GENOMICS **47,** 238–245 (1998)
ARTICLE NO. GE975078

Human NRD Convertase: A Highly Conserved Metalloendopeptidase Expressed at Specific Sites during Development and in Adult Tissues

P. Fumagalli,* M. Accarino,* A. Egeo,† P. Scartezzini,† G. Rappazzo,‡ A. Pizzuti,§ V. Avvantaggiato,¶ A. Simeone,¶ G. Arrigo,‖ O. Zuffardi,‖ S. Ottolenghi,* and R. Taramelli‡,[1]

**Dipartimento di Genetica e di Biologia dei Microrganismi, Universita degei Studi di Milano, 20133 Milano, Italy; †Divisione di Pediatria, E. O. Ospedali Galliera, Genoa, Italy; ‡Dipartimento di Biologia Animale, Catania, Italy; §Clinica Neurologica, Ospedale Policlinico, Milano, and IRCCS.CSS, S. Giovanni Rotondo, Italy; ¶Istituto Internazionale di Genetica e Biofisica, Naples, Italy; and ‖Biologia Generale e Genetica Medica, Pavia, and Lab. Citogenetiica Ospedale S. Raffaele, Milan, Italy*

Received June 16, 1997; accepted October 17, 1997

We report the cloning of the human homologue of the rat metalloprotease N-arginine dibasic convertase (NRD convertase). This endopeptidase is responsible for the processing, at the Arg-Lys dibasic site on the N-terminal side of the arginine residue, of propeptides and proproteins. Comparisons of the human and rat full-length cDNAs show similarity and identity of 94 and 91%, respectively. In humans NRD convertase is predominantly expressed in heart, skeletal muscle, and testis. We have also studied the expression of this gene in mouse at various developmental stages and found that the neural tissue is the almost exclusive site of expression in early development (between E 10.5 and E 16.5). To gain information about the possibility that defects in this gene are linked to inherited neuromuscular disorders, we determined the chromosomal location of the human NRD convertase gene by FISH analysis, showing that the gene resides at 1p32.2.

INTRODUCTION

Processing of precursor molecules into biologically active forms is a widespread mechanism in both lower and higher organisms (Barrett, 1979). For secretory proteins in particular this conversion entails posttranslational modifications, such as limited proteolysis, brought about by presumed selective peptidases (Barrett, 1979; Bond and Butler, 1987). According to the type of reaction catalyzed, these enzymes are known as exopeptidases or endopeptidases (Barrett, 1994). Knowledge of the structure of the active center of the endopeptidases has allowed subdivision of them into serine, cysteine, aspartic, and metalloendopeptidases (Barrett, 1994). Within the latter class of processing enzymes several members have been characterized, mainly on the basis of the conserved pentapeptide Zn-binding motif which is similar to that found in thermolysin (consensus HEXXH) (Barrett, 1994; Rawlings and Barrett, 1993, 1995). Recently, a separate group of metalloendoproteases (belonging to the family M16-pitrilysin) that are structurally unrelated to thermolysin and show a proposed Zn-binding motif HXXEH has been identified (Rawlings and Barrett, 1995). Members of this group include human, rat, and Drosophila insulinase (IDE; EC 3.4.99.45) (Affholter *et al.*, 1988; Baumaister *et al.*, 1993), *Escherichia* coli protease III (pitrilysin; EC 3.4.99.44) (Finch *et al.*, 1986), and endopeptidases involved in the processing of mitochondrial precursor proteins (Pfanner and Neupert, 1990; Meulenberg *et al.*, 1992; Kleiber *et al.*, 1990). In addition, within the M16-pitrilysin family, rat N-arginine dibasic convertase (NRD convertase) was recently added as the single member of a new subfamily characterized by an Asp-Glu-rich segment inserted very close to the catalytic site (Chesneau *et al.*, 1994; Pierotti *et al.*, 1994). We present here the cloning and mapping of a human cDNA encoding a sequence highly similar (93%) to the rat N-arginine dibasic convertase (NRD), implying that it is the human homologue of the rodent enzyme. The expression of the human mRNA in adult tissues is highly restricted to the heart, skeletal muscle, and testes. In fetal tissues the NRD convertase expression is predominant in the central nervous system (CNS), placenta, and heart. We also show that in mouse, the embryonic nervous system is the predominant site of expression of this mRNA at early stages of development.

MATERIALS AND METHODS

Preparation of cDNA and YAC DNA. cDNA was synthesized from 5 ug of poly(A)$^+$ RNA purified from fetal heart using a cDNA synthesis kit (Amersham). Double-stranded cDNA was digested with *Mbo*I and ligated to a linker–adaptor oligonucleotide (5′ GATCGAATTCACTCGAGCATCAGG 3′, 3′ CTTAAGTGAGCTCGTAGTCC 5′) as

[1] To whom correspondence should be addressed at Dipartimento di Genetica e di Biologia dei Microrganismi, Via Celoria 26, 20133 Milan, Italy. Telephone: 39-2-26605225. Fax: 39-2-2664551. E-mail: taramel@imiucca.csi.unimi.it.

0888-7543/98 $25.00

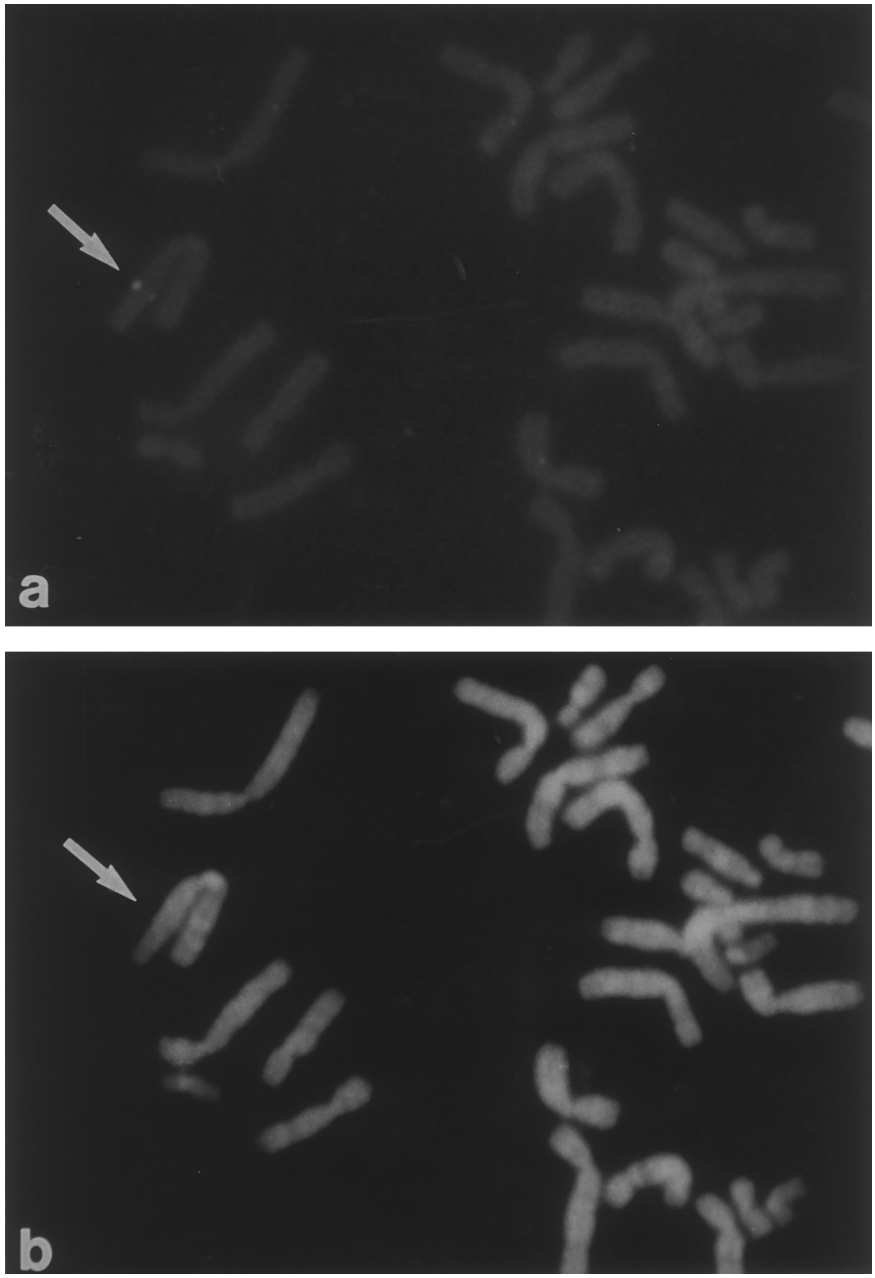

FIG. 1. (**a**) A partial metaphase counterstained with propidium iodide showing chromosome 1 with hybridization signals (arrow); (**b**) the same after DAPI staining.

previously described (Morgan *et al.*, 1992). Approximately 0.5 ng of cDNA was PCR amplified using the relevant primer for the linker–adaptor oligonucleotide.

Chromosome 21 YAC 761B5 was tested for chimerism by FISH analysis and found to be nonchimeric. Pulsed-field gel purified YAC genomic DNA was partially digested with *Mbo*I and ligated to a linker–adaptor oligonucleotide (5′ GATCTCGACGAATTCGTGAGACCA 3′, 3′ AGCTGCTTAAGCACTCTGGT 5′). The linked genomic segments were PCR amplified using a 5′ biotinylated primer whose sequences matched one of the linker–adaptor arms (Morgan *et al.*, 1992). The cDNA selection procedure was essentially as reported in Morgan *et al.* (1992).

Fluorescence in situ hybridization. The human full-length NRD convertase cDNA was biotinylated by nick-translation in the presence of biotin-14–dUTP using the nick-translation kit (Boehringer Mannheim). FISH and the screening of lymphocyte metaphases from a normal male were performed as described (Banfi *et al.*, 1997). For the screening of the slides only those chromosomes with two signals present on both chromatids at the same band position were taken into consideration. The probe was considered mapped when the same chromosome band showed signals in no less than 20% of 200 metaphase examined. The images were taken as seen under microscopy, without any computer elaboration.

Preparation of RNA probes for in situ hybridization. The mouse 606-bp NRD convertase cDNA probe was obtained by screening a mouse skeletal muscle cDNA library with a human cDNA fragment spanning nucleotides 1590 to 2690. Alignment of both sequences revealed 90.5% identity. Transcription reactions with T7 or SP6 polymerase (Riboprobe Kit, Promega Biotech), were carried out in the presence of [^{35}S]CTP (Amersham). The template was then degraded with RNase-free DNase (Pharmacia), and the labeled RNA was purified through a Sephadex G-50 column. The transcripts were progres-

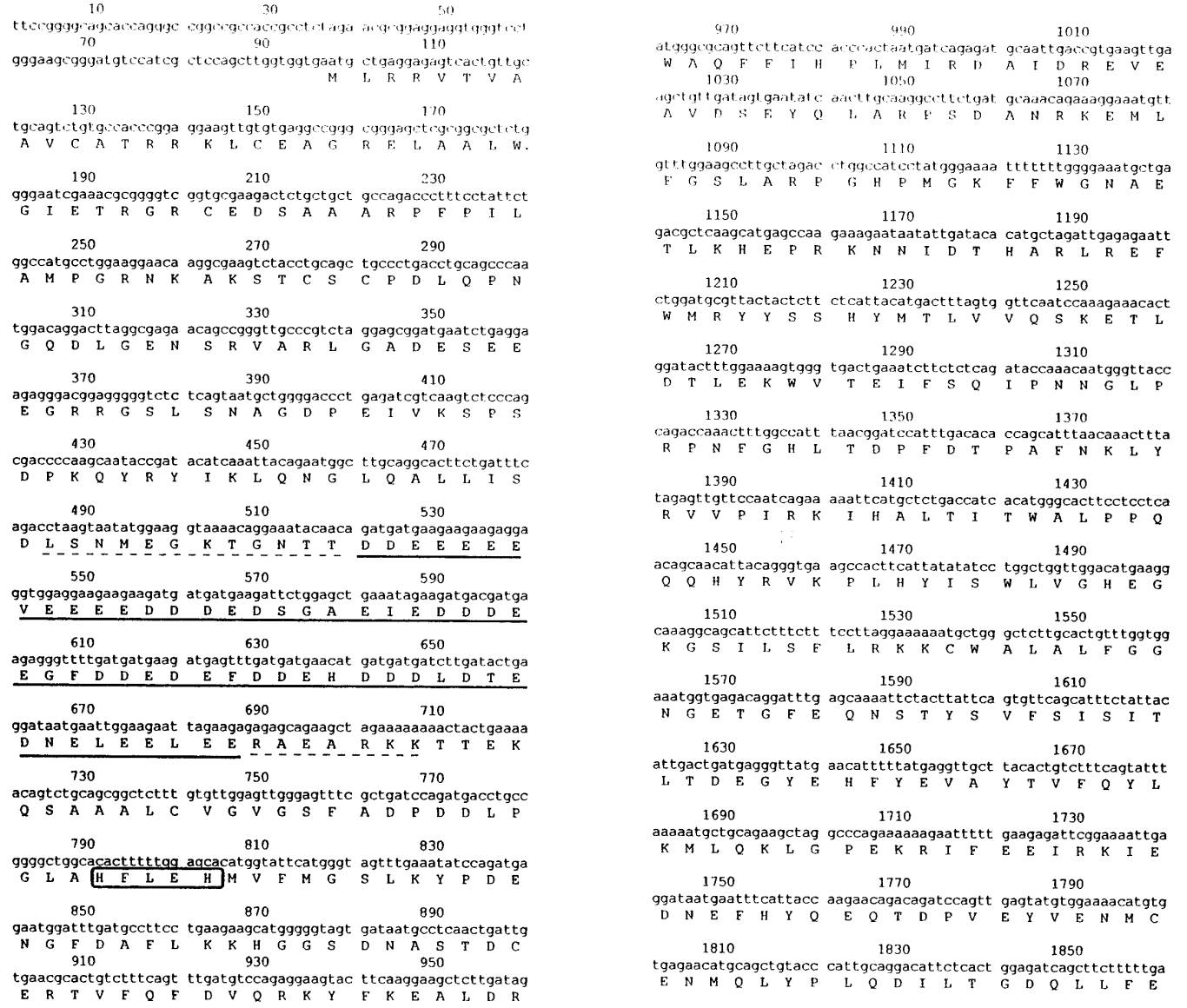

FIG. 2. Nucleotide and deduced amino acid sequences of the human NRD convertase. The two anchor motifs are underlined with broken lines; the catalytic site is boxed; and the acidic stretch is underlined.

sively degraded to an average length of 150 nucleotides by random alkaline hydrolysis to improve access to RNA *in situ*. The probes were dissolved at a working concentration of 1×10^5 cpm/μl in hybridization mix (Wilkinson and Green, 1990).

In situ hybridization. *In situ* hybridization was carried out as described by Wilkinson and Green (1990) with minor modifications. Thirty microliters of the appropriate probe in the hybridization mix was added to each slide. Hybridization was carried out overnight at 55°C. The slides were then washed under stringent conditions (65°C, 2× SSC, 50% formamide) and treated with RNase to remove unhybridized and nonspecifically bound probe. Autoradiography was performed with Kodak NT/B2 emulsion. Exposure times were between 5 and 12 days. After developing, sections were stained in 0.02% toluidine blue and mounted in De Pex (Serva). Sections were examined and photographed using a Zeiss SV11 microscope with both dark- and bright-field illumination.

RESULTS

In the course of a project aimed at cloning genes expressed in fetal heart that map to the Down critical region (21q22.3), cDNA hybridization selection has been applied to YAC 761B5 (CEPH; 950 kb), known to be located in the above genomic region (Chumakov *et al.*, 1992). As a cDNA source fetal myocardium and endocardium tissues were used. Following identification of a 250-bp fragment hybridizing to the YAC recombinant, three overlapping clones constituting a full-length cDNA were obtained and characterized. Using the full-length cDNA, however, Southern blot analysis of a panel of human rodent hybrids (not shown) and more directly FISH analysis (Fig. 1) revealed that the gene encoding this cDNA resides on chromosome 1p32.2. Under the FISH hybridization conditions (37°C, 50% formamide/2× SSC), the absence of other clusters of signals along the chromosomes excluded the presence of homologous loci. Figure 2 shows the combined sequence of the three overlapping clones, which

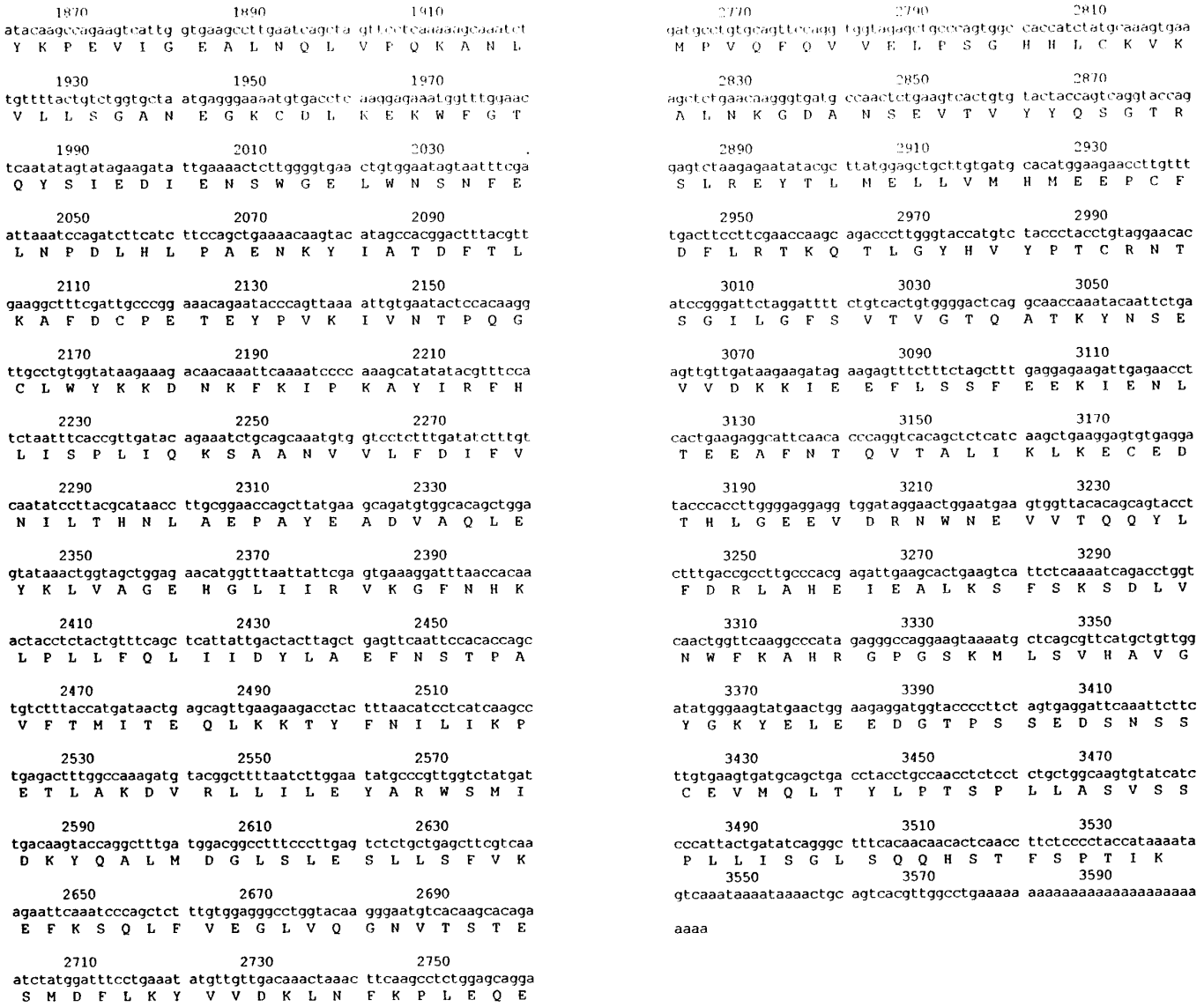

FIG. 2—*Continued*

yielded a cDNA of 3604 bp (Accession No. U64898), in good agreement with the size of the mRNA as detected by Northern analysis (see below). The cDNA encodes an open reading frame of 1147 amino acids plus 97 and 40 nucleotides of 5′ and 3′ untranslated sequences, respectively. Two overlapping polyadenylation sites (AATAAA) are located 8 and 13 nucleotides downstream from the stop codon. Database searches showed that the human cDNA sequences are 91% identical to the rat NRD convertase cDNA. Alignment of amino acid sequences of the human cDNA with rat NRD convertase (data not shown) revealed a striking identity (91.6%) of the two sequences. The sequence comparison confirms that the translation initiation codon in the human sequence is at Met position 98. The major features of the rat NRD convertase are: (i) a 71-bp-long acidic stretch of residues (aa 139–209) where Asp and Glu are highly represented and (ii) a Zn binding domain (aa 244–248) that conforms to the consensus HXXEH as derived from other zinc-metalloendopeptidases. As can be seen from Fig. 2, the Zn binding motif present in the human sequence (aa 244–248) is perfectly conserved, whereas the acidic stretch is shortened by 13 residues. Moreover, the two short anchors that flank the acidic stretch are also conserved between the two species, with only two nonconservative substitutions in the human enzyme: Thr-Ala at position 140 and Ala-Val at position 199 in the upstream and downstream anchor motifs, respectively. Furthermore, a putative 19- to 21-residue signal peptide is present in the human as well as in the rat NDR convertase. Additionally, both human and rat NRD convertase proteins show a substantial homology in the zinc binding motif-containing region with other metalloendopeptidases, in particular the human, rat, and Drosophila insulin degrading enzyme proteins (IDEs) (44% identity) and the protease III from *E. coli* (pitrilysin) (40% identity). Much lower identity scores are found with two other

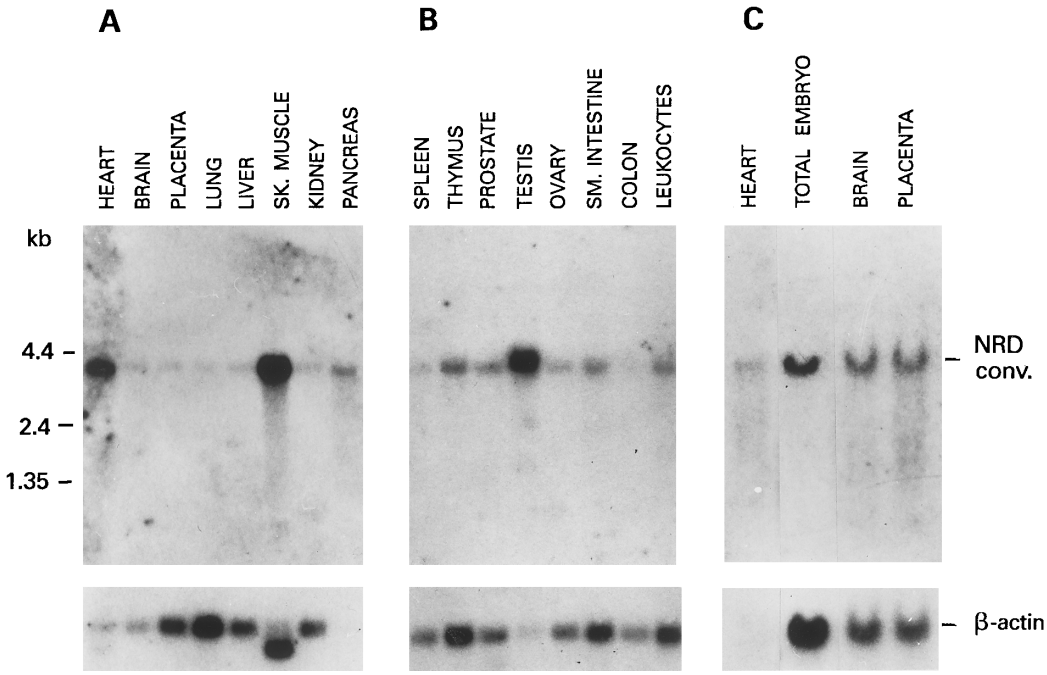

FIG. 3. Northern analysis of human RNA from adult and fetal tissues. Northern blots were hybridized with NRD convertase (**top**) and β-actin (**bottom**) cDNAs. (**A** and **B**) Clontech filters with adult poly(A)$^+$ RNA. Note that the faster actin band in the skeletal muscle lane is due to the muscle-specific isoform. (**C**) Fetal total RNA (tissues were analyzed at the 12th week of gestation).

members of the pitrilysin family, namely the *Klebsiella pneumoniae* open reading frame F of the pyrroloquinoline quinone operon, and *Bacillus subtilis* processing protease (27 and 21% identity, respectively) (data not shown).

Northern blot experiments (Figs. 3A–3B) showed that human NDR convertase is expressed as a 3.6-kb band mainly in adult heart, skeletal muscle, and testes and at much lower levels in other tissues including thymus, prostate, pancreas, ovary, small intestine, and leukocytes. The high level of human NRD mRNA expression in testes is in agreement with the data previously reported in the rat, where the testicular tissue was the starting source for the isolation of the NDR protein and subsequently the cDNA. In addition, the high expression in heart and muscle conforms to the procedure used for the isolation of this cDNA starting from fetal heart mRNA. In human fetal tissues expression is detected in the CNS, placenta, and heart (Fig. 3C). To better investigate the expression of this gene in vertebrate development, we performed both Northern analysis and *in situ* hybridization experiments in mouse embryonic tissues using a highly homologous partial mouse cDNA. The mouse cDNA sequences span the corresponding human sequences between nucleotides 1590 and 2590 (Accession No. U86112; for details see Materials and Methods). NRD convertase transcripts are already present at day 10.5 in the central nervous system and heart, whereas at later stages (12.5 and 14 d.p.c.) the skeletal muscle and liver are also sites of major expression (data not shown).

To have a more accurate expression profile, *in situ* hybridization between 7.5 and 17 days pc (E 7.5 to E 17) of murine embryonic development in sagittal, frontal (Fig. 4), and coronal (not shown) sections was carried out. At early stages (E 7.5 and E 8.5) NRD convertase transcripts are undetectable above the background signal (data not shown).

At E 10.5 NRD convertase mRNA was detectable at high levels in the CNS (Fig. 4) and in the spinal and cephalic ganglia. The high-level expression persisted in the same sites at E 12.5 and E 16.5 (Fig. 4). Transverse sections showing a magnification of the spinal cord and spinal ganglia revealed that NRD convertase transcripts were distributed homogeneously across these structures, involving all cell types (data not shown). In addition, lower levels of NRD convertase transcripts were detected at E 10.5–E 12.5 in the branchial arches (Figs. 4a, 4b, 4d, and 4f), the nasal pits and the olfactory epithelium (Fig. 4f), the epatobiliary complex (Figs. 4a, 4c, 4e, and 4f), the frontonasal prominence, the limb buds (Figs. 4a–4c), the gut, and the genital eminence (Fig. 4f). At E 16.5 NRD convertase expression was ubiquitously detectable, at levels somewhat variable but generally lower than those in the neural system (Fig. 4).

DISCUSSION

We report here the cDNA cloning of the human homologue of rat metalloprotease N-arginine dibasic convertase. This is an endopeptidase responsible for the processing, at the Arg-Lys dibasic site on the N-terminal side of the arginine residue, of propeptides and propro-

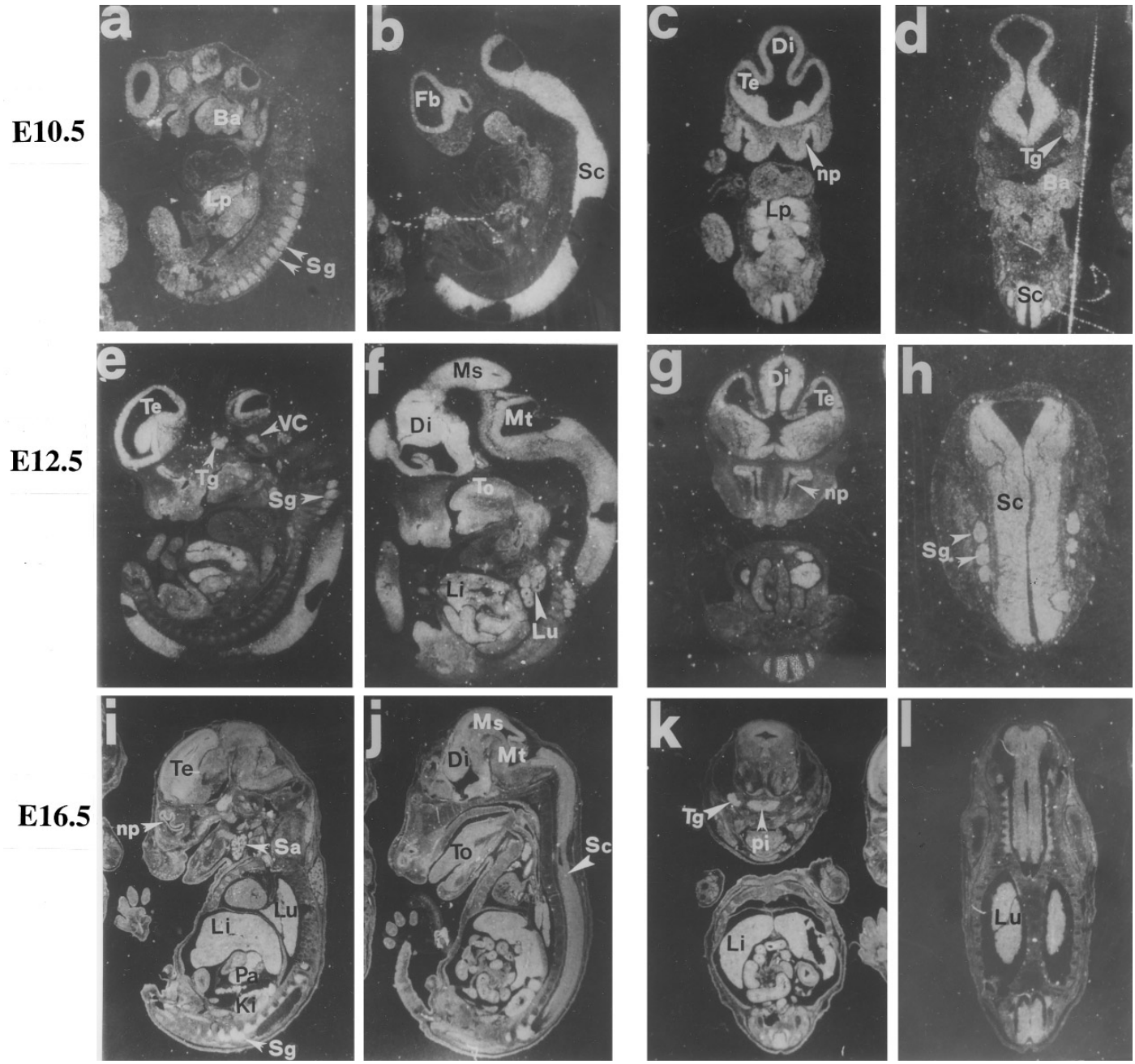

FIG. 4. NRD convertase expression during murine embryonic development. NRD convertase expression in sagittal (a, b, e, f, i, j) and frontal (c, d, g, h, k, l) sections of E 10.5 (a–d) E 12.5 (e–h) and E 16.5 (i–l) mouse embryos. Abbreviations are as follows: Fb, forebrain; Te, telencephalon; Di, diencephalon; Ms, mesencephalon; Mt, metencephalon; Sc, spinal cord; Sg, spinal ganglia; Tg, trigeminal ganglion; Ba, branchial arches; Lp, liver primordium; np, nasal pit; VC, vestibular/cochlear ganglion; To, tongue; Lu, lung; Li, liver; Ki, kidney; Sa, salivary gland; Pa, pancreatic primordium; pi, pituitary gland.

teins (Rawlings and Barrett, 1995; Chesneau *et al.*, 1994; Pierotti *et al.*, 1994). We cloned this cDNA fortuitously during a cDNA selection screening for sequences mapping to the Down region of chromosome 21, by virtue of limited homology of sequences of this cDNA to the chromosome 21 YAC clone employed (this type of artifact is often encountered when cDNA selection procedures are applied for transcript isolation using YAC or cosmid recombinants as templates) (Guimera *et al.*, 1996). The human NRD convertase cDNA was therefore mapped by FISH analysis and was found to be localized to the short arm of chromosome 1 at 1p32.2.

The full-length human NRD convertase cDNA shows features that closely resemble those of the rat cDNA with a similarity and identity of 94 and 91%, respectively (Pierotti *et al.*, 1994). The HXXEH consensus sequence typical of the active center of the M16-pitrilysin family of metalloendopeptidases (Rawlings and Barrett, 1995) is perfectly conserved between human and rat (HFLEH); other motifs are highly conserved,

including a putative signal peptide and a unique Cys residue (Cys-959 in the rat sequence, Cys-935 in the human) proposed to account for thiol dependence. The latter is also common to three other proteins (Drosophila, rat, human DEs) (Affholter *et al.*, 1988; Baumaister *et al.*, 1993). However, a substantial difference between rat and human NRD convertase sequences is found in the acidic segment located in the proximity of the catalytic domain of both proteins. The rat convertase shows a 71-residue acidic stretch rich in Asp and Glu, whereas the human enzyme is shortened by 13 residues. This large difference, within an otherwise highly conserved protein, is intriguing. It has been speculated that this acidic stretch, because of its vicinity to the putative catalytic domain, might be responsible for substrate recognition and binding; alternatively, the motif might play a role in the appropriate routing and targeting to a given subcellular compartment (Pierotti *et al.*, 1994). It will be interesting to compare the properties of the rat and human enzymes in these respects. Differences are also noted in the two short anchor sequences (12 and 7 residues, respectively) that flank the acidic stretch where nonconservative substitutions (Thr-Ala at position 140 and Ala-Val at position 199 in the upstream and downstream anchor motifs, respectively) are found between the two species.

The sites of expression of the rat NRD convertase have been studied to a limited extent (Pierotti *et al.*, 1994). Interestingly, in human NRD convertase mRNA, although ubiquitously expressed at low levels, is particularly prominent in heart, skeletal muscle, pancreas, and testis; in the fetus it is additionally highly expressed in brain. In the mouse *in situ* hybridization at various developmental stages confirms the high-level expression of NRD convertase in the neural system and shows that the neural tissue is almost exclusively the site of expression in early development (between E 10.5 and E 16.5). The highly selective expression suggests that NRD convertase plays specific roles in development, as well as in some few sites of the adult, such as muscle and heart. Rat NRD convertase selectively cleaves at the N-terminal side of Arg from dibasic sites in a number of prohormone fragments, such as somatostatin and opioid peptides (Chesneau *et al.*, 1994; Pierotti *et al.*, 1994; Rholam *et al.*, 1986), some of which are known to be expressed in the testes, a site of high NRD expression. The possible roles played in skeletal and heart muscle and during nervous system development remain elusive. The requirement for metalloendopeptidase activity during the fusion of rat myoblasts as well as the terminal differentiation of mouse myoblasts has been reported (Couch and Strittmatter, 1983; Baldwin and Kayalar, 1986; Yagami-Hiromasa *et al.*, 1995). Additionally, Kuz, a metalloprotease belonging to the ADAM family, has recently been shown to be involved in mediation of signals for neural development in Drosophila (Rooke *et al.*, 1996). Our data suggest NRD convertase as a possible candidate for a role in these processes, which might be testable in functional experiments. Finally, the NRD convertase gene has now been mapped precisely to chromosome 1p32.2. This positional information, taken together with data on *in vivo* expression, might provide a candidate gene, should any inherited neuromuscular disease be mapped in this interval.

ACKNOWLEDGMENTS

This work was supported by a grant from Telethon (EO29) and AIRC (Associazione Italiana Ricerca Cancro) to R.T., grants from AIRC to S.O. and A.S., respectively, and from Progetto Finalizzato Ingegneria Genetica CNR to S.O. The authors thank Linda Cairns for valuable help and discussion.

REFERENCES

Affholter, J. A., Fried, V. A., and Roth, R. A. (1988). Human insulin-degrading enzyme shares structural and functional homologies with *E. coli* protease III. *Science* **242:** 1415–1418.

Baldwin, E., and Kayalar, C. (1986). Metalloendoprotease inhibitors that block fusion also prevent biochemical differentiation in L6 myoblasts. *Proc. Natl. Acad. Sci. USA* **83:** 8029–8033.

Banfi, S., Borsani, G., Rossi, E., Bernard, L., Guffanti, A., Rubboli, F., Marchitiello, A., Giglio, S., Coluccia, E., Zollo, M., Zuffardi, O., and Ballabio, A. (1997). Identification and mapping of human cDNAs homologous to Drosophla mutant gene through EST database searching. *Nature Genet.* **13:** 167–174.

Barrett, A. J. (1994). Classification of peptidases. *Methods Companion Methods Enzymol.* **244:** 1–13.

Barrett, A. J. (1979). "Proteinase in Mammalian Cells and Tissues," North Holland, Amsterdam.

Baumaister, H., Muller, D., Rehbein, M., and Richter, D. (1993). The rat insulin-degrading enzyme, molecular cloning and characterisation of tissue specific transcripts. *FEBS Lett.* **317:** 250–254.

Bond, J. S., and Butler, P. E. (1987). Intracellular proteases. *Annu. Rev. Biochem.* **56:** 333–364.

Chesneau, V., Pierotti, A. R., Barré, N., Creminon, C., Tougard, C., and Cohen, P. (1994). Isolation and characterisation of a dibasic selective metalloendopeptidase from rat testes that cleaves at the amino terminus of arginine residues. *J. Biol. Chem.* **269:** 2056–2061.

Chumakov, I., Rigault, P., Guillou, S., Ougen, P., and Cohen, D. (1992). Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* **359:** 380–387.

Couch, C., and Strittmatter, W. (1983). Rat myoblast fusion requires metalloendoprotease activity. *Cell* **32:** 257–265.

Finch, P. W., Wilson, R. E., Brown, K., Hickson, I. D., and Emmerson, P. T. (1986). Complete nucleotide sequence of the *E. coli* ptr gene encoding protease III. *Nucleic Acids Res.* **14:** 7695–7703.

Guimera, J., Casas, C., Pucharcos, C., Solan, A., Omenech, A., Planas, A., Ashley, J., Lovett, M., Estivill, X., and Pritchard, M. (1996). A human homologue of Drosophila minibrain is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. *Hum. Mol. Genet.* **5:** 1305–1310.

Kleiber, J., Kalousek, F., Swaroop, M., and Rosenberg, L. (1990). The general mitochondrial matrix processing protease from rat liver: structural characterisation of the catalytic subunit. *Proc. Natl. Acad. Sci. USA* **87:** 7978–7982.

Meulenberg, J. J., Sellink, E., Riegman, N., and Postma, P. (1992). Nucleotide sequence and structure of the Klebsiella pneumoniae pqq operon. *Mol. Gen. Genet.* **232:** 284–294.

Morgan, J. G., Dolganof, G. M., Robbins, S. E., Hinton, L. M., and Lovett, M. (1992). The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. *Nucleic Acids Res.* **20:** 5173–5179.

Pfanner, N., and Neupert, W. (1987). Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. *Annu. Rev. Biochem.* **56:** 829–852.

Pierotti, R., Prat, A., Chesneau, V., Gaudoux, F., Leseney, A., Foulon, T., and Cohen, P. (1994). N-arginine dibasic convertase, a metalloendopeptidase as a prototype of a class of processing enzymes. *Proc. Natl. Acad. Sci. USA* **91:** 6078–6082.

Rawlings, N. D., and Barrett, A. J. (1995). Evolutionary families of metallopeptidases. *Methods Companion Methods Enzymol.* **248:** 183–228.

Rawlings, N. D., and Barrett, A. J. (1993). Homologues of insulinase, a new superfamily of metalloendopeptidases. *Biochem. J.* **290:** 205–218.

Rholam, M., Nicolas, P., and Cohen, P. (1986). Precursors for peptide hormones share common secondary structures forming features at the proteolytic processing sites. *FEBS Lett.* **207:** 1–6.

Rooke, J., Pan, D., Xu, T., and Rubin, G. (1996). KUZ, a conserved metalloprotease-disintegrin protein with two roles in Drosophila neurogenesis. *Science* **273:** 1227–1231.

Wilkinson, D. G., and Green, J. (1990). "Postimplantation Embryos: A Practical Approach," IRL Press, Oxford.

Yagami-Hiromasa, T., Sato, T., Kurisaki, T., Kamjio, K., Nabeshima, Y., and Fujisawa-Sehara, A. (1995). A metalloprotease-disintegrin participating in myoblast fusion. *Nature* **377:** 652–656.