

A family of transmembrane proteins with homology to the MET–hepatocyte growth factor receptor

(RON/MSP/X chromosome/neural development)

E. MAESTRINI*†, L. TAMAGNONE†‡, P. LONGATI‡§, O. CREMONA‡, M. GULISANO¶, S. BIONE*, F. TAMANINI*,
B. G. NEEL§, D. TONIOLO*, AND P. M. COMOGLIO‡||

*Institute of Genetics Biochemistry and Evolution, Consiglio Nazionale delle Ricerche, 27100, Pavia, Italy; †Molecular Medicine Unit, Beth Israel Hospital, Boston, MA 02215; ‡Department of Biological and Technological Research, San Raffaele Hospital, 20100, Milano, Italy; and †Institute for Cancer Research (IRCC), University of Torino Medical School, 10126, Torino, Italy

Communicated by Renato Dulbecco, The Salk Institute for Biological Studies, La Jolla, CA, September 15, 1995 (received for review July 10, 1995)

ABSTRACT In hunting for unknown genes on the human X chromosome, we identified a cDNA in Xq28 encoding a transmembrane protein (SEX) of 1871 amino acids. SEX shares significant homology with the extracellular domain of the receptors encoded by the oncogenes MET, RON, and SEA [hepatocyte growth factor (HGF) receptor family]. Further screenings of cDNA libraries identified three additional sequences closely related to SEX; these were named SEP, OCT, and NOV and were located on human chromosomes 3p, 1, and 3q, respectively. The proteins encoded by these genes contain large cytoplasmic domains characterized by a distinctive highly conserved sequence (SEX domain). Northern blot analysis revealed different expression of the SEX family of genes in fetal tissues, with SEX, OCT, and NOV predominantly expressed in brain, and SEP expressed at highest levels in kidney. *In situ* hybridization analysis revealed that SEX has a distinctive pattern of expression in the developing nervous system of the mouse, where it is found in postmitotic neurons from the first stages of neuronal differentiation (9.5 day postcoitus). The SEX protein (220 kDa) is glycosylated and exposed at the cell surface. Unlike the receptors of the HGF family, p220^{SEX}, a MET–SEX chimera or a constitutively dimerized TPR–SEX does not show tyrosine kinase activity. These data define a gene family (SEX family) involved in the development of neural and epithelial tissues, which encodes putative receptors with unexpected enzymatic or binding properties.

Proteins spanning the cell plasma membrane have a major role in the signal exchange between the cytoplasm and the surrounding environment. Single-pass transmembrane proteins (i.e., those having a single transmembrane domain) are involved in cell signaling via specific interactions with soluble, cell surface, or matrix-bound molecules (1–3). The large extracellular region of these proteins is characterized by specific structural motifs (4, 5). Single-pass transmembrane proteins include receptors for growth factors and cytokines, some of which are equipped with intrinsic tyrosine or serine/threonine kinase activity and directly activate signal transduction pathways (for reviews, see refs. 6 and 7). Other receptors associate with cytoplasmic signal transducers with catalytic properties (8).

The family of the MET-related genes (MET, RON, and SEA) encodes single-pass transmembrane proteins sharing structural homology in their extracellular domains and including a conserved intracellular tyrosine kinase (TK) domain (9–13). Upon ligand binding to the extracellular domain, these receptors undergo dimerization and transduce intracellular signals via the large cytoplasmic domain (14–16). Hepatocyte growth

factor (HGF) family growth factor receptors are implicated in a common array of complex biological functions, including proliferation, motility, and differentiation of epithelial, endothelial, and blood-derived cells (17–19). The protooncogene MET is frequently overexpressed in neoplastic cells and, due to its prominent role in the control of motility and invasion, is involved in metastasis formation (20–22).

In this paper, we identify a family of genes encoding putative receptors with structural similarities to the extracellular domain of the MET/HGF receptor. The large cytoplasmic domains of these proteins display no similarity to TKs; instead, they share a highly homologous sequence, suggesting signal transduction by uncharacterized enzymatic or binding properties.**

MATERIALS AND METHODS

Cells and Antibodies. IMR-32, Neuro2a, and COS-7 cell lines were obtained from the American Type Culture Collection. BOSC23 is a subline of 293T human fetal kidney (23). A glutathione S-transferase fusion protein containing the last 38 amino acids of SEX, obtained by PCR amplification and cloning in pGEX-1AT vector (Pharmacia), was purified and used as immunogen to raise anti-SEX antibodies in rabbits, which were then purified by affinity chromatography.

cDNA Library Screening and Cloning of Human and Murine Sequences. The SEX cDNAs were isolated from human fetal brain, skeletal muscle, and total human embryo cDNA libraries (24). OCT and NOV were isolated from a skeletal muscle cDNA library (Clontech) using a Sac II subclone of the SEX cDNA (63-10-S3) as a probe. The cDNAs encoding SEP were isolated from a human fetal brain cDNA library (Stratagene). Hybridization conditions were as described (24). The mouse Sex cDNA probe used for *in situ* hybridization experiments was obtained by reverse transcription (RT)-PCR amplification of mouse brain cDNA as described (25). The primers, derived from the human sequence, were 6.3B52 (5'-CTTCAACTGTGGCTGGTGCATCTC-3', positions 2597–2610) and 6.3B31 (5'-CACTCACAGTGATGGCAGT-GCTTC-3', positions 3324–3301) and the annealing step was performed at 62°C. The murine sequence was 83% and 88% identical to the human SEX sequence at the nucleotide and protein levels, respectively. Nucleotide sequences, obtained by

Abbreviations: HGF, hepatocyte growth factor; RT, reverse transcription; ORF, open reading frame; dpc, days postcoitus; TK, tyrosine kinase.

†E.M. and L.T. contributed equally to this work.

||To whom reprint requests should be addressed at: Institute for Cancer Research, University of Torino, C.so M. D'Azeglio 52, 10126, Torino, Italy.

**The sequences reported in this paper have been deposited in the GenBank data base [accession nos. X87197 (SEX), X87904 (SEP), X87831 (OCT), and X87832 (NOV)].

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

the dideoxynucleotide method, were computer analyzed by BLAST, FASTA, and PROSITE and aligned using CLUSTAL and BESTFIT programs.

Chromosomal Mapping. A panel of DNAs from hamster-human cell hybrids, kindly provided by M. Rocchi (University of Bari, Italy) was used as described (26).

Construction of MET-SEX and TPR-SEX chimeras. (i) Two cDNA fragments corresponding to the extracellular domain of the HGF receptor and to the transmembrane and intracellular domains of SEX were generated by PCR and joined by in-frame *Mlu* I restriction sites. The *MET* antisense primer was 5'-TACCACGCGTTACTTTTCCAAGGAC-3' (positions 2949-2964; accession no. X54559); the *SEX* sense primer was 5'-TCGACGCGTCTGCACACTTCGGCAGA-3'. (ii) Similarly, two cDNA fragments, corresponding to the *TPR* sequence found in the *TPR-MET* oncogene (16) and to the intracellular domain of *SEX*, were joined by in-frame *Spe* I restriction sites. The *TPR* antisense primer was 5'-GGACTAGTTAAGTATTCAAGTTCTTGA-3'; the *SEX* sense primer was 5'-TGACTAGTACCCTCAAGCGTCTGCA-3' (positions 3945-3962).

Northern Blots and in Situ Hybridization. Poly(A)⁺ RNA was hybridized at high stringency using standard methods (27) with the following probes: for *SEX*, the subclone 6.3α-SG, *Eco*RI/*Sac* II; for *SEP*, a fragment from nucleotides 1217-1790 of the cDNA, prepared by RT-PCR; for *OCT*, the whole isolated m9 cDNA; for *NOV*, the 600-bp subclone m13-0.8, from the 5' end of the cDNA to a *Sma* I site. CTP[γ-³⁵S]-labeled antisense RNA probes (105 cpm/ml) were generated from the murine *Sex* cDNA and used for *in situ* hybridization as described (19).

Expression in Eukaryotic Cells. The cDNA constructs were subcloned in pXMT2 or pBAT and transfected by calcium phosphate in COS-7, BOSC23, and Neuro2a cells. Western blots, cell surface biotinylation, and immunoprecipitations were performed as described (28).

RESULTS

Isolation and Sequence Analysis of SEX cDNA. In hunting for genes on the human X chromosome by CpG island screening (24), we identified two overlapping cDNAs in Xq28—6.3 A and 6.3 B. The cDNAs mapped between the 9F and GdX genes and identified a transcript, which was called 6.3. By further screening of cDNA libraries with these sequences, we isolated additional overlapping clones for an approximate total length of 6 kb (Fig. 1). A sequence of 6039 nucleotides was determined; on comparison with computer data bases, it did not show similarity to any known nucleotide sequences. An open reading frame (ORF) encoding a potential protein of 1871 amino acids was identified within the

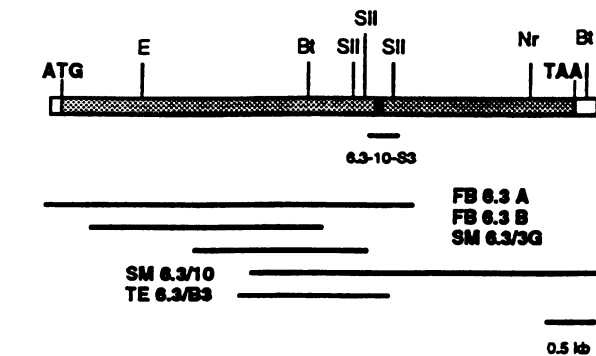


FIG. 1. Map of *SEX* cDNA. Stippled region, ORF; solid box, transmembrane domain. Below are the isolated cDNAs: 6.3-10-S3 is the *Sac* II subclone used for identification of the other members of the *SEX* family. Bt, *Bst*EII; Nr, *Nru* I; E, *Eco*RI; SII, *Sac* II.

sequence (Fig. 2). The corresponding gene was named *SEX*, since it maps on the long arm of the sex chromosome X.

The most upstream initiation site is an ATG codon at position 192, surrounded by a sequence (CGGCCATGC) that represents a good match with the consensus proposed by Kozak (29) for initiation of translation. A termination codon (TAA) was found at nucleotide 5805, followed by several stop codons in all reading frames. None of the cDNAs included a poly(A)⁺ tail or a polyadenylation consensus sequence at its 3' end. A hydrophobic peptide of 19 amino acids downstream from the initiation codon has the features of a signal peptide for membrane localization (30). The hydropathicity plot (31) of the predicted amino acid sequence identified a second hydrophobic stretch from residue 1215 to 1241, which has the characteristics of a transmembrane domain. This is followed by a stretch of basic amino acid residues, the typical feature of stop-transfer sequences (Lys¹²⁴³-Arg-Lys). The expected molecular mass of the *SEX* core protein, after cleavage of the signal peptide, is ≈200 kDa.

The N-terminal region is 1194 amino acids long and includes nine potential N-glycosylation sites. Comparison of the *SEX* sequence with the protein data bases showed a region of similarity to the extracellular domains of the receptors encoded by the *MET* oncogene family, *MET*, *RON*, and *SEA* (9, 12, 13, 32). The overall sequence identity between the extra-

MPSVCLLLL FLAVGGALGN RFFRAFWTD TTLTHLAVHR VTGEVFGAV NRVEKLPANL	60
TELRAHVTFP VEDNARCYPF PSMRVCARHL APVDNINKLL LIDYAAARLV ACGSIWQGIC	120
QFLRLDLDFK LGEFHRKHEH YLSGAQEPDS MAGVIVEVGGQ GPKSLFVUTA VDGGSEYFPT	180
LSSRKLISDE DSADMFLVY QDEFVSSQIK IPSDTLSLYP ALDIYYIYGF VSASFVYFLT	240
LQLDTQQTLL DTAGEKFFTS KIVRMCAQDS EPHYVYVFFI GCSWRGVEYR LVQSAHLAKP	300
GLLLAALQGV PADEDVLFIT FSQQQRNRS PPRQTILCLF TLSNINAHIR RRIQSCYRGE	360
GTLALPMLLN KELPCINTPX QINGNPGGLV LNQPLGLGHV IEGLLPLADS TDGMAVAAV	420
TYRQHSVFI QTRGSLKKV RVDGFQDAHL YETVFPVDS PILRDLFLSP DHRHIYLLSE	480
KQVSLPVET CEYQSCAAC LGSQDPHGW CVLRHRCRE GACLGASAPH GFAEELSKCV	540
QVRVRNVVS VTSRQVQLTV TLHNVPDLA GVSCAFEAFA ENEAVLLPSG ELLCPSFSLQ	600
ELRALTRGHG ATRTVRLQLL SKETGVRFAG ADFVFNCSV LQSCMCSVGS PYPCHWCKYR	660
HTCTSRPHFC SFQEGRVHSP EGCEIILPSG DLLIPVGMVQ PLTLRAKNLP QPQSQGKNYE	720
CVVRVQGRQQ RVPAVRFNSS SVQQQNASYS YEGDEHGDE LDFSVVMGDG FPIDKPFSPR	780
ALLYKQWQR FSCGLCLRAD PRFNCWCIS EHRQLRTHC PAKPTNMHML SQRGTRCSHP	840
RITQIHLVPG KREGTRVTI VGENLGLLR EVGLRVAGVR CNSIPAAYIS AERIVCEMEE	900
SLVPSPPPGP VELCVGDCA DFRQTGEQVY SFVTPFDQV SPSRGPASGG TRLTISGSSL	960
DAGSRVTVTV RDEECQFVRR DAKAIVCISP LSTLGPSEQP ITLAIDRANI SSPGLIYTTT	1020
QDPTVTRLEP TWSIINGSTA ITVSGTHLLT VQEPFRVRAKY RGLIETNTCQ VINDTAMLCK	1080
APGIFLGRFP PRAQGEHPDE FGFLLHVTQ ARSLNRSFT YYPDPFSEPL GPGSVLDVKP	1140
GSHVVLGKKN LIPAAAGSSR LNYTVLIGGG PCSLTVSDTQ LLCDSPSQTG RQPVMVLGG	1200
LEFWLGLTHI SAERALTLP MMGLAAGGL LLLAITAVLV AYKRKTQDAD RTLKRLQLQM	1260
DNLESRVALE CKFAFELQT DINELTNHMD EVQIPFLDYR TYAVRVLPFG IEAHPVLKEL	1320
DFFPVEKAL RLFQQLLHNR AFVLTFTIHTL EAQSFSPMRD RGTVASLTMV ALQSRDYAT	1380
GLLKQLLADL IEKNLESKNH PKLLLRRTES VAERKMLTNP FTLHLKFLKE CAGEPILFLY	1440
CAIKQMEKG PIDAITGEAR YSLSEKLIIR QQIDYKTLTL HCVCPENEGS AQVPRVLCNC	1500
DSITQAKDKL LDTYKIGIPY SQRPKAEMD LEWRQGRMTR IILQDEVDTT KIECDWKRLLN	1560
SLAHYQVTDG SIVALVPEQV SAYNMAASFT PTRSLRSRYES LLRTASSPDS LRSRAMITP	1620
DQETGTLKWH LVKNHHDHADH RBGDRGSKMV SEIYLTRLLA TRGTLQKFPVD DLFETVFTA	1680
HRGSALPLAI KYMFDLDBQ ADQRQISDPD VRHTWKSACL PLRFVWVNIK NPQFVFDIHK	1740
NSITDAQLSV VAQTFMDSCS TSEHRLGKDS PSNKLLYAKD IPNYKSWER YYRIDAKMAS	1800
ISDQMDAYL VEQSRHLHAD FSVLSALNEL FYFYTKYRQE IALTALDRDAS CRKHLRQKQL	1860
EQIISLVSSD S	1871

FIG. 2. Amino acid sequence of the *SEX* protein. The first methionine is followed by a signal sequence (underlined) located between Val⁴ and Gly¹⁹. Thick bar, putative transmembrane domain from Ala¹²¹⁵ to Ala¹²⁴¹. The nine potential N-glycosylation sites are marked with solid circles, while cysteine residues are marked with solid triangles. Boxes mark the region highly conserved among members of the *SEX* family.

cellular domains of these proteins is >20%, and notably the cysteines occupy conserved relative positions (Fig. 3). Furthermore, in *SEX* and in all members of the *MET* family the cysteines are grouped in regularly spaced clusters. One of these is found once in the extracellular domain of the HGF receptor family, whereas it occurs three times in *SEX*, suggesting a triplication of the corresponding DNA sequence. We refer to this cysteine cluster as the MRS motif (*MET*-related sequence).

The large intracellular portion of *SEX* is 630 amino acids long and, surprisingly, it lacks a TK domain; nor does it share homology with any known functional motif. Based on these structural features, we concluded that the newly identified *SEX* gene encodes a single-pass transmembrane protein sharing significant homology with the extracellular domain of the receptors encoded by the *MET* oncogene family.

***SEX* Is the Prototype of a Gene Family Encoding Structurally Related Proteins.** Using as a probe a subclone of *SEX* cDNA spanning the transmembrane domain, two other closely related sequences were identified from a skeletal muscle cDNA library. The cDNAs, named m9 and m13, were 2.2 and 1.5 kb long and proved to be 75% and 72% identical to *SEX*, respectively. A third *SEX*-related cDNA sequence (Fb7) was meanwhile isolated by serendipity from a human fetal brain cDNA library. This was sequenced and found to be closely related to *SEX*. By further screening of cDNA libraries with Fb7, two other overlapping clones were identified, yielding a 5.3-kb cDNA sequence. This included a poly(A) tail at its 3' end and an ORF predicting the intracellular, the transmembrane, and part of the extracellular domain of a protein that shares 37% overall identity with *SEX*. Additional partial clones have also been observed, which indicate that *SEP* exists in alternatively spliced isoforms. The three *SEX*-related sequences were named *SEP* (Fb7), *OCT* (m9), and *NOV* (m13), according to the Latin ordinals following six. The genes corresponding to *SEX*, *OCT*, and *NOV* mapped on different chromosomes. By probing a panel of human-hamster hybrids, we determined that *SEP* maps to the short arm of human chromosome 3 (p1.4-pter), *NOV* maps to the long arm of chromosome 3 (q2.1-qter), and *OCT* maps to human chromosome 1.

The sequence encoding the extracellular domains of *SEP*, *OCT*, and *NOV* could be aligned with that of *SEX*, as well as with the *MET* family of receptors, confirming the distinctive conservation of cysteine residues (Fig. 3). The *SEP* extracellular domain also includes two MRS motifs, described above as typical features of the *MET* and *SEX* protein families.

The intracellular domains of *OCT* and *NOV* are 76% and 81% identical to that of *SEX*, respectively. The identity

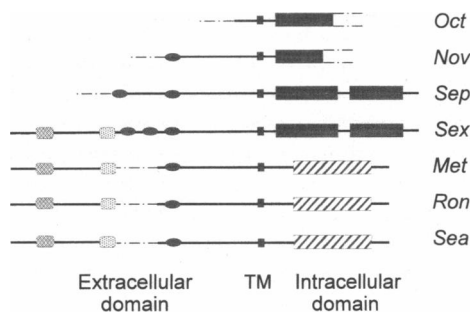


FIG. 3. Close and distant relatives of the *SEX* protein. Transmembrane domains are aligned and indicated with solid squares (TM). Dotted-dashed lines indicate gaps introduced to show the best possible alignment. In the extracellular domains, the repeated MRS motifs are sketched by solid round boxes; the other cysteine-rich motifs are stippled. The intracellular TK domains are indicated with hatched boxes and the conserved *SEX* domain is indicated with solid rectangles.

between *SEX* and *SEP* cytoplasmic domains is 48%; comparing these sequences, two homology regions of approximately 320 and 150 amino acids can be identified. These are separated by a divergent insert of ≈ 70 residues.

The *SEX* Family of Genes Is Predominantly Expressed in Fetal Epithelial and Neural Tissues. Expression of the specific mRNA of *SEX*, and of the other members of the family, was examined by Northern blot analysis in a large panel of human adult and fetal tissues. To avoid cross-hybridization the probes were prepared from the most divergent portions of the cDNAs, corresponding to the extracellular domains of the proteins. The four genes of the *SEX* family encode long mRNAs (7.5–9 kb) and are widely expressed in adult tissues, with prominent levels in brain, kidney, ovary, and testis. Several cell lines of neural origin express the *SEX* transcript, including IMR-32 human neuroblastoma-derived cells (data not shown). Northern blot analysis of RNAs from fetal tissues (Fig. 4) showed that *SEX*, *NOV*, and *OCT* have a predominant expression in fetal neural tissues, whereas *SEP* is most highly expressed in fetal kidney.

***SEX* Expression Is Regulated During Development of the Mouse Nervous System.** We studied the expression of *SEX* during mouse brain development by *in situ* hybridization analysis of mouse embryos from 9.5 to 17.5 days postcoitus (dpc) and of brain of adult mice. A murine *Sex* cDNA fragment was synthesized by RT-PCR, and this was used to prepare an antisense RNA probe. *Sex* transcripts were detectable as early as 9.5 dpc in the ventral region of the spinal cord, where motoneurons start differentiating (Fig. 5A). At 10.5 dpc, hybridization becomes very strong in the whole ventral spinal cord and in spinal ganglia and becomes visible in the neuroepithelia of the encephalic vesicles. At 11.5 dpc, the postmitotic neurons at the base of the telencephalon are positive for expression of *Sex*, as are the cranial ganglia, the olfactory neuroepithelia, and the infundibulum (Fig. 5B). At 12.5 dpc, *Sex* changes its expression pattern and appears uniformly distributed in the ventral and in the dorsal portions of the spinal cord. At later stages (14.5 dpc; Fig. 5C), the whole neuronal layer of the telencephalon is positive and the neuronal layer of the retina shows a very specific and strong hybridization. In the adult mouse brain (Fig. 5D), *Sex* transcription is elevated in the cortex of the frontal lobe and in the olfactory bulbs, spinal cord, and ganglia.

Structure and Biosynthesis of the *SEX* Protein. Three overlapping *SEX* cDNA clones were joined through *Eco*RI and *Hind*III internal restriction sites to assemble the entire ORF encoding *SEX*. The cDNA was subcloned in eukaryotic expression vectors and transfected in human kidney cells (BOSC23) or in murine neural-derived cells (Neuro2a). The synthesized protein was analyzed by Western blotting using

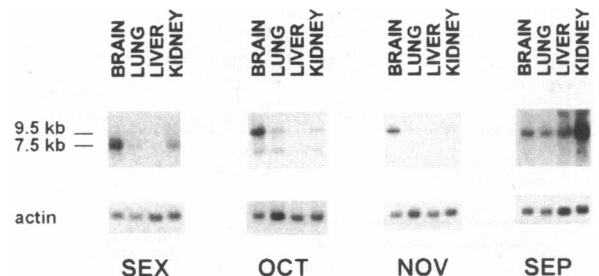


FIG. 4. Expression of the *SEX* family genes in human fetal tissues. Northern blots containing poly(A)⁺ RNA from a panel of fetal tissues were hybridized with cDNA probes from divergent sequences of the four genes of the *SEX* family. The length of the mRNA transcripts is shown alongside. *SEX*, *NOV*, and *OCT* have a predominant expression in neural fetal tissue, whereas *SEP* is eminently expressed by fetal kidney. The amount of β -actin mRNA present in each sample is shown at the bottom.

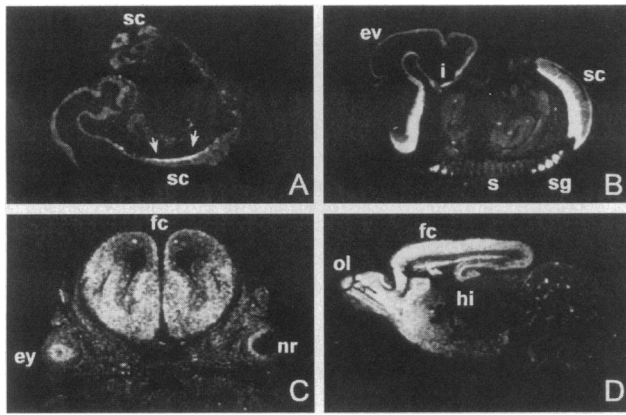


FIG. 5. Expression of *SEX* in the nervous system of the developing and adult mouse. *In situ* hybridization analysis was performed with a murine *Sex* RNA probe. Sections of mouse embryo at 9.5 dpc (A), 11.5 dpc (B), and 14.5 dpc (C); section of adult mouse brain. (D) A, B, and D are sagittal sections; C is a transverse section. sc, Spinal cord; s, somite; sg, spinal ganglia; ev, encephalic vesicles; i, infundibulum; ol, olfactory lobe; fc, frontal cortex; hi, hippocampus; ey, eye; nr, neural layer of the retina.

anti-*SEX* affinity-purified antibodies prepared as described. The protein showed a mass of 220 kDa (p220^{Sex}) when analyzed under reducing (Fig. 6, lane B) and nonreducing (data not shown) conditions. Treatment with tunicamycin, to abolish protein N-glycosylation, yielded a core precursor protein of \approx 200 kDa (lane C). Accordingly, the endogenous *SEX* protein, analyzed by Western blots of the human neuroblastoma-derived cell line IMR-32, consisted of a doublet of specific bands of 220 and 200 kDa corresponding to the mature and to the precursor protein (lane D). The subcellular localization of *SEX*, transfected or endogenously expressed, was studied by using surface-biotinylated cells and immunoprecipitation with anti-*SEX* antibodies (lane G). These data show that *SEX* is a monomeric glycosylated protein exposed at the cell surface.

The Intracellular Domain of *SEX* Is Not Endowed with Kinase Activity. p220^{Sex} was assayed for tyrosine phosphorylation in intact cells and for kinase activity *in vitro* in the presence of [γ -³²P]ATP and various buffers permissive for the reaction (10). Similar analyses were performed upon stimulation with HGF on a MET-*SEX* chimeric protein containing

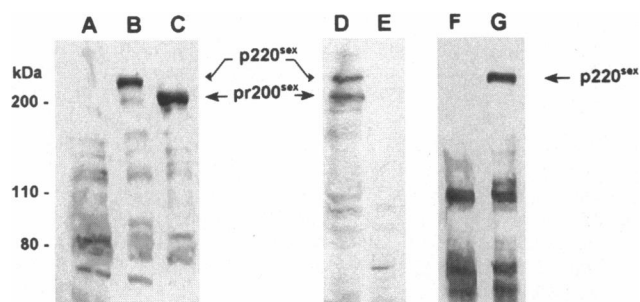


FIG. 6. Biosynthesis and cell surface expression of the *SEX* protein. Lanes A and F, extracts of mock-transfected cells; lanes B, C, and G, extracts of *SEX*-transfected cells; lanes D and E, extracts of IMR32 cells expressing endogenous *SEX*. Western blots were decorated with antibodies raised against the C-terminal sequence of the *SEX* protein expressed in bacteria (lanes A–D) or with preimmune serum (lane E). Transfected cells were left untreated (lane B) or were incubated with tunicamycin in order to abolish protein glycosylation (lane C). Molecular masses of the precursor and of the mature form of *SEX* are indicated. Lanes E–F, cells were surface biotinylated and immunoprecipitated with anti-*SEX* antibodies; Western blot was probed with streptavidin peroxidase.

the extracellular domain of the HGF receptor and the intracellular domain of *SEX*. Neither the full-size nor the chimeric *SEX* was ever found tyrosine phosphorylated or showed detectable kinase activity (data not shown). Notably, since the MET-*SEX* chimera was expressed in cells provided with endogenous MET, heterodimers were detected by immunoprecipitation and Western blotting with either antibody. In these heterodimers, only the β chain of MET was tyrosine phosphorylated. No heterodimers were found when the same cells were transfected with full-size p220^{Sex}.

Moreover, we investigated the function of *SEX* by attempting activation through constitutive dimerization of the intracellular domain, by analogy to what is known for the MET-HGF receptor family. A chimeric fusion protein was used containing the intracellular domain of *SEX* and a sequence derived from 5' of the *TPR* gene (TPR-*SEX*). This sequence includes two leucine zipper motifs, and it is known to induce constitutive dimerization and activation of TKs (33). The TPR-*SEX* homodimers were immunoprecipitated and tested in the kinase assay; also, in this case no phosphorylation was ever observed. After cotransfection of the cDNAs encoding TPR-*SEX* and TPR-MET chimeras (approximately 75 and 65 kDa, respectively), heterodimers of the two cytoplasmic domains were immunoprecipitated and assayed for kinase activity. The TK domain of MET could not phosphorylate the cytoplasmic domain of *SEX* (data not shown).

We finally tested the transforming potential of *SEX* by focus formation assay as described (34). Neither *SEX* nor TPR-*SEX* could induce transformed foci formation (data not shown).

DISCUSSION

This paper identifies a family of human genes (*SEX*, *SEP*, *OCT*, and *NOV*) encoding single-pass transmembrane proteins with large extracellular and cytoplasmic domains. Most of the proteins with these structural features act as receptors for environmental signals and can be grouped according to their structure (5, 35). A large extracellular domain is common to all transmembrane receptors, featuring the binding site for the specific ligand, either a cell matrix-associated molecule or a soluble factor. The size of the intracellular domain varies according to different structural and/or signaling functions. Integrins, which are receptors for matrix-bound proteins, and selections and cadherins, which behave as cell-cell recognition molecules, feature small intracellular domains (see, e.g., ref. 3). Exceptions to this rule are known (36). Hormone and growth factor receptors have a large intracellular domain, which either is endowed with intrinsic enzymatic activity (mostly tyrosine- or serine/threonine-kinase) or associates with catalytic cytoplasmic molecules (8). Although several transmembrane receptors have been described, fewer ligands have so far been identified. Functions have been assigned to the putative receptors based on structural similarities (37, 38).

A putative receptor function for the protein encoded by the *SEX* gene is further suggested by its significant similarity to members of the *MET* gene family; these include the receptor for the growth factor HGF (MET), the receptor for MSP (RON), and the orphan receptor encoded by the protooncogene *SEA* (11–13). When the sequences of *SEX*, *SEP*, *MET*, *RON*, and *SEA* were aligned, it was noted that most of the cysteines in the extracellular domains have conserved their relative positions (Fig. 3). Interestingly, these residues can be grouped in clusters, one of which is repeated three times in *SEX* and twice in *SEP*. The repeated duplication of the corresponding DNA sequences during evolution suggests a functional role for this motif, which we name the MRS motif. The alignment of cysteines was first observed when comparing the sequences of MET and its close relative RON and is considered an important structural feature of the *MET* gene family (12). The relative positions of the cysteine residues are

crucial for forming the intrachain disulfide bridges that stabilize the three-dimensional conformation of the extracellular domain containing the ligand binding site. This may account for the ability of MET and RON receptors to bind HGF and MSP, respectively, which themselves are structurally related growth factors containing conserved kringle domains (39, 40). SEX and its sister proteins SEP, OCT, and NOV could therefore act as receptors for still unidentified soluble factors, possibly related to HGF and MSP and/or characterized by the presence of kringle domains.

The four genes of the SEX family are widely expressed in adult tissues. However, Northern blot analysis of fetal tissues revealed that their expression is much more restricted during development. SEX, OCT, and NOV showed a predominant expression in neural tissue, whereas SEP was most highly expressed in fetal kidney. These data indicate that during development, the functional role of these putative receptors is restricted to specific tissues. This was also confirmed by *in situ* hybridization analysis, showing that the expression pattern of SEX in the mouse nervous system is highly regulated during the initial stages of neuronal differentiation. SEX mRNA is initially transcribed in motoneurons of the spinal cord, and subsequently in the ganglia, in neural epithelia of sensory organs, in the encephalic vesicles, and in the telencephalon. The high level of SEX gene expression in postmitotic neurons suggests the possible involvement of the encoded putative receptor in initiating or maintaining the embryonal neural differentiation.

The putative signaling functions of SEX-related proteins are likely mediated by their conserved cytoplasmic domains. The corresponding amino acid sequences are >50% identical among members of this family, thus establishing a distinctive structural feature totally unrelated to the TK domains of growth factor receptors or to any other known catalytic sequence. We call this unique structure sex domain. A TK with a noncanonical sequence in the catalytic domain has recently been described (41). However, after careful testing, any kinase activity was ruled out. We conclude that the SEX protein exerts its function(s) either via an unexpected enzymatic activity or by association with still unidentified intracellular transducers.

We thank F. liBergolis and L. Palmas for technical assistance, R. M. Freeman, Jr., for isolating the Fb7 cDNA, and Dr. M. Rocchi for kindly providing the DNA used for chromosomal mapping. This work was supported by funds granted to D.T. by the P.F. "Ingegneria Genetica" Consiglio Nazionale delle Ricerche (CNR) and by Telethon; to P.M.C. and O.C. by the Italian Association for Cancer Research and the P.F. ACRO (CNR); and to B.G.N. by National Institutes of Health Grant CA49152. E.M. and S.B. are fellows of the Ph.D. program in Genetics of the University of Pavia; P.L. was recipient of a fellowship from the Italian Association for Cancer Research (AIRC).

1. Cross, M. & Dexter, T. M. (1981) *Cell* **64**, 271–280.
2. Geiger, B. & Ayalon, O. (1982) *Annu. Rev. Cell Biol.* **8**, 307–332.
3. Heynes, R. O. (1992) *Cell* **69**, 11–25.
4. Baron, M., Norman, D. G. & Campbell, I. D. (1991) *Trends Biochem. Sci.* **16**, 13–17.
5. Fantl, W. J., Johnson, D. E. & Williams, L. T. (1993) *Annu. Rev. Biochem.* **62**, 453–481.
6. Ven der Geer, P., Hunter, T. & Lindberg, R. A. (1994) *Annu. Rev. Cell Biol.* **10**, 251–337.
7. Massague, J. (1992) *Cell* **69**, 1067–1070.
8. Stahl, N. & Yancopoulos, G. D. (1993) *Cell* **74**, 587–590.
9. Park, M., Dean, M., Kaul, K., Braun, M. A., Gonda, A. & Vande Woude, G. F. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6379–6383.
10. Giordano, S., Ponzetto, C., DiRenzo, M. F., Cooper, C. S. & Comoglio, P. M. (1989) *Nature (London)* **339**, 155–156.
11. Bottaro, D. P., Rubin, J. S., Faletto, D. L., Chan, A. M., Kmieciak, T. E., Vande Woude, G. F. & Aaronson, S. A. (1991) *Science* **251**, 802–804.
12. Ronsin, C., Muscatelli, F., Mattei, M. G. & Breathnach, R. (1993) *Oncogene* **8**, 1195–1202.
13. Huff, J. L., Jelinek, M. A., Borgman, C. A., Lansing, T. J. & Parsons, J. T. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6140–6144.
14. Naldini, L., Vigna, E., Narsimhan, R., Gaudino, G., Zarnegar, R., Michalopoulos, G. K. & Comoglio, P. M. (1991) *Oncogene* **6**, 501–504.
15. Gaudino, G., Follenzi, A., Naldini, L., Collesi, C., Santoro, M., Gallo, K. A., Godowski, P. J. & Comoglio, P. M. (1994) *EMBO J.* **13**, 3524–3532.
16. Ponzetto, C., Bardelli, A., Zhen, Z., Maina, F., Dalla Zonca, P., Giordano, S., Graziani, A., Panayotou, G. & Comoglio, P. M. (1994) *Cell* **77**, 261–271.
17. Weidner, K. M., Sachs, M. & Birchmeier, W. (1993) *J. Cell Biol.* **121**, 145–154.
18. Bussolino, F., Di Renzo, M. F., Ziche, M., Bocchietto, E., Olivero, M., Naldini, L., Gaudino, G., Tamagnone, L., Coffar, A. & Comoglio, P. M. (1992) *J. Cell Biol.* **119**, 629–641.
19. Galimi, F., Bagnara, G. P., Bonsi, L., Cottone, E., Follenzi, A., Simeone, A. & Comoglio, P. M. (1994) *J. Cell Biol.* **127**, 1743–1754.
20. Di Renzo, M. F., Narsimhan, R. P., Olivero, M., Bretti, S., Giordano, S., Medico, E., Gaglia, P., Zara, P. & Comoglio, P. M. (1991) *Oncogene* **6**, 1997–2003.
21. Di Renzo, M. F., Olivero, M., Giacomini, A., Porte, H., Chastre, E., Mirossay, L., Nordlinger, B., Bretti, S., Bottardi, S., Giordano, S., Plebani, M., Gespach, C. & Comoglio, P. M. (1995) *Clin. Cancer Res.* **1**, 147–154.
22. Rong, S., Segal, S., Anver, M., Resau, J. H. & Vande Woude, G. F. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4731–4735.
23. Pear, S. W., Nolan, G. P., Scott, M. L. & Baltimore, D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8392–8396.
24. Bione, S., Tamanini, F., Maestrini, E., Tribioli, C., Poustka, A., Torri, G., Rivella, S. & Toniolo, D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10977–10981.
25. Tribioli, C., Mancini, M., Plassart, E., Bione, S., Rivella, S., Sala, C., Torri, G. & Toniolo, D. (1994) *Hum. Mol. Genet.* **3**, 1061–1067.
26. Maestrini, E., Patrosso, C., Mancini, M., Rivella, S., Rocchi, M., Rapetto, M., Villa, A., Fratini, A., Zoppé, M., Vezzoni, P. & Toniolo, D. (1993) *Hum. Mol. Genet.* **2**, 761–766.
27. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
28. Crepaldi, T., Pollack, A. L., Prat, M., Zborek, A., Mostov, K. & Comoglio, P. M. (1994) *J. Cell Biol.* **125**, 313–320.
29. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
30. van Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
31. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
32. Ponzetto, C., Giordano, S., Peverali, F., Della Valle, G., Abate, M. L., Vaula, G. & Comoglio, P. M. (1991) *Oncogene* **6**, 553–559.
33. Rodriguez, G. A. & Park, M. (1993) *Mol. Cell. Biol.* **13**, 6711–6722.
34. Zhen, Z., Giordano, S., Longati, S., Medico, E., Campiglio, M. & Comoglio, P. M. (1994) *Oncogene* **9**, 1691–1697.
35. Bazan, J. F. (1993) *Curr. Biol.* **3**, 603–606.
36. Hogervorst, F., Kiuman, I., Von Dem Borne, A. E. G. & Sonnenberg, A. (1990) *EMBO J.* **9**, 765–770.
37. Stitt, T. S., Conn, G., Gore, M., Lai, C., Bruno, J., *et al.* (1995) *Cell* **80**, 661–670.
38. Nagata, S. & Golstein, P. (1995) *Science* **267**, 1449–1456.
39. Nakamura, T., Nishizawa, T., Hagiya, M., Seki, T., Shimonishi, M., Sugimura, A., Tashiro, K. & Shimizu, S. (1989) *Nature (London)* **342**, 440–443.
40. Han, S., Stuart, L. A. & Friezner Degen, S. J. (1991) *Biochemistry* **30**, 9768–9780.
41. Beller, J. F., Larochelle, W. J., Chedid, M., Tronick, S. R. & Aaronson, S. A. (1994) *Mol. Cell. Biol.* **14**, 982–988.