THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 270, No. 42, Issue of October 20, pp. 25150–25158, 1995

Printed in U.S.A.

Expression of a *Drosophila* GATA Transcription Factor in Multiple Tissues in the Developing Embryos

IDENTIFICATION OF HOMOZYGOUS LETHAL MUTANTS WITH P-ELEMENT INSERTION AT THE PROMOTER REGION*

(Received for publication, January 17, 1995, and in revised form, June 29, 1995)

Wen-Hsing Lin‡, Li-Hsuan Huang‡, Jih-Yun Yeh‡, Jörg Hoheisel§¶, Hans Lehrach§, Y. Henry Sun‖, and Shih-Feng Tsai‡**

From the ‡Institute of Genetics, National Yang-Ming University, Taipei 112, Taiwan, Republic of China, the ‖Institute of Molecular Biology, Academia Sinica, Nankang, Taipei 11529, Taiwan, Republic of China, and the §Imperial Cancer Research Fund, P. O. Box 123, Lincoln's Inn Fields, London WC 2A 3PX, United Kingdom

GATA transcription factors are DNA-binding proteins that recognize the core consensus sequence, WGATAR. Previous studies indicated that GATA factors play an important role in the development of tissue-specific functions in vertebrates. Here we report the identification of a new *Drosophila melanogaster* GATA factor, dGATAc, which displays a distinct expression pattern in embryos. The local concentration of dGATAc transcripts varies at different stages, being most prominent in the procephalic region at stages 6–10 and in the posterior spiracles, the gut, and the central nervous system at stages 11–13. On the basis of its predicted sequence, DNA-binding assays were performed to confirm that the dGATAc gene encodes a zinc finger protein that can bind the GATA consensus motif with predicted specificity. Two independent mutants carrying a P-element insertion at the dGATAc gene promoter region were identified that are homozygous lethal at the embryonic stage. Using a genetic scheme, it was demonstrated that the lack of dGATAc function can block normal embryonic development. Our results suggest that the dGATAc protein is a tissue-specific transcription factor that is vital to the development of multiple organ systems in *D. melanogaster*.

The GATA-1 transcription factor was originally identified as a DNA-binding protein that recognizes regulatory elements in the promoter and enhancer regions of several erythroid-expressed genes (1–5). Using cloned mouse GATA-1 (4) as a probe, Yamamoto *et al.* (6) isolated from a chicken cDNA library three different GATA factor genes that encode amino acid sequences homologous to the mouse GATA-1 DNA binding domain. Although these genes were shown to express specifically among different tissues, their protein products apparently share similar DNA binding specificities (6). Subsequently, complementary DNA clones corresponding to these three major GATA factors were identified for humans (7–14), mice (12), and *Xenopus laevis* (15). At the 1990 Globin Switching Conference (Airlie House, Virginia), different members of the GATA gene family were named GATA-1, GATA-2, and GATA-3, and a prefix was given to denote their specific origin (16). Recently, a new member, GATA-4, was added to the vertebrate GATA gene family (17, 18).

Except for the mouse GATA-1 (mGATA-1) and GATA-2 (mGATA-2), shown clearly to play critical roles in controlling erythroid differentiation in the hematopoietic system (for review, see Ref. 19), the exact function of other mammalian GATA factors remains largely unknown. From the fact that only certain tissues express mRNA for the individual GATA factor genes, and because tissue-specific functions are regulated by selective GATA factors, it is generally accepted that GATA proteins, as transcriptional factors, are crucial for both the initial decision and subsequent development of lineage-specific functions. As shown by gene-targeting experiments with the mGATA-1 and mGATA-2 genes (20, 21), disruption of their normal function leads to the failure of blood cell development.

The GATA finger domain is not limited to vertebrates; transcription factors bearing similar amino acid sequences were reported for *Aspergillus nidulans* (22), *Neurospora crassa* (23), *Saccharomyces cerevisiae* (24), and *Caenorhabditis elegans* (25). In these species, proteins were identified that are highly conserved in the so-called "finger domain," which contains a characteristic amino acid sequence: Cys-*X*-Asn-Cys-X_{17}-Cys-Asn-Ala-Cys. Recently, three groups (26–28) have reported the isolation of genomic and cDNA clones encoding the GATA factors in *Drosophila melanogaster*.

To extend the search for evolutionarily related GATA factors to invertebrates, and to study the function of these factors in a model organism, we set out to isolate homologous sequences encoding *Drosophila* GATA factors. Initially, we took advantage of the conserved nature of the GATA zinc finger domain and used the GATA finger sequence as a probe to isolate two different *Drosophila* GATA genes (originally named dGATA-I and dGATA-II). Subsequently, using cloned dGATA cDNA fragment as a probe, we analyzed the RNA distribution in the developing embryos. One of the isolated genes, dGATA-I, is identical to the previously reported dGATAa/*pannier* (27, 28). On the other hand, dGATA-II is unique and was found to be expressed specifically in the head region, the gut, the posterior spiracles, and the central nervous system. Following the nomenclature used by previous works (26, 27) for the *Drosophila* GATA factors, hereafter, our dGATA-II should be named

* This work was supported in part by Grants NSC82-0203-B001-98 (to Y. H. S.) and NSC82-0412-080-102 (to S.-F. T.), and the Clinical Research Center, Veteran's General Hospital-Taipei, Institute of Biomedical Science, Academia Sinica, Republic of China. 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 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) D50542.

¶ Present address: Molecular-Genetic Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120, Germany.

** To whom correspondence should be addressed. Tel.: 886-2-822-5355; Fax: 886-2-826-4930.

GATAc.

Based on sequence comparison of the *Drosophila* GATA factors, the analysis of the dGATAc gene expression, and the genetic studies of the dGATAc mutants, we conclude that 1) GATA factor genes are also present in multiple forms in invertebrate animals; 2) the expression of the dGATAc gene is limited to defined tissues, although it can participate in the development of multiple organ systems; and 3) the function of dGATAc protein is essential to the development of *Drosophila* embryos.

MATERIALS AND METHODS

Isolation of cDNA and Genomic Clones for Drosophila GATA Factors—To isolate GATA factor genes from *D. melagonaster*, we designed degenerate primers for polymerase chain reaction (29) corresponding to the invariant amino acid sequences, ECVNC and LGCANC, of the GATA finger region. The upper strand primer is 5′-GATCAAGCTT-GA(A/G)TG(C/T)GTNAA(C/T)TG(C/T)GG-3′, and the lower strand primer is 5′-GATCTCTAGA(A/G)NCC(A/G)CANGC(G/A)TT(G/A)CA-3′. 50 ng of *Drosophila* genomic DNA were used in a reaction containing 50 mM KCl, 10 mM Tris-HCL, pH 8.4, 1.5 mM $MgCl_2$, 50 μM dNTP, 0.5 μg of each primer, and 2.5 units of *Taq* DNA polymerase (Amplitaq, Cetus). After initial denaturation at 94 °C for 5 min followed by 30 cycles of denaturation (94 °C for 30 s), annealing (50 °C for 30 s), and extension (72 °C for 1 min) and a final extension at 72 °C for 10 min, the amplified product was purified, digested with *Hin*dIII plus *Xba*I, and subcloned into the pUC18 vector for sequence confirmation.

A 102-bp[1] genomic sequence from the positive clone was labeled by radioactive PCR to screen cDNA libraries made from either *Drosophila* embryos or I and II stage instar larvae (30). In summary, 2×10^6 clones were screened at a density of 40,000 plaques/150-mm Petri dish. Hybridization was carried out with 5 × SSPE, 5 × Denhardt's solution, 0.1% SDS, 100 μg/ml heat-denatured salmon sperm DNA, and 1×10^6 cpm/ml probes at 65 °C for 16 h. Stringent washing was done with 0.1 × SSC and 0.1% SDS at 65 °C for 2 h. Autoradiography was carried out using Kodak XAR film with one intensifying screen for 16 h.

Screening for *Drosophila* GATA genes was carried out with cosmid libraries using a reference system developed by Hoheisel *et al.* (31). A *Drosophila* genomic library constructed in the bacteriophage lambda Charon 4A vector (32) was also used. The same conditions for hybridization, washing, and autoradiography were carried out as described above for screening cDNA libraries.

DNA Sequencing Strategy—A combination of subcloning and primer directed sequencing was used to determine the DNA sequence in both strands. Plasmid template DNA was prepared according to the method of Chen and Seeburg (33) for a sequencing reaction using the dideoxy chain termination method (34) with a Sequenase (35) kit (U. S. Biochemical Corp.). DNA sequence editing was carried out with the MicroGenie computer program (Beckman, Palo Alto, CA) (36), and the comparison of compiled DNA sequences was accomplished using the computer programs of the GCG sequence analysis software package (Genetic Computer Group, Madison, WI) (37).

Detection of RNA Transcripts in Whole Mount Embryos—*Drosophila* embryos were collected between 0 and 24 h of development. *In situ* hybridization with nonradioactive probes was performed according to the method of Tautz and Pfeiffle (38). The 2.9-kb *Eco*RI fragment from dGATAc clone 4 was subcloned into pBluescript SK vector. After linearizing with *Sal*I or *Nco*I, an *in vitro* RNA synthesis reaction was carried out with T3 polymerase, using digoxigenin-labeled UTP as substrate. The preparation of embryos for hybridization was carried out as described previously (38), and the development of hybridized signals was performed using the Boehringer Mannheim digoxigenin labeling and detection kit. Staging of embryos was according to Campos-Ortega and Hartenstein (39).

In Vitro Transcription and Translation—The 2.9-kb insert of dGATAc clone 4 was subcloned into pBluescript SK. After linearizing with *Spe*I, 0.5 μg of plasmid DNA was used for *in vitro* transcription (40). *In vitro* translation was carried out with rabbit reticulocyte lysate (Promega, Madison, WI) and [^{35}S]methionine according to the protocol described by Kevin Struhl (41). *In vitro* translated protein was analyzed on a 10% SDS-polyacrylamide gel electrophoresis gel followed by autoradiography.

Expression of Glutathione S-Transferase-GATA Chimeric Proteins in Escherichia coli—The entire open reading frame of dGATAc cDNA sequence was amplified by the polymerase chain reaction (29). The upper strand primer (5′-ATCCCGGGAATGGATATGACCTCAAC-3′) contains a *Sma*I recognition site and the AUG initiator sequence, while the lower strand primer (5′-CCGAATTCTCACGTGTAGTCAGTGC-3′) includes an *Eco*RI site and the terminator sequence. PCR was carried out with a thermal cycler (M-J Research, PTC 100). After initial denaturation at 94 °C for 5 min, DNA was amplified through 30 cycles of denaturation (94 °C for 30 s), annealing (55 °C for 30 s), and extension (72 °C for 1 min). Following a final extension at 72 °C for 10 min, the amplified products were digested with *Sma*I/*Eco*RI and subcloned into the pGEX 2T vector (42). For the zinc finger domain, two PCR primers were synthesized to amplify the DNA sequence between nucleotide 779 and 1350. The sequences for the upper strand primer and the lower strand primer were 5′-GAGGATCCCCGCCAAAGGACTCCACGC-3′ and 5′-GAGAATTCGACTTGGAGCTCAGCTTGC-3′, respectively. Since the amplified fragment contains an internal *Bam*HI site at nucleotide 940 of the dGATAc cDNA sequence, only the 411-bp fragment was subcloned into the *Bam*HI/*Eco*RI-digested pGEX2T vector. To express recombinant proteins for gel mobility shift assays, 1.8 ml of LB culture was seeded with 0.2 ml of an overnight culture and grown to the log-phase (A_{600} = 0.6). At that time, isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.5 mM, and the induced culture was incubated for a further 3 h before harvesting. *E. coli* from the induced culture (1.5 ml) was collected after microcentrifugation for 20 s, and the pellet was resuspended in 0.3 ml of 1× phosphate-buffered saline buffer (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \cdot 7H_2O$, 1.4 mM KH_2PO_4). Lysate was prepared by sonication (3 pulses of 20 s each, 15-second rest between pulses) with a HeatSystems-Ultrasonics, Inc. model W385 sonicator (Farmingdale, NY). After clearing the lysate by microcentrifugation at 4 °C for 5 min, the supernatant was saved for DNA binding assays and SDS-polyacrylamide gel electrophoresis.

Gel Mobility Shift Assay—The gel mobility shift assay was carried out as described previously (3, 4, 43). Homologous (5′-CAGTCGAGTC-CATCTGATAAGACTTATCTGCTGCCCCAGA-3′) and heterologous (5′- AAGGAGGCGGCACACCCCCTCCCCTGCACTGCCCCACCCACT-GGGGCACC-3′) oligonucleotides from the mouse GATA-1 promoter sequence (43) were used as competitors to test DNA binding specificity.

Polytene Chromosome Mapping—Polytene chromosomes were prepared from the squash of larval salivary glands of *D. melanogaster* (Canton-S strain) according to a standard procedure (44). A nonradioactive probe was labeled with biotinylated UTP by the random primed method (45). The 5′ *Eco*RI/*Bam*HI fragment from dGATAc clone 4 was used to avoid the conserved finger domain. After washing with 2× SSC at 53 °C (3 times, 20 min each), the hybridizing DNA was detected by the chromogenic substrate.

Southern Analysis of P1 Clones—Six different P1 clones (46) mapping to the section 84 polytene chromosome region were obtained from T. Yamazaki (Kyushu University, Fukuoka, Japan). These clones have, on average, an insert of 85 kb. DNA was prepared by the alkaline lysis method (47) and subjected to restriction with *Eco*RI, *Sac*I, or *Eco*RI plus *Sac*I. Digested DNA was transferred by alkaline solution (0.4 M NaOH) onto the Hybond N^+ filter in triplicate lanes. Three different probes corresponding to the 5′ (a 3-kb *Sac*I to *Eco*RI fragment from the cosmid clone, c78A9), the finger region (a 0.4-kb *Bam*HI/*Eco*RI fragment from cDNA clone 13.3), and the 3′ end of the dGATAc gene (a 5.3-kb *Eco*RI fragment from the cosmid clone, c41A3) were used for each filter strip. Hybridization and stringent washing (0.1× SSC, 0.1% SDS at 65 °C for 15 min) were carried out according to the manufacturer's instructions. Autoradiography was performed with Kodak XAR film with one intensifying screen for 4 h.

Analysis of the P-element Insertion Strains—Six different P-element insertion lines that are mapped to the 84F region (P971, P227, P609, P614, P812, and 2352) were obtained from the Bloomington Stock Center (University of California, Berkeley, CA). Another line, *l*(3)5930 (48), was kindly provided by G. Rubin. Genomic DNA was isolated by the standard method. Four restriction enzymes (*Eco*RI, *Bam*HI, *Pst*I, and *Sac*I) were used separately to digest 10 μg of each DNA sample, and the restricted DNA was blotted to a Hybond-N^+ (Amersham Corp.) nylon filter by alkaline transfer (0.4 M NaOH). Either a 3.0-kb *Sac*I/*Eco*RI genomic fragment, or the entire dGATAc cDNA, was employed as a probe, and labeled by the random prime method (45). Prehybridization, hybridization, and stringent washing were carried out as described above.

Two approaches were used to determine the flanking sequence of P-insertional mutants. To rescue the 5′-junctional sequence of *l*(3)5930 integration site, genomic DNA isolated from 12 adult flies was digested

[1] The abbreviations used are: bp, base pair(s); PCR, polymerase chain reaction; kb, kilobase pair(s); CS, Canton-S.

with *Xba*I, heat inactivated (65 °C for 20 min), and then self ligated. After transformation into DH5α, kanamycin-resistant colonies were isolated for DNA sequencing. PCR cloning was carried out to isolate the 3′ end junctional fragment of P812 integration site. A *white* gene-specific primer (5′-GCATATATACCCTTCTGAATGC-3′) and a dGATAc gene specific primer (5′-CCGGGAATTCCCATGGCGGTCTAGAGCA-CACTGTTTCAATCACTC-3′) were used in a 100-μl PCR reaction containing 1× PCR buffer with 1.5 mM $MgCl_2$, 50 μM dNTP, 0.5 μg of each primer, and 2.5 units of *Taq* polymerase. 35 cycles of reaction were carried out in a M-J research PTC-100 machine. Annealing was at 55 °C for 1 min, and extension was at 72 °C for 2 min. The PCR product was gel-purified and cloned into the pGEM-T vector (Promega, Wisconsin) for sequence determination.

Genetic Analysis of the dGATAc Mutants—Two lines containing a P-element insertion at the dGATAc promoter, P812 and *l*(3)5930, were crossed (*l*(3)5930/TM6B, Tb × P812/TM3, Sb) to determine whether the two dGATAc mutant alleles could complement each other. The flies that developed into adulthood were examined for the expression of the Tb marker gene and the Sb marker gene. The phenotypes of the marker genes are described in the red book (49).

To isolate *l*(3)5930/*l*(3)5930 homozygous embryos, the original *l*(3)5930 line (*l*(3)5930/TM6B, Tb) was crossed with the CS wild-type strain (+/+). The heterozygous progeny with *l*(3)5930/+ genotype were collected and intercrossed. Embryos were collected for 48 h, and those that failed to develop were isolated for DNA extraction. Genotype determination was performed by Southern analysis as described above using *Bam*HI restriction and a 3.0-kb *Sac*I/*Eco*RI genomic fragment as a probe.

RESULTS AND DISCUSSION

Isolation of cDNA Clones Encoding Drosophila GATA Factors—As the zinc finger domains of different GATA factors are well conserved, we initially designed degenerate primers to amplify homologous sequences from *Drosophila* genome. Two types of GATA zinc finger sequences were identified and named dGATAa and dGATAc, respectively. The dGATAa sequence is identical to those reported by Winick (27) and Ramain (28). To isolate cDNA clones corresponding to the dGATAc genomic sequence, reverse transcription PCR was carried out to determine at which stage the dGATAc gene is expressed. Primers were synthesized according to the genomic sequence (upper strand, 5′-CATAAGATGAACGGAATGAA-3′; lower strand, 5′-ACCGTGCAACTTGTAGTACA-3′) to amplify the *D. melanogaster* GATA finger sequence from pools of cDNA clones. A specific band of 186 bp (62 bp shorter than the genomic PCR product due to the absence of the intervening sequence) could be amplified with cDNA from embryos 0–3, 3–12, and 12–24 h old plus I and II instar larvae, but not from later stages of development (data not shown). This result does not necessarily imply, however, that the expression of dGATAc gene is limited to these stages, as DNA prepared from a cDNA library could easily miss some of the less favorable clones, and not all the cDNA libraries are total representative of transcripts expressed at that particular stage. In fact, an RNase protection assay carried out with a radiolabeled antisense riboprobe detected the expression of dGATAc gene in later stages of *D. melanogaster* development.[2]

We then screened *Drosophila* embryonic cDNA libraries (30) corresponding to 0–3 h and 3–12 h embryonic, and I and II instar larval stages of development. Altogether, eight positive clones were isolated from the 2×10^6 phage plaques screened. Of them, three belong to the dGATAa and were not studied further. Two clones (4 and 11) together represent 3.2 kb of dGATAc cDNA sequence. Another clone, 13.1, is related to dGATAc cDNA, but its 5′ 519-bp sequence is not identical to the other dGATAc clones (Fig. 1*A*). Further analysis of this sequence revealed that the sequence deviation is due to alternative splicing of the dGATAc transcripts.[3]

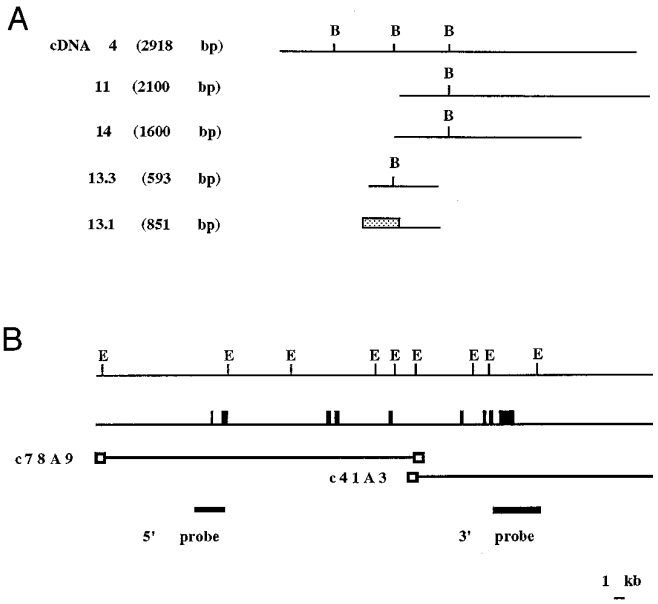

FIG. 1. **Restriction analysis of dGATAc cDNA clones and organization of the dGATAc gene.** *A*, four dGATAc clones were mapped with *Bam*HI. Another clone (13.1) is related to dGATAc but differs in its 5′-half of cDNA sequence (indicated by a *stippled box*). *B*, *Bam*HI. *B*, the position of dGATAc exons is shown relative to *Eco*RI restriction map of the gene. The position of two cosmid clones (c78A9 and c41A3) and probes used for Southern analysis are indicated. The P1 clone (DS01580) mentioned in this paper spans the entire dGATAc gene, but its two ends were not mapped. A detailed description on the isolation of genomic clones, the determination of the cap site and the exon/intron boundaries will be published elsewhere.

Through genomic cloning, restriction mapping, and DNA sequencing, the organization of the dGATAc gene is determined (Fig. 1*B*). The entire transcription unit covers approximately 36 kb of DNA sequence.

Open Reading Frame of the Drosophila GATA cDNA Sequence—The dGATAc cDNA sequence (Fig. 2) was determined by subcloning selected restriction fragments and by the extension of sequences with universal and specific primers. In the 3150-bp insert of dGATAc (including 16 nucleotides from the poly(A) tract), an open reading frame can be predicted by assigning the translation start to the ATG at nucleotide position 224 and the stop codon to position 1682. This open reading frame could encode a polypeptide of 486 amino acids with a predicted molecular mass (M_r) of 51×10^3. As shown in Fig. 2, the predicted dGATAc polypeptide sequence contains two copies of the Cys-Cys type zinc finger that is unique to the GATA factor gene family. Thus, our dGATAc cDNA represents a third expressed GATA gene sequence in *D. melanogaster*.

As the ATG at nucleotide positions 224, 230, and 476 all have an immediate 5′-flanking sequence that is favorable for translation initiation in eukaryotic genes (50), we performed *in vitro* transcription and translation to analyze the translation potential of these open reading frames within the context of the entire cDNA. After subcloning into plasmid pBluescript SK vector, the DNA template was linearized by *Spe*I at nucleotide 2500, downstream from the stop codon at nucleotide 1682. *In vitro* transcription was carried out with T7 polymerase, and the RNA template then used for *in vitro* translation with [^{35}S] methionine. As shown in Fig. 3*A*, SDS-polyacrylamide gel electrophoresis of the synthesized proteins gave a prominent band with an apparent M_r of 52×10^3, a size consistent with that predicted by the open reading frame starting at nucleotide position 224 or 230. Since the gel system we used is unlikely to resolve the small size difference between proteins translated from these two start sites, we provisionally assign the transla-

[2] S.-F. Tsai, unpublished data.
[3] W.-H. Lin and S.-F. Tsai, unpublished results.

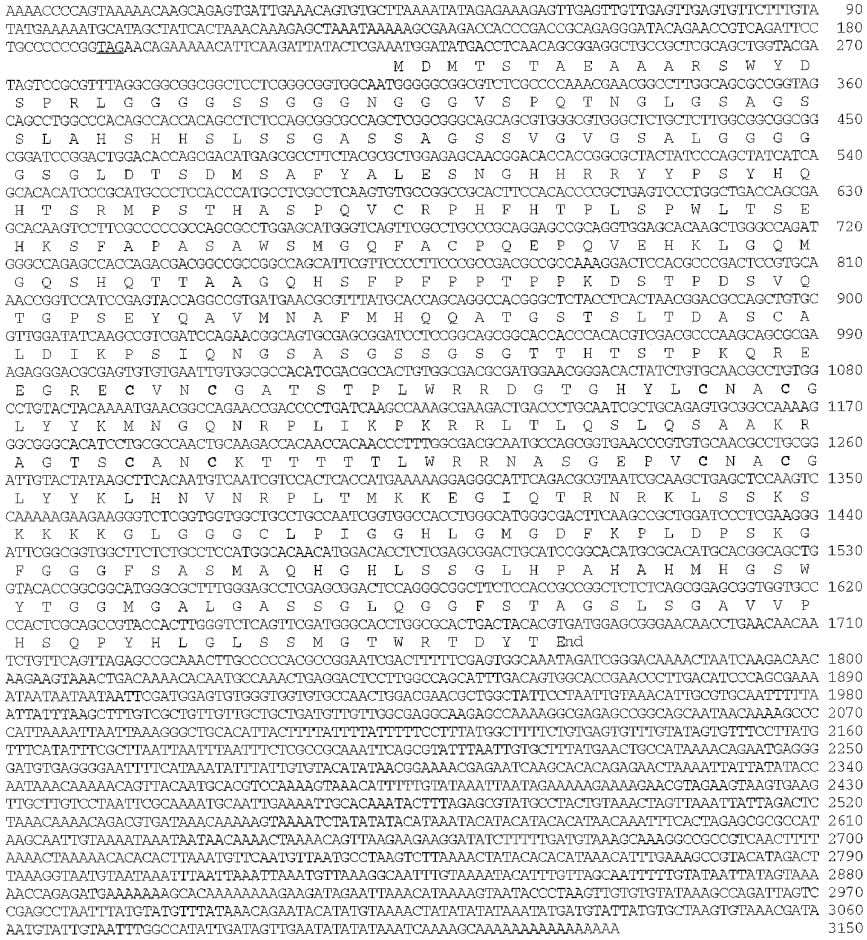

FIG. 2. **DNA sequence of dGATAc and its predicted amino acid sequence.** The ATG at nucleotide position 224–227 is assigned as the translation initiation site. An in-frame termination codon upstream of the initiator ATG is underlined. The cysteine residues of the zinc fingers are in *boldface*.

tion start site to the first AUG at nucleotide position 224.

DNA Binding Activity of Drosophila GATA Proteins—The hallmark for the GATA factors is the so-called GATA zinc finger domain, a 60-amino acid sequence highly conserved among different members of the GATA gene family (for review, see Ref. 51). In vertebrates, the GATA zinc finger sequence is duplicated once, and the two finger domains are separated by 60 amino acids. In contrast, the homologous zinc finger domain of three fungal proteins (22–24) contains only one set of the zinc finger sequence. Previous studies by Martin and Orkin (52) in the mouse GATA-1 demonstrated that, of the two zinc finger domains, the carboxyl finger is necessary for DNA binding and the N-terminal finger contributes to the full specificity and stability of DNA binding. Omichinski *et al.* (53) recently reported that a synthetic peptide containing 59 amino acid residues of chicken GATA-1 (cGATA-1) is sufficient to form a minimal unit capable of interacting with the core DNA consensus sequence. Moreover, NMR structural analysis reveals that a 66-residue fragment from the cGATA-1 zinc finger domain can complex with a 16-bp oligonucleotide containing the GATA target sequence (51). As dGATAa and dGATAc cDNA sequences all predict two zinc finger domains, it would be interesting to test whether *Drosophila* GATA proteins can bind the consensus recognition sequence, particularly that derived from a mammalian globin gene. We, therefore, analyzed the DNA binding activity of the entire dGATAc protein as well as that of the isolated finger region. As suggested by the presence of the conserved GATA finger domain, chimeric proteins containing either the 486-amino acid dGATAc protein or the dGATAc zinc finger domain (nucleotides 940-1350, Fig. 2) can bind γ-globin promoter sequence (Fig. 3, *B* and *C*) and the mouse GATA-1 promoter sequence (data not shown). In both cases, the recombinant proteins specifically bind the GATA sequence. Competition with a homologous GATA sequence effectively abolished the DNA-protein complex, while an unrelated control sequence failed to do so. Similar results were obtained for the dGATAa zinc finger domain (data not shown). Thus, like other animal GATA proteins, the amino acid sequence from the conserved zinc finger region of dGATAa or dGATAc is sufficient to confer DNA binding activity. Moreover, the dGATA finger domains appeared to interact specifically with the recognition sequence derived from a species of distant relationship.

dGATAc Expression in Drosophila Embryos—Although we have shown by cloning that at least two GATA factor genes are present in *D. melanogaster*, and that they are transcribed during embryonic development, the exact functions of their protein products are currently unknown. The fact that the dGATAa and dGATAc genes each encode an amino acid sequence that conforms to the GATA finger consensus sequence suggests that they, too, function as transcriptional factors. Expression of these dGATA factors can be detected as early as within 0–3 h of development. This is consistent with the finding that most GATA factors, as major determinants for tissue development, are expressed in the early embryonic stages. For example, it was shown that mGATA-1 is expressed in the blastocyst stage of mouse embryos (54). Also, the *Xenopus* GATA-1 gene expression was detected around stage 11 of development (55). To clarify the functional role of dGATAc, we used cloned dGATAc fragment as a probe to analyze the distribution of RNA transcripts at different stages of development in whole mount preparations of *Drosophila* embryos. Using samples collected between 0 and 24 h after fertilization, we found that the dGATAc

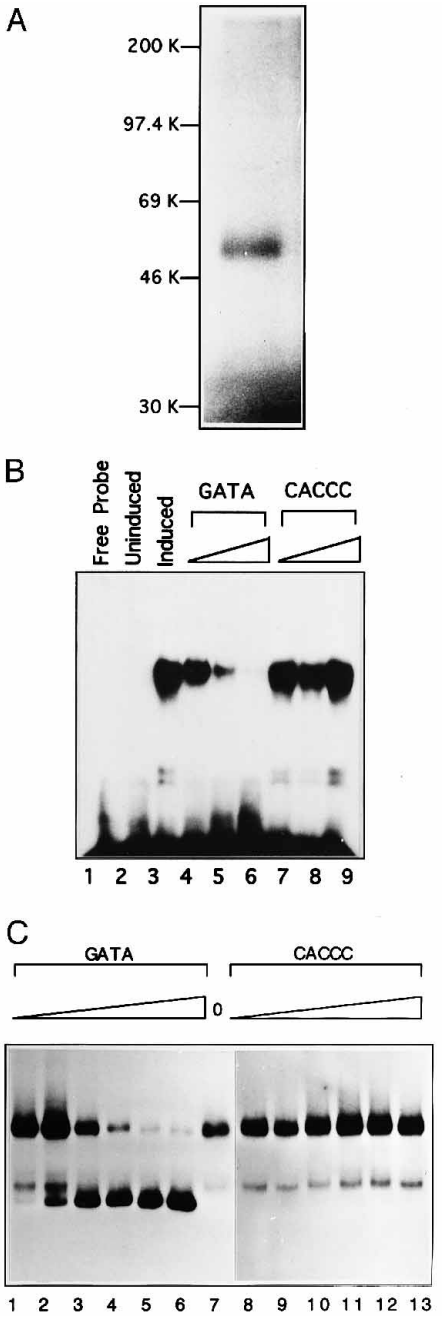

FIG. 3. **Analysis of dGATAc protein.** *A*, *in vitro* transcribed RNA was used to direct the synthesis of dGATAc protein using a rabbit reticulocyte lysate system. The protein product was labeled by [^{35}S] methionine and detected by autoradiography. The size of molecular weight maker proteins is indicated to the left. *B*, specific DNA binding of dGATAc protein to the target site. The entire dGATAc protein was expressed using the pGEX expression system. *E. coli* lysates were prepared from uninduced (*lane 2*) and isopropyl-1-thio-β-D-galactopyranoside-induced (*lane 3*) cultures. Increasing concentration of the GATA-specific competitor (*lanes 4–6*: 5, 25, and 125 ng) and the control CACCC competitor (*lanes 7–9*: 5, 25, and 125 ng) were added to the DNA binding reaction containing approximately 10 ng of end-labeled probe. *Lane 1* contains no protein. *C*, binding of dGATAc finger domain to the target site. Finger domain of dGATAc was expressed using the pGEX vector. Increasing concentration of the GATA-specific competitor (*lanes 1–6*: 6.3, 12.5, 25, 50, 100, and 200 ng) or the CACCC competitor (*lanes 8–13*: 6.3, 12.5, 25, 50, 100, and 200 ng) was added to the DNA binding reaction containing approximately 10 ng of end-labeled probe. *Lane 7* contains no competitor.

gene is expressed in a variety of tissues.

The expression of GATAc transcripts during early *Drosophila* embryos is not detectable until the cellular blastoderm stage. Initially, the RNA transcripts are evenly distributed and concentrated at the basal end of the cells (Fig. 4*A*). Within a short period of time, the transcripts become localized to three regions along the dorsal portion of the embryo (Fig. 4, *B* and *C*). In the procephalic region, the dGATAc gene is abundantly expressed and the transcripts are widely distributed, properly reflecting its later role in the development of the head region. The expressed transcripts are also detectable in the posterior third (15–25% egg length) and middle third (40–60% egg length) of the dorsal embryo. These regions give rise to the precursors of the posterior spiracles and the dorsal epidermis, respectively. In addition, a very faint signal can be seen in a small region of the ventral embryo (Fig. 4*B*, between *two arrows*).

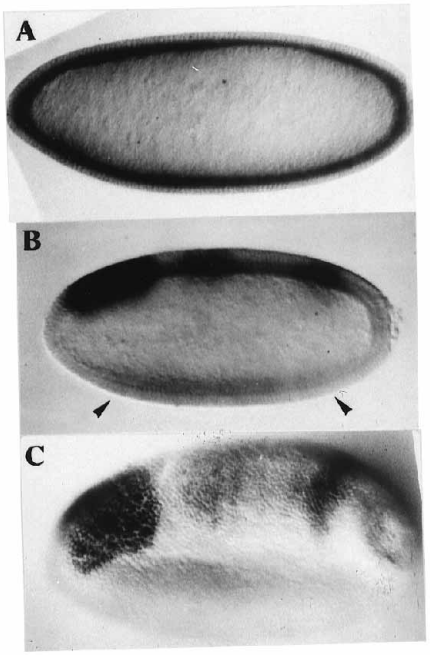

FIG. 4. **Expression of dGATAc transcript during early *Drosophila* development.** Embryos collected from early cellular blastoderm (*A*) and late cellular blastoderm (*B*) are shown on the lateral view. *Left* is anterior and *top* is dorsal. An embryo of early gastrulation stage (*C*) is shown on the dorsolateral view to reveal the distribution of signals in the dorsal portion of the embryo.

As embryonic development reaches stage 11 and beyond, three organ systems clearly stain positive with the dGATAc probe. The developing posterior spiracles are most prominent, and our probe could serve as a useful marker to trace the development of this structure (Figs. 5, *A*, *C*, and *E*). It is noticeable that as germ band shortening occurs, the posterior spiracles moved backward and outward toward their final position. Similarly, the strong but relatively diffuse signals of the anterior and posterior midgut primordia become discrete and approach the middle portion of the embryos (Fig. 5*B*). The expression of dGATAc gene in the developing central nervous system is also seen after stage 11. Distinct signals corresponding to each segment of the embryo become evident (Figs. 5, *B*, *D*, and *G*) at stages 12–13. From the ventral view (Fig. 5*D*), the probe-positive cells for each segment are distributed along both sides of the midline. In the head region, the brain and the developing optic lobes (detail not shown), as well as the anterior tip of the clypeolabrum, are also clearly stained.

Chromosomal Mapping of the dGATAc Gene—The expression of dGATAc in multiple embryonic tissue suggests that dGATAc protein may play an important role in the development of these organ systems and that the lack of dGATAc function could have grave consequences for *Drosophila* embry-

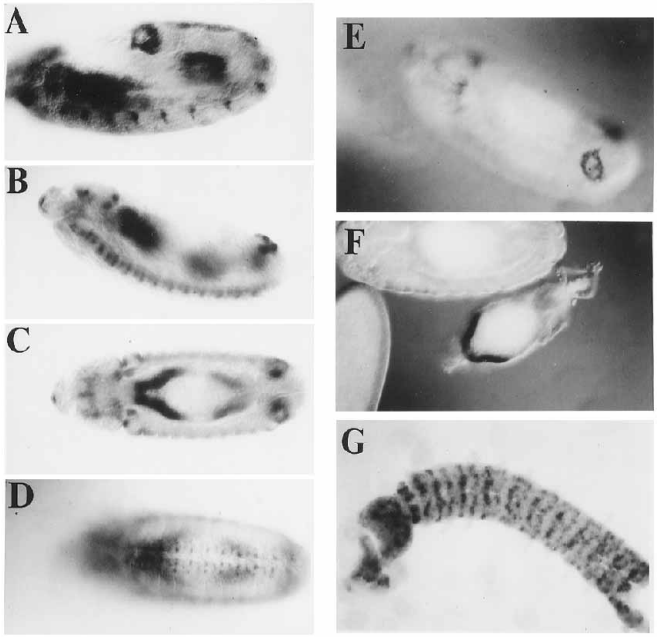

FIG. 5. **Expression of dGATAc transcripts in the developing midgut, posterior spiracles, and the central nervous system.** Embryos of stage 12 (*A*) and stage 13 (*B*) are shown on the lateral view. The same sample shown in *B* is focused on the dorsal region (*C*) to reveal staining in the head region, gut, and posterior spiracles, and on the ventral region (*D*) the developing central nervous system. In *E*, an embryo was focused on one of the posterior spiracles. Additional embryos were dissected to reveal staining of the anterior midgut (*F*) and the central nervous system (*G*).

onic development. To study the functions of dGATAc protein, we have taken the genetic approach and set out to map its chromosomal position by the standard cytogenetic method. Polytene chromosomes were prepared from the larva of *D. melanogaster*. For the probe, a dGATAc cDNA fragment that lies outside the finger region was used to avoid cross-hybridization with other homologous genes. As shown in Fig. 6*A*, a prominent signal was detected at 84F region of the third chromosome. For more exact mapping, several P1 genomic clones (46) corresponding to the same general region (84E–84F) were obtained, and Southern analysis was performed with dGATAc cDNA or genomic probes to identify clones containing the dGATAc gene. Of the five P1 clones, DS01580 derived from bands 84F1–2 was positive with either a probe of the dGATAc finger region (Fig. 6*B*) or with the 3′ genomic probe (Fig. 6*C*). However, when a 5′ genomic probe (3.0-kb *Sac*I-*Eco*RI fragment) containing the first exon was used, it detected *Eco*RI bands of 14 and 8.5 kb in size for the cosmid and P1 DNA, respectively (Fig. 6*D*). Since the same probe also detected a 14-kb band with the total genomic DNA (data not shown), we concluded that the 5′ genomic configuration of the cosmid clone is consistent with that of the genome of the Canton-Special strain *D. melanogaster* we used. We did not pursue this further and investigate the cause of the smaller size of the band detected for the P1 clone. Presumably, however, it could be due to a junctional fragment, a polymorphism, or rearrangement of the cloned sequence. To characterize the promoter region, Southern analysis was also carried out with the P1 and cosmid DNA using either *Sac*I or *Sac*I plus *Eco*RI. In each case, an identical pattern can be seen in these two samples (Fig. 6*E*), suggesting that the difference of genomic structure between the two clones is 5′ to the *Sac*I site and that the entire transcription unit of dGATAc is included in this P1 clone. Therefore, the result of the P1 cloning not only confirmed the polytene *in situ* mapping data, it also localized dGATAc to subsection 84F on the third chromosome.

Disruption of dGATAc Gene Function Causes Embryonic Lethality—On the basis of its chromosomal position, we searched the Flybase for existing strains that bear P-element insertion at and around the 84F region to identify possible mutants that lack the dGATAc function. We then performed genomic Southern analysis with dGATAc cDNA and genomic probes to detect in these strains any rearrangement of the dGATAc gene structure. Of the seven lines obtained from the fly stock center and G. Rubin, we identified two that clearly showed a hybridization pattern that is suggestive of having a P-element integration at the dGATAc locus. As shown in Fig. 7*A*, a 5′ end genomic probe detected a 6.0-kb *Bam*HI band for the wild-type strain (*lane 1*) and all of the P-insertion strains (*lanes 2–8*). Additional bands of 8.0 plus 5.0 kb and 11 plus 4.5 kb were seen for P812 (*lane 2*) and *l*(3)5930 (*lane 7*), respectively. Similarly, in addition to a 5.0 kb common band, extra bands of 7.5 plus 3.5 kb and 15 plus 6.0 kb were detected by the same 5′ genomic probe for *Sac*I-digested samples of *l*(3)5930 and P812, respectively (not shown). As neither the complete cDNA probe nor a 3′ genomic probe revealed any difference among these lines, the structural changes seen in these two strains are limited to the 5′ end, and the P-element insertions probably localize to the dGATAc promoter region. Fortuitously, *l*(3)5930 contains in the P-element construct a lacZ reporter gene, which allows the detection of β-galactosidase expression under the influence of the dGATAc promoter. Thus, by comparing the expression patterns of the lacZ reporter gene and the endogenous dGATAc transcripts, we could obtain functional evidence that the P-element integration is related to dGATAc gene expression.

Both P812 and *l*(3)5930 were originally reported to be homozygous lethal. We collected embryos for 48 h after fertilization to examine their phenotypes. Nearly half of the embryos failed to hatch, and, among these, two groups of dead embryos with distinct morphologies were identified. We presumed that one group was lethal due to the balancer chromosome and the other to the effects of dGATAc gene insertion. To isolate dGATAc homozygous mutant embryos and to confirm that a developmental block can be caused by the dGATAc gene mutation, we generated a *l*(3)5930/+ line by crossing the original stock with the CS wild-type strain. This new line was then self-crossed to obtain a homogeneous pool of mutant embryos carrying two *l*(3)5930 alleles. We presumed, by eliminating the balancer chromosome that also gives a homozygous lethal phenotype, that we could reliably identify the embryonic lethal phenotype due to the *l*(3)5930 mutant chromosome. Indeed, all of the embryos that failed to develop (stopped at stage 17) are homozygous for the mutant allele. As shown in Fig. 7*B*, the genotype of the lethal embryos is distinctly different from that of the original stock and the CS wild-type strain. The 6-kb *Bam*HI band that is diagnostic of the wild-type dGATAc allele is absent in the homozygous mutants (*lane 3*). We conclude by this study that the lethal phenotype of these mutant embryos could be attributed to P-element insertion at the dGATAc gene.

The results of Southern analysis indicated that the P-element insertions in P812 and *l*(3)5930 occurred at the dGATAc promoter region. To precisely map the integration site, we used PCR and plasmid rescue to clone and determine their flanking sequences. As shown in Fig. 8*A*, the 5′ junctional fragment of *l*(3)5930 insertion was obtained through a plasmid rescue experiment. In addition, PCR cloning and sequencing using primers designed for the dGATAc promoter region and the P-element terminal repeat also confirmed the integration site from the 3′ direction. For P812, PCR was carried out using the promoter sequence plus the *white* gene sequence included in the P-element construct (48) to amplify the 3′ junctional se-

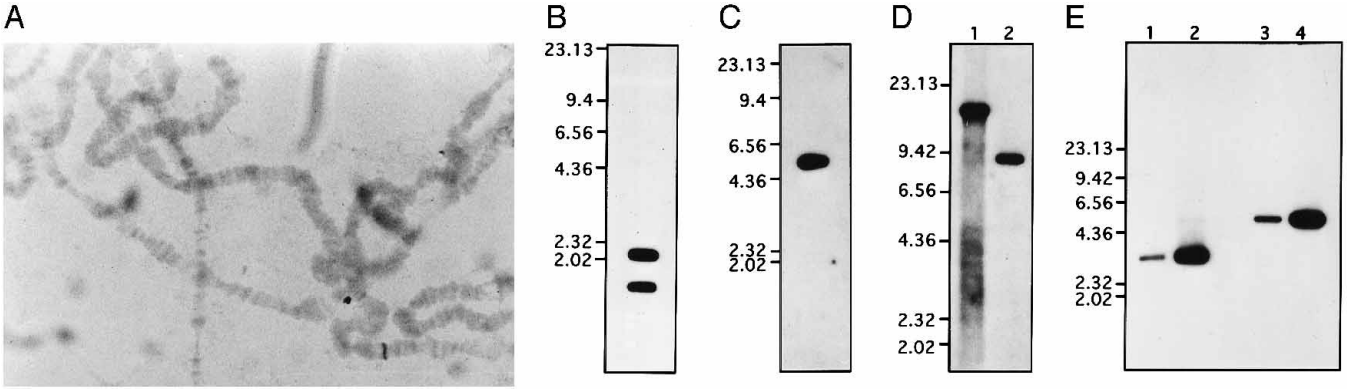

FIG. 6. **Mapping of dGATAc gene and analysis of P1 clones.** *A*, nonradioactive detection of dGATAc chromosomal gene. A prominent band can be seen on the *lower right corner*. *B*, a 420-bp probe from the *Bam*HI/*Eco*RI fragment of dGATA cDNA clone 13.3 detected two *Eco*RI bands of 2.2 kb (exon V) and 1.7 kb (exon VII) for the P1. Another 6.4 kb (exon VI) band is seen with longer exposure. *C*, a 5.3-kb genomic probe (3′ probe, see Fig. 1*B*) derived from the dGATAc cosmid clone detected a 5.3-kb *Eco*RI band for P1. *D*, a 3.0-kb genomic probe (5′ probe, see Fig. 1*B*) detected a 14-kb *Eco*RI band for cosmid DNA (*lane 1*) and an 8.5-kb *Eco*RI band for P1 clone (*lane 2*). *E*, the same 5′ probe detected 3.0 kb band for *Sac*I/*Eco*RI digested cosmid (*lane 1*) and P1 clone (*lane 2*), and 5.0 kb band for *Sac*I-digested cosmid (*lane 3*) and P1 clone (*lane 4*).

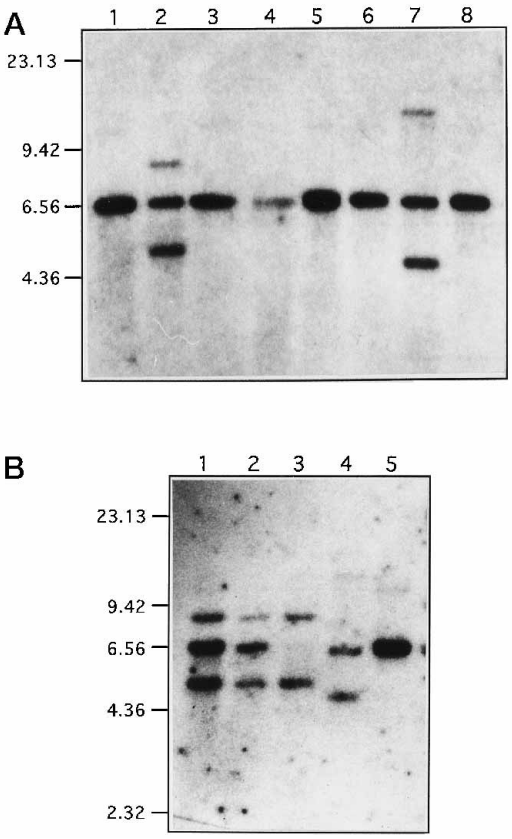

FIG. 7. **Analysis of dGATAc gene mutants.** *A*, identification of P-element insertion mutants of the dGATAc gene. Genomic DNA isolated from eight different *Drosophila* lines were digested with *Bam*HI and probed with a 3.0-kb *Sac*I/*Eco*RI dGATAc genomic fragment. *Lanes 1–8* represent CS wild-type, *l*(3)5930, P971, P227, P609, P614, P812, and 2352, respectively. *B*, correlation of dGATAc mutant genotype with embryonic lethal phenotype. A *l*(3)5930/+ × *l*(3)5930/+ cross was set up to obtain dGATAc homozygous mutants. Southern analysis was performed as in *A* for the original *l*(3)5930 stock (*lanes 1* and *2*), embryos that failed to hatch (*lane 3*), the P812 mutant (*lane 4*), and the CS wild-type strain (*lane 5*).

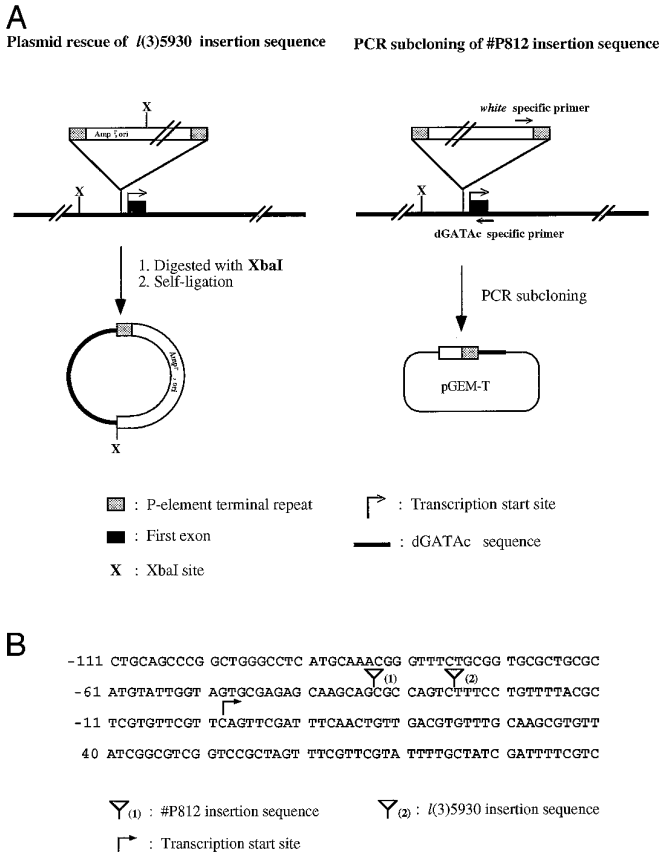

FIG. 8. **Determination of the P-element integration site.** *A*, scheme for cloning the junctional fragments of *l*(3)5930 and P812 P-element insertion strains. *B*, the integration sites for *l*(3)5930 and P812 are marked relative to the dGATAc transcription initiation site. A detailed description of the dGATAc gene structure and the analysis of its promoter will be reported elsewhere.

quence. The results of these studies are summarized in Fig. 8*B*. Interestingly, the two independent P-element insertions are very close to each other, and they are only 26 and 34 bp upstream of the cap site of the dGATAc gene for *l*(3)5930 and P812, respectively.

To validate the assertion that the embryonic lethal phenotype is indeed linked to the lack of dGATAc protein function, we crossed P812 with *l*(3)5930 to determine whether the two mutations are allelic. We expected that the two mutant alleles would not complement each other and that the two insertions, if present in trans, would be lethal to the embryos. The different balancers for P812 and *l*(3)5930 strains provide a convenient way to identify the genotypes of the surviving offspring. All of them, in fact, carry the balancer derived from either one parental strain or both, thus indicating that none of the double mutants were viable (data not shown). We conclude, therefore,

TABLE I
Summary of Drosophila GATA factor genes

Gene	Number of zinc fingers	Map position	Expression pattern (function)	References
dGATAa/*pannier*	2	89A	Along the embryonic dorsoventral axis	(27)
		89B	Regulate *achaete* and *scute* genes	(28)
		89B	ND[a]	Lin and Tsai, unpublished
dGATAb	1	89B	Fat body Regulate *Adh* gene	(26)
dGATAc	2	84F1–2	Optic lobes, central nervous system, midgut, and posterior spiracles	This study

[a] ND, not done.

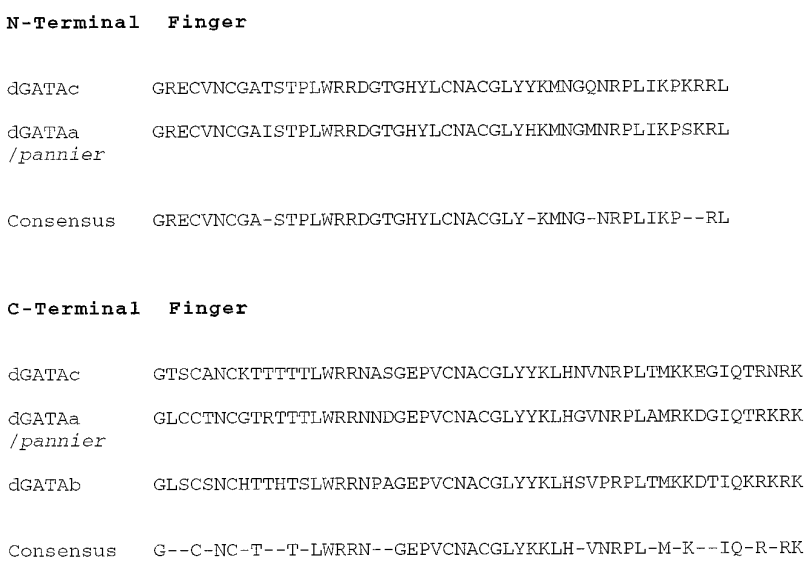

FIG. 9. **Comparison of the amino acid sequences of dGATA zinc fingers.** The predicted zinc finger sequences of dGATAc are aligned with those of dGATAa Winick *et al.* (27) (identical to *pannier* of Ramain *et al.* (28)) and dGATAb cDNA clones (Abel *et al.* (26). The consensus sequence for each finger is shown in the *bottom row*. Identical amino acid residues are indicated by the single-letter abbreviation, and the *dash* represents a conservative change.

that the recessive embryonic lethal mutations on each of the mutant chromosomes belong to the same complementation group. Furthermore, the embryonic lethal phenotype seen in these strains is consistent with the finding that dGATAc transcripts are distributed in multiple yet specific organ systems in the developing embryos (Fig. 5).

The Drosophila GATA Gene Family—So far, at least three functioning *Drosophila* GATA genes have been identified by molecular cloning. Table I gives a summary of these studies. dGATAa (27) (identical to *pannier* (28) and dGATA-I) and dGATAc contain two zinc finger domains, just as is seen in all the vertebrate GATA factor genes. In addition, dGATAb (26) is unique in that, thus far, it is the only single-finger GATA gene known to exist in an animal.

Biochemical studies (52, 53, 56) and NMR structural analysis (51) defined the basic unit of the GATA zinc finger domain that is required for binding to its DNA target sequence. The recent isolation of a single-finger GATA finger protein from *D. melanogaster* (26) strengthened previous findings from *areA* (22), *nit-2* (23), and *gln3* (24) that DNA binding by the single-finger protein requires this 60-amino-acid conserved domain. Comparison of all the available *Drosophila* GATA finger sequences (Fig. 9) reveals that 43 of the 48 (90%) amino acids are identical for the N-terminal fingers of the dGATAa/*pannier* and dGATAc proteins. In contrast, a slightly longer sequence is conserved for the C-terminal dGATAa and C-terminal dGATAc fingers or the single finger of dGATAb. Of the 54 aligned amino acid residues of this region, 38 (70%) are identical, and of the remainder show only conservative change. Interestingly, the four point mutations identified as *pannier* null mutants are restricted to the conserved residues of the N-terminal finger (28).

Function of dGATA Factors—GATA factor genes constitute a unique family whose members encode tissue-specific transcription factors. Their functions are to regulate a group of genes that are important for determining lineage specificity (mGATA-1, for example), or modulating the differential utilization of nutrients under specific growth conditions (*areA*, for example). Taking as precedents the vertebrate GATA factors, we anticipate that different *Drosophila* GATA factors will have distinct roles and that each can regulate the development of specific organ systems. The combined use of molecular cloning and classic genetics provides us with the opportunity to uncover the functions of different GATA factors as well as to understand the mechanism of their action in the *Drosophila* system. Recently, Ramain *et al.* (28) reported the isolation of a GATA-1 finger like sequence from the *pannier* locus. Two types of *pannier* mutants were identified, and one involves the zinc finger domain, causing an over-expression of the *achaete/scute* genes and the development of extra neural precursors. In contrast, mutations affecting a helix at the C-terminal end of the protein generated a hyperactive repressor of the *achaete/scute* complex and caused a loss of neural precursors. They, therefore, concluded that the *pannier* protein is a repressor of the *achaete/scute* gene complex.

Using the expression screening approach, Abel *et al.* (26) isolated the single-finger dGATAb gene from a 9–12-h embryonic cDNA library. Expression of dGATAb is found in the mesoderm-derived fat body, and the encoded protein, ABF, appears to act as an activator of the larval promoter of the alcohol dehydrogenase (*Adh*) gene.

Finally, members of the dGATA factor gene family are expressed in unique but slightly overlapping patterns during embryonic development. Together they could play important roles in tissue-specific gene regulation. However, this begs the question: would there be any functional redundancy or cross-

interaction among different dGATA factors? It is noticeable that dGATAc is expressed in several tissues that were also found to be positive in previous whole mount *in situ* hybridization studies using dGATAa (27) and dGATAb (26) probes. In the late cellular blastoderm and early gastrulation stages, the weak dGATAc signal in the dorsal embryo (40–60% of the egg length) is also distributed in a pattern of stripes (Fig. 4, *B* and *C*), similar to those reported for dGATAa and dGATAb. Interestingly, the dorsal epidermis arising from this region of the embryo is clearly positive with the dGATAc probe at stage 13 of development (data not shown). The signal is particularly strong in segments T1-T3. In addition, dGATAa and dGATAc transcripts are colocalized to the posterior spiracles (Ref. 27 and this work), and the dGATAb gene is expressed transiently in the anterior and posterior midgut primordia (26). We excluded, with two experiments, the possibility that the apparent distribution of dGATAc transcripts in these tissues is an artifact due to cross-hybridization with other homologous dGATA sequences.

Using a shorter probe that lacks the conserved finger sequence, we observed the same pattern of tissue distribution (data not shown). Moreover, we analyzed the enhancer trap line *l*(3)5930 and confirmed by antibody staining that the expression of the lacZ reporter gene is confined to the same multiple organ systems that express dGATAc transcripts.[4] From a biochemical point of view, different dGATA factors share similar DNA binding structures and, through the function of zinc finger domain alone, there would be limited discriminating power to determine target site specificity. Other domains of the dGATA proteins must contribute to the subtle binding site preferences and differential activity necessary for their specific regulatory functions. This could be achieved by intrinsic affinity to different target sequences or by interacting with other regulatory proteins. It remains to be seen whether different dGATA factors can regulate a common gene with different effects. Because *Drosophila* is convenient for genetic studies, the functional role of each dGATA factor can be definitively identified by analyzing the phenotypes of the mutants. The cloning of multiple dGATA genes, the precise mapping of their chromosomal locations, and the identification and characterization of the dGATA gene mutants should facilitate our future studies in this direction.

Acknowledgments—We thank Shiau-Chen Chan and Yuh-Long Chang for excellent technical assistance and Ricky Yu for preparing P1 DNA. We also thank Mel Green, Der-Hua Huang, and Tsung-Sheng Su for critically reviewing the manuscript. We are indebted to Dr. Viatcheslav N. Bolshakov, (Institute of Molecular Biology and Biotechnology, Heraklion, Crete, Greece) for the interpretation of polytene chromosome mapping results and Dr. Tsuneyuki Yamazaki (Kyushu University, Fukuoka, Japan) for providing P1 clones. The help of Der-Hua Huang (Academia Sinica, Taipei), Bill Chia and Xiaohang Yang (Institute of Molecular and Cell Biology, Singapore) in interpreting whole mount embryo *in situ* hybridization results is greatly appreciated. We would also like to thank the Genetic Screen, the National Institutes of Health Genome Center, and the Howard Hughes Medical Institute for the use of *l*(3)5930 line.

REFERENCES

1. Wall, P., deBoer, E., and Grosveld, F. (1988) *Genes & Dev.* **2,** 1089–1100
2. Evans, T. M., Reitman, T., and Felsenfeld, G. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 5976–5980
3. Martin, D. I. K., Tsai, S.-F., and Orkin S. H. (1989) *Nature* **338,** 435–438
4. Tsai, S.-F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G., and Orkin, S. H. (1989) *Nature* **339,** 446–451
5. Evans, T., and Felsenfeld, G. (1989) *Cell* **58,** 877–885
6. Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., and Engel, J. D. (1990) *Genes & Dev.* **4,** 1650–1662
7. Zon, L. I., Tsai, S.-F., Burgess, S., Matsudaira, P., Bruns, G. A. P., and Orkin, S. H. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 668–672
8. Trainor, C. D., Evans, T., Felsenfeld, G., and Boguski, M. S. (1990) *Nature* **343,** 92–96
9. Dorfman, D. M., Wilson, D. B., Bruns, G. A. P., and Orkin, S. H. (1992) *J. Biol. Chem.* **267,** 1279–1285
10. Lee, M.-E., Temizer, D. H., Clifford, J. A., and Quertermous, T. (1991) *J. Biol. Chem.* **266,** 16188–16192
11. Ho, I.-C., Vorhees, P., Marin, N., Oakley, B. K., Tsai, S.-F., Orkin, S. H., and Leiden, J. M. (1991) *EMBO J.* **10,** 1187–1192
12. Ko, L. J., Yamamoto, M., Leonard, M. W., George, K. M., Ting, P., and Engel, J. D. (1991) *Mol. Cell. Biol.* **11,** 2778–2784
13. Joulin, V., Bories, D., Eleouet, J.-F., Labastie, M.-C., Chretien, S., Mattei, M.-G., and Romeo, P. H. (1991) *EMBO J.* **10,** 1809–1816
14. Marine, J., and Winoto, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 7284–7288
15. Zon, L. I., Mather, C., Burgess, S., Bolce, M. E., Harland, R. M., and Orkin, S. H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10642–10646
16. Orkin, S. H. (1990) *Cell* **63,** 665–672
17. Kelly, C., Blumberg, H., Zon, L. I., and Evans, T. (1993) *Development* **118,** 817–827
18. Arceci, R. J., King, A. A. J., Simon, M. C., Orkin, S. H., and Wilson, D. B. (1993) *Mol. Cell. Biol.* **13,** 2235–2246
19. Orkin, S. H. (1992) *Blood* **80,** 575–581
20. Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S.-F., D'Agati, V., Orkin, S. H., and Costantini, F. (1991) *Nature* **349,** 257–260
21. Tsai, F.-Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., and Orkin, S. H. (1994) *Nature* **371,** 221–226
22. Kulda, B., Caddick, M. X., Langdon, T., Martinez-Rossi, N. M., Bennett, C. F., Sibley, S., Davies, R. W., and Arst, H. N., Jr. (1990) *EBMO J.* **9,** 1355–1364
23. Fu, Y.-H., and Marzluf, G. A. (1990) *Mol. Cell. Biol.* **10,** 1056–1065
24. Minehart, P. L., and Magasanik, B. (1991) *Mol. Cell. Biol.* **11,** 6216–6228
25. Spieth, J., Shim, Y. H., Lea, K., Conrad, R., and Blumenthal, T. (1991) *Mol. Cell. Biol.* **11,** 4651–4659
26. Abel, T., Michelson, A. M., and Maniatis, T. (1993) *Development* **119,** 623–633
27. Winick, J., Abel, T., Leonard, M. W., Michelson, A. M., Chardon-Loriaux, I., Holmgren, R. A., Maniatis, T., and Engel, J. D. (1993) *Development* **119,** 1055–1065
28. Ramain, P., Heitzler, P., Haenlin, M., and Simpson, P. (1993) *Development* **119,** 1277–1291
29. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science* **239,** 487–491
30. Poole, S. J., Kauvar, L. M., Drees, B., and Kornberg, T. (1985) *Cell* **40,** 37–43
31. Hoheisel, J. D., Lennon, G. G., Zehetner, G., and Lehrach, H. (1991) *J. Mol. Biol.* **220,** 903–914
32. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., and Efstratiadis, A. (1978) *Cell* **15,** 687–701
33. Chen, E. Y., and Seeburg, P. H. (1985) *DNA* (*N. Y.*) **4,** 165–170
34. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
35. Tabor, S., and Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 4767–4771
36. Queen, C., and Korn, L. J. (1984) *Nucleic Acids Res.* **12,** 581–599
37. Devereux, J. R., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395
38. Tautz, D., and Pfeiffle, C. (1989) *Chromosoma* **98,** 81–85
39. Campos-Ortega, J. A., and Hartenstein, V. (1985) *The Embryonic Development of Drosophila melanogaster.* Springer-Verlag, Berlin, Germany
40. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056
41. Ausubel, F. M., Grent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1993) *Current Protocols in Molecular Biology.* John Wiley and Sons, New York
42. Smith, D. B., and Johnson, K. S. (1988) *Gene* (*Amst.*) **67,** 31–40
43. Tsai, S.-F., Strauss, E., and Orkin, S. H. (1991) *Genes & Dev.* **5,** 919–931
44. Pardue M. L. (1986) in *Drosophila, a Practical Approach* (Roberts, D. B., ed) pp. 111–137, IRL Press, Oxford
45. Feinberg, A. P., and Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13
46. Smoller, D. A., Petrov, D., and Hartl, D. L. (1991) *Chromosoma* **100,** 487–494
47. Brinboim, H. D., and Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523
48. Karpen, G. H., and Spradling, A. C. (1992) *Genetics* **132,** 737–753
49. Lindsley, D., and Zimm, G. (1992) *The genome of Drosophila melanogaster,* Academic Press, San Diego, CA
50. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241
51. Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., and Gronenborn, A. M. (1993) *Science* **261,** 438–446
52. Martin, D. I. K., and Orkin, S. H. (1990) *Genes & Dev.* **4,** 1886–1898
53. Omichinski, J. G., Trainor, C., Evans, T., Gronenborn, A. M., Clore, G. M., and Felsenfeld, G. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 1676–1680
54. Whitelaw, E., Tsai, S.-F., Hogben, P., and Orkin, S. H. (1990) *Mol. Cell. Biol.* **10,** 6596–6606
55. Zon, L. I., Mather, C., Burgess, S., Bolce, M. E., Harland, R. M., and Orkin, S. H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10642–10646
56. Yang, H.-Y., and Evans, T. (1992) *Mol. Cell. Biol.* **12,** 4562–4570

[4] J.-Y. Yeh and S.-F. Tsai, unpublished results.