

The Primary Structure of the VLA-2/Collagen Receptor α^2 Subunit (Platelet GPIa): Homology to Other Integrins and the Presence of a Possible Collagen-binding Domain

Yoshikazu Takada and Martin E. Hemler

Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115

Abstract. VLA-2 (also called gpIa/IIa on platelets) is a collagen receptor with a unique α subunit and a β subunit common to other adhesion receptors in the VLA/integrin family. Multiple cDNA clones for the human VLA-2 α^2 subunit have been selected from a λ gt11 library by specific antibody screening. The 5,374-bp nucleotide sequence encoded for 1,181 amino acids, including a signal peptide of 29 amino acids followed by a long extracellular domain (1,103 amino acids), a transmembrane domain, and a short cytoplasmic segment (22 amino acids). Direct sequencing of purified α^2 protein confirmed the identity of the 15 NH₂-terminal amino acids. Overall, the α^2 amino acid sequence was 18–25% similar to the sequences known for other integrin α subunits. In particular, the α^2 sequence matched other integrin α chains in (a) the positions of 17 of its 20 cysteine residues; (b) the pres-

ence of three metal-binding domains of the general structure DXDXDGXXD; and (c) the transmembrane domain sequence. In addition, the α^2 sequence has a 191-amino acid insert (called the I-domain), previously found only in leukocyte integrins of the β_2 integrin family. The α^2 I-domain was 23–41% similar to domains in cartilage matrix protein and von Willebrand factor, which are perhaps associated with collagen binding. The NH₂-terminal sequence reported here for α^2 does not match the previously reported α^2 NH₂-terminal sequence (Takada, Y., J. L. Strominger, and M. E. Hemler. 1987. *Proc. Natl. Acad. Sci. USA*. 84:3239–3243). Resolution of this discrepancy suggests that there may be another VLA heterodimer that resembles VLA-2 in size but has a different amino acid sequence.

VLA-2 is an α/β -subunit cell surface heterodimer strongly implicated as receptor for collagen because (a) VLA-2 has been shown to be identical (61) to a 150,000/110,000-*M_r* structure recognized by the mAb PIH5, which specifically blocks human fibrosarcoma cell (66) and platelet (31) attachment to collagen; (b) patients deficient in platelet protein Ia (α subunit of VLA-2) also lacked responsiveness to collagen (27, 41); and (c) the mAb 12F1 was used to identify VLA-2 as a 160,000/130,000-*M_r* (nonreduced) platelet protein complex that mediates Mg²⁺-dependent adhesion to collagen (51, 52). Also in this regard, antigens that strongly resemble VLA-2 in size have been implicated in hepatocyte cell attachment to type I collagen (14).

The cell surface heterodimer VLA-2 was initially characterized as a "very late antigen" appearing on activated T cells (19, 20). Later the mAb 12F1 was used to identify VLA-2 on platelets and on many other cell types, including most rapidly growing, adherent cell lines (43). Major portions of the platelet gpIa and IIa are probably the same as the VLA-2

α and β subunits, respectively (31, 44). Thus the new platelet-specific alloantigens Br^a and Br^b, which reside on the platelet Ia-IIa complex (53), may be variable epitopes on VLA-2.

In other studies, an antibody (called 5E8) that recognizes VLA-2 (Hemler, M., and C. Crouse, unpublished results) has been described on most primary human lung tumors (68), and ¹²⁵I-coupled 5E8 antibody has been used to inhibit the growth of lung tumor cell lines in vitro (57).

The expression of VLA-2 can be up regulated on lymphocytes in response to mitogen or antigen stimulation (17) and on fibroblasts in response to serum (10). Conversely, VLA-2 expression slowly diminished when those stimuli were withdrawn and/or cells became quiescent (10, 17). Also, VLA-2 expression can be up regulated by TGF β (15).

Like other cell surface receptors for extracellular matrix components, VLA-2 belongs to the integrin superfamily (26, 49). Three subfamilies of that superfamily that are defined for humans are (a) the six (or more) VLA proteins (16, 23); (b) the leukocyte adhesion molecules, LFA-1, Mac-1, and p150,95 (56); and (c) the cytoadhesins, which include the vitronectin receptor and platelet gpIIb/IIIa (13). Within each

Y. Takada is currently at Biochemistry Division, National Cancer Center Research Institute, 1-1, Tsukiji, 5-chome, Chuo-ku Tokyo, 104 Japan.

family, the member heterodimers are each composed of a unique α subunit associated with a common β subunit. Analyses of human cDNA clones for the β subunits of each subfamily (called β_1 , β_2 , and β_3 , respectively) have revealed that they are 44–47% similar (2, 12, 28, 33, 62), suggesting a common evolutionary origin. The evolutionary conservation of β_1 is emphasized by the homology (82–86%) maintained between β_1 s from widely diverging species, such as human, chicken, and frog (9), and 45% homology between human β_1 and a β_1 -like structure from *Drosophila* (36). Also, complete sequencing of several integrin α subunits—including those from a fibronectin receptor (α^5) (2, 11), the vitronectin receptor (58), platelet gpIIb/IIIa (11, 45), Mac-1 (3, 8, 46), p150,95 (7), and the *Drosophila* PS2 antigen (6)—have revealed 20–60% similarity between any pair. Some of these integrin structures are receptors for ligands containing the amino acid signal sequence Arg-Gly-Asp (RGD) or a closely related sequence (49). Although NH₂-terminal amino acid sequencing has suggested ~40% shared residues among six different VLA α subunits (24, 60), complete sequence information is available for only one VLA α subunit (α^5) (2, 11). Thus, to further establish patterns of similarity and differences among VLA/integrin sequences and to gain basic information for future studies of structure, function, and regulation, the complete sequence of the VLA α^2 subunit was obtained. Also, because the α^2 -subunit NH₂-terminal sequence predicted from cDNA did not match the previously published “ α^2 ” NH₂-terminal sequence, experiments were carried out to resolve this discrepancy.

Materials and Methods

Purification of VLA-2 from Platelets and NH₂-terminal Sequencing

VLA-2 protein was purified from outdated platelets (from the Dana-Farber Cancer Institute [Boston, MA] blood bank or from the American Red Cross [Dedham, MA]) by lectin-Sepharose and then by 12F1 or A-1A5 immunoaffinity chromatography (60). The anti- α^2 mAb 12F1 (43) and the anti- β_1 mAb A-1A5 (18) were obtained as described. For NH₂-terminal amino acid microsequencing of the α^2 subunit, the α^2 subunit was first separated from the β subunit by SDS-gel electrophoresis, transferred to a polyvinylidene difluoride membrane (Immobilon; Millipore Continental Water Systems, Bedford, MA), stained with Coomassie blue, and destained as described (39). Excised membrane pieces containing α^2 material were then sequenced using a gas-phase sequencer (470A; Applied Biosystem, Inc., Foster City, CA) equipped with a phenylhydantoin amino acid analyzer (120A; Harvard Microsequencing Facility, Cambridge, MA).

Preparation of α^2 -Subunit-specific Antiserum

Subunits migrating in the α^2 position (~50 μ g from A-1A5-Sepharose or 12F1-Sepharose) were reduced and separated using preparative 5% SDS-PAGE, stained with Coomassie blue, and destained in 7% acetic acid and 30% methanol. The band corresponding to α^2 peptide was cut out and electroeluted into Tris/glycine/SDS buffer using an Elutrap (Schleicher & Schuell, Inc., Keene, NH). The eluted peptide was used for rabbit immunization as previously described (23). Also, α^2 protein purified from 12F1-Sepharose was used for immobilization onto CNBr-Sepharose (Pharmacia Fine Chemicals, Piscataway, NJ) according to the manufacturer's instructions, except that 0.1% of SDS was included in the coupling buffer. Antibodies from rabbit anti- α^2 -subunit antiserum were then purified using the α^2 -Sepharose affinity column. Briefly, the antiserum was passed through the column several times and the column was washed successively with 3 column volumes each of PBST (0.14 M NaCl, 10 mM sodium phosphate, 0.2% Tween 20), 0.2 M KSCN, and then PBST again. For elution of α^2 -subunit-specific IgG, 3 M KSCN was added, and the eluate was desalted by Sephadex G-25 chromatography into PBST in the presence of 0.1% hemoglobin as carrier. Immunopurified IgG was further incubated with dena-

tured *Escherichia coli* protein coupled to Sepharose, and with glycoproteins from Molt-4 cells coupled to Sepharose to remove additional nonspecific reactivity. The Molt-4 leukemic T cell line does not usually express VLA-2 (22).

Production of Antibodies to an α^2 COOH-terminal Synthetic Peptide

A 22-mer consisting of the COOH-terminal 21 amino acids predicted from the α^2 sequence, with an added cysteine at the NH₂-terminal end of the peptide, was synthesized by Multiple Peptide Systems (San Diego, CA). This peptide was coupled to carrier protein (keyhole limpet hemocyanin) using m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester (Pierce Chemical Co., Rockford, IL) as previously described (29), and then rabbit antibodies were generated using standard techniques (38).

Isolation of the Gene Encoding the α^2 Subunit

A phage λ gt11 expression cDNA library of human lung fibroblast IMR-90 (Clontech Laboratories, Inc., Palo Alto, CA) was screened by using affinity-purified rabbit antibodies against α^2 subunit according to Young and Davis (67), except that alkaline phosphatase-conjugated anti-rabbit IgG was used as a second antibody. Alkaline phosphatase activity was detected by using an immunoscreening system (ProtoBlot; Promega Biotec, Madison, WI) with nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as substrates. Positive clones from the screening were plaque purified, and then phage DNAs were purified by the plate lysate method (37). After the insert was excised from the phage by Eco RI restriction enzyme digestion, the inserts were separated on agarose gels, electroeluted, and subcloned into pBluescript plasmid (Stratagene, La Jolla, CA). A cDNA library made in λ gt11 from endothelial cells was the kind gift from Dr. Tucker Collins (Brigham and Women's Hospital, Boston, MA).

DNA Sequencing

The DNA sequences were determined by the dideoxynucleotide chain-termination method of Sanger et al. (50) using adenosine ³⁵S-5'-[α -thio]triphosphate. To facilitate complete sequencing of both cDNA strands, a series of overlapping deletion clones were made in both directions as described (25) by using the Erase-a-base system (Promega Biotec).

Other Techniques

Southern blot analysis was carried out as described (37). For Northern blot analysis, 5 μ g of each RNA sample was electrophoretically separated on 1% agarose gels in buffer containing formaldehyde, transferred to nylon membrane (Nytran; Schleicher & Schuell) (37), and then probed with a nick-translated 5.4-kb α^2 -subunit cDNA (in pBluescript) using standard hybridization techniques (37). Western blotting (63) and immunoprecipitation (22) were performed as previously described.

Results and Discussion

Cloning of cDNA with α^2 -Subunit-specific Antibody as Probe

Affinity-purified rabbit anti- α^2 antiserum was tested by immunoblotting of whole platelet lysate and found to bind specifically to platelet VLA α^2 subunit (Fig. 1), thus affirming the suitability of the serum for library screening. The affinity-purified antibody was then used for immunoselection of clones from a phage λ gt11 expression cDNA library made from human lung fibroblast (IMR-90) since fibroblasts were known to be good source for α^2 protein (21). After subcloning into plasmid, a representative positive clone (clone 2.38, ~1.2 kb) was directly sequenced and found to have partial homology to known sequences of other α subunits in the integrin superfamily. Subsequently, clone 2.38 was used to probe a human endothelial cell λ gt11 library (another good source for α^2 expression) since that library had longer inserts. Thus, another clone was selected (clone 2.72, ~5.4 kb), completely sequenced, and found to include all of the coding region and

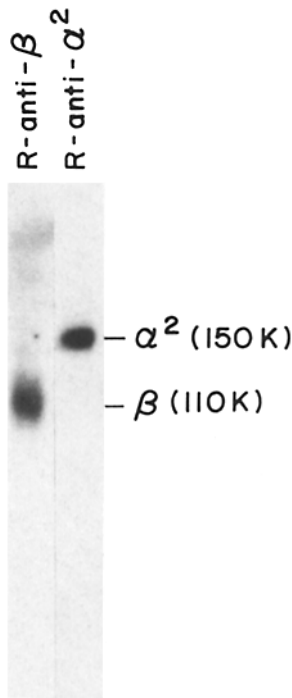


Figure 1. Immunoblotting analysis of rabbit antibody after enrichment for α^2 specificity. After SDS-PAGE separation of purified platelet VLA proteins, the proteins were blotted onto nitrocellulose and then incubated with rabbit anti- β_1 (lane *a*) or anti- α^2 purified (see Materials and Methods) rabbit antiserum (lane *b*). Bound primary antibody was located using ^{125}I -labeled goat anti-rabbit second antibody.

some of the 3' and 5' untranslated regions (Fig. 2). Notably, the sequence of clone 2.38 (from a fibroblast library) was identical to the corresponding region in clone 2.72 (from an endothelial cell library).

Comparison of Clone 2.72 cDNA Sequence with α^2 -Subunit NH₂-terminal Protein Sequence

Surprisingly, the previously published α^2 NH₂-terminal sequence FNLDTXEDNVFRGP (Fig. 3 *A*) only partly resembled the α^2 NH₂-terminal sequence YNVGLPEAKIFSGPS encoded by cDNA (Fig. 3 *C*), with homology in only 5 of the 14 positions. To resolve this discrepancy, additional NH₂-terminal amino acid sequencing was carried out. Using mAb 12F1-Sepharose, VLA-2 material was purified either from platelets or placenta. After preparative SDS-PAGE, highly purified α^2 subunit from either source yielded an NH₂-terminal amino acid sequence of YNVGLPEAKIFSGPS (Fig. 3 *D*), identical to the sequence encoded from cDNA (Fig. 3 *C*). In another experiment, VLA proteins were purified from 25 g of platelets using the anti- β mAb A-1A5, and the resulting mixture of subunits was then separated by reducing SDS-PAGE so that α^5 and α^6 subunits would migrate at 130,000 *M_r*, away from the 160,000-*M_r* α^2 subunit (23). Upon NH₂-terminal amino acid sequencing of the purified β_1 -associated 160,000-*M_r* α^2 -like protein from the reduced gel, a mixture of the YNVGLPEAKIFS and FNLTDXEDNVFR sequences was present, with approximately equivalent molar ratios (Fig. 3 *B*). Thus, it appears that the latter sequence may belong to a protein of 160,000 *M_r* that is distinct from α^2 , though it closely resembles α^2 in size. At present, nothing else is known about this potentially new subunit, other than it appears to coprecipitate with β_1 (when using the anti- β_1 mAb A-1A5). Because the mAb 12F1 (originally used to define VLA-2 [43]) was used to isolate the protein with the YNVGLPEAKIFS... sequence, that protein merits the designation as the authentic α^2 . It remains to be seen if the re-

cently described multiple forms of platelet gpIa (5) bear any relation to the two distinct NH₂-terminal sequences described here.

Northern Blotting Analysis

The distribution of mRNA for the α^2 subunit was studied by Northern blotting with cDNA clone 2.72 as the probe. A single band at ~8 kb was detected in fibroblast RNA, whereas little or no signal at the same position was obtained in HSB or Molt-4 T cell lines (Fig. 4). These results are consistent with the known cell surface expression of VLA-2 on these cells (21, 22). Hybridization of the same blot with a cDNA probe for the human actin gene gave comparable signals in all lanes. Since the RNA size (8 kb) was somewhat larger than the cDNA clone (5.4 kb) and since no poly A tail is present (Fig. 2), it is assumed that the 3' untranslated region (and perhaps also the 5' end) is incomplete.

Authenticity of the α^2 Clone

Despite the unexpected NH₂-terminal sequence results, the α^2 clone obtained is authentic because (*a*) the cDNA was selected using antibodies specific for the α^2 subunit; (*b*) the NH₂ terminus of the predicted sequence was identical to that directly determined from purified α^2 -subunit protein; (*c*) the mRNA distribution agrees with that of the mature protein; (*d*) the size of the predicted protein closely agrees with the experimental value (see below); (*e*) the predicted sequence has several features characteristic of other integrin α subunits (see below); and (*f*) native α^2 protein could be readily immunoprecipitated by antisera prepared against a synthetic COOH-terminal peptide including 21 amino acids deduced from the α^2 cDNA sequence (Fig. 5).

As shown in Fig. 5 *A*, rabbit serum directed against an α^2 COOH-terminal synthetic peptide (predicted from cDNA sequence) recognized an $\alpha\beta$ complex (lane *c*) from the colon carcinoma cell line CCL-228 that closely resembled VLA-2 (lane *b*). Prior removal of all material seen by the rabbit serum (lane *f*) resulted in the simultaneous removal of VLA-2 (lane *e*), but only partially diminished total VLA as seen by mAb A-1A5 (cf. lanes *a* and *d*). Also, anti- α^2 COOH-terminal peptide sera specifically immunoblotted α^2 purified from platelets using either the mAb A-1A5 (Fig. 5 *B*, lane *a*) or 12F1 (lane *b*). In comparison, serum from a rabbit immunized with isolated α^2 protein (before anti- α^2 antibodies were purified) recognized mostly α^2 (but also β_1) in preparations obtained from platelets using A-1A5 (lane *c*) or 12F1 (lane *d*), whereas a negative control serum did not blot any detectable purified platelet protein (lanes *e* and *f*).

Amino Acid Sequence of the α^2 Subunit

Translation of the cDNA sequence of clone 2.72 (Fig. 2) yielded coding sequences of 3,543 bp (encoding 1,181 amino acids) between the 3' and 5' untranslated portions of ~50 bp and 1,800 bp, respectively. Preceding the NH₂-terminal sequence of the mature protein (YNVGLPE...) is a translational start site (30) that codes for a methionine followed by 29 amino acids, fulfilling the requirements for a signal peptide (64). Thus, the open reading frame shown in Fig. 2 encodes a mature protein of 1,152 amino acids predicted to be

| Residue No.: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--------------------------|-----|---|-----|-----|---|---|---|-----|-----|-----|----|----|----|----|----|
| A α^2 (published) | F | N | L | D | T | X | E | D | N | V | F | R | G | P | - |
| B $\alpha^2 + \alpha^2$ | F/Y | N | L/V | D/G | L | P | E | D/A | N/K | I/V | F | /S | - | - | - |
| C α^2 (cDNA) | Y | N | V | G | L | P | E | A | K | I | F | S | G | P | S |
| D α^2 (protein) | Y | N | V | G | L | P | E | A | K | I | F | S | G | P | S |

Figure 3. Comparison of VLA α^2 and other NH₂-terminal sequences. (A) The NH₂-terminal sequence of platelet material previously purified using A-1A5-Sephacel and erroneously published as the α^2 sequence (60); (B) the NH₂-terminal sequence from another preparation of platelet material comigrating with α^2 on SDS-PAGE after A-1A5-Sephacel purification of VLA proteins; (C) the α^2 NH₂-terminal sequence deduced from cDNA; and (D) the α^2 NH₂-terminal sequence obtained after purification of VLA α^2 (from placenta or platelets) using the mAb 12F1 (which is specific for α^2).

126,000 *M_r*. The addition of 10 potential *N*-glycosylation sequences (Asn-Xaa-Ser/Thr, average 2,500 *M_r*) to the core protein would result in an estimated molecular mass of 151,000 D. That value is close to the 155,000–160,000 *M_r* estimated from α^2 -subunit migration on SDS-polyacrylamide gels.

Analysis of the NH₂-terminal portion of the sequence revealed the presence of seven homologous repeating domains (Fig. 6). These domains contain 28–41 amino acids and are 20–30% similar to each other. Also, these domains are spaced 23–32 amino acids apart, except that there is a large additionally inserted "I-domain" of 191 amino acids (see below) between repeating domains II and III.

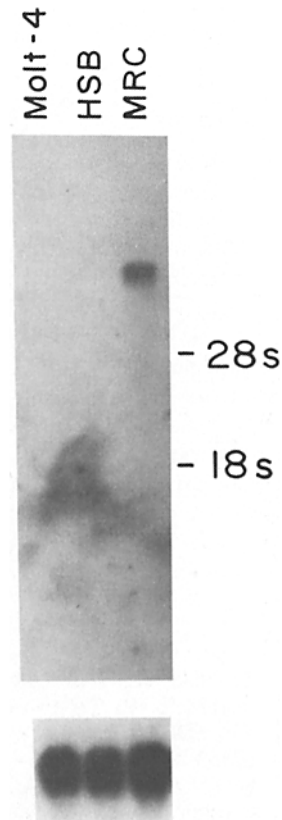


Figure 4. RNA hybridization analysis. Total RNA (5 μ g/lane) from the leukemic cell lines Molt-4 and HSB and from the fibroblast cell line MRC was electrophoretically separated on 1% agarose-formaldehyde gels, transferred to nitrocellulose filters, and probed with full-length α^2 cDNA. In a control experiment (lower panel), the same RNA samples were probed with actin cDNA.

The repeated domains V, VI, and VII each contain sequences of Dx(D/N)xD(G)xxD (Fig. 2, *small boxes*, and Fig. 6, *underline*) that are somewhat similar to the EF-hand consensus metal-binding domains of a number of calcium- and magnesium-binding proteins, including calmodulin, troponin C, parvalbumin (59), thrombospondin (34), myosin light chain (47), and galactose-binding protein (65). These potential divalent cation-binding sites in α^2 are located in a region (between amino acids 470 and 627) devoid of cysteine residues and *N*-glycosylation sites. The presence of divalent cation-binding sites in the α^2 subunit is consistent with divalent cation requirements for the function of VLA-2 as a human collagen receptor (51, 52).

Comparison of α^2 -Subunit Sequence with Other Integrin α Chains

The alignment of the α^2 -subunit sequence with the α -chain sequences of human fibronectin receptor (VLA-5), vitronectin receptor, gpIIb/IIIa, Mac-1, and p150,95 shows that several structural characteristics are shared (Fig. 7). For example, of the 20 α^2 -cysteine residues, 17 are conserved in at least three of the other sequences and 14 are conserved in all six α -subunit sequences. Also, there is 32–45% conservation in the 22-amino acid transmembrane region and 100% conservation of the GFFKR sequence on the cytoplasmic side of transmembrane domain. Additional striking similarities are evident in the region of the α^2 homologous repeats in the NH₂-terminal half of the molecule. Like α^2 , each of the other integrin α subunits has also been noted to have seven homologous repeats, with three or four potential divalent cation sites within repeats IV–VII (2, 3, 7, 8, 11, 45, 46, 58). In these repeat regions, the average similarity between α^2 and the other α subunits ranged from 30–38% (repeats I, II, III, and VII) to 46–52% (repeats IV, V, and VI). The overall similarity between α^2 and the other integrin α subunits is 18–25%, or 22–24% if the I-domain (see below) is excluded. This is in contrast to the higher degree of identity (~45%) between different human integrin β chains (2, 12, 28, 33, 62).

Comparison of the α^2 -Subunit I-Domain with Similar Domains in Other Proteins

In studies of Mac-1 (3, 8, 46), p150,95 (7), and LFA-1 (32), an inserted domain of ~200 amino acids was found that is not present in the α^5 , α^V , or IIb chains. In those studies, the inserted I-domain (32) (formerly called the L-domain) was noted to resemble the von Willebrand factor (vWF)¹ A1, A2, and A3 domains, cartilage matrix protein (CMP)-1 and CMP-2 domains, and domains in complement factors B and C2. Now an I-domain has also been discovered in the α^2 sequence. The sequence of the α^2 I-domain is aligned with other related domains as shown in Fig. 8, and percent similarity calculations from those data were used in the construction of the linkage tree shown in Fig. 9. The α^2 I-domain sequence (allowing for conservative amino acid substitutions) most resembled I-domains from p150 (48%) and Mac-1 (45%) and also resembled the CMP-1 (46%) and CMP-2 (44%) domains. Less similarity was seen with vWF domains

1. Abbreviations used in this paper: CMP, cartilage matrix protein; vWF, von Willebrand factor.

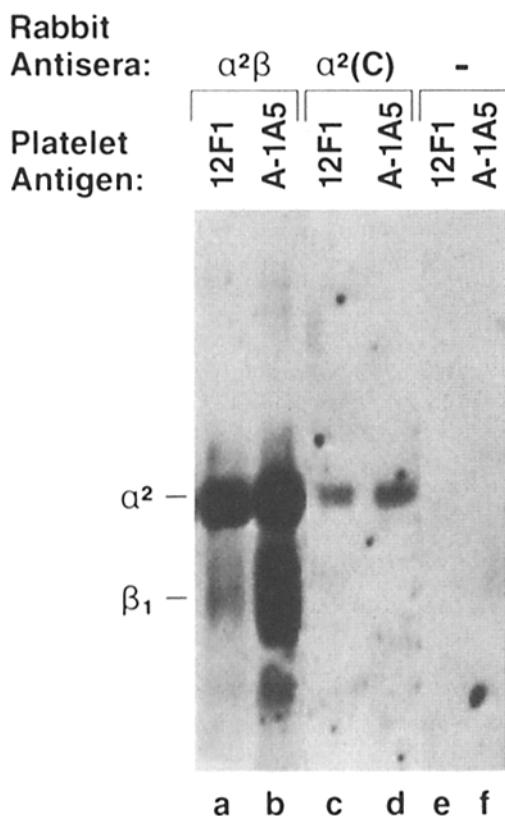
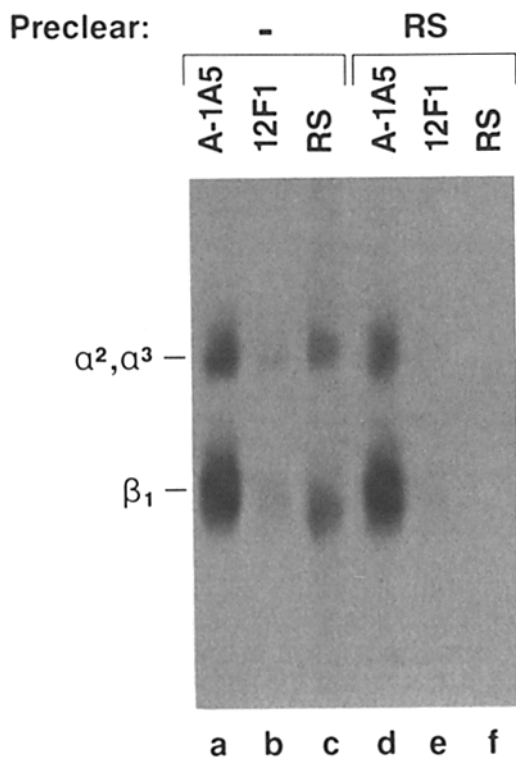


Figure 5. Recognition of α^2 protein by rabbit serum against α^2 COOH-terminal synthetic peptide. (A) Immunoprecipitation of VLA proteins from ^{125}I -radiolabeled CCL-228 cells (a colon carcinoma line) was carried out using the mAb A-1A5 (lanes a and d), 12F1 (lanes b and e), and rabbit anti- α^2 COOH-terminal peptide (RS, lanes c and f) either with (lanes d-f) or without (lanes a-c) prior immunodepletion of all rabbit serum-reactive material. (B) The

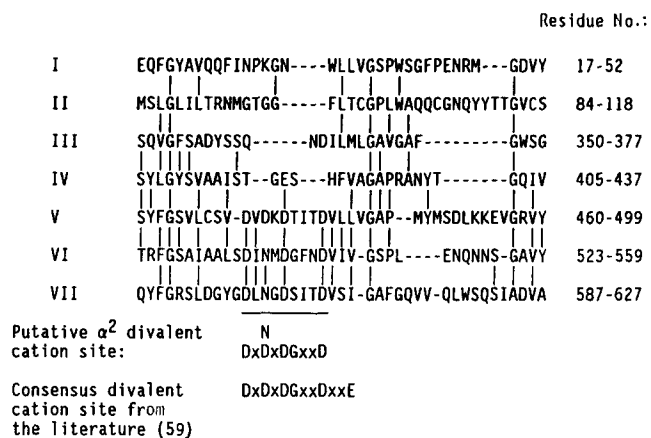


Figure 6. Alignment of seven homologous repeated domains in the α^2 subunit. Residues that are the same in adjacent sequences (including conservative substitutions) are marked with vertical connecting lines. Conservative substitutions are I, V, L; Y, W, F; A, G; S, T; R, K; D, E; N, Q.

A1 (32%), A2 (36%), and A3 (37%), and with complement factors B (29%) and C2 (26%).

The vWF protein contains at least two independent collagen-binding sites, localized to the A1 and A3 domains (42, 48) and, likewise, CMP-1 and CMP-2 domains may be used to mediate interaction with collagen (1). Because VLA-2, vWF, and CMP all share related domains and can interact with collagen, that suggests the I-domain may be important for adhesion to collagen. However, the presence of an I-domain (or related sequence) does not completely correlate with a collagen adhesion function since there is no evidence that p150,95 or Mac-1 interact with collagen and, likewise, complement factors B and C2 are not known to bind collagen.

Grouping of Integrin α Subunits into Subsets

From sequence analysis, together with earlier data, it is clear that there are several structural features that distinguish integrin α subunits containing I-domains from those that do not (Table I). It has previously been noted that the mature forms of integrin α subunits α^5 , α^V , and IIb each contain a cleaved peptide fragment (the COOH-terminal 15% of the protein) that is attached to the rest of the subunit by a disulfide linkage. Consistent with this, dibasic protease cleavage sites have been noted in the appropriate locations in each of those sequences (2, 35, 58). Notably (Table I), the subunits with I-domains (α^2 , α^M , and p150) do not undergo protease cleavage, whereas the cleaved subunits do not have I-domains. In this regard, in the region where protease cleavage occurs for α^5 , α^V , and IIb, the three I-domain subunits have gaps of 20 or more amino acids and no dibasic amino acids (see Fig. 7, α^2 position 1,018).

mAbs 12F1 and A-1A5 were used to purify either VLA-2 or total VLA protein from platelets (see Materials and Methods). Then aliquots of purified proteins were separated by SDS-PAGE, blotted onto nitrocellulose, and probed as described (63) using rabbit anti-intact α^2 plus β_1 (lanes a and b), anti-COOH-terminal synthetic peptide (lanes c and d), or control rabbit serum (lanes e and f).

Table I. Structural Features Distinguishing Integrin α Subunits with I-Domains from α Subunits That Are Cleaved

| Integrin α subunit | I-Domain | Protease cleaved | Divalent cation sites | Residues at α_2 positions | | | |
|---------------------------|----------|------------------|-----------------------|----------------------------------|--------------------|--------------------|--------------------|
| | | | | 110 | 467 | 350 | 645 |
| α^2 | Yes | No | 3 | Cys ₁₁₀ | Cys ₄₆₇ | Ser ₃₅₀ | — ₆₄₅ |
| α^M | Yes | No | 3 | Cys ₉₇ | Cys ₄₄₆ | Ser ₃₃₅ | Gly ₅₃₄ |
| p150 | Yes | No | 3 | Cys ₉₇ | Cys ₄₄₄ | Ala ₃₃₃ | Val ₅₃₁ |
| α^5 | No | Yes | 4 | Lys ₁₂₅ | Ala ₂₈₉ | Cys ₁₆₄ | Cys ₄₈₁ |
| IIb | No | Yes | 4 | Glu ₁₁₇ | Ala ₂₉₅ | Cys ₁₆₇ | Cys ₄₈₄ |
| α^V | No | Yes | 4 | Met ₁₁₂ | Ala ₂₈₀ | Cys ₁₅₅ | Cys ₄₇₂ |
| PS2 α | No | Yes | 4 | Thr ₁₅₇ | Ala ₃₃₇ | Cys ₁₉₈ | Cys ₅₃₆ |

cysteines at positions 110 and 467 in the α^2 sequence are conserved only among α^2 , α^M , and p150. Conversely, at α^2 positions 350 and 645, cysteines are absent from α^2 , α^M , and p150, but present in all of the cleaved subunits. Thus, it might be predicted that each group of subunits (cleaved or with I-domains) may have a characteristic intrachain disulfide loop that distinguishes one group from the other. Although the newly derived sequence for the α^L chain of LFA-1 (32) is not shown here, that sequence has an I-domain and fits the pattern shown in Table I with respect to potential divalent cation sites and conserved cysteine residues. Structural information for the *Drosophila* PS2 α subunit (6) is also included in Table I. Although the source is phylogenically far removed from humans, the properties of that α subunit are fully consistent with the properties of the other cleaved α subunits.

In addition to the "group-specific" cysteine residues mentioned in Table I, there are many other amino acids at posi-

Table II. Comparison of Subsets of α Subunits That Show Amino Acid Identity at the Most Positions*

| Subset of α subunits | Positions with identity (No.) | Common features |
|--------------------------------------|-------------------------------|------------------|
| Part A | | |
| α^5 , IIb, α^V | 99 | Cleaved subunits |
| α^2 , p150, α^M | 64 | I-Domain |
| α^5 , α^2 , α^V | 24 | — |
| IIb, α^M , p150 | 21 | — |
| α^5 , p150, α^M | 19 | — |
| α^V , α^M , p150 | 19 | — |
| α^2 , α^5 , IIb | 13 | — |
| α^2 , IIb, α^V | 13 | — |
| Part B | | |
| p150, α^M | 256 | β_2 |
| α^5 , α^V | 109 | — |
| IIb, α^V | 65 | β_3 |
| α^5 , IIb | 62 | — |
| α^2 , IIb | 40 | — |
| α^2 , α^V | 35 | — |
| α^2 , α^5 | 31 | β_1 |

* Sets of sequences, shown in order of prevalence, were derived from computer-generated lists of positions in which only three (part A) or only two (part B) identical amino acids are shared. Sequence information from I-domains (residues 159–349 in α^2) was omitted for these determinations.

tions throughout the aligned sequences that distinguish I-domain integrins from cleaved integrins. In fact, when the aligned sequences (Fig. 7) were searched by computer to find all positions in which exactly three sequences had identical amino acids, this was found to occur most often (at 99 positions) among the set α^5 , IIb, and α^V and also occurred often (at 64 positions) among the set α^2 , p150, and α^M (Table II, part A). Thus, the I-domain subunits could again be distinguished from the cleaved subunits, this time based on exclusively shared amino acids distributed throughout the coding region. Other sets of three subunits shared amino acids less frequently (at ≤ 24 positions) and had no obvious structural similarities. Thus the results from the data in Table II, part A, support the conclusions drawn from Table I.

Because the α^2 and α^5 subunits both associate with the same β subunit (β_1), it might be assumed that these two sequences would show exclusive amino acid identity at several positions. However, when a computer search was performed (Table II, part B), positions with amino acid identity between only α^2 and α^5 were not very prevalent (seventh in the list of subunit pairs). In fact, α^2 was more likely to share residues with IIb or α^V than with α^5 , and α^5 was more likely to share residues with α^V or IIb than with α^2 . Thus, despite their common β -subunit association, the α^2 and α^5 subunits have relatively few amino acids common to only those sequences, making it difficult to predict potential α -subunit sites that may be specific for β_1 interaction.

Summary

Cloning and sequencing of the α^2 subunit of VLA-2 has revealed its complete primary structure, and confirmed its relatedness to other integrin α subunits. The presence of an α^2 I-domain, which possibly participates in VLA-2 adhesion to collagen has been established, and comparative studies of the α^2 sequence have revealed that subunits containing I-domains have a number of additional structural features that distinguish them from cleaved integrin α subunits. Also, it is clear that β -subunit use does not accurately predict the relative degree of similarity between integrin α subunits. For future studies, the results shown here will provide the basic information and materials needed for (a) additional structural comparisons with other integrins; (b) analysis of VLA-2 function through expression and mutagenesis; and (c) analysis of patient samples to measure deficient or mutated α^2 .

We thank Carol Crouse for assistance with immunoprecipitations and immunoblotting and Donald Faulkner and Temple Smith (Molecular Biology Computer Research Resource, Dana-Farber Cancer Institute, Boston, MA) for assistance with sequence comparisons and alignments.

This work was supported by National Institutes of Health grant GM 38903.

Received for publication 14 December 1988 and in revised form 6 March 1989.

References

1. Argraves, W. S., F. Deak, K. J. Sparks, I. Kiss, and P. F. Goetinck. 1987. Structural features of cartilage matrix protein deduced from cDNA. *Proc. Natl. Acad. Sci. USA.* 84:464–468.
2. Argraves, W. S., S. Suzuki, H. Arai, K. Thompson, M. D. Pierschbacher, and E. Ruoslahti. 1987. Amino acid sequence of the human fibronectin receptor. *J. Cell Biol.* 105:1183–1190.
3. Arnaout, M. A., S. K. Gupta, M. W. Pierce, and D. G. Tenen. 1988. Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3). *J. Cell Biol.* 106:2153–2158.

4. Bentley, D. R. 1986. Primary structure of human complement component C2: homology to two unrelated protein families. *Biochem. J.* 239:339-345.
5. Bienz, D., and K. J. Clemetson. 1989. Human platelet glycoprotein Ia: one component is only expressed on the surface of activated platelets and may be a granule constituent. *J. Biol. Chem.* 264:504-514.
6. Bogaert, T., N. Brown, and M. Wilcox. 1988. The *Drosophila* PS2 antigen is an invertebrate integrin which like the fibronectin receptor becomes localized to muscle attachments. *Cell.* 51:929-940.
7. Corbi, A. L., L. J. Miller, K. O'Connor, R. S. Larson, and T. A. Springer. 1987. cDNA cloning and complete primary structure of the alpha subunit of a leukocyte adhesion glycoprotein, p150,95. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:4023-4028.
8. Corbi, A. L., T. K. Kishimoto, L. J. Miller, and T. A. Springer. 1988. The human leukocyte adhesion glycoprotein Mac-1 (Complement receptor type 3, CD11b) α subunit: cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B. *J. Biol. Chem.* 263:12403-12411.
9. DeSimone, D. W., and R. O. Hynes. 1988. *Xenopus laevis* integrins. *J. Biol. Chem.* 263:5333-5340.
10. Fingermaier, E., and M. E. Hemler. 1988. Regulation of proteins in the VLA cell substrate adhesion family: influence of cell growth conditions on VLA-1, VLA-2 and VLA-3 expression. *Exp. Cell Res.* 177:132-142.
11. Fitzgerald, L. A., M. Poncz, B. Steiner, S. C. Rall, Jr., J. S. Bennett, and D. R. Phillips. 1987. Comparison of cDNA-derived protein sequences of the human fibronectin and vitronectin receptor alpha-subunits and platelet glycoprotein IIb. *Biochemistry.* 26:8158-8165.
12. Fitzgerald, L. A., B. Steiner, S. C. Rall, Jr., S. Lo, and D. R. Phillips. 1987. Protein sequence of endothelial glycoprotein IIIa derived from a cDNA clone: identity with platelet glycoprotein IIIa and similarity to "integrin." *J. Biol. Chem.* 262:3936-3939.
13. Ginsberg, M. H., J. C. Loftus, and E. F. Plow. 1988. Cytoadhesins, integrins and platelets. *Thromb. Hemostasis.* 59:1-6.
14. Gullberg, D., L. Terracio, and K. Rubin. 1988. Membrane glycoproteins involved in hepatocyte adhesion to collagen type I. *Exp. Cell Res.* 175:388-395.
15. Heino, J., R. A. Ignatz, M. E. Hemler, C. Crouse, and J. Massagué. 1989. Regulation of cell adhesion receptors by TGF- β : concomitant regulation of integrins that share a common B1 subunit. *J. Biol. Chem.* 264:380-388.
16. Hemler, M. E. 1988. Adhesive protein receptors on hematopoietic cells. *Immunol. Today.* 41:109-113.
17. Hemler, M. E., and J. G. Jacobson. 1987. Cell matrix adhesion-related proteins VLA-1 and VLA-2: regulation of expression on T cells. *J. Immunol.* 138:2941-2948.
18. Hemler, M. E., C. F. Ware, and J. L. Strominger. 1983. Characterization of a novel differentiation antigen complex recognized by a monoclonal antibody (A-1A5): unique activation-specific molecular forms on stimulated T cells. *J. Immunol.* 131:334-340.
19. Hemler, M. E., J. G. Jacobson, and J. L. Strominger. 1985. Biochemical characterization of VLA-1 and VLA-2: cell surface heterodimers on activated T cells. *J. Biol. Chem.* 260:15246-15252.
20. Hemler, M. E., J. G. Jacobson, M. B. Brenner, D. Mann, and J. L. Strominger. 1985. VLA-1: a T cell surface antigen which defines a novel late stage of human T cell activation. *Eur. J. Immunol.* 15:502-508.
21. Hemler, M. E., C. Huang, Y. Takada, L. Schwarz, J. L. Strominger, and M. L. Clabby. 1987. Characterization of the cell surface heterodimer VLA-4 and related peptides. *J. Biol. Chem.* 262:11478-11485.
22. Hemler, M. E., C. Huang, and L. Schwarz. 1987. The VLA protein family: characterization of five distinct cell surface heterodimers each with a common 130,000 molecular weight beta subunit. *J. Biol. Chem.* 262:3300-3309.
23. Hemler, M. E., C. Crouse, Y. Takada, and A. Sonnenberg. 1988. Multiple very late antigen (VLA) heterodimers on platelets: evidence for distinct VLA-2, VLA-5 (fibronectin receptor) and VLA-6 structures. *J. Biol. Chem.* 263:7660-7665.
24. Hemler, M. E., C. Crouse, and A. Sonnenberg. 1989. Association of the VLA alpha-6 subunit with a novel protein: a possible alternative to the common VLA beta-1 subunit on certain cell lines. *J. Biol. Chem.* 264:6529-6535.
25. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene (Amst.)* 28:351-359.
26. Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48:549-554.
27. Kehrel, B., L. Balleisen, R. Kokott, R. Mesters, W. Stenzinger, K. J. Clemetson, and J. VanDeLoo. 1989. Deficiency of intact thrombospondin and membrane glycoprotein Ia in platelets with defective collagen-induced aggregation and spontaneous loss of disorder. *Blood.* 71:1074-1078.
28. Kishimoto, T. K., K. O'Connor, A. Lee, T. M. Roberts, and T. A. Springer. 1987. Cloning of the beta subunit of the leukocyte adhesion proteins: homology to an extracellular matrix receptor defines a novel supergene family. *Cell.* 48:681-690.
29. Kitagawa, T., and T. Aikawa. 1976. Enzyme coupled immunoassay of insulin using a novel coupling reagent. *J. Biochem. (Tokyo).* 79:233-236.
30. Kozak, M. 1987. An analysis of 5'-noncoding sequence from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125-8132.
31. Kunicki, T. J., D. J. Nugent, S. J. Staats, R. P. Orzechowski, E. A. Wayner, and W. G. Carter. 1988. The human fibroblast class II extracellular matrix receptor mediates platelet adhesion to collagen and is identical to the platelet glycoprotein Ia-IIa complex. *J. Biol. Chem.* 263:4516-4519.
32. Larson, R. S., A. L. Corbi, L. Berman, and T. Springer. 1989. Primary structure of the LFA-1 α subunit: an integrin with an embedded domain defining a protein superfamily. *J. Cell Biol.* 108:703-712.
33. Law, S. K. A., J. Gagnon, J. E. K. Hildreth, C. E. Wells, A. C. Willis, and A. J. Wong. 1987. The primary structure of the beta subunit of the cell surface adhesion glycoproteins LFA-1, CR3, and P150,95 and its relationship to the fibronectin receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:915-919.
34. Lawler, J., and R. O. Hynes. 1986. The structure of human thrombospondin an adhesive glycoprotein with multiple calcium-binding sites and homologues with several different proteins. *J. Cell Biol.* 103:1635-1648.
35. Loftus, J. C., E. F. Plow, L. K. Jennings, and M. H. Ginsberg. 1988. Alternative proteolytic processing of platelet membrane GPIIb. *J. Biol. Chem.* 263:11025-11032.
36. MacKrell, A. J., B. Blumberg, S. R. Haynes, and J. H. Fessler. 1988. The lethal myospheroid gene of *Drosophila* encodes a membrane protein homologous to vertebrate integrin β subunits. *Proc. Natl. Acad. Sci. USA.* 85:2633-2637.
37. Maniatis, T., E. F. Frisch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
38. Marcantonio, E. E., and R. O. Hynes. 1988. Antibodies to the conserved cytoplasmic domain of the integrin β subunit react with proteins in vertebrates, invertebrates and fungi. *J. Cell Biol.* 106:1765-1772.
39. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262:10035-10038.
40. Mole, J. E., J. K. Anderson, E. A. Davison, and D. E. Woods. 1984. Complete primary structure for the zymogen of human complement factor B. *J. Biol. Chem.* 259:3407-3412.
41. Nieuwenhuis, H. K., J. W. N. Akkerman, W. P. M. Houdijk, and J. J. Sixma. 1985. Human blood platelets showing no response to collagen fail to express surface glycoprotein Ia. *Nature (Lond.)* 318:470-472.
42. Pareti, F. I., K. Niya, J. M. McPherson, and Z. M. Ruggeri. 1987. Isolation and characterization of two domains of human von Willebrand factor that interact with fibrillar collagen types I and III. *J. Biol. Chem.* 262:13835-13841.
43. Pischel, K. D., M. E. Hemler, C. Huang, H. G. Bluestein, and V. L. Woods. 1987. Use of the monoclonal antibody 12F1 to characterize the differentiation antigen VLA-2. *J. Immunol.* 138:226-233.
44. Pischel, K. D., H. G. Bluestein, and V. L. Woods. 1988. Platelet glycoproteins Ia, Ic, and IIa are physicochemically indistinguishable from the very late activation antigens adhesion-related proteins of lymphocytes and other cell types. *J. Clin. Invest.* 81:505-513.
45. Poncz, M., R. Eisman, R. Heidenreich, S. M. Silver, G. Vilaire, S. Surrey, E. Schwartz, and J. S. Bennett. 1987. Structure of the platelet membrane glycoprotein IIb: homology to the alpha subunits of the vitronectin and fibronectin membrane receptors. *J. Biol. Chem.* 262:8476-8482.
46. Pytela, R. 1988. Amino acid sequence of the murine Mac-1 α chain reveals homology with the integrin family and an additional domain related to von Willebrand factor. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1371-1378.
47. Reinach, F. C., K. Nagai, and J. Jones-Kendrick. 1986. Site-directed mutagenesis of the regulatory light-chain $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding site and its role in hybrid myosins. *Nature (Lond.)* 322:80-83.
48. Roth, G. J., K. Titani, L. W. Hoyer, and M. J. Hickey. 1986. Localization of binding sites within human von Willebrand factor for monomeric type III collagen. *Biochemistry.* 25:8357-8361.
49. Ruoslahti, E., and M. D. Pierschbacher. 1987. New perspectives in cell adhesion: RGD and integrins. *Science (Wash. DC)* 238:491-497.
50. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.
51. Santoro, S. A. 1986. Identification of a 160,000 dalton platelet membrane protein that mediates the initial divalent cation-dependent adhesion of platelets to collagen. *Cell.* 46:913-920.
52. Santoro, S. A., S. M. Rajpara, W. D. Staatz, and V. L. Woods. 1988. Isolation and characterization of a platelet surface collagen binding complex related to VLA-2. *Biochem. Biophys. Res. Commun.* 153:217-223.
53. Santoso, S., V. Kiefel, and C. Eckhardt-Mueller. 1989. Immunochromatographic characterization of the new platelet alloantigen system Bra/Brb. *J. Clin. Invest.* In Press.
54. Shelton-Inloes, B. B., K. Titani, and J. E. Sadler. 1986. cDNA sequences for human von Willebrand factor reveal five types of repeated domains and five possible protein sequence polymorphisms. *Biochemistry.* 25:3164-3171.
55. Sneath, P. H. A., and Sokal, R. R. 1973. *Numerical taxonomy.* W. H. Freeman & Co. San Francisco, CA. 573 pp.
56. Springer, T. A., M. L. Dustin, T. K. Kishimoto, and S. D. Marlin. 1987. The lymphocyte function-associated LFA-1, CD2, and LFA-3 molecules: cell adhesion receptors of the immune system. *Annu. Rev. Im-*

- munol.* 5:223-252.
57. Sugiyama, Y., F. A. Chen, H. Takita, and R. B. Bankert. 1988. Selective growth inhibition of human lung cancer cell lines bearing a surface glycoprotein gp160 by ¹²⁵I-labeled anti-gp160 monoclonal antibody 1. *Cancer Res.* 48:2768-2773.
 58. Suzuki, S., W. S. Argraves, H. Arai, L. R. Languino, M. Pierschbacher, and E. Ruoslahti. 1987. Amino acid sequence of the vitronectin receptor alpha subunit and comparative expression of adhesion receptor mRNAs. *J. Biol. Chem.* 262:14080-14085.
 59. Szebenyi, D. M. E., S. K. Obendorf, and K. Moffat. 1981. Structure of vitamin D-dependent calcium-binding protein from bovine intestine. *Nature (Lond.)*. 294:327-332.
 60. Takada, Y., J. L. Strominger, and M. E. Hemler. 1987. The very late antigen family of heterodimers is part of a superfamily of molecules involved in adhesion and embryogenesis. *Proc. Natl. Acad. Sci. USA.* 84:3239-3243.
 61. Takada, Y., E. A. Wayner, W. G. Carter, and M. E. Hemler. 1988. The extracellular matrix receptors, ECMRII and ECMRI, for collagen and fibronectin correspond to VLA-2 and VLA-3 in the VLA family of heterodimers. *J. Cell. Biochem.* 37:385-393.
 - 61a. Takada, Y., M. J. Elices, C. Crouse, and M. E. Hemler. 1989. The primary structure of the α^4 subunit of VLA-4: homology to the other integrins and a possible cell-cell adhesion function. *EMBO (Eur. Mol. Biol. Organ.) J.* In press.
 62. Tamkun, J. W., D. W. DeSimone, D. Fonda, R. S. Patel, C. Buck, A. F. Horwitz, and R. O. Hynes. 1986. Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell.* 46:271-282.
 63. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA.* 76:4350-4351.
 64. von Heijne, G. 1984. How signal sequences maintain cleavage specificity. *J. Mol. Biol.* 173:243-251.
 65. Vyas, N. K., M. N. Vyas, and F. A. Quioco. 1987. A novel calcium bonding site in the galactose-binding protein of bacterial transport and chemotaxis. *Nature (Lond.)*. 327:635-640.
 66. Wayner, E. A., and W. G. Carter. 1987. Identification of multiple cell adhesion receptors for collagen and fibronectin in human fibrosarcoma cells possessing unique alpha and common beta subunits. *J. Cell Biol.* 105:1873-1884.
 67. Young, R. A., and R. W. Davis. 1983. Yeast RNA Polymerase II Genes: isolation with antibody probes. *Science (Wash. DC)*. 222:778-782.
 68. Zylstra, S., F. A. Chen, S. K. Ghosh, E. A. Repasky, U. Rao, H. Takita, and R. B. Bankert. 1986. Membrane-associated glycoprotein (gp160) identified on human lung tumor by a monoclonal antibody. *Cancer Res.* 46:6446-6451.