The N-terminal POZ Domain of GAGA Mediates the Formation of Oligomers That Bind DNA with High Affinity and Specificity*

(Received for publication, February 4, 1999, and in revised form, March 15, 1999)

Maria Lluïsa Espinás, Emilio Jiménez-García, Alejandro Vaquero, Silvia Canudas, Jordi Bernués, and Fernando Azorín†

From the Departament de Biologia Molecular i Cel·lular, Institut de Biologia Molecular de Barcelona, Centre d’Investigació i Desenvolupament, Consejo Superior de Investigaciones Científicas, Jordi Girona Salgado 18-26, 08034 Barcelona, Spain

The Drosophila GAGA factor self-oligomerizes both in vivo and in vitro. GAGA oligomerization depends on the presence of the N-terminal POZ domain and the formation of dimers, tetramers, and oligomers of high stoichiometry is observed in vitro. GAGA oligomers bind DNA with high affinity and specificity. As a consequence of its multimeric character, the interaction of GAGA with DNA fragments carrying several GAGA binding sites is multivalent and of higher affinity than its interaction with fragments containing single short sites. A single GAGA oligomer is capable of binding adjacent GAGA binding sites spaced by as many as 20 base pairs. GAGA oligomers are functionally active, being transcriptionally competent in vitro. GAGA-dependent transcription activation depends strongly on the number of GAGA binding sites present in the promoter. The POZ domain is not necessary for in vitro transcription but, in its absence, no synergism is observed on increasing the number of binding sites contained within the promoter. These results are discussed in view of the distribution of GAGA binding sites that, most frequently, form clusters of relatively short sites spaced by small variable distances.

The GAGA factor is a sequence-specific DNA-binding protein, which participates in the regulation of the expression of a variety of different classes of genes in Drosophila such as many developmentally regulated genes, stress induced genes, and cell cycle regulated genes, as well as housekeeping genes (for reviews see Refs. 1 and 2). GAGA binds repeated d(GA

* This work was supported by Grants PB95-64 and PB96-812 from the Spanish Dirección General de Euseñanza Superior, the European Union (BI02-CT94-3069), and the Comissió Interdepartamental de Recerca i Innovació Tecnològica (CIRIT) of the Generalitat de Catalunya (SGR97–55). This work was carried out in the context of the Centre de Referència en Biotecnologia of the CIRIT of the Generalitat de Catalunya. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

† Recipient of doctoral fellowship from the CIRIT.

‡ Recipient of doctoral fellowship from the Dirección General de Euseñanza Superior.

¶ To whom correspondence should be addressed. Dpt. Biologia Molecular i Cel·lular, Institut de Biologia Molecular de Barcelona, Centre d’Investigació i Desenvolupament, Consejo Superior de Investigaciones Científicas, Jordi Girona Salgado 18-26, 08034 Barcelona, Spain. Tel.: 343-4006137; Fax: 343-2045904; E-mail: fambmc@cid.csic.es.

† The abbreviations used are: bp, base pair(s); N12-NTA, nickel nitrotriacetic acid; EMSA, electrophoretic mobility shift assay; M, molecular mass.
FIG. 1. The POZ domain mediates GAGA oligomerization in vivo. A, constructs used in the yeast two-hybrid assays. The POZ245 and I POZ122 constructs were fused to the GAL4 activation domain (AD-GAL4). Full GAL4, the POZ245, and POZ122 constructs were fused to the GAL4 binding domain (BD-GAL4), B, growth in selective medium lacking histidine of strains carrying the POZ245-AD-GAL4 or the POZ122-AD-GAL4 constructs and either the GAL4-BD-GAL4, the POZ245-BD-GAL4, or the POZ122-BD-GAL4 constructs. C, β-galactosidase activities, expressed as Miller units, corresponding to strains carrying several combinations of constructs as indicated. In all cases, the different constructs were fused to the BD-GAL4 and were tested for interaction with the POZ245-AD-GAL4 or the POZ122-AD-GAL4 constructs and the latter to the AD-GAL4.

GAGA Oligomerization

5′-AGC T(TG A)15 CG TAC CGT AC(TG A)15 GTA CGG TAC G(TC)10 A-3′, and 5′-AGC T(GA)15 CGT ACC GGT AAC CGT AC(G A)15(TG A)10GT A CG(T C)10 A3 into the HindIII site of pUC18. Fragments were then excised from the corresponding pUC18 constructs by digestion with EcoRI + PstI.

Yeast Two-hybrid Assays—For the yeast two-hybrid assays the GAGA constructs described in the text were cloned into plasmid pGBT9 (CLONTECH), to express fusion proteins containing the GAL4 DNA binding domain, or into plasmid pGAD424 (CLONTECH), to express fusion proteins containing the GAL4 activation domain. Appropriate plasmids were then transformed into the yeast HF7c strain (MATα, ade2–101, his3–112, trp1–101, leu2–3,112, gal4–542, gal80–528, lys2::GAL1-LacI, GAL1-TATA, HIS3, URA3::GAL4pMBXSS-CYC1-TATA-LacI), and the transformants were tested for their ability to grow on selective medium lacking histidine. For quantitative analysis, transformants were tested for activation of the GAL4-dependent LacZ reporter gene.

EMS A Experiments—For EMSA experiments, the 32P-labeled DNA fragments were incubated with recombinant GAGA, BD-GAL4, or ΔPOZ345 proteins in 50 mM KCl, 15 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 0.5 mM dithiothreitol, and 5% glycerol in the presence of 10 μg of bovine serum albumin and a 50–100-fold excess (w/w) of E. coli DNA for 20 min at room temperature in a final volume of 10 μl. The amount of protein used in each experiment is expressed as units of activity defined as the amount of protein inducing about 50% retardation of 1 ng of fragment DNA at room temperature for 10 min. Fractions were collected from top to bottom, and the presence of the proteins was determined by Western blotting (inserts), using rabbit antiGAGA polyclonal antibodies. Bands of higher electrophoretic mobility arise from proteolytic degradation of the recombinant protein occurring during the dialysis. The positions corresponding to markers of known M: 67, 150, and 232 kDa are indicated at the top of each graph.

DNase I Footprinting Experiments—DNase I digestions were performed with 0.01 units of enzyme (Roche Molecular Biochemicals) for 2 min at room temperature in a final volume of 20 μl. The enzyme was added in 8 mM MgCl2, 4 mM CaCl2, and the reaction stopped by the addition of 200 μl of a solution containing 50 mM Tris-HCl, pH 8, 0.1 mM NaCl, 0.5% SDS. After phenol extraction and ethanol precipitation, samples were analyzed in 7% polyacrylamide, 7 M urea sequencing gels.

Sedimentation Experiments—For sedimentation experiments, purified recombinant proteins were dialyzed overnight against 20 mM Hepes, pH 7.9, 0.1 mM EDTA, 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfon fluoride, 0.1% Nonidet P-40, and the KCl concentration indicated in each case. After dialysis, about 200 μg of protein in 75 μl were loaded onto 10–30% sucrose gradients at the same buffer conditions of the dialysis. Centrifugation was carried out in a Beckman SW41 rotor at 40,000 rpm for 23 h at 4 °C. Fractions of about 0.6 ml were collected from top to bottom and analyzed for the presence of protein by Western blot using a rabbit antiGAGA polyclonal antibody (a gift of Dr. C. Wu) to recognize all full GAGA constructs used in these experiments. Bovine serum albumin (M = 67 kDa) (Amersham Pharmacia Biotech), yeast alcohol dehydrogenase (M = 150 kDa) (Sigma), and bovine liver catalase (M = 232 kDa) (Amersham Pharmacia Biotech) were used as molecular weight markers. When the sedimen-

FIG. 2. Sedimentation profiles through 10–30% sucrose gradients of full GAGA (A and B) or 5–20% sucrose gradients of ΔPOZ345 (C). Gradients were run in the presence of 0.1% Nonidet P-40 at either 50 mM KCl (A and C) or 200 mM KCl (B). Fractions were collected from top to bottom, and the presence of the proteins was determined by Western blotting (inset), using rabbit antiGAGA polyclonal antibodies. Bands of higher electrophoretic mobility arise from proteolytic degradation of the recombinant protein occurring during the dialysis. The positions corresponding to markers of known M: 67, 150, and 232 kDa are indicated at the top of each graph.

The sedimentation behavior of the GAGA-DNA complexes was analyzed, about 40 ng of 32P-labeled GAGA fragment was incubated with 35 μg of protein in the presence of a 50-fold excess (w/v) of E. coli DNA at room temperature
for 30 min at the buffer conditions indicated in each case in a final volume of 200 µL. Samples were then loaded onto the gradients and, after centrifugation, the radioactivity recovered in each fraction was determined by Cerenkov counting.

**Transcription Experiments**—The templates used for in vitro transcription assays were derived from a minimal promoter containing a TATAAA box followed by a G-less cassette as described in (15). A double-stranded oligonucleotide of sequence TTG GGA GCG AGA GGG AA, encompassing the GAGA binding site found in region C of the engrailed promoter (16), was multimerized by self-ligation and inserted at the AccI site at position 240 of the minimal promoter described above. Transcription reactions were carried out as described (17) using 200 ng of each supercoiled template, 50 µg of nuclear HeLa cell extract, and increasing amounts of either recombinant GAGA or DPOZ245 proteins. Quantitative analysis was performed by laser densitometry of the corresponding autoradiographs in a Molecular Dynamics densitometer and normalized using the recovery controls for each case. For each template, maximal activation rates with respect to controls receiving no protein were determined and compared.

**RESULTS**

The POZ Domain of GAGA Mediates Oligomerization Both in Vitro and in Vivo—The ability of the POZ domain of GAGA to mediate protein oligomerization in vivo was determined in a yeast two-hybrid assay. As judged from its homology to other POZ domain containing proteins, the POZ domain of GAGA extends for about the first 120 amino acid residues. Fig. 1A summarizes the constructs used in these experiments. Two different peptides, ΔPOZ122 and POZ245, were fused to the GAL4 activation domain (ADGAL4). ΔPOZ122 is missing the first 122 amino acid residues of GAGA, whereas POZ245 contains the POZ domain plus about half of the region located between the POZ domain and the DNA binding domain. These two constructs were tested for interaction with fusion proteins containing the GAL4 DNA binding domain (BDGAL4) and either the 122 amino acid long POZ domain of GAGA (POZ122), the POZ245 construct, or full GAGA. None of these fusion proteins were able to activate by themselves the GAL4-dependent expression of the HIS3 or LacZ genes used as reporters (not shown). However, as judged from the growth obtained in the absence of histidine (Fig. 1B), the ΔPOZ122::ADGAL4 fusion interacts with all three POZ245::BDGAL4, POZ245::BDGAL4, and GAGA-BDGAL4 fusion proteins. On the other hand, none of these three fusion proteins were found to interact with the ΔPOZ122::ADGAL4 fusion. Similar results were obtained when

**FIG. 3.** The interaction of GAGA with d(GA/TC)n sequences of increasing length. A, the binding of GAGA to fragments GA22, GA10, and GA5 is analyzed by EMSA in 0.8% agarose gels as a function of increasing protein concentration expressed as units of activity (see “Experimental Procedures”). All lanes contain 0.5 ng of the labeled DNA fragment. Panel GA22: 0 units (lane 0); 0.3 units (lane 1); 0.6 units (lane 2); 1.1 units (lane 3); 1.5 units (lane 4); and 2.25 units (lane 5). Panels GA10 and GA5: 0 units (lane 0); 0.6 units (lane 1); 1.1 units (lane 2); 1.5 units (lane 3); 2.25 units (lane 4); and 3 units (lane 5). Quantitative analysis of the results is shown on the right: △, GA5; ○, GA10; and [], GA22. B, DNase I footprinting analysis of the binding of GAGA to fragments GA22, GA10, and GA5 as a function of increasing protein concentration: 0 units (lanes 0); 1.5 units (lanes 1); and 3 units (lanes 2). Lanes L correspond to G + A sequencing ladders. The position of the d(GA/TC)n sequence is indicated. The 5′ to 3′ direction is also indicated.
the *LacZ* gene was used as reporter (Fig. 1C). Also in this case, all three constructs interact with the POZ245-AD<sub>GAL4</sub> fusion but not with the ΔPOZ<sub>122</sub>-AD<sub>GAL4</sub> fusion. These results show that the POZ domain mediates GAGA oligomerization *in vivo*. In agreement with these results, a positive interaction is also detected between POZ<sub>122</sub>-AD<sub>GAL4</sub> and POZ<sub>122</sub>-BD<sub>GAL4</sub> (Fig. 1C). This interaction is of the same magnitude as the interaction between POZ<sub>245</sub>-AD<sub>GAL4</sub> and POZ<sub>245</sub>-BD<sub>GAL4</sub> but weaker than between POZ<sub>245</sub>-AD<sub>GAL4</sub> and POZ<sub>122</sub>-BD<sub>GAL4</sub> which is the strongest. Terminal deletions of the POZ domain have a strong effect on its oligomerization potential, and no interaction is detected upon removal of the last 21 amino acids (POZ<sub>1</sub>-101), the first 30 amino acids (POZ<sub>30</sub>-122), or simply the first 6 amino acids (POZ<sub>6</sub>-122) (Fig. 1C).

The sedimentation behavior through 10–30% sucrose gradients was analyzed to determine the stoichiometry of the GAGA oligomers. As shown in Fig. 2, the sedimentation profile of GAGA shows three relatively well defined peaks when the gradients are performed in the presence of 50 mM KCl (Fig. 2A). A first peak is found at the light fractions of the gradient that must correspond to GAGA monomers, because it shows an apparent molecular mass (M) close to that of GAGA (67 kDa). A second peak is detected at the middle of the gradient showing an apparent M close to what would be expected for a GAGA dimer, and a third peak is found at the dense fractions of the gradient that, from its apparent M, is likely to correspond to tetramers. In addition to these three peaks, part of the molecules are found at the bottom of the gradient corresponding to complexes of stoichiometry higher than tetramers. The sedimentation experiments described above were performed in the presence of 0.1% Nonidet P-40. When the gradients are run in the absence of detergent, all the molecules are found at the dense fractions corresponding to oligomers of high stoichiometry (not shown). The stability of the complexes described above depends on the ionic strength so that, at 200 mM KCl, GAGA molecules are found all throughout the gradient (Fig. 2B). Under these conditions, the monomeric forms predominate and the complexes of high stoichiometry are less abundant. When the sedimentation behavior of the ΔPOZ<sub>122</sub> peptide was analyzed, a single peak was detected at the light region of the gradient (Fig. 2C), indicating that in the absence of the POZ domain only monomers are formed. This behavior is independent of the KCl concentration at which the gradients are run and of the presence or absence of Nonidet P-40 (not shown). Similar results were obtained with the ΔPOZ<sub>245</sub> peptide. These results demonstrate that GAGA oligomerization also occurs *in vitro* and the formation of dimers, tetramers, and complexes of higher stoichiometry is detected at low ionic strength.

The Interaction of GAGA Oligomers with DNA Is of High Affinity and Specificity—Fig. 3 shows the interaction of GAGA with d(GAT<sub>n</sub>C<sub>n</sub>) sequences of increasing length as analyzed by EMSA and DNaSe I footprinting. The DNA fragments used in these experiments were described in detail elsewhere (14). Briefly, fragments GA5, GA10, and GA22 are all about 180-bp long, differing in the size of the repeated d(GAT<sub>n</sub>C<sub>n</sub>) sequence they contain, n = 5, 10, and 22, respectively. Several features of the EMSA experiments presented in Fig. 3A indicate that the protein-DNA complexes detected arise from the binding of GAGA oligomers. First of all, the GAGA-DNA complexes formed with all three DNA fragments show an anomalous low electrophoretic migration, corresponding to a very large apparent M, much higher than would be expected for a polypeptide of 67 kDa such as GAGA. These protein-DNA complexes can only be resolved in low percentage (0.8%) agarose gels where they show a slow electrophoretic mobility. They cannot be resolved in native polyacrylamide gels, barely entering gels of very low percentage (3.5%) (not shown). As a comparison, complexes formed with the same DNA fragments by proteins of similar M are easily resolved in 4–6% polyacrylamide gels (14). Furthermore, regardless of the actual length of the d(GAT<sub>n</sub>C<sub>n</sub>) sequence contained within the fragment, the formation of only one protein-DNA complex is detected, indicating that increasing the length of the binding site 4-fold, from 10 bp in fragment GA5 to 44 bp in fragment GA22, does not result in an increased number of GAGA molecules entering the complex. As judged by DNaSe I footprinting (Fig. 3B), the interaction of GAGA oligomers with all three DNA fragments takes place at the repeated d(GAT<sub>n</sub>C<sub>n</sub>) sequence. In all three cases, the footprints are centered around the d(GAT<sub>n</sub>C<sub>n</sub>) sequence (Fig. 3B), spanning similar regions. Even at high protein concentration, the footprints do not significantly extend beyond the d(GAT<sub>n</sub>C<sub>n</sub>) sequence, and no footprints are detected in other regions of the DNA fragments (Fig. 3B, lanes 2). These results indicate that GAGA oligomers bind DNA sequence specifically. The affinity
of GAGA for fragment GA22, carrying long d(GA-TC)n sequences, is significantly higher than for fragments GA10 or GA5 containing shorter GAGA binding sites (Fig. 3A, right panel). This difference in affinity is better reflected by the amount of protein required to induce 50% retardation, which is about 3-fold higher for fragments GA10 and GA5 than for fragment GA22. The higher affinity of GAGA for fragment GA22 is consistent with the protein being an oligomer capable of establishing a higher number of productive interactions with long d(GA-TC)n DNA sequences than with short ones.

The EMSA experiments reported above were performed under conditions (50 mM KCl in the absence of Nonidet P-40) where, according to its sedimentation behavior, GAGA exists as complexes of high stoichiometry. Similar results are obtained when the EMSA experiments are carried out in the presence of Nonidet P-40 at either 50 mM KCl (Fig. 4A) or 200 mM KCl (Fig. 4B) where, according to the sedimentation experiments described in Fig. 2, a percentage of the GAGA molecules exists as complexes of low stoichiometry, monomers, and dimers. Also under these conditions, the formation of protein-DNA complexes of high M is detected, with both the GA22 and GA5 fragments, and the affinity of GAGA for the fragments carrying long d(GA-TC)n DNA sequences is higher. In addition to the high M complexes, the formation of complexes of fast electrophoretic mobility can also be observed. Although these complexes are detected at different experimental conditions, they appear to be more abundant when the binding is performed under conditions favoring disassembly of the GAGA oligomers (200 mM KCl in the presence of Nonidet P-40) (Fig. 4B), suggesting that they arise from the binding of the GAGA species of low stoichiometry. Similar results were obtained when sedimentation experiments, similar to those described in Fig. 2, were performed in the presence of radioactively labeled GA5 fragment. When the gradients are run at 200 mM KCl, the 32P-labeled fragment is distributed into two peaks (Fig. 4C, center panel). One peak is found at the heavy region of the
gradient, as expected for the binding of GAGA oligomers of high stoichiometry, but in addition a second peak is found at the light region of the gradient, sedimenting significantly faster than the free DNA fragment (Fig. 4C, left panel), as would be expected for the binding of GAGA forms of low stoichiometry. Consistent with this interpretation, when the gradients are run at 50 mM KCl, most of the radioactivity is recovered at the heavy fractions and only a low percentage is found at the light region of the gradient (Fig. 4C, right panel). Under these low ionic conditions, a slow sedimenting peak is also detected at the position corresponding to the free DNA fragment.

The characteristic DNA binding behavior of GAGA described above is not observed in the absence of the POZ domain. When the interaction of the $\Delta$POZ$_{245}$ peptide with the GA5, GA10, and GA22 DNA fragments was studied, a totally different binding behavior was observed (Fig. 5). In this case, the apparent M of the protein-DNA complexes corresponds to what would be expected from the size of the polypeptide. These protein-DNA complexes are resolved in native 4.5% polyacrylamide gels where they show an electrophoretic mobility consistent with the binding of the 274-amino acid residue-long $\Delta$POZ$_{245}$ peptide (Fig. 5A). In the presence of low amounts of $\Delta$POZ$_{245}$, a single protein-DNA complex is detected, but on increasing the protein concentration, the formation of additional bands of slower electrophoretic mobility is observed reflecting the binding of additional $\Delta$POZ$_{245}$ peptides to the DNA fragments. As judged from the number of complexes detected, fragment GA5 accommodates basically only one $\Delta$POZ$_{245}$ molecule, whereas fragments GA10 and GA22 can accommodate up to two and three independent molecules, respectively. When the interaction of the DNA binding domain of GAGA (BD$_{GAGA}$) with the same DNA fragments described above was studied, a binding behavior very similar to that of the $\Delta$POZ$_{245}$ peptide was observed (not shown). As observed with full GAGA, binding of $\Delta$POZ$_{245}$ results in intense DNase I footprints centered around the repeated d(GA-TC)$_n$ sequence (Fig. 5B). The binding of $\Delta$POZ$_{245}$ to fragment GA22 is slightly more efficient than to fragment GA5, as would be expected from the higher number of binding sites present in fragment GA22 (Fig. 5A, right panel), but the difference in relative affinity in this case is much lower than that observed with full GAGA. These results indicate that, contrary to full GAGA, which binds DNA as an oligomer, $\Delta$POZ$_{245}$ binds as a monomer whose interaction with DNA fragments containing long d(GA-TC)$_n$ sequences, capable of accommodating several $\Delta$POZ$_{245}$ molecules, is noncooperative.

Altogether these results show that GAGA oligomers bind DNA with high affinity and specificity and corroborate the important contribution of the N-terminal POZ domain to the formation of these oligomers and, thereby, to their DNA binding properties.

A Single GAGA Oligomer Can Interact with Two Adjacent Binding Sites Spaced by as Many as 20 bp—In general, GAGA binding sites at promoters show a peculiar distribution in which relatively short sites, spaced by small distances, form clusters (5). Most frequently, the individual binding sites of each cluster are spaced between 10 and 30 bp. Promoters might contain several clusters of GAGA binding sites, which are then spaced by larger distances. To analyze the ability of GAGA to interact with adjacent binding sites, the binding of GAGA to DNA fragments containing two d(GA-TC)$_n$ sequences spaced by either 5 bp (GA10/5), 10 bp (GA10/10), or 20 bp (GA10/20) was determined. As shown by the EMSA experiments in Fig. 6A, the affinity of GAGA oligomers for fragments GA10/5, GA10/10 and GA10/20 is higher than for fragment GA10, which
contains a single d(GA-TC)_10 site, but very similar to the affinity for fragment GA22, which for this purpose can be considered as containing two immediately adjacent sites of about the same size. Furthermore, regardless of the distance between the two binding sites, the formation of only one protein-DNA complex of high M is detected. This behavior depends on the presence of the POZ domain, because binding of ΔPOZ_245 to fragments GA10/5, GA10/10, and GA10/20 is noncooperative showing a similar affinity as to fragments GA10 or GA22 (Fig. 6B). BD_{GAGA} shows a similar binding behavior as ΔPOZ_245 (not shown). These results indicate that a single GAGA oligomer is binding to the fragments containing two independent d(GA-TC)_10 sequences even when spaced by 20 bp, suggesting that GAGA oligomers can bind simultaneously to at least two, and likely more, adjacent binding sites. Interestingly, binding of GAGA does not protect from DNase I cleavage the region located between the two adjacent d(GA-TC)_10 sequences of fragments GA10/5, GA10/10, and GA10/20 (Fig. 7A), indicating that, in these DNA fragments, the region between the two GAGA binding sites is not in intimate contact with the protein. Actually, the footprints obtained with full GAGA are very similar to those obtained with the ΔPOZ_245 peptide (Fig. 7B). These results show that the simultaneous interaction of GAGA with adjacent sites is also highly specific.

**Multiple GAGA Binding Sites Are Required for Efficient Transcription Activation in Vitro—**GAGA is known to enhance transcription from promoters containing d(GA-TC)_n sequences, both in vitro and in vivo (5, 18). To analyze the contribution of the presence of multiple binding sites to the transcription activity of GAGA, we determined the rate of GAGA-dependent transcription activation from promoters containing an increasing number of GAGA binding sites. For these experiments, the GAGA binding site found at the C-region of the engrailed promoter (16) was multimerized and fused to a minimal promoter, which was shown earlier to efficiently drive transcription of a G-less cassette (15). As shown in Fig. 8A, the constructs used in these experiments contained from 1 to 6 copies of this engrailed site. The extent of maximal activation obtained in the presence of GAGA strongly depends on the number of binding sites present at the promoter (Fig. 8B). No significant activation is observed from constructs containing only one or two GAGA binding sites, and only a moderated 3-fold activation is observed in the presence of three binding sites. However, a strong increase in activation, to about 8–9-fold, is seen from constructs containing five or six binding sites. Similar results were reported earlier by others (5). This behavior depends on the presence of the POZ domain. When the transcription activity of the ΔPOZ_245 peptide was analyzed, a significant activation was observed in the presence of two GAGA binding sites, which increases only slightly as the number of binding sites do (Fig. 8B). In this case, a low though reproducible activation is detected even in the presence of a single site. The synergism in transcription activation detected upon increasing the number of binding sites is consistent with the higher affinity of GAGA oligomers for fragments carrying multiple GAGA sites. Consistent with this hypothesis, this synergism depends on the presence of the POZ domain.

**DISCUSSION**

We have shown that GAGA forms oligomers both in vitro and in vivo. Others (19) have also reported the formation of GAGA-DNA complexes compatible with the formation of GAGA oligomers. Here, however, we have been able to show that formation of these oligomers, which exist in solution in the absence of DNA binding, requires the contribution of the N-terminal POZ domain and that GAGA can form different types of oligomers, dimers, but also tetramers and complexes of high stoichiometry. Several results indicate that these oligomeric forms do not arise from the unspecific in vitro self-aggregation of the protein. First, they bind d(GA-TC)_n sequences with high specificity as reflected by the fact that their DNase I footprints do not extend beyond the d(GA-TC)_n sequence, even when binding to DNA fragments carrying two independent binding sites spaced by as few as 5 bp (Fig. 7). Second, the formation of these GAGA oligomers requires the contribution of specific protein-protein interactions involving the N-terminal POZ domain, deletion of the first 6 or the last 21 amino acid residues of the domain abolishing the interaction in vivo. Finally, these GAGA oligomers are functionally active, being transcriptionally competent in vitro. Altogether, these observations indicate that GAGA oligomers are functionally active forms of the protein arising from specific protein-protein interactions.

The crystal structure of the POZ domain of the human promyelocytic leukemia zinc finger protein was recently reported (20). In the crystal, the promyelocytic leukemia zinc finger POZ

---

**Fig. 7.** DNase I footprinting analysis of the interaction of GAGA (A) and ΔPOZ_245 (B) with fragments GA22, GA10/5, GA10/10, GA10/20, and GA10 as a function of increasing protein concentration. Panel A: 0 units (lanes 0), 1.1 units (lanes 1), and 3 units (lanes 2). Panel B: 0 units (lanes 0), 0.6 units (lanes 1), and 1.5 units (lanes 2). Lanes L correspond to G + A sequencing ladders. The 5’ to 3’ direction is indicated.
domain forms a dimer that is held together by interactions involving the most N- and C-terminal regions of the monomers. In particular, the five most N-terminal residues of one monomer form a β1-strand that interacts with an internal β5’-strand on the second monomer. One face of this β1/β5’ sheet is then packed against the α6’ helix of the second monomer, which spans the last 15 residues, establishing a number of stabilizing hydrophobic interactions. These interactions provide a reasonable interpretation for the strong effect that terminal deletions have on GAGA oligomerization in vivo, because given its high degree of conservation, it is reasonable to assume that all POZ domains will have a similar fold.

Our results indicate that in solution GAGA can form dimers and tetramers, as well as oligomers of higher stoichiometry. This behavior is not influenced by the source of the protein used that, in the experiments reported here, was expressed in E. coli as a His-tagged protein. Non-His-tagged recombinant GAGA also binds DNA as an oligomer (21), and native GAGA protein obtained directly from Drosophila cells (22, 23) or by cell-free in vitro transcription-translation (24) behaves also as an oligomer of high M. Other POZ domain-containing proteins have also been shown to form oligomers (7–13). Although, in general, the stoichiometry of these complexes was not determined, the formation of both dimers and tetramers was observed in some cases (12). Interestingly, though the promyelocytic leukemia zinc finger POZ domain appears to form only dimers in solution, the formation of a short four-stranded antiparallel β-sheet between two symmetry related dimers is observed in the crystal (20). This interaction involves four different peptide chains and, therefore, can give rise to the formation of tetramers and oligomers of higher stoichiometry.

The binding experiments reported here were performed under conditions where GAGA exists mainly as tetramers and oligomers of higher stoichiometry. Consistent with these results, the GAGA-DNA complexes formed under these conditions are of high M. However, when the binding was performed under conditions where a significant percentage of the GAGA molecules exists as oligomers of low stoichiometry, the formation of protein-DNA complexes of low M could also be observed (Fig. 4), indicating that the GAGA species of low stoichiometry, monomers and dimers, can also bind DNA. Nevertheless, the formation of GAGA-DNA complexes of high M is also observed under these conditions, predominating at high protein concentration (Fig. 4B). These results suggest that the GAGA oligomers of high stoichiometry bind DNA very efficiently or, alternatively, that DNA binding facilitates GAGA oligomerization.

Others (7) had reported earlier that GAGA, through its POZ domain, forms in vitro heteromeric complexes of unknown stoichiometry with other POZ domain-containing proteins, such as tramtrack. Several other Drosophila proteins, including the enhancer of variegation E(var)3-39D, contain similar POZ domains (13, 25), but it is not known whether they can also interact with GAGA. These observations indicate that the POZ domain is likely to play a key role in the control of GAGA function(s) in vivo, by regulating its oligomerization state and its interaction with other proteins. Interestingly, GAGA associates with the centromeric (AAGAG), heterochromatin during mitosis (3, 4). The formation of large oligomers might play a role in this context. The POZ domain might also regulate DNA binding. Earlier studies indicate that the POZ domain inhibits the interaction of GAGA with a short 17-bp-long DNA fragment carrying a single GAGA binding site (7). We have also observed that, compared with the BDGAGA peptide, GAGA binds with high difficulty to short synthetic oligonucleotides (not shown). However, as reported here and elsewhere (5, 26), GAGA binds efficiently to DNA fragments carrying longer or multiple GAGA binding sites. Our results also show that the affinity of GAGA for DNA depends strongly on the length and number of binding sites, being significantly lower for fragments carrying single short d(GA\textgreek{TC})\textsubscript{n} sequences than for fragments carrying either several short sites or a long one, providing a reasonable interpretation for the apparently contradictory results mentioned above.

We have shown here that a single GAGA oligomer can bind two adjacent GAGA binding sites spaced by as many as 20 bp. Others have recently reported similar results (94). As judged from EMSA and DNase I footprinting experiments, this interaction is of high affinity and specificity. As mentioned before, most frequently, promoters contain clusters of independent d(GA\textgreek{TC})\textsubscript{n} sequences that, as derived from our results, could be bound by a single GAGA complex. Previously reported results suggest that this situation might occur both in vitro (16, 27) and in vivo (28). Moreover, a single short d(GA\textgreek{TC})\textsubscript{n} sequence is not a good substrate for GAGA binding in vitro. These results suggest that several adjacent d(GA\textgreek{TC})\textsubscript{n} sequences are re-
required to create a GAGA binding site of high affinity. Given the repetitive character of the sequences that GAGA recognizes, this mode of interaction, requiring the presence of several independent d(GA-TC)₃ sites for efficient protein binding, provides an additional factor that increases the specificity of the interaction of GAGA with DNA.

Several observations suggest that, to some extent, GAGA functions at the chromatin level, participating in the formation of an open chromatin structure. GAGA is the product of the Trithorax-like (Trl) gene (29) which, being a member of the Trithorax group, antagonizes the chromatin-mediated repression that Polycomb genes induce on the expression of the homo gene. A more direct link to chromatin structure is indicated by the fact that Trl is an enhancer of position effect variegation (29). Moreover, in collaboration with nucleosome remodelling factor, GAGA was shown to help nucleosome disruption at specific regions of the hsp70 promoter, encompassing GAGA binding sites (26, 30, 31). At present, little is known about the specific contribution of GAGA to chromatin remodelling and, although other DNA-binding proteins such as HSF (heat shock factor) or GAL4-HSF can also help disrupting chromatin organization (29, 31, 32), GAGA appears to be particularly efficient in this respect (31, 33). Although a direct interaction with the chromatin remodelling machinery cannot be excluded, the simultaneous interaction of GAGA oligomers with multiple adjacent sites could significantly contribute to the higher efficiency of GAGA in disrupting nucleosomes. In this context, it would be interesting to know whether a functional POZ domain is required for efficient nucleosome disruption. GAGA can also activate transcription in vitro suggesting a possible interaction with the basal transcription machinery. Our results indicate that the presence of several independent GAGA sites is required for efficient transcription activation in vitro, indicating that the oligomeric character of GAGA might also be functionally relevant in this context. Interestingly, in the case of the ΔPOZ₂₄₅ peptide, significant transcription activation is detected in the presence of a single binding site, and no synergism is observed upon increasing the number of GAGA binding sites. These results suggest that the synergism observed with full GAGA arises from specific features of the GAGA-DNA complex rather than from the simple recruitment of multiple GAGA molecules to the promoter.

Acknowledgments—We are thankful to Dr. C. Wu for providing us with αGAGA antibodies and to Dr. M. Ortiz-Lombardía for purified GAGA protein. We are also thankful to Dr. B. Piña for helpful discussions and advice.

REFERENCES

1. Granok, H., Leibovitch, B. A., Shaffer, C. D., and Elgin, S. C. R. (1995) Curr. Biol. 5, 238–241
2. Wilkins, R. C., and Lia, J. T. (1997) Nucleic Acids Res. 25, 3963–3968
3. Platero, J. S., Csink, A. K., Quintanilla, A., and Henikoff, S. (1998) J. Cell Biol. 140, 1297–1306
4. Raff, J. W., Kellum, R., and Alberts, B. (1994) EMBO J. 13, 5977–5983
5. Soeller, W. C., Oh, C. E., and Kornberg, T. B. (1995) Mol. Cell. Biol. 15, 7961–7970
6. Pedone, P. V., Ghirlando, R., Clere, G. M., Grönenborn, A. M., Felsenfeld, G., and Omichinski, J. G. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 2822–2826
7. Bardwell, V. J., and Tsevimas, R. (1994) Genes Dev. 8, 1664–1677
8. Chen, W., Zollman, S., Couderc, J. L., and Laski, F. A. (1995) Mol. Cell. Biol. 15, 3424–3429
9. Dhordain, P., Albagl, O., Ansieau, S., Koken, M. H., Deweindt, C., Quiét, S., Lantoine, D., Leutz, A., Kerckaert, J. P., and Leprince, D. (1995) Oncogene 11, 2689–2697
10. Dong, S., Zhu, J., Reid, A., Strutt, P., Guidez, F., Zhong, H.-J., Wang, Z.-Y., Licht, J., Waxman, S., Chomienne, C., Chon, Z., Zelent, A., and Chen, S.-J. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 3624–3629
11. Godt, D., Couderc, J. L., Cramton, S. E., and Laski, F. A. (1993) Development 119, 799–812
12. Robinson, D. N., and Cooley, L. (1997) J. Cell Biol. 138, 799–810
13. Zollman, S., Godt, D., Prive, G. G., Couderc, J. L., and Laski, F. A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 10717–10721
14. espinas, M. L., Jiménez-Garcia, E., Martínez-Balbis, A., and Azorín, P. (1996) J. Biol. Chem. 271, 31807–31812
15. Bernués, J., Carrera, P., and Azorín, F. (1996) Nucleic Acids Res. 24, 2950–2958
16. Soeller, W. C., Poole, S. J., and Kornberg, T. B. (1988) Genes Dev. 2, 68–81
17. Bernués, J., Simmen, K. A., Lewis, J. D., Gunderson, S. I., Polycarpou-Schwarz, M., Moncini, V., Egy, J.-M., and Mattaj, I. W. (1993) EMBO J. 12, 3573–3585
18. Croston, G. E., Kerrigan, L. A., Lira, L. M., Marshall, D. R., and Kadonaga (1997) J. Cell Biol. 251, 1297–1306
19. Zollman, S., Godt, D., Prive, G. G., Couderc, J. L., and Laski, F. A. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 799–811
20. Ahmad, K. F., Engel, C. K., and Prive, G. G. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 12123–12128
21. Jiménez-García, E., Vaquero, A., Espínas, M. L., Soliva, R., Orozco, M., Bernués, J., and Azorín, F. (1998) J. Biol. Chem. 273, 24640–24648
22. Biggin, M. D., and Tjian, R. (1988) Cell 53, 689–711
23. Gilmour, D. S., Thomas, G. H., and Elgin, S. C. R. (1989) Science 245, 1487–1490
24. Lehmann, M., Siegmund, T., Løntermann, K.-G., and Korge, G. (1998) J. Biol. Chem. 273, 24554–24559
25. Dorn, R., Kraus, V., Reuter, G., and Saumweber, H. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 11376–11380
26. Tsukiyama, T., Becker, F., and Wu, C. (1994) Nature 367, 525–522
27. Read, D., Nishigaki, T., and Manley, J. L. (1990) Mol. Cell. Biol. 10, 4334–4344
28. Weber, J. A., and Gilmour, D. S. (1995) Nucleic Acids Res. 23, 3327–3334
29. Farkeas, G., Gazze, J., Gallon, M, Reuter, G., Gyurkorics, H., and Karr, F. (1994) Nature 371, 806–808
30. Okada, M., and Hirose, S. (1998) Mol. Cell. Biol. 18, 2455–2461
31. Tsukiyama, T., and Wu, C. (1995) Cell 82, 1011–1020
32. Wall, G., Varga-Weisz, P. D., Sandaltzopoulos, R., and Becker, P. (1995) EMBO J. 14, 1727–1736
33. Mizuguchi, G., Tsukiyama, T., Wissniewski, J., and Wu, C. (1997) Mol. Cell 1, 141–150
34. Katsau, K. R., Hajibagheri, M. A. N., and Verrijzer, C. P. (1999) EMBO J. 18, 698–708