11. Knudson, A. *Cancer Res.* **45,** 1437–1443 (1985).
12. Saiki, R. K. *et al. Science* **239,** 487–491 (1988).
13. Kondoleon, S. *et al. Nucleic Acids Res.* **15,** 10605 (1987).
14. Nakamura, Y. *et al. Nucleic Acids Res.* **16,** 5707 (1988).
15. Buchman, V. L. *et al. Gene* **70,** 245–252 (1988).
16. Collins, F. S., Ponder, B. A. J., Seizinger, B. R. & Epstein, C. J. *Am. J. hum. Genet.* **44,** 1–5 (1989).
17. Finlay, C. A. *et al. Molec. cell. Biol.* **8,** 531–539 (1988).
18. Eliyahu, D. *et al. Oncogene* **3,** 313–321 (1988).
19. Soussi, T. *et al. Oncogene* **1,** 71–78 (1987).
20. Lee, J. H. *et al. Proc. Am. Ass. Cancer Res.* **30,** 442 (1989).
21. Atkin, N. B. & Baker, M. C. *Cancer Genet. Cytogenet.* **37,** 229–233 (1989).
22. Yano, T. *et al. J. natn. Cancer Inst.* **181,** 518–523 (1989).
23. Tsai, Y. C. *et al. Cancer Res.* (in the press).
24. Herskowitz, I. *Nature* **329,** 219–222 (1987).
25. Kraiss, S., Quaiser, A., Oren, M. & Montenach, M. *J. Virol.* **62,** 4737–4744 (1988).
26. Kern, S. E. *et al. J. Am. Med. Ass.* **261,** 3099–3103 (1989).
27. Haymerle, H. *et al. Nucleic Acids Res.* **14,** 8615–8624 (1986).
28. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. *Nucleic Acids Res.* **16,** 7583–7600 (1988).

ACKNOWLEDGEMENTS. We thank Y. Nakamura, R. White and M. Litt for providing the RFLP probes and T. Gwiazda for preparation of the manuscript. The work was supported in part by The Clayton Fund, the McAshan Fund, The National Neurofibromatosis Foundation and the NIH.

Putative transcription activator with alternative isoforms encoded by human *ZFX* gene

Ansbert Schneider-Gädicke, Peggy Beer-Romero, Laura G. Brown, Graeme Mardon, Shiuh-Wen Luoh & David C. Page

Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE *ZFY* gene in the sex-determining region of the human Y chromosome encodes a protein with 13 zinc fingers, and may determine whether an embryo develops as a male or female[1]. *ZFX*, a related gene on the human X chromosome, may also function in sex determination; it encodes a protein with a very similar zinc-finger domain and escapes X inactivation[1,2]. *ZFY* and *ZFX* diverged from a common ancestral gene before the radiation of placental mammals, and retain a similar genomic organization[2]. Analysis of complementary DNAs from the mouse Y-chromosomal homologues of *ZFY* indicates that these genes encode probable transcription activators[3,4]. Here, we report that *ZFX* encodes a protein composed of a highly acidic amino-terminal domain, a basic putative nuclear-localization signal, and a carboxy-terminal zinc-finger domain. This combination of features, also found in the *ZFY* gene product, is typical of transcription activators. Alternative splicing generates *ZFX* transcripts encoding isoforms of 575 and 804 amino acids. These ZFX protein isoforms differ in the length of their acidic domains and may be functionally distinct.

ZFX is transcribed in all human cells analysed[2]. We cloned *ZFX* cDNAs from a male lymphoblastoid cell line. Analysis by restriction mapping and hybridization with genomic probes revealed three distinct types of cDNAs. Whereas two types were represented by a single clone each (cDNAs 1 and 3), a third type was represented by three clones (cDNAs 2, 4 and 5).

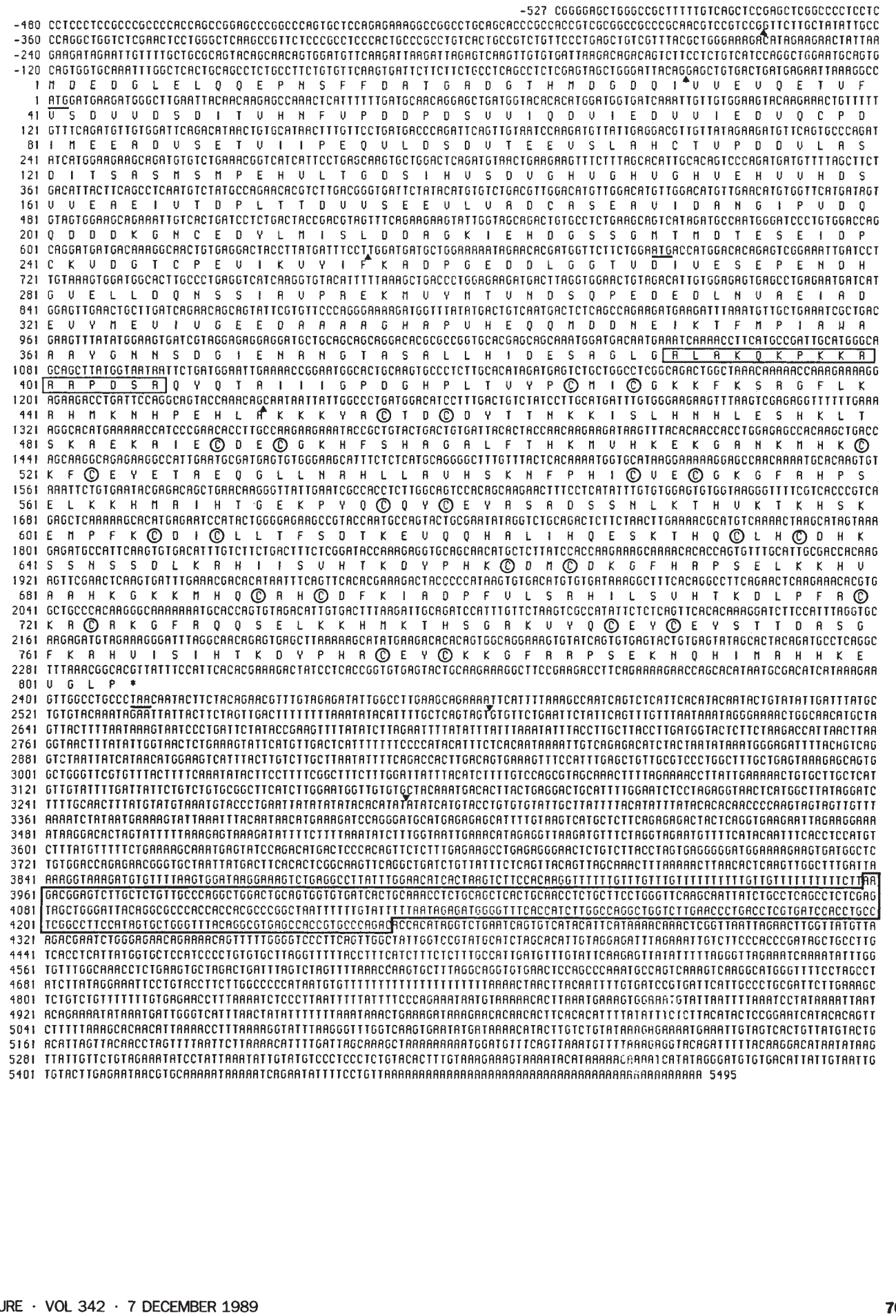
We determined the nucleotide sequence of one cDNA of each type (Fig. 1). Comparison of cDNAs 1, 2, and 3 showed that alternative splicing and polyadenylation had produced structurally distinct 5′ untranslated, coding, and 3′ untranslated regions (Fig. 2*a*). Differential splicing involved an invariant donor site (nucleotide −378) and alternative acceptor sites (at nucleotides +647, −28 and −377, in cDNAs 1, 2 and 3, respectively). Each cDNA contains a single long open reading frame (ORF), and in each case the first ATG occurs in a sequence context favourable for initiation of translation[5]. Complementary DNAs 2 and 3, with identical ORFs, encode a protein of 804 amino acids (ZFX^{804}), whereas cDNA 1 encodes an isoform of 575 amino acids (ZFX^{575}). (Alternatively, with cDNA 1, translation initiation at a second ATG, whose context is highly favourable, would result in production of an isoform of 573 amino acids.) *In vitro* transcription and translation of cDNAs 3 and 1 yielded the predicted full-length and truncated proteins, respectively (data not shown). Both ZFX protein isoforms contain three domains—an N-terminal acidic portion (25% aspartic and glutamic acid), a small basic domain, and a C-terminal run of 13 zinc fingers, each with two cysteines and two histidines (Cys-Cys/His-His zinc fingers). The isoforms differ in that the acidic domain of ZFX^{575} is half that of ZFX^{804}.

Comparison of *ZFX* cDNAs with the genomic locus by restriction mapping and oligonucleotide hybridization gave an overview of the intron–exon organization (Fig. 2*a*). The C-terminal zinc-finger domain and 3′ untranslated region (UTR) are encoded by a single exon, whereas the N-terminal acidic domain is encoded by a minimum of four exons for ZFX^{575} (cDNA 1) and a minimum of six exons for ZFX^{804} (cDNAs 2 and 3). The cDNAs span 67 kilobases (kb) in the genome.

Of the three types of cDNAs, type 2, encoding ZFX^{804}, seems to be most representative of the 6.3- and 8-kb transcripts observed on northern blots[2]. First, the coding exon defined by oligonucleotides '150' and '637' is present in cDNAs 2 and 3 but not in cDNA 1 (Fig. 2*a*). Northern analysis using the corresponding genomic DNA fragment revealed that this exon is present in the 6.3- and 8-kb transcripts (not shown). Second, northern analysis indicated that the polyadenylation site used in cDNA 2 corresponds, at least roughly, to the 3′ end of the main *ZFX* transcripts (data not shown). Clone 2, which is 5.6 kb long, could represent a 6.3- or 8-kb transcript that is incomplete at the 5′ end.

Transcripts corresponding to cDNA 1, encoding ZFX^{575}, have not been detected by northern analysis. Using polymerase chain reaction (PCR) amplification, however, we confirmed the differential splicing predicted from cDNA analysis and crucial to the generation of ZFX protein isoforms (Fig. 2*b*). We designed splice-specific 5′ primers spanning splices of the invariant donor site at nucleotide −378 to alternative acceptor sites at nucleotides +647, −28 and −377, as in cDNAs 1, 2 and 3, respectively. The

FIG. 1 Composite nucleotide sequence of *ZFX* cDNAs 1, 2 and 3 (Fig. 2*a*) and predicted amino-acid sequences. Numbering of nucleotides and amino acids is with reference to the first in-frame ATG codon in cDNAs 2 and 3. Known splice sites are indicated by upward arrows. Alternative starts of poly(A) tails are indicated by downward arrows. Complementary DNA 1 begins at nucleotide −527, uses a splice donor site at −378 and an acceptor site at +647; a tail of 20 adenosines follows nucleotide 2,585. Complementary DNA 2 starts at nucleotide −426, uses the same donor, but a different acceptor at −28; a tail of 46 adenosine residues follows nucleotide 5,450. Complementary DNA 3 starts at nucleotide −426, uses the same donor, but yet another different acceptor at −377; a tail of 26 adenosines follows nucleotide 3,298. Complementary DNAs 2 and 3 contain the same ORF with the first in-frame ATG codon at nucleotide 1; the predicted protein of 804 amino acids is composed of an acidic N-terminal domain, a small basic domain (residues 391–406, boxed), and 13 zinc fingers (cysteines of the Cys-X-X-Cys repeats are circled). Complementary DNA 1 has a shorter ORF with the first in-frame ATG codon at nucleotide +688, encoding a protein of 575 residues. The 3′ UTR of cDNA 2 contains an *Alu* sequence[16] (nucleotides +3,959 to +4,250, boxed). An AATAAA polyadenylation signal[17] occurs 20 nucleotides 5′ of the poly(A) tail of cDNA 2. Similar sequences are present 5′ of the poly(A) tails in cDNAs 1 and 3.
METHODS. Complementary DNA libraries were prepared[3] using poly(A)$^+$ RNA prepared[2] from human male lymphoblastoid cell line WHT1659. Five million recombinant phages from unamplified libraries were screened using plasmids pDP1007, pDP1041 and pDP1006 (*ZFY* genomic fragments which cross-hybridize to *ZFX*)[1,2] as probes. Complementary DNA inserts of seven phages were subcloned into Bluescript vectors (Stratagene). The full DNA sequences of three inserts and partial sequences of the other four inserts were determined[2]. As judged by comparison with genomic sequences, five cDNAs originated from *ZFX* and two from *ZFY*.



3′ primer was common to all three cDNAs. Bulk cDNA generated by reverse transcription of male or female lymphoblastoid RNA was used as template. Fragments predicted from the sequences of cDNA clones 1, 2 and 3 were amplified (Fig. 2*b*), confirming the differential splicing shown in Fig. 2*a*. Some unanticipated fragments were simultaneously amplified (Fig. 2*b*).

The ZFX^{804} protein has a primary sequence strikingly similar to those encoded by the mouse *ZFY* homologues (*Zfy*-1 and *Zfy*-2)[3,4], with which it shares virtually identical N and C termini. ZFX^{804} and ZFX^{575} proteins are also similar to the human ZFY protein (ref. 1; and P.B.-R. and A.S.-G., unpublished results). Each of these proteins is composed of an acidic domain, a small basic domain, and 13 Cys-Cys/His-His zinc fingers. This combination of features is reminiscent of those of eukaryotic transcription activators[6]. By analogy, we assume that ZFX and ZFY proteins activate transcription in a sequence-specific fashion. When fused to the DNA-binding domain of GAL4, the acidic domains of mouse Zfx and Zfy-2 do activate transcription in yeast[7].

Are the various ZFX and ZFY proteins functionally distinct? All share a small basic domain, which could localize the proteins to the nucleus[8,9]. Cys-Cys/His-His zinc-finger domains[10,11] bind to nucleic acids in a sequence-specific fashion. The zinc-finger domains of ZFX^{804} and ZFX^{575} are identical and differ from the zinc-finger domain of human ZFY at only 10 of 393 amino acids[2]. Thus ZFX^{804}, ZFX^{575} and ZFY proteins could all recognize the same nucleic-acid target sequences, and thereby perhaps regulate transcription of the same genes.

The N-terminal portions of the ZFX and ZFY proteins are highly acidic. Acidic domains mediate activation in several eukaryotic transcription factors[12–15]. The potency of such activating domains may depend more on the number and density of acidic residues than on particular primary sequences[5]. In ZFX^{804} and mouse Zfx, Zfy-1 and Zfy-2, the acidic domains are of similar size, each with ~25% of residues acidic. By contrast, ZFX^{575} contains an abbreviated acidic domain. ZFX^{804} and ZFX^{575} might be expected to have quantitatively different activating potencies. Furthermore, acidic domains of transcription factors are believed to interact with other proteins in transcription complexes. Truncation of the acidic domain in ZFX^{575} could alter its interactions with other factors, resulting in qualitatively different regulatory properties. (Internal ATG codons in human *ZFY* and mouse *Zfx*, *Zfy-1* and *Zfy-2* could allow production of similarly truncated isoforms.)

We compared the human *ZFX* cDNA sequence with that of a murine *Zfx* cDNA (ref. 7; Fig. 3). The coding regions of *ZFX* and *Zfx* show a high degree of similarity. It is of interest that the 3′ UTR is also conserved. Except for an *Alu* insertion in the human sequence, the entire human 3′ UTR of 3 kb shows over 80% nucleotide identity with the murine 3′ UTR.

Given the similarity in the primary structure of the ZFX protein and its Y chromosome-encoded homologues, it is conceivable that the products of the *ZFX* and *ZFY* genes are

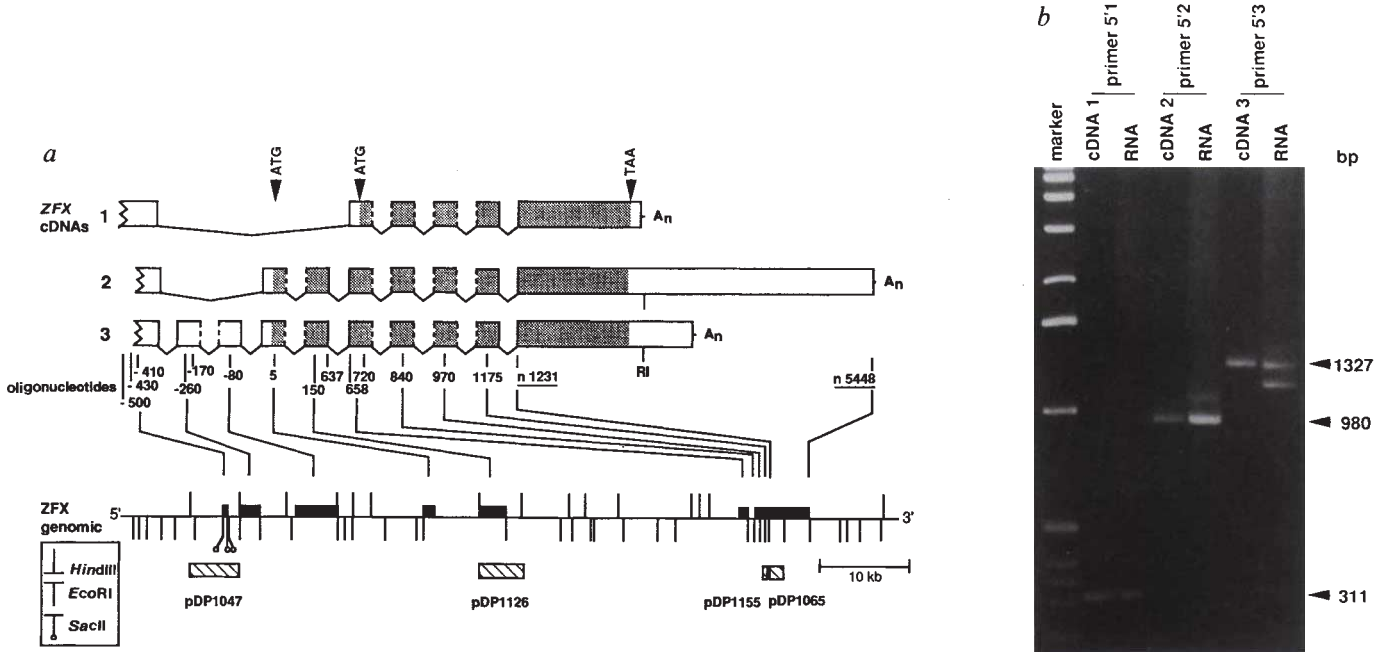
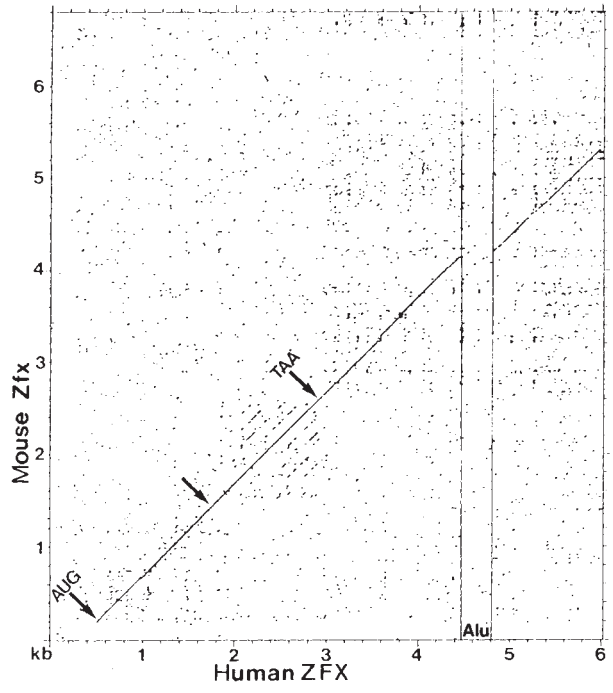


FIG. 2 Analysis of alternative ZFX transcripts. *a*, Organization of cDNAs and genomic locus. Top: boxes indicate the minimum number of exons and, with the exception of the 3′-most exon, are not to scale; coding regions are shaded. Several splice sites (solid vertical lines of boxes) were defined by DNA sequence comparison among cDNA clones 1, 2 and 3 or by comparison with genomic sequences. Additional splices (dashed vertical lines of boxes) are not precisely defined but inferred from comparison of cDNA and genomic restriction maps. Boxes bounded by dashed lines may contain more than one exon each. Genomic DNA fragments[2] hybridizing (5–10 °C below calculated melting temperature) with *ZFX* oligonucleotides (numbers denote the 11th nucleotide in 20-mer oligonucleotides) are shown by black segments; each contains one or more exons. *b*, PCR analysis of differential splicing. Three different 5′ primers (20mers) spanning splices unique to cDNAs 1, 2 or 3 have in common the invariant donor site at nucleotide −378, but each includes a unique acceptor (−389 to −378/+647 to +654 (primer 5′1); −383 to −378/−28 to −15 (primer 5′2); −381 to −362 (primer 5′3)). An invariant 3′ primer (+923 to +946) was used. Using cDNA clones 1, 2 and 3 as templates, the predicted products were amplified (cDNA 1, 311 base pairs (bp); cDNA 2, 980 bp; cDNA 3, 1,327 bp). Using 'RNA' (bulk cDNA prepared from male lymphoblastoid WHT1659 RNA by reverse transcription) as template, products of the same lengths were obtained. In addition, with bulk cDNA, primer 5′2 amplified a 1.1-kb fragment (which may incorporate additional coding exon(s) not detected by cDNA cloning), whereas primer 5′3 amplified a barely detectable 1.4-kb fragment (possibly incorporating additional coding or 5′ UTR exon(s)) and a 1.2-kb fragment (which may lack coding or 5′ UTR exon(s) present in cDNA 3). Controls not shown: (1) The 5′ primers are splice specific, yielding products only with the corresponding cDNAs; (2) amplification of cell line RNA, without reverse transcriptase, yielded no products; (3) by using bulk cDNA from female lymphoblastoid cell line WHT1628 as template, products of the same lengths were obtained; and (4) hybridization of Southern blots of these gels verified that the ethidium bromide-detected products originate from *ZFX*.

METHODS. Poly(A)$^+$ RNA (200 ng) was reverse-transcribed using Moloney reverse transcriptase (Bethesda Research). Forty cycles of amplification were performed using *Taq* polymerase and a DNA Thermo Cycler (Cetus). Each PCR reaction (10%) was separated on an agarose gel and stained with ethidium bromide. The marker lane contains a '1-kb ladder' (Bethesda Research).



Homology between human ZFX and mouse Zfx

	% Identity DNA	protein	No of gaps
Coding region:			
N-terminal domain	90	92	2
Zinc finger domain	93	99.5	0
3'untranslated region:			
5' of *Alu* insertion	85	-	19
3' of *Alu* insertion	81	-	22

FIG. 3 Nucleotide and corresponding amino-acid similarity between human *ZFX* (Fig. 1) and murine *Zfx* (ref. 7). Striking similarity is found not only in the coding region, but also across the entire 3′ UTR (except for an *Alu* insertion, indicated by vertical lines, in the human), and ~20 nucleotides 5′ of the ATG codon. Arrows indicate the start codon in *ZFX* cDNAs 2 and 3, the splice site preceding the zinc-finger exon, and the TAA stop codon. Numbering (in kb) is with reference to the 5′ ends of the cDNAs. The analysis[18] employed a 'window' of 21 and a 'stringency' of 14.

functionally interchangeable. But the occurrence of isoforms—at least of ZFX—indicates that *ZFX* and *ZFY* could produce transcription factors with qualitatively or quantitatively distinct functions. The 3′ UTRs of *ZFX* transcripts are longer and more highly conserved than those found in *Zfy* transcripts[3,4], and could provide a basis for differential post-transcriptional regulation. Detailed studies of *ZFX* and *ZFY* transcripts during development should further elucidate their postulated roles in sex determination. □

Received 7 August; accepted 17 October 1989.

1. Page, D. C. *et al. Cell* **51,** 1091–1104 (1987).
2. Schneider-Gädicke, A. *et al. Cell* **57,** 1247–1258 (1989).
3. Mardon, G. & Page, D. C. *Cell* **56,** 765–770 (1989).
4. Ashworth, A., Swift, S. & Affara, N. *Nucleic Acids Res.* **17,** 2864 (1989).
5. Kozak, M. *Cell* **44,** 283–292 (1986).
6. Ptashne, M. *Nature* **335,** 683–689 (1988).
7. Mardon, G. *et al. Molec. cell. Biol.* (in the press).
8. Dang, C. V. & Lee, W. M. F. *Molec. cell. Biol.* **8,** 4048–4054 (1988).
9. Dingwall, C. & Laskey, R. A. *A. Rev. Cell Biol.* **2,** 367–390 (1986).
10. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4,** 1609–1614 (1985).
11. Brown, R. S., Sander, C. & Argos, P. *FEBS Lett.* **186,** 271–274 (1985).
12. Ma, J. & Ptashne, M. *Cell* **48,** 847–853 (1987).
13. Hope, I. A. & Struhl, K. *Cell* **46,** 885–894 (1986).
14. Hollenberg, S. M. & Evans, R. M. *Cell* **55,** 899–906 (1988).
15. Godowski, P. J., Picard, D. & Yamamoto, K. R. *Science* **241,** 812–816 (1988).
16. Schmid, C. W. & Jelinek, W. R. *Science* **216,** 1065–1070 (1982).
17. Fitzgerald, M. & Shenk, T. *Cell* **24,** 251–260 (1981).
18. Maizel, J. V. & Lenk, R. P. *Proc. natn. Acad. Sci. U.S.A.* **78,** 7665–7669 (1981).

ACKNOWLEDGEMENTS. We thank Nazneen Aziz, Cornelius Murre and Sang-Ho Park for technical advice, and our colleagues, especially Eric Lander, for comments on the manuscript. This work was supported by the NIH, the Whitaker Health Sciences Fund, the Lucille P. Markey Charitable Trust, and the Searle Scholars Program/Chicago Community Trust. A.S-G. was supported by a fellowship from EMBO.

PDGF induction of tyrosine phosphorylation of GTPase activating protein

Christopher J. Molloy, Donald P. Bottaro, Timothy P. Fleming, Mark S. Marshall*, Jackson B. Gibbs* & Stuart A. Aaronson

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Building 37, Room 1E24, Bethesda, Maryland 20892, USA
* Department of Molecular Biology, Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486, USA

THE cascade of biochemical events triggered by growth factors and their receptors is central to understanding normal cell-growth regulation and its subversion in cancer. Ras proteins ($p21^{ras}$) have been implicated in signal transduction pathways used by several growth factors, including platelet-derived growth factor (PDGF)[1]. These guanine nucleotide-binding Ras proteins specifically interact with a cellular GTPase-activating protein (GAP)[2]. Here we report that in intact quiescent fibroblasts, both AA and BB homodimers of PDGF rapidly induce tyrosine phosphorylation of GAP under conditions in which insulin and basic fibroblast growth factor (bFGF) are ineffective. Although GAP is located predominantly in the cytosol, most tyrosine-phosphorylated GAP is associated with the cell membrane, the site of $p21^{ras}$ biological activity[3,4]. These results provide a direct biochemical link between activated PDGF-receptor tyrosine kinases and the $p21^{ras}$–GAP mitogenic signalling system.

In human malignancies, *ras* genes are commonly activated as oncogenes and their p21 products are potently mitogenic when microinjected into quiescent fibroblasts[5,6]. Yet the function of $p21^{ras}$ proteins in normal mitogenic signal transduction remains unknown. It is well established that the GTP-bound forms of $p21^{ras}$ proteins are active and that the GDP-bound forms are inactive[7–9]. The GAP protein of relative molecular mass 125,000 (M_r ~ 125K) catalyses the conversion of p21-GTP to p21-GDP so that the rate of this recycling is more than 100-fold faster than that intrinsic for p21 (ref. 2). GAP has no effect on oncogenic p21 proteins, allowing them to remain in their active GTP-bound states.

To investigate whether GAP serves as a substrate for activated growth factor-receptor kinases in intact cells, confluent NIH3T3 cells were rendered quiescent by incubating them overnight in serum-free medium. Cells were exposed to saturating concentrations of PDGF BB homodimer for 5 min, after which cell lysates were subjected to immunoblot analysis with an antiphosphotyrosine antiserum (anti-P-Tyr) or anti-GAP peptide serum 638 (ref. 10). PDGF BB homodimer induced the specific tyrosine phosphorylation of several proteins (Fig. 1*a*). These included the α- and β-PDGF receptors each of M_r 185K (ref. 11), a protein of M_r 145K that has been identified as phospholipase C_γ (ref. 12), and several other proteins of M_r ranging from 50–125K. Anti-GAP peptide serum 638, raised against GAP amino-acid residues 139–152 (ref. 10), identified a single band corresponding to a protein of M_r ~ 125K, which was not observed in the presence of excess competing peptide (Fig. 1*a*, lane 3).

When quiescent cells or cells treated with PDGF BB homodimer were first immunoprecipitated with either polyclonal or monoclonal anti-P-Tyr antibodies followed by immunoblotting with anti-GAP antiserum 638, GAP was detectable only in cells stimulated with PDGF BB homodimer (Fig. 1*b*, lane 3). Similar results were observed using anti-GAP peptide serum 677 raised against GAP residues 968–981 (ref. 10). Moreover, recovery of GAP by anti-P-Tyr antiserum was specifically inhibited by an excess of competing phosphotyrosine