Multisubunit erythroid complexes binding to the enhancer element of the chicken histone H5 gene

C. Gail PENNER and James R. DAVIE*

Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Manitoba, 770 Bannatyne Avenue, Winnipeg, Manitoba, Canada R3E 0W3

We identified the factor(s) that bind to the chicken erythroid-cell-specific histone H5 enhancer region which is located on the 3' end of the gene. In DNAase I footprinting and u.v. cross-linking experiments with nuclear extracts from adult chicken immature erythrocytes, we determined that the *trans*-acting factor GATA-1 was the predominating protein interacting with the histone H5 enhancer. GATA-2 and GATA-3 were not detected. In contrast, gel-mobility-shift assays and competition experiments demonstrated that several specific complexes formed with the histone H5 enhancer region. Gel-mobility-shift assays with 23 bp oligonucleotides containing the GATA-binding site (AGATAA) of the histone H5 enhancer or of the β -globin enhancer showed that the GATA sequence was sufficient for the formation of at least five complexes. Diagonal mobility-shift assays demonstrated that multisubunit complexes were forming with the GATA-1 protein. Our interpretation of the results is that GATA-1 interacts with a protein of approx. 105 kDa which, in turn, can associate with protein or protein complexes of approx. 26 kDa, 146 kDa and a protein(s) of molecular mass greater than 450 kDa. The different multisubunit complexes formed via the *trans*-acting factor GATA-1 may impart different transcriptional responses to the promoter and enhancer elements of the histone H5 and globin genes.

INTRODUCTION

Gene expression in eukaryotes is thought to be regulated by (at least) the interaction of *trans*-acting factors with *cis*-acting DNA sequences and the availability of the template to the transcription machinery. It is well established that transcriptionally active gene chromatin has an altered structure which is preferentially sensitive to DNAase I (Gross & Garrard, 1987; Locklear *et al.*, 1990). Within the DNAase I-sensitive chromatin domains there exist regions that are hypersensitive to DNAase I digestion. These DNAase I-hypersensitive (DH) sites are regions of the chromatin that lack nucleosomes and often mark the chromatin for the presence of *cis*-acting regulatory DNA sequences and *trans*-acting factors (Elgin, 1988). Interactions between *trans*-acting factors (sequence-specific DNA-binding proteins) and/or co-activators appear to be important in enhancing the activity of RNA polymerase II promoters (Lewin, 1990; Pugh & Tjian, 1990). For example, one mechanism proposed to explain enhancer action is that proteins associated with the promoter interact with those proteins bound to the enhancer, resulting in the juxtapositioning of the enhancer and promoter elements and a looping-out of the intervening sequences (Ptashne, 1986; Li *et al.*, 1991; Su *et al.*, 1991).

Histone H5 is a H1-like linker histone that is expressed only in nucleated erythrocytes of birds, fish, reptiles and amphibians (Neelin *et al.*, 1964). Levels of histone H5 protein increase during the differentiation and maturation of the chicken erythroid cell, and this increase is correlated with chromatin condensation, the shut down of replication and the repression of the transcription of most, but not all, genes [Sun *et al.* (1989) and references cited therein]. Erythroid-specific β -globin and histone H5 genes, for example, are transcriptionally active in adult chicken immature erythrocytes (Affolter *et al.*, 1987).

The chicken erythroid histone H5 gene chromatin has several DH sites located in the 5' and 3' flanking regions of the gene

(Renaud & Ruiz-Carrillo, 1986). DH site 5 maps at or near a tissue-specific upstream promoter element (Trainor & Engel, 1989; Rousseau *et al.*, 1989), whereas the DH site 7 maps with a 3' enhancer element (Trainor *et al.*, 1987; Rousseau *et al.*, 1989). Two positive elements of the histone H5 promoter are a GC box and an upstream activating sequence (UPE) which has strong sequence similarity to the H4 gene subtype-specific element (Rousseau *et al.*, 1989). The GC box appears to bind a Sp1-like protein, while the UPE has the potential to bind transcription factor H4TF2. The histone H5 3' enhancer element possesses the sequence AGATAA that is recognized by the erythroid protein factor GATA (Rousseau *et al.*, 1989). The GATA motif (WGA-TAR: W = A or T; R = A or G) is also found in the promoters and enhancers of globin and other erythroid-specific genes (Evans *et al.*, 1988, 1990; Perkins *et al.*, 1989; Evans & Felsenfeld, 1991). Both the histone H5 and β -globin enhancers also have similar 34 bp DNA sequences which differ by seven nucleotides and a single nucleotide gap (Trainor *et al.*, 1987).

Yamamoto *et al.* (1990) recently reported the existence of multiple GATA-binding proteins: GATA-1, GATA-2 and GATA-3. All three proteins can serve as positive *trans*-acting factors. The three GATA-binding proteins are highly conserved within the DNA-binding domain, which consists of two putative zinc fingers, but they are distinct from each other outside of this region. GATA-1 (previously called Eryf1, NF-E1, GF-1, EF-1 and EF-ra; Orkin, 1990) is a DNA-binding protein that is expressed at all stages of erythropoiesis. Furthermore, GATA-1 has a major role in erythroid development (Pevny *et al.*, 1991). In avian adult reticulocytes, GATA-1 mRNA is abundant, whereas GATA-2 and GATA-3 transcripts are present at low levels (Yamamoto *et al.*, 1990).

Here we identified nuclear proteins isolated from adult chicken immature erythrocytes that interact with the histone H5 enhancer element. Gel-mobility-shift analysis and competition experiments revealed that several specific complexes were formed. However,

Abbreviations used: GATA-1, 2 and 3, proteins (factors) binding the nucleotide motif GATA; DH, DNAase I-hypersensitive; PMSF, phenylmethanesulphonyl fluoride; UPE, upstream activating element.

* To whom correspondence should be sent.

DNAase I protection assays displayed only one strong DNA-binding activity, corresponding to the GATA protein. Proteins binding to the 34 bp segment were not detected. U.v. cross-linking of protein-DNA complexes presented evidence that the DNA-binding protein was GATA-1. Diagonal gel-mobility-shift assays indicated that proteins of approx. 105 kDa, 26 kDa, 146 kDa and a protein(s) with a molecular mass greater than 450 kDa interacted directly or indirectly with GATA-1, forming multisubunit complexes. The ability of GATA-1 to interact with an array of other proteins expands the regulatory potential of target promoters and enhancers of the erythroid-specific genes.

MATERIALS AND METHODS

Preparation of nuclear extracts

Adult White Leghorn chickens were made anaemic by injections of phenylhydrazine hydrochloride. Blood was collected and nuclei were isolated as described previously (Ridsdale & Davie, 1987). Nuclei were extracted with RSB [10 mM-Tris/HCl (pH 7.5)/10 mM-NaCl/3 mM-MgCl₂/10 mM-sodium butyrate/1 mM-phenylmethanesulphonyl fluoride (PMSF)] containing 0.3 M-NaCl as described by Emerson & Felsenfeld (1984). The nuclear extract was made 0.35 g/ml in (NH₄)₂SO₄, and the precipitate was resuspended in Buffer D [20 mM-Hepes pH 7.9/20% (v/v) glycerol/0.1 M-KCl/0.2 mM-EDTA/0.5 mM-PMSF/0.5 mM-dithiothreitol] and dialysed against buffer D. The sample was cleared by centrifugation and the supernatant was stored at -75 °C. Nuclear extracts were also prepared by the method described by Dignam *et al.* (1983).

Oligonucleotides, clone construction and fragment labelling

Two oligonucleotides with the sequences 5'-GAGGCTGG-AGATAACAGTGC GG-3' (H5-GATA-for) and 5'-GCCGCA-CTGTTATCTCCAGCCT-3' (H5-GATA-rev) were synthesized. These two complementary sequences were annealed to form double-stranded DNA (H5-GATA) covering the region from +1045 to +1067 downstream of the start of transcription of the chicken H5 gene and which covers the GATA-protein binding site (Rousseau *et al.*, 1989). Oligonucleotides (coding strand 5'-CAGGTTGCAGATAAACATTTTG-3' + non-coding strand 5'-GCAAAATGTTTATCTGCAACCT-3') containing the GATA sequence of the β -globin gene enhancer located from +1888 to +1910 downstream of the start of transcription (Evans *et al.*, 1988) were synthesized and annealed (β -globin-GATA).

DNA containing either the histone H5 3' enhancer region (H5-3'E) from *Ddel* (+893) to *Xma*III (+1064), the histone H5 5' promoter (H5-5'P) from *Sau*3AI (-350) to *Sau*3AI (+212), or the β -globin 3' enhancer region (β -globin-3'E) from *Pvu*II (+1877) to *Pvu*II (+2178) was subcloned into pGEM-3. The insert was excised with appropriate restriction enzymes from the multiple cloning site and purified. Therefore the insert contained parts of the pGEM multiple cloning site at either end. For some gel-mobility-shift assays, the insert DNA (either a 203 bp *Xba*I-*Eco*RI fragment or a 240 bp *Hind*III-*Kpn*I fragment) containing H5-3'E was further digested with either *Mva*I or *Hha*I. DNA was labelled selectively at the 3' end of either the non-template or template strand using Klenow enzyme and [α -³²P]dATP or [α -³²P]dCTP whichever was appropriate to the end that was to be filled.

DNA-binding assays

Labelled DNA (approx. 10 pmol) was incubated on ice for 30 min with 5-15 μ g of extracted nuclear protein in the presence of 0.36 μ g of poly (dl·dC)/ μ l in 10 mM-Tris/HCl (pH 7.5)/

40 mM-NaCl/1 mM-EDTA/1 mM- β -mercaptoethanol/4% (w/v) glycerol in a total volume of 10 μ l (Schindler & Cashmore, 1990). The protein was added to the incubation last. For competition assays the appropriate amount of the competitor DNA was added to the incubation mix before the addition of protein. After incubation, the samples were electrophoresed through 4% acrylamide (acrylamide/bisacrylamide, 20:1) (w/w) in 0.5 $\times$ TBE (1 $\times$ TBE is 89 mM-Tris/89 mM-borate/2 mM-EDTA) at 12 V $\cdot$ cm⁻¹ at 4 °C (the gel-mobility-shift assay). After electrophoresis the gels were dried and autoradiographed.

The molecular-mass-determination experiments (Bading, 1988) and diagonal gel-mobility-shift assay (Schaufele *et al.*, 1990) were done as described in the references cited, except that the protein-DNA binding conditions described above were used. The molecular-mass determinations were done with 5%-polyacrylamide gels as described by Bading (1988). The molecular-mass determination of the DNA-bound protein is based on the assumptions that: (1) the reduced ion mobility of the protein-DNA complex is not influenced by protein structure; (2) the charge of the protein is negligible compared with the charge of the DNA fragment; and (3) the DNA is not bent by the protein. The linear range of this method is 0-450 kDa, with an error of about 10%. For the diagonal mobility-shift assay, 0.5 $\times$ TBE was used for the electrophoresis buffer. The DNAase I protection (footprinting) assay was done as described previously (Emerson *et al.*, 1985). In order to determine the DNAase I footprint for specific complexes, the DNA-protein mixture was treated with DNAase I (Promega). The reaction was stopped with 20 mM-EDTA and the reaction mixture was then electrophoresed in the usual manner for a gel-shift assay. The individual bands were excised and the DNA was extracted into 20 mM-Hepes (pH 7.4)/0.5 M-NaCl/0.2% (w/v) SDS/1 mM-EDTA by overnight incubation at 37 °C. The DNA was purified by phenol and chloroform extractions, ethanol precipitation and then electrophoresed on a DNA sequencing gel.

Cross-linking of protein-DNA complexes

The probe was prepared by hybridizing the oligonucleotide H5 GATA-for to a 8 bp complimentary primer (5'-CCGCACTG-3') or by hybridizing the oligonucleotide H5 GATA-rev to a 8 bp complimentary primer (5'-AGGCTGGA-3'). The oligonucleotide was made completely double-stranded by incubation with the Klenow polymerase in the presence of dCTP, dGTP, dTTP and [α -³²P]dATP. The labelled probe was allowed to associate with nuclear proteins under standard assay conditions. Portions (20 μ l) were cross-linked to a total of 7 J/cm² in a UV Stratalinker 2400 (254 nm) apparatus. After cross-linking, a portion of the sample was electrophoresed under standard gel-retardation conditions. The remainder of the irradiated sample was incubated with DNAase I and micrococcal nuclease as described by Chodosh (1988). The sample was then electrophoresed in SDS/10% (or 12%) polyacrylamide gels as described by Davie (1982).

RESULTS

Several protein complexes are formed with the histone H5 enhancer DNA fragment, but only one DNA-binding activity is detected

Gel-mobility-shift assays were used to determine what nuclear proteins bind to the histone H5 enhancer. A 203 bp *Xba*I-*Eco*RI DNA fragment containing the histone H5 3' enhancer (H5-3'E) activity (Trainor *et al.*, 1987; Rousseau *et al.*, 1989) was incubated with a nuclear extract from adult chicken immature erythrocytes. Fig. 1 shows that nine or more shifted bands were generated, indicating the presence of several binding activities. Complexes of greater size than C8 did not always separate well on the gels

Fig. 1. Specific interactions between the histone H5 enhancer region and factors from adult chicken immature erythrocyte nuclear extracts

A 203 bp *XbaI-EcoRI* (H5-3'E) end-labelled DNA fragment was incubated in the presence or absence of nuclear proteins extracted from adult chicken immature erythrocytes. The left lane in each panel did not contain any protein and the other lanes each contained 10 μ g of protein and 1 ng of labelled DNA. The second lane from the left does not contain any competitor DNA, and the next lanes contain increasing amounts of competitor DNA as indicated below each lane. Competitor DNA amounts are in molar excess for a 203 bp *XbaI-EcoRI* (H5-3'E) DNA fragment (left panel) and excess by weight for salmon sperm DNA (right panel). F is the free DNA fragment, and C1–C10 are the protein–DNA complexes.

Fig. 2. DNAase I footprint analysis of the histone H5 enhancer regions

Two 3' end-labelled restriction fragments, a 203 bp *XbaI-EcoRI* [labelled template (T) strand] and a 240 bp *HindIII-KpnI* [labelled non-template (NT) strand], containing the histone H5 enhancer region were incubated with (+) or without (–) proteins (P) (10 μ g of protein/ng of DNA) extracted from adult-chicken immature-erythrocyte nuclei. The DNAase I-resistant region generated by the GATA protein (Rousseau *et al.*, 1989) is shown. The arrowheads indicate the DH sites of the GATA–DNA complex.

and thus they were not studied in detail. Complex C3, which does not appear in the gel-shift pattern shown in Fig. 1, was thought to be a proteolytic degradation product of one of the higher-molecular-mass complexes. Complex C3 was not observed in freshly prepared nuclear extracts, and the abundance of complex C3 increased as the extracts aged. All of the complexes showed

specific binding to the 203 bp *XbaI-EcoRI* H5-3'E DNA fragments. A 100-fold molar excess of the unlabelled 203 bp H5-3'E DNA fragment competed effectively for all of these complexes, whereas the non-specific competitor salmon sperm DNA did not effectively compete, even at high levels. The order of addition of protein to the binding mixture also did not alter the number of complexes observed.

The locations of specific protein–DNA interactions in the histone H5 enhancer region were determined by the DNAase I protection (footprinting) assay. Although several shifted bands were observed in the gel-mobility-shift assay, only one footprinting activity was detected (Fig. 2). The region protected had the sequence AGATAA, which is recognized by the DNA-binding protein, GATA. A distinguishing feature of GATA binding to its recognition sequence in the histone H5 enhancer was the generation of DH sites on the non-template strand (see the arrows in Fig. 2). A similar observation was reported by Rousseau *et al.* (1989). Protection of the 34 bp DNA segment, which is similar to a sequence found in the β -globin enhancer (Trainor *et al.*, 1987), was not apparent.

The majority of the complexes are associated with a DNA segment containing the GATA-binding site

To determine where the complexes were forming in the histone H5 enhancer region, gel-mobility-shift assays were performed with various segments of the histone H5 enhancer. By comparing the intensity of the complexes formed with the DNA fragment with different nuclear extracts, as well as their position in the gel pattern, we were able to identify the complexes formed with the DNA fragments. Fig. 3(a) shows that, with two different nuclear extracts, complexes C1, C2, a weak C3 and C4–C8 were formed with the 203 bp *XbaI-EcoRI* H5-3'E DNA fragment. Complexes C3, C7 and C9 were associated with the 207 bp *HindIII-HhaI* (Fig. 3b) and 146 bp *HindIII-MvaI* (Fig. 3e) DNA fragments, whereas complexes C1, C2, C4–C6 and C8 were located on the 41 bp *HhaI-EcoRI* (Fig. 3c) and 102 bp *MvaI-EcoRI* (Fig. 3f) DNA fragments, both of which contained the GATA sequence. Since the same complexes (C3, C7 and C9) were formed with the 207 bp *HindIII-HhaI* and 146 bp *HindIII-MvaI* fragments, this demonstrates that none of the complexes were formed on the DNA located between the *MvaI* and *HhaI* restriction sites. This suggests that nuclear factors were not binding to the 34 bp DNA segment of the histone H5 3' enhancer. A 23 bp oligonucleotide (H5-GATA) containing the sequence AGATAA was also used in the gel-mobility-shift assay. Fig. 3(d) shows that complexes C1, C2, C4–C6 and C8 formed with this oligonucleotide. Thus the GATA DNA-binding sequence (AGATAA) and as little as 8 bp on either side of it were sufficient to generate complexes C1, C2, C4–C6 and C8. These complexes were independent of C3, C7 and C9, since the two groups of complexes (C1, C2, C4–C6 and C8) and (C3, C7 and C9) were seen with the two separate regions of DNA.

Multiple factor complexes are formed with the GATA-binding site

The results presented in Fig. 3 demonstrated that several complexes were formed with an oligonucleotide containing the GATA-binding site AGATAA. We determined whether DNA fragments with the GATA sequence would compete for the complexes C1, C2, C4–C6 and C8. Fig. 4 shows that the 23 bp H5-GATA oligonucleotide selectively and effectively competed for complexes C1, C2, C4–C6 and C8. Complexes C7, C9 and C10 were not affected. A 328 bp *XbaI-EcoRI* DNA fragment that harbours the β -globin 3' enhancer and contains two binding sites for GATA (AGATAA and TGATAG) also selectively and effectively competed for the same complexes as the H5-GATA

Fig. 3. Multiple gel-shifted complexes are formed with the GATA region of the histone H5 enhancer

The lower part of this Figure shows the DNA fragments that were used in the gel-mobility-shift assays. Restriction-endonuclease sites (*HindIII*, *XbaI*, *KpnI* and *EcoRI*) in the multiple cloning region of pGEM-3 and restriction sites (*MvaI* and *HhaI*) in the histone H5 enhancer region are shown. The arrows show the position of the *DdeI* (+893) to *XmaIII* (+1064) restriction fragment containing the histone H5 3' enhancer region (H5-3'E). The restriction-endonuclease site that was labelled is indicated with an asterisk. Note that the 41 bp *HhaI*-*EcoRI* fragment (c) was contaminated with a longer fragment. The 23 bp oligonucleotide H5-GATA was synthesized as described in the Materials and methods section. The DNA fragments used for each of the gel-mobility-shift assays are indicated at the bottom of each panel. In each case the left lane is control DNA with no added protein. The middle and right lanes show the shifted DNA using two different nuclear extracts (10 µg of protein/ng of DNA) of adult chicken immature erythrocytes. F is the free DNA fragment, and C1-C9 are the protein-DNA complexes. Gel-mobility-shift assays were performed with six different extracts and each of the DNA fragments shown. By comparing the intensities of various complexes formed with each DNA fragment and nuclear extract, we were able to assign which of the complexes were formed with each of the DNA fragments.

Fig. 4. DNA fragments containing the GATA sequence compete for complexes C1, C2, C4-C6 and C8

The left lane in each panel did not contain any protein and the other lanes each contained 10 µg of nuclear extracted protein from chicken immature erythrocytes and 1 ng of DNA (3' end-labelled 203 bp *XbaI*-*EcoRI* DNA fragment with H5-3'E). The second lane from the left does not contain any competitor DNA, and the next lanes contain increasing amounts of competitor DNA as indicated in molar excess. Competitor DNA was the synthesized 23 bp oligonucleotide H5-GATA, a 328 bp *XbaI*-*EcoRI* restriction fragment containing the β-globin-3'E and a 528 bp *XbaI*-*EcoRI* restriction fragment containing the H5-5'P. F is the free DNA fragment, and C1-C10 are the protein-DNA complexes.

oligonucleotide. It should be noted that the 328 bp β-globin enhancer fragment contained the DNA segment that showed sequence similarity to the 34 bp region of the histone H5 enhancer. However, this DNA fragment did not effectively compete with complexes C7 and C9. A 528 bp *XbaI*-*EcoRI* DNA fragment containing the histone H5 5' promoter region (H5-5'P) was also used in the competition experiments. This DNA fragment does not contain a GATA consensus site. Fig. 4 shows

Fig. 5. Complexes C1, C2, C4-C6 and C8 form with oligonucleotides containing the GATA-binding site

In each of the panels the left lane does not contain protein and the right lane has nuclear extracted protein from chicken immature erythrocytes (10 µg of protein/ng of DNA) added to the assay. The 23 bp oligonucleotides with the GATA sequence were synthesized as described in the Materials and methods section. The oligonucleotides contain 23 bp segments of the histone H5 enhancer (H5 GATA) or β-globin enhancer (β-globin GATA) regions.

that the higher-molecular-mass complexes C8 and C9, but not C1-C7, were competed for at low levels of H5-5'P DNA.

To investigate further whether the nucleotides flanking the GATA-binding site influenced the generation of multiple complexes, another 23 bp GATA-containing oligonucleotide with the sequence of one of the GATA-binding sites in the β-globin 3' enhancer was synthesized. Fig. 5 shows the results of gel-mobility-shift experiments with labelled oligonucleotides H5-GATA (template strand 5'-GAGGCTGGAGATAACAGTGCGG-3') and β-globin-GATA (template strand 5'-CAGGTTGCAGATAAA

Fig. 6. U.v. cross-linking of the protein-DNA complexes reveals that the *trans*-acting factor GATA-1 is bound to the histone H5 enhancer region

The internally labelled 23 bp oligonucleotide (H5-GATA) containing the GATA sequence of the histone H5 enhancer region was incubated with the nuclear-extracted protein of chicken immature erythrocytes. The incubation mixture was (+) or was not (-) irradiated with u.v. light and analysed by the gel-mobility-shift assay (shown in the left panel). F is the free DNA fragment, and C1, C2, C4-C6 and C8 are the protein-DNA complexes. The protein-DNA complexes were digested with DNAase I and micrococcal nuclease, and the affinity-labelled proteins were analysed on 10%- or 12%-polyacrylamide/SDS gels. Molecular-mass (*M*) markers are Bio-Rad Low Molecular Weight Markers. Molecular masses of the proteins were calculated from a log(molecular mass)-versus-migration plot.

CATTTTG-3'). The same complexes were formed with both oligonucleotides, demonstrating that the DNA sequence surrounding the consensus GATA sequence was not important in forming the complexes C1, C2, C4-C6 and C8.

The DNA-binding location of each of the complexes C1, C2, C4-C6 and C8 was determined by DNAase I footprinting. The non-template strand of the 240 bp *Hind*III-*Kpn*I H5-3'E DNA fragment was end-labelled and incubated with a nuclear extract from adult chicken immature erythrocytes and digested with DNAase I. After isolation of each complex by electrophoresis on a preparative gel, the DNA of each complex was isolated and run on a sequencing gel. Each of the complexes yielded (results not shown) the characteristic GATA protein DNA footprint shown in Fig. 2. These results provide evidence that the only protein in the complexes that is binding to DNA is the GATA protein, suggesting that these complexes are generated by the interaction of other proteins with GATA.

The GATA-1 DNA-binding protein is common to at least six complexes

Recently, Yamamoto *et al.* (1990) demonstrated that there was a family of GATA-binding proteins, GATA-1, -2, and -3, which have molecular masses of about 39.5, 56, and 55 kDa on SDS-containing gels. Of these three GATA-binding proteins, GATA-1 was the most abundant in immature erythrocytes. It was conceivable that some of these complexes arose from the interaction of these different GATA-binding proteins with the GATA-binding site. To elucidate which GATA-binding protein was interacting with the histone H5 enhancer region, a uniformly labelled H5-GATA oligonucleotide was incubated with a nuclear extract isolated from chicken immature erythrocytes, and then the sample was irradiated with u.v. light (254 nm). After irradiation, the sample was incubated with DNAase I and micrococcal nuclease, and the affinity-labelled proteins were identified on SDS-containing gels. Fig. 6 shows that proteins of approximately 29 and 43 kDa were labelled. The results were the same for an H5-GATA oligonucleotide containing a labelled

Fig. 7. Diagonal gel electrophoresis of GATA-1 complexes

The labelled 23 bp oligonucleotide H5-GATA was incubated with nuclear extracted protein of chicken immature erythrocytes. The diagonal gel-mobility-shift assay was done as described in the Materials and methods section. As a reference, gel slices from the first and second gels are placed on the top and left side of the diagonal gel shift pattern respectively. (a) and (b) are autoradiograms of the same two-dimensional gel pattern, except that the autoradiogram in (b) was exposed for a longer time than in (a). The arrowheads show the locations of the C2 complexes that had dissociated from the complexes C4-C6 and C8. (c) is a schematic representation of the gel pattern shown in (a) and (b). The symbols represent individual proteins present in each complex. The symbol Δ represents GATA-1, and 'F' shows the DNA fragment. Other symbols are explained in the text.

template or non-template strand. In a similar experiment using purified GATA-1, Perkins *et al.* (1989) observed that a protein of about 40 kDa was labelled. The 29 kDa labelled protein was likely a proteolytic product of GATA-1 that is sensitive to degradation (Evans & Felsenfeld, 1989; Perkins *et al.*, 1989; Evans & Felsenfeld, 1991). Proteins of molecular mass greater than the 43 kDa GATA-1 protein were not observed, even after longer exposure of the autoradiogram. This result suggested that GATA-1 was the major DNA-binding protein of the GATA-related complexes.

The approximate molecular mass of the DNA-bound proteins in the complexes was determined by the method described by Bading (1988). This method is subject to a number of assumptions, including that the protein does not bend DNA and the charge of the protein is negligible compared with the charge of the DNA fragment. Thus this method provides a rough estimate of the molecular mass of the DNA-binding protein. The end-labelled 240 bp *Hind*III-*Kpn*I H5-3'E DNA was incubated with a nuclear extract isolated from chicken immature erythrocytes, and the complexes were resolved on a 5%-polyacrylamide gel (results not shown). The migration of the complex relative to that

of the free DNA was used to calculate the molecular mass of the bound protein(s) (Bading, 1988). The protein in complex C2 had a molecular mass of approx. 45 kDa, which is consistent with the bound protein being GATA-1. The molecular mass of the DNA-binding protein of complex C1 was approx. 31 kDa. The protein(s) of complexes C4, C5, and C6 had molecular masses of approx. 150 kDa, 176 kDa and 322 kDa respectively. The molecular masses of the DNA-binding proteins of complexes C7, C8, and C9 were greater than 450 kDa. Complexes with bound proteins of approx. 56 kDa (GATA-2) or 55 kDa (GATA-3) were not detected.

The novel diagonal gel-mobility-shift assay described by Schaufele *et al.* (1990) was used to determine the protein components of the GATA-related complexes. The labelled 23 bp H5-GATA oligonucleotide was incubated with a nuclear extract isolated from chicken immature erythrocytes, and the sample was electrophoresed for such a time that complex C2 was running off of the non-denaturing gel. A lane from the first gel was placed horizontally on to a second non-denaturing gel and electrophoresed. During electrophoresis in the second gel, a portion of each of the complexes dissociated as it entered the second dimension. Complexes remaining on the diagonal were intact, whereas those below the diagonal had dissociated. It should be noted that only those dissociated complexes containing a DNA-binding protein will be detected. From the pattern of dissociation it was possible to determine which complexes were part of a larger complex. Fig. 7 shows the results of this experiment as well as schematic representation (Fig. 7c) of the gels shown in Figs. 7(a) and 7(b). The diagonal gel-mobility-shift assay was also done with the labelled 203 bp *XbaI-EcoRI* H5-3'E DNA. The gel-pattern results indicate that the pattern of dissociation of the complexes was as follows: C8 → C4 and C2; C6 → C5 and C2; C5 → C4 and C2; C4 → C2. Complex C2, the GATA-1-DNA complex, was a common component of complexes C4-C6 and C8 (indicated by arrowheads in Figs. 7a and 7b). None of the complexes dissociated to C1, the complex containing the GATA-1 proteolytic product. Our results suggest that C4 was also common to the larger complexes C5, C6 and C8. Our interpretation of the results shown in Fig. 7(c) was that there were at least four proteins in addition to GATA-1 involved in forming the complexes. GATA-1 (symbol Δ) was the only DNA-binding protein. Protein A (symbol ∇) was interacting with GATA-1, forming complex C4. Protein B (symbol $\circ$) interacted with protein A to generate complex C5. Protein C (symbol $\bullet$) bound to protein A and/or B to form complex C6, whereas protein(s) D (symbol $\ominus$) associated with protein A generating complex C8.

DISCUSSION

The chicken histone H5 gene has an erythroid-specific enhancer located at its 3' end. The enhancer activity has been shown to be present in the region between +1042 to +1185 by Rousseau *et al.* (1989) and between +851 to +1064 by Trainor *et al.* (1987). We used a DNA fragment which spans the region +893 to +1064 to detect adult immature erythrocyte nuclear DNA-binding proteins that interacted with the enhancer DNA element. The results of the DNAase I footprinting (or protection) assays demonstrated that the erythroid protein binding most avidly to the histone H5 enhancer DNA was the *trans*-acting factor GATA. This result was in agreement with the observations of Rousseau *et al.* (1989), who used a DNA fragment encompassing the region +980 to +1087. These investigators detected only one DNA-binding activity by DNAase I protection, binding interference and gel-mobility-shift assays.

Trainor *et al.* (1987) noted that a 34 bp sequence with seven nucleotide mismatches and a single base-pair gap was common

to the histone H5 and β -globin enhancers. A factor isolated from adult erythrocyte nuclei binds to this region of the β -globin enhancer (Emerson *et al.* 1987). The results of our gel-mobility-shift assays, DNAase I protection assays and competition studies indicated that this factor was unable to bind the 34 bp segment of the histone H5 enhancer. It should be noted that the factor protects a 18 bp region of the β -globin enhancer (Emerson *et al.*, 1987) and that the sequence of this 18 bp segment differs by five mismatches in the histone H5 enhancer. This suggests that these changes in the sequence of this region are sufficient to prevent the factor from binding to the histone H5 enhancer.

In adult chicken immature erythrocytes there are several proteins that recognize the general consensus motif WGATAR. The levels of the transcripts encoding these proteins belonging to the GATA family indicate that, in adult immature erythroid cells, GATA-1 is the most abundant, followed by GATA-2, then GATA-3 (Yamamoto *et al.*, 1990). The results of the u.v.-cross-linking experiments presented evidence that GATA-1 was the only member of the GATA-binding proteins that was affinity-labelled, indicating that GATA-1 is the major protein binding to the GATA sequence of the histone H5 enhancer region. Furthermore, complexes containing proteins of 55–70 kDa were not detected in the gel-mobility-shift assays.

In apparent contradiction to the results of the DNAase I protection assay, we observed that the immature erythroid nuclear proteins could form several specific complexes with the histone H5 enhancer DNA. By using a 23 bp oligonucleotide containing the GATA sequence in competition experiments, gel-mobility-shift assays and diagonal gel-mobility-shift assays, we demonstrated that several of these complexes were formed by proteins associating with the GATA-1 *trans*-acting factor. It is unlikely that these complexes were formed by GATA-1 dimers, since complexes of approx. 80–90 kDa were not seen. Moreover, Evans & Felsenfeld (1989) reported that chicken GATA-1 did not dimerize. The estimated molecular masses of these GATA-related complexes and the results of the diagonal gel-mobility-shift assay suggest the following arrangement of proteins in the complexes. A protein of approx. 105 kDa associates with GATA-1, forming complex C4. The larger complexes are generated by other proteins associating with the 105 kDa protein as follows: complex C5, a 26 kDa protein; complex C6, a 26 kDa and 146 kDa protein; complex C8, a protein(s) with molecular mass greater than 450 kDa. Note that the molecular masses of these proteins are rough estimates. The proteolytic form of GATA-1, which forms complex C1, obviously has the ability to bind its recognition sequence. However, we did not detect the presence of the proteolytic GATA-1 form in the complexes C4–C6 and C8, suggesting that the GATA-1 binding proteins do not interact with the DNA-binding domain of GATA-1.

In transient transfection studies using histone H5 expressing cells, Rousseau *et al.* (1989) demonstrated that to activate transcription the histone H5 3' enhancer required the presence of histone H5 promoter elements (–38 to –90). One mechanism proposed for the β -globin 3' enhancer activity is that this DNA element is juxtapositioned next to the β -globin promoter elements through protein–protein interactions (Choi & Engel, 1988). A similar model would be applicable for the histone H5 gene. Thus proteins interacting with the histone H5 enhancer may interact directly or indirectly via 'bridging' proteins or co-activators with proteins interacting with the promoter elements localized between –38 and –90. This region contains two positive elements: a GC-box that binds the chicken homologue of Sp1 and a UPE, a sequence common to H1 genes coding for replacement variants and similar to the H4 subtype-specific element recognized by transcription factor H4TF2 (Rousseau *et al.*, 1989). It is conceivable that the proteins forming multisubunit complexes with

GATA-1, which is bound to the enhancer element, and/or with proteins bound to the segment (+893 to +980), interact with either Sp1 and/or H4TF2. In this regard it is noteworthy that the histone H5 5' promoter fragment competed effectively for GATA-related complex C8 and non-GATA-related complex C9 (Fig. 4).

GATA-1 binding sites are distributed throughout the promoters and enhancers of the chicken globin genes (Evans *et al.*, 1988, 1990). We have demonstrated that a segment of the β -globin gene enhancer (+1888 to +1910) harbouring the GATA-1 recognition sequence formed multisubunit complexes C4-C6 and C8. The results presented here suggest that GATA-1 interacts with a 105 kDa protein which, in turn, interacts with an array of other proteins. Conceivably the different multisubunit complexes formed via the *trans*-acting factor GATA-1 may impart different transcriptional responses to the regulatory element. Thus the proteins binding to GATA-1 may play a role in modulating the activity of not only the histone H5 gene enhancer but also the promoter and enhancer elements of the globin genes.

This project was supported by a grant from the Medical Research Council of Canada and by a Medical Research Council Scientist award to J.R.D.

REFERENCES

- Affolter, M., Cote, J., Renaud, J. & Ruiz-Carrillo, A. (1987) *Mol. Cell. Biol.* **7**, 3663-3762
- Bading, H. (1988) *Nucleic Acids Res.* **16**, 5241-5248
- Chodosh, L. A. (1988) in *Current Protocols in Molecular Biology* (Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds.), pp. 12.5.1-12.5.8, Greene Publishing and Wiley-Interscience, New York
- Choi, O.-R. B. & Engel, J. D. (1988) *Cell (Cambridge, Mass.)* **55**, 17-26
- Davie, J. R. (1982) *Anal. Biochem.* **120**, 276-281
- Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475-1489
- Elgin, S. C. R. (1988) *J. Biol. Chem.* **263**, 19259-19262
- Emerson, B. M. & Felsenfeld, G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 95-99
- Emerson, B. M., Lewis, C. D. & Felsenfeld, G. (1985) *Cell (Cambridge, Mass.)* **41**, 21-30
- Emerson, B. M., Nickol, J. M., Jackson, P. D. & Felsenfeld, G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4786-4790
- Evans, T. & Felsenfeld, G. (1989) *Cell (Cambridge, Mass.)* **58**, 877-885
- Evans, T. & Felsenfeld, G. (1991) *Mol. Cell. Biol.* **11**, 843-853
- Evans, T., Reitman, M. & Felsenfeld, G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5976-5980
- Evans, T., Felsenfeld, G. & Reitman, M. (1990) *Annu. Rev. Cell Biol.* **6**, 95-124
- Gross, D. S. & Garrard, W. T. (1987) *Trends Biochem. Sci.* **12**, 293-297
- Lewin, B. (1990) *Cell (Cambridge, Mass.)* **61**, 1161-1164
- Li, R., Knight, J. D., Jackson, S. P., Tjian, R. & Botchan, M. R. (1991) *Cell (Cambridge, Mass.)* **65**, 493-505
- Locklear, Jr., L., Ridsdale, J. A., Bazett-Jones, D. P. & Davie, J. R. (1990) *Nucleic Acids Res.* **18**, 7015-7024
- Neelin, J. M., Callahan, P. X., Lamb, D. C. & Murray, K. (1964) *Can. J. Biochem.* **42**, 1743-1752
- Orkin, S. H. (1990) *Cell (Cambridge, Mass.)* **63**, 665-672
- Perkins, N. D., Nicolas, R. H., Plumb, M. A. & Goodwin, G. H. (1989) *Nucleic Acids Res.* **17**, 1299-1314
- Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S.-F., D'Agati, V., Orkin, S. H. & Costantini, F. (1991) *Nature (London)* **349**, 257-260
- Ptashne, M. (1986) *Nature (London)* **322**, 697-701
- Pugh, B. F. & Tjian, R. (1990) *Cell (Cambridge, Mass.)* **61**, 1187-1197
- Renaud, J. & Ruiz-Carrillo, A. (1986) *J. Mol. Biol.* **189**, 217-226
- Ridsdale, J. A. & Davie, J. R. (1987) *Nucleic Acids Res.* **15**, 1081-1096
- Rousseau, S., Renaud, J. & Ruiz-Carrillo, A. (1989) *Nucleic Acids Res.* **17**, 7495-7511
- Schaufele, F., Cassill, J. A., West, B. L. & Reudelhuber, T. (1990) *J. Biol. Chem.* **265**, 14592-14598
- Schindler, U. & Cashmore, A. R. (1990) *EMBO J.* **9**, 3415-3427
- Su, W., Jackson, S., Tjian, R. & Echols, H. (1991) *Genes Dev.* **5**, 820-826
- Sun, J.-M., Wiaderkiewicz, R. & Ruiz-Carrillo, A. (1989) *Science* **245**, 68-71
- Trainor, C. D. & Engel, J. D. (1989) *Mol. Cell. Biol.* **9**, 2228-2232
- Trainor, C. D., Stamler, S. J. & Engel, J. D. (1987) *Nature (London)* **328**, 827-830
- Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H. & Engel, J. D. (1990) *Genes Dev.* **4**, 1650-1662

Received 23 August 1991/24 October 1991; accepted 30 October 1991