MOLECULAR AND CELLULAR BIOLOGY, Mar. 1994, p. 1909–1919
0270-7306/94/$04.00+0

Vol. 14, No. 3

Structural and Functional Aspects of the Multiplicity of Neu Differentiation Factors

DUANZHI WEN,[1]* SIDNEY V. SUGGS,[1] DEVARAJAN KARUNAGARAN,[2] NAILI LIU,[1] ROD L. CUPPLES,[1] YI LUO,[1] ANN M. JANSSEN,[1] NOA BEN-BARUCH,[2] DAVID B. TROLLINGER,[1] VICTORIA L. JACOBSEN,[1] SHI-YUAN MENG,[1] HSIENG S. LU,[1] SYLVIA HU,[1] DAVID CHANG,[1] WEINING YANG,[1] DONNA YANIGAHARA,[1] RAYMOND A. KOSKI,[1]† AND YOSEF YARDEN[2]

Amgen, Inc., Thousand Oaks, California 91320,[1] and Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel[2]

Received 13 October 1993/Accepted 1 December 1993

We used molecular cloning and functional analyses to extend the family of Neu differentiation factors (NDFs) and to explore the biochemical activity of different NDF isoforms. Exhaustive cloning revealed the existence of six distinct fibroblastic pro-NDFs, whose basic transmembrane structure includes an immunoglobulin-like motif and an epidermal growth factor (EGF)-like domain. Structural variation is confined to three domains: the C-terminal portion of the EGF-like domain (isoforms α and β), the adjacent juxtamembrane stretch (isoforms 1 to 4), and the variable-length cytoplasmic domain (isoforms a, b, and c). Only certain combinations of the variable domains exist, and they display partial tissue specificity in their expression: pro-NDF-α2 is the predominant form in mesenchymal cells, whereas pro-NDF-β1 is the major neuronal isoform. Only the transmembrane isoforms were glycosylated and secreted as biologically active 44-kDa glycoproteins, implying that the transmembrane domain functions as an internal signal peptide. Extensive glycosylation precedes proteolytic cleavage of pro-NDF but has no effect on receptor binding. By contrast, the EGF-like domain fully retains receptor binding activity when expressed separately, but its β-type C terminus displays higher affinity than α-type NDFs. Likewise, structural heterogeneity of the cytoplasmic tails may determine isoform-specific rate of pro-NDF processing. Taken together, these results suggest that different NDF isoforms are generated by alternative splicing and perform distinct tissue-specific functions.

The *neu* proto-oncogene (also known as *HER-2* or c-*erbB-2*) encodes a 185-kDa receptor tyrosine kinase. The transmembrane glycoprotein is present in many epithelial and neural tissues (6, 13, 18, 22, 27, 28, 33, 34). Although adult tissues generally exhibit a lower level of p185^{neu} expression than corresponding fetal tissues, p185^{neu} expression levels are frequently elevated in certain human neoplasms. Overexpression of p185^{neu} occurs in approximately 20% of breast, stomach, pancreatic, bladder, and ovarian carcinomas (20) and is associated with poor prognosis for breast and ovarian cancers. Increased p185^{neu} expression also occurs in certain nonmalignant neoplasias, such as adenomatous polyps (5, 7), Barrett's esophagus (17), and polycystic kidneys (14).

Many growth factors, cytokines, and neurotrophic factors activate receptor tyrosine kinases. The activated receptors become phosphorylated on tyrosine residues, thereby initiating intracellular signaling, leading to cellular responses (4, 35). Phosphorylation of the p185^{neu} receptor on tyrosine residues can be stimulated by proteins purified from several sources (9, 15, 16, 21, 30). Purification of the rat and human p185^{neu} stimulatory proteins led to the isolation of cDNAs encoding novel epidermal growth factor (EGF)-related proteins (15, 30, 36). The 44-kDa rat factor, named Neu differentiation factor (NDF), stimulates p185^{neu} tyrosine phosphorylation and induces the production of milk components (casein and lipids) in certain breast carcinoma cell lines (30). The homologous human factors, termed heregulins (HRGs), were found to be mitogenic for other mammary tumor cells (15). The NDF cDNA sequences predicted a transmembrane glycoprotein precursor (pro-NDF) with an EGF-like domain, an immunoglobulin homology unit, and a variable-length cytoplasmic domain. Recombinant rat NDF was found to interact with p185^{neu} and to stimulate tyrosine phosphorylation of the receptor in human tumor cells of breast, colon, and neuronal origin (31). However, no interaction of NDF with Neu-overexpressing ovarian and fibroblastic cells could be detected, suggesting that a still-unknown membrane molecule functions as a coreceptor (31).

In situ hybridization experiments with a NDF probe identified the central and peripheral nervous systems as prominent sites of NDF expression in mouse embryos and suggested a neuronal function for this factor (29). Such functions have been recently demonstrated for two NDF-related molecules. These are a chicken brain-derived 33- to 42-kDa protein, termed ARIA, that induces elevated expression of the acetylcholine receptor at the neuromuscular junction (12), and a group of brain and pituitary proteins, termed glial growth factors (GGFs), that exert a mitogenic response on cultured Schwann cells (23). We have previously reported that multiple RNA species of NDF exist, they display a tissue-specific pattern of expression, and the levels of some transcripts are significantly elevated upon transformation with the *ras* oncogene (36). The structures of several distinct forms of HRGs and GGFs have been described (15, 23). In an effort to address the structural basis and biological role of this multiplicity, we present here the characterization of rat and human NDF cDNA clones and the pro-NDF isoforms encoded by them. This enabled us to confine the structural variations to three domains, namely, the C-terminal portion of the EGF-like domain, the adjacent juxtamembrane stretch, and the intracellular tails of the precursor molecules. Expression of the various

* Corresponding author. Mailing address: Amgen Center, 1840 Dehavilland Drive, Thousand Oaks, CA 91320-1789. Phone: (805) 499-5725, ext. 3864. Fax: (805) 499-7464.

† Present address: 7 Meetinghouse La., Old Lyme, CT 06371.

rat pro-NDF molecules or their human homologs in mammalian and bacterial cells permitted initial determination of the biochemical significance of the observed structural heterogeneity.

MATERIALS AND METHODS

Cell culture. COS-7, T47D, A-704, HT-1080, HS294T, 5637 (ATCC HTB9), and MDA-MB-453 cells were obtained from the American Type Culture Collection (Rockville, Md.). COS-7 and T47D cells were cultured in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS). MDA-MB-453 cells were grown in RPMI 1640 medium with 15% FBS. Rat-1-EJ cells, generated by transfection of the human EJ *ras* oncogene into Rat-1 fibroblasts, were previously described (19, 30).

Isolation and sequencing of rat NDF cDNA clones. Ten NDF cDNA clones were previously isolated from a cDNA library prepared from Rat-1-EJ cell poly(A)$^+$ mRNA (36). These clones were designated clones 4, 19, 20, 22, 38, 40, 41, 42A, 42B, and 44. The nucleotide sequences of both strands of the NDF coding region were determined for all ten cDNAs by using synthetic oligonucleotide primers with fluorescence-based dideoxy-DNA sequencing as previously described (36).

Isolation of human NDF cDNA clones. Double-stranded cDNA was synthesized from poly(A)$^+$ RNA extracted from A-704 (ATCC HTB45) cells (American Type Culture Collection). Column-fractionated cDNA was ligated to *Sal*I- and *Not*I-digested plasmid vector pJT-2 and used to transform *Escherichia coli* DH10B (GIBCO-BRL) or *E. coli* MC1061 (Bio-Rad, Hercules, Calif.) by electroporation, using a cDNA kit and a procedure recommended by the manufacturer (GIBCO-BRL). Approximately 700,000 primary transformants were screened with the ^{32}P-labeled rat clone 44 NDF cDNA probe. Hybridization was at medium stringency (30% formamide, 6× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 2× Denhardt's solution, 100 μg of salmon sperm DNA per ml, 1 mM EDTA, 0.2% sodium dodecyl sulfate (SDS), 0.1% sodium PP_i, 50 mM NaH_2PO_4 [pH 6.8]) at 42°C overnight. Filters were washed in a solution of 0.5× SSC, 2 mM EDTA, and 0.2% SDS at 23°C for 30 min and then at 42°C for 90 min. The filters were exposed overnight to X-ray film at −75°C. Seven positive clones were identified after several rounds of colony hybridization. Plasmid DNA was purified from each clone, and cDNA sequences were determined. To obtain additional variants of NDF, a second cDNA library was synthesized with poly(A)$^+$ RNA extracted from A-704 cells. Double-stranded cDNA was ligated to *Bst*XI linkers and ligated into the *Bst*XI site on plasmid vector pCDNA II (Invitrogen, San Diego, Calif.). The ligation mixture was used to transform *E. coli* DH10B. Approximately 260,000 primary transformants were screened with a ^{32}P-labeled human NDF-α2b cDNA probe (clone 43) that was isolated in the first screening. Hybridization was at high stringency (50% formamide, 6× SSC, 2× Denhardt's solution, 100 μg of salmon sperm DNA per ml, 1 mM EDTA, 0.2% SDS, 0.1% sodium PP_i, 50 mM NaH_2PO_4 [pH 6.8]) at 42°C overnight. The filters were washed in a solution of 0.1× SSC, 2 mM EDTA, and 0.2% SDS at 23°C for 30 min and then at 70°C for 60 min. The filters were then exposed to X-ray films at −75°C overnight. Six positive clones were identified after several rounds of colony hybridization. Plasmid DNA was purified from the clones, and the cDNA sequences were determined by the method previously described. Among the positive cDNA clones was a 920-bp cDNA encoding a partial human NDF-β3 (clone 33). In addition, two human pituitary cDNA libraries (Clontech) were similarly screened. Among the positive clones, clone P1 encoded human NDF-α1a and clone P13 encoded human NDF-β1a. Lastly, a PCR cloning strategy was used to amplify NDF-specific DNA sequences from tumor cell lines that express NDF mRNAs. Five human tumor cell lines were used: HT-1080, HS 294T, 5637 (ATCC HTB9), MDA-MB-231, and A-704. Poly(A)$^+$ RNAs were extracted from the five cell lines and reverse transcribed to generate first-strand cDNA templates by using a cDNA kit from GIBCO-BRL. NDF cDNA sequences (~1 kb) corresponding to the region encoding amino acids 119 to 410 in human NDF-α2b (clone 43) were amplified. This region contains the spacer domain, the variable EGF-like domain (α and β forms), the variable sequences between the EGF-like domain and the transmembrane domain, the conserved transmembrane domain, and a portion of the cytoplasmic tail. The sense primer was 5′-AGG AAA TGA CAG TGC CTC T-3′. The antisense primer was 5′-TCT CTG GCA TGC CTG AGG-3′. PCR was carried out for 40 cycles. The reaction conditions were 94°C, 1 min; 50°C, 2 min; and 72°C, 3 min. The PCR-amplified DNA fragments were subcloned into the pSPORT plasmid vector (GIBCO-BRL), and their DNA sequences were determined.

PCR analyses of rat pro-NDF mRNA. NDF cDNA sequences encoding the variable portion of the EGF-like domain, namely, pro-NDF-α2a codon positions 200 to 241, were amplified by using the synthetic oligonucleotide primers 441-17 (5′-GCGTCTAGATGAAGGACCTGTCAAACCC-3′) (sense) and 441-18 (5′-GCGGGATCCCTTCTGGTAGAGT TCCTCC-3′) (antisense). The underlined sequences add a *Xba*I or a *Bam*HI restriction site to the 5′ ends of the primers. DNA sequences encoding the variable region of the pro-NDF cytoplasmic domains (NDF-α2a and NDF-α2b forms) were amplified with primers 403-17 (5′-CCCCTGAGCCTACAG AG-3′) (sense) and 403-16 (5′-GGGGCGATTTAGGGG AG-3′) (antisense). RNA samples (1 μg) from normal adult female Sprague-Dawley rat tissues (Charles River Laboratory, Wilmington, Mass.) and from the Rat-1 and Rat-1-EJ cell lines (36) were reverse transcribed to generate first-strand cDNA with cDNA cycle kits (Invitrogen). PCRs employed Perkin Elmer-Cetus GeneAmp PCR kits. The reaction mixtures were in a final volume of 50 μl and contained 10% of the reverse transcription reaction products or approximately 100 ng of cloned cDNAs as positive controls. Twenty-five cycles in a Perkin Elmer-Cetus GeneAmp 9600 thermocycler amplified the PCR products. Each cycle included 20 s at 94°C, 20 s at 55°C, and 20 s at 72°C. Ten-microliter aliquots of each reaction mixture were analyzed by electrophoresis in 2% agarose gels and then by ethidium bromide staining. DNA molecular weight markers were from Boehringer Mannheim Corp. (Indianapolis, Ind.). PCR products from brain, spinal cord, muscle, heart, and other tissues were subcloned into the pSPORT vector (GIBCO-BRL) and transfected into *E. coli* cells by electroporation with a Gene-Pulser apparatus (Bio-Rad) according to the manufacturer's directions. Plasmid DNA from the transfected colonies was analyzed by DNA sequencing.

Bacterial expression of NDF molecules and generation of an antibody to NDF. Recombinant rat and human NDF molecules were produced in *E. coli* and purified to greater than 98% homogeneity by subjecting a clarified *E. coli* lysate to anion-exchange, cation-exchange, hydrophobic interaction, and hydroxyapatite column chromatography. Extra methionine residues were added at the N terminus of NDF as initiation sites for protein translation. To raise antibodies to recombinant rat NDF, rabbits were immunized with an initial injection of 200 μg of recombinant rat NDF-$\alpha 2_{14\text{-}241}$. The antigen was emulsified with complete Freund's adjuvant and injected subcutane-

ously at multiple sites. At four-week intervals, the rabbits were boosted with 200 μg of antigen in incomplete Freund's adjuvant injected at multiple subcutaneous sites. Sera were collected by ear vein bleed 10 days after each injection and evaluated for reactivity with rat NDF-$\alpha 2_{14-241}$ by enzyme-linked immunosorbent assay. Anti-NDF antibodies were purified by affinity chromatography with an Actigel ALD kit as recommended by the manufacturer (Sterogene Bioseparations, Inc., Arcadia, Calif.). The affinity gel was prepared by overnight coupling of 2 mg of recombinant rat NDF-$\alpha 2_{14-241}$ to 3 ml of monoaldehyde-agarose. Unreacted aldehyde groups were deactivated with 0.1 M ethanolamine. Five milliliters of antiserum was passed through a column (1 by 10 cm) containing the affinity gel. The column was washed with 50 ml of phosphate-buffered saline (PBS) and eluted with 3 mM $MgCl_2$, 80 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES; pH 6.0), and 25% ethylene glycol. The eluted antibodies were dialyzed overnight against PBS at 4°C and concentrated with a Centriprep unit (Amicon) to an ~1-mg/ml concentration.

Transient expression of NDF cDNA clones in COS-7 cells. Transient expression of pro-NDFs employed our previously described method (36). NDF cDNA clones inserted in the pJT-2 eukaryotic expression vector are under the control of the simian virus 40 early promoter and are 3′ flanked with the simian virus 40 termination and polyadenylation signals. COS-7 cells were transfected with the pJT-2/NDF cDNA plasmids by electroporation as follows. A total of 2×10^6 cells in 0.8 ml of DMEM and 10% FBS were transferred to a 0.4-cm cuvette and mixed with 10 μg of plasmid DNA in 10 μl of TE solution (10 mM Tris-HCl [pH 8.0], 1 mM EDTA). Electroporation was performed at room temperature at 1,600 V and 25 mF by using a Bio-Rad Gene Pulser with the pulse controller unit set at 200 Ω. The cells were then diluted into 5 ml of DMEM–10% FBS and transferred into a 60-mm-diameter dish. After 14 h of incubation at 37°C, the medium was replaced with DMEM–1% FBS. Conditioned media were harvested 48 h later.

Purification of recombinant mammalian NDF produced in CHO cells. Recombinant rat NDF-α2 expressed by CHO cells transfected with pDSRα2/NDF-α2 was purified by the following procedure. Pooled serum-free conditioned media, from harvests of roller bottles containing the transfected CHO cells expressing recombinant rat NDF-α2, were cleared by filtration through 0.2-μm-pore-size filters and concentrated with a Pellicon diafiltration system with a 10-kDa molecular mass cutoff membrane. Concentrated material was directly loaded onto a column of heparin-Sepharose preequilibrated with 20 mM sodium phosphate buffer (pH 7.2) containing 25 mM NaCl. The column was then washed with the same buffer containing 0.25 M NaCl until A_{280} fell below 0.05. Bound proteins were eluted with a continuous gradient of NaCl from 0.25 to 1 M. Recombinant NDF in the collected fractions was detected with a NDF-specific antibody and by the Neu receptor phosphorylation assay. Active fractions were pooled and dialyzed against 23 mM sodium phosphate buffer containing 20 mM NaCl. Insoluble particulates present in the sample were cleared by centrifugation. The purification yield was about 70%. The dialyzed and clarified sample was then loaded onto a DEAE-Sepharose 6B fast flow column which had been preequilibrated with the above-described dialysis buffer. The column was extensively washed until no A_{280} could be detected. The column was then developed with a 200-ml gradient of 0.02 to 0.5 M NaCl in 20 mM sodium phosphate buffer (pH 7.2). Fractions were collected and assayed for NDF content. Ammonium sulfate was added to the pooled NDF fraction obtained from DEAE-Sepharose chromatography to achieve a concentration of 2 M. The material was loaded on a phenyl-Sepharose 4B column (Pharmacia) preequilibrated with 20 mM sodium phosphate (pH 7.2) containing 2 M ammonium sulfate. After loading, the column was washed and developed with a gradient of ammonium sulfate (from 2 M to no salt) in 20 mM sodium phosphate (pH 7.1). The main peak of activity was pooled, extensively dialyzed against PBS, and concentrated by centrifugation with a Centriprep 10 cartridge.

$p185^{neu}$ tyrosine phosphorylation assay. The media conditioned by transfected COS-7 cells were filtered through a 0.2-mm-pore-size sterile filter unit (Costar) and concentrated 16-fold with a Centriprep 10 unit (Amicon). The concentrated media were added to individual wells of a 48-well dish that contained 3×10^5 MDA-MB-453 human breast cancer cells per well. Following a 5-min incubation at 37°C, the samples were aspirated and the cells were processed for immunoblotting with an anti-phosphotyrosine monoclonal antibody (ICN). The protocol for cell lysis and enhanced chemiluminescence Western blotting (immunoblotting) was as previously described (30).

Labeling and immunoprecipitation of NDF proteins. COS-7 cells transfected with pro-NDF cDNA expression plasmids were grown in 60-mm-diameter dishes in DMEM containing 10% FBS at 37°C for 48 h. Cells were washed and placed in 1.5 ml of medium without methionine and cysteine with 1% dialyzed FBS. After 60 min, 200 μCi of [^{35}S]methionine and 200 μCi of [^{35}S]cysteine (Tran^{35}S-label; ICN) was added, and the cells were incubated for ~17 h. The media were collected, made 1 mM in phenylmethylsulfonyl fluoride, and clarified of cell debris in a microcentrifuge. The cells were harvested from culture dishes by scraping them into 1 ml of PBS. The cells were recovered by centrifugation and lysed by addition of 200 μl of 1% SDS–1 mM phenylmethylsulfonyl fluoride and heating at 100°C for 3 min. The cell lysates were clarified by centrifugation and diluted 1:4 in immunoprecipitation dilution buffer (1.25% Triton X-100, 190 mM NaCl, 60 mM Tris-HCl [pH 7.4], 6 mM EDTA, 10 U of Trasylol per liter). The radiolabeled COS-7 cell-conditioned media and the diluted cell lysates were pretreated at 4°C for 2 h with 20 μl of normal rabbit serum and 20 μl of protein A-Sepharose CL-4B (Sigma) diluted 1:1 in PBS. Protein A-Sepharose was then removed by centrifugation. The pretreated samples were gently agitated at 4°C overnight with affinity-purified anti-NDF antibody (20 μg/ml) and 20 μl of protein A-Sepharose. The protein A-Sepharose beads were pelleted and washed four times in a washing solution (0.1% Triton X-100, 0.02% SDS, 150 mM NaCl, 50 mM Tris-HCl [pH 7.5], 5 mM EDTA, 10 U of Trasylol per liter). The washed beads were heated with 60 μl of SDS-gel electrophoresis sample buffer (Novex) at 100°C for 3 min. The beads were removed by centrifugation, and the supernatants were analyzed by electrophoresis in SDS–10% polyacrylamide gels (Novex). Prestained protein molecular weight markers were from Novex.

Enzymatic deglycosylation. Immunoprecipitated NDF proteins (5 μl) in SDS gel sample buffer (Novex) were diluted into 20 μl of H_2O with 3-[(3-cholamidopropyl)-dimethyl-ammonio]-1-propanesulfonate (CHAPS) added to a 2% final concentration. The samples were incubated at 37°C for 2 h after 0.5 U of *N*-glycanase, *O*-glycanase, or neuraminidase (Genzyme) was added. The digested proteins were mixed with 5 μl of concentrated SDS gel sample buffer, heated at 100°C for 3 min, and analyzed by electrophoresis in a SDS–10% polyacrylamide gel (Novex). Prestained molecular weight markers were from Novex.

Ligand binding analysis. Monolayers of 1×10^5 to 2×10^5

TABLE 1. Rat-1-EJ pro-NDF cDNA clones

cDNA clone	Size (kb)	Pro-NDF isoform	Bioassay results[a]
4[b]	1.1	β3	−
19	2.9	α2b	+
20	2.5	α2b	+
22	3.4	β2a	+
38	2.4	α2a, truncated	+
40[c]	1.3	β2	+
41	2.7	α2a	+
42A	2.4	β4a	+
42B	3.2	α2a	+
44	1.9	α2c	+

[a] COS-7 cells were transfected with plasmid vectors directing expression of pro-NDF cDNAs. The conditioned cultured media were assayed for stimulation of p185^{neu} tyrosine phosphorylation.

[b] Truncated at 5′ end, starts at codon 11. The 5′ end was replaced with the 5′ end of clone 44 for expression studies.

[c] Truncated at codon 300 in the intracellular domain.

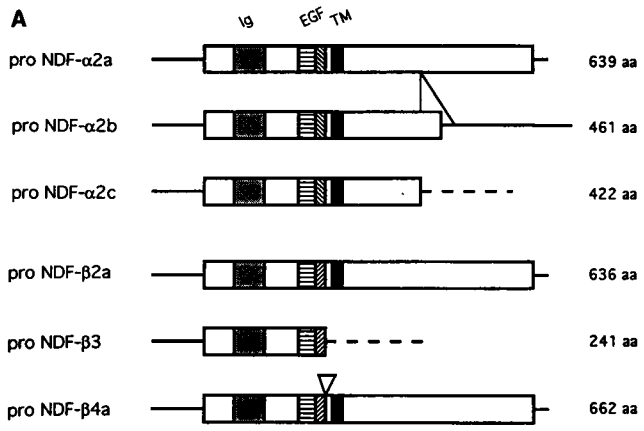
cells per well in 24-well dishes were washed once with binding buffer (DMEM containing 20 mM HEPES and 1% bovine serum albumin) and then incubated with 5 ng of ^{125}I-NDF per ml in the same buffer. Different concentrations of unlabeled NDF were coincubated with the radiolabeled ligand. Nonspecific binding was determined by the addition of a 100-fold excess of unlabeled NDF together with ^{125}I-NDF in the same experiment. After incubation for 2 h at 4°C, the cells were washed three times with ice-cold binding buffer. Labeled cells were lysed in 0.5 ml of 0.1 N NaOH–0.1% SDS for 15 min at 37°C, and the radioactivity was determined by using a gamma counter.

Nucleotide sequence accession number. cDNA clones 4, 19, 20, 22, 38, 40, 41, 42A, 42B, and 44 were submitted to GenBank under accession numbers U02315 to U02324, respectively. The accession numbers of the human NDF cDNA sequences are U02325 to U02330.

RESULTS

Structural heterogeneity of pro-NDF molecules. We previously isolated 10 independent clones from a cDNA library prepared from *ras*-transformed Rat-1-EJ cells (36). The cDNAs ranged in size from 1.1 to 3.2 kb (Table 1) and presumably are derived from the 6.8-, 2.6-, and 1.7-kb NDF mRNAs found in Rat-1-EJ cells (36). All 10 clones were sequenced, revealing that they encode six distinct NDF precursor proteins (pro-NDFs) (Table 1). The open reading frames in the 10 cDNAs ranged from 0.7 kb (clone 4) to 2.0 kb (clone 42A). Two of the cDNAs were partial-length clones. The sequence of clone 4 begins at codon 11. Clone 40 is truncated at codon 300 in the intracellular portion, so that it encodes only the first 36 amino acid residues of the cytoplasmic domain.

The 10 rat cDNAs encode six homologous NDF precursor proteins comprising 241 to 662 amino acid residues (Fig. 1 and Table 1). The six predicted pro-NDFs are identical in the first 213 amino acid residues (Fig. 1B). This identical region includes a basic N terminus which is proteolytically processed at amino acid 14 in naturally occurring rat NDF from the Rat-1-EJ cell line (30, 36), an immunoglobulin-like domain, a



B

```
proNDF-α2a  MSERKEGRGKGKGKKKDRGSRGKPGPAEGDPSPALPPRLKEMKSQESAAGSKLVLRCETSSEYSSLRFKWFKNGNELNRKNKPENIKIQKKPGKSELRIN  100

proNDF-α2a  KASLADSGEYMCKVISKLGNDSASANITIVESNEFITGMPASTETAYVSSESPIRISVSTEGANTSSSTSTSTTGTSHLIKCAEKEKTFCVNGGECFTVK  200

proNDF-α2a  DLSNPSRYLCKCQPGFTGARCTENVPMKVQTQEK----------AEELYQKRVLTITGICIALLVVGIMCVVAYCKTKKQRQKLHDR  277
proNDF-β1                PNEFTGDRCQNYVMASFYKHLGIEFME----------
proNDF-β2a               PNEFTGDRCQNYVMASFYK----------
proNDF-β3                PNEFTGDRCQNYVMASFYSTSTPFLSLPE*                                                    241
proNDF-β4a               PNEFTGDRCQNYVMASFYMTSRRKRQETEKPLERKLDHSLVKESK

proNDF-α2a  LRQSLRSERSNLVNIANGPHHPNPPPENVQLVNQYVSKNVISSEHIVEREVETSFSTSHYTSTAHHSTTVTQTPSHSWSNGHTESVISESNSVIMMSSVE  377

proNDF-α2a  NSRHSSPAGGPRGRLHGLGGPRDNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSPKSPPSEMSPPVSSMTVSMPSVAVSPFVEEERPL  477
proNDF-α2b                                          HNLIAELRRNKAYRSKCMQIQLSATHLRPSSITHLGFIL*                  461
proNDF-α2c                                          *                                                         422

proNDF-α2a  LLVTPPRLREKKYDHHPQQLNSFHHNPAHQSTSLPPSPLRIVEDEEYETTQEYESVQEPVKKVTNSRRAKRTKPNGHIANRLEMDSNTSSVSSNSESETE  577

proNDF-α2a  DERVGEDTPFLGIQNPLAASLEVAPAFRLAESRTNPAGRFSTQEELQARLSSVIANQDPIAV*                                        639
```

FIG. 1. (A) Rat pro-NDF structures predicted from the rat cDNA sequences. Boxed areas indicate protein coding regions. Ig, EGF, and TM, immunoglobulin-like, EGF-like, and transmembrane domains, respectively. The number of amino acid (aa) residues in each predicted precursor sequence is given in the right-hand column. Dashed lines represent divergent 3′ untranslated DNA sequences. (B) Amino acid sequences encoded by rat NDF cDNAs. The complete pro-NDF-α2a amino acid sequence is shown. Divergent sequences in pro-NDF variants are aligned with the pro-NDF-α2a sequence. Pro-NDF-β2a, -β3 and -β4a structures were deduced from Rat-1-EJ cDNAs (Table 1). The NDF-β1 sequence was obtained from a PCR-amplified cDNA derived from rat brain and spinal cord tissues. Asterisks, COOH-terminal amino acids; dashes, gaps introduced to facilitate sequence alignment; underlining, putative transmembrane domain; dots, cysteines in predicted extracellular domain (the two N-terminal cysteines are part of an immunoglobulin-like domain; the six other marked cysteines are part of the EGF-like domain).

TABLE 2. Frequency of NDF isoforms in cDNA amplified from human tumor cells[a]

Cell line	Frequency of following NDF isoform:			
	α2	β1	β2	β3
A-704	19	1	3	
MDA-MB231	5			
Hs294T	16	2	2	
HT-1080	2			1
5637	1			

[a] Poly(A)$^+$ RNAs were extracted from the indicated five human tumor cell lines and reverse transcribed to generate first-strand cDNA templates. PCR was used to amplify a 1-kb DNA segment that includes the coding region of the EGF-like domain. The amplified DNA fragments were subcloned into a plasmid vector, and their nucleotide sequences were determined. This enabled identification of the encoded isoforms and determination of their relative expression levels.

spacer domain with sites for N- and O-linked glycosylations, and two disulfide loops of the EGF-like domain (36). Variable sequences C terminally flank this constant portion of NDF. Structural variation is confined to three domains and defines the identity of each precursor protein. Thus, two variant C-terminal portions of the EGF-like domain characterize the α and β isoforms. An adjacent juxtamembrane processing stretch displays structural heterogeneity that is denoted by a number (1 to 4 isoforms). Lastly, the length of the cytoplasmic tail, which lies downstream to an invariant segment that includes the constant transmembrane domain and 157 amino acids, determines the identity of the isoform as a, b, or c. This nomenclature is an extension of the terminology that was suggested for human HRGs (15).

In order to isolate human homologs of the rat isoforms of NDF, we used rat cDNA clone 44 as a probe to analyze Northern (RNA) blots. This analysis included human kidney adenocarcinoma cell line A-704, MDA-MB231 human breast cancer cells, Hs294T human melanoma cells, 5637 human bladder carcinoma cells, and HT1080 human fibrosarcoma cells. Whereas NDF was found to be expressed in all these cell lines, the kidney adenocarcinoma cells exhibited the highest level of mRNA (data not shown). Therefore, we used this cell line as a source of mRNA and constructed a cDNA library that was screened with rat cDNA clone 44 as a probe. Among the seven positive cDNA clones were one cDNA encoding full-length NDF-α2b (clone 43) and two partial clones that encode NDF-α2b (clone 17) and NDF-α3 (clone 19). Rescreening of the cDNA library with clone 43 as a probe yielded a cDNA clone (920-bp insert) encoding a partial human NDF-β3 (clone 33). In addition, two human pituitary cDNA libraries were similarly screened and yielded NDF-α1a (clone P1) and NDF-β1a (clone P13). To identify additional independent human NDF cDNA clones, we employed a PCR cloning strategy that amplifies NDF-specific sequences from the above-mentioned five tumor cell lines. NDF cDNA sequences (~1 kb) corresponding to the region encoding amino acids 119 to 410 in human NDF-α were amplified (see Materials and Methods). This region includes the variable portion of the EGF-like domain and the adjacent juxtamembrane stretch. Nucleotide sequencing of the amplified DNA segments enabled identification of the corresponding isoforms of NDF. Table 2 summarizes the results of this analysis. Evidently, in addition to the previously identified NDF-α2, -β1, and -β3 sequences, DNA encoding an additional isoform (NDF-β2) was identified. Although only one α form was identified, namely, NDF-α2, this variant is apparently the most abundant isoform in the examined tumors.

The variable region of the pro-NDF ectodomain begins in the third disulfide loop of the EGF-like domain at amino acid position 213. The α/β variation alters the sequence between the fifth and the sixth cysteines of this domain. Pro-NDF-β proteins share a common sequence at positions 213 to 230 that is distinct from the corresponding NDF-α sequence (Fig. 1B). Variation between members of the pro-NDF-β family occurs carboxy terminally to the EGF domain, where the β1 and β4 proteins contain additional amino acid residues. Pro-NDF-β3 cDNA encodes a stop codon in this region, resulting in a smaller protein that lacks the transmembrane and cytoplasmic domains. Human cDNAs encoding HRGs α, β1, β2, and β3 were previously reported (15). The pro-NDF amino acid sequences are greater than 90% identical to their human HRG counterparts. A human cDNA corresponding to pro-NDF-β4 has not been reported.

The most variable part of pro-NDF is the stalk that connects the EGF-like domain with the transmembrane stretch. This domain presumably undergoes proteolysis while pro-NDF is released from the cell. Not only does this stretch display completely different amino acid sequences, but also its length is variable: the shortest block is present in pro-NDF-β2a, whereas the longest (27 amino acids) is found in NDF-β4a. A tetrad of four basic amino acids (Arg-Arg-Lys-Arg) is present in the latter sequence, suggesting that it functions as a proteolysis site.

Cytoplasmic domains with 374, 196, and 157 amino acid residues distinguish pro-NDFs α2a, α2b, and α2c, respectively. These isoforms are identical in the first 422 amino acid residues. The pro-NDF-α2a cytoplasmic domain extends for an additional 217 amino acid residues. This extended carboxy terminus is found in pro-NDFs β2a and β4a (Fig. 1); in pro-HRGs α, β1, and β2 (15); and in chicken pro-ARIA (12). Pro-NDF-α2b cDNA encodes a different 39-amino-acid-long sequence following position 422 that results from a 146-bp insert at codon 422 in the pro-NDF-α2b cDNA. This insert adds 39 amino acid residues and then a stop codon. Pro-NDF-α2c has the shortest cytoplasmic domain. Its cDNA has a stop codon at position 423 and then a different 3′-untranslated sequence. Pro-HRG-β2 variant clone 84 has a similar structure (15). Interestingly, in contrast with NDF-α isoforms, no βb and βc isoforms were identified in our screening.

PCR analyses of rat pro-NDF mRNAs. Eleven rat tissues were analyzed for tissue-specific expression of mRNAs encoding variant pro-NDF proteins. RNA samples were reverse transcribed and analyzed for divergent pro-NDF sequences by PCR. Primers that flank the variable EGF-like domain sequences (pro-NDF-α2a codons 200 to 241 [Fig. 1B]) were used to amplify the products shown in Fig. 2. Most of the 11 tissues yielded PCR products that comigrated with products obtained from cloned pro-NDF-α2a and pro-NDF-β2 cDNAs (lanes 13 and 14, respectively). Their identities were confirmed by DNA sequencing. The slightly larger PCR products (~160 bp) from both brain and spinal cord tissues (lanes 11 and 12, respectively) were subcloned into a plasmid vector and sequenced. This distinct brain and spinal cord PCR product encoded a rat NDF variant that corresponds to human HRG-β1 (15). Specifically, it is identical to NDF-β2 sequence but contains eight additional amino acids inserted at codon position 231 (Fig. 1B). PCR products corresponding to pro-NDF-β4 are not evident in Fig. 2, but they were detectable by Southern blotting (data not shown). Heart (lane 2) and muscle (lane 9) tissues did not yield significant amounts of the expected products. Sequencing data showed that the larger PCR products in lanes 2 and 9 were not related to NDF.

Additional PCR experiments using other primers detected

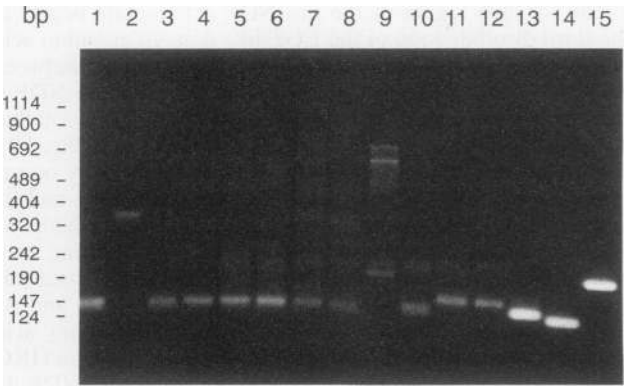


FIG. 2. PCR analysis of NDF mRNAs expressed in rat tissues. Reverse transcription was used to prepare first-strand cDNA from rat tissue mRNA samples. NDF cDNAs were amplified in PCRs with primers 441-17 and 441-18, which flank the variable EGF-like domains (pro-NDF-α2a codons 200 to 241 [Fig. 1B]). Lanes: 1, Rat-1-EJ cells; 2, heart; 3, skin; 4, ovary; 5, lung; 6, stomach; 7, spleen; 8, liver; 9, muscle; 10, kidney; 11, brain; 12, spinal cord; 13, cDNA clone 20 (pro-NDF-α2a); 14, cDNA clone 40 (pro-NDF-β2); 15, cDNA clone 42A (pro-NDF-β4a). Positions of DNA size markers are indicated.

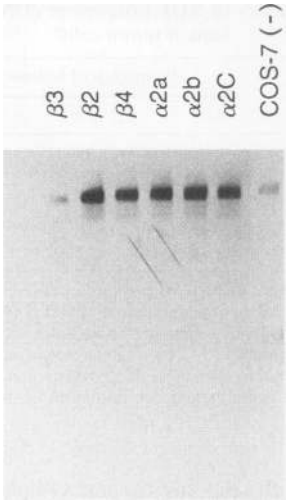


FIG. 3. Stimulation of p185^{neu} tyrosine phosphorylation by recombinant NDF proteins. Monolayers of MDA-MB-453 cells were incubated for 5 min at 37°C with media conditioned by untransfected COS-7 cells (labeled −) or transfected COS-7 cells. These cells were transfected with expression plasmids containing the indicated cDNAs clones. Cell lysates were prepared from the stimulated MDA-MB-453 cells and subjected to antiphosphotyrosine Western blot analysis. The major tyrosine-phosphorylated band detected in the Western blot corresponds to p185^{neu} (30).

NDF cytoplasmic domain sequences in mRNAs prepared from the 11 tissues. Most of the tissues yielded both products, namely, α2a and α2b forms, that comigrated with products obtained from Rat-1-EJ cells (data not shown). The PCR primers would not detect pro-NDF-α2c mRNA, which encodes the shortest cytoplasmic domain. The PCR products were undetectable in heart and muscle tissue.

Functional analyses of recombinant NDF proteins. We used transient COS-7 cell expression of pro-NDF cDNAs to determine whether all 10 rat cDNA clones encode biologically active NDF proteins. The cDNAs were transfected into COS-7 cells under the control of the simian virus 40 promoter. Media conditioned by the transfected cells were assayed for their ability to stimulate tyrosine phosphorylation of p185^{neu}. The conditioned media were incubated for 5 min with cultured MDA-MB-453 breast cancer cells, which express high levels of p185^{neu}. Cell lysates were prepared and analyzed by Western blotting for p185^{neu} tyrosine phosphorylation (Fig. 3 and data not shown). Nine of the 10 cDNA clones directed the synthesis and secretion of biologically active recombinant NDF molecules into the COS-7 cell-conditioned media (Table 1). Clone 4, which encodes pro-NDF-β3, did not produce a biologically active conditioned medium.

In order to quantitatively compare receptor binding characteristics of the different isoforms of NDF, we expressed certain rat or human isoforms, or their truncation mutants, in either bacterial cells or in Chinese hamster ovary (CHO) cells. The purified proteins were analyzed for their binding to T47D human mammary carcinoma cells. Ligand displacement analysis was performed with radiolabeled human NDF-$\beta1_{177\text{-}246}$. The latter differs from the corresponding rat protein in only one amino acid. When expressed in and purified from bacterial cells, this isolated EGF-like domain displayed an apparent dissociation constant of 0.2×10^{-9} to 0.4×10^{-9} M (Fig. 4A). This value is in agreement with those from a previous report (15), but it is 10-fold lower than the dissociation constant of native NDF-α2 that was determined with SKBR-3 cells (31). Comparative ligand displacement analysis confirmed that the α2 and β1 isoforms differ in their apparent affinities by a factor of 8 to 10 (Fig. 4A). The observed difference is not due to inactivation of the recombinant α2 isoform, since both a truncated CHO cell-derived rat NDF-$\alpha2_{5\text{-}228}$ and a bacterially made rat NDF-$\alpha2_{14\text{-}241}$ displayed similar low affinities. Interestingly, the bacterially made α2 isoform contains no sugar residues, but it nevertheless displayed binding affinity comparable to that of the fully glycosylated truncated protein (Fig. 4A). It was, therefore, concluded that the sugar moieties of NDF, as well as its whole N-terminal half (residues 1 to 176), including the immunoglobulin-like domain, do not significantly affect receptor binding. This was supported by expressing an N-terminal half of NDF that lacks the third disulfide loop of the EGF-like domain. This recombinant protein, $NDF_{14\text{-}211}$, displayed no specific cellular binding when its radiolabeled form was analyzed (data not shown). In addition, it could not displace cell-bound NDF-$\beta1_{177\text{-}246}$ (Fig. 4B). The relatively low binding affinity that was displayed by the rat α isoform was evident also by analyzing the corresponding human protein, namely, human NDF-$\alpha2_{14\text{-}241}$ (Fig. 4B). This result excludes the possibility that the observed differences (Fig. 4A) are species specific. However, the observation that NDF-α2 and NDF-β1 differ in their binding affinity could be due either to the different C-terminal portions of the EGF-like domains or to the adjacent variant juxtamembrane sequences (Fig. 1B). In order to address this question, we used the truncated version of the α2 isoform, which lacks the variant juxtamembrane sequence. This recombinant protein, rat NDF-$\alpha_{5\text{-}228}$, displayed ligand displacement activity that was very similar to that of NDF-$\alpha2_{14\text{-}241}$ (Fig. 4A), indicating that the C-terminal portion of the EGF-like domain, rather than the juxtamembrane stretch, affects receptor interactions.

Processing of pro-NDF isoforms. Synthesis of NDF proteins in transfected COS-7 cells was analyzed by immunoprecipitation of ^{35}S-labeled proteins. For these experiments, antibodies were raised by immunizing rabbits with a recombinant rat NDF-$\alpha2_{14\text{-}241}$. The NDF antigen was immobilized on a column and used to affinity purify the anti-NDF antibodies. The

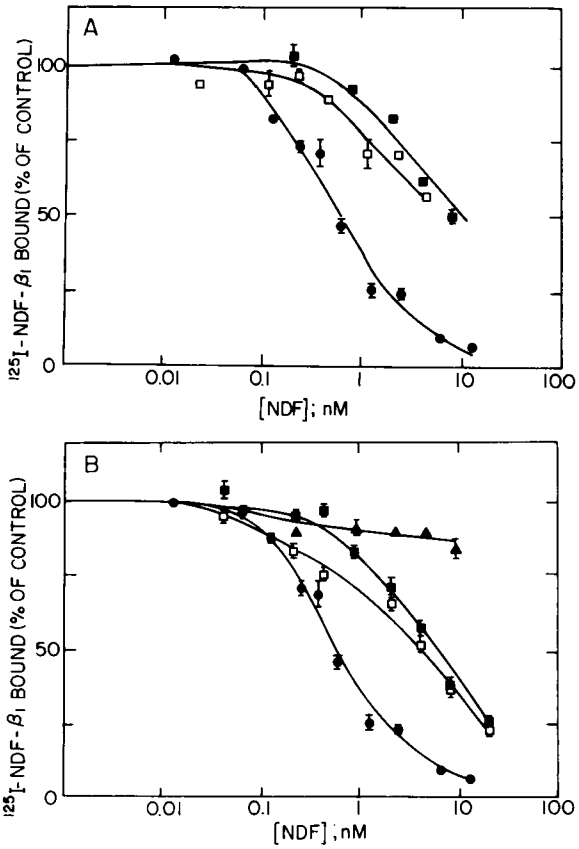


FIG. 4. Binding of NDF isoforms to T47D human mammary carcinoma cells. The ability of various NDF proteins to displace radiolabeled NDF-$\beta1_{177\text{-}246}$ was analyzed on monolayers of T47D human breast cancer cells. Binding reactions were carried out with 5 ng of ^{125}I-NDF $\beta1_{177\text{-}246}$ per ml for 2 h at 4°C. This was followed by extensive washing of the cell monolayers and determination of bound radioactivity. The amount of bound NDF-$\beta1_{177\text{-}246}$ is expressed relative to ligand binding in the absence of competitor unlabeled protein. (A) Rat NDF proteins. The following proteins were used as unlabeled competitors: bacterially made rat NDF-$\alpha2_{14\text{-}241}$ (filled squares), CHO cell-derived rat NDF-$\alpha_{5\text{-}228}$ (open squares), and human NDF-$\beta1_{177\text{-}246}$ (closed circles). The latter protein differs in only one amino acid residue from the corresponding sequence of rat NDF-β1. (B) Human NDF isoforms. The following unlabeled proteins were used to displace cell-bound human NDF-$\beta1_{177\text{-}246}$: human NDF-$\beta1_{177\text{-}246}$ (closed circles), a bacterially expressed human $NDF_{14\text{-}211}$ (triangles), and two completely independent preparations of human NDF-$\alpha2_{14\text{-}241}$ that were derived from bacterial cells (open squares and filled squares). Averages of duplicate determinations and the corresponding ranges (bars) are shown. Each experiment was repeated three times.

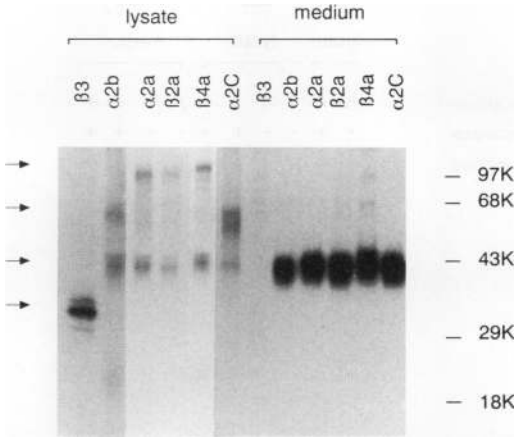


FIG. 5. Transient expression of recombinant pro-NDF cDNAs in COS-7 cells. COS-7 cells were transfected with pro-NDF cDNA expression plasmids and labeled for 17 h with [^{35}S]methionine and [^{35}S]cysteine. Radiolabeled COS-7 cell media and lysates were immunoprecipitated with affinity-purified rabbit antibody raised against recombinant rat NDF-$\alpha2_{14\text{-}241}$. The immunoprecipitates were analyzed by electrophoresis. NDF proteins are indicated by arrows on the left. The positions (in kilodaltons) of the molecular mass markers are shown on the right.

purified antibody immunoprecipitated ^{35}S-labeled NDF proteins from COS-7 cells expressing the different pro-NDFs (Fig. 5). NDF proteins were undetectable in cell lysates and in conditioned media from COS-7 cells transfected with irrelevant cDNAs and also when samples were immunoprecipitated in the presence of excess unlabeled recombinant NDF (data not shown). The affinity-purified antibody immunoprecipitated 40- to 44-kDa proteins from COS-7 cells transfected with cDNAs encoding pro-NDF proteins α2a, α2b, α2c, β2a, and β4a (Fig. 5). The immunoprecipitated proteins correspond in size to the processed form of NDF purified from media conditioned by Rat-1-EJ cells (30). The 40- to 44-kDa recombinant NDF proteins were found both in cell lysates and in conditioned media. Larger NDF precursor proteins were detected as high-molecular-weight proteins immunoprecipitable from the cell lysates but not from conditioned media. As predicted from pro-NDF cDNA structures (Fig. 1), the precursors of NDF-α2a, -β2a, and -β4a are significantly larger than the NDF-α2b and -α2c precursor proteins (Fig. 5). On the basis of molecular weight estimation, these larger proteins are most likely unprocessed, glycosylated NDF precursor proteins.

Importantly, the ratio between the various precursor forms (p60 and p95) and the fully processed form of NDF (p44) displayed variation in cell lysates. This was especially prominent in the case of pro-NDF-α2c, in which the precursor is the major cellular form. Nevertheless, the medium of pro-NDF-α2c-expressing cells contained the processed p44 molecule, implying that NDF release takes place despite a lower rate of cellular processing.

In contrast to other cDNA clones, the pro-NDF-β3 clone directed the synthesis of a major 30-kDa cellular protein that appeared as a sharp protein band on the gel (Fig. 5). No product was detectable in the medium of transfected COS-7 cells, in agreement with the absence of NDF bioactivity (Fig. 3). The diffuse nature and the size of the other immunoprecipitated NDF proteins suggested that they are highly glycosylated, unlike NDF-β3, which appeared to be unglycosylated. A large percentage of the mass of naturally occurring rat NDF is due to N- and O-linked glycosylation (30). Enzymatic deglycosylation of the 40- to 44-kDa and 60- to 75-kDa NDF-α2c proteins reduced their size by approximately 15 kDa but did not affect pro-NDF-β3 (Fig. 6). Comparison of the mobility shift that was induced by *N*-glycanase treatment, with the combined effect of *O*-glycanase and neuraminidase, suggested that O-linked sugars are more abundant than N-linked carbohydrates (Fig. 6). In addition, the susceptibility of pro-NDF-α2c to glycanase treatment implies that glycosylation, and perhaps also transmembrane localization of the precursor protein, precedes its proteolytic cleavage and release from the cell. Consistent with this possibility, pro-NDF-β3, which has no transmembrane domain, was resistant to glycanase treatment.

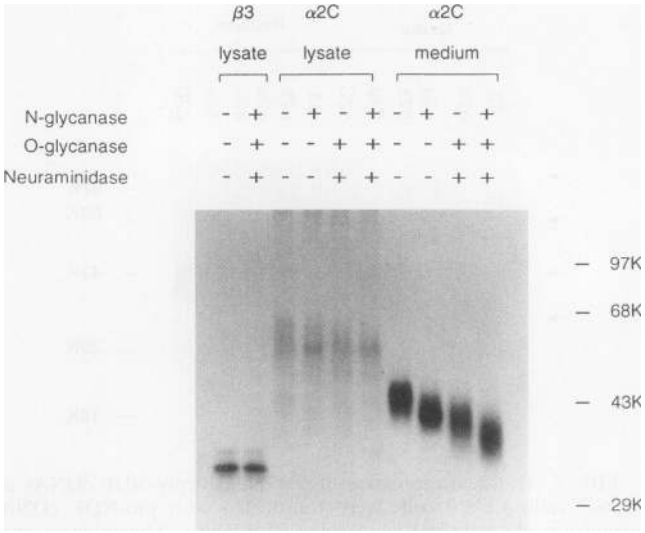


FIG. 6. Endoglycosidase treatment of NDF proteins synthesized by transfected COS-7 cells. [^{35}S]methionine- and [^{35}S]cysteine-labeled recombinant NDF-α2c and NDF-β3 were immunoprecipitated from COS-7 cell media or lysates as described in the legend to Fig. 5. The samples were incubated with the indicated enzymes and then separated by gel electrophoresis. The positions (in kilodaltons) of molecular mass marker proteins are indicated on the right.

This result and the exclusively intracellular localization of NDF-β3 (Fig. 5) indicate that this shortest NDF isoform does not enter the secretory pathway and imply that the hydrophobic transmembrane domain that is present in all the secreted isoforms of NDF is essential for glycosylation and secretion.

DISCUSSION

Members of the EGF family include EGF, transforming growth factor α (TGF-α), heparin-binding EGF, amphiregulin (AR), and viral growth factors (24). These polypeptide factors are synthesized as larger precursors with N-terminal signal peptide sequences, a six-cysteine EGF-like domain, a hydrophobic transmembrane domain, and a cytoplasmic domain of unknown function. Neu differentiation factors belong to this family and share similar molecular architecture and membrane topology. However, unlike other members of the EGF family, which are present in single forms, multiple variants of the basic structure of NDF exist. Thus, the NDF family includes, in addition to the seven rat NDF forms that were described in this work (α2a, α2b, α2c, β1, β2a, β3, and β4a), seven versions of the GGFs that were characterized as cDNA clones (23) and a single form of ARIA (12). Taken together, it appears that the minimal number of NDF isoforms is 12. As only one NDF gene was detected in the human genome (29), it is very likely that the multiple isoforms are generated by alternative splicing from a single NDF gene. This multiplicity of isoforms is with no precedence in other families of growth-regulatory secreted proteins and raises questions that relate to expression patterns, the structural basis, and the functional role of many related factors.

Expression patterns of NDF isoforms. In contrast with EGF, whose expression is limited to submaxillary gland, digestive tract, and kidney, NDFs display wide distribution in many tissues and organs (29). Moreover, exhaustive cloning of NDF cDNAs from fibroblasts and epithelial cells and comparison with GGF (23) and ARIA (12) sequences indicate that the expression patterns of some isoforms are tissue specific. For example, the α2 isoform is the predominant form in mesenchymal cells and in nonneuronal organs (Table 2 and Fig. 2). The second most enriched isoform in these tissues, NDF-β2, also lacks the juxtamembrane stretch. By contrast, the β1 isoform is enriched in brain tissue and spinal cord (Fig. 2), but it is a minor form in mesenchymal tissues. A related isoform, designated ARIA-1, was identified in embryonic chick brain tissue (12). Neuronal and nonneuronal tissues and organs differ also in splicing of additional putative exons. Thus, a kringle domain-containing isoform has not been isolated from tissues other than brain (23), whereas spacer domain-containing NDFs were identified only in mesenchymal cells, and they are absent in the neuronal isoforms, namely, GGFs and ARIA. This observation suggests that the sugar residues and the glycosaminoglycan attachment site of the spacer domain are involved in a mesenchyme-specific function. Taken together, the differential proportion of the α2 and β1 isoforms and the mutually exclusive presence of the kringle and spacer domains suggest that tissue-specific, alternatively spliced forms evolved in order to perform cell-type-specific functions. Some of these functions are reflected in vitro, and they are discussed below.

Structural aspects of NDF multiplicity. The alternative NDF sequences beginning at codon 213 could result from differential splicing of at least four different exons. Extensive splicing is common in mRNAs encoding the EGF family of growth factors. The human EGF gene has 24 exons (1), whereas human TGF-α and AR precursors are encoded by genes with 6 exons (8, 32). Alignment of the AR, EGF, and TGF-α sequences identified an identically placed intron disrupting the coding sequences between the second and the third disulfide loops of the EGF-like domains (32). The α/β divergency of NDF starts near this position. Differential exon splicing can also account for divergent NDF 3′ coding and noncoding sequences: conceivably, the cytoplasmic domain of NDF is encoded by three exons that when present together encode NDF-α2a and NDF-β2a, but the combination of only two exons yields isoform 2b (Fig. 1A). The amino-terminal half of the cytoplasmic domain and the flanking transmembrane domain appear to be encoded by a single exon that is shared by all the transmembrane isoforms. This predicted multiplicity of exons is extended by at least six exons, whose existence can be inferred from the reported structures of NDF-related cDNA clones. These include exons encoding the kringle domain, together with a signal peptide, the immunoglobulin-like motif, and the glycosylated spacer domain. We predict that the minimal number of exons in the NDF gene is 13, which is in agreement with an estimation that was made on the basis of the structure of the coding regions of GGF cDNA clones (23). Nevertheless, detailed genomic mapping and nucleotide sequencing will be needed in order to confirm this prediction. Such analysis may reveal new exons whose expression may be limited to specific cell lineages or to certain sites in the nervous system. Another prediction that arises from the structural heterogeneity of NDF isoforms relates to the possibility that transcription of the gene is regulated by two promoters. Accordingly, one promoter is specific to mesenchymal and other nonneuronal cells and undergoes activation by certain retroviral oncogenes, including *ras* and *sis* (36), whereas the other promoter is presumably localized upstream to the kringle domain coding region and selectively controls neural expression of certain NDF isoforms.

Functional aspects of the multiplicity of NDF isoforms. Perhaps the most interesting question that is raised by the structural multiplicity of NDFs relates to the possibility that it fulfills certain functional requirements which are dictated by

the multisite expression of the family. A major hint as to possible functional roles of the structural variation is provided by the localization of this heterogeneity. Certainly, the most variable region in the NDF family comprises the third loop of the EGF-like domain, the C-terminal tail of this motif, and the short stretch that connects it with the transmembrane domain. A region of secondary heterogeneity includes the variable-length cytoplasmic tail of NDF. Lastly, the domain structure of the whole region that N terminally flanks the EGF-like domain displays tissue specificity. At present, only limited information on the function of each structural domain and its variants is available.

EGF domain variants. The α/β variation partially alters the sequence between the fifth and the sixth cysteine residues of the EGF-like domain but completely changes the sequence of the 10-amino-acid-long C terminus of this motif. Nevertheless, the overall configuration and spacing of cysteine residues is conserved in all forms of NDF. Our ligand displacement analyses indicate that these changes significantly affect the affinity of NDF to its receptor: the β isoforms bind to mammary cells with an affinity that is approximately 10-fold better than that of the α isoform (Fig. 4). This observation is consistent with mutational analyses of the binding of human EGF to its receptor (3, 10, 11, 25). Evidently, both the amino- and the carboxy-terminal domains of EGF participate in receptor interactions. Site-directed mutagenesis identified hydrophobic residues in the carboxy-terminal domain that reduce receptor binding, with the most significant decrease observed in mutations at the highly conserved Leu-47 (10, 25). In addition, the guanidinium group of the conserved Arg-41 has been shown to be of critical importance (11). Whereas the analogous Arg residue is found in NDF, the homologs of Leu-47 are a methionine in NDF-α and a proline in NDF-β. Presumably, the variation of the carboxy-terminal tail is responsible for the differences in affinities, but conservation of most of the EGF-like domain confers the same receptor specificity and similar basal affinity to all NDF isoforms. The observation that interaction between NDF and its putative receptor, p185^{neu}, requires the presence of a still-unknown membrane molecule (31) raised the possibility that NDF isoforms differ in the necessity of the putative coreceptor. However, we observed no differences in cell type specificity of the α and β isoforms. In other words, both isoforms specifically bind to mammary cells but display no specific interaction with fibroblasts and ovarian cells (data not shown).

Juxtamembrane variants. The observation that a truncated version of NDF that lacks the whole juxtamembrane stretch shares the same receptor affinity with the longer NDF-α2 isoform (Fig. 4) implies that the region proximal to the EGF domain is not involved in receptor binding and may have a different function. A putative proteolysis site (Lys-Arg) is shared by all the transmembrane forms of NDF, but additional cleavage sites may be present within the variable part of the juxtamembrane region. The fact that pro-NDF-β2a, which carries no variable juxtamembrane sequence, is nevertheless processed in COS-7 cells (Fig. 5) indicates that the shared cleavage site is indeed functional and that the variable part is dispensable for proteolytic cleavage and glycosylation of the precursor molecule. Nevertheless, the rate and tissue specificity of pro-NDF processing may be isoform specific. For example, NDF-β4a carries the longest juxtamembrane variable sequence (27 amino acids long), which contains five acidic residues and nine basic amino acids. The latter are clustered in one tetrad and one pair, which may function as processing sites (Fig. 1B). Metabolic labeling, however, revealed no accelerated processing of this isoform in COS-7 cells (Fig. 5). It is worthwhile, however, to note that the similarly localized cleavage sites of the precursors of AR, EGF, and TGF-α have different sequences and may be processed by distinct proteases. It is, therefore, not unlikely that the variable juxtamembrane sequences of NDF isoforms function as targets for tissue-specific proteases that control the balance between soluble and membrane-bound forms of the factors. These proteases may not exist in COS-7 cells.

Cytoplasmic tail variants. The three variants of the cytoplasmic tail of NDF appear to undergo proper glycosylation and release when expressed in COS-7 cells. However, their rates of proteolytic cleavage, as inferred from the relative abundance of the precursor forms, display variation (Fig. 5). The longest isoform, NDF-α2a, is apparently the most rapidly processed form, whereas the shortest precursor, NDF-α2c, undergoes relatively slow maturation. Other NDF isoforms that carry the long 374-amino-acid tail, namely, NDF-β2a and NDF-β4a, also displayed efficient processing. This raises the possibility that the length, or specific sequence, of the cytoplasmic tail determines the rate of NDF processing. Another possibility is that the identity of the carboxy-terminal amino acid affects the rate of proteolytic processing. Pro-TGF-α has a carboxy-terminal valine that is critical for regulated release of TGF-α, but leucine or isoleucine can substitute for valine in this requirement (2). Similarly, the a isoforms of NDF have a terminal valine residue, and the 196-amino-acid-long tail of pro-NDF-α2b has a terminal leucine. By contrast, the 157-amino-acid-long cytoplasmic tail of pro-NDF-α2c ends with two hydrophilic residues, which may account for its low rate of processing. Lastly, the cytoplasmic tails of rat, human, and chicken NDF molecules display more than 85% sequence identity, suggesting that this curiously long domain has a functional role other than determining turnover rate of the precursor molecules.

Amino terminus. The EGF-like domain of NDF isoforms is preceded by a variable-length sequence that contains one invariant motif, an immunoglobulin-like domain, and two mutually exclusive domains that flank it at both sides. These are the neuronal cell-specific kringle domain and the apparently mesenchyme-specific glycosylated spacer domain. Since isoforms that contain only the EGF-like domain displayed receptor affinities comparable to those of their amino-terminally extended forms, it is safe to assume that the amino-terminal portion of NDF is not involved in receptor recognition. This conclusion is supported by the lack of receptor binding interference by a recombinant amino-terminal portion of NDF (Fig. 4B). Instead, the functional role of the variable amino-terminal portion of NDF isoforms may involve interaction with other tissue-specific molecules and, therefore, may not be revealed by analyses that are limited to cultured cells.

Transmembrane domain. Unlike the transmembrane domain-containing forms of NDF, isoform β3 undergoes no glycosylation (Fig. 6) and is not released from transfected COS-7 cells (Fig. 5). Intracellular localization of NDF-β3 is consistent with the absence of biological activity in the medium of cells that were transfected with the corresponding cDNAs of NDF, GGF, and HRG (Fig. 3) (15, 23). Pro-NDF-β3 has essentially the entire pro-NDF extracellular domain (Fig. 1). In contrast, cDNA clone 40 encodes a truncated pro-NDF-β2 comprising the extracellular domain, the 23-amino-acid-long hydrophobic domain, and the first 36 amino acid residues of the cytoplasmic domain. This truncated cDNA clone directed expression of biologically active NDF that accumulated in media conditioned by transfected cells. Thus, the capacity for secretion may be localized to the internal 23-amino-acid-long hydrophobic domain. The role of the intracellular NDF-β3 is

unknown. This NDF isoform may be released from cells by a different mechanism. The NDF-β3 isoform may function as a sequestered growth factor that is released when tissues are injured. A similar role has been proposed for interleukin-1β and fibroblast growth factors, which also lack signal peptide sequences (26). Alternatively, NDF-β3 could have intracellular functions. Consensus nuclear targeting sequences are present in AR and pro-HRG (15, 32). Homologous sequences in the N-termini of pro-NDFs may direct the intracellular pro-NDF-β3 to the nucleus.

In summary, our study attempted to resolve the structural and functional bases of the multiplicity of NDF isoforms. Exhaustive cDNA cloning characterized the structures of six fibroblastic and epithelial isoforms, but genomic cloning and other approaches may be required in order to determine whether additional forms exist in these cells. We now distinguish between mesenchyme- and neuronal cell-specific groups of isoforms that differ in the structure of their extracellular portions. At present, the functional implications of the N-terminal heterogeneity is less understood than the variability that is displayed by the carboxy-terminal portions of NDF isoforms. This latter variation is presumably responsible for two functional differences: variation within the EGF-like domain confers differences in binding affinity, whereas the different structures of the juxtamembrane and cytoplasmic domains apparently affect the rate of processing of the transmembrane precursor into a soluble molecule. These conclusions, however, were derived from experiments that were performed with only a few isoforms, and they are limited by the possibility that recombinant molecules only partially represent their endogenous counterparts. It is, nevertheless, tempting to speculate that the NDF family plays multiple biological roles that have spatial and temporal specificities which are achieved by the combinatorial design of the mosaic structure of NDF.

ACKNOWLEDGMENTS

We thank Burt Goodman, Jennifer Hsu-Chen, and Sara Lavi for technical assistance, Elior Peles for preliminary binding experiments, and Miriam Fagan for typing the manuscript.

This work was supported by Amgen, Inc., and in part by the Laub Foundation. Y.Y. is an incumbent of a Research Career Development Award from The Israel Cancer Research Fund.

REFERENCES

1. **Bell, G. I., N. M. Fong, M. M. Stemplen, M. A. Wormsted, D. Caput, L. L. Ku, M. S. Urdea, L. B. Rall, and R. Sanches-Pescador.** 1986. Human epidermal growth factor precursor: cDNA sequence, expression *in vitro* and gene organization. Nucleic Acids Res. **14**:8427–8446.
2. **Bosenberg, M. W., A. Pandiella, and J. Massague.** 1992. The cytoplasmic carboxy-terminal amino acid specifies cleavage of membrane TGFα into soluble growth factor. Cell **71**:1157–1165.
3. **Campion, S. R., M. K. Geck, and S. K. Niyogi.** 1993. Cumulative effect of double-site mutations of human epidermal growth factor on receptor binding. J. Biol. Chem. **268**:1742–1748.
4. **Cantley, L. C., K. R. Auger, C. Carpenter, B. Duckworth, A. Graziani, R. Kapeller, and S. Soltoff.** 1991. Oncogenes and signal transduction. Cell **64**:281–302.
5. **Cohen, J. A., D. B. Weiner, K. F. More, Y. Kokai, W. V. Williams, H. J. Maguire, V. A. LiVolsi, and M. I. Greene.** 1989. Expression pattern of the neu (NGL) gene-encoded growth factor receptor protein ($p185^{neu}$) in normal and transformed epithelial tissues of the digestive tract. Oncogene **4**:81–88.
6. **Cohen, J. A., A. T. Yachnis, M. Arai, J. G. Davis, and S. S. Scherer.** 1992. Expression of the *neu* proto-oncogene by Schwann cells during peripheral nerve development and Wallerian degeneration. J. Neurosci. Res. **31**:622–634.
7. **D'Emilia, J., K. Bulovas, K. D'Ercole, B. Wolf, G. J. Steele, and I. C. Summerhayes.** 1989. Expression of the c-*erb*B-2 gene product (p185) at different stages of neoplastic progression in the colon. Oncogene **4**:1233–1239.
8. **Derynck, R., A. B. Roberts, D. H. Eaton, M. E. Winkler, and D. V. Goeddel.** 1985. Human transforming growth factor-α: precursor sequence, gene structure, and heterologous expression. Cancer Cells **3**:79–86.
9. **Dobashi, K., J. G. Davis, Y. Mikami, J. K. Freeman, J. Hamuro, and M. I. Greene.** 1991. Characterization of a *neu*/c-*erb*B-2 protein-specific activating factor. Proc. Natl. Acad. Sci. USA **88**:8582–8586.
10. **Dugeon, T. J., R. M. Cooke, M. Baron, I. D. Campbell, R. M. Edwards, and A. Fallon.** 1990. Structure-function analysis of epidermal growth factor: site directed mutagenesis and nuclear magnetic resonance. FEBS Lett. **261**:392–396.
11. **Engler, D. A., S. R. Campion, M. R. Hanser, J. S. Cook, and S. K. Niyogi.** 1992. Critical functional requirements for the guanidinium group of the arginine 41 side chain of human epidermal growth factor as revealed by mutagenic inactivation and chemical reactivation. J. Biol. Chem. **267**:2274–2281.
12. **Falls, D. L., K. M. Rosen, G. Corfas, W. S. Lane, and G. D. Fischbach.** 1993. ARIA, a protein that stimulates acetylcholine receptor synthesis, is a member of the Neu ligand family. Cell **72**:801–815.
13. **Gullick, W. J., M. S. Berger, P. L. Bennett, J. B. Rothbard, and M. D. Waterfield.** 1987. Expression of the c-*erb*B-2 protein in normal and transformed cells. Int. J. Cancer **40**:246–254.
14. **Herrera, G. A.** 1991. c-*erb*B-2 amplification in cystic renal disease. Kidney Int. **40**:509–513.
15. **Holmes, W. E., M. X. Sliwkowski, R. W. Akita, W. J. Henzel, J. Lee, J. W. Park, D. Yansura, N. Abadi, H. Raab, G. D. Lewis, M. Shepard, W. I. Wood, D. V. Goeddel, and R. L. Vandlen.** 1992. Identification of heregulin, a specific activator of p185*erb*B2. Science **256**:1205–1210.
16. **Huang, S. S., and J. S. Huang.** 1992. Purification and characterization of the *neu*/*erb*B-2 ligand-growth factor from bovine kidney. J. Biol. Chem. **267**:11508–11512.
17. **Jankowski, J., G. Coghill, D. Hopwood, and K. G. Wormsley.** 1992. Oncogenes and onco-suppressor genes in adenocarcinoma of the oesophagus. Gut **33**:1033–1038.
18. **Kokai, Y., J. A. Cohen, J. A. Drebin, and M. I. Greene.** 1987. Stage- and tissue-specific expression of the neu oncogene in rat development. Proc. Natl. Acad. Sci. USA **84**:8498–8501.
19. **Land, H., L. F. Parada, and R. A. Weinberg.** 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two operating oncogenes. Nature (London) **304**:596–602.
20. **Lofts, F. J., and W. J. Gullick.** 1992. Breast cancer: cellular and molecular biology. Kluwer Academic Publishers, Boston.
21. **Lupu, R., R. Colomer, G. Zugmaier, J. Sarup, M. Shepard, D. Slamon, and M. Lippman.** 1990. Direct interaction of a ligand for the *erb*B-2 oncogene product with EGF receptor and p185*erb*B-2. Science **249**:1552–1555.
22. **Maguire, H. C. J., C. Jaworsky, J. A. Cohen, M. Hellman, D. B. Weiner, and M. I. Greene.** 1989. Distribution of neu (c-*erb*B-2) protein in human skin. J. Invest. Dermatol. **92**:786–790.
23. **Marchionni, M. A., A. D. J. Goodearl, M. S. Chen, O. Bermingham-McDonogh, C. Kirk, M. Hendricks, F. Denehy, D. Misumi, J. Sudhalter, K. Kobayashi, D. Wroblewski, C. Lynch, M. Baldassare, I. Hiles, J. B. Davis, J. J. Hsuan, N. F. Totty, M. Otsu, R. N. McBury, M. D. Waterfield, P. Stroobant, and D. Gwynne.** 1993. Glial growth factors are alternatively spliced *erb*B-2 ligands expressed in the nervous system. Nature (London) **362**:312–318.
24. **Massagué, J.** 1990. Transforming growth factor-α: a model for membrane anchored growth factors. J. Biol. Chem. **265**:21393–21396.
25. **Matsunami, R. K., M. L. Yette, A. Stevens, and S. K. Niyogi.** 1991. Mutational analysis of leucine 47 in human epidermal growth factor. J. Cell. Biochem. **46**:242–249.
26. **Mitzutani, H., R. Black, and T. S. Kupper.** 1991. Human keratinocytes produce but do not process pro-interleukin-1β: different strategies of IL-1 production and processing in monocytes and keratinocytes. J. Clin. Invest. **87**:1066–1071.
27. **Mori, S., T. Akiyama, Y. Yamada, Y. Morishita, I. Sugaware, K.**

Toyoshima, and T. Yamamoto. 1989. c-erbB-2 gene product, a membrane protein commonly expressed on human fetal epithelial cells. Lab. Invest. **61**:93–97.

28. **Natali, P. G., M. R. Nicotra, A. Bigotti, I. Venturo, D. J. Slamon, B. M. Fendly, and A. Ullrich.** 1990. Expression of the p185 encoded by HER-2 oncogene in normal and transformed human tissues. Int. J. Cancer **45**:457–461.
29. **Orr-Urtreger, A., L. Trakhtenbrot, R. Ben-Levi, D. Wen, G. Rechavi, P. Lonai, and Y. Yarden.** 1993. Neural expression and chromosomal mapping of Neu differentiation factor to 8p12-p21. Proc. Natl. Acad. Sci. USA **90**:1746–1750.
30. **Peles, E., S. S. Bacus, R. A. Koski, H. S. Lu, D. Wen, S. G. Ogden, R. Ben-Levy, and Y. Yarden.** 1992. Isolation of the *neu*/HER-2 stimulatory ligand: a 44 kd glycoprotein that induces differentiation of mammary tumor cells. Cell **69**:205–216.
31. **Peles, E., R. Ben-Levy, E. Tzahar, L. Naili, D. Wen, and Y. Yarden.** 1993. Cell-type specific interaction of Neu differentiation factor (NDF/heregulin) with Neu/HER-2. EMBO J. **12**:961–971.
32. **Plowman, G. D., G. S. Whitney, M. G. Neubauer, J. M. Green, V. I. McDonald, G. J. Todaro, and M. Shoyab.** 1990. Molecular cloning and expression of an additional epidermal growth factor receptor-related gene. Proc. Natl. Acad. Sci. USA **87**:4905–4909.
33. **Press, M. F., C. Cordon-Cardo, and D. J. Slamon.** 1990. Expression of the HER-2/*neu* proto-oncogene in normal human adult and fetal tissues. Oncogene **5**:953–962.
34. **Quirke, P., A. Pickles, N. Tuzi, O. Mohamdee, and W. Gullick.** 1989. Pattern of expression of c-*erb*B-2 oncoprotein in human fetuses. Br. J. Cancer **60**:64–69.
35. **Ullrich, A., and J. Schlessinger.** 1990. Signal transduction by receptors with tyrosine kinase activity. Cell **61**:203–212.
36. **Wen, D., E. Peles, R. Cupples, S. V. Suggs, S. S. Bacus, Y. Luo, G. Trail, S. Hu, S. M. Silbiger, R. B. Levy, Y. Luo, and Y. Yarden.** 1992. Neu differentiation factor: a transmembrane glycoprotein containing an EGF domain and an immunoglobulin homology unit. Cell **69**:559–572.