The Structure of the Globin Genes in Chromatin†

Richard Axel, Howard Cedar, and Gary Felsenfeld*

ABSTRACT: The distribution of proteins in the neighborhood of the globin genes of duck reticulocyte chromatin has been studied. This chromatin is first shown to be an active template for transcription in vitro of globin messenger-like RNA. The chromatin is then treated with staphylococcal nuclease and the DNA fragments protected from nuclease attack ("covered DNA") are isolated. Alternatively, the chromatin is titrated with poly-D-lysine, and by successive treatment with Pronase and nuclease, the DNA regions accessible to polylysine are isolated ("open DNA"). In order to determine the distribution of globin gene sequences in open and covered DNA, these two fractions are annealed to globin cDNA (globin probe). It is found that while all globin gene sequences are represented in covered DNA, a specific portion of the globin gene is missing from open DNA, corresponding to about 20% of the gene length. It is concluded that specific regions of the globin genes of reticulocyte chromatin are partly covered by proteins in such a way as to render them inaccessible to polylysine. In contrast, no difference is observed in the annealing properties of open and covered regions to globin probe using DNA isolated from erythrocyte chromatin, which is a poor template in vitro for production of globin message. The annealing of open and covered DNA to each other has also been studied. It is found that open and covered DNA have identical sequence populations. Thus, in contrast to the special arrangement of proteins in the neighborhood of the globin gene, there does not appear to be any sequence-specific arrangement of the bulk of the chromatin proteins on chromatin DNA.

Chromatin isolated from eucaryotic nuclei retains the information required to direct the transcription of RNA specific to the tissue from which the chromatin was extracted (Axel et al., 1973a,b; Gilmour and Paul, 1973). We can only speculate at this time as to the nature of the controlling elements, but studies on the reconstitution of chromatin proteins with DNA suggest that the non-histone proteins include specific transcriptional control factors (Paul et al., 1973).

In previous studies (Clark and Felsenfeld, 1971), we have measured the accessibility of the DNA in chromatin to various chemical probes. About half the DNA is accessible to titration with divalent cations, histones, or polylysine and to digestion by staphylococcal nuclease. The DNA resistant to digestion consists principally of a series of double-stranded fractions of discrete lengths, ranging between about 150 and 45 base pairs (Axel et al., 1974; Sollner-Webb and Felsenfeld, 1975; Axel, 1975). Although the structural features giving rise to this behavior are not yet understood, it seems clear that a substantial fraction of the DNA is either free of protein or covered by protein segments that are readily displaced by the probes. The relationship between the chemically accessible DNA and the actively transcribed genes of the chromatin is obviously not straightforward; the amount of DNA reactive to nuclease and other probes is far too great to account for the striking restriction of transcription observed with chromatin templates.

In this paper, we investigate further the relationship between the distribution of proteins on DNA and the biological role of chromatin in the transcription process. We find that the mere presence of protein on a particular DNA sequence is not sufficient to restrict its transcription. We focus our attention on the globin gene of avian blood cells and show that actively transcribed genes are partially covered by protein in a site-specific manner. However, the majority of chromatin proteins appear to be randomly arranged along the DNA and do not have unique binding sites.

† From the Laboratory of Molecular Biology, National Institute of Arthritis, Metabolism and Digestive Diseases, National Institutes of Health, Bethesda, Maryland 20014. *Received December 18, 1974.*

Materials and Methods

(a) Preparation of Chromatin and DNA. Duck reticulocytes were obtained by injection of phenylhydrazine hydrochloride (5 mg/kg) for 5 days and harvested by cardiac puncture on day 8 (Attardi et al., 1966). Reticulocyte chromatin was prepared from Triton-washed nuclei by a stepwise reduction in ionic strength as previously described (Axel et al., 1973). The final preparation was sheared to an average DNA molecular weight of 6–8 $\times 10^6$ in a Virtis homogenizer (1 min at 80 V); but was not sonicated. Chromatin prepared in this way had a protein/DNA ratio of 1.3 g/g and was stable for up to 6 weeks at 4°. DNA was prepared from isolated nuclei as described (Axel et al., 1973).

(b) Transcription of Chromatin in Vitro. RNA was synthesized in 10-ml reaction mixtures containing: 10 m*M* Tris-HCl (pH 7.9), 1 m*M* $MnCl_2$, 0.08 m*M* each of ATP, GTP, UTP, and [α-^{32}P]CTP (specific activity 3 cpm/pmol), 300 units of *Escherichia coli* RNA polymerase (fraction V of Berg et al., 1971), and 0.75 mg (DNA content) of chromatin. Where necessary highly radioactive RNA was synthesized by the addition of 0.01 m*M* UTP at a specific activity of 20,000 cpm/pmol. The reaction was incubated at 37° for 45 min and the newly synthesized RNA was then purified free of protein, DNA, and triphosphates (Axel et al., 1973). It is important to note that under these reaction conditions the chromatin is soluble and less than 3% of the histones undergo exchange (Cedar and Felsenfeld, 1973). A typical 10-ml reaction mixture with chromatin as template yields from 50 to 100 μg of newly synthesized RNA.

(c) Synthesis of Globin cDNA. Globin 9S mRNA was purified from duck reticulocyte polyribosomes by cellulose chromatography and sucrose gradient centrifugation as de-

scribed (Axel et al., 1973). RNA prepared in this way results in the predominant synthesis of duck globin components when added to a Krebs-ascites cell-free translating system. [^{3}H]DNA complementary to this globin mRNA (globin cDNA) was synthesized in a 1-ml reaction containing: 50 m*M* Tris-HCl (pH 8.0), 6 m*M* $MgCl_2$, 0.2 m*M* dTTP, dATP, dGTP, and 0.05 m*M* [^{3}H]dCTP (26.8 Ci/mmol), 0.05 *M* KCl, 25 μg of actinomycin D, 1 μg of $(dT)_{10}$, 10 μg of mRNA, and 50 units of RNA-dependent DNA polymerase (a gift of Dr. T. Papas) purified from avian myeloblastosis virus (Kacian et al., 1971). The reaction was incubated at 37° for 2 hr and the newly synthesized [^{3}H]DNA was freed of protein and template RNA as described (Kacian et al., 1972). The size of this DNA was estimated from alkaline sucrose gradient sedimentation to be about 70% of the length of globin mRNA.

(d) Preparation of "Open" and "Covered" DNA from Chromatin. About half of the DNA in chromatin resists digestion by the enzyme staphylococcal nuclease. This fraction of DNA (covered DNA) is prepared in large quantities as follows: 100 ml of a solution of 0.1 m*M* $CaCl_2$–2 m*M* Tris-HCl (pH 8.0) are added dropwise with stirring to 100 ml of chromatin at a concentration of 0.5 mg/ml (DNA content). Staphylococcal nuclease (Worthington Biochemicals) is then added at a concentration of 8 μg/ml and the reaction is incubated at 37° for 1 hr with gentle stirring. Toward the end of the reaction the protected DNA precipitates and is collected by centrifugation at 10,000 rpm in a Sorvall SS-34 rotor. This precipitate contains half of the original complement of DNA and virtually all of the chromatin proteins. The precipitate is suspended in 0.01 *M* Tris-HCl (pH 8), 0.4 *M* NaCl, 0.005 *M* EDTA, and 0.5% sodium dodecyl sulfate and freed of protein by extraction with phenol–chloroform. The covered DNA is then precipitated from the aqueous phase by the addition of two volumes of ethanol.

The regions of DNA in chromatin susceptible to nuclease attack and titratable by poly-D-lysine (open DNA) are prepared as follows: 100 ml of a solution containing 70 μmol of poly-D-lysine in 2 m*M* Tris-HCl (pH 8.0) are added dropwise with stirring to 100 ml of chromatin, DNA content 0.5 mg/ml (150 μmol of DNA) to yield a lysine/DNA phosphate ratio of 0.46. Pronase is then added to a concentration of 8 U/ml and the reaction is incubated at 37° for 3 hr and then at 4° for 16 hr to permit autodigestion of the Pronase. This reaction results in complete digestion of the chromatin proteins leaving in solution a complex of poly-D-lysine bound to DNA regions previously accessible in intact chromatin. The DNA complexed to poly-D-lysine is now isolated by nuclease digestion in a manner identical with that described for the preparation of native covered DNA, except that the limit digest was first adsorbed on hydroxylapatite in 0.5 *M* NaCl and the DNA eluted with 0.5 *M* phosphate buffer (pH 6.5).

Formaldehyde cross-linked chromatin, prepared according to the method of Brutlag et al. (1969), was digested with staphylococcal nuclease in the manner described above for untreated chromatin. The covered DNA was freed of protein by dissolving the digest in 0.4 *M* NaCl, 0.01 *M* Tris (pH 8), 5 m*M* NaEDTA, and 0.2% sodium dodecyl sulfate. The solution was treated first with 0.5 mg/ml of Pronase (Calbiochem), 3 hr, 37°, then with 0.1 mg/ml of Proteinase K (E. Merck) for 3 hr at 37°. The solution was extracted with phenol and the DNA was precipitated with ethanol.

(e) In Vitro Labeling of Open and Covered DNA with DNA Polymerase III. Either open or covered region DNA was 30% digested with exonuclease III of *Escherichia coli* (a gift of Dr. Motohiro Fuke) under conditions described by Richardson (1965). This exonuclease treated DNA was now used as template for DNA synthesis by *E. coli* DNA polymerase III (Kornberg and Gefter, 1972). [^{3}H]DNA was synthesized in a 0.1-ml reaction containing: 30 m*M* Tris-HCl (pH 7.4), 13 m*M* $MgCl_2$, 50 m*M* β-mercaptoethanol, 0.1 m*M* dATP, dGTP, dTTP, and 0.1 m*M* [^{3}H]dCTP (13 Ci/mmol), 4 μg of either exonuclease III treated open or covered DNA, and 5 units of *E. coli* DNA polymerase III (kindly supplied by Dr. Sue Wickner). The reaction was incubated for 45 min at 25° and the DNA was isolated by phenol extraction and Sephadex G-50 chromatography. DNA with a specific activity of 10^6 cpm /μg and with annealing properties identical with unreacted template DNA can readily be obtained by this procedure.

(f) In Vitro Labeling of DNA with Polynucleotide Kinase. Open and covered regions of DNA from chromatin prepared with staphylococcal nuclease can serve as substrate for the enzyme polynucleotide kinase (Richardson, 1965). The 5′-hydroxyl end of open or covered DNA was labeled in a 0.1-ml reaction containing: 70 m*M* Tris-HCl (pH 7.4), 10 m*M* $MgCl_2$, 5 m*M* dithiothreitol, 0.066 m*M* [α-^{32}P]ATP (15 Ci/mmol), and 5 units of polynucleotide kinase (a gift of Dr. Motohiro Fuke). The reaction was incubated for 2 hr at 37° and the DNA was then purified by phenol extraction and Sephadex G-50 chromatography. This procedure resulted in DNA with a specific activity of 2×10^5 cpm/μg. From the specific activity of the ATP and the weight average molecular weight of the DNA we calculate that about 70% of the 5′ termini are labeled. This is a slightly high estimate, since the weight average molecular weight is likely to be somewhat higher than the number average molecular weight.

(g) Annealing reactions were carried out in 0.001 *M* Tris-HCl (pH 8.0), 0.1 m*M* EDTA, and 0.4 *M* NaCl at 65°. The concentration of the reactants and reaction volumes are specified in the figure legends. Duplex molecules formed in reactions between excess DNA and radioactive DNA probes were measured with staphylococcal nuclease under conditions in which only single-stranded DNA is digested (Kacian and Spiegelman, 1974). In experiments involving annealing to globin cDNA designed to assure the presence of large sequence excess, globin cDNA (0.2 ng) was annealed to 5 mg of open region reticulocyte DNA in a volume of 0.2 ml. At saturation (C_0t = 15,000) the reaction was divided into two 0.1-ml aliquots; 3 mg of either open or covered region denatured DNA was added. The salt concentrations were readjusted to 0.4 *M* NaCl, and the reactions were incubated at 69° for 48 hr. Duplex formation was assayed by single-strand specific nuclease digestion.

Annealings in which retention on hydroxylapatite was used as an assay were carried out by diluting 10-μl aliquots of the reaction mixture into 3 ml of 0.01 *M* sodium phosphate buffer (pH 6.5) at 60°, and applying to a 1 × 1 cm column of hydroxylapatite maintained at 60°. The column was washed with 0.15 *M* phosphate buffer (pH 6.5) to remove single-stranded material, and then with 0.5 *M* phosphate to elute double stranded molecules.

(h) Materials. Nucleoside triphosphates were purchased from P-L Biochemicals. [^{3}H]UTP (29.8 Ci/mmol), [^{3}H]dCTP (26.8 Ci/mmol), [α-^{32}P]ATP (30 Ci/mmol), and [α-^{32}]CTP (0.8 Ci/mmol) were purchased from New England Nuclear Corp. $(dT)_{10-12}$ was obtained from Miles

Laboratories and actinomycin D from Schwarz/Mann Corp. Poly-D-lysine hydrobromide, degree of polymerization 50 or 600, was purchased from Pilot Laboratories.

Results

In Vitro Transcription of the Globin Genes. Attempts to relate the structure of chromatin to its proposed biological role require that we possess a measurable in vitro activity related to its activity in vivo. It has been demonstrated in our own and other laboratories (Axel et al., 1973; Gilmour and Paul, 1973; Steggles et al., 1974), that chromatin isolated from reticulocytes or other tissue actively synthesizing globin mRNA retains the capacity to synthesize globin specific RNA in vitro at levels similar to that observed in the highest molecular weight nuclear RNA isolated from avian erythroblasts (Imaizumi et al., 1973). Since it is known that as the reticulocyte matures to the erythrocyte, globin RNA synthesis ceases, it was of obvious interest to determine whether this pattern of RNA synthesis could be reproduced in vitro using chromatin from these sources in transcription reactions.

We have employed the DNA complement (cDNA) of duck globin mRNA in annealing reactions to detect the presence of globin specific RNA in such transcripts. The highly radioactive cDNA was annealed to RNA transcribed in vitro utilizing *E. coli* RNA polymerase and duck reticulocyte and erythrocyte chromatin as template (Figure 1). Kinetic analysis of these data using the equations for RNA–DNA hybridization (Axel and Felsenfeld, manuscript in preparation) reveals that about 0.01% of the reticulocyte transcript is homologous to globin RNA, but only one-seventh as much globin RNA is present in the transcript from erythrocyte chromatin.

To rule out the possibility that the globin RNA was an endogenous contaminant, reticulocyte chromatin was incubated in the absence of RNA polymerase; when RNA isolated from this reaction mixture reacted with globin cDNA, virtually no evidence of annealing was observed (Figure 1) at any Cr_0t values tested. These data demonstrate that the globin sequences detected in the transcript population are synthesized de novo and are not the result of globin RNA contaminating the chromatin isolates.

Accessibility of DNA in Chromatin. We have established that reticulocyte chromatin is an active template in vitro for the transcription of the globin genes. We now wish to examine the distribution of proteins along the DNA of the chromatin, and to ask whether specific nucleotide sequences are covered by protein either in the neighborhood of the globin gene, or in the bulk of the chromatin. In previous papers we have described the use of the enzyme staphylococcal nuclease and the polycation polylysine as probes for chemically accessible DNA in chromatin (Clark and Felsenfeld, 1971, 1974). When chromatin isolated from various duck tissues is exposed to staphylococcal nuclease, about 50% of the DNA is accessible to digestion, while the rest is apparently covered by protein in such a way as to render it inaccessible to the enzyme. The values obtained for chromatin from reticulocytes, erythrocytes, and liver vary from 51 to 54%. Protein exchange is an unlikely explanation for these results. If an equivalent amount of radioactively labeled DNA is added to the chromatin, and the digestion carried out at low ionic strength, less than 5% of the DNA in the limit digest is labeled. Furthermore, identical digestion results (54% acid soluble) are obtained with chromatin in which the protein and DNA have been covalently cross-

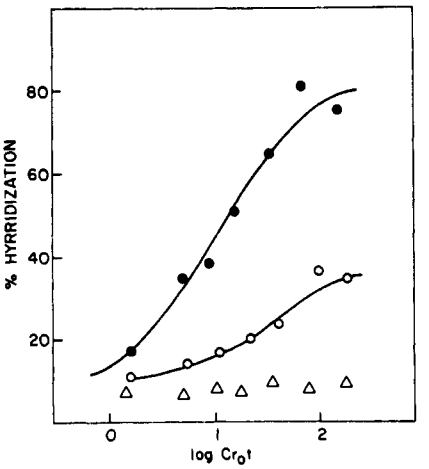

FIGURE 1: Kinetics of annealing of globin cDNA to in vitro RNA. Globin specific [^{3}H]DNA (0.2 ng) was annealed to 10 μg of reticulocyte chromatin templated RNA (●), 10 μg of erythrocyte chromatin templated RNA (○) synthesized in vitro, and 10 μg of *E. coli* DNA templated RNA added to control preparations of reticulocyte chromatin incubated in the absence of RNA polymerase (△). Hybridization reactions were performed in a volume of 50 μl and 5-μl aliquots were removed and assayed by single strand specific nuclease digestion. Cr_0t = moles of ribonucleotides × seconds/liter.

linked by formaldehyde (Brutlag et al., 1969).

The DNA of chromatin which is tightly bound to protein and resists nuclease digestion is readily obtained since, toward the end of the digestion reaction, the protected DNA fragments precipitate. After centrifugation this precipitate contains half of the original DNA complement and all of the original chromatin proteins. We call the DNA purified from this fraction "covered DNA".

The synthetic polycation, polylysine, may be used to titrate accessible phosphodiester groups on the DNA of chromatin. Titration of chromatin with polylysine reveals that about half the phosphodiester groups of DNA are free to react. Beyond this point further added polylysine fails to react (Figure 2). The regions accessible to polylysine can be isolated by titrating the chromatin with poly-D-lysine. If this complex is then treated with Pronase, the chromatin proteins are hydrolyzed, but the poly-D-lysine is resistant, leaving only poly-D-lysine bound to regions of DNA with which it was able to react. This polylysine–DNA complex is now subjected to nuclease digestion and the DNA which resists digestion (50%) is purified free of polylysine. We refer to the product as "open DNA".

Detection of Globin Genes in Open and Covered DNA. Since reticulocyte chromatin is an active template in vitro for the production of globin RNA, it seems reasonable to ask whether this activity is correlated with some special arrangement of proteins in the neighborhood of the globin genes. In particular, we can ask whether the globin genes are localized within either the open or covered regions of chromatin, by annealing the globin probe (globin cDNA) to a vast excess of either open or covered region DNA (Figure 3). We find that globin cDNA anneals to covered DNA with a $C_0t_{1/2}$ of 220, reaching a plateau at 81% duplex formation. When similar annealing is performed with open DNA a $C_0t_{1/2}$ of 300 is obtained but only 56% of the globin probe reassociates even at genomic DNA to cDNA ratios exceeding 10^7, corresponding to a tenfold sequence excess. This experiment has been performed with several samples of open and covered DNA, isolated from different preparations of reticylocyte chromatin, and with different prepara-

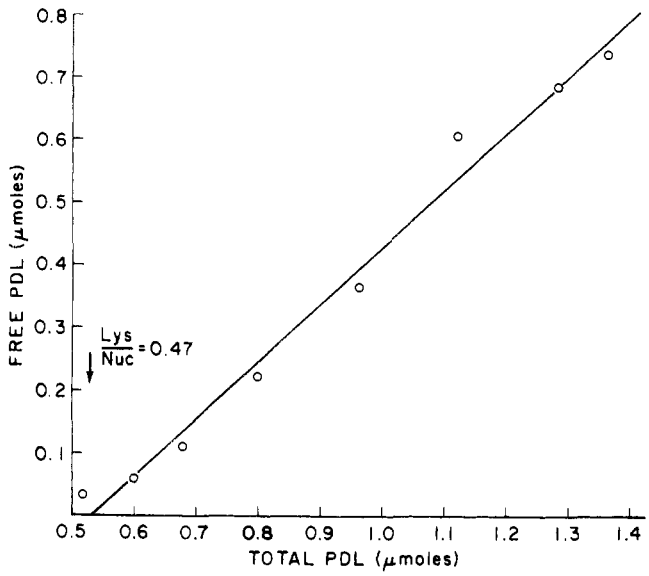

FIGURE 2: Binding of poly-D-lysine to duck reticulocyte chromatin. Varying amounts of polylysine (0.40 μmol/ml) are added dropwise to 2-ml aliquots of duck reticulocyte chromatin (0.57 μmol of DNA/ml). Both polylysine and chromatin are dissolved in 1 m*M* Tris buffer (pH 8) and the polylysine is lightly labeled by reaction with [^{3}H]fluorodinitrobenzene (Clark and Felsenfeld, 1971). Data are collected beginning at a polylysine/DNA ratio where chromatin is completely precipitated. The insoluble polylysine–chromatin complex is centrifuged from solution, and the concentration of polylysine in the supernate measured.

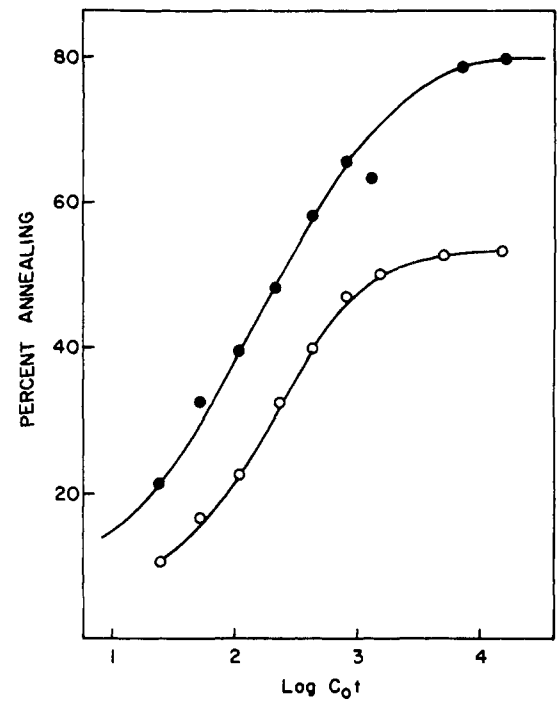

FIGURE 3: Kinetics of annealing of globin cDNA to open and covered region DNA from duck reticulocyte chromatin; 0.1 ng of [^{3}H]globin DNA was reacted with 5 mg of either covered DNA (●) or open region (○) DNA from reticulocyte chromatin. Reactions were performed in 0.2 ml and duplex formation was assayed as described (see Materials and Methods).

tions of globin cDNA. The results did not vary significantly from those shown in Figure 3. Depending upon the ratio of unlabeled DNA to globin cDNA probe, the plateau values of probe annealing to covered DNA were found to vary between 60% and a limiting value of 80 ± 3%. The latter values were obtained with a sufficiently large excess of covered DNA. In every case, the plateau values for annealing of an equivalent amount of open DNA to globin probe were lower by 15–25% of the globin probe present. It appears, therefore, that the distribution of globin specific sequences differs in the two DNA fractions derived from chromatin.

There are four possible explanations of this result. (1) The difference in annealing behavior reflects an inherent specificity of poly-D-lysine for certain sites on DNA. It is known, for example, that at high salt concentration polylysine binds selectively to A-T rich regions of DNA (Shapiro et al., 1969). In order to rule out this possibility, purified reticulocyte DNA was half-titrated with poly-D-lysine and digested with nuclease. The DNA protected by polylysine was purified and annealed to globin DNA. Using this DNA fraction, more than 80% of the probe reassociates, suggesting that in the absence of chromatin proteins polylysine is able to bind to the entire globin sequence. The specific globin sequences absent from open region DNA of chromatin must therefore be generated by the distribution of the proteins.

(2) The results shown in Figure 3 might arise from a simple quantitative difference in the abundance of complete globin genes in the two DNA fractions. To eliminate this possibility, open DNA was again annealed to globin cDNA. When the reaction reached its plateau (56%) the mixture was split and a further excess of either covered or open region DNA was added. If the difference in plateau values observed with the two DNA fractions simply represents a difference in the relative abundance of globin genes in open and covered DNA, then the addition of excess denatured open region DNA should now permit the reassociation of the remainder of the unannealed cDNA. We find that the addition of open DNA to a cDNA-open DNA annealing that has reached its plateau (see Materials and Methods) only results in a 6% increase (to 62%) in reassociation of the globin cDNA. Addition of covered region DNA, however, permits virtually complete reassociation (83%) of the globin probe.

Since the difference in annealing characteristics of open and covered region DNA is preserved, it cannot arise from a relatively small difference in the abundance of complete globin genes in the two fractions.

(3) The above discussion makes it clear that a distinct set of sequences present in the globin probe is present in covered region DNA, and completely absent from open region DNA. It is possible that these missing sequences are not globin sequences, but represent contamination of our globin cDNA by a subclass of separate molecules that are restricted to the covered DNA regions of chromatin. To exclude this possibility, we show that the sequences of the globin probe missing from the open DNA regions are covalently linked to the sequences of the globin probe which do anneal to open DNA. This is done by using hydroxylapatite chromatography to measure the amount of duplex formed between globin probe and excess open or covered region DNA. Under the conditions of the assay (see Materials and Methods) any probe molecule which is even partly involved in duplex formation will be counted as a double-stranded molecule. This is in contrast to the single-strand nuclease assay used in the earlier experiment (Figure 3), which detects only those segments of the globin probe which are actually double stranded.

We find that at C_0t values of 15,000, about 80% of the cDNA probe reassociates with covered DNA as measured either by nuclease digestion or hydroxylapatite chromatog-

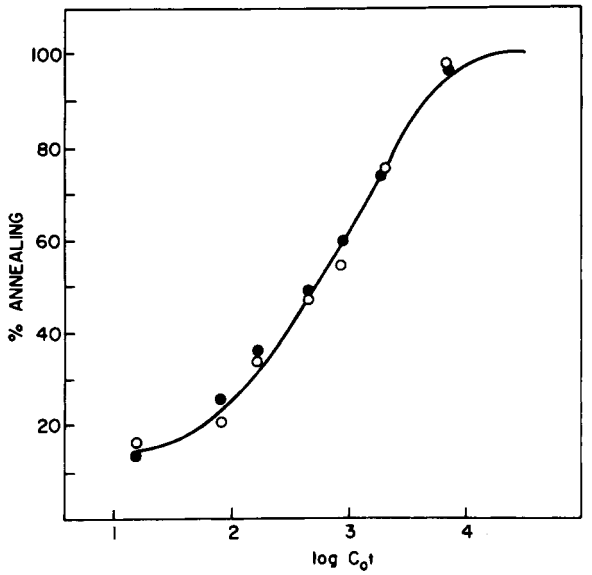

FIGURE 4: Kinetics of renaturation of unlabeled open and covered region DNA; 3 mg of either open (○) or covered (●) region reticulocyte DNA were denatured in a volume of 0.2 ml. Renaturation was monitored by hydroxylapatite chromatography. Saturation values of 67% for both open and covered DNA fractions were normalized to 100% in this figure only.

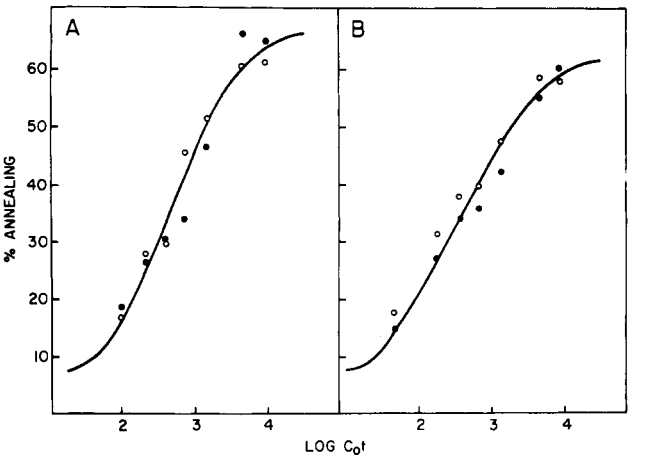

FIGURE 5: Kinetics of annealing of open and covered DNA labeled in vitro with DNA polymerase III. Open and covered region DNA from reticulocyte chromatin were treated with exonuclease III and repaired with DNA polymerase III to yield DNA with a specific activity of 4×10^5 cpm/μg. (A) ^{3}H-covered region DNA (0.02 μg) was reacted with 3 mg of either covered (●) or open (○) region reticulocyte DNA; (B) ^{3}H-open region DNA (0.02 μg) was reacted with 3 mg of either covered (●) or open (○) region DNA. Annealing reactions were performed in a volume of 0.15 ml and duplex formation was monitored by single strand specific nuclease digestion.

raphy. The saturation values observed upon annealing of Hb DNA to open region DNA are 79% when scored by hydroxylapatite and only 56% when determined by nuclease. The fact that most of the globin probe is retained on hydroxylapatite using either open or covered region DNA means that the sequences missing from open region DNA do not correspond to a class of contaminant molecules in the globin probe, but must be covalently linked to the globin gene sequences which anneal to open DNA.

(4) The last, and we believe the correct, explanation of our results is that while the entire complement of globin genes is present in covered region DNA, a distinct region of the globin genes is absent from open DNA fractions. There appears to be a specific portion of the globin gene which is always covered by proteins in chromatin even in a tissue where these genes are actively transcribed.

It was now of interest to determine whether this characteristic distribution of proteins covering the globin genes in reticulocyte chromatin is also observed in erythrocyte chromatin. Open and covered region DNA were therefore prepared from erythrocyte chromatin and annealed to globin cDNA. Globin cDNA anneals to either open or covered regions of erythrocyte chromatin with virtually identical kinetics, and with 80% of the probe reassociated at completion with both DNA fractions. The pattern of protein distribution along the globin genes is therefore different in erythrocytes and reticulocytes.

Comparison of Sequences in Open and Covered DNA. These experiments using globin cDNA and cRNA to study the association of proteins with the transcribed sequences limit us to a study of only a small percentage of the total genomic DNA. To study the patterns of proteins associated with the entire genome without regard for transcriptional capacity, we examined the properties of total open and covered DNA.

Covered region DNA isolated from reticulocyte chromatin had a weight average degree of polymerization of 150, as compared with 210 base pairs for open region DNA. In separate experiments, unlabeled open or covered region DNA was denatured and allowed to self-anneal. Duplex formation was monitored by hydroxylapatite chromatography. As seen in Figure 4, the two fractions have identical kinetics of reassociation with a $C_0t_{1/2}$ of 920. No significant difference in the relative abundance of reiterated sequences can be noted in the two DNA fractions. It should be noted that in Figure 4 the values have been normalized to 100% reassociation. The actual saturation values observed were 67%. The most likely explanation for the inability of these fractions to reassociate completely is that a substantial portion of these low molecular weight DNA fractions form duplexes shorter than the minimum length required to be retained on hydroxylapatite.

We now asked whether open and covered DNA consist of distinct sequences. If this were the case, it would appear that all the covered regions are comprised of specific and fixed sequences on DNA. If, however, the sequences of covered region DNA are present in open DNA as well, it is likely that the majority of the chromatin proteins are distributed along the genome randomly with respect to nucleotide sequence.

Radioactive covered and open DNA were prepared in vitro by subjecting the DNA to partial digestion by exonuclease III followed by repair synthesis in the presence of DNA polymerase III and labeled deoxynucleoside triphosphates. The small size of the DNA probe and the fidelity of DNA synthesis observed with DNA polymerase III make this an exceedingly useful procedure for in vitro labeling of DNA at high specific activity. These radioactive DNA fractions can now be used as a probe in annealing reactions with open and covered DNA. The kinetics of annealing of labeled covered region DNA with a vast excess of either open or covered unlabeled DNA are identical (Figure 5A). Similar results are obtained in reassociation experiments with labeled open region DNA (Figure 5B). These results indicate that the majority of the sequences in the duck genome are shared by both open and covered regions of chromatin.

Because it was impossible to determine whether the DNA labeled in vitro with DNA polymerase was uniformly

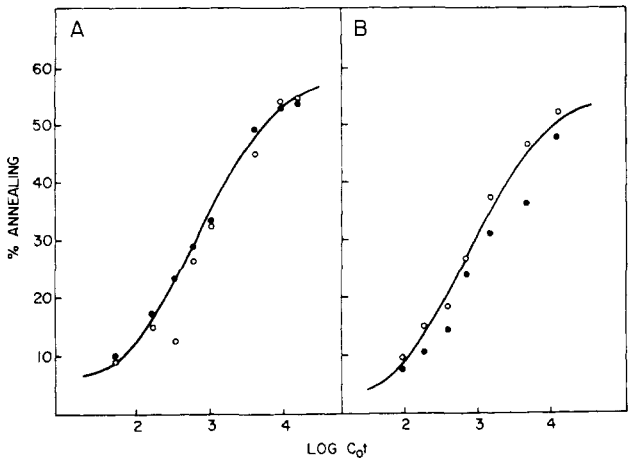

FIGURE 6: Kinetics of annealing of open and covered region DNA labeled with polynucleotide kinase. Open and covered DNA fractions were labeled in vitro with [λ-^{32}P]ATP and polynucleotide kinase to a specific activity of 2 × 10^5 cpm/μg. (A) ^{32}P-covered region DNA (0.05 μg) was reacted with 3 mg of either covered (●) or open (○) region DNA. (B) ^{32}P-open DNA (0.05 μg) was reacted with 3 mg of either covered (●) or open region reticulocyte DNA. Reactions were performed in 0.15 ml and were assayed by hydroxylapatite chromatography.

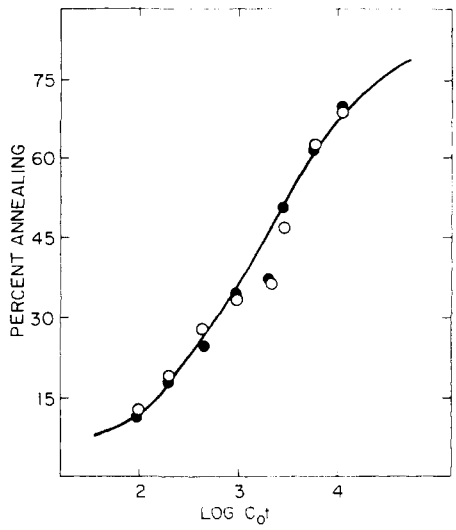

FIGURE 7: Kinetics of annealing of open and covered region DNA prepared from cross-linked chromatin. Duck reticulocyte chromatin was treated with formaldehyde to cross-link the proteins and DNA (Brutlag et al., 1969). Covered region DNA (4 mg) was prepared from this chromatin and annealed with 0.02 μg of ^{3}H-covered DNA (●) or ^{3}H-open region (○) labeled in vitro with DNA polymerase.

labeled in all sequences, radioactive open and covered DNA fractions were also prepared with γ-^{32}P-labeled ATP and polynucleotide kinase (see Materials and Methods). These radioactive DNA fractions were then annealed with an excess of unlabeled DNA from either covered or open regions as described above (Figure 6). The results confirm that the majority of the DNA of open and covered regions consists of identical sequences.

It seemed possible that many of the proteins of chromatin might be free to exchange or slide along the DNA during the digestion processes used in preparation of open and covered DNA, and that the random association of protein with DNA is not a characteristic of unperturbed chromatin. To exclude this possibility, chromatin was treated with formaldehyde prior to the preparation of covered regions. This treatment results in the formation of linkages between protein and DNA and prevents protein exchange or sliding during the digestion process. The formaldehyde-treated chromatin can be digested with staphylococcal nuclease, and behaves in the digestion process like untreated chromatin. The proteins are all covalently linked to the DNA of the limit digest, and there is no DNA in the digest that is free of protein, as judged by isopycnic banding behavior in CsCl in the analytical ultracentrifuge. Recent studies (Sollner-Webb et al., unpublished) suggest that the size distribution of DNA in the limit digest is similar to that obtained from untreated chromatin, so that it seems unlikely that formaldehyde treatment itself causes histone rearrangement. Covered DNA was prepared from the digest and annealed in vast excess to radioactive open and covered DNA. The two labeled DNA fractions reassociated with this covered DNA with identical kinetics, saturating at greater than 70% duplex formation (Figure 7). It appears then that protein exchange does not occur during the preparation of covered regions, leading us to the conclusion that the majority of proteins of "native" chromatin are not sequence selective in their binding sites.

Discussion

Our observations of the specific synthesis of globin RNA by avian reticulocyte chromatin (Axel et al., 1973) in vitro and those of Gilmour and Paul (1973) and Steggles et al. (1974) using mammalian erythroid chromatin suggest that the transcriptional controls operating in vivo are retained in our chromatin preparations and can be recognized by a bacterial polymerase. It is becoming increasingly clear from reconstitution experiments (Felsenfeld et al., 1974; Paul et al., 1973) that those factors responsible for the specific enchancement of globin RNA synthesis reside within the protein fraction of chromatin. We do not know, however, whether the polymerase is capable of recognizing a natural initiator sequence adjacent to the globin genes or whether these genes are simply more accessible to polymerase binding in reticulocyte chromatin.

In our initial studies of the accessibility of DNA in chromatin to a variety of chemical and biological probes, we found that about half of the DNA in chromatin is susceptible either to attack by staphylococcal nuclease or to titration by the polycation, poly-D-lysine (Clark and Felsenfeld, 1971). These studies demonstrate that the DNA of chromatin can be structurally divided into two distinct classes: one which is tightly bound to protein (covered DNA) and another which is either free or associated with chromatin protein in such a way as to be chemically reactive (open DNA). Our experimental methods do not presently permit us to distinguish between the protective effects of histone and non-histone proteins. Most of the covered DNA certainly arises from interactions with histones, which constitute the major protein fraction (80–90%) of our chromatin preparations, but we do not know what kinds of proteins give rise to the covered regions of the globin gene.

The isolation of open and covered region DNA permits a characterization of the DNA sequences present in the two fractions. Annealing experiments (Figures 5 and 6) reveal that the majority of the sequences of the duck genome are common to both the open and covered DNA fractions. We conclude from these data that there are no specific polynucleotide recognition sequences for the majority of the protein segments bound to covered DNA; the bulk of these proteins are therefore distributed randomly along the DNA

backbone. It is unlikely that this phenomenon arises during the digestion process used to isolate the two DNA fractions since similar experiments performed with formaldehyde cross-linked chromatin yield identical results. The data suggest that a given DNA sequence in one cell could be bound tightly by protein while the identical sequence in another cell of the same type may be chemically reactive. This lack of specificity is consistent with the observation that the amino acid sequence of many of the histone proteins is highly conserved from species to species (DeLange and Smith, 1971) while the DNA of these species diverges to the point where sequence homology is no longer observed. Thus any theory of chromatin structure based upon interaction of all of the chromatin proteins with highly specific recognition sequences on DNA is implausible.

Our results do not preclude the possibility that the proteins are arranged in a regular array of repeating units along a given DNA molecule, nor do they mean that all of the genomic DNA is subject to random association with the chromatin proteins. Recent studies (Olins and Olins, 1974) suggest that histone complexes are in fact regularly spaced along the DNA relative to one another. The annealing behavior of open and covered DNA (Figures 5 and 6) shows that no long special sequences of nucleotides are involved in generating this spacing. However, very short special sequences (less than 10–20 base pairs) would go undetected in annealing experiments, and we cannot exclude the possibility that they exist. Furthermore, if some of the proteins of chromatin do act as transcriptional control factors regulating gene expression, then we would expect that specific proteins must exist which are capable of interacting with specific regulatory sequences on the DNA. We would hope therefore that at least in the neighborhood of the globin genes the arrangement of proteins on reticulocyte chromatin bears some relationship to its transcriptional activity. Annealing experiments with globin cDNA and open and covered region DNA from reticulocyte chromatin (Figure 3) show that covered DNA contains most if not all of the globin sequences, while a substantial portion (perhaps 20%) of the globin information is always absent in open region DNA. These findings suggest that while most of the globin sequences are shared by open and covered region DNA and are therefore randomly covered by protein, there exists a specific class of globin sequences which are always covered or chemically restricted by protein and on which the protein distribution is nonrandom. Since reticulocyte chromatin serves as a template for globin transcription we conclude that transcription through protein covered regions must occur and that the mere presence of protein on a specific genome is not sufficient to restrict polymerase action. Experiments with open and covered DNA from erythrocyte chromatin show that the pattern of protein distribution on the globin genes from these cells differs from that observed in reticulocytes, and support the idea that this distribution may in some way be related to transcriptional activity.

References

Attardi, G., Parnas, H., Hwang, M-I. H., and Attardi, B. (1966), *J. Mol. Biol. 20,* 145–182.

Axel, R. (1975), *Biochemistry,* in press.

Axel, R., Cedar, H., and Felsenfeld, G. (1973a), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2029–2032.

Axel, R., Cedar, H., and Felsenfeld, G. (1973b), *Cold Spring Harbor Symp. Quant. Biol. 38,* 737–783.

Axel, R., Melchior, W., Jr., Sollner-Webb, B., and Felsenfeld, G. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 4101–4105.

Berg, D., Barrett, K., and Chamberlin, M. (1971), *Methods Enzymol. 24,* 506–519.

Brutlag, D., Cameron, S., and Bonner, J. (1969), *Biochemistry 8,* 3214–3218.

Cedar, H., and Felsenfeld, G. (1973), *J. Mol. Biol. 77,* 237–254.

Clark, R. J., and Felsenfeld, G. (1971), *Nature (London), New Biol. 229,* 101–106.

Clark, R. J., and Felsenfeld, G. (1974), *Biochemistry 13,* 3622–3628.

DeLange, R., and Smith, E. (1971), *Annu. Rev. Biochem. 40,* 279–314.

Felsenfeld, G., Axel, R., Cedar, H., and Sollner-Webb, B. (1974), Ciba International Symposium on Chromatin, in press.

Gilmour, R. S., and Paul, J. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 3440–3442.

Imaizumi, T., Diggelman, H., and Scherrer, K. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 1122–1126.

Kacian, D. L., and Spiegelman, S. (1974), *Anal. Biochem. 58,* 534–540.

Kacian, D. L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L., and Marks, P. (1972), *Nature (London), New Biol. 235,* 167–169.

Kacian, D. L., Watson, K., Buruy, A., and Spiegelman, S. (1971), *Biochim. Biophys. Acta 246,* 365–383.

Kornberg, T., and Gefter, M. (1972), *J. Biol. Chem. 247,* 5369–5375.

Melli, M., Whitefield, C., Rao, V. K., Richardson, M., and Bishop, J. D. (1971), *Nature (London) 231,* 8–12.

Olins, A., and Olins, D. (1974), *Science 181,* 330–332.

Paul, J., Gilmour, R. S., Affara, N., Birnie, G., Harrison, P., Hell, A., Humphries, S., Windass, J., and Young, B. (1973), *Cold Spring Harbor Symp. Quant. Biol. 38,* 773–783.

Reeder, R. H. (1973), *J. Mol. Biol. 80,* 229–241.

Richardson, C. (1965), *Proc. Natl. Acad. Sci. U.S.A. 54,* 158–165.

Richardson, C. (1966), *Proc. Nucleic Acid Res. 1,* 212–223.

Shapiro, J. T., Leng, M., and Felsenfeld, G. (1969), *Biochemistry 8,* 3219–3232.

Simpson, R. (1973), *Adv. Enzymol. Relat. Areas Mol. Biol. 38,* 41.

Sollner-Webb, B., and Felsenfeld, G. (1975), *Biochemistry,* in press.

Steggles, A. W., Wilson, G., Kantor, J., Picciano, D., Falney, A., and Anderson, W. F. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 1219–1223.