Muller, W. E. G., & Zahn, R. K. (1976) *Prog. Biochem. Pharmacol. 11*, 28-47.
Muller, W. E. G., Yamazaki, Z., Breter, H., & Zahn, R. K. (1972) *Eur. J. Biochem. 31*, 518-525.
Newman, S. A. (1979) *J. Theor. Biol. 79*, 55-56.
Niwa, O., & Moses, R. E. (1981) *Biochemistry 20*, 238-243.
Olins, A. L., & Olins, D. E. (1974) *Science (Washington, D.C.) 183*, 330-332.
Oudet, P., Gross-Bellard, M., & Chambon, P. (1975) *Cell (Cambridge, Mass.) 4*, 281-300.
Panayotatos, N., & Wells, R. D. (1981) *Nature (London) 289*, 466-470.
Piper, P. W., Celis, J., Kaltaft, K., Leer, J. C., Nielson, O. F., & Westergaard, O. (1976) *Nucleic Acids Res. 3*, 493-505.
Povirk, L. F., Hogan, M., & Dattagupta, N. (1979) *Biochemistry 18*, 96-101.
Renz, M., Niehls, P., & Hozier, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 74*, 1879-1883.
Sastry, K. S. R., & Hallee, G. J. (1978) *International Symposium of Biomolecular Structure, Conformation, Function and Evolution*, Madras, India (presented paper).
Sausville, E. A., Peisach, J., & Horwitz, S. B. (1976) *Biochem. Biophys. Res. Commun. 91*, 871-877.
Setlow, P., Bratlag, D., & Kornberg, A. (1972) *J. Biol. Chem. 247*, 224-231.
Sobell, H. M., Tsai, C. C., Jain, S. G., & Gilbert, S. G. (1977) *J. Mol. Biol. 114*, 333-365.
Strong, J. E., & Crooke, S. T. (1978) *Cancer Res. 38*, 3322-3326.
Sutcliffe, J. G. (1978) *Cold Spring Harbor Symp. Quant. Biol. 41*, 77-90.
Suzuki, H., Nagai, K., Yamaki, H., Tanaka, N., & Umezawa, H. (1969) *Jpn. J. Antibiot. 22*, 446-449.
Takeshita, M., Horwitz, S. B., & Grollman, A. P. (1974) *Virology 60*, 455-465.
Takeshita, M., Grollman, A. P., Ohtsubo, E., & Ohtsubo, H. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5983-5987.
Umezawa, H., Maeda, K., Takeuchi, T., & Oakami, Y. (1966) *J. Antibiot., Ser. A 19*, 200-209.
Weintraub, H., & Groudine, M. (1976) *Science (Washington, D.C.) 193*, 848-856.
Worcel, A. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 313-324.

Characterization of Deoxyribonucleic Acid Sequences at the 5′ and 3′ Borders of the 100 Kilobase Pair Ovalbumin Gene Domain†

William E. Stumph, Melvyn Baez, Wanda G. Beattie, Ming-Jer Tsai, and Bert W. O'Malley*

ABSTRACT: The ovalbumin and the ovalbumin-related X and Y genes are coordinately expressed in the chicken oviduct and are linked within a 100 kilobase pair (kbp) chromosomal domain of DNA which is preferentially sensitive to DNase I [Lawson, G. M., Knoll, B. J., March, C. J., Woo, S. L. C., Tsai, M.-J., & O'Malley, B. W. (1982) *J. Biol. Chem. 257*, 1501-1507; Lawson, G. M., Tsai, M.-J., & O'Malley, B. W. (1980) *Biochemistry 19*, 4403-4411]. At each end of this domain, the chromatin undergoes a structural transition such that the DNA becomes relatively more resistant to DNase I. In order to understand the mechanisms underlying this structural transition, we have been studying the characteristics of DNA sequences which are found at the 5′ and 3′ borders of the domain. We have discovered that members of a dispersed repetitive DNA sequence family (termed the CR1 family) are located near or within the transition regions of DNase I sensitivity; in addition, the CR1 sequences found at opposite ends of the domain exist in inverse orientations with respect to each other. These observations suggest that the CR1 sequences may possibly play a role in defining the ends of the ovalbumin gene DNase I sensitive domain. A consensus CR1 sequence is presented on the basis of a comparison of four homologous genomic sequences which are members of the CR1 family. We also searched for, but failed to obtain, any evidence that specific DNA rearrangements are involved in the expression of the ovalbumin gene. However, a region at the far 3′ end of the domain, approximately 30 kbp downstream from the ovalbumin gene, is highly enriched for *Hpa*II/*Msp*I restriction sites compared to most regions of the chicken genome. More importantly, this same region exhibits a pattern of methylation which is highly variable from one chicken tissue to another. In particular, DNA in the oviduct is specifically undermethylated in this region when compared to other tissues. Thus, this DNA at the 3′ border may serve as a control region involved in determining the structural state of the chromatin domain in which the ovalbumin gene family resides.

Evidence has accumulated over the past several years that active gene sequences are preferentially digested by the enzyme DNase I. For example, when compared to the bulk of the chromatin DNA sequences, globin gene sequences are preferentially sensitive to DNase I in erythroid cell nuclei but not in nuclei isolated from other cell types (Weintraub & Groudine, 1976). Similarly, in the chicken oviduct where the ovalbumin gene is expressed, those DNA sequences which code for ovalbumin are preferentially sensitive to DNase I (Garel & Axel, 1976). Such preferential sensitivity to DNase I is not observed in other chicken tissues where the ovalbumin gene is not expressed. Thus, the DNase I sensitivity of a gene is correlated with the ability of that gene to be expressed in any given cell type [see Lawson et al. (1982) for additional references and examples]. This preferential sensitivity, which is observed for active genes when nuclei are digested with DNase I until about 20% of the DNA is rendered acid soluble, presumably reflects a modification in the packaging or the configuration of the nucleosomes in those regions of the chromatin which contain active genes. The altered chromatin structure and preferential DNase I sensitivity measured in this

† From the Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030. *Received August 19, 1982.* This work was supported by the National Institutes of Health (Grant HD-08188) and the Baylor Center for Reproductive Biology (Grant HD-07495).

0006-2960/83/0422-0306$01.50/0

way are not confined to the transcribed sequences themselves, however, but also extend into the flanking DNA as well. This has been shown for the chicken ovalbumin and ovomucoid genes (Lawson et al., 1980), globin genes (Stalder et al., 1980), and the glyceraldehyde-3-phosphate dehydrogenase gene (M. Alevy, unpublished results).

We have recently reported that in the case of the ovalbumin gene this preferential DNase I sensitivity extends over a region approximately 100 kilobase pairs (kbp)[1] in length (Lawson et al., 1982). This chromosomal domain contains the ovalbumin gene and the ovalbumin-related X and Y genes which are coordinately expressed in the chicken oviduct in response to steroid hormones. At each end of this 100-kbp domain, the chromatin undergoes a structural transition and becomes relatively resistant to DNase I [see Lawson et al. (1982) and Figure 6 below]. We are interested in determining at the molecular level what signals are responsible for establishing the boundaries of DNase I sensitive genetic domains. We also would like to know what mechanisms are responsible for conferring the active DNase I sensitive configuration upon those specific segments of the chromatin which contain active genes. In particular, we would like to understand how this is accomplished for the ovalbumin gene domain chromatin during the normal development and differentiation of the chicken oviduct.

As a starting point to address these questions, we have examined the DNA near the structural transition regions of the chromatin for possible signals which could be involved in defining or establishing the ovalbumin gene DNase I sensitive domain. We have found that members of a certain middle repetitive DNA sequence family [termed the CR1 family (Stumph et al., 1981)] are specifically located at each end of the ovalbumin gene domain. Since these CR1 sequences are located either near or within the transition regions of chromatin structure, it is possible that they could play a role in defining or opening up the domain. We have also determined the individual nucleotide sequences of these related elements, and we present a consensus sequence for the CR1 family based upon the available sequence data.

In addition, we have also explored the possibility that establishment of the DNase I sensitive domain in oviduct cells is dependent upon tissue-specific rearrangements of the DNA in the boundary regions of the domain. However, we have not obtained any evidence to support that hypothesis. We have also examined the methylation of cytidine residues at *Hpa*II/*Msp*I sites throughout the ovalbumin gene domain. Interestingly, a region at the far 3′ border of the domain, approximately 30 kbp downstream from the ovalbumin gene, is specifically undermethylated in the oviduct as compared to other chicken tissues. Thus, it is possible that this region could represent a distal controlling element that may be important for determining the structural configuration of the ovalbumin domain chromatin as a function of cellular differentiation and development.

Experimental Procedures

DNA Cloning and Sequencing. The isolation of clones containing overlapping segments of chicken DNA from the ovalbumin region of the genome has been previously described (Lawson et al., 1982). All clones shown in Figure 1, except CL734, were isolated from a partial *Alu*I/*Hae*III library (Dodgson et al., 1979). CL734 was isolated from a partial *Eco*RI library. In all cases, the chicken genomic DNA was cloned into the *Eco*RI site of λ Charon 4A. For sequencing and other detailed studies, selected fragments from the Charon 4A clones were subcloned into the plasmid pBR322. Once the fragments were subcloned, more detailed restriction maps were obtained from double- and triple-digestion patterns. So that the CR1 sequences within the subcloned fragments could be more precisely mapped, the clones were restricted with combinations of enzymes, run on 2% agarose gels, and transferred to nitrocellulose filters (Southern, 1975). Filter hybridizations were then carried out by using the CR1 sequence specific probe described below. On the basis of those results, subfragments were then selected and isolated for sequencing. DNA sequencing was carried out by using the method of Maxam & Gilbert (1977).

Preparation of DNA from Various Chicken Tissues. For comparison of DNA from different tissues, in any particular experiment a single hen was sacrificed and various organs were removed. High molecular weight DNA was then purified from isolated nuclei after extensive proteinase K digestion and subsequent extraction with phenol/chloroform/isoamyl alcohol. Traces of RNA were removed by treating with RNase A that had previously been heated at 95 °C to destroy any contaminating DNase activity. These DNA preparations barely migrated into 0.8% agarose gels prior to any restriction enzyme treatment.

The following steps were taken in order to assure that the *Hpa*II and *Msp*I digestions of genomic DNA were complete. An aliquot of each digestion mixture was combined with 1 μg of φX174 DNA without adding any additional enzyme. The φX174 DNA, which could be observed after the gel was stained with ethidium bromide, thus served as an internal standard for complete digestion. As a further control in one experiment, DNAs from different tissues were digested with 2 units of enzyme per μg of DNA for 4 h, and the digestions were then interrupted by cooling to 0 °C. An aliquot was then taken and further digested with a second addition of the same enzyme (2 units/μg of DNA) for 16 h. After Southern transfer and hybridization, no change was noted in the banding patterns of the 16-h digests compared to those of the 4-h digests.

Southern Hybridizations. For Southern filter hybridizations (Southern, 1975), samples were electrophoresed in agarose gels and transferred to nitrocellulose. Approximately 1 μg of cloned DNA or approximately 6–12 μg of restriction enzyme digested total nuclear DNA was loaded per gel lane. The filters were hybridized overnight to labeled probe at 68 °C in 6 × SSC, 1 mM EDTA, 0.5% $NaDodSO_4$, and 0.04% each of bovine serum albumin, Ficoll-400, and poly(vinylpyrrolidone) 360 in the presence of either 100 μg/mL *Escherichia coli* DNA or 100 μg/mL yeast RNA to further reduce background. The nitrocellulose filters were then washed at 68 °C in three changes of 1 × SSC and 0.5% $NaDodSO_4$ over a period of 4–6 h prior to being dried and mounted for autoradiography.

Hybridization Probes. Restriction fragments were isolated from plasmids or directly from the phage clones by either agarose or acrylamide gel electrophoresis. In most cases, the specific fragments utilized as probes are described in the figures and figure legends. The CR1 sequence hybridization probe used in Figure 1 (and also for mapping the subclones) was the 5′-terminal *Hin*fI fragment from a previously described clone which was designated pU1 2.5 (Roop et al., 1981; Stumph et al., 1981). Clone pU1 2.5 contains a CR1 family repetitive DNA sequence upstream from a gene which codes

[1] Abbreviations: kbp, kilobase pair(s); SSC, 0.15 M NaCl and 0.015 M sodium citrate, pH 7; EDTA, ethylenediaminetetraacetic acid; $NaDodSO_4$, sodium dodecyl sulfate; bp, base pair(s); HMG, high mobility group.

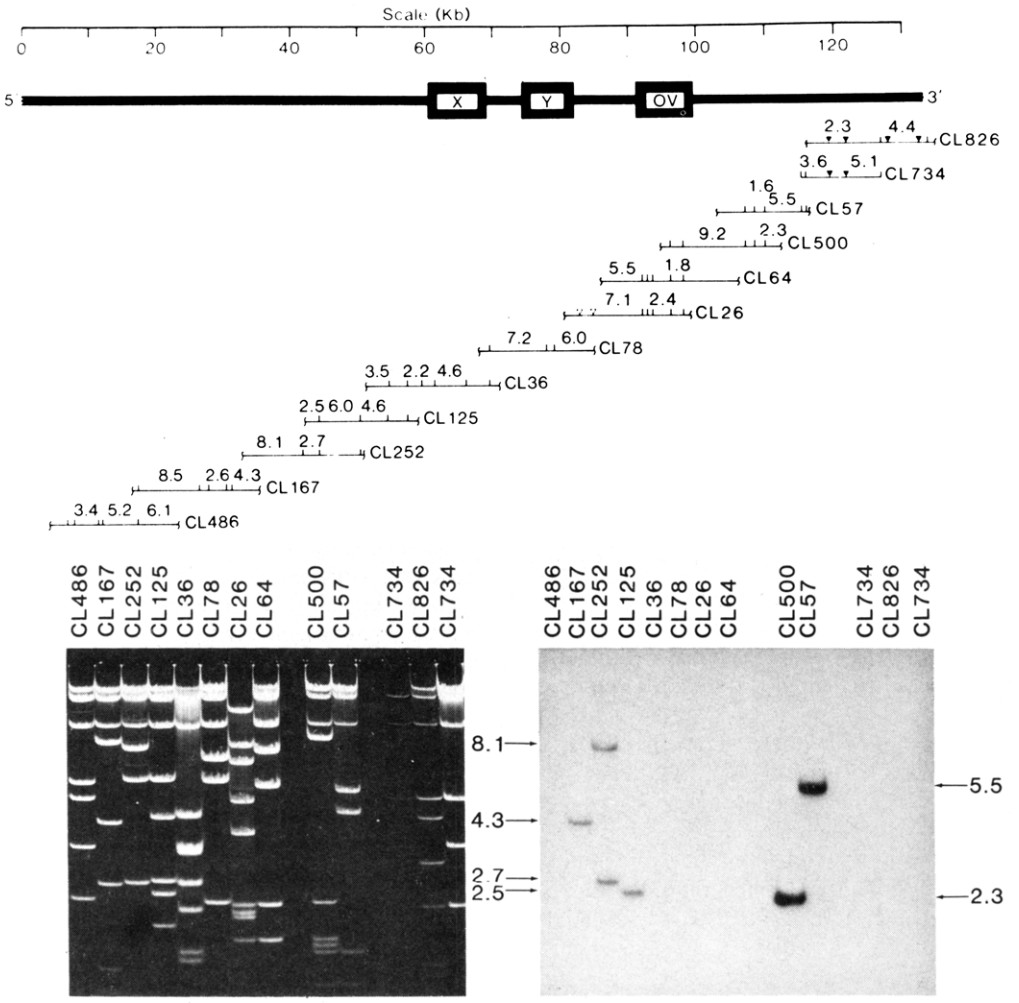

FIGURE 1: Location of CR1 family repetitive DNA sequences in the ovalbumin gene domain. The upper part of the figure is a schematic diagram of the ovalbumin gene region of the chicken genome. Transcription of the X, Y, and ovalbumin genes occurs from left to right. Overlapping clones are depicted which contain genomic DNA from this region of the genome. The squiggles at the ends of the cloned fragments represent artificial *Eco*RI sites that were introduced for cloning into Charon 4A vectors. Vertical bars denote naturally occurring *Eco*RI sites. These clones were digested with *Eco*RI and electrophoresed on a 1% agarose gel. In selected clones where sites for *Bam*HI (open triangles) or *Xho*I (closed triangles) are shown, digestions were carried out with those enzymes in addition to *Eco*RI before being run on the gel. The lower left panel is a photograph of the gel after being stained with ethidium bromide. (CL734 was run in two lanes because of an uncertainty in the DNA concentration of that sample.) The lower right-hand panel is a Southern blot of the gel after hybridization to a CR1 family repetitive DNA sequence probe as described under Experimental Procedures. The fragments which hybridized to the CR1 probe are identified by the arrows.

for chicken U1 RNA. A *Hin*fI restriction map of this clone and the location of the CR1 sequence have been previously described in detail (Stumph et al., 1981).

The various restriction fragments were labeled by nick translation with [α-^{32}P]dCTP and [α-^{32}P]TTP as previously described (Roop et al., 1978). Depending upon the particular experiment, we used between 8×10^6 and 80×10^6 cpm in each Southern hybridization reaction.

Results

CR1 Family Repetitive DNA Sequences in the Ovalbumin Gene Domain. We have recently reported the initial characterization of a family of short dispersed repetitive DNA sequences present in the chicken genome (Stumph et al., 1981). We have termed this family the CR1 family and have reported nucleotide sequences for two family members. One of those family members was present in a clone which also contained a gene coding for chicken U1 RNA (Stumph et al., 1981; Roop et al., 1981). A fragment containing specifically the CR1 sequence was isolated from that clone and used as a hybridization probe in an experiment to determine whether CR1 family sequences might also be present at certain locations within the ovalbumin gene domain. This was done by digesting clones from the ovalbumin region of the genome (shown in the upper part of Figure 1) with the designated restriction enzymes and separating the fragments on a 1% agarose gel. A photograph of this gel after being stained with ethidium bromide is shown in the lower left panel of Figure 1. The DNA in the gel was transferred to a nitrocellulose filter by the method of Southern (1975), and the filter was then hybridized to the CR1 sequence probe. The hybridization pattern shown in the lower right-hand panel of Figure 1 shows that the CR1 probe hybridized to specific fragments whose locations within the ovalbumin gene domain are known. It can be concluded from these results that CR1 sequences are present in at least three distinct locations within the ovalbumin gene domain. The six bands which hybridize represent three different pairs of fragments in which the two members of each pair overlap with each other. For example, the 4.3 (CL167) and the 8.1 (CL252) fragments overlap each other, so presumably the same sequence is being observed in each case. The 2.7 (CL252) and the 2.5 (CL125) fragments overlap, and the 2.3 (CL500) and the 5.5 (CL57) fragments overlap. The similarity of intensities observed for each of the fragments

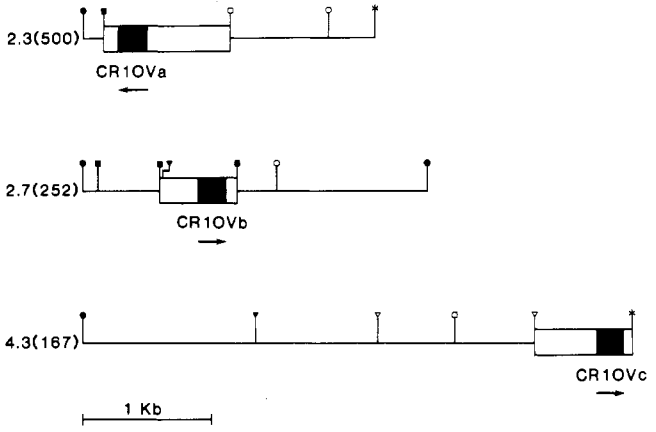

FIGURE 2: Restriction maps of subcloned fragments which contain CR1 sequences. The three fragments are oriented in the same direction as shown in Figure 1. Complete restriction maps are shown for each of the enzymes *Eco*RI (●), *Pst*I (■), *Ava*I (□), *Hin*dIII (○), *Bgl*II (▼), and *Kpn*I (▽), except in the case of the 4.3 (167) fragment which contains multiple *Pst*I sites which have not been precisely mapped. Asterisks indicate artificial *Eco*RI sites. The enlarged boxes denote the regions that were selected for sequencing based upon their hybridization to a CR1 sequence probe. The darkened areas within the boxes denote the portions that are presented in the sequence comparisons shown in Figure 3. The horizontal arrows indicate the orientations of the CR1 sequences within each fragment.

within a pair makes it unlikely that additional CR1 sequences are present in the nonoverlapping sequences that are not shared. The specific hybridization of the CR1 probe to the above-mentioned fragments was confirmed by results from several independent experiments similar to the one that is shown in Figure 1. Therefore, in order to carry out more detailed studies on the CR1 sequences, we subcloned into the plasmid pBR322 three fragments [4.3 (CL167), 2.7 (CL252), and 2.3 (CL500)] which correspond to the three locations of CR1 sequence hybridization.

Nucleotide Sequence and Structure of CR1 Family Members within the Ovalbumin Gene Domain. It appears from the results shown in Figure 1 that the CR1 sequences are not scattered randomly throughout the domain. Rather, they are localized toward each end of the domain where there is a transition in chromatin structure as measured by a change in DNase I sensitivity (Lawson et al., 1982). Since this correlation suggested a potential functional role for the CR1 sequences as possibly being involved in opening up the domain or in defining its boundaries, we have investigated the CR1 sequences in more detail. First, detailed restriction maps were obtained for each subcloned fragment. The individual clones were then digested by using appropriate restriction enzymes, and the fragments which contain CR1 sequences were identified by Southern blot hybridizations (not shown) by using the CR1 sequence probe described under Experimental Procedures. Fragments were then chosen for sequencing, and the results are summarized in Figures 2 and 3.

Figure 2 shows restriction maps of the fragments that were subcloned. The fragments are oriented in the same direction as shown in Figure 1, such that left to right is in the same direction as the direction of transcription of the X, Y, and ovalbumin genes in the adjoining DNA of the domain. The fragments that were sequenced are depicted by the enlarged boxes. The solid areas of those boxes represent the portions of the sequenced fragments that are shown in the comparison in Figure 3. It should be noted that the CR1OVa sequence, located at the 3′ end of the domain, exists in the genome in an orientation opposite to that of the CR1OVb and CR1OVc sequences, which are found at the 5′ end of the domain. That is, the sequences for CR1OVb and CR1OVc shown in Figure 3 are the sequences from the noncoding strand with respect to the X, Y, and ovalbumin genes, whereas the CR1OVa sequence is taken from the coding strand with respect to those structural genes. Figure 3 presents a detailed comparison at the nucleotide level of the homologous regions of the CR1OVa, CR1OVb, and CR1OVc sequences. For additional comparative purposes, Figure 3 also includes the nucleotide sequence of a CR1 element (CR1U1a) which is found in a clone containing a gene for U1 RNA. This CR1U1a sequence is what was used as the hybridization probe to identify the CR1 sequences in the ovalbumin gene domain. The consensus sequence shown along the top line of each row was formulated by choosing the most common nucleotide at any one position after the four sequences were aligned for maximum homology.

The CR1 sequence family has an interesting structure that is not commonly observed in families of short dispersed repetitive DNA sequences that have been characterized in other organisms. The CR1 repeat seems to be divided into two separate regions of high homology that are separated by a

```
          1                   20                  40                  60                  80
Consensus TNTA--CTTTG---A GAGTGA-CATG AGGTCACTGG AACAGGCTGC CCAGAGAGGT TGTGGATGCC ----CCATCCCTGG AGGYATTCAA
CR1U1a     T     A          G --       -G          -                              CCGT            TG
CR1OVa     G-    G -     CT  T         -            T      A                                       C
CR1OVb     A  TT   ACAC     G C A          CA  C       T        A C     -  G                      C C
CR1OVc     C              G  A   -    --                    G             GT T      T  T     AAT

          81                 100                 120                 140                 160
Consensus RGCCAGGYTG GATGTGGC---CC TGGGCARCCT GRGCTGGTNC T-A--G-GGAGTCT GATGNCCCTG CNCANAGCAT GGGGGANTGG
CR1U1a    G  G T    C  -            G    G     A  G AT T          G      -- ---- -      GG-
CR1OVa    G    C           T   AA  G    GT  A -- - - - - -  G A      A  T   -      T  A
CR1OVb    A    T      G            A    A     C   -  - --  - - -  --     C  C       --T
CR1OVc    AA  T CC    --  - TTT    T   A   -A-  ----           -   -- ----  - ------ T G  A T  G T

          161                180                 200                 220                 245
Consensus RRCTGCAGAT G-RTGNCCATT GAGGTCCCTT CCAACCCAGR CCATTCTATG ATTCT--AT-NAN CT-ACATCACT TAGACAGGGTCAGAT
CR1U1a    GGT        TTG  A  C                    G       G       -    T C   G -  -   T -     -- T
CR1OVa    AA  --       A ---      T    T   T       G                    --- -- -G    -       TT   T T
CR1OVb    GG   -  -    G- -- T                    AA   T          AG TA  AA T                TC  -  A   CC
CR1OVc    AA   -  -  A A CT   --                  TCTA  A           C     ---   A  GG-C       -        A
```

FIGURE 3: Nucleotide sequence comparison of four chicken CR1 sequences. The regions of highest homology shared by four CR1 sequences are shown. The consensus sequence was determined by choosing the most common nucleotide at any one position when the four individual sequences were aligned for maximum homology. Only differences from the consensus sequence are shown for the individual sequences in each of the lines below the consensus sequence. R stands for purine, Y stands for pyrimidine, and N was used when there was no clear consensus at a position. Dashes in the individual sequences indicate deletions of nucleotides in the individual sequences when compared to the consensus sequence. Dashes in the consensus sequence, on the other hand, indicate that nucleotides are present at those positions in only one of the individual sequences examined.

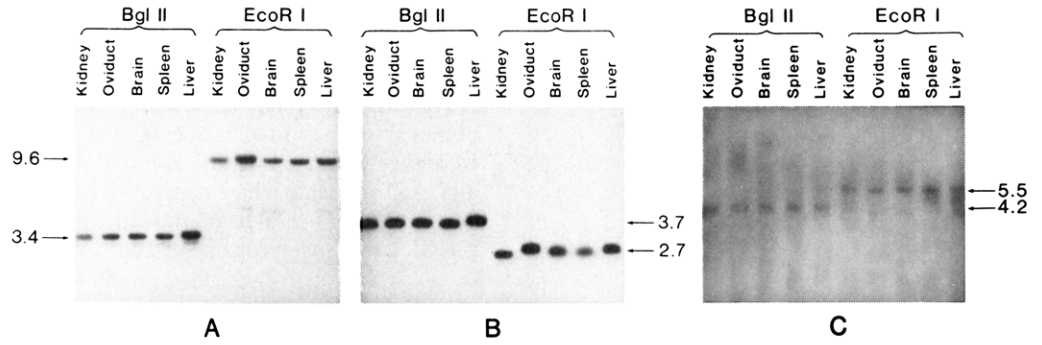

FIGURE 4: No evidence for tissue-specific DNA rearrangements at the ends of the ovalbumin gene domain in the vicinity of CR1 sequences. Total nuclear DNA was isolated from various tissues of a single hen and then digested with either *Bgl*II or *Eco*RI. The samples were electrophoresed in 0.7% agarose gels and blotted onto nitrocellulose filters (Southern, 1975). The filters were then hybridized to ^{32}P-labeled restriction fragments isolated from various CR1-containing clones. In the experiment in panel A, the probe was the 1.2-kbp internal *Kpn*I/*Kpn*I fragment isolated from the clone 4.3 (167) (see Figure 2 for restriction maps). In panel B, the probe was the 1.2-kbp *Hin*dIII to *Eco*RI fragment from the right-most end of the clone 2.7 (252). In panel C, the probe was the 0.4-kbp *Hin*dIII to *Eco*RI fragment from the right-most end of the clone 2.3 (500). In all cases, neither *Bgl*II nor *Eco*RI possesses cutting sites between the probe and the neighboring CR1 sequence. Therefore, any major rearrangements of the DNA which occur tissue specifically near the CR1 sequences should be detected as a variation in band size in different tissues. No such variations are observed.

divergent region of relatively low homology. The four sequences share approximately 80% sequence homology with each other, beginning at position 18 of the consensus sequence and continuing to about position 118. However, this is followed by a region extending from about 119 to 178 where the homology is significantly lower and consensus sequence is very poorly defined. This divergent region is characterized by numerous deletions in the individual sequences when they are compared to the relatively poorly defined consensus sequence within this region. Between positions 179 and 217, however, all four sequences are again very highly homologous (80–90%). Following these 39 bp of highly conserved sequence, the homology gradually decreases. It is not possible to identify any precisely defined positions as marking the boundaries between the CR1 sequences and the flanking DNA. Preceding position 18 of the consensus sequence and following position 217, there is a certain degree of homology evident, but at a lower level. When sequences further upstream and downstream are examined (not shown), no significant degree of homology among the different sequences is evident in the flanking DNA.

No Evidence for DNA Rearrangements Involved in the Expression of the Ovalbumin Gene. It is conceivable that tissue-specific DNA rearrangements could occur in the ovalbumin gene domain during the process of oviduct differentiation and cellular commitment. If so, this could be part of the mechanism responsible for the establishment of the DNase I sensitive state. We have previously shown (Stumph et al., 1981) that CR1 sequences share a certain degree of sequence homology with mammalian dispersed repetitive DNA sequences such as the human Alu family (Rubin et al., 1980). Because many Alu-type sequences exhibit structural analogies to known transposable elements (Duncan et al., 1981; Haynes et al., 1981; Grimaldi & Singer, 1982), it has been suggested by a number of groups that Alu sequences are movable genetic elements and that they might play a role in promoting DNA rearrangements (Haynes et al., 1981; Van Arsdell et al., 1981; Calabretta et al., 1982). It was therefore conceivable that CR1 sequences could be involved in promoting such DNA rearrangements in the transition regions of the ovalbumin gene domain during oviduct cellular differentiation.

We have done a series of experiments to test that hypothesis. Nuclear DNAs were prepared from the kidney, oviduct, brain, spleen, and liver of a single hen. These samples were separately digested with either *Bgl*II or *Eco*RI and blotted onto nitrocellulose filters following agarose gel electrophoresis. These filters were then hybridized with single-copy DNA probes from regions of the chicken genome very close to, but not including, the three CR1 sequences located in the ovalbumin gene domain. The restriction fragments comprising these probes are described in the legend to Figure 4. It can be noted from Figure 2 that there are no cutting sites for either the enzyme *Bgl*II or the enzyme *Eco*RI between the probes and their corresponding nearby CR1 sequences. There are also no cutting sites for these enzymes within the CR1 sequences themselves (Figure 2). Therefore, any tissue-specific rearrangements in the regions detected by these probes should appear as a variation in band size in the autoradiograms. As demonstrated by the results shown in Figure 4, however, no tissue-specific variations are observed when any of the probes are used. In particular, there are no specific differences observed for oviduct DNA as compared to other tissues. (The very slight migrational differences among the 2.7-kbp bands are most likely due to variations inherent to the Southern blotting technique and are not deemed to represent actual size differences in the DNA.) Although these experiments do not rule out the possibility of DNA rearrangements occurring in other regions of the domain, the results do indicate that major DNA rearrangements involving the CR1 sequences are probably not a mechanism involved in the tissue-specific expression of the ovalbumin gene.

Tissue-Specific DNA Methylation at the 3′ Boundary of the Ovalbumin Gene Domain. DNA sequences within and immediately flanking active genes have often been characterized as being undermethylated at cytidine residues in the dinucleotide sequence CpG. This undermethylation within and around active genes is tissue specific. That is, specific sites are undermethylated in DNA isolated from tissues where a gene is expressed relative to those same sites in DNA isolated from tissues where that gene is not expressed. We have extended such studies by examining methylation patterns not only in the immediate vicinity of expressed genes but also throughout the entire ovalbumin gene domain, including regions quite far removed from the known transcriptional units. The results of some of these experiments are shown in Figure 5. DNA preparations from various tissues were digested with either of the restriction enzymes *Hpa*II or *Msp*I, and Southern transfers were carried out. The same DNA sequence (CCGG) is recognized by both of these enzymes; however, *Hpa*II does not cut at this sequence if the internal C is methylated, whereas

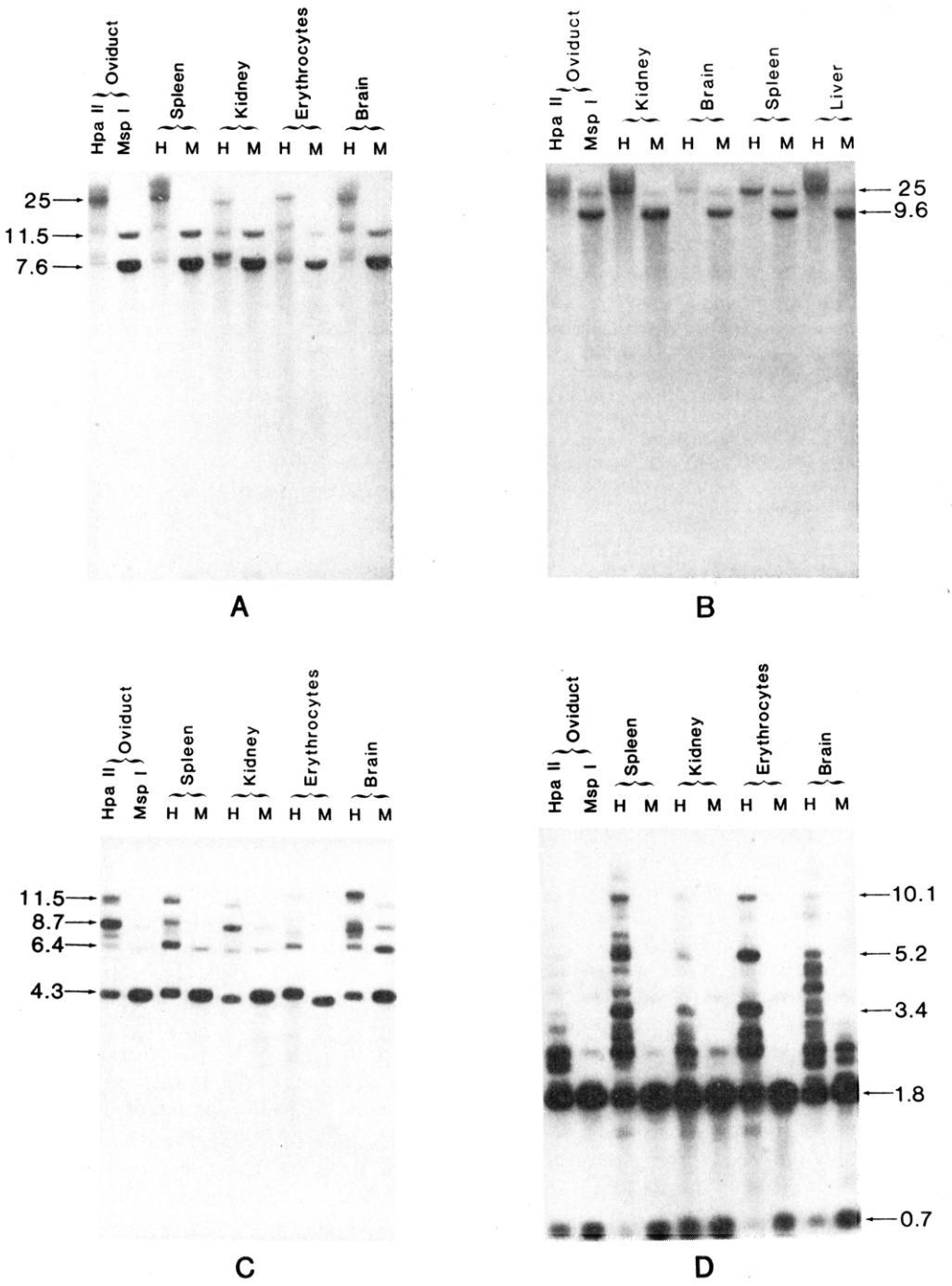

FIGURE 5: DNA methylation at *Hpa*II/*Msp*I sites near the 5′ and 3′ boundaries of the ovalbumin gene domain. DNA samples isolated from various hen tissues were digested to completion with either *Hpa*II or *Msp*I, electrophoresed in 0.7% agarose gels, transferred to nitrocellulose filters (Southern, 1975), and hybridized to ^{32}P-labeled probes from the 5′ (panels A and B) and 3′ (panels C and D) ends of the ovalbumin domain. The various tissue DNA samples used in the experiments in panels A, C, and D were all from the same hen, whereas the DNA samples used in panel B were from a second hen. In panel A, the hybridization probe was the 8.5-kbp *Eco*RI fragment from CL167 (see Figure 1 for restriction mapping data). In panel B, the probe was the 8.1-kbp *Eco*RI fragment from CL252 [in this experiment, a large excess of the cloned 4.3 (167) fragment was included in the hybridization reaction in order to competitively eliminate hybridization by the repetitive sequence, CR1OVc, that exists within the probe]. In panel C, the hybridization probe was the 2.3-kbp *Xho*I fragment from CL734. In panel D, the probe was a 1.8-kbp *Sst*I/*Eco*RI subfragment from the 3′ end of the 5.1-kbp *Xho*I/*Eco*RI fragment of CL734.

*Msp*I does. Figure 5 compares the extent of methylation of these sites in regions at the 5′ and 3′ boundaries of the domain. The probes used in panels A and D are from regions at the extreme 5′ and 3′ ends, respectively. In these regions, the domain is closed and shows no preferential sensitivity to DNase I in any of the tissues examined, including the oviduct. The probes used in panels B and C are from regions which exhibit an intermediate level of DNase I sensitivity within the 5′ and 3′ transition regions of the oviduct chromatin (Lawson et al., 1982). The results indicate that the internal C within CCGG sequences at the 5′ end of the domain (panels A and B) is highly methylated in all of the tissues examined, including the oviduct. These results should be contrasted with the patterns shown in panels C and D which were obtained by using the two probes from the 3′ end of the domain. These probes reveal a complex pattern of tissue-specific methylation. DNA from each of the tissues examined is only partially methylated in this region, and each of the tissues also has a unique methylation pattern. Notably, the probe used in panel D reveals that oviduct DNA is the most undermethylated in this region

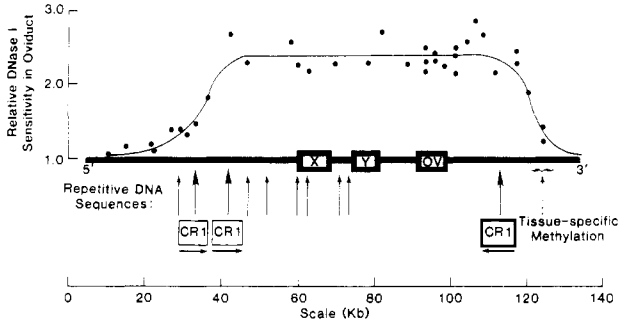

FIGURE 6: Summary of the ovalbumin gene domain. The solid circles indicate the relative DNase I sensitivity in oviduct nuclei of DNA sequences at various positions in the ovalbumin gene domain as determined by Lawson et al. (1982). Those data were obtained by hybridizing cloned probes from throughout the domain to DNA isolated from oviduct nuclei that had been digested to 15–20% acid solubility. For the calculation of the relative DNase I sensitivity, those hybridization rates were compared to the corresponding rates obtained in parallel hybridizations utilizing total nuclear DNA as the driver. The solid vertical arrows denote the locations of repetitive DNA sequences within the domain (Lawson et al., 1982; Heilig et al., 1980). The bolder vertical arrows point out the specific locations of CR1 family repetitive DNA sequences. The orientations of these CR1 sequences are denoted by the horizontal arrows. The dashed arrow indicates a region where distinctive methylation patterns are noted for different chicken tissues, and where oviduct DNA is specifically undermethylated.

when compared to the other tissues. This is shown by the fact that *Hpa*II is able to digest oviduct DNA from this region of the genome to smaller fragments than what is observed for the same DNA sequences in other tissues. In addition, the number of bands observed in panel D indicates that there is a greater density of CCGG sequences in this 3′-boundary region than what is observed in the rest of the domain or than what is normally observed in chicken DNA in general. This observation is further supported by direct restriction site mapping and nucleotide sequence analysis of cloned fragments from this region (data not shown). This increased occurrence of the rare dinucleotide CpG, taken together with the observed undermethylation at these sites in the oviduct and the variation in the methylation patterns observed in different tissues, could indicate that this 3′ border of the domain may be a control region involved in regulating the tissue-specific expression of the ovalbumin gene.

Discussion

Many of our findings in regard to the ovalbumin gene domain are summarized in Figure 6. The X, Y and ovalbumin genes are specifically expressed in the chicken oviduct upon exposure to steroid hormones. The domain which contains these genes is composed of approximately 100 kbp of DNA which is preferentially sensitive to DNase I in nuclei isolated from chicken oviduct cells (Lawson et al., 1982). These same DNA sequences exhibit no preferential sensitivity to DNase I in nuclei isolated from other chicken tissues. Since many examples of preferential DNase I sensitivity correlated with active gene expression have been reported [see Lawson et al., (1982) for references], and since no evidence has been reported to the contrary, it may be presumed that all actively expressed genes in higher organisms exist in domains of preferential DNase I sensitivity which reflect an altered chromatin structure. We feel that it is important to understand what molecular mechanisms are involved in the establishment of DNase I sensitive domains in those defined regions of the chromatin which contain active genes. Several lines of evidence indicate that it is not ongoing transcription of the genes themselves which is responsible for the observed DNase I sensitivity. This is demonstrated by earlier results which show that the DNase I sensitivity of the ovalbumin gene domain is maintained in the oviduct even under conditions of hormonal withdrawal (Shepherd et al., 1980; Lawson et al., 1980, 1982), a state in which transcription of the ovalbumin gene ceases (Roop et al., 1978; Swaneck et al., 1979). An analogous result is obtained for the globin gene in mature chicken erythrocytes (Weintraub & Groudine, 1976) which are no longer actively synthesizing RNA. In addition, the DNase I sensitivity is not confined to the genes themselves but also extends into the flanking DNA (Lawson et al., 1980, 1982; Stalder et al., 1980) which is not transcribed (Tsai et al., 1980). Therefore, there must be mechanisms and signals other than transcription of the structural genes which are responsible for rendering specific regions of the chromatin preferentially DNase I sensitive.

There is evidence that the HMG proteins 14 and 17 are responsible for conferring a highly DNase I sensitive state upon nucleosomes within active genes (Weisbrod et al., 1980; Gazit et al., 1980). However, this highly elevated DNase I sensitivity and associated binding of HMG proteins 14 and 17 appear to be confined to the immediate vicinity of the transcribed sequences (Weisbrod & Weintraub, 1981). In addition, smaller regions of even greater sensitivity (DNase I hypersensitive sites) can be observed at specific locations in chromatin, most often at the 5′ ends of active genes [for a review, see Elgin (1981)]. Our data presented here do not address the question of the cause of the extreme DNase I sensitivity found at the hypersensitive sites, nor do our data deal with the elevated sensitivity that has been associated with HMG protein binding and is observed within and immediately adjacent to transcribed sequences. Rather, our DNase I sensitivity data are derived from experiments employing a far more exhaustive digestion of the chromatin with DNase I. Thus, our data are concerned with the more generalized DNase I sensitivity which extends for relatively great distances into DNA regions flanking active genes. This generalized DNase I sensitivity presumably reflects an alteration in the packaging of the chromatin fibrils into higher order structures. We have initiated studies to determine what signals are involved in defining such extensive domains of altered chromatin structure.

In the case of the chicken ovalbumin gene domain, it is possible that the CR1 family repetitive DNA sequences could be involved in such a process. These sequences are found in the ovalbumin gene domain close to where the transition in chromatin structure occurs. This correlation may be coincidental, or it could have an important functional significance. As a next step, it will be necessary to examine CR1 sequences in other chromatin domains and to determine whether their locations there correlate with transitions in DNase I sensitivity.

Because of the potentially important functional role of the chicken CR1 sequences, we have characterized at the nucleotide sequence level three members of this family that are present in the ovalbumin gene domain (Figure 3). We have previously assigned a 5′ to 3′ polarity to the chicken CR1 sequences based upon their partial nucleotide sequence homology to human Alu sequences and other mammalian Alu-type repeats (Stumph et al., 1981). [The Alu sequences have been assigned a polarity based upon the direction in which they are transcribed in vitro by RNA polymerase III (Duncan et al., 1981).] It is intriguing that the two CR1 sequences at the 5′ end of the domain (CR1OVb and CR1OVc) are in the same orientation as the direction of transcription of the X, Y, and ovalbumin genes, whereas the CR1OVa sequence at the 3′ end of the domain is in the inverse orientation. These orientations

are depicted by the horizontal arrows in Figures 2 and 6. Thus, the CR1 sequences in the ovalbumin gene domain have inward orientations pointing toward the central portion of the domain which contains the structural genes. We have also characterized one other CR1 repeat (CR1U1a) at the nucleotide sequence level (Figure 3). The CR1U1a sequence comes from a clone which contains a gene coding for chicken U1 RNA. Interestingly, its orientation is in agreement with the above observation (Stumph et al., 1981). That is, it also "points" toward the downstream structural U1 RNA gene. This inward orientation may be an important structural feature if CR1 sequences actually do play a role in establishing gene domains. However, additional examples of CR1 sequences in other genetic domains need to be studied before the generality of this feature can be established. It has been noted that in the rat a pair of inverted repeats flank a 12-kbp length of DNA which contains a gene for rat preproinsulin (Lomedico et al., 1979). Also, approximately 35 kbp of rabbit DNA, which contains the rabbit β-globin gene cluster, is flanked by inverted repeats (Shen & Maniatis, 1980). It would be interesting to know if the locations of these inverted repeats are correlated with a change in DNase I sensitivity in these other systems.

We have previously reported a preliminary characterization of the chicken CR1 repetitive DNA sequence family including nucleotide sequences for the CR1U1a and CR1OVa family members (Stumph et al., 1981). On the basis of the new data obtained from the sequences of the CR1OVb and CR1OVc repeats, plus additional sequencing data on CR1U1a, we are now able to present a more complete structure of the chicken CR1 sequence family as depicted in Figure 3. Our previous results included sequence data for CR1U1a up to only position 129 of the consensus sequence as shown in Figure 3. At that point, the CR1U1a and CR1OVa sequences had become divergent and showed no significant homology. Thus, we did not detect the second tract of high homology which is shared by the four sequences and which extends between positions 179 and 217 of the consensus sequence. Only when additional sequencing data for CR1OVb and CR1OVc became available did the second homology block become evident. This split structure of the CR1 family (two highly conserved tracts interrupted by a region of relatively low homology) is an interesting feature. Presumably, the more conserved regions may play important roles in CR1 sequence function, whatever that function might be. On the other hand, the conservation of nucleotide sequence in the central portion might not be essential for CR1 sequence function. Rather, it may be that simply maintaining a certain distance between the two putative functional regions may be the only requirement involved. In fact, it has been shown that promoters for tRNA genes are split into two independent regions (Hofstetter et al., 1981; Sharp et al., 1981) and that transcription is optimally promoted when these two regions are about 40–50 nucleotides apart (Ciliberto et al., 1982). An analogous situation may be the case in the CR1 sequence family. In the case of all four of the CR1-containing clones, we have sequenced a minimum of 90 bp on each side of the points where the homology between family members begins to decrease, and we have not discovered any additional large regions of homology when these flanking sequences are compared. Therefore, from our present data, it appears likely that the most highly conserved regions of homology shared by the four CR1 family members that have been characterized can be accounted for by the sequence data presented in Figure 3.

As mentioned above, the locations and orientations of the CR1 sequences in the ovalbumin gene domain suggest that the CR1 sequences potentially may be involved in establishing or defining DNase I sensitive domains in chicken cell chromatin. If such a role is proposed for CR1 sequences, what kinds of mechanisms might be involved? The results of the experiments shown in Figure 4 make it very unlikely that CR1 sequences in the ovalbumin gene domain are active in promoting tissue-specific rearrangements of the DNA in oviduct cells. However, a number of other possible mechanisms can be proposed. It has been demonstrated that some mammalian repetitive DNA sequences act as promoters for RNA polymerase III in in vitro systems (Duncan et al., 1981; Haynes & Jelinek, 1981). In one case, a class of small RNAs has been identified in Chinese hamster cells which could represent in vivo transcripts of such repetitive DNA sequences (Haynes & Jelinek, 1981). By analogy, it is possible that the chicken CR1 sequences could act as promoters for RNA polymerase in vivo. If so, perhaps just the very act of transcribing the CR1 sequences might be sufficient to result in the opening up of the domain by affecting nucleosome configuration or by relieving constraints in the DNA. This effect could then be propagated throughout the domain. Alternatively, a certain class of in vivo RNA transcripts could possibly interact with the CR1 sequences in a sequence-specific manner to bring about the opening up of the domain. For example, a class of small nuclear RNAs in meiotic cells has been reported to interact with specific repetitive sequences in lily DNA (Hotta & Stern, 1981). By a mechanism as yet unknown, these interactions apparently render specific DNA sequences preferentially accessible to both endogenous and exogenous nucleases (Hotta & Stern, 1981).

Alternatively, the CR1 sequences could act as specific binding sites for a particular class of proteins. One possibility is that these proteins would then induce conformational changes in the chromatin. More specifically, the CR1's may serve as reference sites for attachment of the chromatin to the nuclear matrix, which is an architectural structure of the nucleus (Berezney & Coffey, 1974) which continues to exist after removal of most of the DNA, RNA, and chromosomal proteins. Along this line of reasoning, several groups have reported that a large part of the chromatin in eucaryotic cells exists as supercoiled loops (Cook & Brazell, 1975; Benyajati & Worcel, 1976; Paulson & Laemmli, 1977; Vogelstein et al., 1980). Most of these loops contain about 50–150 kbp of DNA, and the two ends of each individual loop appear to be anchored to adjacent points in the nuclear matrix or chromosomal scaffold. Perhaps the CR1 sequences are found at these anchorage points, and the chromatin is bound to the matrix via interactions between the CR1 sequences and intrinsic protein components of the matrix. In fact, specific proteins that are bound to matrix-associated DNA have been described (Razin et al., 1981). In addition, evidence has accumulated that transcriptionally active genes and newly synthesized RNA are closely associated with the nuclear matrix (Miller et al., 1978; Herman et al., 1978; Nelkin et al., 1980; Robinson et al., 1982; Ciejek et al., 1982). It is therefore possible that CR1's could be utilized as binding sites only when the chromatin is to be closely associated with the matrix in an active conformation.

In many instances, tissue-specific undermethylation of cytidine residues at the sequence CpG has been correlated with active gene expression (McGhee & Ginder, 1979; Van der Ploeg & Flavell, 1980; Weintraub et al., 1981; Groudine et al., 1981; Bird et al., 1981; Clough et al., 1982). Using DNA from different tissues, we have extensively studied methylation patterns at CCGG sequences (*Hpa*II/*Msp*I sites) in the ovalbumin gene domain. *Hpa*II/*Msp*I sites occur rarely

throughout most of the ovalbumin gene domain, thereby reflecting the underrepresentation of the CpG dinucleotide in eucaryotic cells. As described earlier by other investigators (Mandel & Chambon, 1979), we have also observed a limited degree of undermethylation at certain sites near to the ovalbumin gene itself. However, using a variety of probes, we have not been able to detect any tissue-specific undermethylation at the 5′ end of the domain. However, a region at the 3′ boundary of the domain contains a highly increased density of *Hpa*II/*Msp*I sites compared to the rest of the domain. More importantly, methylation at these sites is highly variable from tissue to tissue, and oviduct DNA appears to be specifically undermethylated at the extreme 3′ end of the domain compared to the other tissues (Figure 5, panel D). It is very intriguing that a region approximately 30 kbp downstream from the ovalbumin gene exhibits such a tissue-specific pattern of methylation. It is highly unlikely that there is any transcriptional activity from the DNA in this region because in all tissues studied, including the oviduct, this part of the genome is found in the DNase I resistant fraction (Lawson et al., 1982). It is, therefore, tempting to postulate that this area may be a control region involved in governing the structural state of the ovalbumin gene domain. One possibility is that DNA sequences exist in this region which could undergo a Z $\rightleftarrows$ B conformational transition as a result of tissue-specific demethylation (Behe & Felsenfeld, 1981; Nordheim et al., 1981). Since transitions of this type would markedly alter the supercoiling of the DNA, such a mechanism seems plausible for the activation of large domains of chromatin structure. In this regard, it has been hypothesized, based upon practical considerations, that a 3′-distal element is likely to be involved in the control of immunoglobulin gene expression and that this control could be carried out by an influence on chromatin structure (Mather & Perry, 1981). Perhaps a 3′ controlling region could be responsible also for the observed differences in the rates of transcription of the X, Y and ovalbumin genes. In stimulated oviduct, the most 3′ gene, ovalbumin, is the most rapidly transcribed of the three related genes in the domain, whereas the X gene, located toward the 5′ end of the domain, is transcribed at the lowest rate (Colbert et al., 1980; LeMeur et al., 1981). It is potentially interesting that kidney DNA also appears to be less methylated in the 3′-flanking region when compared to the other nonoviduct tissues (Figure 5, panel D). We are currently investigating the question of whether the ovalbumin gene domain might exhibit preferential DNase I sensitivity in kidney tissue as well as in the oviduct.

The results reported here represent our initial efforts to identify signals which could be responsible for establishing the ovalbumin gene DNase I sensitive domain. It is clear that such giant structural domains exist in eucaryotic interphase chromosomes and are likely related to the molecular mechanisms of the differentiation process. We suggest that tissue-specific genes which are to be expressed in the lifetime of a given cell type will be found in the 10–20% of the nuclear DNA which exists in such structures. Genes not expressed in a given cell type will be found in the 80–90% of the DNA which is not contained in DNase I sensitive domains. Although the inclusion in such domains appears to be a necessary requirement for gene expression, it is apparently not a sufficient condition for expression. In other words, the DNase I sensitive structure provides the potential for a gene to be activated through the further action of additional cellular regulatory and inducer molecules (e.g., hormone receptors). In the present study, a number of interesting correlations have been observed possibly relating chromatin structure to signals in the DNA, but the generality and significance of these findings still need to be proven. We are currently investigating the structure of the DNase I sensitive domains which contain the ovomucoid and glyceraldehyde-3-phosphate dehydrogenase genes. This area of research should prove fertile over the next several years for scientists interested in the mechanisms of cellular differentiation and gene activation.

Acknowledgments

We thank Serlina Robinson and Valerie McMullian for excellent technical assistance. Homologies between the various nucleotide sequences were determined by using the NIH-Aim computer facilities of the Stanford MOLGEN Project.

Supplementary Material Available

Nucleotide sequences of CR1 family members plus flanking DNA: CR1OVa, 1003 bp; CR1OVb, 608 bp; CR1OVc, 770 bp (3 pages). Ordering information is given on any current masthead page.

Registry No. CR1U1a, 84009-42-7; CR1OVa, 84009-39-2; CR1OVb, 84009-41-6; CR1OVc, 84009-40-5.

References

Behe, M., & Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1619–1623.

Benyajati, C., & Worcel, A. (1976) *Cell (Cambridge, Mass.) 9*, 393–407.

Berezney, R., & Coffey, D. S. (1974) *Biochem. Biophys. Res. Commun. 60*, 1410–1417.

Bird, A., Taggart, M., & McLeod, D. (1981) *Cell (Cambridge, Mass.) 26*, 381–390.

Calabretta, B., Robberson, D. L., Barrera-Saldana, H. A., Lambrou, T. P., & Saunders, G. F. (1982) *Nature (London) 296*, 219–225.

Ciejek, E. M., Nordstrom, J. L., Tsai, M. J., & O'Malley, B. W. (1982) *Biochemistry 21*, 4945–4953.

Ciliberto, G., Traboni, C., & Cortese, R. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1921–1925.

Clough, D. W., Kunkel, L. M., & Davidson, R. L. (1982) *Science (Washington, D.C.) 216*, 70–73.

Colbert, D. A., Knoll, B. J., Woo, S. L. C., Mace, M. L., Tsai, M. J., & O'Malley, B. W. (1980) *Biochemistry 19*, 5586–5592.

Cook, P. R., & Brazell, I. A. (1975) *J. Cell Sci. 19*, 261–279.

Dodgson, J. B., Strommer, J., & Engel, J. D. (1979) *Cell (Cambridge, Mass.) 17*, 879–887.

Duncan, C. H., Jagadeeswaran, P., Wang, R. R. C., & Weissman, S. M. (1981) *Gene 13*, 185–196.

Elgin, S. C. R. (1981) *Cell (Cambridge, Mass.) 27*, 413–415.

Garel, A., & Axel, R. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3966–3970.

Gazit, B., Panet, A., & Cedar, H. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1787–1790.

Grimaldi, G., & Singer, M. F. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1497–1500.

Groudine, M., Eisenman, R., & Weintraub, H. (1981) *Nature (London) 292*, 311–317.

Haynes, S. R., & Jelinek, W. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6130–6134.

Haynes, S. R., Toomey, T. P., Leinwand, L., & Jelinek, W. R. (1981) *Mol. Cell. Biol. 1*, 573–583.

Heilig, R., Perrin, F., Gannon, F., Mandel, J. L., & Chambon, P. (1980) *Cell (Cambridge, Mass.) 20*, 625–637.

Herman, R., Weymouth, L., & Penman, S. (1978) *J. Cell Biol. 78*, 663–674.

Hofstetter, H., Kressmann, A., & Birnstiel, M. L. (1981) *Cell (Cambridge, Mass.) 24*, 573–585.

Hotta, Y., & Stern, H. (1981) *Cell (Cambridge, Mass.) 27*, 309–319.

Lawson, G. M., Tsai, M.-J., & O'Malley, B. W. (1980) *Biochemistry 19*, 4403–4411.

Lawson, G. M., Knoll, B. J., March, C. J., Woo, S. L. C., Tsai, M.-J., & O'Malley, B. W. (1982) *J. Biol. Chem. 257*, 1501–1507.

LeMeur, M., Glanville, N., Mandel, J. L., Gerlinger, P., Palmiter, R., & Chambon, P. (1981) *Cell (Cambridge, Mass.) 23*, 561–571.

Lomedico, P., Rosenthal, N., Efstradiadis, A., Gilbert, W., Kolodner, R., & Tizard, R. (1979) *Cell (Cambridge, Mass.) 18*, 545–558.

Mandel, J., & Chambon, P. (1979) *Nucleic Acids Res. 7*, 2081–2103.

Mather, E. L., & Perry, R. P. (1981) *Nucleic Acids Res. 9*, 6855–6867.

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.

McGhee, J. D., & Ginder, G. D. (1979) *Nature (London) 280*, 419–420.

Miller, T. E., Huang, C. Y., & Pogo, A. O. (1978) *J. Cell Biol. 76*, 675–691.

Nelkin, B. D., Pardoll, D. M., & Vogelstein, B. (1980) *Nucleic Acids Res. 8*, 5623–5633.

Nordheim, A., Pardue, M. L., Lafer, E. M., Moller, A., Stollar, B. D., & Rich, A. (1981) *Nature (London) 294*, 417–422.

Paulson, J. R., & Laemmli, U. K. (1977) *Cell (Cambridge, Mass.) 12*, 817–828.

Razin, S. V., Chernokhvostov, V. V., Roodyn, A. V., Zbarsky, I. B., & Georgiev, G. P. (1981) *Cell (Cambridge, Mass.) 27*, 65–73.

Robinson, S. I., Nelkin, B. D., & Vogelstein, B. (1982) *Cell (Cambridge, Mass.) 28*, 99–106.

Roop, D. R., Nordstrom, J. L., Tsai, S. Y., Tsai, M.-J., & O'Malley, B. W. (1978) *Cell (Cambridge, Mass.) 15*, 671–685.

Roop, D. R., Kristo, P., Stumph, W. E., Tsai, M.-J., & O'Malley, B. W. (1981) *Cell (Cambridge, Mass.) 23*, 671–680.

Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T., & Schmid, C. W. (1980) *Nature (London) 284*, 372–374.

Sharp, S., DeFranco, D., Dingermann, T., Farrell, P., & Soll, D. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6657–6661.

Shen, C.-K. J., & Maniatis, T. (1980) *Cell (Cambridge, Mass.) 19*, 379–391.

Shepherd, J. H., Mulvihill, E. R., Thomas, P. S., & Palmiter, R. D. (1980) *J. Cell Biol. 87*, 142–151.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

Stalder, J., Larsen, A., Engel, J. D., Dolan, M., Groudine, M., & Weintraub, H. (1980) *Cell (Cambridge, Mass.) 20*, 451–460.

Stumph, W. E., Kristo, P., Tsai, M.-J., & O'Malley, B. W. (1981) *Nucleic Acids Res. 9*, 5383–5397.

Swaneck, G. E., Nordstrom, J. L., Kreuzaler, F., Tsai, M.-J., & O'Malley, B. W. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1049–1053.

Tsai, S. Y., Roop, D. R., Stumph, W. E., Tsai, M.-J., & O'Malley, B. W. (1980) *Biochemistry 19*, 1755–1761.

Van Arsdell, S. W., Denison, R. A., Bernstein, L. B., Weiner, A. M., Manser, T., & Gesteland, R. F. (1981) *Cell (Cambridge, Mass.) 26*, 11–17.

Van der Ploeg, L. H. T., & Flavell, R. (1980) *Cell (Cambridge, Mass.) 19*, 947–958.

Vogelstein, B., Pardoll, D. M., & Coffey, D. S. (1980) *Cell (Cambridge, Mass.) 22*, 79–85.

Weintraub, H., & Groudine, M. (1976) *Science (Washington, D.C.) 193*, 848–856.

Weintraub, H., Larsen, A., & Groudine, M. (1981) *Cell (Cambridge, Mass.) 24*, 333–344.

Weisbrod, S., & Weintraub, H. (1981) *Cell (Cambridge, Mass.) 23*, 391–400.

Weisbrod, S., Groudine, M., & Weintraub, H. (1980) *Cell (Cambridge, Mass.) 19*, 289–301.