Volume 7 Number 8 1979 Nucleic Acids Research

DNA methylation: organ specific variations in the methylation pattern within and around ovalbumin and other chicken genes

J.L.Mandel and P.Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, Institut de Chimie Biologique, Strasbourg, France

Received 8 November 1979

ABSTRACT

The restriction enzymes HhaI and HpaII, whose activity is inhibited by cytosine methylation within their recognition sites, have been utilised as probes to study methylation in the vicinity of the ovalbumin gene in DNA from various chicken tissues. This was complemented by a preliminary study of methylation in the regions of chicken ovotransferrin (conalbumin), ovomucoïd and β-globin genes. From our data we conclude that HhaI or HpaII sites can be divided in 3 classes according to their pattern of methylation in different tissues. In the first class of sites (m^V class) the extent of methylation varies in different tissues. The patterns obtained show that methylation at the sites located within and around the 3 genes which code for egg white proteins is in general lowest in oviduct of laying hen, where these genes are expressed. However some sites are not methylated (m^- class) and others are 95 to 100% resistant (m^+ class) to digestion by HhaI or HpaII in the DNAs of all the tissues which were tested. Our study has also revealed a remarkable number of allelic variants for the presence of HhaI or HpaII sites in the region of the ovalbumin gene.

INTRODUCTION

Cytosine methylation occurs specifically at CpG dinucleotides in animal DNA (1, 2). The possible roles in cellular differentiation or in regulation of gene activity of this important post-synthetic (3) modification, has been the object of intensive speculation (4, 5, 6). Experimental data concerning possible functions of DNA methylation have been scarce, and contradictory reports have appeared on tissue specific variations in methylcytosine content of DNA (7, 8) or on correlations between DNA methylation and chromatin structure (9, 10). However, satellite DNAs and in general repetitive sequences have been shown to be highly methylated (11, 12). The finding that some restriction nucleases

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

like HhaI (GCGC) and HpaII (CCGG) ("CpG enzymes") are inhibited by cytosine methylation at their recognition site (13) has allowed one for the first time to study methylation at specific sites in defined regions of the genome. This approach was first described by Bird and Southern for ribosomal DNA of Xenopus laevis (14). They showed that in somatic rDNA most of the restriction sites were methylated but that some unmethylated sites could be found. Amplified rDNA, on the contrary, contains no detectable methylcytosine (15).

We have applied this last approach to genes whose expression in highly differentiated cells is under well known transcriptional regulation. In parallel with our work on the organization of the ovalbumin gene (16, 17), we have performed a systematic study of the reactivity of HhaI and HpaII sites in the region of this gene in DNAs purified from various chicken tissues (referred, as cellular DNAs in this study). We have also investigated in a more preliminary way the methylation patterns in the regions of three other chicken genes, coding for conalbumin (ovotransferrin), ovomucoid and β-globin.

We show here that there are tissue-specific variations in the extent of methylation of several HhaI and HpaII sites. The genes coding for ovalbumin, conalbumin, and ovomucoid appear to be in general less methylated at these sites in the oviduct of laying hen, where they are actively expressed, than in all the other tissues studied, sperm DNA being the most highly methylated. However, certain HhaI and HpaII sites appear always unmethylated even in sperm DNA, and other sites have been found to be always resistant to restriction nuclease digestion irrespective of the tissue origin. Our study has also revealed a striking number of allelic variants for HhaI and HpaII sites in the region of the ovalbumin gene.

MATERIEL AND METHODS

Cellular DNA was purified as described by Gross-Bellard et al. (18), starting from a crude nuclear fraction. Restriction nuclease EcoRI was prepared by the method of Sumegi et al. (19), HhaI and HpaII were gifts from Dr. P. Humphries and P. Gerlinger, KpnI and MspI were from Biolabs. Single and double digestions

were always monitored using SV40 DNA as an internal standard. Agarose gel electrophoresis, transfer onto nitrocellulose filters and hybridization conditions were as previously described (16). Autoradiography was performed using Kodak XRI films and Dupont Cronex Li plus intensifying screens.

Densitometry scanning (using a Vernon apparatus) was performed on films exposed for shorter times than those shown in the figures to avoid saturation of the film. Due to the many steps involved in these experiments the values which are given are only indicative, but were reproducible from one experiment to another.

RESULTS

I. PRESENCE AND REACTIVITY OF HhaI AND HpaII SITES WITHIN AND AROUND THE OVALBUMIN GENE.

Several HhaI and HpaII sites in the ovalbumin gene region were localized by restriction mapping of cloned DNA segments (20-22) or of cellular DNA (this study). Their positions with respect to the exon and intron sequences of the gene are given in Fig. 1.

To study the reactivity of these sites, high molecular

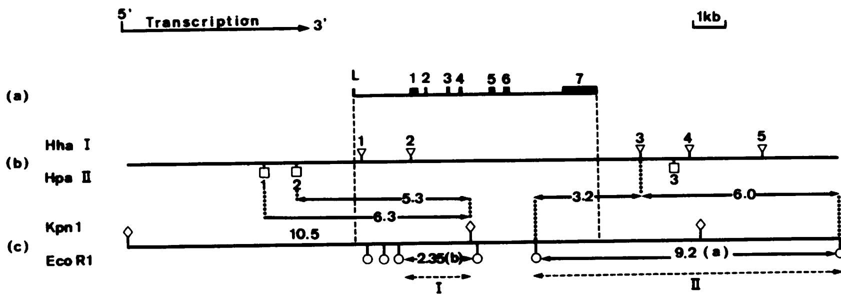


Fig. 1 : Map of chicken DNA in the ovalbumin gene region. (a) Location in the ovalbumin gene of the leader-coding sequence (L) and of the 7 other exons (22,23). The vertical dashed lines indicate the position of the ovalbumin gene in (b) and (c). (b) Location of HhaI (▽,1-5) and HpaII (□,1-3) sites. (c) Location of KpnI (◊) and EcoRI (o) sites, and length (in kb) of restriction fragments which are discussed in the text. The sequences covered by the hybridization probes I and II are indicated by the horizontal dashed lines.

weight DNA samples isolated from various organs were digested both with the appropriate "CpG enzyme" and with an enzyme (EcoRI or KpnI) which yields fragments previously mapped in the ovalbumin gene. These digests were analysed by the method of Southern, using the hybridization probes described in Fig. 1 (c).

I.1. Hhal sites.

Site Hha1 : This site has been found in only one of two recently cloned DNA segments which contain the genomic sequences coding for the ovalbumin mRNA leader (22). Site Hha1 has not been detected in genomic DNA in a preliminary screening and further studies are needed to find whether this is due to a low frequency of occurence of this site in the population studied, or to its presence in a completely methylated state.

Site Hha2 : This site is located in exon 1 of one cloned ovalbumin gene (21, 23), but there are allelic variants which do not possess it. It is found in the ovalbumin cDNA clone pCRIov2.1 constructed in our laboratory (24), but is absent from another cDNA clone (25) and from other cloned ovalbumin genes [the cosmid clone pAR2 and the λ clone λC4-ov5 (see Ref. 22)].

We have examined directly the occurence and reactivity of this site in oviduct DNA digested with EcoRI and HhaI using the hybridization probe (Probe I) which corresponds to the 2.35 kb EcoRI fragment "b" of the ovalbumin gene [Fig. 1 (c)]. Analyzing the DNA of 16 laying hens, we did not detect any digestion at the position corresponding to site Hha2 in the DNA from 11 animals (only 4 are shown in Fig. 2A, lanes 1, 2, 4, 6), while partial digests were obtained in the 5 other cases (Fig. 2A, lanes 3, 5 and 7). The extent of digestion was evaluated by densitometry scanning of the autoradiographs. Digestion at site Hha2 varied from 25% to 33% and a mean value of 30% was found for the 5 hens. DNAs from various organs of the same hens were prepared and analyzed in a similar way. As shown in Fig. 2B digestion at site Hha2 was much less efficient in liver DNA (lane 3) (about 4%) than in oviduct DNA (lane 2), and was barely detectable in erythrocyte, kidney, spleen or brain DNA (similar patterns were obtained with the other hens). A study of DNA from roosters

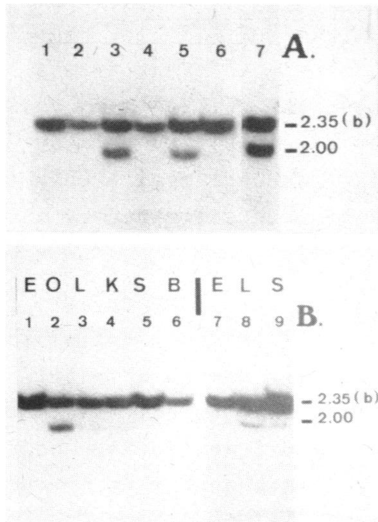


Fig. 2 : (A). Detection of site Hha2 in oviduct DNA. Samples of oviduct DNA from 7 different laying hen were digested with EcoRI and HhaI, separated on a 1% agarose gel and analysed using ovalbumin probe I (see Fig. 1c). Extents of digestion, as measured by densitometry scanning was 30%, 25% and 33% for chicken 3,5 and 7, respectively). Size of the fragments are given in kilo base pairs. (B). Reactivity of site Hha2 in DNA from different tissues. DNAs from erythrocyte (E), oviduct (O), liver (L), kidney (K), spleen (S), or brain (B) were digested and analysed as in (A). Lanes 1-6, same animal as in A lane 3. Lanes 7-9, DNA prepared from a rooster. Extents of digestion : 27, 3.5 and 3.5 % for lanes 2, 3 and 8, respectively. Digestion for other samples was detectable only on overexposed films and could not measured.

showed that, when the site was present, the extent of its digestion was also very low in liver DNA, and even lower in erythrocyte or spleen DNA (Fig. 2B, lanes 7-9, Fig. 3, lane 4). No digestion was detectable in sperm DNA (Fig. 3, lane 3). The site is most likely absent in the other animal analyzed (Fig. 3, lanes 1 and 2) since it cannot be detected even in liver DNA (lane 2).

These results confirm the allelic polymorphism for the presence of site Hha2, which was previously suggested by the study of cloned fragments (23, 25). When present, this site

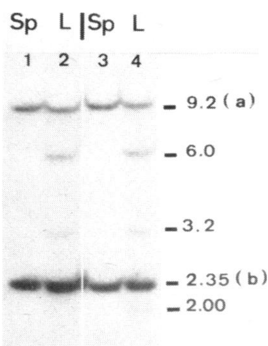


Fig. 3 : Reactivity of sites Hha2 and 3 in sperm and liver rooster DNA. Sperm (Sp) and liver (L) DNAs from two roosters were digested with EcoRI and HhaI, separated on a 1% agarose gel and analysed using a mixture of the hybridization probes I and II (Fig. 1c). Extents of digestion : site Hha3 (6.0 and 3.2 kb fragments) lane 2, 50% : lane 4, 52%; site Hha2 (2.0 kb fragment) lane 4, 7%; no digestion could be detected in lanes 1 and 3.

appears preferentially digested in oviduct DNA. Since about 60% of the chickens appear totally devoid of Hha2 site in the population studied (24 animals, including roosters), chickens heterozygous for site Hha2 are expected to outnumber chicken homozygous for the site by a factor of 6. Thus, the digestibility of about 30% found for site Hha2 in oviduct DNA corresponds most likely to heterozygous animals and may therefore be underestimated by a factor of 2.

Sites Hha3, 4 and 5 : these three sites were mapped (our unpublished results) in a cloned 9.2 kb EcoRI fragment "a" of the ovalbumin gene (20, see Fig. 1). Site Hha3 is located at about 1200 bp downstream from the sequence coding for the polyA addition site of the ovalbumin mRNA and sites Hha4 and 5 are further downstream from the gene (Fig. 1b). All three sites appear to be located outside the ovalbumin transcription unit (26). The presence of these sites was tested in EcoRI-digested erythrocyte or oviduct DNAs of 5 laying hen, using the EcoRI fragment "a" as hybridization probe (Probe II, see Fig. 1c). The EcoRI 9.2 kb cellular DNA fragment is partially digested by HhaI to yield two fragments 6.0 and 3.2 kb long (Fig. 4A). Only the latter hybridizes with an ovalbumin cDNA probe (result not shown), which prove that these bands arise from digestion at the site Hha3. No other bands are seen which indicates that no significant digestion occurs at the location of sites Hha4 and 5 (see below). As shown in Fig. 4A, the level of digestion at site Hha3 is about 85% in oviduct DNA, (lanes 6 to 10) and about 30% in erythrocyte DNA (lanes 1 to 5). This difference is not due to incomplete digestion as shown in Fig. 4B, since the digestion pattern did not change when the amount of HhaI enzyme and the incubation time were both doubled. The digestibility of site Hha3 was studied in DNA samples isolated from different organs of a laying hen (Fig. 5, lanes 1 to 10) or of a rooster (Fig. 5, lanes 11 to 14). Results show that the reactivity of this site is different in various organs, but it is always lower than in oviduct DNA. Liver DNA from either females (Fig. 5, lane 2) or males (Fig. 5, lane 12 and Fig. 3, lanes 2 and 4) is digested to a greater extent (up to 50%) than erythrocyte, spleen or brain DNA. As was found for site Hha2, no digestion could be detected in sperm DNA (Fig. 3, lanes 1 and 3).

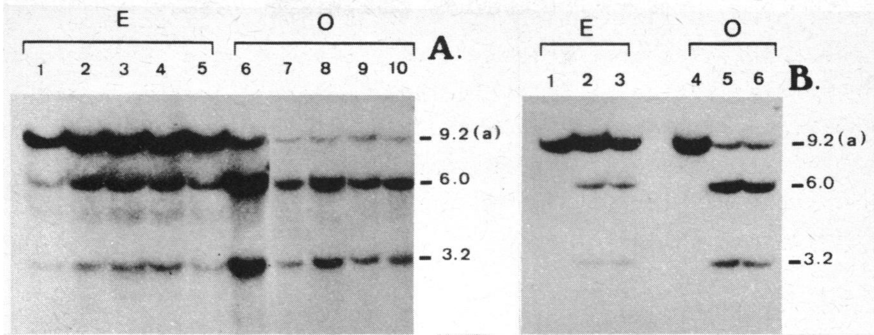


Fig. 4 : Digestion of site Hha3 in DNA from erythrocyte (E) and oviduct (O) (A). Erythrocyte DNA (lanes 1-5) and oviduct DNA (6-10) from 5 laying hens were separated on a 1% agarose gel and analysed using hybridization probe II (Fig. 1c). Extent of digestion : lanes 1 to 5 : 31, 34.5, 29.5, 32, 25% respectively ; lanes 6 to 10 : 85, 87, 84,81 and 87 %, respectively. (B). Effect of increased exposure to nuclease HhaI. Erythrocyte (E) and oviduct (O) DNAs from a laying hen were digested with EcoRI alone (lanes 1 and 4) or together with a single dose of HhaI for 1 hour (lanes 2 and 5) or with a double dose for 2 hours (lanes 3 and 6). The digests were analysed as in (A). Extents of digestion : 20, 30, 77 and 74% for lanes 2, 3, 5 and 6, respectively.

The high level of digestion in oviduct DNA indicates that in the DNA of all the laying hen tested site Hha3 occurs in the homozygous state. On the contrary, we have never detected any digestion at sites Hha4 and 5 in cellular DNA (more than 15

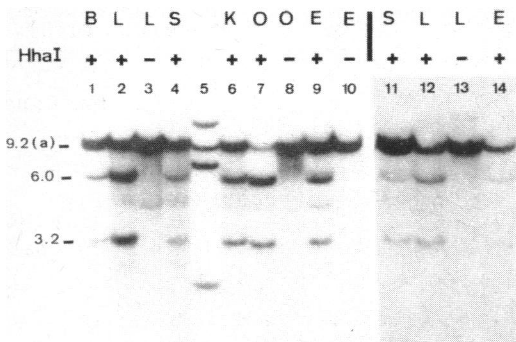


Fig. 5 : Reactivity of site Hha3 in DNA from different tissues. DNA samples from different tissues (same symbols as in Fig. 2B) from a single laying hen (lanes 1 to 10) or a rooster (lanes 11 to 14) were digested with EcoRI and HhaI (lanes 3, 8, 10 and 13 are controls without HhaI), and analysed as in Fig. 4A. Extents of digestion : 15, 51, 25.5, 34, 84, 20.5, 11, 31 and 13% for lanes 1, 2, 4, 6, 7, 9, 11, 12 and 14, respectively. A molecular weight marker is shown in lane 5.

animals were tested). However, these two sites are both present in the cloned EcoRI fragment "a" of Garapin et al. (20) and in the genomic clone λC4-ov5 (22) isolated from a chicken library constructed by Dodgson et al. (27) in California ; they should thus exist at least in some of the animals studied here. We tentatively conclude that when they are present, they must both be completely resistant to the action of the HhaI nuclease.

I.2. HpaII sites.

Sites Hpa1 and 2 : restriction mapping of genomic DNA from several laying hens, using both HpaII and its isoschizomer MspI, which is not inhibited by cytosine methylation at its recognition site (28-29)has allowed us to detect and map two sites, Hpa1 and 2. These sites are located upstream from the ovalbumin mRNA leader coding sequence, in a 10.5 kb KpnI fragment (Fig. 1). The presence and reactivity of sites Hpa1 and 2 were studied in KpnI-HpaII or KpnI-MspI double digests hybridized to probe I (Fig. 1c) and some of the results are shown in Fig. 6. In DNA from laying hen A, the 10.5 kb KpnI fragment (lane 6) was digested by MspI to a 5.3 kb hybridizing fragment (lane 1) indicating the presence of site Hpa2 (Fig. 1b) in the homozygous state in this animal. With HpaII however, less than 8 % digestion occured at this site in erythrocyte, oviduct, liver or kidney DNA from the same animal

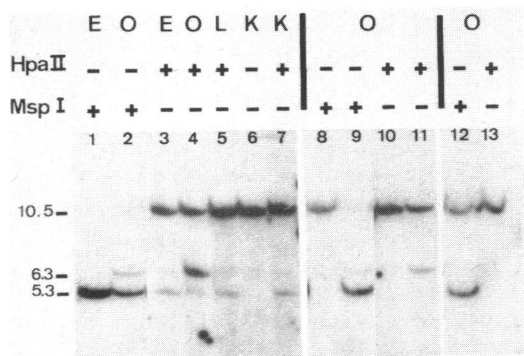


Fig. 6 : Presence and reactivity of sites Hpa 1 and 2. DNA from different tissues (symbols as in Fig. 2B) and from different laying hen were digested with KpnI and either HpaII or MspI as indicated in the figure. The digests were electrophoresed on a 0.7% agarose gel and hybridized to probe I (Fig. 1c). Chicken A, (lanes 1 to 7), B,(lanes 8 and 10), C,(lanes 9 and 11), D, (lanes 12 and 13). Persistance of the band at 6.3 kb was reproducibly observed in independent MspI digests of oviduct DNA from chicken A (see also Fig. 7). Extents of digestion : site Hpa1 42 and 22% for lanes 4 and 11 respectively; site Hpa2, 8, 6 and 6% for lanes 3, 5, and 7 respectively.

(lanes 3, 4, 5 and 7). On the other hand in oviduct DNA, HpaII converted 40 % of the 10.5 kb KpnI fragment to a 6.3 kb hybridizing fragment (lane 4) revealing the digestibility of site Hpa1 (Fig. 1b). This site was not digested by HpaII in other tissues (lanes 3, 5, 7).

A similar pattern was obtained with DNA from two other laying hens, C (lanes 9 and 11) and E (results not shown). In both animals site Hpa2, although present in the homozygous state as indicated by MspI digestion (lane 9), was resistant to the action of HpaII in oviduct DNA (lane 11) or in liver DNA (not shown). Site Hpa1 was also present in both animals, and was sensitive to enzyme HpaII only in oviduct DNA (lane 11) although the extent of digestion (20 to 25 % for both animals) was reproducibly lower than for oviduct DNA from chicken A. (It should be noted that our experiments do not allow us to determine whether chickens A, C or E are homozygous or heterozygous for site Hpa1, since in MspI digests, this site is separated by the site Hpa2 from the sequences hybridizing to probe I). A different situation was found in the case of chicken B and D. Hpa1 and 2 are both absent in chicken B, since the 10.5 kb KpnI fragment is completely resistant to MspI or HpaII digestion (lanes 8 and 10). In chicken D, site Hpa2 is present only in the heterozygous state (lane 12) and there is no detectable digestion at site Hpa1 by either MspI or HpaII (lanes 12 and 13). It is thus likely that site Hpa1 is absent in chicken D (although we cannot exclude at the present time that it is present in a completely methylated state on the same allele as site Hpa2).

From these data, which are summarized in Table 1, we conclude that sites Hpa1 and 2 define two additional allelic polymorphisms in the vicinity of the ovalbumin gene. Site Hpa1 is preferentially digested by HpaII in oviduct DNA, but we do not know the true extent of its digestion because of the uncertainty on the dosage of this site in the genome of the animals studied. Site Hpa2, on the contrary, appears very poorly digested by HpaII, irrespective the origin of the tissue.

Site Hpa3 : We found very recently that clone λC4-ov5 (22) contains an HpaII site located in the middle of the EcoRI fragment "a". This site was absent from the previously cloned fragment "a"

TABLE 1

Site	Occurence	HpaII digestion (in %) oviduct	other tissues
Hpa1	Hen A : present	42	≤3
	Hen B : absent	-	-
	Hens C and E : present	22 (C), 26 (E)	≤3
Hpa 2	Hen A : present (homozygote)	≤3	≤8
	Hen B : absent (homozygote)	-	-
	Hens C and E : present (homozygote)	nd	≤6
	Hen D : present (heterozygote)	nd	nt

Occurence and digestibility of sites Hpa1 and 2. This table summarizes data obtained on 5 different hen (Fig. 6 and results not shown). nd : no digestion was detected (<1%). nt : not tested. When unknown, the allelic state is not indicated.

(20). A search (using MspI) for this site in 5 laying hens and 2 roosters, revealed its presence in 3 animals (two laying hens and one rooster). Site Hpa3 is not digested by HpaII in oviduct (Fig. 7, lanes 5 and 6), liver or sperm DNA (not shown). It is however digested by HpaII in erythrocyte DNA (Fig. 7, lanes 1 and 2). This digestion is not complete, as indicated by the comparison with the pattern obtained with MspI (Fig. 7, lanes 3, 4). It is interesting to note that MspI does not digest to completion at

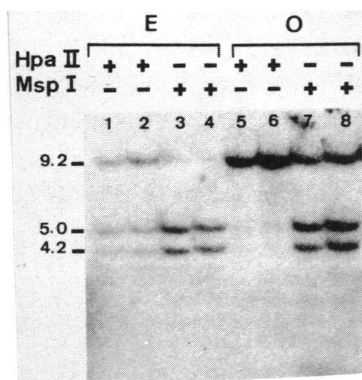


Fig. 7 : Reactivity of site Hpa3. DNA from erythrocyte (E) or oviduct (O) from a single laying hen were digested with EcoRI and HpaII or MspI as indicated in the figure. In lanes 2, 4, 6 and 8, samples were digested with twice the amount of enzymes for twice as long as in lanes 1, 3, 5 and 7. Digests were analysed on a 0.9% agarose gel and hybridized to probe II (Fig. 1c). Extents of digestion : 53, 52, 91, 90, 64 and 64% in lanes 1 to 4 and 7, 8, respectively. Extent of digestion in lanes 5 and 6 was too low to be measured.

the site Hpa3 of oviduct DNA, which suggests that this site might be further modified, possibly by methylation (R. Flavell, personal communication) on the first C of the recognition sequence (For a possible similar situation at site Hpa2, see Fig. 6, lane 2).

II. HhaI AND HpaII DIGESTION PATTERNS WITHIN AND AROUND THREE OTHER CHICKEN GENES.

The above results suggest that there may be a correlation between DNA undermethylation in a gene region and its expression, since the ovalbumin gene is actively transcribed in hen oviduct but not in the other chicken organs. In order to probe the generality of such a correlation between DNA undermethylation and gene expression, we studied the methylation patterns of unmapped HpaI and HpaII sites in 3 other genes for which cDNA clones have been constructed in our laboratory, and which have different patterns of tissue-specific expression.

II.1. The ovotransferrin (conalbumin) gene.

The ovotransferrin gene is expressed both in oviduct, under steroid hormone control (30) and, at a lower rate, in the liver where both iron and steroids modulate its transcription (31). This gene is cleaved by EcoRI into 3 fragments, "a", "b", "c", of 10.7, 4 and 2.5 kb (32, 33; see also Figs 8 and 9). We analyzed EcoRI-HhaI or EcoRI-HpaII double digests of DNA from various tissues, using as hybridization probe a segment of the cloned conalbumin cDNA (34) which corresponds to nucleotides 20 to about 1900 of the 2400 nucleotide-long conalbumin mRNA. The results in Figs 8 and 9 show complex patterns of digestion, with striking differences between the various tissues examined. It is apparent that in all cases, including sperm DNA (Fig. 9, lanes 2, 3, 6 and 7) the EcoRI fragment "a" is totally digested by both HhaI and HpaII, into smaller fragments. In erythrocyte (Fig. 8, lanes 2 and 3), spleen (Fig. 8, lanes 14 and 15) and sperm DNA (Fig. 9, lanes 2, 3, 6 and 7) the principal product of digestion of EcoRI fragment "a" by each of these two nucleases is 8.2 kb long. In oviduct (Fig. 8, lanes 5, 6), liver (Fig. 8, lanes 8, 9, and

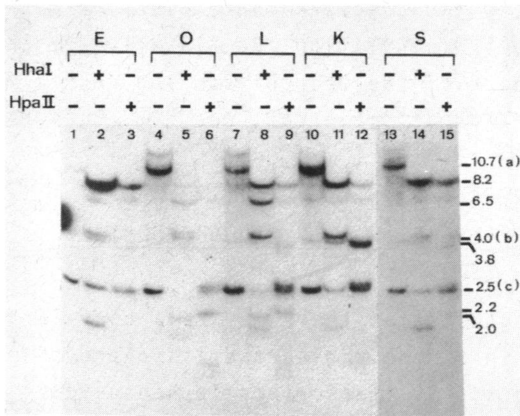


Fig. 8 : Analysis of HhaI and HpaII digestion patterns within and around the conalbumin gene. DNA from various tissues from a single laying hen were digested with EcoRI alone or in combination with HhaI or HpaII as indicated in the figure. Symbols are as in Fig. 2B. Electrophoresis was on an 1% agarose gel and the hybridization probe used was a mixture of the HhaI fragments A and B from the cloned conalbumin dscDNA (32,34). The presence of bands above band "a" in EcoRI digests (lanes 4, 7 and 13) might indicate some incomplete digestion although when samples from the same digests were analysed with an ovalbumin probe a pattern of complete digestion was obtained.

Fig. 9, lanes 4, 5, 8, 9) and kidney (Fig. 8, lane 12) fragments smaller than 8.2 kb appear. We conclude from these patterns that at least one HhaI site and one HpaII site are found in EcoRI fragment "a" which are unmethylated in all tissues examined. Preliminary mapping studies locate these two sites downstream from the polyA addition site of the conalbumin gene. On the other hand, other sites are methylated to a variable extent in different tissues. Here too, it is in oviduct DNA that HhaI or HpaII digestions are the most efficient, as shown by the almost complete disappearance of the 8.2 kb fragment with HhaI (Fig. 8, lane 5) and of fragments longer than 2.6 kb with HpaII (Fig. 8, lane 6).

In liver DNA, the persistance of part of the 8.2 kb fragment after HhaI or HpaII digestion indicates that at least some sites are less sensitive to these enzymes than in oviduct DNA. On the other hand, some fragments which are common to liver and oviduct DNA are not found in other tissues, eg. a 6.5 kb and 2 kb bands in the HhaI digests, a 2.2 kb band in the HpaII digests, Fig. 8, lanes 5, 6, 8, 9). This suggests that there are sites which are preferentially undermethylated in both oviduct and liver DNA. In kidney, at least one HpaII site also appears to be undermethylated, as indicated by the relative decrease of the 8.2 kb fragment and the appearance of a 3.8 kb fragment (Fig. 8, lane 12).

It is also apparent that the pattern found in liver DNA from laying hen (Fig. 8, lanes 8, 9) is very similar to that found in liver DNA from roosters (Fig. 9, lanes 4, 5, 8, 9). The digestion patterns in oviduct, liver and erythrocyte DNAs are very reproducible for different chicken indicating that they reflect a stable property of the gene in a given tissue.

II. 2. The ovomucoid gene.

This gene is expressed in oviduct under steroid hormone control (35-37), but not in liver (38). Methylation in the ovomucoid gene region was studied using an ovomucoid cDNA clone constructed in our laboratory (A. Krust et al., unpublished results, and ref. 26) as hybridization probe. The ovomucoid gene sequences are found in two EcoRI fragments : "a", 14 kb; and "b", 6.5 kb. [fragment "b" contains only 160 bp of mRNA-coding sequence (39, and P. Gerlinger unpublished results), and is generally more difficult to detect in our experiments]. In all tissues (including spleen and brain, not shown) both HhaI and HpaII cleave the EcoRI fragments "a" in hybridizing fragments of 8.4 and 8.2 kb, respectively, or in even smaller fragments (Fig. 10). This suggests that as in the case of the conalbumin gene region, at least one HhaI and HpaII sites are unmethylated in all tissues tested. Oviduct DNA

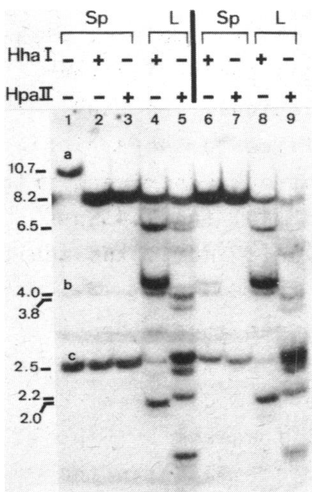


Fig. 9 : HhaI and HpaII digestion patterns in the region of the conalbumin gene, in sperm and liver DNA. DNA from sperm (Sp) and liver (L) from two roosters were analysed as described in legend to Fig. 8 (but electrophoresis was on an 0.8% agarose gel).

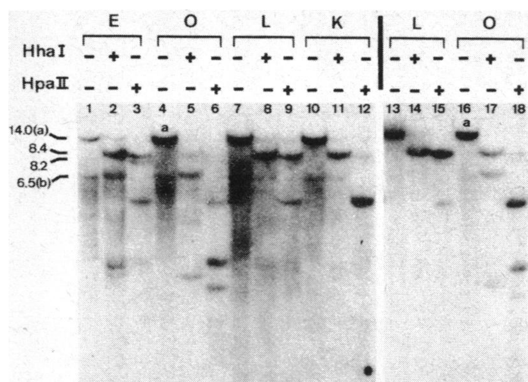


Fig. 10 : HhaI and HpaII digestion patterns in the region of the ovomucoid gene. DNA from various tissues (symbols as in Fig. 2B) were digested and analysed as described in the legend to Fig. 8, but the hybridization probe was a DNA fragment containing the ovomucoid cDNA sequence of plasmid pBovoml (A. Krust et al., unpublished). Samples in lanes 1 to 12 and in lanes 13 to 18 are from two different laying hens.

contains additional HhaI and HpaII sites available for digestion as indicated by the almost complete disappearance of the 8.2 kb fragment in the HpaII digests (Fig. 10, lanes 6 and 18), and the presence of fragments of 2.7 kb or less in both HhaI and HpaII digests (lanes 5, 6, 17 and 18). Contrary to what was observed for the conalbumin region, the ovomucoid region appears less sensitive to digestion by HhaI and HpaII in liver DNA (lanes 8, 9, 14 and 15) than in erythrocyte DNA (lanes 2 and 3).

II. 3. β-globin genes.

A chicken globin cDNA has been cloned in our laboratory (A. Krust, unpublished results) and one of the recombinant plasmids, has been shown to contain the sequences coding for the adult β-globin (D. Engel, personal communication). Two EcoRI fragments ("a", 9.4, and "b", 6.0 kb) hybridize to the β-globin probe under the relatively low stringency of our experiment (Fig. 11, lane 4). According to Engel and Dodgson (40) the 9.4 and the 6.0 kb fragments correspond to an embryonic and to the adult β-globin gene, respectively.

After HhaI digestion, 4 bands are seen in erythrocyte DNA. The faint 6.0 kb fragment is unlikely to correspond to the "adult" undigested EcoRI fragment "b" since in 3 different chickens it migrated slightly ahead from EcoRI fragment b (Fig. 11, lane 1, and results not shown). The three other bands in the

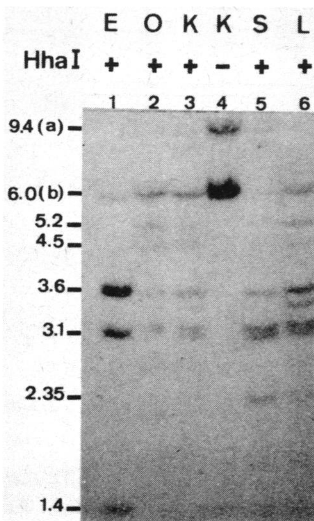


Fig. 11 : HhaI digestion pattern in the region of embryonic and adult β-globin genes. DNA from various tissues from a single laying hen were digested with EcoRI and HhaI as indicated. Electrophoresis was on an 0.8% agarose gel and the excised β-globin sequence of a recombinant β-globin cDNA plasmid (see text) was used as an hybridization probe. Symbols are as in Fig. 2B.

erythrocyte track are 3.6, 3.1. and 1.4 kb-long. HhaI produces a more complex pattern in other tissues (lanes 2, 3, 5 and 6) with some undigested EcoRI fragment "b" and two bands 5.2- and 4.5 kb-long, which are not found in erythrocyte DNA digests and which indicate the presence of sites partially or totally resistant to HhaI digestion. It is interesting to note that a 2.35 kb fragment was easily detected in DNA from spleen, but not in DNA from other tissues (Fig. 11, lane 5 and results not shown on one other animal). HhaI did not digest EcoRI fragments "a" and "b" in sperm DNA (results not shown). We conclude that some HhaI sites in the region of the adult β-globin gene are less methylated in DNA from mature erythrocyte than in DNA from other tissues indluding oviduct.

DISCUSSION

If the biological significance of cytosine methylation at CpG dinucleotides is to fulfill a role in differentiation or in the regulation of gene expression, as previously proposed (4, 5) then one should be able to find variations in the methylation pattern at specific sequences in different cell types or under different functional states of the same cell. We have used

the restriction nuclease approach, first described by Bird and Southern (14) to study the extent of DNA methylation in the regions of four chicken genes : the ovalbumin, conalbumin (ovo-transferrin) and ovomucoid genes, which are expressed at high rate in oviduct under steroid hormone control (the ovotransferrin gene being also expressed constitutively in liver), and the β-globin gene.

The efficiency of digestion by HhaI and HpaII nucleases was examined for 7 restriction sites in the region of the oval-bumin gene and our results are summarized in Table II. We found that two HhaI sites (Hha2 and 3) and one HpaII site (Hpa1) are less methylated in laying hen oviduct than in other tissues. Site Hha2 is located within the ovalbumin gene whereas sites Hpa1 and Hha3 are situated upstream and downstream from the gene, respectively. There are two difficulties in interpreting the degree of

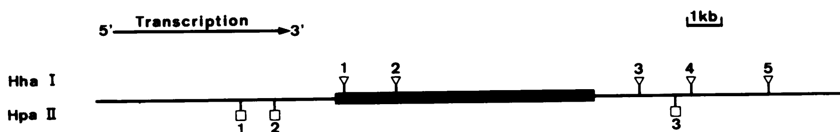
TABLE 2

OCCURENCE AND DIGESTIBILITY OF HhaI AND HpaII SITES IN THE OVALBUMIN GENE REGION

Site	Occurence	Digestion (in %)			
		oviduct	liver	erythrocyte	sperm
Hpa1 (Fig. 6)	Present[(a)]	22 - 42	⩽3	⩽3	nt
Hpa2 (Fig. 6)	Homozygous	nd or ⩽3	4 - 8	4 - 8	nd
Hha2 (Fig. 2,3)	Heterozygous?[(a)]	25 - 32	4 - 7	⩽3	nd
Hha3 (Fig. 3-5)	Homozygous	75 - 85	31 - 51	13 - 32	nd
Hpa3 (Fig. 7)	Homozygous	⩽3	⩽3	40 - 52	nd
Hha4 (Fig. 3-5)	?(b)	nd	nd	nd	nd
Hha5 (Fig. 3-5)	?(b)	nd	nd	nd	nd

(a) Our experiments do not allow to determine the dosage of this site in the animals studied (see text).

(b) This site was detected in two independently cloned ovalbumin genes and might therefore be present in the analyzed cellular DNAs (see text).



This table incorporates data presented in the figures indicated and data not shown obtained on DNA from other animals. The numbers represent the range of the values obtained. nd: no digestion detected; nt : not tested. The map indicates the location of HhaI and HpaII restriction sites with respect to the ovalbumin gene transcription unit (heavy line, see Fig. 1).

undermethylation of these three sites. First, it is likely, but not proven, that sites Hpa1 and Hha2 exist in the heterozygous state in the animals studied (see Result section). A digestibility of about 30% could therefore indicate that 60% of the existing sites are unmethylated. Second, it is known that the tubular gland cells in which ovalbumin is synthesized represent 70 to 80% of the cells of laying hen oviduct (36). Thus, the 80% digestibility of the homozygote site Hha3 most likely reflects complete unmethylation of this site in cells actively synthesizing ovalbumin [see accompanying paper (41) for evidence supporting this interpretation]. It appears therefore that all three sites Hpa1, Hha2 and Hha3 are mostly, if not totally, unmethylated in oviduct cells where the ovalbumin gene is expressed. On the contrary, Hpa1 and Hha2 sites are almost completely, or totally, methylated in all other cells and site Hha3 is partially methylated in these cells, with the exception of sperm where it is totally methylated.

The ovalbumin results indicate that undermethylation, at some sites at least, could be correlated with tissue-specific gene expression. This correlation is also observed in the regions of the conalbumin and ovomucoid genes. These regions which are actively transcribed in oviduct are preferentially undermethylated in oviduct DNA. Furthermore, in liver the ovomucoid region is methylated whereas the conalbumin gene, which is expressed in this organ, is undermethylated. Further studies are required to establish whether the different extents of undermethylation in oviduct and liver are related to different mechanisms of control of conalbumin gene expression or to a greater cell heterogeneity in liver. The β-globin genes are digested best by HhaI in erythrocyte DNA and are not preferentially undermethylated in oviduct when compared to other tissues. This eliminates the possibility that the patterns observed in the other genes could be due to a general deficiency of methyl groups in oviduct DNA, caused for instance by a low methylase activity in this tissue.

Although there is undoubtedly a correlation between gene expression and under- or unmethylation of HhaI and HpaII sites, there are several examples showing that this is not true for all of these sites (Table 2). First, in the ovalbumin gene region, the Hpa3 site is fully methylated in oviduct and in liver, but undermethylated

in the erythrocyte. Knowing that DNaseI sensitivity is a property of "active" chromatin (42, for other refs., see ref. 43), this observation is particularly striking, since the region of the genome containing site Hpa3 is DNaseI-sensitive in chromatin of oviduct, but not in erythrocyte and liver cells (M. Bellard, personal communication and ref. 41). Second, some sites appear to be always methylated, irrespective of the origin of the tissue. This is the case of site Hpa2. It is interesting to note that this site is located between two sites (Hpa1 and Hha2) which are undermethylated in oviduct DNA. As already discussed in the Results section, it is also very likely that sites Hha4 and Hha5 are always fully methylated. Third, in the conalbumin and ovomucoid regions, there are HpaII and HhaI sites which always appear to be completely unmethylated, even in sperm DNA.

Taken together, our results suggest the existence in the chicken genome of three classes of methylatable sites (assuming that what is observed with HhaI or HpaII can be extended to the other methylatable sites). Some sites appear always unmethylated and would correspond to the m^- class as defined by Bird et al. (44); others are always completely (or at least 95%) methylated in all tissues and belong to the m^+ class of Bird et al.; on the other hand, most of the sites in the genes studied are characterized by their variable level of methylation in different tissues, and could belong to a third variable class (m^v). Undermethylation at these sites is in general correlated with tissue-specific gene expression. Taking into account the cell heterogeneity of the tissues which were studied, it is most likely that some of these "variable" sites are not methylated at all in the genome regions which are transcribed [see above and (41)]. On the other hand, it is also clear from our results that some undermethylation can be found in cells in which a given gene is not transcribed (see for instance the site Hha3 in the ovalbumin gene region and the β-globin digestion pattern in Fig. 11, for further discussion of this point see ref. 41).

The methylation patterns that we have observed in the chicken genome appear very different from those described by Bird et al. (44) for the sea urchin genome. They showed that m^+ and m^- sites are compartmentalized to different long sequence tracts. Histone

DNA, rDNA and 5S DNA are found in the m^- compartment irrespective of the tissue (even in tissues which do not express the majority of the histone genes). These authors thus proposed that methylation was an all or none phenomenon. This is clearly not the case for the chicken genome and there is in fact some evidence that the compartmentalization phenomenon could be restricted to some species (Bird, personal communication). On the other hand, recent results on methylation of β-globin genes in various species appear to support our conclusions. The single HpaII site in the rabbit β-globin gene, studied by Waalwijk and Flavell (45) would appear to belong to the m^v class, since it was found 50% methylated in several erythroïd or non erythroïd tissues, but 80% and 100% methylated in brain and sperm DNA, respectively. In that case, however, no correlation was found between the methylation level and the β-globin gene activity. While this paper was being completed, a report by McGhee and Ginder (46) appeared, showing that some unmapped HpaII sites in the chick β-globin gene regions are found unmethylated in erythrocyte DNA, but are at least partially modified in oviduct and brain DNA. Some correlation between expression and undermethylation has also been found in the case of human globin genes (R. Flavell, personal communication) and for integrated adenovirus sequences in transformed cells (47).

The complete methylation in sperm of all of the m^v sites that we have analysed suggest that the tissue-specific methylation patterns are generated during differentiation by loss of methyl groups at given sites, rather than by addition of methyl groups to sites previously unmethylated. Since no demethylase appears to exist (3, 6) this loss would occur by absence of post-replicative modification. The absence of methyl groups at sites belonging to the m^- class, which most likely corresponds to a total impossibility to methylate at these sites, might be related either to a particular higher order structure of the genome around these sites, or to the existence of DNA sequences which cannot be recognized by the DNA methylase. In this respect, it is interesting to note that the SV40 DNA which is synthesized during a lytic infection is unmethylated and cannot be methylated *in vitro* using a purified methylase preparation (48, 49).

Although gene expression and undermethylation at m^v sites

are correlated, it is unlikely that this DNA modification is directly involved in the rapid modulation of gene transcription, since in the absence of cell division, methylation appears to be essentially an irreversible process (3). In fact, we have not observed any modification in the HhaI and HpaII digestion patterns of chicken oviduct DNA 48 hr after oestradiol withdrawal under conditions where ovalbumin mRNA synthesis is completely stopped (50). Our results (Fig. 11) and those of McGhee and Ginder (46) on the chicken β-globin gene indicate also that at least some sites in this gene region are unmethylated in DNA from adult erythrocytes which no longer synthesize β-globin mRNA. It appears therefore that, as previously suggested (4, 5), differences in methylation at the m^V sites could be related to the process of the commitment of a cell to express a given sequence of the genome. Wheter undermethylation is a cause or a consequence of cell differentiation is unknown at the present time. It could be viewed as a process which would create or delete signals recognized by proteins involved either directly in the regulation of transcription or indirectly by modifying the higher order structure of chromatin. In this respect, it is worth mentioning that methylation of a single cytosine alters dramatically the binding of lac repressor to operator DNA (51). Alternatively, undermethylation could be a consequence of the differentiation process which by obliterating (by protein binding for instance) certain DNA sites will prevent their methylation. In any case, it is important to remember that methylation at some sites appears to be dispensable, since the m^V site Hha2 can be replaced in some animals by the non-methylable sequence GTGC (25).

Finally, there is in the ovalbumin gene a striking number of allelic variants for HhaI and HpaII sites (at least 5 variant sites out of 8 sites which have been characterized). We do not know whether the same is true for the other genes which have been studied less thoroughly. Allelic variants have been found, but not in such a high proportion, for other restriction sites in various genes(17, 52, 53) and they can be useful markers in genetic studies, notably for prenatal diagnosis (54). If confirmed by further work, such a high variability in restriction sites for "CpG" nucleases would support the hypothesis (55) that the stri-

kingly low "CpG" content in DNA from vertebrates might be due to a higher mutation rate for this dinucleotide [methyl cytosine being deaminated to thymine (4,56)]. Methyl cytosines have been shown recently to be sites of mutational hot spots in E.coli (57). Analysis of sequences at variant restriction sites will indicate whether a similar mechanism also operates in higher cells.

Biohazards associated with the experiments described in this publication were examined previously by the French National Control Committee. The experiments were carried out under L3 B1 conditions (Le Progrès Scientifique, N° 191, Nov. Dec. 1977).

ACKNOWLEDGEMENTS

We wish to thank M.C. Chanal for excellent technical assistance, Drs P. Gerlinger and P. Humphries for generous gifts of restriction nucleases and Drs. F. Gannon and R. Palmiter for their editorial help. This work was supported by grants from the INSERM (CRT 76.5.468 and 76.5.462, the CNRS (ATP n° 3558) and the Fondation pour la Recherche Médicale Française.

REFERENCES

1. Doskocil, J. and Sorm, F. (1962) Biochim. Biophys. Acta 55, 953-959.
2. Grippo, P., Iaccarino, M., Parisi, E., and Scarano, E. (1968) J. Mol. Biol. 36, 195-208.
3. Burdon, R.H., and Adams, R.L.P. (1968) Biochim. et Biophys. Acta 174, 322-329.
4. Holliday, R., and Pugh, J.E. (1975) Science 187, 226-232.
5. Riggs, A.D. (1975) Cytogenet. Cell Genet. 14, 9-25.
6. Taylor, J.H. (1979) in Molecular Genetics Vol. III Chromosomal Structure, 89-115, Acad. Press New-York.
7. Vanyushin, B.F., Mazin, A.L., Vasilyev, V.K., and Belozersky, A.N. (1973) Biochim. Biophys. Acta 299, 397-403.
8. Pollock, J.M., Swihart, M.Jr., and Taylor, J.H. (1978) Nucleic Acids Res. 5, 4855-4864.
9. Adams, R.LP., McKay, E.L., Douglas, J.T., and Burdon R.H., (1977) Nucleic Acids Res. 4, 3097-3108.
10. Solage, A., and Cedar, H. (1978) Biochemistry 17, 2934-2938.
11. Gautier, F., Bünemann, H., and Grotjahn, L. (1977) Eur. J. Biochem. 80, 175-183.
12. Browne, M.J., Cato, A.C.B., and Burdon, R.H. (1978) Febs Letters 91, 69-73.
13. Mann, M.B., and Smith, H.O. (1977) Nucleic Acids Res. 4, 4211-4221.
14. Bird, A.P. and Southern E.M., (1978) J. Mol. Biol. 118, 27-47.
15. Dawid, I.B., Brown, D.D., and Reeder, R.H. (1970) J.Mol. Biol. 51, 341-360.

16. Breathnach, R., Mandel, J.L., and Chambon, P. (1977), Nature 270, 314-319.
17. Mandel, J.L., Breathnach, R., Gerlinger, P., LeMeur, M., Gannon, F., and Chambon, P. (1978) Cell 14, 641-653.
18. Gross-Bellard, M., Oudet, P., and Chambon, P. (1973) Eur. J. Biochem. 36, 32-38.
19. Sumegi, J., Breedveld, D., Hossenlopp, P., and Chambon, P. (1977) Biochem. Biophys. Res. Commun. 76, 78-86.
20. Garapin, A.C., LePennec, J.P., Roskam, W., Perrin, F., Cami, B., Krust, A., Breathnach, R., Chambon, P., and Kourilsky, P. (1978) Nature 273, 349-354.
21. Garapin, A.C., Cami, B., Roskam, W., Kourilsky, P., LePennec, J.P., Perrin, F., Gerlinger, P., Cochet, M., and Chambon, P. (1978) Cell 14, 629-639.
22. Gannon, F., O'Hare, K., Perrin, F., LePennec, J.P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B., and Chambon, P. (1979) Nature 278, 428-434.
23. Breathnach, R., Benoist, C., O'Hare, C., Gannon, F., and Chambon, P. (1978) Proc. Natl. Acad. Sci. USA 75, 4853-4857.
24. Humphries, P., Cochet, M., Krust, A., Gerlinger, P., Kourilsky, P., and Chambon, P. (1977) Nucleic Acids Res. 4, 2389-2406.
25. McReynolds, L., O'Malley, B.W., Nisbet, A.D., Fothergill, J.E., Givol, D., Fields, S., Robertson, M. and Brownlee, G.G. (1978) Nature 273, 723-728.
26. Chambon, P., Benoist, C., Breathnach, R., Cochet, M., Gannon, F., Gerlinger, P., Krust, A., LeMeur, M. LePennec, J.P., Mandel, J.L., O'Hare, K., and Perrin, F. in Proceedings of the 11th Miami Winter Symp. "From Gene to Protein : Transfer in normal and abnormal cells". Academic Press (New-York) 1979. in press.
27. Dodgson, J.B., Strommer, J., Engel, J.D. (1979) Cell, 17, 879-887.
28. Waalwijk, C., and Flavell, R.A. (1978) Nucleic Acids Res. 5, 3231-3236.
29. Singer, J., Roberts-Ems, J., Riggs, A.D. (1979) Science 203, 1019-1020.
30. Lee, D.C., McKnight, S.G., and Palmiter, R.D. (1978) J. Biol. Chem. 253, 3494-3503.
31. McKnight, S.G., Lee, D.C., and Palmiter, R.D., J. Biol. Chem. in press.
32. Perrin, F. Cochet, M., Gerlinger, P., Cami, B., LePennec, J.P., and Chambon, P. (1979), Nucleic Acids Res. 6, 2731-2748.
33. Cochet, M., Gannon, F., Hen, R., Maroteaux, L., Perrin, F., and Chambon, P. Nature, in press.
34. Cochet, M., Perrin, F., Gannon, F., Krust, A., Chambon, P., and McKnight, G.S., Lee, D.C., Mayo, K.E., and Palmiter, R. (1979) Nucleic Acids Res. 6, 2435-2453.
35. O'Malley, B.W., McGuire, W.L., Kohler, P.O. and Korenman, S.G. (1969) Recent Prog. Horm. Res. 25, 105-160.
36. Palmiter, R.D. (1972) J. Biol. Chem. 247, 6450-6461.
37. Hynes, N.E., Groner, B., Sippel, A.E., Nguyen-Huu, M.C., and Schütz, G. (1977) Cell 11, 923-932.
38. Nordstrom, J.L., Roop, D.R., Tsai, M.J., and O'Malley, B.W. (1979) Nature 278, 328-331.
39. Catterall, J.F., Stein, J.P., Lai, E.C., Woo, S.L.C., Dugaiczyk, A., Mace, M.L., Means, A.R., and O'Malley, B.W.

(1979) Nature 278, 323-327.
40. Engel, J.D., and Dodgson, J.B. (1978) J. Biol. Chem. 253, 8239-8246.
41. Kuo, T., Mandel, J.L., and Chambon, P. accompanying paper.
42. Weintraub, H., and Groudine, M. (1976) Science 193, 848-856.
43. Mathis, D., Oudet, P., and Chambon P. (1980) Progress in Nucleic Acids Research and Molecular Biology, 24, in press.
44. Bird, A.P., Taggart, M.H., and Smith, B.A. (1979) Cell 17, 889-901.
45. Waalwijk, C., and Flavell, R.A. (1978) Nucleic Acids Res. 5, 4631-4641.
46. McGhee, J.D., and Ginder, G.D. (1979) Nature 280, 419-420.
47. Sutter, D., and Doerfler, W. (1979) Cold Spring Harbor Symp. Quant. Biol. 44, in press.
48. Turnbull, J.F., and Adams, R.L.P. (1976) Nucleic Acids Res. 3, 677-695.
49. Simon, D., Grunert, F., Acken, U.V., Döring, H.P., and Kröger, H. (1978) Nucleic Acids Res. 5, 2153-2167.
50. McKnight, G.S., and Palmiter, R.D. (1979) J. Biol. Chem., in press.
51. Fisher, E.F., and Caruthers, M.H. (1979) Nucleic Acids Res. 7, 401-416.
52. Lai, E.C., Woo, S.L.C., Dugaiczyk, A., and O'Malley, B.W. (1979) Cell 16, 201-211.
53. Jeffreys, A.J. (1979) Cell 18, 1-10.
54. Kan, Y.W., and Dozy, A.M. (1978) Proc. Natl. Acad. Sci USA 75, 5631-5635.
55. Salser, W. (1977) Cold Spring Harbor Symp. Quant. Biol. 42, 985-1002.
56. Scarano, E. (1971) Adv. Cytopharmacol. 1, 13-24.
57. Coulondre, C., Miller, J.H., Farabaugh, P.J., and Gilbert, W. (1978) Nature 274, 775-780.