MOLECULAR AND CELLULAR BIOLOGY, May 1991, p. 2778–2784
0270-7306/91/052778-07$02.00/0

Vol. 11, No. 5

Murine and Human T-Lymphocyte GATA-3 Factors Mediate Transcription through a *cis*-Regulatory Element within the Human T-Cell Receptor δ Gene Enhancer

LINDA J. KO,[1] MASAYUKI YAMAMOTO,[2] MARK W. LEONARD,[1] KATHLEEN M. GEORGE,[1] PAUL TING,[1] AND JAMES DOUGLAS ENGEL[1]*

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208-3500,[1] and Department of Biochemistry, Tohoku University School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai 981, Japan[2]

Received 14 December 1990/Accepted 26 February 1991

A family of transcriptional activators has recently been identified in chickens; these transcriptional activators recognize a common consensus motif (WGATAR) through a conserved C_4 zinc finger DNA-binding domain. One of the members of this multigene family, cGATA-3, is most abundantly expressed in the T-lymphocyte cell lineage. Analysis of human and murine GATA-3 factors shows a striking degree of amino acid sequence identity and similar patterns of tissue specificity of expression in these three organisms. The murine and human factors are abundantly expressed in a variety of human and murine T-cell lines and can activate transcription through a tissue-specific GATA-binding site identified within the human T-cell receptor δ gene enhancer. We infer that the murine and human GATA-3 proteins play a central and highly conserved role in vertebrate T-cell-specific transcriptional regulation.

Cellular differentiation during hematopoiesis is the result of a series of hierarchical responses to a complex extracellular, cytoplasmic, and nuclear signaling pathway finally leading to the induction of genes unique to the erythroid, myeloid, or lymphoid cell lineages. Terminal lineage specification appears to be mediated by the binding of nuclear *trans*-acting regulatory protein factors, some of which are uniquely expressed in distinct hematopoietic cell types, to *cis*-regulatory elements of a given lineage-specific gene. One of the best understood of these lineage-specific loci is that encoding β-globin, where both distal and proximal *cis*-regulatory sequences have been demonstrated to bind tissue-specific and ubiquitous *trans*-acting factors to elicit the characteristic erythroid-specific expression of this gene.

Recently a great deal of interest has arisen from the discovery that transcription factor GATA-1 (29) (previously designated GF-1, NF-E1, Eryf1, or EF-1 [8, 22, 31, 44]) appears to be a major regulator of erythroid-specific transcription. GATA-1 was originally identified as an erythroid-specific DNA-binding protein (8, 22) and implicated in the transcriptional activation of both globin and nonglobin (24, 30, 31) erythrocyte-specific genes. The cloning of GATA-1 cDNAs has provided further insight into the properties of this factor (7, 43). The protein binds to a WGATAR consensus sequence site through two similar $C\text{-}X_2\text{-}C\text{-}X_{17}\text{-}C\text{-}X_2\text{-}C$ zinc fingers which differentially contribute to the specificity for and affinity of DNA binding (21). Domains both amino and carboxy terminal to these zinc fingers contribute to transcriptional activation by mGATA-1 (21). Although the factor was originally identified as erythroid specific, it has subsequently been found to also be expressed in megakaryocytes and mast cells of myeloid origin (23, 35).

GATA-1 has recently been shown to be but one member of a multigene family in chickens (originally designated cNF-E1a, cNF-E1b, and cNF-E1c; now termed cGATA-1, cGATA-2, and cGATA-3, respectively [29]), with each family member exhibiting a unique pattern of tissue-restricted expression (47). Members of the cGATA (chicken GATA) family of proteins are related by virtue of their zinc fingers, which show a high degree of identity to other members of the family both within species and between species (42, 47, 48). All of the factors recognize the common consensus binding site (WGATAR) and function as transcriptional activators in vivo (47).

Of particular interest was the observation that cGATA-3 mRNA is predominantly expressed in chicken T lymphocytes and embryonic brain tissue (47). Here we demonstrate that the human and murine homologs of this factor (hGATA-3 and mGATA-3, respectively) are abundantly expressed in a variety of T-lymphocyte cell lines and that the conceptually translated proteins share a striking degree of amino acid sequence identity with cGATA-3 and with one another. We show that the mRNA expression patterns of these factors are also highly conserved, suggesting that the proteins fulfill a critical function in the activation of T-lymphocyte-specific genes. The identification of GATA-binding sites in the human T-cell receptor (TCR) δ gene enhancer, as well as the further demonstration that these sites are capable of conferring GATA-3-dependent transcriptional induction to reporter genes linked in *cis*, underscores the likelihood that GATA-3 plays a highly conserved, functionally important role in T-lymphocyte-specific transcriptional regulation in vertebrates.

MATERIALS AND METHODS

Cell lines. Cell lines were purchased from the American Type Culture Collection (T-cell lines A0DH, BW5147.3, MOLT-3, CCRF-CEM, CCRF-HSB-2, and HuT78) or were gifts from V. Patel (murine lines MEL, BaF3, 38B9, PD31, and MPC11), S. K. Pierce (murine lines TP*c* and CH27), or S. T. Rosen (human B-lymphoma lines IM9 and U266). The murine TCR $\alpha\beta^+$ T-cell lines A0DH and TP*c* and the TCR

* Corresponding author.

$\alpha\beta^-$ T-cell line BW5147.3 were grown in Dulbecco's modified Eagle's medium (DMEM). The human TCR^- T-cell lines MOLT-3 and HuT78 were grown in RPMI 1640 plus 10% fetal bovine serum (FBS), and the human TCR^- T-cell line CCRF-HSB-2 was grown in Iscove's DMEM plus 10% FBS; the TCR $\alpha\beta^+$ human T-cell line CCRF-CEM was grown in RPMI 1640 plus 20% FBS. QT6 (transformed quail fibroblast) cells were grown in DMEM plus 5% FBS plus 2% chicken serum.

Isolation and characterization of mGATA-3 and hGATA-3 cDNAs. cDNA libraries prepared from RNA derived from the murine EL4 and human Jurkat TCR $\alpha\beta^+$ T-lymphocyte cell lines (Stratagene and Clontech, respectively) were screened (standard hybridization, but reduced washing stringency: 0.5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate] plus blot wash buffer at 37°C [47]) with random primer-radiolabeled cGATA-3 (p30c [47]) as the hybridization probe (2). Approximately 5×10^6 PFU of each library was screened; 68 and 93 human and murine positives, respectively, were identified. Twenty-four of each were plaque purified, and those with the largest cDNA inserts were subcloned and sequenced. The DNA sequence of the largest isolated murine (cDNA clone mc5b) and human (cDNA clone p8h) homologs of the cGATA-3 factor was determined by the dideoxy-chain termination method on serially deleted subclones in pGEM7Zf(+) (Promega, Inc.) (17, 47).

RNA blot analysis. Either 8 μg of total RNA isolated from various mouse organs, tissues, or cell lines or 10 μg of total RNA from either human B-lymphoma or T-lymphocyte cell lines was denatured, electrophoresed, and blotted to nylon filters (34). The filters were then hybridized to random primer-labeled mGATA-3 cDNA clone mc5b or hGATA-3 clone p8h2, respectively, washed at moderate stringency (50 mM Tris-HCl [pH 7.5]–1 mM EDTA–1 mM sodium phosphate–0.1% sodium dodecyl sulfate at 55°C three times for 1 h each time), and exposed for autoradiography. Exposure times are given in the figure legends.

Gel mobility shift assays. Whole-cell extracts were prepared as previously described (43, 47) and incubated with radiolabeled oligonucleotides (amounts specified in figure legends) corresponding to the GATA site sequences of either the mouse α1 globin promoter (MαP [31]) or the human TCR δ gene enhancer (δE4 [33]; see Results). The δE4 sequence is as follows:

```
TCGACACTTGATAACAGAAAGTGATAACTCT
    GTGAACTATTGTCTTTCACTATTGAGAAGCT
```

DNA binding reactions and electrophoresis were performed as described previously (47).

Cotransfection *trans*-activation assays. The rabbit β-globin TATA box and transcription initiation site was used to direct transcription of the human growth hormone gene (see Fig. 6A) in plasmid p0GH (Nichols Institute Diagnostics, San Juan Capistrano, Calif. [43]). To this minimal promoter construct, either one or three copies of an oligonucleotide corresponding in sequence to the TCR δ gene enhancer footprint 4 (δE4 [33]) or three copies of an oligonucleotide corresponding to the chicken β-globin enhancer footprint 4 (βE-F4 [11]) were inserted directly 5′ to the TATA box in the orientation specified in Fig. 6 to produce the H1δGH, H3δGH and C3βGH reporter plasmids, respectively (see Fig. 6A). Activator plasmids R/mGATA-3 and R/hGATA-3 were prepared by sense insertion of the cDNA segments of mGATA-3 clone mc5b or hGATA-3 clone p8h2, respectively, into the unique *Hin*dIII site of RSV · CAT (14) after the chloramphenicol acetyltransferase-coding sequences were removed. The transcription of the recombinant GATA-3 cDNAs is therefore directed by the Rous sarcoma virus promoter and enhancer.

Immortalized quail fibroblasts (QT6 cells [25]) were used as recipient cells in activation assays as previously described (13, 47). The cells were transfected with 8 μg of reporter gene plasmid and 2 μg of (potential) activator plasmid by the $CaPO_4$ method, and the supernatants were assayed for secreted human growth hormone 4 days posttransfection by radioimmunoassay (RIA) (using the Allegro hGH system from Nichols Institute Diagnostics [47]). After the supernatants were collected for RIA, the cells were trypsinized, counted, and lysed to prepare extracts for gel mobility shift assays (43). Counts derived from the hGH RIA were normalized for cell numbers, and *trans*-activation was calculated (47) according to the equation *trans*-activation = $[(A + R) - M]/(R - M)$, where $A + R$ represents the counts recovered per minute in the RIA after transfection with activator and reporter plasmids, M represents the counts per minute in the medium alone, and R represents the counts recovered per minute after transfection with the reporter plasmid alone.

Nucleotide sequence accession numbers. The nucleotide sequence in Fig. 2 is listed in the GenBank data base under accession number X55123; the sequence in Fig. 3 is listed in GenBank under accession number X55122.

RESULTS

Murine and human T lymphocytes contain an abundant GATA-binding protein. With the identification of T lymphocytes as the major sites of cGATA-3 expression (47), we initially sought to determine whether human and murine T lymphocytes express a GATA-binding protein. Extracts were prepared from a variety of murine T-lymphocyte cell lines and incubated with a 30-bp oligonucleotide probe containing a consensus GATA-binding site from the mouse α1 globin promoter (MαP [43]) in a gel mobility shift assay (9, 12). The complex produced with a protein present in the extracts has a slightly lower mobility than that formed with the cGATA-3 protein and is specifically eliminated by unlabeled MαP oligonucleotide in the binding reaction (Fig. 1A). Extracts were also prepared from several human T-cell lines and tested in mobility shift assays with the MαP probe. The strong signal produced by complex formation indicates the presence of an abundant GATA-binding protein in human T lymphocytes; this complex is also specifically inhibited by an excess of unlabeled MαP oligonucleotide in the binding reaction (Fig. 1B). Thus, murine and human T-lymphocyte cell lines expressing either an αβ TCR or no TCR synthesize an abundant protein which binds with high affinity to the GATA consensus sequence.

Cloning and characterization of murine and human GATA-3 cDNAs. With the knowledge that a prominent GATA-binding protein is expressed in mammalian T cells, we then isolated the murine and human homologs of cGATA-3. cDNA libraries prepared from murine EL4 and human Jurkat T-lymphocyte cell lines were screened with the full-length cGATA-3 cDNA clone p31a (2, 47). Independent clones were isolated from each library, and mGATA-3 and hGATA-3 cDNA clones of 2,054 and 2,030 nucleotides, respectively, were sequenced (Fig. 2 and 3, respectively) (17). Both cDNA clones contain complete coding regions; the open reading frame of the mGATA-3 cDNA clone mc5b is predicted to encode a protein of 443 amino acids with a

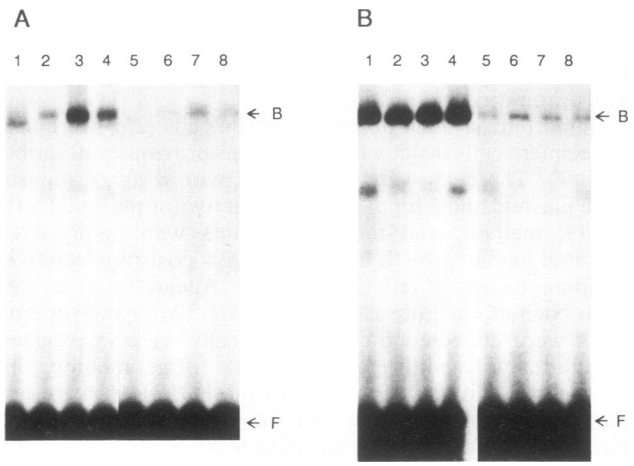

FIG. 1. A protein binding to the WGATAR motif is prevalent in murine and human T-lymphocyte cell lines. (A) Whole-cell extracts were prepared from murine T-cell lines and examined by gel mobility shift assay for binding to 1 ng of radiolabeled MαP oligonucleotide (43, 47). The T-lymphocyte cell lines used for extract preparation were A0DH (lanes 2 and 6), BW5147.3 (lanes 3 and 7), and TP*c* (lanes 4 and 8). For comparison, cGATA-3 as expressed in chicken embryo fibroblast cells from a recombinant retrovirus (6a) is shown (lanes 1 and 5). Lanes 1 to 4 contained 1.5 μg of nonspecific poly(dI-dC)/poly(dA-dT) competitor, while lanes 5 to 8 contained the same plus 150 ng of unlabeled MαP oligonucleotide. B and F refer to the positions of bound and free MαP oligonucleotide, respectively. (B) Extracts were prepared from various human T-cell lines (43) and examined for specific DNA binding to the radiolabeled MαP oligonucleotide by gel mobility shift assay. The cell lines used were MOLT-3 (lanes 1 and 5), CCRF-CEM (lanes 2 and 6), CCRF-HSB-2 (lanes 3 and 7), and HuT78 (lanes 4 and 8). Lanes 1 to 4 contained 1.5 μg of nonspecific competitor DNA, while lanes 5 to 8 contained the same plus 150 ng of unlabeled MαP oligonucleotide. Once again, B and F indicate the positions of bound and free radiolabeled oligonucleotide, respectively.

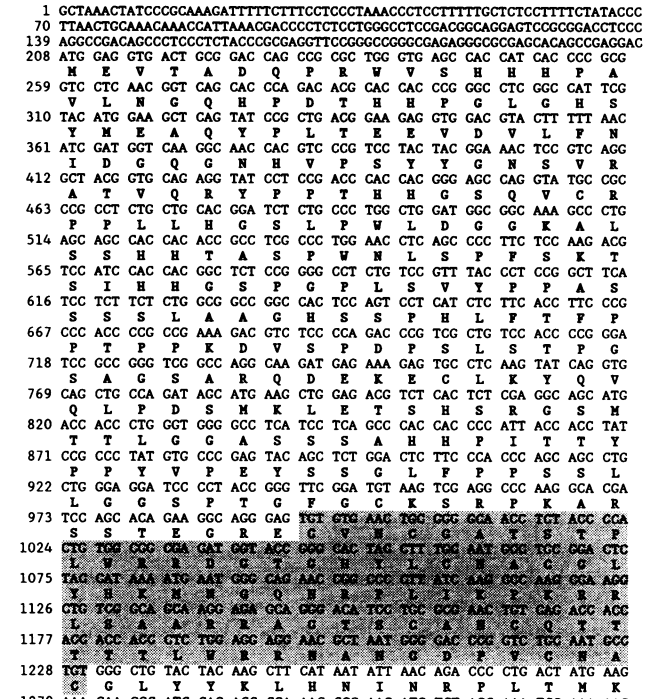

FIG. 2. Sequence of the mGATA-3 cDNA. The confirmed nucleotide sequence of cDNA clone mc5b is shown. Numbers on the left indicate nucleotides beginning at the 5′ end of the insert. The shaded area corresponds to the zinc finger domain and the sequences encoding it.

molecular mass of 47,964 Da (pI, 9.89), and the hGATA-3 clone p8h contains a single long open reading frame encoding a 444-amino-acid protein with a predicted molecular mass of 47,973 Da (pI, 9.84). Both of these are similar in size and isoelectric point to the 48,194-Da molecular mass and pI of 9.89 predicted for the cGATA-3 protein (47).

Vertebrate GATA-3 proteins are highly conserved. A comparison of the deduced amino acid sequences of the chicken, human, and murine proteins (Fig. 4) illustrates the striking identity between the GATA-3 factors of these species. cGATA-3 and hGATA-3 have >92% overall identity; the amino-terminal residues (1 to 215) share 85% identity, while the carboxy-terminal amino acids (216 to 444) share 97% identity. Similarly, cGATA-3 and mGATA-3 show 92% overall sequence identity. In contrast to this high degree of conservation, the erythroid/megakaryocytic/mast (23, 35) cell-specific family member GATA-1 is far less well conserved between these same species, having only 41% overall identity between chickens and humans (42). The GATA-3 proteins also share far greater identity across species boundaries than they do with the GATA-1 or GATA-2 protein within the same species (data not shown).

Tissue distribution of the murine and human GATA-3 transcripts. Expression of murine and human GATA-3 was examined by RNA blot analysis. As shown in Fig. 5A, mGATA-3 is expressed as an approximately 4-kb mRNA and accumulates to the greatest abundance in T lymphocytes, as is the case with cGATA-3 (47). A large number of other tissues and cell types, including the murine B-cell lines 38B9 and PD31 and the myeloma MPC11 cell line, do not express detectable mGATA-3. mGATA-3 mRNA is found in murine placenta and fetal brain at an approximately threefold lower abundance than in T cells, whereas it is undetectable in RNA isolated from adult brain. Thus, the GATA-3 expression profiles in murine and chicken cells are similar in that the sites of the most abundant expression are the embryonic brain and T lymphocytes (47).

A more restricted examination was undertaken to investigate the activity and tissue specificity of the human GATA-3 mRNA. A 3.6-kb mRNA is detected in human T, but not B, lymphocytes by RNA blot analysis (Fig. 5B). hGATA-3 is also expressed in at least one human neuroblastoma cell line, SKNSH (data not shown). Thus, on the basis of examination of a limited number of human cell types, we conclude that the expression characteristics of chicken, human, and murine GATA-3 mRNAs appear to be highly conserved.

hGATA-3 and mGATA-3 *trans*-activate expression from a

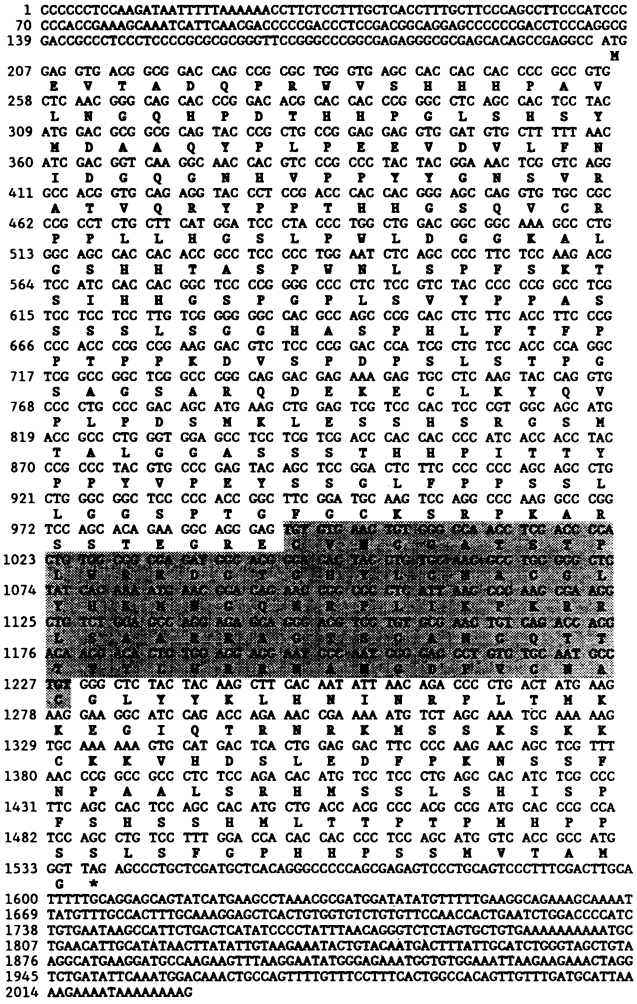

FIG. 3. Sequence of the hGATA-3 cDNA. The confirmed nucleotide sequence of cDNA clone p8h is shown. Numbers on the left indicate nucleotides beginning at the 5′ end of the cDNA insert. The shaded area corresponds to the zinc finger domain and the sequences encoding it.

tissue-specific human TCR δ gene enhancer GATA-binding site. Given the degree of amino acid sequence identity between the chicken, human, and murine GATA-3 proteins, we anticipated that they should each be capable of stimulating transcription from consensus GATA-binding sites. Redondo et al. have described detailed genetic and biochemical experiments examining the human TCR δ gene enhancer (33). Within that *cis*-regulatory domain, one footprint (δE4) lies over two consensus GATA-binding sites. We therefore sought to assess the ability of the GATA-3 proteins to stimulate transcription of a reporter gene containing δE4 to ask whether this site is capable of being regulated by the hGATA-3 protein in vivo.

For cotransfection *trans*-activation assays, the human and murine GATA-3 cDNAs were placed under the transcriptional control of the Rous sarcoma virus long terminal repeat. Reporter plasmids with GATA-binding sites upstream of the rabbit β-globin TATA box were cotransfected with presumptive activator plasmids into quail fibroblast (QT6) cells. We then determined whether the encoded

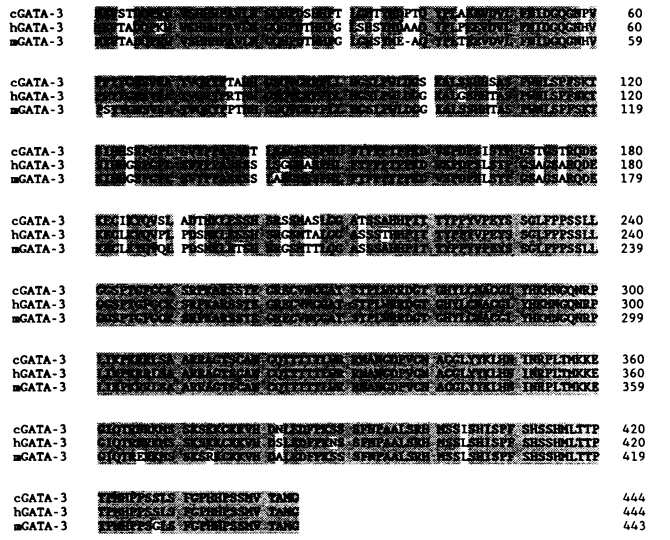

FIG. 4. Amino acid sequence comparison of the chicken, human, and murine GATA-3 factors. The predicted amino acid sequences of cGATA-3, hGATA-3, and mGATA-3 were aligned, and identities with the conceptually translated cGATA-3 protein (47) are shaded. Numbers on the right correspond to the position of the last amino acid on each line.

proteins (mGATA-3 or hGATA-3) could stimulate expression of the human growth hormone gene in *trans*.

The H3δGH and H1δGH reporter gene constructs contain either three or one of the δE4 sites of the human TCR δ gene enhancer, respectively; each δE4 oligonucleotide contains

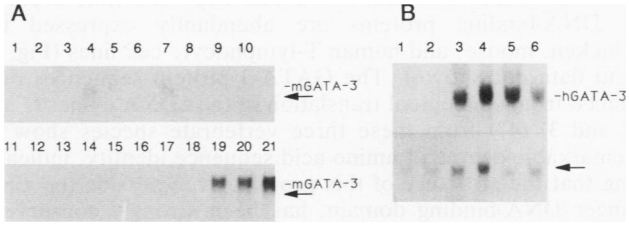

FIG. 5. Expression of mGATA-3 and hGATA-3 mRNAs. (A) RNA isolated from mouse organs, tissues, and cell lines was probed with random primer-labeled mGATA-3 cDNA clone mc5b. The RNA samples shown were isolated from the following tissues or cells: uterus (lane 1); ovary (lane 2); testes (lane 3); placenta (lane 4); liver (lane 5); heart (lane 6); fetal brain (lane 7); maternal brain (lane 8); ES cells (lane 9); 10-day-old embryoid bodies (lane 10); uninduced MEL cells (lane 11); MEL cells induced with dimethyl sulfoxide for 4 days (lane 12); BaF3 (pre-B/pre-T lymphocytes; lane 13); the B-cell lines CH27 (lane 14), 38B9 (lane 15), PD31 (lane 16), and MPC11 (lane 17); normal murine B lymphocytes (lane 18); and the T-cell lines A0DH (lane 19), BW5147.3 (lane 20), and TPc (lane 21). The position of hybridization of the upper band to the murine β-tubulin internal control is indicated by the arrow. The exposure time was 48 h. (B) Blot analysis was performed with RNA isolated from either human B-lymphoma (IM9 [lane 1] and U266 [lane 2]) or T-lymphocyte (CCRF-CEM [lane 3], CCRF-HSB-2 [lane 4], HuT78 [lane 5], and MOLT-3 [lane 6]) cell lines. The position of hybridization of radiolabeled cDNA clone p8h is indicated at the right. Hybridization to the upper band of the human β-tubulin RNA (probed with the murine β-tubulin cDNA clone) internal marker is indicated by the arrow in the lower panel; the size of hGATA-3 was also determined by comparison with an RNA ladder (Bethesda Research Laboratories). The exposure time was 46 h.

two GATA-binding sites in the same orientation (as encountered in the human TCR δ enhancer [33]). The C3βGH reporter construct has three copies of the chicken β-globin enhancer footprint 4 region (which also contains two GATA-binding sites in inverted orientation [6, 11, 47]). The levels of growth hormone secreted by the transfected cells into the tissue culture medium were determined by RIA.

Both mGATA-3 and hGATA-3 are capable of stimulating transcription of the growth hormone reporter gene linked in *cis* to GATA-binding sites from each of the different reporter constructs (Fig. 6A). Although six GATA-binding sites (as found in H3δGH and C3βGH) activate transcription to a greater extent than two sites (in H1δGH), the number of sites alone cannot be the sole determinant of activity, since the degree of stimulation was significantly greater with the C3βGH reporter. The ability of the mammalian GATA-3 factors to *trans*-activate the various reporter genes was comparable to that of the chicken GATA-3 protein tested in parallel experiments (data not shown and reference 47).

In order to demonstrate that both hGATA-3 and mGATA-3 proteins were expressed at comparable levels in transfected cells, extracts were made and incubated with a radiolabeled δE4 oligonucleotide; the resulting DNA-protein complexes were then assayed in gel mobility shift experiments. Specific complexes were formed with proteins present in extracts of cells transfected with the hGATA-3 and mGATA-3 cDNAs, but not with extracts from control transfectants (Fig. 6B). These complexes were inhibited by an excess of unlabeled δE4 oligonucleotide, thereby demonstrating that hGATA-3 and mGATA-3 can specifically bind to the GATA sites present in the human TCR δ gene enhancer.

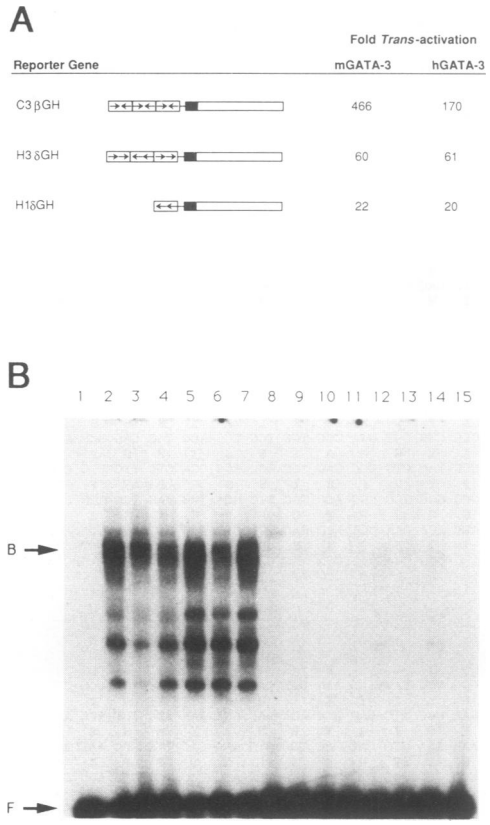

FIG. 6. *trans*-activation by the mGATA-3 and hGATA-3 factors. (A) mGATA-3 and hGATA-3 factors activate transcription from the chicken β-globin enhancer or the human TCR δ gene enhancer GATA-binding sites. mGATA-3 and hGATA-3 cDNAs placed under the transcriptional control of the Rous sarcoma virus promoter and enhancer (see Materials and Methods) were cotransfected with reporter plasmid C3βGH, H3δGH, or H1δGH (see text) into QT6 cells. *trans*-activation was calculated on the basis of human growth hormone secretion (as assayed by RIA) into the cell culture medium upon activation of the reporter gene plasmid (see Materials and Methods). Numbers given for *trans* activation are the results obtained after averaging at least two independent transfections. (B) mGATA-3 and hGATA-3 are expressed at similar levels in transfected cells. Extracts were prepared from the same transfected cells used for the RIA described for panel A and then examined for GATA-3 expression by gel mobility shift assays (see Materials and Methods) using 0.5 ng of the δE4 oligonucleotide as a probe (lane 1). The extracts used were prepared either from cells transfected with mGATA-3 (lanes 2 to 4 and 9 to 11) or hGATA-3 (lanes 5 to 7 and 12 to 14) expression plasmids plus reporter genes or from cells transfected with reporter plasmids and RSV · CAT (lanes 8 and 15). The reporter genes were H3δGH (lanes 2, 5, 8, and 15), C3βGH (lanes 3 and 6), and H1δGH (lanes 4 and 7). Lanes 2 to 8 contained 0.5 μg of nonspecific competitor DNA in the binding reaction, while lanes 9 to 15 contained the same plus 50 ng of unlabeled δE4 oligonucleotide. B and F indicate the positions of bound and free radiolabeled oligonucleotide, respectively.

DISCUSSION

GATA-3 is a vertebrate T-cell transcription factor. GATA-3 DNA-binding proteins are abundantly expressed in chicken, mouse, and human T-lymphocyte cell lines (Fig. 1 and data not shown). The GATA-3 protein sequences deduced from conceptual translation of the cDNA clones (Fig. 2 and 3) (47) from these three vertebrate species show a remarkable degree of amino acid sequence identity, indicating that the structure of this protein, even outside the zinc finger DNA-binding domain, has been strongly conserved through evolution (Fig. 4). With the most abundant levels of expression found in T lymphocytes and the developing brain (Fig. 5), the tissue distribution of GATA-3 mRNA expression has also been conserved between species, implying that it fulfills a critical function in those cells in which it is expressed.

The *trans*-activation data presented here (Fig. 6) further demonstrate that hGATA-3 and mGATA-3 can recognize and stimulate transcription from a pair of consensus WGAT AR binding sites found in the human T-cell receptor δ gene enhancer. These binding sites lie within a previously identified human T-lymphocyte-specific footprint (TCR footprint δE4 [33]), consistent with a central role for this factor in transcriptional regulation of the TCR δ gene. The WGATAR consensus has also been identified within murine and human TCR α and β gene enhancer footprints (15, 18, 41, 46), although it is not possible to determine the tissue specificity of the interactions of the presumptive GATA proteins with these regulatory sequences from the data presented in these reports. The observation that the factor is abundantly expressed in human and murine TCR $\alpha\beta^+$ T lymphocytes (for example, in cell lines CCRF-CEM, A0DH, and Jurkat) and TCR$^-$ T lymphocytes (e.g., CCRF-HSB-2, HuT78, and MOLT-3) strongly implies that the TCR δ gene enhancer site identified here is but one element within T cells which serves as a target for regulation by this transcription factor.

Vertebrate GATA-3 factor expression in the nervous system. In addition to their expression in T cells, mGATA-3 and cGATA-3 mRNAs are also expressed in embryonic (fetal) brain, while hGATA-3 mRNA is strongly expressed in the

human neuroblastoma cell line SKNSH (data not shown). While we have no indication at present as to how this factor may be utilized in the nervous system, there is precedence for tissue-specific transcription factor expression in the brain. Transcription factor Oct-2, originally thought to be exclusively expressed in B lymphocytes, has since been shown to be expressed in the brain as well (16). There are also other examples of temporally distinct developmental patterns of expression similar to that which we find for GATA-3. The *Drosophila* homeodomain gene fushi tarazu (5) plays a role in early pattern formation and in later development is expressed in the nervous system; furthermore, Pit-1, originally identified as a pituitary gland-specific regulatory factor, is expressed in the embryonic brain but not in the adult brain, much like mGATA-3 (16).

GATA gene family regulation during development. The identification of multiple members of a transcription factor family is not a unique observation. A number of other transcription factor gene families (including Oct [3, 28, 36, 39, 40], CCAAT [4, 19, 20, 26, 27, 38], and AP-1 [32, 37]) in which individual members of the family display distinct tissue and temporal expression profiles have also been identified. The cloning of hGATA-3 and mGATA-3 adds two to the growing list of transcription factors containing a $C\text{-}X_2\text{-}C\text{-}X_{17}\text{-}C\text{-}X_2\text{-}C$ zinc finger motif. This DNA-binding domain has been evolutionarily conserved from fungi (1, 10) to humans (42, 48) and is utilized repeatedly within individual organisms as the conserved element defining the GATA transcription factor family. GATA protein families have now been identified in *Xenopus laevis* (47a), chickens (47), mice, and humans (references 42, 45, and 48 and this report). These vertebrate factors all bind to the WGATAR sequence by using this highly conserved, duplicated C_4 zinc finger structure (21, 47). As in chickens, the various human proteins that bind to the GATA consensus sequence are expressed in distinct cell types and are therefore inferred to direct tissue-specific transcription in those lineages. mGATA-1 and hGATA-1 have been shown to be expressed specifically in the erythroid, megakaryocytic, and mast cell lineages (23, 35). A distinct hGATA protein found in endothelial cells may, on the basis of its pattern of expression, be the human homolog of cGATA-2 (45, 47). hGATA-3 and mGATA-3 (identified here), like cGATA-3 (47), are most abundantly expressed in T lymphocytes.

We conclude from this study that the vertebrate GATA-3 transcription factors, having virtually the same DNA sequence recognition properties, primary amino acid sequence, and mRNA developmental expression profiles, act to stimulate specific sets of genes in T lymphocytes and the developing brain in chickens, mice, and humans. The differential expression of specific members of the vertebrate GATA gene families may, by cooperative activity with other cell-specific and/or ubiquitous factors, act to direct the tissue-specific transcription of different genes in a number of distinct and developmentally unrelated cell lineages. Functional analysis of the activities of each member of the GATA gene family of activator proteins within an appropriate in vitro or in vivo transcriptional regulatory environment should determine whether or not this expectation is correct.

ACKNOWLEDGMENTS

We thank F. Grosveld, D. I. H. Linzer, and V. Patel for the generous donation of murine RNA samples and S. K. Pierce and S. T. Rosen for the gifts of a variety of human and murine cell lines. We also thank M. Davis for T-cell tutelage and S. H. Orkin for continued enlightening discussions.

This work was supported by NIH grants GM 28896 and HL 24415 (J.D.E.), by an NIH NRSA predoctoral training award (GM 08061) to Northwestern University (L.J.K.), and by a postdoctoral fellowship from the Leukemia Society of America (M.W.L.).

REFERENCES

1. **Arst, H. N., Jr., B. Kudla, and N. Martinez-Rossi.** 1989. *Aspergillus* and mouse share a new class of 'zinc finger' protein. Trends Genet. **5**:291.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science **196**:180–183.
3. **Clerc, R. G., L. M. Corcoran, J. H. LeBowitz, D. Baltimore, and P. A. Sharp.** 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes Dev. **2**:1570–1581.
4. **Descombes, P., M. Chojkier, S. Lichtsteiner, E. Falvey, and U. Schibler.** 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. Genes Dev. **4**:1541–1551.
5. **Doe, C. Q., Y. Hiromi, W. J. Gehring, and C. Goodman.** 1988. Expression and function of the segmentation gene *fushi tarazu* during *Drosophila* neurogenesis. Science **239**:170–175.
6. **Emerson, B. M., J. M. Nickol, P. D. Jackson, and G. Felsenfeld.** 1987. Analysis of the tissue-specific enhancer at the 3′ end of the chicken adult β-globin gene. Proc. Natl. Acad. Sci. USA **84**:4786–4790.

6a. **Engel, J. D.** Unpublished data.

7. **Evans, T., and G. Felsenfeld.** 1989. The erythroid-specific transcription factor Eryf1: a new finger protein. Cell **58**:877–885.
8. **Evans, T., M. Reitman, and G. Felsenfeld.** 1988. An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. Proc. Natl. Acad. Sci. USA **85**:5976–5980.
9. **Fried, M. G., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
10. **Fu, Y.-H., and G. A. Marzluf.** 1990. *nit-2*, the major nitrogen regulatory gene of *Neurospora crassa*, encodes a protein with a putative zinc finger DNA-binding domain. Mol. Cell. Biol. **10**:1056–1065.
11. **Gallarda, J. L., K. P. Foley, Z. Yang, and J. D. Engel.** 1989. The β-globin stage selector element factor is erythroid-specific promoter/enhancer binding protein NF-E4. Genes Dev. **3**:1845–1859.
12. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. Nucleic Acids Res. **9**:3047–3060.
13. **Giguere, V., S. M. Hollenberg, M. G. Rosenfeld, and R. M. Evans.** 1986. Functional domains of the human glucocorticoid receptor. Cell **46**:645–652.
14. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79**:6777–6781.
15. **Gottschalk, L. R., and J. M. Leiden.** 1990. Identification and functional characterization of the human T-cell receptor β gene transcriptional enhancer: common nuclear proteins interact with the transcriptional regulatory elements of the T-cell receptor α and β genes. Mol. Cell. Biol. **10**:5486–5495.
16. **He, X., M. N. Treacy, D. M. Simmons, H. A. Ingraham, L. W. Swanson, and M. G. Rosenfeld.** 1989. Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature (London) **340**:35–42.
17. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
18. **Ho, I.-C., L.-H. Yang, G. Morle, and J. M. Leiden.** 1989. A T-cell-specific transcriptional enhancer element 3′ of Cα in the human T-cell receptor α locus. Proc. Natl. Acad. Sci. USA

86:6714–6718.
19. **Landschulz, W. H., E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
20. **Lum, L. S. Y., L. A. Sultzman, R. J. Kaufman, D. I. H. Linzer, and B. J. Wu.** 1990. A cloned human CCAAT-box-binding factor stimulates transcription from the human hsp70 promoter. Mol. Cell. Biol. **10**:6709–6717.
21. **Martin, D. I. K., and S. H. Orkin.** 1990. Transcriptional activation and DNA-binding by the erythroid factor GF-1/NF-E1/Eryf1. Genes Dev. **4**:1886–1898.
22. **Martin, D. I. K., S.-F. Tsai, and S. H. Orkin.** 1989. Increased γ-globin expression in a nondeletion HPFH mediated by an erythroid-specific DNA-binding factor. Nature (London) **338**: 435–438.
23. **Martin, D. I. K., L. I. Zon, G. Mutter, and S. H. Orkin.** 1990. Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. Nature (London) **344**:444–447.
24. **Mignotte, V., J. F. Eleouet, N. Raich, and P.-H. Romeo.** 1989. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. Proc. Natl. Acad. Sci. USA **86**:6548–6552.
25. **Moscovici, C., M. G. Moscovici, H. Jimenez, M. M. C. Lai, M. J. Hayman, and P. K. Vogt.** 1977. Continuous tissue culture cell lines derived from chemically induced tumors of Japanese quail. Cell **11**:95–103.
26. **Mueller, C. R., P. Maire, and U. Schibler.** 1990. DBP, a liver-enriched transcription activator, is expressed late in ontogeny and its tissue specificity is determined posttranscriptionally. Cell **61**:279–291.
27. **Muller, M. M., S. Ruppert, W. Schaffner, and P. Matthias.** 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. Nature (London) **336**:544–551.
28. **Okamoto, K., H. Okazawa, A. Okuda, M. Sakai, M. Muramatsu, and H. Hamada.** 1990. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell **60**:461–472.
29. **Orkin, S. H.** 1990. Globin gene regulation and switching: circa 1990. Cell **63**:665–672.
30. **Perkins, N. D., R. H. Nicolas, M. A. Plumb, and G. H. Goodwin.** 1989. The purification of an erythroid protein which binds to enhancer and promoter elements of haemoglobin genes. Nucleic Acids Res. **17**:1299–1314.
31. **Plumb, M., J. Frampton, H. Wainwright, M. Walker, K. Macleod, G. Goodwin, and P. Harrison.** 1989. GATAAG; a cis-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. Nucleic Acids Res. **17**:73–92.
32. **Rauscher, F. J., III, D. R. Cohen, T. Curran, T. J. Bos, P. K. Vogt, D. Bohmann, R. Tjian, and J. B. R. Franza.** 1988. *Fos*-associated protein p39 is the product of the *jun* proto-oncogene. Science **240**:1010–1016.
33. **Redondo, J. M., S. Hata, C. Brocklehurst, and M. S. Krangel.** 1990. A T cell-specific transcriptional enhancer within the human T cell receptor δ locus. Science **247**:1225–1229.
34. **Riddle, R. D., M. Yamamoto, and J. D. Engel.** 1989. Expression of δ-aminolevulinate synthase in avian cells: separate genes encode erythroid-specific and non-specific isozymes. Proc. Natl. Acad. Sci. USA **86**:792–796.
35. **Romeo, P.-H., M.-H. Prandini, V. Joulin, V. Mignotte, M. Prenant, W. Vainchenker, G. Marguerie, and G. Uzan.** 1990. Megakaryocytic and erythrocytic lineages share specific transcription factors. Nature (London) **344**:447–449.
36. **Rosner, M. H., M. A. Vigano, K. Ozato, P. M. Timmons, F. Poirier, P. Rigby, and L. M. Staudt.** 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature (London) **345**:686–692.
37. **Ryder, K., L. F. Lau, and D. Nathans.** 1988. A gene activated by growth factors is related to the oncogene v-*jun*. Proc. Natl. Acad. Sci. USA **85**:1487–1491.
38. **Santoro, C., N. Mermod, P. C. Andrews, and R. Tjian.** 1988. A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) **334**:218–224.
39. **Schöler, H. R., S. Ruppert, N. Suzuki, K. Chowdhury, and P. Gruss.** 1990. New type of POU domain in germ line-specific protein Oct-4. Nature (London) **344**:435–439.
40. **Sturm, R. A., G. Das, and W. Herr.** 1988. The ubiquitous octamer protein Oct-1 contains a POU domain with a homeo subdomain. Genes Dev. **2**:1582–1599.
41. **Takeda, J., A. Cheng, F. Mauxion, C. A. Nelson, R. D. Newberry, W. C. Sha, R. Sen, and D. Y. Loh.** 1990. Functional analysis of the murine T-cell receptor β enhancer and characteristics of its DNA-binding proteins. Mol. Cell. Biol. **10**:5027–5035.
42. **Trainor, C. D., T. Evans, G. Felsenfeld, and M. S. Boguski.** 1990. Structure and evolution of a human erythroid transcription factor. Nature (London) **343**:92–96.
43. **Tsai, S.-F., D. I. K. Martin, L. I. Zon, A. D. D'Andrea, G. G. Wong, and S. H. Orkin.** 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. Nature (London) **339**:446–451.
44. **Wall, L., E. deBoer, and F. Grosveld.** 1988. The human β-globin gene 3′ enhancer contains multiple binding sites for an erythroid-specific protein. Genes Dev. **2**:1089–1100.
45. **Wilson, D. B., D. M. Dorfman, and S. H. Orkin.** 1990. A nonerythroid GATA-binding protein is required for function of the human preproendothelin-1 promoter in endothelial cells. Mol. Cell. Biol. **10**:4854–4862.
46. **Winoto, A., and D. Baltimore.** 1989. A novel, inducible and T cell-specific enhancer located at the 3′ end of the T cell receptor α locus. EMBO J. **8**:729–733.
47. **Yamamoto, M., L. J. Ko, M. W. Leonard, H. Beug, S. H. Orkin, and J. D. Engel.** 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. Genes Dev. **4**:1650–1662.

47a. **Zon, L., and S. Orkin.** Personal communication.

48. **Zon, L. I., S.-F. Tsai, S. Burgess, P. Matsudaira, G. A. P. Bruns, and S. H. Orkin.** 1990. The major human erythroid DNA-binding protein (GF-1): primary sequence and localization of the gene to the X chromosome. Proc. Natl. Acad. Sci. USA **87**:668–672.