THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 51, Issue of December 21, pp. 48206–48212, 2001

Printed in U.S.A.

Selective Binding of Collagen Subtypes by Integrin α_1I, α_2I, and α_{10}I Domains*

Received for publication, May 7, 2001, and in revised form, September 6, 2001
Published, JBC Papers in Press, September 25, 2001, DOI 10.1074/jbc.M104058200

Mira Tulla‡, Olli T. Pentikäinen§, Tiina Viitasalo‡, Jarmo Käpylä‡, Ulla Impola‡, Petri Nykvist‡, Liisa Nissinen‡¶, Mark S. Johnson§, and Jyrki Heino‡¶‖

From the ‡Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä FIN-40351, the ¶Department of Medical Biochemistry and MediCity Research Laboratory, University of Turku, Turku FIN-20520, and the §Department of Biochemistry and Pharmacy, Åbo Akademi University, Turku FIN-20521, Finland

Four integrins, namely $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_{10}\beta_1$, and $\alpha_{11}\beta_1$, form a special subclass of cell adhesion receptors. They are all collagen receptors, and they recognize their ligands with an inserted domain (I domain) in their α subunit. We have produced the human integrin α_{10}I domain as a recombinant protein to reveal its ligand binding specificity. In general, α_{10}I did recognize collagen types I–VI and laminin-1 in a Mg^{2+}-dependent manner, whereas its binding to tenascin was only slightly better than to albumin. When α_{10}I was tested together with the α_1I and α_2I domains, all three I domains seemed to have their own collagen binding preferences. The integrin α_2I domain bound much better to fibrillar collagens (I–III) than to basement membrane type IV collagen or to beaded filament-forming type VI collagen. Integrin α_1I had the opposite binding pattern. The integrin α_{10}I domain was similar to the α_1I domain in that it bound very well to collagen types IV and VI. Based on the previously published atomic structures of the α_1I and α_2I domains, we modeled the structure of the α_{10}I domain. The comparison of the three I domains revealed similarities and differences that could potentially explain their functional differences. Mutations were introduced into the αI domains, and their binding to types I, IV, and VI collagen was tested. In the α_2I domain, Asp-219 is one of the amino acids previously suggested to interact directly with type I collagen. The corresponding amino acid in both the α_1I and α_{10}I domains is oppositely charged (Arg-218). The mutation D219R in the α_2I domain changed the ligand binding pattern to resemble that of the α_1I and α_{10}I domains and, vice versa, the R218D mutation in the α_1I and α_{10}I domains created an α_2I domain-like ligand binding pattern. Thus, all three collagen receptors appear to differ in their ability to recognize distinct collagen subtypes. The relatively small structural differences on their collagen binding surfaces may explain the functional specifics.

Collagens are abundant structural proteins in the extracellular matrix. So far, 19 different triple helical protein trimers have been classified as a collagen subtype (1). The collagens can be grouped into subclasses according to their structural details. Many collagen subtypes (namely types I, II, III, V, and XI) have long continuous triple helices, and they can form large fibrils. In other collagens the triple helix has interruptions. Some collagens form networks (types IV, VIII, and X) or beaded filaments (type VI). Other collagen subclasses include fibril-associated collagen with short interruptions in the triple helices (collagen types IX, XII, XIV, XVI, and XIX), anchoring fibril-forming collagen (type VII), and transmembrane collagen (types XIII and XVII). Collagen types XV and XVIII are found in association with basement membranes (the multiplexins; see Ref. 2).

The integrins form a large family of heterodimeric cell surface receptors involved in cell-extracellular matrix as well as in cell-cell adhesion and communication. Of the 24 different integrin receptors presently known, 5 function as a collagen receptor (3). The collagen receptors are composed of the β_1 subunit in complex with either an α_1, α_2, α_3, α_{10}, or an α_{11} subunit. Integrin $\alpha_1\beta_1$ is the collagen receptor of many mesenchymal cells. α_1 null mice are viable (4), but they seem to have defects in the feedback regulation of collagen synthesis (5), in the regulation of matrix metalloproteinase expression (5, 6), in tumor-related angiogenesis (6), and in lymphocyte function (7). Integrin $\alpha_2\beta_1$ is an important collagen receptor on platelets and epithelial cells (8). Little is known about the biology of the recently identified $\alpha_{10}\beta_1$ and $\alpha_{11}\beta_1$ integrins. $\alpha_{10}\beta_1$ is one of the collagen receptors on chondrocytes (9), and $\alpha_{11}\beta_1$ was originally found in fetal muscle (10). Integrin $\alpha_3\beta_1$ can bind some collagens, but it may function as an assisting collagen receptor rather than as a primary receptor (11).

One notable feature common to the primary collagen receptors is the existence of an I ("inserted") domain at the N terminus of the α subunit. No other extracellular matrix receptor contains this independently folding domain, but five integrin-type cell-cell adhesion receptors do have it. The I domain is built up of β sheets surrounded by amphipathic α helices (12). I domains are homologous to the A domains found in von Willebrand factor and in cartilage matrix protein, in some collagen subtypes, and in components of the complement system. The A and I domains are commonly involved in molecular interactions, and they are responsible for the collagen binding activity of von Willebrand factor and collagen receptor integrins (13–15).

At the "top" of the I domain, where ligands bind, resides a metal ion in a conserved coordination site called MIDAS[1]

* This work was supported by the Academy of Finland, the Finnish Cancer Association, the Sigrid Jusélius Foundation, and the Technology Development Center in Finland. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed: Dept. of Biological and Environmental Science, University of Jyväskylä, P. O. Box 35, Jyväskylä FIN-40351, Finland. Tel.: 358-14-260-2240; Fax: 358-14-260-2221; E-mail: jyrki.heino@utu.fi.

[1] The abbreviations used are: MIDAS, metal ion-dependent adhesion site; PCR, polymerase chain reaction; MBP, maltose-binding protein; GST, glutathione *S*-transferase; PBS, phosphate-buffered saline; BSA, bovine serum albumin.

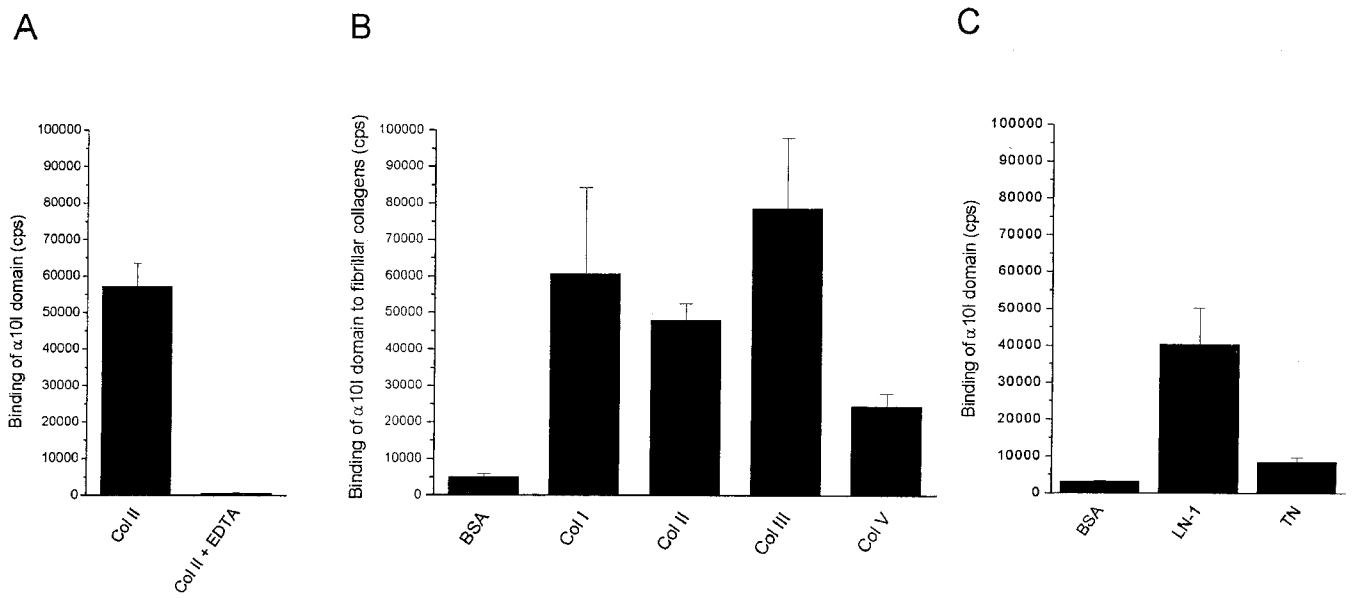

FIG. 1. **Binding of α_{10}I domain to extracellular matrix ligands.** Microtiter plates were coated with 5 μg/cm^2 collagen types I (*Col I*, rat), II (*Col II*, bovine), III (*Col III*, human), V (*Col V*, human), laminin-1 (*LN-1*, mouse), or tenascin (*TN*, chicken) overnight. Diluent containing BSA was used as a background control and to block the wells. The MBP (*panels A* and *B*) or GST fusion (*panel C*) α_{10}I domains (300 nM) were allowed to bind for 1 h in the presence of 2 mM $MgCl_2$ or 5 mM EDTA (*panel A*). Wells were washed three times. Bound αI domains were detected with anti-MBP or anti-GST antibody and Europium^{3+}-labeled protein G. Time-resolved fluorescence measurements were used. The data are the means of four parallel determinations (±S.D.).

(metal ion-dependent adhesion site; see Ref. 16). Five amino acid side chains bind the magnesium ion, directly or through water molecules, and a water molecule or glutamate from a collagenous ligand may complete the coordination to the metal ion. The MIDAS site is centered on a groove restricted on one side by a helix called the αC helix. This αC helix can be found only in collagen binding integrin αI domains. Ligand binding induces a conformational change in the α_2I domain resulting in the unwinding of the αC helix and opening of the binding site (17). Recombinant α_2I domain missing the αC helix has been shown to have altered kinetics in binding type I collagen (18). A prominent feature in the αI domain binding surface is the presence of several charged amino acids surrounding the MIDAS site. The binding surface of α_2I is more negatively charged than the binding surfaces of α_1I and α_{10}I (19).

Despite the structural similarity of the integrins $\alpha_1\beta_1$ and $\alpha_2\beta_1$, their binding specificity differs from each other. Integrin $\alpha_1\beta_1$ prefers network-forming collagen type IV over fibrillar collagen type I, whereas $\alpha_2\beta_1$ binds type I collagen more strongly than collagen type IV (20). Moreover, $\alpha_1\beta_1$ is a much better receptor for type XIII collagen than $\alpha_2\beta_1$ (21). The same binding pattern can also be seen with the corresponding recombinant I domains. Here we have tested the binding of α_1I and α_2I domains to various collagen subtypes. We have also unveiled the ligands recognized by the α_{10}I domain. The similarities and differences in the ligand binding patterns were correlated with the atomic structures, and based on experiments with mutated αI domains, we suggest that Asp-219 in the α_2I domain, and the corresponding amino acid Arg-218 in the α_1I and α_{10}I domains, is a critical residue in the determination of collagen binding specificity.

MATERIALS AND METHODS

Cloning and Mutagenesis of the Human Integrin α_{10}I Domain—α_{10}I domain cDNA was generated by reverse transcription-PCR from RNA isolated from KHOS-240 cells (human Caucasian osteosarcoma). Total cellular RNA was isolated by using an RNeasy Mini Kit (Qiagen). Reverse transcription-PCR was done using the Gene Amp PCR kit (PerkinElmer Life Sciences). The α_{10}I forward primer (5′-CAG GGA TCC CCA ACA TAC ATG GAT GTT GTC-3′) contained a *Bam*HI restriction site at the 5′-end. The reverse primer (5′-GGC TGA ATT CCC CTT CAA GGC CAA AAA TCC G-3′) also contained an *Eco*RI restriction site at the 5′-end. The primers were delivered by CyberGene, Sweden. Forty cycles of PCR amplification were done in 2 mM $MgCl_2$ using the following protocol: denaturation for 1 min at 94 °C, annealing for 1 min at 67 °C, and extension for 2 min at 72 °C. The amplified α_{10}I domain cDNA was digested along with the pMAL-c expression vector (New England Biolabs) or pGEX-2T expression vector (Amersham Pharmacia Biotech) using the *Bam*HI and *Eco*RI restriction enzymes (Promega). The cDNA was ligated to the pMAL-c vector with T4 DNA ligase (Promega). To the pGEX-2T vector the α_{10}I domain cDNA was ligated with the SureClone ligation kit (Amersham Pharmacia Biotech). The constructs were transformed into the *Escherichia coli* BL21 strain for production. The DNA sequences of the constructs were checked with DNA sequencing and compared with the published α_{10} DNA sequence (10) (data not shown). Site-directed mutation of the α_{10}I domain cDNA in a pGEX-2T vector was made using PCR according to Stratagene's QuikChange mutagenesis kit instructions. Briefly, mutagenesis was based on primers that contained the planned mutations (primers were delivered by CyberGene). At 68 °C the proofreading *Pfu* DNA polymerase (Promega) makes only one copy of the plasmid without replacing the primers. The mutated strands were then selected from parental ones by digesting the template with *Dpn*I endonuclease (Promega), which digests only methylated and hemimethylated DNA. By that means only mutated strands remained. The presence of mutations was checked by DNA sequencing. Mutant constructs were then transformed into *E. coli* strain BL21 for production of recombinant protein.

Cloning and Mutagenesis of the Human Integrin α_1I Domain—The α_1I domain was cloned into the pGEX-4T vector (Amersham Pharmacia Biotech) as described earlier by Nykvist *et al*. (21). The full-length integrin α_1 cDNA was a gift from Dr. E. Marcantonio (Columbia University, New York). The α_1I domain cDNA sequence differed from the originally published sequence (23) at one position, and as a result the amino acid lysine was replaced with glutamate at position 170 (numbered from the beginning of the mature peptide). Site-directed mutations to α_1I domain cDNA were made as stated above for the α_{10}I domain.

Cloning and Mutagenesis of the Human Integrin α_2I Domain—Human integrin α_2I domain was generated as described earlier by Ivaska *et al*. (24). The integrin α_2 cDNA was a gift from Dr. M. Hemler, Dana Farber, Boston. The α_2I domain was cloned into the pGEX-2T vector (Amersham Pharmacia Biotech). The site-directed mutations were made as stated above for the α_{10}I domain.

Production and Purification of MBP Fusion Protein—To produce the wild type α_{10}I domain as a fusion with the maltose-binding protein

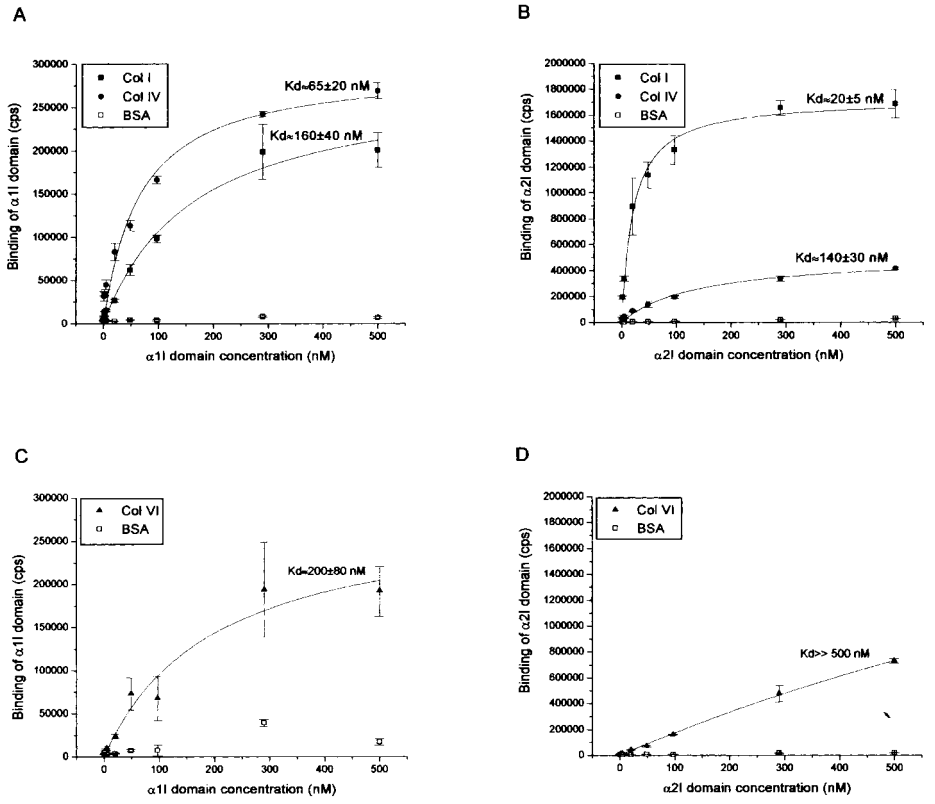

FIG. 2. **Binding curves for α_1I and α_2I domains to collagen types I, IV, and VI.** Microtiter plates were coated with 5 μg/cm^2 collagen types I (*Col I*, rat; *panels A* and *B*), IV (*Col IV*, mouse; *panels A* and *B*), or VI (*Col VI*, human; *panels C* and *D*) overnight. Diluent containing BSA was used as a background control and to block the wells. The αI domains (1–500 nM) were allowed to bind for 1 h in the presence of 2 mM $MgCl_2$. Wells were washed three times. Bound I domains were detected with anti-GST antibody and Europium^{3+}-labeled protein G. Time-resolved fluorescence measurements were used. The data are the means of three parallel determinations (±S.D.). K_d values were obtained by fitting binding data for αI domain concentrations in the range of 1–500 nM using the Michaelis-Menten equation.

(MBP), 500 ml of LB medium containing 50 μg/ml ampicillin was inoculated with an overnight culture of 10 ml of *E. coli* BL21 cells transformed with the pMAL-cα_{10}I vector. Cells were grown at 37 °C for about 4 h. Protein production was induced with isopropyl-β-D-thiogalactoside (Amersham Pharmacia Biotech) and allowed to continue for 2 h. The cells were pelleted by centrifugation (4,000 × *g*, 20 min, 4 °C). Cell pellets were resuspended in column buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol) and stored at −20 °C until purification. The cell suspension was thawed on ice and sonicated in 15-s bursts for 3 min. Cell debris was removed by centrifugation (9,000 × *g*, 30 min, 4 °C). The supernatant was incubated with amylose resin (New England BioLabs) in an end-over-end rotor at 4 °C overnight. Amylose resin with bound MBP-α_{10}I fusion was transferred to 10-ml columns (Bio-Rad) and washed three times with 10 ml of column buffer. The bound fusion protein was eluted with 5 mM maltose in column buffer after a 2–3-h incubation at room temperature. Maltose was removed using hydroxyapatite, first washed with column buffer. One ml of washed hydroxyapatite was mixed with eluted protein and incubated in an end-over-end rotor for 1 h at 4 °C, after which the fusion protein was attached to the hydroxyapatite. Soluble maltose was washed away with column buffer, repeated three times. The fusion protein was eluted from the hydroxyapatite with sodium phosphate buffer after a 30-min incubation at 4 °C. The protein concentrations were measured using Bradford's method (22). The purity and folding of the proteins were checked by SDS- and native polyacrylamide gel electrophoresis (PhastSystem, Amersham Pharmacia Biotech; data not shown).

Production and Purification of GST Fusion Proteins—The wild type and mutated α_1I, α_2I, and α_{10}I domains were produced as glutathione *S*-transferase (GST)-αI domain fusion proteins. 500 ml of LB medium (with 50 μg/ml ampicillin) was inoculated with an overnight culture of *E. coli* BL21 cells transformed with either wild type or mutant plasmid. The cultures were grown at 37 °C until the A_{600} was in the range of 0.5–1. Next, an inducer of the tac promoter, isopropyl β-D-thiogalactoside, was added to a final concentration of 0.4 mM, and the temperature was lowered to room temperature. The bacteria were allowed to produce the recombinant protein for 3–4 h after which the cultures were centrifuged (5,550 rpm, 10 min, 4 °C) to collect the cells. Cells were stored until purification at −20 °C overnight and at −70 °C for longer times.

The bacterial cells were resuspended in PBS for purification. The cell suspension was sonicated and detergent (Triton X-100) was added to a final concentration of 1%. To dissolve the recombinant protein, the disrupted cells were incubated with agitation and Triton X-100 for 1 h. Cell debris was removed by centrifugation (17,000 rpm, 10 min, 4 °C). Glutathione-Sepharose 4B (Amersham Pharmacia Biotech) was added to the supernatant, and the mixture was incubated for 1–1.5 h in gentle agitation. The glutathione-Sepharose 4B was then transferred to columns (Bio-Rad), and unbound proteins were washed away with PBS. The bound fusion protein was eluted with reduced glutathione (Sigma). Protein concentrations were measured with Bradford's method (22). The purity and folding of the fusion proteins were checked with SDS- and native polyacrylamide gel electrophoresis (data not shown).

Solid Phase Binding Assays—The binding of wild type or mutated I domains to collagen was studied using solid phase binding assays. Ninety-six-well microtiter plates (Wallac) were coated with each collagen type: I (rat tail, Sigma), II (bovine, Chemicon), III (human, Chemicon), IV (Engelbreth-Holm-Swarm mouse sarcoma basement membrane, Sigma), V (human, Chemicon), or VI (human, Biodesign International) or with laminin-1 (Engelbreth-Holm-Swarm mouse sarcoma basement membrane, Sigma), tenascin (chicken, Chemicon) 5 μg/cm^2 in PBS and control wells with Diluent II (containing BSA, Wallac) 1:2 in PBS at 4 °C overnight. Before carrying out the binding assays, wells were blocked with Diluent II for 1 h at 37 °C. Dilutions of the I domains were made with assay buffer (Wallac) and added to the wells. A concentration series of 1–500 nM was used for the I domains. The I domains were allowed to attach to the wells for 1 h at 37 °C in the presence of 2 mM $MgCl_2$, unless stated otherwise. The unattached I domain was washed away with PBS containing 2 mM $MgCl_2$, repeated three times. Detection of the I domain was based on the existence of the GST or MBP fusion partner. Goat anti-GST antibody (Amersham Pharmacia Biotech) or anti-MBP (New England Biolabs) was added to the wells in a 1:8,000 or 1:1,000 dilution, respectively, in assay buffer and incubated for 1 h at 37 °C. The wells were then washed three times with PBS containing 2 mM $MgCl_2$. Europium^{3+}-labeled protein G (Wallac) (1:100 in assay buffer) was then added to the wells and incubated for 1 h at 37 °C. Wells were washed three times with PBS containing $MgCl_2$. Enhancement solution (Wallac) was used to dissociate the highly fluorescent Europium^{3+} label. The fluorescence was measured with a time-resolved fluorometer (Victor^2, Wallac). All assays were performed at least in triplicate.

Structures and Structure Modeling—Three-dimensional structures of the integrin α_1I domain in the apo form (1qcy, see Ref. 25; Mg^{2+} present but no bound ligand) and the apo and holo (bound hexapeptide collagen fragment) forms (1aox, Ref. 26; 1dzi, Ref. 17) were obtained from the Protein Data Bank (27). The sequence alignment of human α_1I, α_2I, and α_{10}I domains was made using MALIGN (28) in the BODIL

modeling package[2] using a structure-based sequence comparison matrix (29). MALIGN constructs a multiple sequence alignment from pairwise alignments according to a tree relating the sequences being matched. The alignment between the three I domains contains no gaps and for α_1I and α_2I matches their structural alignment. The sequence identity between the α_2I domain and the α_1I domain is 52% and between the α_2I domain and the α_{10}I domain is 46%, thus leading to high quality reliable models. The program MODELLER 4.0 (30) was used to construct three-dimensional model structures of the holo form of the α_1I domain and both the apo and holo forms of the α_{10}I domain.

RESULTS

α_{10}I Domain Binds Type I–VI Collagens and Laminin-1 in a Magnesium-dependent Manner—α_{10}I domain was produced in *E. coli* as a recombinant protein fused to MBP or GST. The recombinant proteins were purified with affinity chromatography and characterized with SDS- and native polyacrylamide gel electrophoresis (data not shown). The proteins appeared as one band in native polyacrylamide gel electrophoresis, suggesting the presence of a single molecular form. The binding of purified recombinant α_{10}I domain to collagen was studied using a solid phase binding assay. Wells were coated with different collagens, laminin-1, or tenascin (5 μg/cm^2). α_{10}I fusion protein (300 nM) was allowed to bind to the coated wells. The α_{10}I domain was found to bind to cartilage-derived type II collagen in a magnesium-dependent manner (Fig. 1*A*) because 5 mM EDTA inhibited binding completely. Thus, the basic binding mechanism of α_{10}I is cation-dependent as are the other integrin αI domains (16, 17). To investigate whether other fibrillar collagens, namely subtypes I, II, III, and V, also can function as ligands for α_{10}I, a binding assay was carried out (Fig. 1*B*). The α_{10}I domain bound to the collagen subtypes I–III nearly equally, but in repeated experiments the binding to type V collagen was weaker. Collagen receptors $\alpha_1\beta_1$ and $\alpha_2\beta_1$ can both bind to laminin-1 (31, 32). $\alpha_2\beta_1$ is also able to bind tenascin (33). The binding of α_{10}I to these extracellular matrix proteins was also studied. The α_{10}I domain was found to bind to laminin-1, but binding to tenascin was only slightly better than to BSA (Fig. 1*C*). The ligand recognition patterns of α_{10}I-GST and α_{10}I-MBP were identical. In some experiments the α_{10}I-MBP showed higher unspecific binding to BSA than α_{10}I-GST (data not shown).

The Binding Specificities of α_1I and α_2I Domains Differ from Each Other—Integrins $\alpha_1\beta_1$ and $\alpha_2\beta_1$ and their corresponding αI domains differ in collagen binding specificity even though they are structurally highly similar. Overall about half of their αI domain sequence is identical, and the αI domain binding surfaces resemble each other. The α_1I domain prefers type IV collagen over type I collagen, whereas the preference of α_2I is reversed (20). To study the binding of these αI domains to collagen type VI in addition to collagen types I and IV, binding assays with an αI domain concentration series of 1–500 nM were carried out (Fig. 2). The results were fit to the Michaelis-Menten equation to obtain estimates of the K_d of binding. In accordance with previous studies, the integrin α_1I domain bound better to type IV collagen ($K_d \approx 65 \pm 20$ nM; *panel A*) than type I collagen ($K_d \approx 160 \pm 40$ nM). The integrin α_2I domain, in contrast, associated considerably better with collagen type I ($K_d \approx 20 \pm 5$ nM) than with collagen type IV ($K_d \approx 140 \pm 30$ nM; *panel B*).

Type VI collagen forms beaded filaments, and it has been reported to be a ligand of both $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins (34). However, in chondrocytes $\alpha_1\beta_1$ may be the preferred receptor for collagen type VI (35). Here, the binding of the α_1I and α_2I domains to type VI collagen was tested (Fig. 2). Integrin α_1I domain showed the strongest binding ($K_d \approx 200 \pm 80$ nM; Fig. 2*C*), whereas the binding of the α_2I domain was so weak that no reasonable estimate of the half-maximal value for binding could be obtained (Fig. 2*D*).

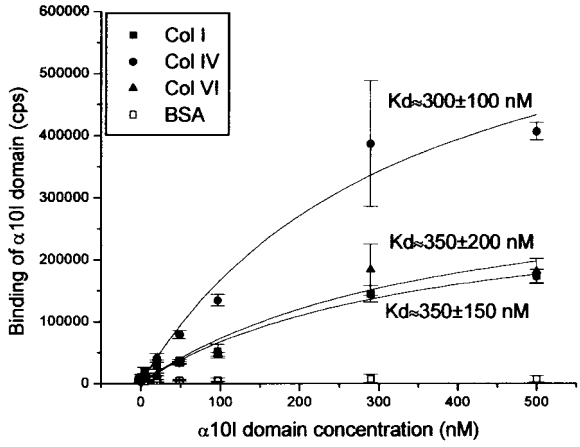

FIG. 3. **Binding of the α_{10}I domain to collagen types I, IV, and VI.** Microtiter plates were coated with 5 μg/cm^2 collagen types I (*Col I*, rat), IV (*Col IV*, mouse), or VI (*Col VI*, human) overnight. Diluent containing BSA was used as a background control and to block the wells. The GST fusion α_{10}I domain (1–500 nM) was allowed to bind for 1 h in the presence of 2 mM $MgCl_2$. Wells were washed three times. Bound αI domain was detected with anti-GST antibody and Europium^{3+}-labeled protein G. Time-resolved fluorescence measurements were used. The data are the means of three parallel determinations (±S.D.). K_d values were obtained by fitting binding data for αI domain concentrations in the range of 1–500 nM using the Michaelis-Menten equation.

Integrin α_{10}I Domain Binds Collagen Types I, IV, and VI with Nearly Equal Affinity—To test whether the α_{10}I domain can also bind collagens other than the fibrillar subtypes and to get quantitative information, a binding assay with collagen types I, IV, and VI was carried out. In the solid phase assay the range of α_{10}I domain concentrations spanned 1–500 nM. The α_{10}I domain was found to bind type IV collagen slightly better ($K_d \approx 300 \pm 100$ nM) than collagen types I and VI (Fig. 3). The estimated K_d for binding to fibrillar collagen type I was 350 ± 150 nM, similar to the value for collagen type VI binding ($K_d \approx 350 \pm 200$ nM). In terms of binding properties, these results place the α_{10}I domain closer to the α_1I domain, which prefers collagen type IV over collagen type I or VI. Moreover, the fact that the α_{10}I domain binds well to collagen type VI, in contrast to the α_2I domain, suggests that the binding preferences of the α_1I and α_{10}I domains are similar.

The Sequence Surrounding the MIDAS Site Differs in the α_1I, α_2I, and α_{10}I Domains—Sequence differences in the vicinity of the binding site were identified by using the sequence alignment of the α_1I, α_2I, and α_{10}I domains, and the x-ray structures of the apo form of the α_1I domain (25) and both the apo and holo forms (1aox, Ref. 26; 1dzi, Ref. 17) of the α_2I domain. To make it easier to visualize the differences among the surface regions surrounding the MIDAS site, both the α_1I (holo form) and α_{10}I (apo and holo forms) domain structures were modeled using MODELLER 4.0. The electrostatic potentials were mapped to the solvent accessible surfaces of the apo and holo forms (Fig. 4) for each of the αI domains, and the sequence differences are listed in Table I.

The α_1I, α_2I, and α_{10}I domains have among them 11 amino acid positions (Table I), in the vicinity of the MIDAS site in both the apo and holo forms, where amino acid replacements show large differences in either the chemistry and/or the size of the amino acids at those sites. In the α_1I domain, 9 of these 11 residues are polar, of which 2 are positively charged and 1 is negatively charged; α_2I has a total of 6 polar residues (1 posi-

[2] J. Lehtonen, V.-V. Rantanen, D.-J. Still, M. Gyllenberg, and M. S. Johnson, unpublished results.

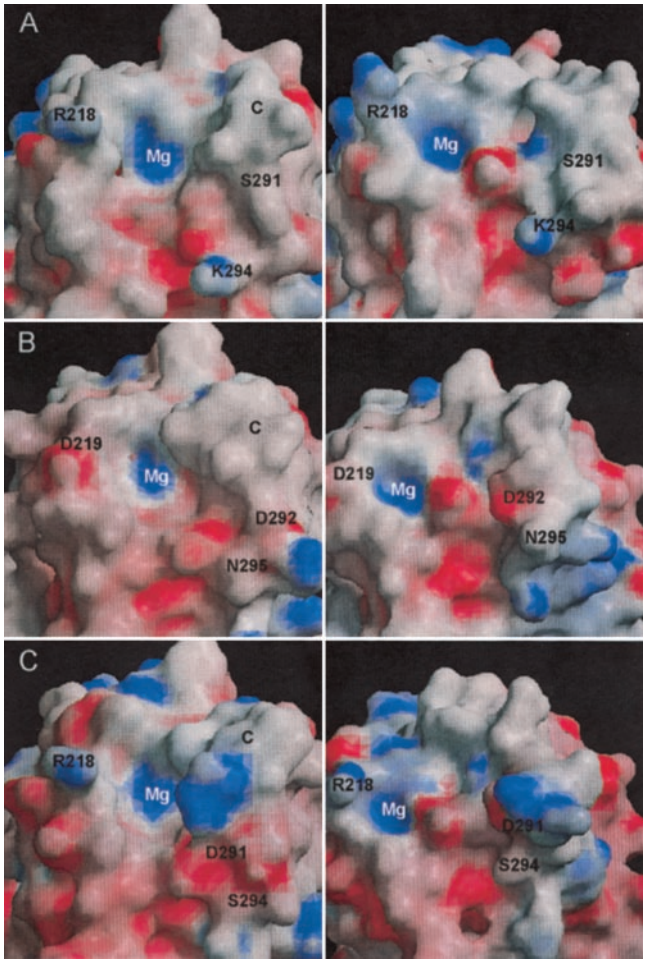

FIG. 4. **Solvent accessible surfaces surrounding the MIDAS site α_1I (*panel A*), α_2I (*panel B*), and α_{10}I (*panel C*) domains.** For each I domain, the *left image* corresponds to the apo form, and the *right image* corresponds to the holo form. The surface charge distributions were calculated using the program GRASP (see Ref. 17) and rendered using RASTER3D (Ref. 45): *blue*, positive charge; *red*, negative charge; and *white*, neutral regions. The bound metal (*Mg*) and some key residues are *labeled*; helix C, which unwinds during collagen binding, is labeled *C* in the *left images*.

tively and 2 negatively charged residues); but α_{10}I has, in addition to glycine, a total of 10 polar residues, and 7 of them are charged residues: 4 positively and 3 negatively.

Differences between α_{10}I and α_2I—Asp-219 in α_2I may participate directly in ligand binding (17). The corresponding amino acid in α_{10}I is Arg-218. Of the 11 positions that exhibit chemical and size changes among the three integrins, α_2I and α_{10}I have several residues in common: Gly-259 (α_{10}I) and Gly-260 (α_2I); Asp-291 and Asp-292; Arg-297 and Lys-298. Of the 4 positively charged residues in α_{10}I, 2 (Arg-288 (Asn-289 in α_2I) and Arg-290 (Leu-291 in α_2I)) are not found in α_2I but are located in α_{10}I within a region that undergoes a conformational change on collagen binding in α_2I (residues 283–293, α_2 numbering). The model of the α_{10}I holo form places these residues away from the MIDAS site. The structure of α_2I with a bound collagen fragment, a GFOGER hexapeptide complex (17), shows that this region is not directly involved in collagen binding, suggesting that if binding occurs in a similar manner in α_{10}I then these 2 residues may not be important for collagen binding in α_{10}I either.

Differences between α_{10}I and α_1I—Arginine is found at position 218 in both of these αI domains. Thus, the differences in their collagen binding patterns must be the result of interactions with other residues. Sequence differences in the vicinity

TABLE I
Sequence differences in the vicinity of the MIDAS site in the α_1I, α_2I, and α_{10}I domains

The first two columns contain the corresponding amino acid numbering for the α_1I/α_{10}I domains and α_2I domain, respectively. The next three columns contain the three-letter code of the amino acid at the corresponding position in each I domain.

Amino acid numbering		Three-letter code for an amino acid		
$\alpha_{1,10}$	α_2	α_1	α_2	α_{10}
218	219	Arg	Asp	Arg
219	220	Qln	Leu	Glu
259	260	Asn	Gly	Gly
260	261	His	Ser	Glu
285	286	Tyr	Leu	Tyr
288	289	Gly	Asn	Arg
290	291	Leu	Leu	Arg
291	292	Ser	Asp	Asp
294	295	Lys	Asn	Ser
297	298	Glu	Lys	Arg
301	302	Ser	Ala	Thr
Amino acids that are conservatively replaced among the integrins or are less likely to be involved in ligand binding				
$\alpha_{1,10}$	α_2	α_1	α_2	α_{10}
284	285	Ser	Tyr	His
286	287	Asn	Asn	Leu
289	290	Asn	Ala	Qln
292	293	Thr	Thr	Pro
293	294	Glu	Lys	Ser
296	297	Val	Ile	Leu
295	296	Phe	Leu	Phe
300	301	Lys	Lys	Arg

of MIDAS are: Gln-219 (α_1) → Glu-219 (α_{10}), Asn-286 → Leu-286, Gly-288 → Arg-288, Leu-290 → Arg-290, Ser-291 → Asp-291, Thr-292 → Pro-292, Glu-293 → Ser-293, Lys-294 → Ser-294, and Glu-297 → Arg-297. The amino acids at positions 286, 292, and 293 are located too distant from the collagen binding area seen in the α_2I complex to be involved directly in binding of the collagen fragment.

Differences between α_1I and α_2I—Asp-219 in αI is replaced by Arg-218 in the α_1I domain. Among the other sequence differences in the vicinity of MIDAS are: Gln-219 (α_1) → Leu-220 (α_2), Tyr-285 → Leu-286, Gly-288 → Asn-289, Asn-289 → Ala-290, Ser-291 → Asp-292, Glu-293 → Lys-294, Lys-294→ Asn-295, and Glu-297 → Lys-298. Tyr-285 (α_2) is not located close to the ligand in the α_2I complex structure (17); Ala-290 (α_2) is located far from the collagen binding area in both the unbound and bound α_2I crystal structures.

Mutation Analysis Indicates a Critical Role for Asp-219 in α_2I and Arg-218 in α_1I/α_{10}I in the Discrimination between Different Collagen Subtypes—Based on sequence comparisons and molecular modeling, mutations were introduced into the α_1I, α_2I, and α_{10}I domains. To test the idea that Arg-218 in α_1I is important for the collagen binding specificity, this residue was swapped using site-directed mutagenesis to the corresponding amino acid of α_2I, an aspartate, thus reversing the charge. The αI domain was produced as a GST fusion protein, and binding assays for collagen types I, IV, and VI were performed (Fig. 5*A* and Table II). The ligand binding profile of α_1I R218D resembles that of α_2I. The binding to collagen types IV and VI weakened considerably compared with type I collagen binding, which remained nearly the same (Fig. 5*A* and Table II). These results indicate that Arg-218 in the α_1I domain is one of the key amino acids functioning to discriminate among collagen subtypes. The equal weakening in the binding to type IV and type VI collagen suggests that the binding mechanisms of these collagen subtypes may be similar. A corresponding mutation in the α_{10}I domain had very similar effects. The R218D α_{10}I domain bound to type I collagen in the same way as the

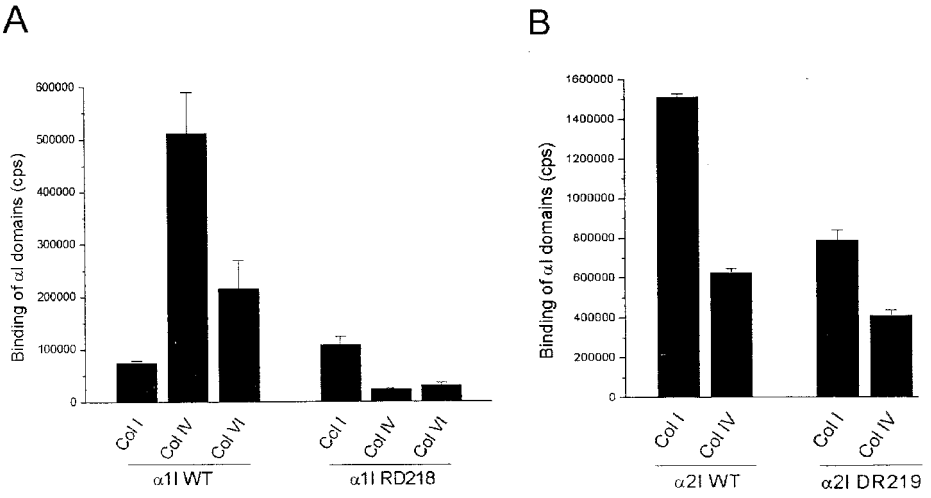

FIG. 5. **Binding of α_1I RD218 (*panel A*) and α_2I DR219 (*panel B*) mutants and wild types (*WT*) to collagen types I, IV, and VI.** Microtiter plates were coated with 5 μg/cm^2 collagen types I (*Col I*, rat), IV (*Col IV*, mouse), or VI (*Col VI*, human) overnight. Diluent containing BSA was used as a background control and to block the wells. The αI domains (290 nM) were allowed to bind for 1 h in the presence of 2 mM $MgCl_2$. Wells were washed three times. Bound αI domains were detected with anti-GST antibody and Europium^{3+}-labeled protein G. Time-resolved fluorescence measurements were used. The data are the means of three parallel determinations (±S.D.).

TABLE II

Approximate K_d values for mutant α_1I, α_2I, and α_{10}I domains to collagen type I, IV, and VI binding

K_d values (nM) for collagen binding were obtained by fitting binding data for αI domain concentrations in the range of 1–500 nM using the Michaelis-Menten equation.

	α_1I R218D	% of wild type binding (K_d)	α_2I D219R	% of wild type binding (K_d)	α_{10}I R218D	% of wild type binding (K_d)
Col I	≈200 ± 50	80	>500	<10	≈400 ± 100	90
Col IV	≫500	<10	≈250 ± 25	60	>500	
Col VI	≫500		≫500		≈500 ± 100	70

wild type α_{10}I (Table II), whereas its binding to type IV collagen was considerably weaker (Table II).

To investigate the importance of the α_1I and α_{10}I Arg-218 equivalent amino acid Asp-219 in α_2I, mutant I domains were produced. Asp-219 was mutated to arginine, changing the negatively charged residue to a positively charged residue. Binding assays using collagen types I, IV, and VI were performed. The D219R mutation was found to retain the collagen type IV binding level better, whereas collagen type I binding weakened significantly (Fig. 5*B* and Table II). The effect of mutation to collagen type VI binding is more difficult to evaluate because the binding of the α_2I domain wild type is also so weak (Table II). These data stress the importance of Asp-219 in the α_2I domain for collagen type I recognition, although it is obvious that one single amino acid cannot explain all of the differences seen in collagen binding among the αI domains.

The collagen binding pattern of the D219R α_2I domain resembles the binding pattern of the α_{10}I domain more than it resembles the collagen binding pattern of α_1I. Five positions at the collagen binding surface of the I domains are more similar (identical, chemically similar, or of more similar size) in the α_2I and α_{10}I domains than in the α_1I domain. These positions are: 259/260 (α_1-N α_2, α_{10}-G), 288/289 (α_1-G, α_2-N, α_{10}-R), 291/292 (α_1-S, α_2, α_{10}-D), 294/295 (α_1-K, α_2-N, α_{10}-S), and 297/298 (α_1-E, α_2-K, α_{10}-R). To date only the α_1I S291D mutation has been made, but no effect on the collagen binding pattern was observed (data not shown).

DISCUSSION

Integrins are found in all metazoans from marine sponges to man (36). However, the I domains containing collagen-binding integrins are not present in the published genomes of *Drosophila melanogaster* or *Caenorhabditis elegans*, suggesting that they may have evolved significantly later (37). The four collagen receptors are structurally very similar, and their tissue distribution is partially overlapping. Especially in bone- and cartilage-derived cells, all four receptors may be present at the same time.[3] Because the short intracellular domains of the collagen-binding α subunits are very different, it has been assumed that the four receptors have distinct signaling functions after binding ligand. Opposite effects on cells have been observed for the $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins. Cell adhesion to collagen, when mediated by $\alpha_1\beta_1$ integrin, promotes cell growth and suppresses collagen synthesis (3, 38–40). In contrast, $\alpha_2\beta_1$ integrin inhibits the cell cycle and promotes matrix remodeling by increasing both collagen synthesis and degradation (3, 38, 39, 41, 42). It is not known whether $\alpha_{10}\beta_1$ or $\alpha_{11}\beta_1$ binding to collagen leads to alterations in cell behavior.

Another reason for the simultaneous existence of multiple collagen receptors may be the fact that the collagens, despite their common basic structure, have significant structural and functional differences. The collagen receptor integrins recognize their ligands by means of a special αI domain. Integrin α_1I and α_2I domains have been reported to bind to various collagen subtypes (20), but the ligand pattern of α_{10}I has been unknown. We have produced a human α_{10}I domain as a recombinant protein and tested it together with α_1I and α_2I domains. In general, α_{10}I binds to collagen types I–VI with a K_d that is within the range typical of α_1I and α_2I domains. Collagen binding required the presence of magnesium. Integrin α_{10}I could also recognize laminin-1, but its ability to bind to tenascin was weaker. A more detailed analysis with different collagen subtypes revealed interesting similarities and differences among the three αI domains investigated here. The α_1I and α_{10}I domains resemble each other in the sense that they both bind to non-fibrillar collagen types IV and VI, whereas α_2I prefers fibrillar collagens. The difference between α_1I and α_2I domains in the recognition of type I and IV collagen had been noted previously (20). A recent study suggested that in chondrocytes, $\alpha_1\beta_1$ is the predominant type VI collagen receptor (35). In support of this observation, our results demonstrate

[3] L. Nissinen and J. Heino, unpublished results.

that α_1I binds type VI collagen significantly better than does the α_2I domain. Type VI collagen is expressed in various tissues, including cartilage. The fact that the α_{10}I domain also binds type VI collagen suggests that chondrocytes may have several receptors mediating interactions with the collagen in beaded filaments.

A crystal structure of a collagen-like triple helix in complex with α_2I identified amino acids in α_2I which can make a direct contact with the ligand (17). Among others, Asp-219 was found to form a salt bridge with an arginine residue in the collagenous peptide. Accordingly, earlier modeling and mutational studies had suggested that Asp-219 is important for α_2I binding to type I collagen (18) and to peptides derived from snake venom (19). No crystal structure of the α_1I domain in contact with any ligand is yet available. However, it is obvious that there must be differences in the ligand binding mechanism because in the α_1I domain the amino acid corresponding to Asp-219 (α_2I) is Arg-218 and therefore has an opposite charge. Here, the mutation D219R in the α_2I domain dramatically decreased the ratio of type I collagen binding to type IV and to type VI collagen binding. Correspondingly, the mutation R218D in the α_1I and α_{10}I domains increased this ratio. Thus, we suggest that Asp-219 (α_2I) and Arg-218 (α_1I/α_{10}I) are critical amino acids in the determination of binding preferences. It is obvious that in addition to this single amino acid position other differences in the structure of the three αI domains must contribute to their different functions. We have no direct evidence (complex structures) that Arg-218 in the α_1I domain or the α_{10}I domain interacts directly with their ligands. The fact that both types I and IV collagen have multiple integrin binding sites (43, 44) suggests that a single αI domain can bind to different collagens in different ways.

In conclusion, we have shown that the α_{10}I domain can recognize other collagen subtypes, in addition to type II collagen (9), including fibrillar and non-fibrillar collagens. α_{10}I may also function as a receptor for laminin-1. We have identified an essential determinant of ligand binding specificity of the integrin αI domains. The amino acid Arg-218 in the α_1I and the α_{10}I domain seems to promote binding to non-fibrillar collagen types IV and VI, whereas an opposite charge at the same position shifts the preference of the αI domain toward fibrillar collagens, partially explaining the function of the α_2I domain.

Acknowledgments—We thank Drs. E. Marcantonio and M. Hemler for cDNAs. We gratefully acknowledge Arja Mansikkaviita for technical assistance.

REFERENCES

1. Prockop, D. J., and Kivirikko, K. I. (1995) *Annu. Rev. Biochem.* **64,** 403–434
2. Oh, S. P., Warman, M. L., Seldin, M. F., Cheng, S. D., Knoll, J. H., Timmons, S., and Olsen, B. R. (1994) *Genomics* **19,** 494–499
3. Heino, J. (2000) *Matrix Biol.* **19,** 319–323
4. Gardner, H. A., Kreidberg, J., Koteliansky, V., and Jaenisch, R. (1996) *Dev. Biol.* **175,** 301–313
5. Gardner, H. A., Broberg, A., Pozzi, A., Laato, M., and Heino, J. (1999) *J. Cell Sci.* **112,** 263–272
6. Pozzi, A., Moberg, P. E., Miles, L. A., Wagner, S., Soloway, P., and Gardner, H. A. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 2202–2207
7. Meharra, E. J., Schön, M., Hassett, D., Parker, C., Havran, W., and Gardner, H. (2000) *Cell. Immunol.* **201,** 1–5
8. Zutter, M. M., and Santoro, S. A. (1990) *Am. J. Pathol.* **137,** 113–120
9. Camper, L., Hellman, U., and Lundgren-Åkerlund, E. (1998) *J. Biol. Chem.* **273,** 20383–20389
10. Velling, T., Kusche-Gullberg, M., Sejersen, T., and Gullberg, D. (1999) *J. Biol. Chem.* **274,** 25735–25742
11. DiPersio, C. M., Shah, S., and Hynes, R. O. (1995) *J. Cell Sci.* **108,** 2321–2336
12. Lee, J.-O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995) *Cell* **80,** 631–638
13. Cruz, M. A., Yuan, H., Lee, J. R., Wise, R. J., and Handin, R. I. (1995) *J. Biol. Chem.* **270,** 10822–10827
14. Kamata, T., and Takada, Y. (1994) *J. Biol. Chem.* **269,** 26006–26010
15. Kern, A., Briesewitz, R., Bank, I., and Marcantonio, E. E. (1994) *J. Biol. Chem.* **269,** 22811–22816
16. Mischishita, M., Videm, V., and Arnaout, M. A. (1993) *Cell* **72,** 857–867
17. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000) *Cell* **100,** 47–56
18. Käpylä, J., Ivaska, J., Riikonen, R., Nykvist, P., Pentikäinen, O., Johnson, M., and Heino, J. (2000) *J. Biol. Chem.* **275,** 3348–3354
19. Pentikäinen, O., Hoffren, A.-M., Ivaska, J., Käpylä, J., Nyrönen, T., Heino, J., and Johnson, M. S. (1999) *J. Biol. Chem.* **274,** 31493–31505
20. Kern, A., Eble, J., Golbik, R., and Kuhn, K. (1993) *Eur. J. Biochem.* **215,** 151–159
21. Nykvist, P., Tu, H., Ivaska, J., Käpylä, J., Pihlajaniemi, T., and Heino, J. (2000) *J. Biol. Chem.* **275,** 8255–8261
22. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
23. Briesewitz, R., Epstein, M. R., and Marcantonio, E. E. (1993) *J. Biol. Chem.* **268,** 2989–2996
24. Ivaska, J., Käpylä, J., Pentikäinen, O., Hoffren, A.-M., Hermonen, J., Huttunen, P., Johnson, M. S., and Heino, J. (1999) *J. Biol. Chem.* **274,** 3513–3521
25. Salminen, T. A., Nymalm, Y., Kankare, J., Käpylä, J., Heino, J., and Johnson, M. S. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 1365–1367
26. Emsley, J., King, S. L., Bergelson, J. M., and Liddington, R. C. (1997) *J. Biol. Chem.* **272,** 28512–28517
27. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28,** 235–242
28. Johnson, M. S., and Overington, J. P. (1993) *J. Mol. Biol.* **233,** 716–738
29. Johnson, M. S., May, A. C. W., Rodionov, M. A., and Overington, J. P. (1996) *Methods Enzymol.* **266,** 575–598
30. Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol.* **234,** 779–815
31. Tomaselli, K. J., Hall, D. E., Flier, L. A., Gehlsen, K. R., Turner, D. C., Carbonetto, S., and Reichardt, L. F. (1990) *Neuron* **5,** 651–662
32. Languino, L. R., Gehlsen, K. R., Wayner, E., Carter, W. G., Engvall, E., and Ruoslahti, E. (1989) *J. Cell Biol.* **109,** 2455–2462
33. Sriramarao, P., Mendler, M., and Bourdon, M. A. (1993) *J. Cell Sci.* **105,** 1001–1012
34. Loeser, R. F. (1997) *Arthritis Rheum.* **40,** 270–276
35. Loeser, R. F., Sadiev, S., Tan, L., and Goldring, M. B. (2000) *Osteoarthr. Cartil.* **8,** 96–105
36. Müller, W. E. G. (1998) *Naturwissenschaften* **85,** 11–25
37. Hynes, R. O., and Zhao, Q. (2000) *J. Cell Biol.* **150,** F89–F95
38. Riikonen, T., Westermarck, J., Koivisto, L., Broberg, A., Kähäri, V.-M., and Heino, J. (1995) *J. Biol. Chem.* **270,** 13548–13552
39. Langholz, O., Rockel, D., Mauch, C., Kozlowska, E., Bank, I., Krieg, T., and Eckes. B. (1995) *J. Cell Biol.* **131,** 1903–1915
40. Pozzi, A., Wary, K. K., Giancotti, F. G., and Gardner, H. A. (1998) *J. Cell Biol.* **142,** 587–594
41. Koyama, H., Raines, E. W., Bornfeldt, K. E., Roberts, J. M., and Ross, R. (1996) *Cell* **87,** 1069–1078
42. Ivaska, J., Reunanen, H., Westermarck, J., Koivisto, L., Kähäri, V.-M., and Heino, J. (1999) *J. Cell Biol.* **147,** 401–416
43. Setty, S., Kim, Y., Fields, G. B., Clegg, D. O., Wayner, E. A., and Tsilibary, E. C. (1998) *J. Biol. Chem.* **273,** 12244–12249
44. Xu, Y., Gurusiddappa, S., Rich, R. L., Owens, R. T., Keene, D. R., Mayne, R., Höök, A., and Höök, M. (2000) *J. Biol. Chem.* **275,** 38981–38989
45. Nichols, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296