

# cDNA and amino acid sequences of the cell adhesion protein receptor recognizing vitronectin reveal a transmembrane domain and homologies with other adhesion protein receptors

(molecular cloning/cytoplasmic domain/peptidolytic processing/lymphocyte function-associated antigen 1/macrophage antigen 1 homology)

SHINTARO SUZUKI, W. SCOTT ARGRAVES, ROBERT PYTELA\*, HIROHARU ARAI, TOM KRUSIUS†, MICHAEL D. PIERSCHBACHER, AND ERKKI RUOSLAHTI

Cancer Research Center, La Jolla Cancer Research Foundation, La Jolla, CA 92037

Communicated by Renato Dulbecco, July 28, 1986

**ABSTRACT** Cells adhere to vitronectin substrates through a cell surface receptor that recognizes an Arg-Gly-Asp sequence in vitronectin. The receptor is a glycoprotein composed of a 150-kDa  $\alpha$  and a 115-kDa  $\beta$  subunit. The  $\alpha$  subunit consists of two disulfide-bonded chains of 125 kDa and 25 kDa. cDNA clones were isolated for the  $\alpha$  subunit of the vitronectin receptor from a phage  $\lambda$ gt11 expression library made with RNA from a human fibroblast cell line, IMR-90. The identity of the clones that had been selected from the library based on immunological criteria was verified by comparison of DNA and protein sequences. NH<sub>2</sub>-terminal sequences were obtained for each of the  $\alpha$ -subunit chains. The sequence of the 25-kDa chain of the  $\alpha$  subunit was found in a cDNA clone, and the amino acid sequence deduced from the cDNA establishes the complete amino acid sequence of the 25-kDa chain. This chain contains a membrane-spanning domain as well as a putative intracytoplasmic region that is 32 amino acids long and consists mostly of polar amino acids. Comparison of the cDNA and protein sequences shows that the 25-kDa chain is generated by proteolytic cleavage of an  $\alpha$ -subunit precursor, the partial sequence of which is contained in the cDNA clones. These clones contain 1910 base pairs of open reading frame and a 3' untranslated sequence. RNA blot hybridization detected one transcript of about 7 kilobases in RNA from fibroblastic and epithelial cells. Together, the cDNA clones cover 4442 bases of this RNA. The  $\alpha$ -subunit sequence showed strong homology with the sequence of the  $\alpha$  subunit of fibronectin receptor. Moreover, the NH<sub>2</sub>-terminal protein sequence of the 125-kDa chain was homologous with the NH<sub>2</sub>-terminal sequences of two other cell surface proteins, lymphocyte function-associated antigen 1 (LFA-1) and macrophage antigen 1 (Mac-1), which have been implicated as receptors for adhesion proteins of leukocytes. These results establish several of the structural features in the vitronectin receptor and suggest the existence of a superfamily of receptors for cell adhesion proteins.

Adhesive interactions between cells and their environment are of fundamental importance in processes such as embryonic development, organ function, and tumor-cell invasion as well as in the function of cells comprising the immune system. It has become clear that the major mediators of such interactions are cell surface glycoproteins that specifically bind extracellular protein ligands. Many of these receptors belong to a group of receptors that recognize an Arg-Gly-Asp (RGD in the single-letter code for amino acids) sequence in their ligand protein (1, 2). This receptor group includes the fibronectin receptor (FNR) (3), the vitronectin receptor (VNR) (4), and the gpIIb/IIIa glycoprotein of platelets (5),

which recognizes the Arg-Gly-Asp sequence of fibrinogen, fibronectin, and vitronectin (6, 7) as well as that of von Willebrand factor (8). In addition, a chicken receptor appears to recognize fibronectin and laminin in an Arg-Gly-Asp-dependent manner (9, 10).

The VNR is present in various types of cells and can be isolated in chemical amounts by affinity chromatography (4, 7). VNR consists of two major subunits, one of about 150 kDa and another of about 115 kDa. We refer to these subunits as the  $\alpha$  and  $\beta$  subunit, respectively. Upon reduction, the larger ( $\alpha$ ) subunit loses about 25 kDa of its apparent molecular mass and a small polypeptide can be seen on gels, suggesting the existence of a small chain disulfide bonded to the rest of the subunit. VNR is thought to be an integral membrane protein because it can be incorporated into liposomes (4). The structures of FNR and gpIIb/IIIa are similar to those of VNR. VNR and gpIIb/IIIa are especially closely related to each other with regard to the molecular sizes of the subunits (7). All of these structural and functional similarities suggest that the receptors are homologous membrane proteins with a shared evolutionary origin. However, since no amino acid sequence data have been available for these receptors, such relationships remain unproven.

We report here protein and cDNA sequencing results that reveal a typical transmembrane domain in VNR, establish sequence homologies within the Arg-Gly-Asp receptor family, and suggest that a group of leukocyte proteins is related to the Arg-Gly-Asp receptor family.

## MATERIALS AND METHODS

**Isolation of cDNA Clones.** A phage  $\lambda$ gt11 expression cDNA library (11) was made from the poly(A)-containing mRNA of IMR-90 human fibroblasts as described (12) and was screened by using antibodies affinity-purified from a rabbit anti-VNR antiserum. The antiserum was prepared by immunizing with VNR isolated from human placenta (13). Anti-VNR antibodies were purified from the antiserum by chromatography on VNR-Sepharose, and antibodies specific to the  $\alpha$  and  $\beta$  subunits of VNR were affinity-purified by using each of the subunits immobilized on nitrocellulose filters (14). Positive clones from the antibody screening were plaque-purified and subjected to further immunological analysis in which the hybrid proteins produced by the positive clones were used to purify antibodies. About  $1 \times 10^5$  plaque-forming units were

Abbreviations: kbp, kilobase pair(s); FNR, fibronectin receptor; LFA-1, lymphocyte function-associated antigen 1; Mac-1, macrophage antigen 1; VNR, vitronectin receptor.

\*Present address: Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland.

†Present address: Department of Medical Chemistry, University of Helsinki, Siltavuorenpenger 10, 00170 Helsinki, Finland.

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

plated on a 7.5-cm Petri dish and incubated for 2.5 hr at 42°C until small plaques became visible. A nitrocellulose filter paper soaked with isopropyl  $\beta$ -D-thiogalactopyranoside was put on top of the agarose and incubated for 6 hr at 37°C. The resultant filter was removed and washed four times with 50 mM Tris-HCl, pH 7.5/150 mM NaCl/0.05% Tween-20 and then incubated with antiserum. The bound antibody was eluted with 5 mM glycine-HCl, pH 2.3/0.5 M NaCl/0.5% Tween 20/0.01% bovine serum albumin (14). The reactivity of the resultant antibodies was studied by immunoblotting (15). VNR (0.5  $\mu$ g) was separated by NaDodSO<sub>4</sub>/PAGE using 7% gels. The separated polypeptides were blotted on nitrocellulose filters and treated with the appropriate antibodies. The bands were visualized with horseradish peroxidase-labeled goat anti-rabbit IgG (Bio-Rad).

**DNA Sequencing.** Phage DNAs were purified by the plate lysate method (16). After *Eco*RI digestion, the inserts were separated on agarose gels, extracted, and subcloned into phage M13mp19. The DNA sequences were determined by the dideoxynucleotide chain-termination method of Sanger *et al.* (17) using adenosine 5'-[ $\alpha$ -thio]triphosphate labeled at the  $\alpha$ -thio position with <sup>35</sup>S. To facilitate the sequencing, a series of overlapping deletion clones of both strands were made as described by Dale *et al.* (18). In both strands, several undetermined parts remained, which were then sequenced by using an oligonucleotide priming method as described (19).

**NH<sub>2</sub>-Terminal Amino Acid Sequencing.** The NH<sub>2</sub>-terminal amino acid sequences of the VNR subunits were determined as follows. Approximately 300  $\mu$ g of the isolated VNR, containing <sup>125</sup>I-labeled VNR as a radioactive tracer, was separated by NaDodSO<sub>4</sub>/PAGE under reducing conditions. The subunits were localized on the dried gel by autoradiography, excised, and electrophoretically eluted from the rehydrated gel. After extensive dialysis against 0.01 M NH<sub>4</sub>HCO<sub>3</sub>/0.02% NaDodSO<sub>4</sub> and lyophilization, the protein was subjected to gas-phase microsequencing using an Applied Biosystems (Foster City, CA) model 470A sequencer and the trifluoroacetic acid-based chemistry provided by the manufacturer.

**Other Procedures.** mRNA was isolated from cultured cells as described by Rowe *et al.* (20). RNA and DNA blot analyses were performed as described by Thomas (21). The probes were labeled by nick-translation or by random oligonucleotide primer extension (22).

## RESULTS

**Identification of a cDNA Coding for the Vitronectin Receptor.** Screening of 10<sup>6</sup> plaques from a  $\lambda$ gt11 library constructed from the human fibroblast cell line IMR-90 revealed 14 clones that reacted with anti-VNR antibodies. These clones were classified into four different groups based on DNA cross-hybridization and reactivity with antibodies against the subunits of VNR.

One of the groups consisted of two clones ( $\lambda$ VNR10 and  $\lambda$ VNR11) that appeared to code for the  $\alpha$  subunit. Each of the clones reacted with the affinity-purified anti-VNR antibody in blotting experiments; when subunit-specific antibodies were used, the antibody against the  $\alpha$  subunit of VNR recognized the clones, whereas the anti- $\beta$  subunit antibody did not. Moreover, antibodies affinity-purified on the hybrid proteins produced by these clones reacted with the  $\alpha$  subunit of VNR but not with the  $\beta$  subunit. The result for the  $\lambda$ VNR10 fusion protein is shown in Fig. 1. These results indicate that the  $\alpha$  subunit of VNR and the two hybrid proteins share one or more epitopes.

**DNA Sequence of the VNR  $\alpha$ -Subunit Clones and the Deduced Amino Acid Sequence.** The combined DNA sequence of the  $\lambda$ VNR10 and  $\lambda$ VNR11 inserts and the deduced amino acid sequence are shown in Fig. 2. The size of the

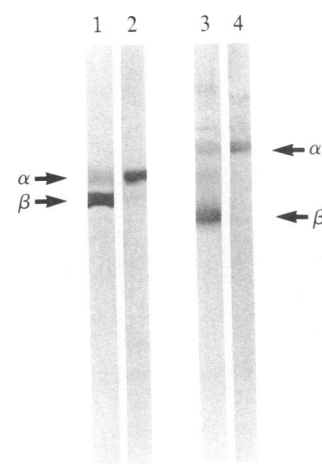


FIG. 1. Staining of electrophoretically separated VNR polypeptides by antibodies purified with a hybrid protein encoded by cloned cDNA. Antibody bound from anti-VNR antiserum to the hybrid protein encoded by  $\lambda$ VNR10 was eluted and used for immunoblotting analysis of VNR separated under reducing (lanes 1 and 2) and nonreducing (lanes 3 and 4) conditions. The antibody preparation obtained with the hybrid protein stains the upper band of the VNR receptor (lanes 2 and 4), whereas the original antibody preparation stains both bands (lanes 1 and 3). The 25-kDa subunit was not reproducibly detected.

$\lambda$ VNR10 insert is about 1.3 kilobase pairs (kbp) and that of  $\lambda$ VNR11 about 3.9 kbp. The two inserts share 773 bp of overlapping sequence. One open reading frame of 1910 bp extends from the 5' end of the  $\lambda$ VNR10 insert to the middle of the  $\lambda$ VNR11 insert. The 3' half of the  $\lambda$ VNR11 insert corresponds to an untranslated region, but since no poly(A) tail is present, it may not be complete.

A possible membrane-spanning domain consisting of hydrophobic amino acid residues is seen near the COOH terminus of the open reading frame. At the COOH terminus is a putative intracytoplasmic region that is 32 amino acids long and has a highly hydrophilic character. Ten possible N-linked glycosylation sites are present throughout the open reading frame; they are all located in the region believed to be extracellular.

A protein homology search carried out by using the Bionet system showed that the VNR sequences established from the cDNA clones are unique, with no significant homologies with other proteins in the database. Protein sequencing performed to confirm the identification of the clones and to correlate the cDNA sequences with the two chains of the  $\alpha$  subunit did reveal homologies with known sequences (see below).

**NH<sub>2</sub>-Terminal Amino Acid Sequences of the  $\alpha$ -Subunit Chains.** NH<sub>2</sub>-terminal amino acid sequences determined for the 25-kDa and 125-kDa chains of the  $\alpha$  subunit are shown in Fig. 3. The NH<sub>2</sub>-terminal sequence of the 25-kDa chain was found in the cDNA-derived sequence (bases 1437–1469 in Fig. 2). This confirms the identity of our cDNA clones as VNR  $\alpha$ -subunit clones and shows that the complete sequence of the light chain is contained within the sequences derived from these clones. The NH<sub>2</sub> terminus of the 25-kDa chain is preceded in the cDNA sequence by a sequence coding for two basic amino acids, which presumably serve as a recognition signal for the cleavage, as is the case with many other proteins. Also shown in Fig. 3 is the homology of the NH<sub>2</sub>-terminal sequence of the 125-kDa chain with the NH<sub>2</sub>-terminal sequences of the leukocyte adhesion proteins Mac-1 (macrophage antigen 1) and LFA-1 (lymphocyte function-associated antigen 1) reported earlier (23).

**mRNA Analysis.** RNA blotting with the VNR clones revealed one strongly hybridizing transcript of about 7 kilobases in the IMR-90 fibroblasts (Fig. 4). The same

⤴λVNR-10  
 GC TTG AAC GCA GTC CCA TCT CAA ATC CTT GAA GGG CAG TGG GGT GCT CGA AGC ATG CCA CCA AGC TTT GGC TAT TCA ATG AAA GGA GCC 89  
 Leu Asn Ala Val Pro Ser Gln Ile Leu Glu Gly Gln Trp Ala Ala Arg Ser Met Pro Pro Ser Phe Gly Tyr Ser Met Lys Gly Ala  
  
 ACA GAT ATA GAC AAA AAT GGA TAT CCA GAC TTA ATT GTA GGA GCT TTT GGT GTA GAT CGA GCT ATC TTA TAC AGG GCC AGA CCA GTT ATC 179  
 Thr Asp Ile Asp Lys Asn Gly Tyr Pro Asp Leu Ile Val Gly Ala Phe Gly Val Asp Arg Ala Ile Leu Tyr Arg Ala Arg Pro Val Ile  
  
 ACT GTA AAT GCT GGT CTT GAA GTG TAC CCT AGC ATT TTA AAT CAA GAC AAT AAA ACC TGC TCA CTG CCT GGA ACA GCT CTC AAA GTT TCC 269  
 Thr Val Asn Ala Gly Leu Glu Val Tyr Pro Ser Ile Leu Asn Gln Asp Asn Lys Thr Cys Ser Leu Pro Gly Thr Ala Leu Lys Val Ser  
  
 TGT TTT AAT GTT AGG TTC TGC TTA AAG GCA GAT GGC AAA GGA GTA CTT CCC AGG AAA CTT AAT TTC CAG GTG GAA CTT CTT TTG GAT AAA 359  
 Cys Phe Asn Val Arg Phe Cys Leu Lys Ala Asp Gly Lys Cys Leu Arg Lys Leu Asn Phe Gln Val Glu Leu Leu Leu Asp Lys  
  
 CTC AAG CAA AAG GGA GCA ATT CGA GCA GCA CTG TTT CTC TAC AGC AGG TCC CCA AGT CAC TCC AAG AAC ATG ACT ATT TCA AGG GGG GGA 449  
 Leu Lys Gln Lys Gly Ala Ile Arg Arg Ala Leu Phe Leu Tyr Ser Arg Ser Pro Ser His Ser Lys Asn Met Thr Ile Ser Arg Gly Gly  
  
 CTG ATG CAG TGT GAG GAA TTG ATA GCG TAT CTG GGG GAT GAA TCT GAA TTT AGA GAC AAA CTC ACT CCA ATT ACT ATT TTT ATG GAA TAT 539  
 Leu Met Gln Cys Glu Glu Leu Ile Ala Tyr Leu Arg Asp Glu Ser Glu Phe Arg Asp Lys Leu Thr Pro Ile Thr Ile Phe Met Glu Tyr  
  
 CGG TTG GAT TAT AGA ACA GCT GCT GAT ACA ACA GGC TTG CAA CCC ATT CTT AAC CAG TTC ACG CCT GCT AAC ATT AGT CGA CAG GCT CAC 629  
 Arg Leu Asp Tyr Arg Thr Ala Ala Asp Thr Thr Gly Leu Gln Pro Ile Leu Asn Gln Phe Thr Pro Ala Asn Ile Ser Arg Gln Ala His  
  
 ATT CTA CTT GAC TGT GGT GAA GAC AAT GTC TGT AAA CCC AAG CTG GAA GTT TCT GTA AGT GAT CAA AAG AAG ATC TAT ATT GGG GAT 719  
 Ile Leu Asp Asp Cys Gly Glu Asp Asn Val Lys Cys Pro Lys Leu Glu Val Ser Val Asp Ser Asp Gln Lys Lys Ile Tyr Ile Gly Asp  
  
 GAC AAC CCT CTG ACA TTG ATT GTT AAG GCT CAG AAT CAA GGA GAA GGT GCC TAC GAA GCT GAG CTC ATC GTT TCC ATT CCA CTG CAG GCT 809  
 Asp Asn Pro Leu Thr Leu Ile Val Lys Ala Gln Asn Gln Gly Ala Tyr Glu Ala Glu Leu Ile Val Ser Ile Pro Leu Gln Ala  
  
 GAT TTC ATC GGG GTT GTC CGA AAT GAA GCC TTA GCA AGA CTT TCC TGT GCA TTT AAG ACA GAA AAC CAA ACT CGC CAG GTG GTA TGT 899  
 Asp Phe Ile Gly Val Val Arg Asn Asn Glu Ala Leu Ala Arg Leu Ser Cys Ala Phe Lys Thr Glu Asn Gln Thr Arg Gln Val Val Cys  
  
 GAC CTT GGA AAC CCA ATG AAG GCT GCA ACT CAA CTC TTA GCT GGT CTT CGT TTC AGT GTG CAG CAG TCA GAG ATG GAT ACT TCT TCT GTG 989  
 Asp Leu Gly Asn Pro Met Lys Ala Gly Thr Gln Leu Leu Ala Gly Leu Arg Phe Ser Val His Gln Gln Ser Glu Met Asp Thr Ser Val  
  
 AAA TTT GAC TTA CAA ATC CAA AGC TCA AAT CTA TTT GAC AAA GTA AGC CCA GTT GTA TCT CAC AAA GTT GAT CTT GCT GTT TTA GCT GCA 1079  
 Lys Phe Asp Leu Gln Ile Gln Ser Ser Asn Leu Phe Asp Lys Val Ser Pro Val Val Ser His Lys Val Asp Leu Ala Val Leu Ala Ala  
  
 GTT GAG ATA AGA GGA GTC TCG AET GCT CAT ATC TTT CTT CCG ATT CCA AAC TGG GAG CAG AAG GAG AAC CCT GAG ACT GAA GAA GAT 1169  
 Val Glu Ile Arg Gly Val Ser Ser Pro Asp His Ile Phe Leu Pro Ile Pro Asn Trp Glu His Lys Glu Asn Pro Glu Thr Glu Glu Asp  
  
 GTT GGG CCA GTT GTT CAG CAC ATC TAT GAG CTG AGA AAC AAT GGT CCA AGT TCA TTC AGC AAG GCA ATG CTC CAT CTT CAG TGG CCT TAC 1259  
 Val Gly Pro Val Val Gln His Ile Tyr Glu Leu Arg Asn Asn Gly Pro Ser Ser Phe Ser Lys Ala Met Leu His Leu Gln Trp Pro Tyr  
  
 λVNR-10 ←  
 AAA TAT AAT AAT AAC ACT CTG TTG TAT ACT CTT CAT TAT GAT ATT GAT GGA CCA ATG AAC TGC ACT TCA GAT ATG GAG ACT AAC CCT TTG 1349  
 Lys Tyr Asn Asn Thr Leu Leu Tyr Ile Leu His Tyr Asp Ile Asp Gly Pro Met Asn Cys Thr Ser Asp Met Glu Ile Asn Pro Leu  
  
 AGA ATT AAG ATC TCA TCT TTG CAA ACA ACT GAA AAG AAT GAC ACG GTT GCC GGG CAA GGT GAG CGG GAC CAT CTC ACT AAG CGG GAT 1439  
 Arg Ile Lys Ile Ser Ser Leu Thr Glu Thr Val Ala Gly Gln Gly Glu Arg Asp His Leu Ile Thr Lys Arg Asp  
  
 CTT GCC CTC AGT GAA GGA GAT ATT CAC ACT TTG GGT TGT GGA GTT GCT CAG TGC TTG AAG ATT GTC TGC CAA GTT GGG AGA TTA GAC AGA 1529  
 Leu Ala Leu Ser Glu Gly Asp Ile His Thr Leu Gly Cys Gly Val Ala Gln Cys Leu Lys Ile Val Cys Gln Val Gly Arg Leu Asp Arg  
  
 GGA AAG AGT GCA ATC TTG TAC GTA AAG TCA TTA CTG TGG ACT GAG ACT TTT ATG AAT AAA GAA AAT CAG AAT CAT TCC TAT TCT CTG AAG 1619  
 Gly Lys Ser Ala Ile Leu Tyr Val Lys Ser Leu Leu Trp Thr Glu Thr Phe Met Asn Lys Glu Asn Gln Asn His Ser Tyr Ser Leu Lys  
  
 TCG TCT GCT TCA TTT AAT GTC ATA GAG TTT CCT TAT AAG AAT CTT CCA ATT GAG GAT ATC ACC AAC TCC ACA TTG GTT ACC ACT AAT GTC 1709  
 Ser Ser Ala Ser Phe Asn Val Ile Glu Phe Pro Tyr Lys Asn Leu Pro Ile Glu Asp Ile Thr Asn Ser Thr Leu Val Thr Thr Asn Val  
  
 ACC TGG GGC ATT CAG CCA GGC CCC ATG CCT GTG CCT GTG TGG GTG ATC ATT TTA GCA GTT CTA GCA GGA TTG TTG CTA CTG GCT GTT TTG 1799  
 Thr Trp Gly Ile Gln Pro Ala Pro Met Pro Val Pro Val Trp Val Ile Ile Leu Ala Val Leu Ala Val Leu Leu Leu Ala Val Leu  
  
 GTA TTT GTA ATG TAC AGG ATG GGC TTT TTT AAA GGG GTC CGG CCA CCT CAA GAA GAA CAA GAA AGG GAG CAG CTT CAA CCT CAT GAA AAT 1889  
 Val Phe Val Met Tyr Arg Met Gly Phe Phe Lys Arg Val Arg Pro Pro Gln Glu Glu Gln Glu Arg Glu Gln Leu Gln Pro His Glu Asn  
  
 GGT GAA GGA AAC TCA GAA ACT TAA CTG CAG TTT TTA AGT TAT GCT ACA TCT TGA CCC ACT AGA ATT AGC AAC TTT ATT ATA GAT TTA AAC 1979  
 Gly Glu Gly Asn Ser Glu Thr  
  
 TTT CTT CAT GAG GAG TAA AAA TCC AAG GCT TTA CTG CTG ATA GTG CTA ATT GGC ATT AAC CAA AAA ATG AGA ATT ATA TTT GTC AAC CTT 2069  
 CTC CTT ATA AAT AAG TTC AGA CAT ACA TTT AAT AAC ATA GGG TGA CTT GTG TTT TTA GGT ATT TAA ATA ATA AAA TTT CAA GGG ATA GTT 2159  
 CTT TTT AAA TGT ATA TAA GAC AGG TAG TGC CTG ATT TAC TAC TTA TAA AAT AGT ACC TCC TTC AGT TAC TGT TCT TGA TTT AAT GTA 2249  
 CGG AAC TTT ATT TGT 2339  
 GTA CTT AAT GTT AAT ACA TAT TAC ACT ACA GTT TAC TTT TCA GAA TAC TAA AGA CTT TAT AAC TGC ATG AAC TTG GAT TTT TTT AAT CAC 2429  
 TCA TAT GGT AGA ATT TTA TAA ACA ACT TCA TGA CAT CCA AAT TCT TGC TTT TAA TAA CAA AGG TAC AAT ATT TTG TTT TAG TAT GAA 2519  
 AAT CTG GTA GAT CCT ATT ACA CTT CTG TTT ATA TTA AAT CCA CAA TTT ATT ACA TTT TTA ATT ACA TTT TTA ATT ACA TTT TTA ATT ACA 2609  
 TCA AGC CAA CCT ATA CTA AAA ATT AGT TCC ATA ATC ACA AAT GGC TCT TTT GTG TAA TTG TTT AAT TTC ACC TGA ATA TCA TAA TGC TTA 2699  
 AAC CCA TAT GGA GTT GGA AAT TAT TTT CAA AGC ATA TTT ATT CGA TTG TTT TAG TCT GGC TAT TTA CAG TAT AAA AAA AGC ATT TTA TTA 2789  
 AAA TAC TGT GTA GTT CTT TGA AGT GGT TGA TGC ATA TAG TAA TGA TTA CAT TCT TAG AGT GGA GCA GAG TTT TTA GTT AGT ATT AAT 2879  
 TTA TTT TCC TCC ATT CAT GTA CTT TTT CTT ATA TTT CCA AAA CTG TTA CTG AGA ATG GGT CAA GAT CAG TGA GAA ATC TTT ACA GTT GAC 2969  
 AGG AAC CTG GAC CCC TTA CCC CAA CTT TAT GAG TAA TGC TTG GAA TAA AAA ACT CTT AAG GCA ACT CAC TGA TTT ACT TCT AGC AAT AGC 3059  
 ATG ATG TTA CAG GAA TAT TAC CTC TGT TTA AGC AAG GTA ATG TGT AAA ATC AGT CTT GGC TGT CAG AAT AAC TTC TAA AAG GTA TTT TTA 3149  
 TAA GCA GTT CAA GTT ACT GAA ACA CTT TTA AAC CTT TCT GAA GTT CTT TAG TAT AAA TTA CTT TTC TAG GAT TAT TAA TAA AAG CCA CAT 3239  
 AGG TGG CAA GGT GTA GTT TTA TAT GCT TCT GTA GAG TGG TGA ACC TTC TAG AGG AAT ATA TGA TTT ATT CAC AGT TCC TCA AGG CCT GGG 3329  
 GAT GAT GAT CAG TTA TAC CTA TTT TTG TGC AAT TAC ATC ATG TTG TAC ATT AGA AAT GGA GAG TTT AAT AGC TCT TTA ACT GCT GTC CTC 3419  
 ATT AGG TAA TGA TAA ATA TTT CCC TTA AAT AAT TGA CTA TTT TGC TGT GTT TTA AAA ATG ATT GAA ATT TAT CTT GCC ATA TCT CAT AAT 3509  
 TTC ATG CAG AAG TTG ACT GAG CTA ATC TTG AGA ATA TAT TCG TAA AAT AGG AGC ACA TTT AGT TGA GGT ATA CAA AGG ACT CTA GAC 3599  
 AAA ACC TTT TAT TTT AGC TTT AGT GAA TTT CAA AAG TAA TGG GTC TTG GAG TAT AGA TTT TTA TTA GGT GCT TGA AAG AGC TTA ATC ATA 3689  
 TGC AGT AAG TAT TTT TAT TAC CAA TAA ATT TAA AAT TTT TTA AGA AAA ATA TTT TTA TCC TAG GGC CAA GTG TTG CTT GCC ACC AAT CAG 3779  
 TAA GTT AGT CTA TAA CAA ATT TTA CCC TAA CAG TTT TAC CAC CTA GCA ACA GTC ATT TTT GAA AAT ATG TTG GAT AGA AGG TCA CTC TTT 3869  
 GGG AAA AGT GTT AGA ATT TGC TTT TGT GGC ATC TAT TCC TTT TAT GGC ACT TAT CTT GAA AGT AAT CTT GTA TTG GAG ATT GAA AGC TGC 3959  
 TGT AAT TTA GAA ATT AAC ATG ATA TCT AAT ACC TTT ATG AAA TAT AGT TTT GTA TAG CAT AGA TTT TCC TTC AAA AAA TGA ACA 4049  
 TTT ATA TAT CTA CAA AAA TAT GGA GAA GAG CAA TTT GAA AGC CTA CTT TCT GAA GAA AAT GGT GGG ATT TTT TTT TAT CAT GAT TAA ATA 4139  
 TCA AAA AAT TGC CCT ATG AAA ACT TTA AAT CTC TAA AAC ATT TGA AAT ACT ACC ATA TTT GTG ATT TAT TGA GAA TAA AAA TCC ATT TTG 4229  
 AAA TGT AAA ATT TTT ATG ATC TGA TCT AGT TTT AAG AAA ACA TGA ATG AAC TAG AAG ATA TTA AAA ACA TTT GAC ATT GGT AAG AAA TAT 4319  
 TGC TAC TGA TGT TTT TTA TAT AGG TAT TTT CAG AAT TGA TAT TTT GAG AAA AAT ACA TGT GAG TCA TTT TTT CTG TTT CTC TTT 4409  
 TCT CTT AAC GAT TAT CAC TGT AAT TCT GAT TCT  
 λVNR-11 ←

Fig. 2. DNA sequence and deduced amino acid sequence of cDNA clones λVNR10 and λVNR11. Both strands of λVNR10 and λVNR11 were sequenced by the dideoxynucleotide chain-termination method (17). The ends of the λVNR10 and λVNR11 clones are indicated. The cleavage site at the NH<sub>2</sub> terminus of the 25-kDa chain is marked by the arrow. The possible N-linked glycosylation sites are indicated by asterisks, and the hydrophobic stretch of amino acids contained in the putative membrane-spanning domain is underlined.

7-kilobase transcript was found in smaller amounts in MG-63 osteogenic sarcoma cells and in A431 carcinoma cells.

**DISCUSSION**

We have isolated two overlapping VNR cDNA clones and have derived the complete amino acid sequence of the

α-subunit 25-kDa chain and a portion of the α-subunit 125-kDa chain sequence from them. That our cDNA clones belong to the α subunit of VNR is based on two pieces of evidence. First, the hybrid proteins produced by the clones cross-react immunologically with the α subunit of VNR. Secondly, the cDNA-derived sequence contains the NH<sub>2</sub>-terminal protein sequence of the α-subunit 25-kDa chain.

	1	2	3	4	5	6	7	8	9	10	11
VNR $\alpha$ 25 kDa	X	Leu	Ala	Leu	Ser	Glu	Gly	Asp	Ile	X	Thr
VNR $\alpha$ 125 kDa	Phe	Asn	Leu	Asp	Val	X	Ser	Pro	Ala	Glu	Tyr
Mac-1 $\alpha$	Phe	Asn	Leu	Asp	Thr	Glu	His	Pro	Met	Thr	Phe
LFA-1 $\alpha$	Tyr	Asn	Leu	Asp	Thr	Arg	Pro	Thr	Gln	Ser	Phe

FIG. 3. NH<sub>2</sub>-terminal amino acid sequences of the 25-kDa and 125-kDa chains of VNR  $\alpha$  subunit and homology of the 125-kDa chain with the  $\alpha$  subunits of Mac-1 and LFA-1. VNR polypeptides (1150 pmol) were separated on NaDodSO<sub>4</sub>/PAGE under reducing conditions, and the  $\alpha$ -subunit polypeptides were cut out, eluted from the gels electrophoretically, and subjected to gas-phase amino acid sequencing. The yields for the 25-kDa and 125-kDa sequences were 300 and 580 pmol as determined from the recoveries for the second and first residues, respectively. Residues 1 and 10 of the 25-kDa polypeptide and residue 6 of the 125-kDa polypeptide could not be identified. The Mac-1 and LFA-1 sequences are from ref. 23. The boxes indicate identical residues and the conservative replacements, Tyr/Phe, Ser/Thr, or Thr/Val. The Thr/Val substitution is considered conservative because the side chains of Thr and Val are similar in size and have been found to be interchangeable in functionally active synthetic peptides (24).

Comparison of this 10-amino acid sequence with the corresponding sequence in the FNR (Fig. 5; ref. 25) shows that, although the two receptors show considerable overall homology, this region is not strongly homologous in the VNR and FNR. Since the 25-kDa chain NH<sub>2</sub>-terminal sequence we have used in the identification of the cDNA clones is poorly conserved in the VNR and FNR, it is unlikely that some related as-yet-unknown receptor would contain the same sequence as the VNR, supporting the assumption that the cDNA clones are VNR clones.

The sequence corresponding to the 25-kDa chain is near the 3' end of the open reading frame, suggesting that the 25-kDa chain is formed by proteolytic processing of a precursor. The remaining part of the cDNA corresponds to the 125-kDa chain of the  $\alpha$  subunit. Although the cDNA sequence does not extend far enough to include the NH<sub>2</sub>-terminal protein sequence of the 125-kDa chain, our immunological results show that the 125-kDa chain of the  $\alpha$  subunit and the hybrid proteins coded for by these clones are immunologically crossreactive. Thus, we can conclude that the cDNA sequences contain the complete amino acid sequence of the 25-kDa chain and a large part of the 125-kDa chain sequence. The open reading frame contained in the cDNA clones codes for  $\approx 70$  kDa of polypeptide sequence. The molecular mass of the VNR  $\alpha$  subunit is  $\approx 150,000$  (4, 7). Depending on the carbohydrate content of this subunit, which is not known, the sequence presented here includes about half or more of the polypeptide. NaDodSO<sub>4</sub>/PAGE

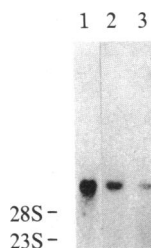


FIG. 4. RNA hybridization analysis. RNA from IMR-90, MG-63, and A431 cells (lanes 1-3, respectively) was electrophoresed on agarose gels, blotted onto nitrocellulose, and probed with <sup>32</sup>P-labeled  $\lambda$ VNR10 insert. The amount of RNA was 10  $\mu$ g in each case, and the RNAs in all three lanes were electrophoresed and probed together.

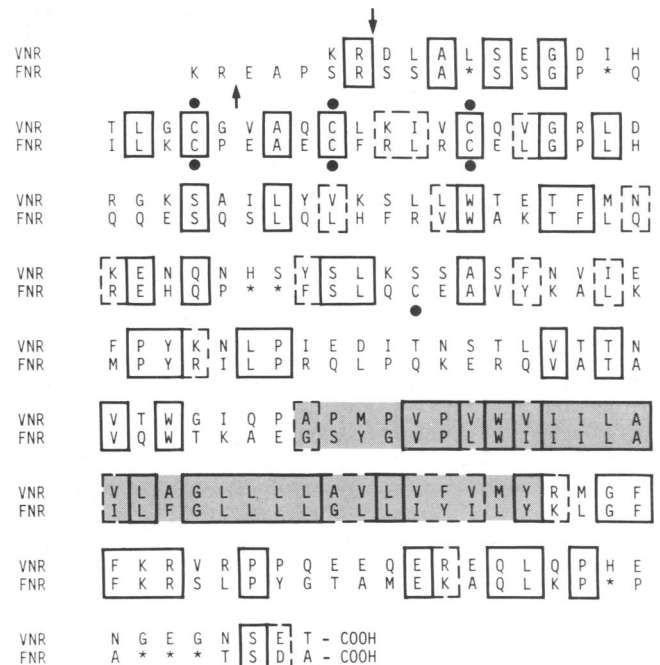


FIG. 5. Comparison of the deduced amino acid sequence of the 25-kDa chain of the VNR  $\alpha$  subunit with the corresponding portion of the FNR  $\alpha$  subunit (standard single-letter amino acid code). The boxes with solid lines indicate identical residues. Conservative replacements are boxed with dashed lines. The hydrophobic stretches believed to be transmembrane domains have been shaded. The boundaries of these domains were chosen by making them coincide with the first charged residue in either sequence. The FNR sequence is from ref. 25. A cleavage site analogous to the one in VNR (arrow pointing down) may generate a small chain from the  $\alpha$  subunit of the FNR (arrow pointing up) at the KRD and KRE sequences, respectively.

analysis of VNR presented here as well as earlier results (4, 7) indicates that the two chains that make up the  $\alpha$  subunit of VNR are linked by disulfide bonding. The extracellular domain of the 25-kDa chain contains three cysteine residues; at least one of these apparently forms a disulfide bond to the 125-kDa chain.

The 25-kDa chain contains a domain consisting of hydrophobic amino acid residues. This domain has features shared by transmembrane domains of other proteins, such as the uninterrupted row of hydrophobic amino acids and the basic residues commonly present on the cytoplasmic side of transmembrane domains (26-29). These structural features, added to the fact that the cell surface localization, the receptor function, and the capacity to be incorporated into liposomes suggest that VNR is an integral membrane protein, make it highly likely that this domain spans the cell membrane. Membrane insertion of the putative transmembrane domain could leave the 32 amino acids at the COOH terminus of the 25-kDa domain excluded from the membrane, if the transmembrane domain ended at the first arginine residue that follows the hydrophobic stretch. The high content of hydrophilic amino acids among this 32-amino acid domain is characteristic of cytoplasmic domains of membrane proteins, some of which are short like ours (27-29). This domain is likely to be on the cytoplasmic side of the membrane. Therefore, the structural arrangement of the  $\alpha$  subunit appears to be that the 25-kDa chain anchors the  $\alpha$  subunit of the VNR to the cell membrane and the 125-kDa chain is attached to the 25-kDa chain through a disulfide bond.

The  $\alpha$  subunits of the VNR and FNR are clearly related to one another and both appear to be related to the gpIIb subunit of the platelet adhesive protein receptor gpIIb/IIIa. The

relatedness of the VNR and FNR is apparent from the amino acid sequences of the VNR 25-kDa chain and the corresponding portion of the FNR sequence, which we have also elucidated from cDNA (25). The sequence comparison in Fig. 5 shows that the sizes of their putative membrane-spanning and cytoplasmic domains are nearly identical and there is an overall amino acid sequence homology of 34% when the two sequences are analyzed on the basis of similarity (30). If conservative replacements are considered, the overall homology approaches 48%. The gpIIb subunit is composed of two chains, and the molecular sizes of these chains are closely similar to those of the VNR  $\alpha$  subunit (31), and there is some evidence to suggest that gpIIb is made from a single mRNA and that its light chain has a transmembrane domain (32). Moreover, the VNR and gpIIb/IIIa are immunologically crossreactive (our unpublished results with M. Ginsberg and E. Plow). The structural similarities of the vitronectin, fibronectin, and platelet receptors as well as the similarity of their functions as Arg-Gly-Asp receptors (6) clearly define them as members of a family of proteins. Our results also suggest that this family may have more than these three members.

Examination of the NH<sub>2</sub>-terminal sequence of the 125-kDa chain of the VNR  $\alpha$  subunit (Fig. 3) reveals a high degree of homology with the recently published NH<sub>2</sub>-terminal sequences of Mac-1 and LFA-1 (23). This homology is especially striking with Mac-1, which shares the first four residues with VNR (Fig. 3). The homology is less pronounced for residues 6–11, but when conservative replacements are allowed, the overall degree of homology between VNR and Mac-1 is 63%. The same extent of homology exists between Mac-1 and LFA-1.

These results provide a link connecting two previously described families of adhesion proteins, the VNR, FNR, and gpIIb/IIIa family and the one consisting of LFA-1 (33), Mac-1 (34), and gp150/95 (35), which shares a subunit with the other two (36). The suggestion that these two protein families are homologous is based not only on their related NH<sub>2</sub>-terminal sequences which we report here, but also on their similar subunit structure and size and their related functions in adhesion interactions at the cell surface. Two members of this superfamily (VNR and FNR) are found in various cell types, including fibroblasts and lymphocytes (37). The other four (gpIIb/IIIa, Mac-1, LFA-1, and gp150/95) appear to occur primarily in blood cells. The existence of a connection between the Mac-1 and LFA-1 proteins and the Arg-Gly-Asp-directed receptors VNR, FNR, and gpIIb/IIIa has also been suggested recently by Cosgrove *et al.* (38), who found that a single genomic clone could lead to the expression of a protein or proteins that reacted with antibodies to Mac-1, LFA-1, and gpIIb/IIIa. It remains to be determined to what extent the ligand specificity of LFA-1, Mac-1, and gp150/95 might resemble that of the Arg-Gly-Asp-directed receptors.

We thank Khanh Nguyen for technical assistance. This work was supported by Grants CA 28896, CA 38352, and Cancer Center Support Grant CA 30199 from the National Cancer Institute.

- Pierschbacher, M. D. & Ruoslahti, E. (1984) *Nature (London)* **309**, 30–33.
- Ruoslahti, E. & Pierschbacher, M. D. (1986) *Cell* **44**, 517–518.
- Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. (1985) *Cell* **40**, 191–198.
- Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5766–5770.
- Hagen, I., Nurden, A., Bjerrum, O. J., Solum, N. O. & Caen, J. (1980) *J. Clin. Invest.* **65**, 722–731.
- Gardner, J. M. & Hynes, R. O. (1985) *Cell* **42**, 439–448.
- Pytela, R., Pierschbacher, M. D., Ginsberg, M. H., Plow, E. F. & Ruoslahti, E. (1986) *Science* **231**, 1559–1562.
- Plow, E. F., Pierschbacher, M. D., Ruoslahti, E., Marguerie, G. A. & Ginsberg, M. H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8057–8061.
- Horwitz, A., Duggan, K., Greggs, R., Decker, C. & Buck, C. (1985) *J. Cell Biol.* **101**, 2134–2144.
- Akiyama, S. K., Yamada, S. S. & Yamada, K. M. (1986) *J. Cell Biol.* **102**, 442–448.
- Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1194–1198.
- Krusius, T. & Ruoslahti, E. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7683–7687.
- Pytela, R., Pierschbacher, M. D., Argraves, W. S., Suzuki, S. & Ruoslahti, E. (1986) *Methods Enzymol.*, in press.
- Smith, D. E. & Fisher, P. A. (1984) *J. Cell Biol.* **99**, 20–28.
- Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31–40.
- Suzuki, S., Oldberg, A., Hayman, E. G., Pierschbacher, M. D. & Ruoslahti, E. (1985) *EMBO J.* **4**, 2519–2524.
- Rowe, D. W., Moen, R. C., Davidson, J. M., Byers, P. H., Bornstein, P. & Palmiter, R. D. (1978) *Biochemistry* **17**, 1581–1590.
- Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
- Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
- Springer, T. A., Teplow, D. B. & Dreyer, W. J. (1985) *Nature (London)* **314**, 540–542.
- Cai, R.-Z., Szoke, B., Lu, R., Fu, D., Redding, T. W. & Schally, A. V. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1896–1900.
- Argraves, W. S., Pytela, R., Suzuki, S., Millan, J. L., Pierschbacher, M. D. & Ruoslahti, E. (1986) *J. Biol. Chem.*, in press.
- Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982) *J. Cell Biol.* **92**, 1–22.
- Schneider, C., Owen, M. J., Banville, D. & Williams, J. G. (1984) *Nature (London)* **311**, 675–678.
- Leonard, W. J., Depper, J. M., Crabtree, G. R., Rudikoff, S., Pumphrey, J., Robb, R. J., Kronke, M., Svetlik, P. B., Peffer, N. J., Waldmann, T. A. & Greene, W. C. (1984) *Nature (London)* **311**, 626–631.
- Brown, M. S. & Goldstein, J. L. (1986) *Science* **232**, 34–47.
- Dayhoff, M. O. (1978) *Atlas of Protein Sequence & Structure* (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 1–3.
- Phillips, D. R. & Agin, P. P. (1977) *J. Biol. Chem.* **252**, 2121–2126.
- Bray, P. F., Rosa, J.-P., Lingappa, V. R., Kan, Y. W., McEver, R. P. & Shuman, M. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1480–1484.
- Davignon, D., Martz, E., Reynolds, T., Kurzinger, K. & Springer, T. A. (1981) *J. Immunol.* **127**, 590–595.
- Beller, D. I., Springer, T. A. & Schreiber, R. D. (1982) *J. Exp. Med.* **156**, 1000–1009.
- Beatty, P. G., Ledbetter, J. A., Martin, P. J., Price, T. H. & Hansen, J. A. (1983) *J. Immunol.* **131**, 2913–2918.
- Sanchez-Madrid, F., Nagy, J. A., Robbins, E., Simon, P. & Springer, T. A. (1983) *J. Exp. Med.* **158**, 1785–1803.
- Cardarelli, P. M. & Pierschbacher, M. D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2647–2651.
- Cosgrove, L. J., Sandrin, M. S., Rajasekariah, P. & McKenzie, I. F. C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 752–756.