same variation as the size of replication origins that are in the range of 35-bp for the smallest reiterated sequence found by Fangman and coworkers [1989] to be autonomously replicated in yeast mitochondria, to 75-bp or longer for some yeast ARS [Broach et al., 1983], to 160-bp and 290-bp for the SV40 and polyoma virus ORIs, respectively [Hay and DePamphilis, 1982; Hendrickson et al., 1987], to 6 kb for the origin of replication of the amplified *THFR* gene in Chinese hamster cells [Caddle et al., 1990] [although the core origin of this gene seems to lie within a 450 bp fragment; see Burhans et al., 1990]. Similarly, the two origins used for the amplification of chorion genes in *Drosophila* map within a rather broad zone of the chorion gene domain [Delidakis and Kafatos, 1989].

The nuclear matrix DNA harbors AT-rich sequences [e.g., Mirkovitch et al., 1984; see Table 1] with intrinsically curved motifs [Anderson, 1986; Homberger, 1989]. The studies of Razin and collaborators [1978], Matsumoto [1981] and Neuer-Nitsche and co-workers [1988] imply that satellite DNA might be involved in the nuclear matrix structure. The studies of Chimera and Musich [1985] indicate that a subset, not all sequences, of the *Kpn*I long interspersed repetitive DNA is associated with the nuclear matrix. Some *Alu* sequences may also be MAR elements [Small et al., 1982]. On the other hand, the blot hybridization experiments of Mirkovitch and coworkers [1986] suggest that the scaffold attachment sites of the *Drosophila* chromosome are within unique sequences.

HISTONE H1 HAS NUCLEAR MATRIX ACTIVITIES

An unexpected twist in the nuclear matrix field is the finding that histone H1 recognizes the same type of sequences as total nuclear matrix proteins, causing their precipitation in vitro [Izzauralde et al., 1989; Boulikas et al., 1992]. Histone H1 locks the two helical turns of the DNA around the nucleosome [Boulikas et al., 1980] and maintains higher-order chromatin structures [Thoma et al., 1979; see Boulikas, 1992]. Histone H1 is the most abundant repressor of gene activity [Croston et al., 1991; Laybourn and Kadonaga, 1991]. Even physically unconnected mononucleosomes are assembled into higher-order chromatin structures by exogenous H1 [Boulikas, 1986a]. This is so in spite of the presence of histone H1 within active genes [Grossbach et al., 1990], but in stoichiometric amounts about two-fold lower than in inactive genes [Kamakaka and Thomas, 1990; Postnikov et al., 1991].

Supposedly one copy of histone H1 per nucleosome [Reudelhuber et al., 1980] promotes higher-order chromatin structures where the 1–72 amino acid domain of histone H1 interacts with three core histones of neighboring nucleosomes and the 73–106 amino acid segment of H1 contacts histone H2A of its "own" nucleosome [Boulikas et al., 1980; Boulikas, 1992]; lower stoichiometric amounts of histone H1 might break the cooperative H1-H1 interactions [Kamakaka and Thomas, 1990] and the H1-neighboring nucleosome interactions [Boulikas, 1992] that stabilize the 30 nm chromatin fiber. Interactions between histone H1 and protein transcription factors or nuclear matrix proteins have not been studied to a significant extent. These interactions could participate in the domain chromatin structure and in the remodeling of the attachment points of chromatin loops to the nuclear matrix believed to occur during development.

HMG I, an activator of gene expression and histone H1 which compete with one another for binding to nucleosomes [Boulikas and Deschênes, in preparation], share an 11 amino acid DNA binding domain peptide (TPKRPRGRPKK) called AT-hook, which is directed toward the minor groove of DNA [Reeves and Nissen, 1990; Reeves et al., 1991]. This property may determine the specificity of histone H1 for binding to distinct restriction fragments from gene clones [Berent and Sevall, 1984; Ristiniemi and Oikarinen, 1989] that seem to be the same fragments preferred by nuclear matrix proteins [Izaurralde et al., 1989; Boulikas et al., 1992]. It remains to be determined what type of interactions between histone H1, HMG I, and matrix proteins prevail at attachment sites of chromatin loops of specific genes in vivo.

HOMEOTIC PROTEINS AND DEVELOPMENT

The homeodomain was first described as a conserved 61 amino acid segment of *Drosophila* proteins that regulate development [McGinnis et al., 1984a,b; Laughon and Scott, 1984; Scott and Weiner, 1984]. Most homeobox genes that encode for homeotic proteins possess, in their far upstream, immediate upstream, or downstream regulatory regions, binding sites for ho-

TABLE II. Homeodomain Protein Recognition and Binding Sites From DNase I Footprints

DNA Sequence	Protein which binds	Reference
ATTA	Core sequence in recognition sites of homeodomain proteins. Recognized by the conserved amino acids Arg-3, Arg-5, Ile-47, and Asn-51	Odenwald et al. [1989]
CAATTA CCATTA	Homeodomain proteins with Glu-50	Hanes and Brent [1991]
GGATTA	Homeodomain proteins with Lys-50	Hanes and Brent [1991]
TAATGAnAT	Oct-1	Herr et al. [1988]
TTAAAATTCA	Oct-3, Oct-4	Schöler et al. [1989], Okamoto et al. [1990]
AACAATTACAAA GGCAATTAAACT (*Xenopus*Hbox 2 gene promoter)	Hbox 1 *Xenopus* homeotic protein	Cho et al. [1988]
CGTTTTATTAGG CATTAATC TATAATC CATAAAATTTTTATTG AGGCATAATATCATTAC (human HOX4C homeotic gene promoter)	HOX4D human homeotic protein	Zappavigna et al. [1991]
GGTTAATnATTAAC	Hepatic nuclear factor 1 (HNF 1) homeodomain protein	Frain et al. [1989], De Simone et al. [1991]
TAAATTAAATGTCAATTAAATATCAATCAATT-21nt-ACATTTAACTGGTTAATTGAAG 5′ flanking region (~ −700) of the *Drosophila engrailed* gene) Consensus: TCAATTAAAT	Engrailed *Drosophila* homeotic protein	Desplan et al. [1988]
TCAATTAAAT-59bp-CCAATTAGCC-19bp-CTAATTAGAG within intron 1 of *engrailed* gene	Engrailed fusion protein binding site	Desplan et al. [1988]
(TAATAA)n (TAATCG)n in clusters of 40–90 bp far upstream of the *Antennapedia* gene	Ultrabithorax *Drosophila* homeotic protein	Beachy et al. [1988]
GGGATTAGA upstream of *hunchback* gene	*Drosophila bicoid* gene product, (Bicoid protein) necessary for establishing anterior-posterior polarity	Driever and Nusslein-Volhard [1989]
GATTTTTTAATG (*bithorax* region enhancer)	*Drosophila hunchback* gene product	Qian et al. [1991]
CAATTAAATATCAATCAATTTC CATTTAACTGGTTAATTG TCATTTAAATTAAAAC (5′ regulatory region of *Drosophila engrailed* gene)	*Drosophila even-skipped* gene product	Hoey and Levine [1988]

continued

TABLE II. Homeodomain Protein Recognition and Binding Sites From DNase I Footprints(continued)

DNA Sequence	Protein which binds	Reference
ATGGTCATAAATCAAATTGT Distal upstream enhancer (−1,400) of dopa decarboxylase gene	Footprint of the *Drosophila Cf1α* POU domain gene product, a neuron-specific transcription factor	Johnson and Hirsh [1990]
GGTACATTAATGGATTATCCCTTTAAATGT GCGA Upstream region of the a-specific STE6 yeast gene	Binding site of the yeast α2 protein, product of the MATα2 gene a negative reductor of a-specific genes	Johnson and Herskowitz [1985]

meotic proteins (see Table II). This strongly suggests the presence of a complex network of auto- and cross-regulatory interactions among homeotic proteins and homeobox genes. A relatively large number of other cellular genes, especially those whose expression pattern is altered during development, are expected to contain homeotic protein binding sites in their regulatory regions. The interplay of homeotic proteins with other transcriptional protein factors, their number of copies in the nucleus, and their affinity for a given regulatory site would determine which proteins will occupy the regulatory regions of a gene. Variations in this are expected among different cell types and stages of development [McGinnis and Krumlauf, 1992].

The structure of the homeodomain-DNA complex has been solved by X-ray crystallography [Kissinger et al., 1990]. The homeodomain is responsible for the recognition of the ATTA core nucleotide sequence on DNA involving two arginine residues at positions 3 and 5 which make contacts with the two consecutive thymines of the ATTA motif in the minor groove. Contacts in the major groove of DNA involve an isoleucine at position 47 (which makes a hydrophobic contact with the methyl group of a thymine), an asparagine at position 51 (which establishes two hydrogen bonds with an adenine), and a glutamine at position 50 (which makes van der Waals contacts with a thymine methyl group but also with several different positions near the 5′ side of the ATTA motif, mostly including the CAATTA and CCATTA motifs) [Odenwald et al., 1989; Kissinger et al., 1990]. Replacement of glutamine at position 50 with lysine directs the homeodomain to the GGATTA core element [Hanes and Brent, 1991]. Thr-6, Tyr-25, Agr-31, Trp-48, Arg-53, Lys-55, and Lys-57 establish contacts with the six phosphate groups of DNA over a stretch of 9 bp [Kissinger et al., 1990].

DNase I footprinting experiments using purified homeotic proteins and regulatory sequences of homeotic genes have identified their cognate binding sites on the DNA (Table II). One major class of homeodomain proteins include in their core recognition sequences the ATTA [see Scott et al., 1989], ATTTA, and ATTTTA motifs, as well as their complementary TAAT, TAAAT, and TAAAAT motifs (Table II). For example, such motifs are recognized by the gene products of the *Drosophila engrailed* and *even-skipped* [Desplan et al., 1988; Hoey and Levine, 1988], *Antennapedia* [Müller et al., 1988], and *hunchback* [Qian et al., 1991]. The homeotic gene complex *ultrabithorax* (*Ubx*), located in the bithorax complex of *Drosophila,* encodes a family of closely related proteins that direct the fates of one another, of other homeotic genes, and very possibly of a number of other developmentally controlled or cell type-specific genes. One of the Ubx proteins has been shown to bind to tandem repeats of the trinucleotide TAA or the related hexanucleotide TAATCG which both occur in clusters, 40–90 bp in size, far upstream of the *Antennapedia* gene [Beachy et al., 1988].

In addition, several mammalian homeotic proteins recognize the above-mentioned core motifs including the Oct-1 transcription factor [Herr et al., 1988], the mouse Hox-1.3 homeotic protein [Odenwald et al., 1989], the homeodomain-containing hepatic nuclear transcription factor 1 (HNF 1) [Frain et al., 1989], and the human HOX 4D homeotic protein [Zappavigna et al., 1991]. The *Xenopus* Hbox 1 homeotic protein displays a similar specificity [Cho et al., 1988].

The helix-turn-helix motif that is present in the homeodomain of homeotic proteins [Qian et al., 1989] binds to the major and minor groove of DNA [see Triesman et al., 1992]. Homeotic proteins form dimers [Kissinger et al., 1990], like other transcription factors [e.g., Marmorstein et al., 1992], with the second helix of each helix-turn-helix motif fixed on either half in the major groove of a palindromic recognition sequence [Pabo and Sauer, 1984]. X-ray crystallography

of homeotic protein-DNA complexes [Kissinger et al., 1990] has revealed that the 61 amino acid homeodomain interacts directly with about 12 bp of DNA centered around the ATTA core motif. However, the C- and N-terminal parts of the intact homeotic protein might bind a stretch of DNA longer than 12 bp [Kissinger et al., 1990].

The size of the DNase I footprint of a single transcription or replication protein factor on the DNA is about 24 nucleotides (nt) for the E1 protein on the ORI of bovine papillomavirus [Ustav et al., 1991], 26 nt for the T antigen on SV40 ORI [Hay and DePamphilis, 1982], 24–30 nt for the Xenopus Hbox 1 protein [Cho et al., 1988], and 18 nt for the Oct-3 factor [Okamoto et al., 1990]. The six proteins that interact with one another on the promoter region of the mouse albumin gene occupy all together a stretch in length of 133 bp of DNA, with individual sites 19–25 nt in some instances overlapping by 2–6 nt [Lichtsteiner et al., 1987]. DNA methylation protection experiments show that a protein factor binds tightly to and protects guanines and adenines over a stretch of 9 nt against dimethyl sulfate [e.g., Müller et al., 1988]. X-ray crystallography of the yeast transcriptional activator GAL4 shows that the protein binds as a dimer to a symmetrical 17 bp sequence [Marmorstein et al., 1992] and that the single 61 amino acid homeodomain recognizes 9 bp of DNA [Kissinger et al., 1990]. It might thus be expected that the center-to-center distance between ATTA and ATTTA motifs would be in the range of 10–30 nt, to allow for interaction of the proteins bound to them with one another. This is consistent with the spacing of these motifs on MARS (Table I).

ORIGINS OF REPLICATION ARE ATTACHED TO THE NUCLEAR MATRIX

Origins of replication appear associated with the nuclear matrix [Berezney and Coffey, 1975; Pardoll et al., 1980; Aelen et al., 1983; Razin et al., 1986; Carri et al., 1986; Dijkwel and Hamlin, 1988; Vaughn et al., 1990]. The activation of a fraction of the 60,000 or so potential ORIs present in a single mammalian nucleus are tightly linked to the transcriptional activity of neighboring genes. Active genes are thought to be replicating from an origin of replication that is 5′ to the gene in cell types expressing this gene; on the contrary, an origin of replication downstream to the gene is used for its duplication in cell types not expressing this gene [Trempe et al., 1988; Leffak and James, 1989]. Replicons and supercoiled chromatin loop domains seem to coincide in their sizes [Buongiorno-Nardelli et al., 1982]. Examples may be found, however, where a large domain known to be a single replicon is divided into smaller looped DNA domains. Brown and collaborators [1987] have shown that the 300 kb murine immunoglobulin heavy chain gene locus is a single replicon; yet the studies of Cockerill [1990] clearly show that 200 kb of this domain are divided into four looped domains of 30, 20, 30, and greater than 70 kb in length.

The data in Table III suggest that specific regions, shown in some cases to interact with regulatory proteins and found within the minimal sequence that functions as an origin of replication, contain ATTA and ATTTA motifs. These data suggest that the differential activation of origins which occurs during development might indeed be regulated on the nuclear matrix.

CONCLUSION

The observation presented here is that nuclear matrix attachment sites found in specific AT-rich regions of genes that are responsible for the domain structure of chromatin possess the ATTA core DNA sequence element and other similar core DNA sequence elements that are recognition sites of homeodomain proteins. The same core motifs are found in the origins of replication of yeast, viruses, mitochondria, chloroplasts, and some defined mammalian origins of replication (Table III). There is a lack of vigorous DNase I footprinting studies of individual (or of a mixture of) nuclear matrix proteins bound to the DNA showing that some of them specifically interact with or include in their site of interaction, the ATTA core motif. The present study suggests that such proteins exist. Indeed the MAR of the human apolipoprotein B gene is almost entirely composed of a contiguous stretch of 555 bp made of a mosaic of the TAAT, TAAAT, ATTA, ATTTTA, TAAAAT, and ATTTA motifs [Table I; Levy-Wilson and Fortier, 1989]. This stretch needs to interact with unidentified nuclear matrix proteins in order to anchor the 3′ flanking region of the gene to the nuclear matrix.

The conclusion of this study is that the nuclear matrix controls the differential gene expression during development. This adds a new dimen-

TABLE III. Origins of Replication

DNA Sequence	Type of origin	Reference
ATTTATATTTA T G T	*S.cerevisiae* autonomously replicating sequence (ARS)	Broach et al. [1983]
TAAT-31nt-ATTAATATAT-35nt-ATTTTTTTAT GTTTTTTTAAAACATTAAAG Thick bar shows a palindrome; the 11 bp stretch homologous to ARS consensus is underlined	Yeast ARS	Figure 1 in Broach et al. [1983]
ATTTATTTG-31nt-TATTAAA-24nt-ATTTAATACCTAAATATAAAAAA TGTTATTATATTG	Yeast ARS	Figure 2 in Broach et al. [1983]
ATAAATATTAATTAA ATTAATATTTAATTAA	*S.cerevisiae* HMR-I ARS	Amati and Gasser [1988]
TTTTATATTTAGGTA	*S.cerevisiae* HMR-E ARS	Amati and Gasser [1988]
ATTTATATTTAGTAA	*S.cerevisiae* ARS2	Amati and Gasser [1988]
AATTTATTTAA	*S.pombe* ARS	Umek et al. [1989]
CATCATCAATAATATACCTTATTTTGGAT TGAAGCCAATATGATAATGAGGGGG	Adenovirus 2 and 4 ORS	Hay [1985], Pruijn et al. [1986]
TTTATTTATTAATTATTAATATTTT GGGGAATTTTAGG	Sequences within the ARS2 of *Tobacco* chloroplast genome	Shinozaki et al. [1986]
ATAATTTGGAAAAT CTAATAAAAATGTGATTTTG	Sequences within the ARS1 of *Tobacco* chloroplast genome	Shinozaki et al. [1986]
AATAAATAAA AGTAAATAAA	Yeast histone H4 gene origin of replication	Amati and Gasser [1988]
TGAATAATTGTTGTTAACAATAATC	Bovine papillomavirus (BPV) origin of replication	Ustav et al. [1991]
GAATTATTTC CCATTACCGG GTTCATTAGC GGTTAATTTTCA	Motifs from the origin of replication of the human *c-myc* gene	Iguchi-Ariga et al. [1988]
GAAAAAATAATAATAATTG GATTTACTGG	Putative origin of replication ORS 17 of an unknown gene from monkey CV-1 cells	Landry and Zannis-Hadjopoulos [1991]
TTCTTTATAACTATATTATG GAAAAAATTATTTTG	ORS24	Landry and Zannis-Hadjopoulos [1991]
CTTAAATTGGG TGATAAATCA ATTCTTAATATTTTG	ORS 25	Landry and Zannis-Hadjopoulos [1991]
550 TAAAATAAATTAAAAAATTA ATTAATTAAAAAAG 1816 GTTAATATTAT-42nt-CATTATAAG-31nt-ACATTTAGATTAG-9nt-TAAAATACA 3412 GTAAATAGTCTTAAAAGCCCATAAAAT GACTCAAAACTAGTTTTTTTATTATTATTATTAG (motifs directing DNA curvature are underlined) 3765 $(TA)_{23}$-10nt-TATTTATGCATATAAAAA TAAAATAAATT$_6$CT$_{10}$	Sequences from the 6 kb ORI of the amplified Chinese hamster DHFR	Caddle et al. [1990]

continued

TABLE III. Origins of Replication (continued)

DNA Sequence	Type of origin	Reference
GATTATATATATAATATATTATTGG	ORI of Marek's disease virus (MDV) causing malignant T-cell lymphoma in chickens	Camp et al. [1991]
CCAATATATATATATATTATTAGG	HSV-1 oris	See Camp et al. [1991]
CCAATAATATATATATATTATTAGG	HSV-1 oriL	See Camp et al. [1991]
TATTATTATATATATATATATTATAT TATATAATTTATTATTTTATATATATATATATA TAATTTATTATTATATATATATATATTATAT TATATAATTTATTATTTATATATATATATATA TAATTTATTA The 35nt smallest reiterated sequence which is replicated in yeast mitochondria is in grey	Footprint of ABF2 (HMG1-like protein) from yeast mitochondria on the yeast chromosomal origin of replication ARS1	Diffley and Stillman [1992]
14 GTATTTATTTTTTTTAATC	Core ORI of SV40 between the two T-antigen binding sites	Hay and De Pamphilis [1982]
5,077 CTTAAAATAGAAAATGT (region α) 5,149 CTTTTAATTAGTTG (region β) 5,269 TGTTTTTTTTAGTATTAAGC ori-core 101 CCAAATTGATATAATTAAGC T-antigen binding site	DNA sequences motifs from polyoma virus core origin and enhancers	Hendrickson et al. [1987]

sion to the demonstrated role of the nuclear matrix in active DNA replication (see above). It can be speculated that nuclear matrix might be involved in the differential activation of origins of replication during development. Of the total number of potential origins of replication that are present in each cell, only of fraction are activated at a given developmental stage [Spradling and Orr-Weaver, 1987]. A decrease in the number of active ORIs takes place during development [Callan, 1973]. The differential activation of ORI sequences among established cell types of an organism is thought to play a role in the differential gene expression [Trempe et al., 1988; Leffak and James, 1989].

In addition, gene activity is responsible for the differential timing of replication of the various active and inactive genes during the S-phase of the cell cycle [Holmquist, 1987; Ferguson et al., 1991]. Since both transcription and replication seem to occur on the nuclear matrix [see Boulikas, 1987], nuclear matrix might also be involved in determining the timing of replication of the various chromatin fractions during the S-phase. The order of replication of genes during S-phase might have an importance during development: active genes which are replicated first might deplete the cell nucleus of the transcription factors that will become less available for binding to the regulatory regions of the late-replicating genes. Thus, the roads between the timing of gene replication during S-phase, the differential activation of ORIs during development and the involvement of homeodomain proteins in differential gene expression and in pattern formation during development may indeed cross within the nuclear matrix.

ACKNOWLEDGMENTS

Special thanks to Emile Zuckerkandl for valuable input, to Diane Read for her patience and kindness in preparing the manuscript, to Jolanta Walichiewicz for computer graphics, and to Rosemary Babcock for literature searches.

REFERENCES

Aelen JMA, Opstelten RJG, Walka F (1983): Organization of DNA replication in *Physarum polycephalum*. Attachment of origins of replicons and replication forks to the nuclear matrix. Nucl Acids Res 11:1181–1195.

Amati BB, Gasser SM (1988): Chromosomal ARS and CEN

elements bind specifically to the yeast nuclear scaffold. Cell 54:967–978.
Anderson JN (1986): Detection, sequence patterns and function of unusual DNA structures. Nucleic Acids Res 14: 8513–8533.
Archer TK, Cordingley MG, Wolford RG, Hager GL (1991): Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary tumor virus promoter. Mol Cell Biol 11:688–698.
Bachvarov D, Moss T (1991): The RNA polymerase I transcription factor xUBF contains 5 tandemly repeated HMG homology boxes. Nucleic Acids Res 19:2331–2335.
Beachy PA, Krasnow MA, Gavis ER, Hogness DS (1988): An Ultrabithorax protein binds sequences near its own and the Antennapedia P1 promoter. Cell 55:1069–1081.
Belgrader P, Dey R, Berezney R (1991): Molecular cloning of matrin 3: A 125-kiloDalton protein of the nuclear matrix contains an extensive acidic domain. J Biol Chem 266: 9893–9899.
Berent SL, Sevall JS (1984): Histone H1 binding at the 5′ end of the rat albumin gene. Biochemistry 23:2977–2983.
Berezney R (1991): The nuclear matrix: A heuristic model for investigating genomic organization and function in the cell nucleus. J Cell Biochem 47:109–123.
Berezney R, Coffey DS (1975): Nuclear protein matrix: Association with newly synthesized DNA. Science 189:291–293.
Berrios M, Osheroff N, Fisher PA (1985): In situ localization of DNA topoisomerase II, a major polypeptide component of the *Drosophila* nuclear matrix fraction. Proc Natl Acad Sci USA 82:4142–4146.
Blasquez VC, Sperry AO, Cockerill PN, Garrard WT (1989b): Protein:DNA interactions at chromosomal loop attachment sites. Genome 31:503–509.
Blasquez VC, Xu M, Moses SC, Garrard WT (1989a): Immunoglobulin κ gene expression after stable integration. I. Role of the intronic MAR and enhancer in plasmacytoma cells. J Biol Chem 264:21183–21189.
Bode J, Maass K (1988): Chromatin domain surrounding the human interferon-β gene as defined by scaffold-attached regions. Biochemistry 27:4706–4711.
Boulikas T, Kong CF, Spandidos DA (1992): An unusual segment of the human *ras* gene is preferentially associated with nuclear matrix proteins and with histone H1. Submitted.
Boulikas T, Wiseman JM, Garrard WT (1980): Points of contact between histone H1 and the histone octamer. Proc Natl Acad Sci USA 77:127–131.
Boulikas T (1986a): Nucleosomes are assembled into discrete size structures by histone H1 in vitro. Biochem Cell Biol 64:463–473.
Boulikas T (1986b): Protein-protein and protein-DNA interactions in calf thymus nuclear matrix using cross-linking by ultraviolet irradiation. Biochem Cell Biol 64:474–484.
Boulikas T (1987): Nuclear envelope and chromatin structure. Int Rev Cytol Suppl 17:599–684.
Boulikas T (1992): Poly(ADP-ribosyl)ation, repair, chromatin and cancer. Curr Persp Mol Cell Oncol 1:1–109.
Broach JR, Li Y-Y, Feldman J, Jayaram M, Abraham J, Nasmyth KA, Hicks JB (1983): Localization and sequence analysis of yeast origins of DNA replication. Cold Spring Harbor Symp Quant Biol 47:1165–1173.
Brown EH, Iqbal MA, Stuart S, Hatton KS, Valinsky J, Schildkraut CL (1987): Rate of replication of the murine immunoglobulin heady-chain locus: Evidence that the region is part of a single replicon. Mol Cell Biol 7:450–457.
Burhans WC, Vassilev LT, Caddle MS, Heintz NH, DePamphilis ML (1990): Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell 62:955–965.
Buongiorno-Nardelli M, Micheli G, Carrî MT, Marilley M (1982): A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. Nature 298:100–102.
Caddle MS, Lussier RH, Heintz NH (1990): Intramolecular DNA triplexes, bent DNA and DNA unwinding elements in the initiation region of an amplified dihydrofolate reductase replicon. J Mol Biol 211:19–33.
Callan HG (1973): DNA replication in the chromosomes of eukaryotes. Cold Spring Harbor Symp Quant Biol 38:195–203.
Camp HS, Coussens PM, Silva RF (1991): Cloning, sequencing, and functional analysis of a Marek's disease virus origin of DNA replication. J Virol 65:6320–6324.
Carri MT, Micheli G, Graziano E, Pace T, Buongiorno-Nardelli M (1986): The relationship between chromosomal origins of replication and the nuclear matrix during the cell cycle. Exp Cell Res 164:426–436.
Chimera JA, Musich PR (1985): The association of the interspersed repetitive *Kpn*I sequences with the nuclear matrix. J Biol Chem 260:9373–9379.
Cho KWY, Goetz J, Wright CVE, Fritz A, Hardwicke J, De Robertis EM (1988): Differential utilization of the same reading frame in a *Xenopus* homeobox gene encodes two related proteins sharing the same DNA-binding specificity. EMBO J 7:2139–2149.
Ciejek EM, Tsai M-J, O'Malley BW (1983): Actively transcribed genes are associated with the nuclear matrix. Nature 306:607–609.
Cockerill PN (1990): Nuclear matrix attachment occurs in several regions of the IgH locus. Nucleic Acids Res 18: 2643–2648.
Cockerill PN, Garrard WT (1986): Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell 44:273–282.
Croston GE, Kerrigan LA, Lira LM, Marshak DR, Kadonaga JT (1991): Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription. Science 251:643–649.
De Simone V, De Magistris L, Lazzaro D, Gerstner J, Monaci P, Nicosia A, Cortese R (1991): LFB3, a heterodimer-forming homeoprotein of the LFB 1 family, is expressed in specialized epithelia. EMBO J 10:1435–1443.
Delidakis C, Kafatos FC (1989): Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila*. EMBO J 8:891–901.
Desplan C, Theis J, O'Farrell PH (1988): The sequence specificity of homeodomain-DNA interaction. Cell 54: 1081–1090.
Diffley JFX, Stillman B (1992): DNA binding properties of an HMG1-related protein from yeast mitochondria. J Biol Chem 267:3368–3374.
Dijkwel PA, Hamlin JL (1988): Matrix attachment regions are positioned near replication initiation sites, genes, and an interamplicon junction in the amplified dihydrofolate reductase domain of Chinese hamster ovary cells. Mol Cell Biol 8:5398–5409.
Driever W, Nüsslein-Volhard C (1989): The bicoid protein is

a positive regulator of *hunchback* transcription in the early *Drosophila* embryo. Nature 337:138–143.
Fangman WL, Henly JW, Churchill G, Brewer BJ (1989): Stable maintenance of a 35-base-pair yeast mitochondrial genome. Mol Cell Biol 9:1917–1921.
Farache G, Razin SV, Targa FR, Scherrer K (1990): Organization of the 3′ boundary of the chicken α globin gene domain and characterization of a CR 1-specific protein binding site. Nucleic Acids Res 18:401–409.
Ferguson BM, Brewer BJ, Reynolds AE, Fangman WL (1991): A yeast origin of replication is activated late in S phase. Cell 65:507–515.
Fisher DZ, Chaudhary N, Blobel B (1986): cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. Proc Natl Acad Sci USA 83:6450–6464.
Frain M, Swart G, Monaci P, Nicosia A, Stämpfil S, Frank R, Cortese R (1989): The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. Cell 59:145–157.
Gasser SM, Laemmli UK (1986): Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. Cell 46:521–530.
Grossbach ER, Björkroth B, Daneholt B (1990): Presence of histone H1 on an active Balbiani ring gene. Cell 60:78–83.
Hakes DJ, Berezney R (1991): Molecular cloning of matrin F/G: A DNA binding protein cf the nuclear matrix that contains putative zinc finger motifs. Proc Natl Acad Sci USA 88:6186–6190.
Hancock R, Boulikas T (1982): Functional organization in the nucleus. Int Rev Cytol 79:165–214.
Hanes S, Brent R (1991): A genetic model for interaction of the homeodomain recognition helix with DNA. Science 251:426–430.
Hay RT, DePamphilis ML (1982): Initiation of SV40 DNA replication in vivo: Location and structure of 5′ ends of DNA synthesized in the *ori* region. Cell 28:767–779.
Hay RT (1985): Origin of adenovirus DNA replication: Role of the Nuclear Factor I binding site in vivo. J Mol Biol 186:129–136.
He D, Nickerson JA, Penman S (1990): Core filaments of the nuclear matrix. J Cell Biol 110:569–580.
Hendrickson EA, Fritze CE, Folk WR, DePamphilis ML (1987): The origin of bidirectional DNA replication in polyoma virus. EMBO J 6:2011–2018.
Herr W, Sturm RA, Clerc RG, Corcoran LM, Baltimore D, Sharp PA, Ingraham HA, Rosenfeld MG, Finney M, Ruvkun G, Horvitz HR (1988): The POU domain: A large conserved region in the mammalian *pit*-1, *oct*-1, *oct*-2, and *Caenorhabditis elegans unc-86* gene products. Genes Dev 2:1513–1516.
Hoey T, Levine M (1988): Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. Nature 332:858–861.
Hofmann JF-X, Gasser SM (1991): Identification and purification of a protein that binds the yeast ARS consensus sequence. Cell 64:951–960.
Holmquist GP (1987): Role of replication time in the control of tissue-specific gene expression. Am J Hum Genet 40: 151–173.
Homberger HP (1989): Bent DNA is a structural feature of scaffold-attached regions in *Drosophila melanogaster* interphase nuclei. Chromosoma 98:99–104.
Iguchi-Ariga SMM, Okazaki T, Itani T, Ogata M, Sato Y, Ariga H (1988): An initiation site of DNA replication with transcriptional enhancer activity present upstream of the *c-myc* gene. EMBO J 7:3135–3142.
Izaurralde E, Käs E, Laemmli UK (1989): Highly preferential nucleation of histone H1 assembly on scaffold-associated regions. J Mol Biol 210:573–585.
Jackson DA, Dickinson P, Cook PR (1990): Attachment of DNA to the nucleoskeleton of HeLa cells examined using physiological conditions. Nucleic Acids Res 18:4385–4393.
Jantzen H-M, Admon A, Bell SP, Tjian R (1990): Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature 344:830–836.
Johnson AD, Herskowitz I (1985): A repressor (*MAT α2* product) and its operator control expression of a set of cell type specific genes in yeast. Cell 42:237–247.
Johnson WA, Hirsh J (1990): Binding of a *Drosophila* POU-domain protein to a sequence element regulating gene expression in specific dopaminergic neurons. Nature 343: 467–470.
Kamakaka RT, Thomas JO (1990): Chromatin structure of transcriptionally competent and repressed genes. EMBO J 9:3997–4006.
Käs E, Chasin LA (1987): Anchorage of the Chinese hamster dihydrofolate reductase gene to the nuclear scaffold occurs in an intragenic region. J Mol Biol 198:677–692.
Kissinger CR, Liu B, Martin-Blanco E, Kornberg TB, Pabo CO (1990): Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: A framework for understanding homeodomain-DNA interactions. Cell 63:579–590.
Klehr D, Maass K, Bode J (1991): Scaffold-attached regions from the human interferon β domain can be used to enhance the stable expression of genes under the control of various promoters. Biochemistry 30:1264–1270.
Landry S, Zannis-Hadjopoulos M (1991): Classes of autonomously replicating sequences are found among early-replicating monkey DNA. Biochim Biophys Acta 1088:234–244.
Laughon A, Scott MP (1984): Sequence of a *Drosophila* segmentation gene: Protein structure homology with DNA-binding proteins. Nature 310:25–31.
Laybourn PJ, Kadonaga JT (1991): Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science 254:238–245.
Leffak M, James CD (1989): Opposite replication polarity of the germ line *c-myc* gene in HeLa cells compared with that of two Burkitt lymphoma cell lines. Mol Cell Biol 9:586–593.
Levy-Wilson B, Fortier C (1989): The limits of the DNase I-sensitive domain of the human apolipoprotein B gene coincide with the locations of chromosomal anchorage loops and define the 5′ and 3′ boundaries of the gene. J Biol Chem 264:21196–21204.
Lichtsteiner S, Wuarin J, Schibler U (1987): The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell 51:963–973.
Marmorstein R, Carey M, Ptashne M, Harrison SC (1992): DNA recognition by GAL4: Structure of a protein-DNA complex. Nature 356:408–414.
Marilley M, Gassend-Bonnet G (1989): Supercoiled loop organization of genomic DNA: A close relationship between loop domains, expression units and replicon organization in rDNA in *Xenopus laevis*. Exp Cell Res 180:475–489.

Matsumoto LH (1981): Enrichment of satellite DNA on the nuclear matrix of bovine cells. Nature 294:481–482.

McCready SJ, Cook PR (1984): Lesions induced in DNA by ultraviolet light are repaired at the nuclear cage. J Cell Sci 70:189–196.

McGinnis W, Krumlauf R (1992): Homeobox genes and axial patterning. Cell 68:283–302.

McGinnis W, Hart CP, Gehring WJ, Ruddle FH (1984a): Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila.* Cell 38:675–680.

McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ (1984b): A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. Nature 308:428–433.

McKeon FD, Kirschner MW, Caput D (1986): Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. Nature 319: 463–468.

Mielke C, Kohwi Y, Kohwi-Shigematsu T, Bode J (1990): Hierarchical binding of DNA fragments derived from scaffold-attached regions: Correlation of properties in vitro in function in vivo. Biochemistry 29:7475–7485.

Mirkovitch J, Mirault M-E, Laemmli UK (1984): Organization of the higher-order chromatin loop: Specific DNA attachment sites on nuclear scaffold. Cell 39:223–232.

Mirkovitch J, Spierer P, Laemmli UK (1986): Genes and loops in 320,000 base-pairs of the *Drosophila melanogaster* chromosome. J Mol Biol 190:255–258.

Mullenders LHF, van Kasteren van Leeuwen AC, van Zeeland AA, Natarajan AT (1988): Nuclear matrix associated DNA is preferentially repaired in normal human fibroblasts, exposed to a low dose of ultraviolet light but not in Cockayne's syndrome fibroblasts. Nucleic Acids Res 16:10607–10623.

Müller MM, Ruppert S, Schaffner W, Matthias P (1988): A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. Nature 336:544–551.

Neuer-Nitsche B, Lu X, Werner D (1988): Functional role of a highly repetitive DNA sequence in anchorage of the mouse genome. Nucleic Acids Res 16:8351–8360.

Odenwald WF, Garbern J, Arnheiter H, Tournier-Lasserve E, Lazzarini RA (1989): The *Hox-1.3* homeo box protein is a sequence-specific DNA-binding phosphoprotein. Genes Dev 3:158–172.

Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H (1990): A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell 60:461–472.

Osman M, Paz M, Landesman Y, Fainsod A, Gruenbaum Y (1990): Molecular analysis of the *Drosophila* nuclear lamin gene. Genomics 8:217–224.

Pabo CO, Sauer RT (1984): Protein DNA recognition. Annu Rev Biochem 53:293–321.

Pardoll DM, Vogelstein B, Coffey DS (1980): A fixed site of DNA replication in eukaryotic cells. Cell 19:527–536.

Paulson JR, Laemmli UK (1977): The structure of histone-depleted metaphase chromosomes. Cell 12:817–828.

Phi-Van L, Strätling WH (1988): The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J 7:655–664.

Phi-Van L, von Kries JP, Ostertag W, Strätling WH (1990): The chicken lysozyme 5′ matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. Mol Cell Biol 10:2302–2307.

Postnikov YV, Shick VV, Belyavsky AV, Khrapko KR, Brodolin KL, Nikolskaya TA, Mirzabekov AD (1991): Distribution of high mobility group proteins $^{1}/_{2}$,E and $^{14}/_{17}$ and linker histones H1 and H5 on transcribed and non-transcribed regions of chicken erythrocyte chromatin. Nucleic Acids Res 19:717–725.

Pruijn GJM, van Driel W, van der Vliet PC (1986): Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. Nature 322:656–659.

Qian YQ, Billeter M, Otting G, Müller M, Gehring WJ, Wüthrich K (1989): The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution. Comparison with prokaryotic repressors. Cell 59:573–580.

Qian S, Capovilla M, Pirrotta V (1991): The *bx* region enhancer, a distant *cis*-control element of the *Drosophila Ubx* gene and its regulation by *hunchback* and other segmentation genes. EMBO J 10:1415–1425.

Razin SV, Mantieva VL, Georgiev GP (1978): DNA adjacent to attachment points of deoxyribonucleoprotein fibril to chromosomal axial structure is enriched in reiterated base sequences. Nucleic Acids Res 5:4737–4751.

Razin SV, Kekelidze MG, Lukanidin EM, Scherrer K, Georgiev GP (1986): Replication origins are attached to the nuclear skeleton. Nucleic Acids Res 14:8189–8207.

Reudelhuber TL, Boulikas T, Garrard WT (1980): A non-amer of histones in chromatin. J Biol Chem 255:4511–4515.

Reeves R, Nissen MS (1990): The A · T-DNA-binding domain of mammalian high mobility group I chromosomal proteins: A novel peptide motif for recognizing DNA structure. J Biol Chem 265:8573–8582.

Reeves R, Langan TA, Nissen MS (1991): Phosphorylation of the DNA-binding domain of nonhistone high-mobility group I protein by cdc2 kinase: Reduction of binding affinity. Proc Natl Acad Sci USA 88:1671–1675.

Richard-Foy H, Hager GL (1987): Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. EMBO J 6:2321–2328.

Ristiniemi J, Oikarinen J (1989): Histone H1 binds to the putative nuclear factor I recognition sequence in the mouse α_2(I) collagen promoter. J Biol Chem 264:2164–2174.

Scheuermann RH (1991): The tetrameric structure of NF-μNR provides a mechanism for cooperative binding to the immunoglobulin heavy chain μ enhancer. J Biol Chem 267:624–634.

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986): The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049.

Schöler HR, Hatzopoulos AK, Balling R, Suzuki N, Gruss P (1989): A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. EMBO J 8:2543–2550.

Scott MP, Weiner AJ (1984): Structural relationships among genes that control development: sequence homology between the *Antennapedia, Ultrabithorax,* and *fushi tarazu* loci of *Drosophila.* Proc Natl Acad Sci USA 81:4115–4119.

Scott MP, Tamkun JW, Hartzell GW (1989): The structure and function of the homeodomain. Biochim Biophys Acta 989:25–48.

Small D, Nelkin B, Vogenstein B (1982): Nonrandom distribution of repeated DNA sequences with respect to supercoiled loops and the nuclear matrix. Proc Natl Acad Sci USA 70:5911–5915.

Spradling A, Orr-Weaver T (1987): Regulation of DNA replication during *Drosophila* development. Annu Rev Genet 21:373–403.

Stief A, Winter DM, Strätling WH, Sippel AE (1989): A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature 341:343–345.

Swanson MS, Carlson M, Winston F (1990): SPT6, an essential gene that affects transcription in *Saccharomyces cerevisiae,* encodes a nuclear protein with an extremely acidic amino terminus. Mol Cell Biol 10:4935–4941.

Swanson MS, Malone EA, Winston F (1991): SPT5, an essential gene important for normal transcription in *Saccharomyces cerevisiae,* encodes an acidic nuclear protein with a carboxy-terminal repeat. Mol Cell Biol 11:3009–3019.

Sykes RC, Lin D, Hwang SJ, Framson PE, Chinault AC (1988): Yeast ARS function and nuclear matrix association coincide in a short sequence from the human HPRT locus. Mol Gen Genet 212:301–309.

Thoma F, Koller T, Klug A (1979): Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol 83:403–427.

Treisman J, Harris E, Wilson D, Desplan D (1992): The homeodomain: A new face for the helix-turn-helix? BioEssays 14:145–150.

Tremethick DJ, Molloy PL (1986): High mobility group proteins 1 and 2 stimulate transcription in vitro by RNA polymerases II and III. J Biol Chem 261:6986–6992.

Tremethick DJ, Molloy PL (1988): Effects of high mobility group proteins 1 and 2 on initiation and elongation of specific transcription by RNA polymerase II in vitro. Nucleic Acids Res 16:11107–11123.

Trempe JP, Lindstrom YI, Leffak M (1988): Opposite replication polarities of transcribed and nontranscribed histone H5 genes. Mol Cell Biol 8:1657–1663.

Umek RM, Linskens MHK, Kowalski D, Huberman JA (1989) Review: New beginnings in studies of eukaryotic DNA replication origins. Biochim Biophys Acta 1007:1–14.

Ustav M, Ustav E, Szymanski P, Stenlund A (1991): Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. EMBO J 10:4321–4329.

Vaughn JP, Dijkwel PA, Mullenders LHF, Hamlin JL (1990): Replication forks are associated with the nuclear matrix. Nucleic Acids Res 18:1965–1969.

von Kries JP, Buhrmester H, Strätling WH (1991): A matrix/scaffold attachment region binding protein: Identification, purification and mode of binding. Cell 64:123–135.

Walker JM, Gooderham K, Hastings JRB, Mayes E, Johns EW (1980): The primary structures of non-histone chromosomal proteins HMG 1 and 2. FEBS Lett 122:264–270.

Xu M, Hammer RE, Blasquez VC, Jones SL, Garrard WT (1989): Immunoglobulin κ gene expression after stable integration. II. Role of the intronic MAR and enhancer in transgenic mice. J Biol Chem 264:21190–21195.

Yuan J, Simos G, Blobel G, Georgatos SD (1991): Binding of lamin A to polynucleosomes. J Biol Chem 266:9211–9215.

Zappavigna V, Renucci A, Izpisúa-Belmonte J-C, Urier G, Peschle C, Duboule D (1991): HOX4 genes encode transcription factors with potential auto- and cross-regulatory capacities. EMBO J 10:4177–4187.

Zuckerkandl E, Villet R (1988): Generation of high specificity of effect through low-specificity binding of proteins to DNA. FEBS Lett 231:291–298.