Mäkinen, K. K., Mäkinen, P.-L., Wilkes, S. H., Bayliss, M. E., & Prescott, J. M. (1982) *J. Biol. Chem. 257*, 1765–1772.
Malkinson, A. M., & Hardman, J. K. (1969) *Biochemistry 8*, 2769–2776.
Miles, E. W. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol. 49*, 127–186.
Miles, E. W., & Moriguchi, M. (1977) *J. Biol. Chem. 252*, 6594–6599.
Miles, E. W., & Higgins, W. (1980) *Biochem. Biophys. Res. Commun. 93*, 1152–1159.
Miles, E. W., & Eun, H.-M. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 1346 (Abstr. 756).
Patthy, L., & Thész, J. (1980) *Eur. J. Biochem. 105*, 387–393.
Peters, R. G., Jones, W. C., & Cromartie, T. H. (1981) *Biochemistry 20*, 2564–2571.
Platt, T. (1978) in *The Operon* (Miller, J. H., & Regnikoff, W. S., Eds.) pp 263–302, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Ray, W. J., Jr., & Koshland, D. E., Jr. (1961) *J. Biol. Chem. 236*, 1973–1979.
Riddles, P. W., Blakeley, R. L., & Zerner, G. (1983) *Methods Enzymol. 91*, 49–60.
Riordan, J. F. (1973) *Biochemistry 12*, 3915–3923.
Riordan, J. F., McElvany, K. D., & Borders, G. L., Jr. (1977) *Science (Washington, D.C.) 195*, 884–886.
Schubert, M. P. (1936) *J. Biol. Chem. 114*, 341–350.
Takahashi, K. (1968) *J. Biol. Chem. 243*, 6171–6179.
Takahashi, K. (1977) *J. Biochem. (Tokyo) 81*, 395–402.
Tanizawa, K., & Miles, E. W. (1983) *Biochemistry 22*, 3594–3603.
Tschopp, J., & Kirschner, K. (1980) *Biochemistry 19*, 4514–4521.
Tsou, C.-L. (1962) *Sci. Sin. (Engl. Ed.) 11*, 1535–1558.
Tuszynski, G. P., Knight, L., Piperno, J. R., & Walsh, P. N. (1980) *Anal. Biochem. 106*, 118–122.
Vandenbunder, B., Dreyfus, M., Bertrand, O., Dognin, M. J., Sibilli, L., & Buc, H. (1981) *Biochemistry 20*, 2354–2360.
Weischet, W. O., & Kirschner, K. (1976) *Eur. J. Biochem. 65*, 365–373.
Weng, L., Heinrikson, R. L., & Westley, J. (1978) *J. Biol. Chem. 253*, 8109–8119.
Werber, M. M., Moldovan, M., & Sokolovsky, M. (1975) *Eur. J. Biochem. 53*, 207–216.
Wilson, D. A., & Crawford, I. P. (1965) *J. Biol. Chem. 240*, 4801–4808.
Yamasaki, R. B., Vega, A., & Feeney, R. E. (1980) *Anal. Biochem. 109*, 32–40.
Yamasaki, R. B., Shimer, D. A., & Feeney, R. E. (1981) *Anal. Biochem. 111*, 220–226.
Yanofsky, C., & Crawford, I. P. (1972) *Enzymes, 3rd Ed. 7*, 1–31.
Yanofsky, C., Platt, T., Crawford, I. P., Nichols, B. P., Christie, G. E., Horowitz, H., Van Cleemput, M., & Wu, A. M. (1981) *Nucleic Acids Res. 9*, 6647–6668.
Zalkin, H., & Yanofsky, C. (1982) *J. Biol. Chem. 257*, 1491–1500.

Isolation of a Protein Fraction That Binds Preferentially to Chicken Middle Repetitive DNA†

Michael Sanzo,‡ Bryn Stevens, Ming-Jer Tsai, and Bert W. O'Malley*

ABSTRACT: We have fractionated oviduct tissue extracts by using a combination of ion-exchange and DNA–Sephadex chromatography. By comparing the electrophoretic patterns of proteins eluted from competing specific and nonspecific DNA columns, we isolated a fraction which bound with specificity to columns containing the chicken middle repetitive sequence "CR1". This fraction showed a clear preference for binding to separate, cloned CR1 fragments derived from either the 5′ or the 3′ transition region of the ovalbumin gene domain when examined by using nitrocellulose filter binding assays. To localize the protein binding site, a CR1 clone was digested with various restriction enzymes, and the resulting fragments were examined for preferential protein binding. Results suggest that the binding site lies within a 39-nucleotide sequence which is highly conserved among different CR1 elements. This finding represents the first isolation of a protein which demonstrates a preference for binding to a middle repetitive sequence and suggests that this interaction may have a biological role. The DNA column competition adsorption method should have general application to the isolation of other gene-regulating proteins possessing DNA sequence preference.

The DNA of a large number of species has been shown to contain certain nucleotide sequences which are repeated many times and whose copies are interspersed among structural genes (Davidson et al., 1975, 1977; Schmid & Deininger, 1975). In chicken cells, a family of short (250 nucleotides) repetitive elements, designated "CR1", is present in about 7000 copies per genome (Stumph et al., 1981). CR1 sequences have been shown to contain regions homologous with Alu repetitive elements from the human and with B1 repetitive elements from the mouse and to contain short direct repeats, a characteristic common to many transposable elements (Stumph et al., 1981). Of particular interest is the positioning of CR1 elements in relation to the ovalbumin gene. In oviduct tissue, the ovalbumin gene is highly transcribed and is located within a region

† From the Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030. *Received May 21, 1984.* This work was supported by grants from the National Institutes of Health (HD-08188 and HD-07495).
‡ Present address: Division of Laboratory Medicine, Washington University School of Medicine, St. Louis, MO 63110.

0006-2960/84/0423-6491$01.50/0

of DNA that is preferentially sensitive to digestion by DNase I (Garel & Axel, 1976; Lawson et al., 1980, 1982). The DNase I sensitive domain extends over a region about 100 kilobases (kb) in length and contains two other ovalbumin related genes which are coordinately expressed with ovalbumin in response to steroid hormones (Lawson et al., 1982; Colbert et al., 1980). At the 5′ and 3′ borders of this domain, there are transition regions in which the DNA becomes progressively more resistant to nuclease digestion, reflecting structural changes in the chromatin. CR1 elements are located within both the 5′ transition region and the 3′ transition region, suggesting that they may play a role in conferring upon the DNA a conformation which is both nuclease sensitive and amenable to transcription (Stumph et al., 1983). There are several ways in which CR1 elements could mediate such an effect: First, it has been reported that repetitive sequences of the Alu family contain promoter sites for RNA polymerase III and that Alu transcripts are present in hamster cells as discrete, primary products (Duncan et al., 1981; Haynes & Jelinek, 1981). This led to the consideration that CR1 sequences may also contain promoters which, in the act of initiating transcription, or through the action of their RNA products, relieve constraints on the DNA in surrounding regions (Stumph et al., 1983).

A second possibility is that CR1 segments are acting as transposable elements, as has been suggested for other repetitive DNA sequences (Haynes et al., 1981; Calabretta et al., 1982; Van Arsdell et al., 1981). The incorporation of such elements at specific sites may cause rearrangements in the DNA which result in conformational changes leading to the expression of genes in specific tissues. However, this mechanism appears unlikely in that hybridization studies using DNA sequences near repetitive elements of the ovalbumin gene domain as probes have failed to detect any major DNA rearrangements in various tissues (Stumph et al., 1983).

A third possibility is that the function of CR1 elements is mediated through a protein, or set of proteins, which binds to a specific nucleotide segment within the CR1 sequence. We have examined oviduct tissue extracts for the presence of CR1-specific proteins. Our procedure involves fractionating oviduct tissue extracts by ion exchange and then applying each fraction to two competing columns of DNA–Sephadex linked in tandem as developed by Bearden (1980). One column has *Escherichia coli* (nonspecific) DNA, and the other column contains a cloned, CR1-containing (specific) DNA fragment. Proteins are recirculated in the presence of a decreasing salt gradient, and then bound proteins are eluted from each column separately. By comparing electrophoretic patterns of proteins eluted from the two columns, we have established a protocol in which proteins that bind selectively to the CR1 DNA column can be recovered. When examined with nitrocellulose filter binding assays, these proteins show a clear preference for DNA fragments containing CR1 elements. No preference is seen in proteins eluted from the *E. coli* DNA column. A comparison of nucleotide sequences in different preferentially bound CR1 fragments reveal a highly conserved region which is about 39 nucleotides long and which appears to contain a binding site for CR1-specific proteins.

Experimental Procedures

Recombinant DNA Clones. The preparation of the recombinant plasmids used in these studies as well as restriction maps and nucleotide sequences for the clones has been previously reported (Stumph et al., 1981, 1983). Restriction enzymes were purchased from either Bethesda Research Labs or New England Biolabs.

Labeling of DNA. Restriction fragments used in nitrocellulose filter binding assays were end labeled by using either the Klenow fragment of *E. coli* DNA polymerase I (Boehringer) and [α-^{32}P]dATP (Amersham, 3000 Ci/mmol). Preparations were extracted with phenol–chloroform and then chromatographed on columns of Sephadex G-50 to remove unincorporated label.

Construction of DNA–Sephadex Columns. p2.3/500 was digested with *Dde*I in order to yield a 920 base pair CR1-containing fragment. This was isolated by preparative gel electrophoresis and was used in constructing CR1 DNA–Sephadex columns. *E. coli* DNA (Sigma) was deproteinized by repeated extraction with phenol–chloroform and was used for constructing nonspecific DNA columns. Two separate samples were prepared: the first contained approximately 1 mg of CR1 DNA, 6 mL of deionized, distilled H_2O (dd-H_2O), and 2 g of Sephadex G-10, and the second contained 100 mg of deproteinated *E. coli* DNA, 33 mL of dd-H_2O, and 11 g of Sephadex G-10. Each of the samples was placed in a beaker and covered with parafilm. The samples were mixed by using a wrist-action shaker for 4–6 h at room temperature in order to allow the G-10 to swell and were then completely dried by using a vacuum desiccator. Samples were reconstituted in 100% ethanol (10 mL for CR1 DNA preparations and 40 mL for the *E. coli* DNA preparation) and were irradiated with UV light (85-W, low-pressure mercury vapor lamp; General Electric Model H85A) for 20 min at a distance of 15 cm. During irradiation, samples were continuously mixed by using a stir motor and magnetic stir bar. G-10 was pelleted by centrifugation and was washed twice with water to remove ethanol. The packing was allowed to swell for about 4 h at room temperature and was then washed and equilibrated with cold high-salt buffer [20 mM *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid (Hepes) (pH 7.4), 10 mM ethylenediaminetetraacetic acid (EDTA), 1.5 M NaCl, and 1 mM dithiothreitol (DTT)]. Finally, *E. coli* and CR1 preparations were packed into columns at 4 °C and washed thoroughly with high-salt buffer.

Determination of the Amount of DNA Bound. Forty milligrams of dried packing from *E. coli* and CR1 preparations and 40 mg of irradiated G-10 were weighed out and placed in screw-top test tubes with Teflon-lined caps. One milliliter of 1 N HCl was added to each tube, and samples were hydrolyzed for 30 min at 95 °C. After hydrolysis, 1 mL of 1 N NaOH was added to each sample to neutralize HCl, and test tubes were centrifuged at low speed to pellet Sephadex. Supernatants from the *E. coli* and CR1 DNA–Sephadex preparations were removed, and their absorbance at 260 nm was determined by using the supernatant from the G-10 sample as the blank. The concentration of DNA in the samples was determined from the optical density.

Isolation of CR1-Specific Binding Proteins. Oviducts from diethylstilbestrol-stimulated White Leghorn chicks were removed and immediately frozen in liquid nitrogen. These were stored at −70 °C until use. A 250-g sample of frozen oviduct tissue was pulverized in a Waring blender and washed 3 times with 500-mL aliquots of ice-cold phosphate-buffered saline containing 2.5 mM DTT and 2 mM phenylmethanesulfonyl fluoride (PMSF). All subsequent steps were carried out at 0–4 °C. The washed tissue was reconstituted with 2 volumes (500 mL) of buffer [20 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) (pH 7.9), 2.5 mM $MgCl_2$, 0.2 mM EDTA, 0.25 M sucrose, 2.5 mM DTT, and 2 mM PMSF] and was tissuemized 3 times with 30-s bursts (Tekmar Tissuemizer, Cincinnati, OH). The suspension was homo-

genized (two strokes) by using a motor-driven homogenizer with a Teflon pestle and then centrifuged at 6000 rpm (4360*g*) for 15 min. The pellet was reconstituted with 1.5 volumes (375 mL) of buffer [10 mM Tris-HCl (pH 7.9), 1 mM EDTA, 2.5 mM DTT, and 2 mM PMSF] and homogenized (2 strokes). The homogenate was incubated at 4 °C for 45 min with constant stirring and then rehomogenized (7–10 strokes). One-half volume (125 mL) of buffer [50 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 25% sucrose, 30% glycerol, 2.5 mM DTT, and 2 mM PMSF] was added, and the preparation was placed on a stir motor at 4 °C. Proteins were slowly extracted with ammonium sulfate until the preparation was 10% saturated. At that point, chromatin containing unextracted proteins was removed by centrifugation at 50 000 rpm for 2 h at 2 °C in a Beckman 60 Ti rotor. The supernatant was removed, and solid ammonium sulfate was added until the preparation was 65% saturated. Precipitated protein was collected by centrifugation at 18 000 rpm for 30 min in a JA 20 rotor. Pellets were reconstituted in about 50 mL of buffer [20 mM Hepes (pH 7.9), 0.2 mM EDTA, 5 mM $MgCl_2$, 20% glycerol, 2.5 mM DTT, and 2 mM PMSF] and dialyzed against the same buffer overnight.

The extract prepared as described above was separated by ion exchange using basically the same procedure previously described for the isolation of oviduct transcription factors (Tsai et al., 1981), except that larger columns were used [300 mL for the DEAE-Sephadex column (4 × 24 cm) and 60 mL for the phosphocellulose column (2.4 × 13 cm)], the concentration of DTT was 2.5 mM in all elution buffers, and no spermidine was included in the phosphocellulose elution buffer. The fraction most extensively used in our studies was MP-1 which was derived from a protein fraction that was unadsorbed by DEAE-Sephadex and which eluted from phosphocellulose with 1 M KCl. This ion-exchange fraction was given the designation P-1000. In terms of recovery, we applied 250–300 mg of protein to the DEAE-Sephadex column and obtained 200–240 mg in the flow-through fraction (DE50). This fraction was then applied to the phosphocellulose column. At the end, we recovered 12–20 mg of protein eluted in the P-1000 fraction.

Ion-exchange fractions were separated on columns of DNA–Sephadex by using a procedure similar to that reported by Bearden (1981) and as described in the text.

Nitrocellulose Filter Binding Assays. Nitrocellulose filter binding assays were performed by using a procedure similar to that previously described (Compton et al., 1983). Nitrocellulose sheets (Millipore, 0.45-μm pore size) were cut into 0.5 cm × 0.5 cm squares. These were soaked in 0.5 N KOH for 15 min and then equilibrated with assay buffer [50 mM potassium phosphate (pH 7.4), 0.1 mM EDTA, 10% glycerol, and 50 μg/mL bovine serum albumin (BSA)]. Protein–DNA binding reactions proceeded at 4 °C for a period of 30 min in a total reaction volume of 400 μL. Reactions were terminated by filtering samples through nitrocellulose squares in 200-μL portions. Filters were rinsed with two 200-μL aliquots of assay buffer minus BSA in order to remove protein-free DNA. In competition experiments, filters were placed in 22-mL scintillation vials, mixed with 0.5 mL of 0.5 N HCl, and then dissolved in 1 mL of ethyl acetate. Ten milliliters of Aquasol counting fluid (New England Nuclear) was added to each vial, and radioactivity was determined by liquid scintillation counting. In preferential binding experiments, filters were cut into small pieces and incubated in 200 μL of extraction buffer [10 mM Tris-HCl (pH 7.4), 0.5 M ammonium acetate, and 0.5 M sodium acetate] at 50 °C for 30 min.

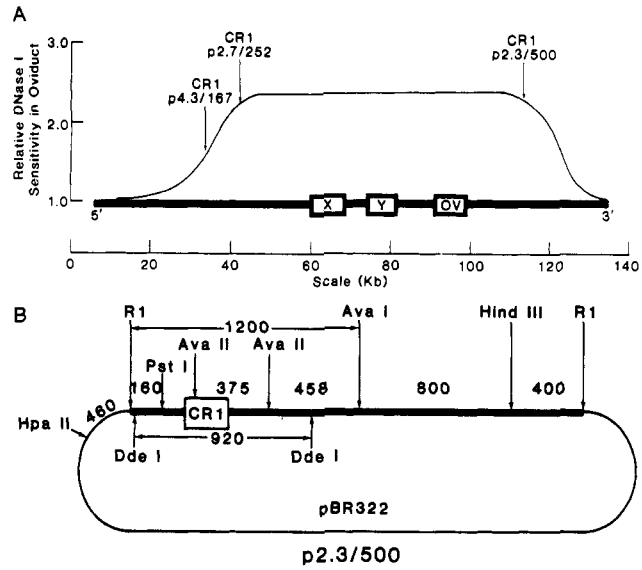

FIGURE 1: Location of CR1 segments within ovalbumin DNase I transition regions. The location of CR1 elements with respect to the ovalbumin and the X and Y multiple gene domains has been determined (Stumph et al., 1983) and is shown in (A). The CR1 elements have been cloned and given the designations p4.3/167, p2.7/252, and p2.3/500. A restriction map of p2.3/500 (the fragment used most commonly in our studies) is also shown (B).

Samples were centrifuged, supernatants were removed, and 200 μL more of extraction buffer was added to each filter. Samples were incubated at 37 °C for 30 min and centrifuged and the resulting supernatants placed with those from the first extraction. Filter extracts received 5 μL of 1 mg/mL carrier RNA and were then precipitated with 2 volumes of ethanol. Resulting DNA pellets were washed twice with 70% ethanol and electrophoresed on agarose gels. Bands were visualized by autoradiography.

Sodium Dodecyl Sulfate (SDS) Gel Electrophoresis of Proteins. Proteins were treated with SDS and separated on 12.5% polyacrylamide gels according to the procedure of Laemmli (1970).

Protein Assays. The concentration of protein in various preparations was determined by the method of Lowry (Lowry et al., 1951) using bovine serum albumin (Sigma) as a standard.

Results

Location of CR1 Segments in the Ovalbumin Gene Domain. The location of CR1 segments with respect to the ovalbumin gene has been previously reported (Stumph et al., 1983) and is shown in Figure 1A. The gene lies within a 100-kilobase (kb) section of DNA that is preferentially sensitive to digestion by DNase I (Lawson et al., 1982). This region is bordered at the 5′ and 3′ ends by transition regions in which the DNA becomes increasingly resistant to digestion. Two CR1 segments are found within the 5′ transition region and one within the 3′ region. DNA fragments containing each of the CR1 segments have been cloned, and their positions are shown in the figure. The clones p2.3/500 and p2.7/252 were used in the nitrocellulose filter binding assays described below. In addition, fragments obtained by digesting p2.3/500 with restriction endonucleases were used in isolating CR1-specific proteins and in studies designed to localize protein binding sites. A restriction map of this clone is shown in Figure 1B.

Isolation of CR1-Specific Protein. Chick oviduct tissue was extracted with 10% ammonium sulfate. The extract was applied to a DEAE-Sephadex column in buffer containing 50 mM ammonium sulfate. Proteins adsorbed by the column

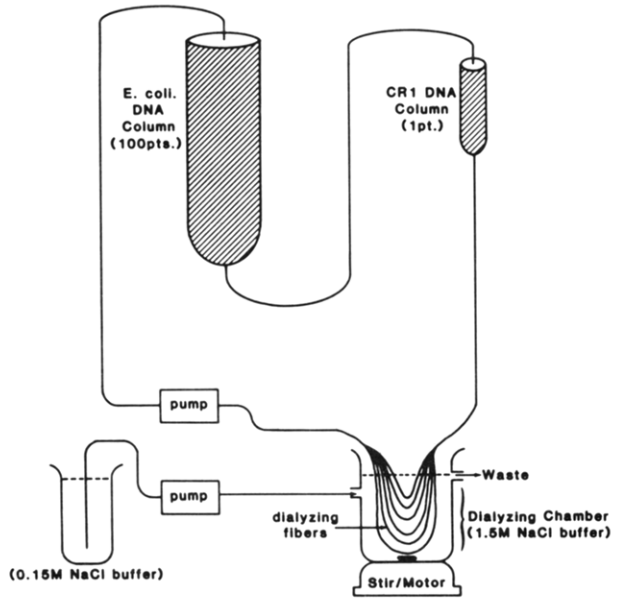

FIGURE 2: Isolation of specific CR1 DNA binding proteins. Chick oviduct tissue was extracted with 10% ammonium sulfate and the extract fractionated by ion exchange on columns of DEAE-Sephadex and phosphocellulose. Proteins preferentially binding to CR1-containing DNA were isolated from an ion-exchange fraction eluting from phosphocellulose with 1 M KCl using columns of DNA arranged as shown. The CR1 DNA column was prepared from a 920-nucleotide fragment obtained from p2.3/500 by digestion with *Dde*I (Figure 1B).

were eluted in steps of 175 mM ammonium sulfate and 500 mM ammonium sulfate. Material unadsorbed by the anion-exchange column was applied to a phosphocellulose column in buffer containing 100 mM KCl. The unadsorbed fraction was collected, and the column was then eluted in steps of 350 mM KCl and 1000 mM KCl.

Between 12 and 20 mg of protein from each ion-exchange fraction was dialyzed against "high-salt" buffer [20 mM Hepes (pH 7.4), 10 mM EDTA, 1.5 M NaCl, and 1 mM dithiothreitol (DTT)] and applied to the DNA columns illustrated in Figure 2. One column contained approximately 60 mg of *E. coli* DNA covalently bound to Sephadex G-10 (1.6 × 15 cm) by UV irradiation. Hybridization studies have failed to reveal the presence of CR1 segments in *E. coli* DNA (W. Stumph, personal communication). This nonspecific DNA column was linked in tandem to a smaller, CR1-containing DNA–Sephadex column (0.9 × 10 cm). The second column had about 600 μg of a 920 base pair fragment derived from clone p2.3/500 by digestion with *Dde*I (Figure 1B). Effluent from the CR1 DNA column was pumped through the hollow fiber dialyzing system shown at the bottom of Figure 2 and then recirculated back onto the *E. coli* DNA column. The hollow fibers allow the passage of small inorganic salts but retain larger molecular weight species such as proteins. At the start of experiments, columns were equilibrated with high-salt buffer, and the reservoir surrounding the fibers was filled with this buffer. As proteins were recirculated through the columns, low-salt buffer [20 mM Hepes (pH 7.4), 10 mM EDTA, 0.15 M NaCl, and 1 mM DTT] was pumped into the hollow fiber buffer chamber and mixed by means of a magnetic stir bar and motor. This created a gradient of decreasing ionic strength within the column system.

Proteins were recirculated at a flow rate of 1.5–2 mL/min for a period of time sufficient to allow the complete equilibration of columns with low-salt buffer (approximately 24 h). At the end of this time, columns were uncoupled, and each was washed with 5 column volumes of low-salt buffer to remove unbound protein. Adsorbed proteins were eluted with 1.5 column volumes of high-salt buffer. Eluted protein was stabilized by the addition of an equal volume of a second buffer [0.05 M Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 25% sucrose, 30% glycerol, 2.5 mM DTT, and 2 mM PMSF] and immediately concentrated in a dialysis bag by placing the bag in a container of dry Sephadex G-10 or G-25. Preparations were dialyzed to remove salt, lyophilized, and examined by SDS–polyacrylaide gel electrophoresis. Proteins with a specificity for CR1 DNA should appear in gels as bands present in "CR1-adsorbed" lanes at positions not occupied by prominent bands in "*E. coli* adsorbed lanes". Of the various protein preparations examined, only that derived from the ion-exchange fraction eluting from phosphocellulose with 1000 mM KCl (the "P-1000" fraction) had such bands (Figure 3). The CR1-adsorbed lane shows a band with a molecular weight of about 42 000, which is much more intense than what would be expected of a nonspecific binder on the basis of an observation of bands in the same position in the *E. coli* adsorbed lane. Several other bands of relatively weak intensity are also present in the CR1-adsorbed lane. The dramatic enrichment of the 44 000-dalton protein in relation to these other bands is a reflection of its greater perference for the CR1-containing DNA column.

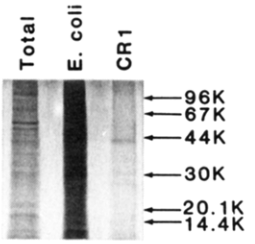

FIGURE 3: SDS gel electrophoresis profile of proteins adsorbed by *E. coli* DNA and CR1-containing columns. Proteins in the P-1000 ion-exchange fraction were allowed to partition between columns of *E. coli* DNA and CR1 DNA as described in the text. Proteins bound to each column were electrophoresed on 12.5% polyacrylamide gels and visualized by silver staining. The lane marked "total" shows the proteins present in the P-1000 precursor fraction. The following two lanes show the proteins bound by the *E. coli* and CR1 DNA columns.

The amount of protein recovered from *E. coli* and CR1 DNA columns varied from preparation to preparation. In general, about 500 μg of protein was recovered from the *E. coli* column when 12 mg of P-1000 was applied. Protein levels in CR1 DNA fractions were often too low for an accurate determination using the Lowry assay; however, a rough estimate would place levels at between 5 and 10 μg.

Nitrocellulose Filter Binding Assays of CR1-Specific Protein Fractions. The P-1000-derived protein fraction isolated from the CR1 column was given the designation "MP-1" and was examined for specificity in DNA preferential binding assays. In these assays, ^{32}P end-labeled DNA fragments and proteins are incubated together and then passed through a nitrocellulose filter. Protein–DNA complexes are retained by the filter whereas free DNA is not. DNA fragments are then extracted from the filter, separated by electrophoresis, and visualized by autoradiography. By comparing the intensity of the signals recorded at various positions on the X-ray film, it can be determined if the proteins have shown a preference for binding to a particular DNA fragment. Figure 4 shows the results of an assay performed by using the insert and plasmid DNA obtained from the digestion of p2.3/500 with *Eco*RI. MP-1 protein concentration was roughly 1 ng/μL. The preparation shows a clear preference for binding to the CR1-containing DNA at doses of between 20 and 80 μL. Scans of autoradiographic films using a densitometer showed

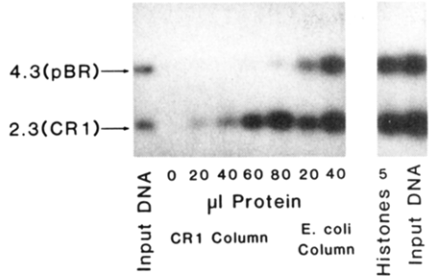

FIGURE 4: Preferential binding of CR1-adsorbed protein in the nitrocellulose filter binding assay. Clone p2.3/500 was digested with *Eco*RI to yield a 2.3-kb CR1-containing fragment and a 4.3-kb fragment of pBR322 plasmid DNA. These were ^{32}P end labeled and incubated with fraction MP-1 at 4 °C for 30 min. Samples were passed through nitrocellulose filters which retain protein–DNA complexes. DNA was extracted from the filters and electrophoresed on 0.7% agarose gels. For comparison, protein adsorbed by the *E. coli* DNA column and also core histones (10 μg/mL) isolated from chick oviduct DNA were also examined in the assay. The labeled DNA fragments are shown at both low (left panel) and high (right panel) doses.

preparations of MP-1 to have a preference for the CR1 insert of p2.3/500 as high as 18-fold, with an average value for all experiments of greater than 10-fold. Neither *E. coli* adsorbed proteins nor histone proteins showed any similar preference relative to input DNA.

The specificity of CR1-adsorbed proteins was also studied by using competitive nitrocellulose filter binding assays (data not shown). Incubations were performed in the presence of the labeled 2.3-kb insert and increasing doses of unlabeled competitor DNA. Unlabeled 2.3 was found to be an effective competitor for MP-1 while chick DNA fragments from clones devoid of CR1 sequence and plasmid fragments were relatively poor competitors. In contrast, 2.3, CR1-free chick DNA, and pBR322 all competed equally well for proteins isolated from the *E. coli* DNA column, as would be expected for nonspecific DNA binding proteins.

Preferential Binding of MP-1 to the Insert of p2.7/252 and Localization of the Protein Binding Site. The results from the above experiments indicate that MP-1 selectively binds the CR1-containing insert of p2.3/500. However, since the CR1 region contains only about 11% of the nucleotides in this fragment and only about 27% of the nucleotides in the fragment which was used in isolating MP-1, it was possible that the preferential binding site for proteins in MP-1 lay outside the CR1 region. To test this possibility, preferential binding assays were performed using p2.7/252, a separate CR1-containing clone derived from the 5′ transition region of the ovalbumin gene domain (Figure 1A). Analyses of partial nucleotide sequences from the inserts of p2.3/500 and p2.7/252 as well as Southern hybridization analysis have indicated that these fragments share no regions of homology outside of the CR1 sequence. It can be seen from the results shown in Figure 5 that MP-1 has a preference for the 2.7-kb insert (17-fold as determined by densitometry) similar to that seen with the 2.3-kb insert. Since the CR1 segments are the only regions which these fragments have in common, it appears that the binding site for MP-1 proteins lies within the CR1 sequence.

A similar line of reasoning can be used in considering the location of the binding site within the CR1 region. CR1 repetitive elements are about 250 nucleotides in length whereas a typical protein binding site would be expected to be comprised of about 20 nucleotides. A sequence comparison made between the CR1 segments from four cloned fragments, including those from p2.7/252 and p2.3/500, has shown that these repetitive elements have two regions of high homology

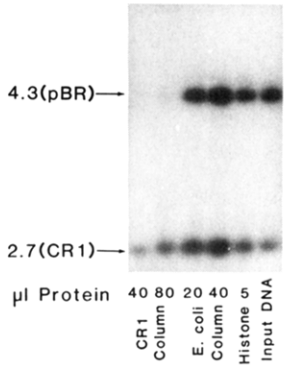

FIGURE 5: Preferential binding of MP-1 to the insert of p2.7/252. The insert of clone p2.7/252 is derived from the 5′ transition region of the ovalbumin gene domain and shares no detectable homology with the insert of p2.3/500 outside of the CR1 region. The results of a preferential binding assay performed with this fragment and MP-1 are shown. The labeled DNA fragments used in the incubations are shown on the far right. Also shown is the result of an incubation performed with core histones (10 ng/μL) and the results of assays performed with protein adsorbed by the *E. coli* DNA column.

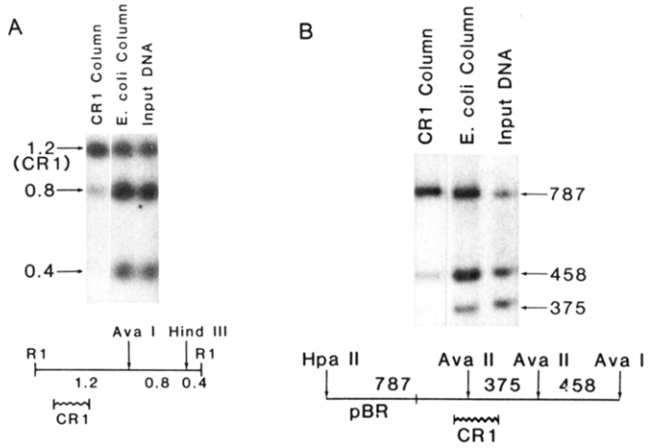

FIGURE 6: Localization of protein binding site in CR1-adsorbed protein. The insert of p2.3/500 was digested with *Ava*I and *Hin*dIII to give fragments of 1.2, 0.8, and 0.4 kb. These were used in preferential binding assays performed with protein obtained from either the CR1 DNA column (MP-1) or the *E. coli* DNA column. Results are shown in panel A. A second fragment was isolated from p2.3/500 after digestion with *Hpa*II and *Ava*I. *Ava*II was then used to generate fragments of 787, 375, and 458 nucleotides. Approximately 60 base pairs of the CR1 segment were located on the 787-nucleotide fragment and the rest on the 375-nucleotide fragment. The result of a preferential binding assay performed with these fragments is shown in panel B. The protein fractions were the same as those in panel A.

separated by a region of low homology (Stumph et al., 1983). The first homologous region is 100 nucleotides long and has about 80% of its nucleotides conserved among the clones. There is then a stretch of about 60 nucleotides of low homology, followed by a second highly homologous region (about 85%) approximately 39 nucleotides long. From Figures 4 and 5, the binding site for MP-1 proteins would be expected to be in one of the highly conserved sequences. In order to determine which homologous region is more likely to contain the MP-1-specific protein binding site, we cleaved a CR1 segment at a site between the two and performed preferential binding assays on the resulting fragments. p2.3/500 was first digested with *Eco*RI, *Hin*dIII, and *Ava*I to give a 1.2-kb CR1 fragment whose entire sequence was known. A preferential binding assay was performed which confirmed that MP-1 selectively bound the CR1-containing 1.2-kb fragment (Figure 6A).

A similar fragment was then isolated from p2.3/500 after digestion with *Hpa*II and *Ava*I. The use of *Hpa*II instead of

```
Consensus: C A T T G A G G T C C C T T C C A A C C C A N A C C A T T C T A T G A T T C T
p2.3/500:          T          T    T              G G
p2.7/252:   T                                     A       T                  A G
p4.3/167:     - -                                T C T     A                      C
```

FIGURE 7: Homology between CR1 sequences in the region of DNA preferentially bound by MP-1. MP-1 preferentially bound to a DNA fragment derived from p2.3/500 which contained approximately 60 nucleotides from the CR1 segment. Included in this sequence was a 39-nucleotide segment which is highly conserved among different CR1 elements (between nucleotides 177 and 215 of the published consensus sequence; Stumph et al., 1983) and which is shown above for the three CR1 elements in the vicinity of the ovalbumin gene. The consensus sequence was determined by choosing the most common nucleotide at each position when the individual sequences are aligned for maximum homology. Deviations from the consensus sequence for each individual sequence are shown on the lines below. The letter "N" was used to indicate that there was no agreement among the three individual sequences at this position. The dashes in the sequence of p4.3/167 indicate positions that were deleted in aligning sequences for maximum homology.

*Eco*RI in this digestion resulted in the inclusion of about 400 extra nucleotides from pBR322 on the 1.2-kb piece. This was done in order to facilitate the separation of fragments after a final digestion was performed, this time using *Ava*II. The procedure resulted in the generation of three fragments; a 787 base pair fragment which contained the smaller region of homology and 21 additional nucleotides from the CR1 segment; a 375 base pair piece containing the other region of homology and the rest of the CR1 sequence; and a 458 base pair piece which was devoid of CR1 sequence. The results of preferential binding experiments performed with the three fragments are shown in Figure 6B. The preference of MP-1 for the 787 base pair fragment over either the 375 or the 458 base pair fragments is evident from these results. Although *E. coli* adsorbed proteins also appear to bind slightly better to the larger, 787 nucleotide fragment than to the 375 base pair fragment, densitometric scans show this difference to be minimal (about 2-fold and proportional to size) when compared to the preference (>10-fold) exhibited by the MP-1 proteins. The results suggest that the specific binding site for proteins in MP-1 lies within the first 60 nucleotides of the CR1 segment in p2.3/500 and is probably within the 39 base pair sequence shared by the CR1 fragment from p2.7/252. The nucleotide sequences of this region in the three CR1 segments associated with the ovalbumin gene are shown in Figure 7. It can be seen that when compared with a consensus sequence determined by taking the most common nucleotide at each position, 87% of the nucleotides in p2.3/500 and p2.7/252 and 82% of the nucleotides in p4.3/167 are conserved.

Discussion

The results presented in this paper show that oviduct tissue contains proteins that bind with specificity to DNA clones which have been derived from either the 3′ or the 5′ transition region of the ovalbumin gene domain and which contain the CR1 middle repetitive sequence. These proteins were identified in a fraction, MP-1, which was isolated from oviduct tissue extracts by using a DNA–Sephadex column prepared from cloned CR1-containing fragments. Neither pBR322 plasmid DNA nor DNA fragments cloned from a part of the ovalbumin gene domain free of CR1 sequence could effectively compete with CR1 fragments for MP-1 proteins when examined by nitrocellulose filter binding assays. No preferential binding was ever observed to fragments which did not contain at least part of the CR1 sequence, and no preferential binding was observed in protein fractions isolated from DNA columns not enriched in CR1 segments.

The view that MP-1 proteins are binding somewhere within the CR1 sequence is supported by the observation that the inserts of p2.7/252 and p2.3/500, two clones which contain no detectable regions of homology outside of the CR1 sequence, are both preferentially bound. Within the CR1 segment, there are two regions which are highly conserved. When a CR1 clone was digested in such a way as to separate the conserved regions, it was found that preferential binding occurred to fragments containing the smaller, 39-nucleotide, region of homology. This is the most likely site for the binding of preferential proteins in MP-1 and is shown in Figure 7. It is interesting to note that the segment from p2.3/500 contains the sequence GGCCATTCT which is identical with a sequence found 70–80 base pairs upstream from the transcription start site of a number of cellular and viral genes (Benoist et al., 1980; Efstratiadis et al., 1980). However, the positions of these nucleotides appear to be the most poorly conserved within the region of homology, making the significance of their presence uncertain.

There are several protein bands in our MP-1 preparation, one major band at 42 000 daltons and several minor bands (Figure 3). At present, we do not have any direct evidence concerning which of these bands is responsible for the specific binding to CR1 DNA.

We have shown that the CR1 sequence has a short sequence homology to the Alu family in the human genome (Stumph et al., 1981). Since Alu sequences are transcribed by the RNA polymerase III transcription system, it is possible that CR1 sequences may also be transcribed by the same enzyme. If this is the case, our MP-1 protein(s) may be identical with or similar to TFIIIa factor which binds to the 5S RNA gene promoter (Engelke et al., 1980). However, DNA sequence analysis was not able to show any sequence homology between the split promoter of polymerase III transcribed genes (5S gene, tRNA genes, and Alu sequences) and CR1 sequences. Furthermore, we were unable to detect transcription of the CR1 sequences located within the ovalbumin gene chromatin domain using an in vitro transcription system. In addition, Northern analysis of in vivo RNA which corresponds to CR1 sequences did not yield a low molecular weight RNA species characteristic of RNA polymerase III transcription products. Finally, we have sequenced seven CR1 sequences and noticed that these sequences closely resemble the LTRs of retroviruses (unpublished results). Thus, it is likely that CR1 sequences are transcribed (if they are transcribed at all) by RNA polymerase II.

It is known that the DNA in the region of many actively transcribed genes exists in a conformation which is preferentially sensitive to digestion by DNase I (Lawson et al., 1982; Stalder et al., 1980; Wu & Gilbert, 1981; Tata et al., 1980), and certain proteins, for example, high mobility group (HMG) proteins 14 and 17, appear to be involved in conveying upon the DNA in regions surrounding transcribed genes a conformation which is sensitive to the enzyme and suitable for transcription (Weisbrod et al., 1980; Gazit et al., 1980). Although we have no direct evidence relating to the function of CR1-specific DNA binding proteins, the location of the

repetitive elements with respect to the ovalbumin gene suggests that such proteins could play a role in altering the conformation of the DNA in this region. Like middle repetitive sequences in other species, the CR1 elements of the chick are interspersed among structural genes throughout the genome (Stumph et al., 1981, 1983). This arrangement could provide a means whereby a single sequence-specific protein could coordinate activities in different chromosomal regions.

Isolation of Sequence-Specific DNA Binding Proteins. The procedure for identifying CR1-specific DNA binding proteins which we have presented in this paper may be used in attempts to isolate proteins specific for other DNA sequences. Tissue extracts are first fractionated by ion exchange in the hope of enriching one fraction in a specific DNA binding protein, in our case a protein with a preference for CR1 DNA. Each fraction is applied to the tandemly linked specific and nonspecific DNA columns under conditions which minimize the electrostatic interactions of prime importance in nonspecific DNA–protein binding (Record et al., 1981; Jovin, 1976). The proteins are then recirculated in the presence of a gradient of decreasing ionic strength. During this procedure, sequence-specific binding proteins, P_1, will bind and distribute as $[P_1S_1]/[P_1S_2] = (K_{d,12}/K_{d,11})([S_1]/[S_2])$ where S_1 and S_2 represent specific and nonspecific binding sites, respectively, and $K_{d,11}$ and $K_{d,12}$ are the dissociation constants of P_1 for these sites. In the experiments presented here, about 99% of the total S_2 sites are located on the *E. coli* DNA column. Under conditions of DNA excess, it would be expected that if $K_{d,12}$ were 100 times higher than $K_{d,11}$, P_1 would distribute evenly between the columns. At $K_{d,12}/K_{d,11}$ ratios greater than 100, P_1 would be found preferentially bound to the smaller CR1 column. Nonspecific DNA binding proteins, P_2, will also bind to the columns. However, since the dissociation constant of these proteins for specific binding sites is equal to the constant for nonspecific binding sites, the majority will be found on the larger *E. coli* DNA column.

Initially, column eluants were examined for the presence of specific DNA binding proteins by comparing the electrophoretic patterns of *E. coli* adsorbed and CR1-adsorbed fractions on SDS gels. The advantage of screening preparations in this way is that it is rapid, and unlike nitrocellulose filter binding assays does not demand that the structural integrity of a specific binder be maintained. Once a potentially interesting fraction has been identified, steps can be taken to maintain the native conformation of proteins in future eluants prior to filter binding assays. It is particularly important that eluants be concentrated as quickly as possible, and the addition of dry Sephadex to preparations has proven quite useful in this regard. The inclusion of a protein which does not bind to DNA (e.g., bovine serum albumin) in column buffers may also help to prevent denaturation. The overall purification obtained by the procedure is equal to the purification due to ion-exchange chromatography plus the purification due to the elimination of non-DNA binders plus the purification due to the partitioning of proteins between DNA columns.

There are a number of modifications which can be introduced into the use of tandemly linked DNA columns that might prove useful in attempts to purify other sequence-specific proteins. We have used column packing made by linking DNA to Sephadex with ultraviolet light mainly because such columns can be easily prepared and have very good flow properties. However, a number of other types of support can be used and may have dramatic effects on purifications obtained [for reviews, see Allfrey & Inoue (1978) and Stein (1978)]. Allfrey et al. have prepared DNA–Sephadex columns using carbodiimide under conditions which favor the formation of linkages between dextran hydroxyls and monoesterified phosphates [Allfrey et al., 1973; see also Gilham (1968) and Weissbach & Poonian (1974)]. This method of linking DNA to Sephadex has the advantage of avoiding damage to DNA caused by ultraviolet light.

If a particular segment of DNA is suspected to contain specific binding sites, in certain instances it may be possible to clone a concatamer of that segment. Columns prepared from such concatamers should have improved capacities for specific binding proteins. We have recently used concatamer DNA columns to isolate proteins that appear to be specific for sequences immediately 5′ to the ovalbumin gene. Improved purifications may also be achieved by changing the type of salt used in the columns. It has been shown that the conformation of DNA is dependent not only on ionic strength but also on the nature of the salt present (Anderson & Bauer, 1978; Hanlon et al., 1975; Chan et al., 1979). Since specific binding is known, at least in some instances (Record et al., 1977; Jovin, 1976), to depend primarily on hydrophobic interactions, chaotropic salts or hydrophobic reagents such as ethylene glycol may prove effective as eluants. Finally, if the specificity of a protein for a certain DNA is high, it may be possible to reapply proteins to columns to achieve greater purification.

Finally, the availability of this technique to isolate sequence-specific DNA binding protein will facilitate the identification and isolation of other protein molecules important for regulation of gene expression and therefore cellular differentiation. For example, the promoter DNA binding proteins, the enhancer DNA binding protein(s), RNA chain termination protein(s), or Z DNA binding protein(s) may be isolated by using this approach. Actually, in our laboratory, we already have succeeded in identifying and isolating a protein fraction which binds specifically to the 5′ flanking sequence of the ovalbumin gene, and this fraction can stimulate initiation of transcription on ovalbumin as well as globin genes (M. Sanzo et al., unpublished results).

Acknowledgments

We express our thanks to David C. Bruce, Debra D. Brown, and Serlina Robinson for excellent technical assistance.

References

Allfrey, V. G., & Inoue, A. (1978) *Methods Cell Biol. 17*, 253–270.

Allfrey, V. G., Inoue, A., Karn, J., Johnson, E. M., & Vidali, G. (1973) *Cold Spring Harbor Symp. Quant. Biol. 38*, 785–801.

Anderson, P., & Bauer, W. (1978) *Biochemistry 17*, 594–601.

Bearden, J. C., Jr. (1980) *J. Biochem. Biophys. Methods 2*, 37–47.

Benoist, C., O'Hare, K., Breathnach, R., & Chambon, P. (1980) *Nucleic Acids Res. 8*, 127–142.

Calabretta, B., Robberson, D. L., Barrera-Saldana, H. A., Lambrou, T. P., & Saunders, G. F. (1982) *Nature (London) 269*, 219–225.

Chan, A., Kilkuskie, R., & Hanlon, S. (1979) *Biochemistry 18*, 84–91.

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.

Compton, J. G., Schrader, W. T., & O'Malley, B. W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 16–20.

Davidson, E. H., Hough, B. R., Klein, W. H., & Britten, R. J. (1975) *Cell (Cambridge, Mass.) 4*, 217–238.

Davidson, E. H., Klein, W. H., & Britten, R. J. (1977) *Dev. Biol. 55*, 69-84.
Duncan, C. H., Jagadeesivaran, P., Wang, R. R. C., & Weissman, S. N. (1981) *Gene 13*, 185-196.
Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. N., Shoulder, C. C., & Proudfoot, N. J. (1980) *Cell (Cambridge, Mass.) 21*, 653-668.
Engelke, D. R., NG, S. Y., Shastry, B. S., & Roeder, R. G. (1980) *Cell (Cambridge, Mass.) 19*, 717-728.
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.
Gilham, P. T. (1968) *Biochemistry 7*, 2809-2813.
Hanlon, S., Brudno, S., Wu, T. T., & Wolf, B. (1975) *Biochemistry 14*, 1648-1660.
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.
Jovin, T. M. (1976) *Annu. Rev. Biochem. 45*, 889-920.
Laemmli, U. K. (1970) *Nature (London) 227*, 680-685.
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., Tsai, M.-J., & O'Malley, B. W. (1982) *J. Biol. Chem. 257*, 1501-1507.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265-275.
Record, M. T., Jr., de Haseth, P. L., & Lohman, T. M. (1977) *Biochemistry 16*, 4791-4796.
Record, M. T., Jr., Mazur, S. J., Melancon, P., Roe, J.-H., Shaner, S. L., & Unger, L. (1981) *Annu. Rev. Biochem. 50*, 997-1024.
Schmid, C. W., & Deininger, P. L. (1975) *Cell (Cambridge, Mass.) 6*, 345-358.
Stalder, J., Larsen, A., Engel, J. D., Dolan, M., Groudine, M., & Weintraub, H. (1980) *Cell (Cambridge, Mass.) 20*, 451-460.
Stein, G. (1978) *Methods Cell Biol. 17*, 271-283.
Stump, W. E., Kristo, P., Tsai, M.-J., & O'Malley, B. W. (1981) *Nucleic Acids Res. 9*, 5383-5397.
Stumph, W. E., Baez, M., Beattie, W. G., Tsai, M.-J., & O'Malley, B. W. (1983) *Biochemistry 22*, 306-315.
Tata, J. R., Baker, B. S., & Deeley, J. V. (1980) *J. Biol. Chem. 255*, 6721-6726.
Tsai, S. Y., Tsai, M.-J., Kops, L. E., Minghetti, P. P., & O'Malley, B. W. (1981) *J. Biol. Chem. 256*, 13055-13059.
Van Arsdell, S. W., Denison, R. A., Bernstein, L. B., & Weiner, A. M. (1981) *Cell (Cambridge, Mass.) 26*, 11-17.
Weisbrod, S., Groudine, M., & Weintraub, H. (1980) *Cell (Cambridge, Mass.) 19*, 289-301.
Weissbach, A., & Poonian, M. (1974) *Methods Enzymol. 34B*, 463-475.
Wu, C., & Gilbert, W. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1577-1580.

Glycolipid Transfer Protein from Bovine Brain†

Martin Wong,‡ Rhoderick E. Brown, Y. Barenholz,§ and Thomas E. Thompson*

ABSTRACT: Glycolipid transfer protein from bovine brain has been purified partially by ammonium sulfate precipitation, CM-52 ion-exchange, and Sephadex G-75 column chromatography. Both pyrene-labeled and tritium-labeled glucocerebrosides have been used to study the kinetics of protein-mediated transfer between donor and acceptor vesicles. Protein accelerates glucocerebroside transfer but does not accelerate phospholipid transfer. In colyophilized small sonicated vesicles (10% glucocerebroside, 90% 1-palmitoyl-2-oleoylphosphatidylcholine) about two-thirds of the glycolipid is transferred in 2 h and the remaining one-third does not transfer (up to 5 h). For donor and acceptor vesicles made of dipalmitoylphosphatidylcholine or 1-palmitoyl-2-oleoylphosphatidylcholine, glucocerebroside (10% in donors) is transferred rapidly only when both the donor and acceptor matrix phospholipids are in the liquid-crystalline state. If either donor or acceptor vesicles are in the gel state, transfer protein mediated transfer is much reduced. The amount of transfer protein bound specifically to glucocerebroside-containing vesicles is nearly equal above and below the matrix phospholipid phase transition temperature. Bound protein transfers glucocerebroside upon addition of acceptor vesicles.

Glycolipids are components of biological membranes located primarily on the extracellular surface (Steck & Dawson, 1974; Gamberg & Hakomori, 1975; Moss et al., 1977; Critchley et al., 1981). They are thought to be receptors for lectins, hormones, and toxins (Fishman & Brady, 1976; Mullin et al., 1976; Hanson et al., 1977).

Small amounts of glycolipids have a relatively large effect on the physical properties of phospholipid membranes (Correa-Freire et al., 1979; Barenholz et al., 1983). The lateral organization of glycolipids in phospholipid membranes has been studied by using freeze–etch electron microscopy. Glycolipids have been localized in phospholipid bilayers and in biological membranes by attaching electron dense ferritin or colloidal gold particles to an affinity label (Tillack et al., 1982, 1983). These studies suggest that glycolipids are laterally organized in domains that change in size and geometry with temperature. Further evidence comes from measurements in which the spontaneous intervesicular transfer rates of glycosphingolipids in several systems have been shown to be very

† From the Department of Biochemistry, University of Virginia School of Medicine, Charlottesville, Virginia 22908. *Received April 25, 1984.* Supported by U.S. Public Health Service Grants GM-23573 and GM-14628 and U.S.-Israel Grant BSF 2772.
‡ Present address: Abbott Laboratories, North Chicago, IL 60064.
§ Y.B. is also associated with the Department of Biochemistry, Hebrew University, Hadassah Medical School, Jerusalem, Israel.

0006-2960/84/0423-6498$01.50/0