Determinants of Ligand Binding Specificity of the \(\alpha_1\beta_1\) and \(\alpha_2\beta_1\) Integrins*

(Received for publication, July 28, 1999)

S. Kent Dickeson‡, Nancy L. Mathis‡, Mariam Rahman§, Jeffrey M. Bergelson§§, and Samuel A. Santoro¶

From the ‡Department of Pathology, Washington University School of Medicine, St. Louis, Missouri 63110 and the §Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104

The \(\alpha_1\beta_1\) and \(\alpha_2\beta_1\) integrins are cell surface collagen receptors. Cells expressing the \(\alpha_1\beta_1\) integrin preferentially adhere to collagen IV, whereas cells expressing the \(\alpha_2\beta_1\) integrin preferentially adhere to collagen I. Recombinant \(\alpha_1\) and \(\alpha_2\) integrin I domains exhibit the same collagen type preferences as the intact integrins. In addition, the \(\alpha_1\) integrin I domain binds echovirus 1; the \(\alpha_2\) I domain does not. To identify the structural components of the I domains responsible for the varying ligand specificities, we have engineered several \(\alpha_1/\alpha_2\) integrin I domain chimeras and evaluated their virus and collagen binding activities. Initially, large secondary structural components of the \(\alpha_2\) I domain were replaced with corresponding regions of the \(\alpha_1\) I domain. Following analysis in echovirus 1 and collagen binding assays, chimeras with successively smaller regions of \(\alpha_1\) I were constructed and analyzed. The chimeras were analyzed by ELISA with several different \(\alpha_2\) integrin monoclonal antibodies to assess their proper folding. Three different regions of the \(\alpha_1\) I domain, when present in the \(\alpha_2\) I domain, conferred enhanced collagen IV binding activity upon the \(\alpha_1\) I domain. These include the \(\alpha_3\) and \(\alpha_5\) helices and a portion of the \(\alpha_6\) helix. Echovirus 1 binding was lost in a chimera containing the \(\alpha_2\)-\(\alpha_6\) loop; higher resolution mapping identified Asn289 as playing a critical role in echovirus 1 binding. Asn289 had not been implicated in previous echovirus 1 binding studies. Taken together, these data reveal the existence of multiple determinants of ligand binding specificities within the \(\alpha_1\) and \(\alpha_2\) integrin I domains.

The integrins constitute a large family of cell adhesion molecules that are involved in both cell-cell adhesion as well as the adhesion of cells to the extracellular matrix (for review see Ref. 1). The integrins are involved in many important physiologic processes including development and differentiation, cell migration, wound healing, thrombosis and hemostasis, metastasis, and immune system function. Structurally, integrins are heterodimeric glycoproteins composed of two noncovalently associated integral membrane subunits. The \(\alpha\) subunits range in size from 120 to 180 kDa, and the \(\beta\) subunits range in size from 90 to 110 kDa. Many integrin subunits associate with more than one \(\alpha\) or \(\beta\) subunit resulting in a large number of different integrins.

The \(\beta\) subunit associates with at least 11 different \(\alpha\) subunits. Members of the resulting \(\beta_1\) subfamily of integrins are cell surface receptors for specific components of the extracellular matrix. The \(\alpha_1\beta_1\) integrin is a cell surface receptor for several different collagens and laminin-1 (2). \(\alpha_1\beta_1\) integrin-dependent adhesion of cells to collagens or laminin requires the presence of divalent cations (3). Similarly, the \(\alpha_2\beta_1\) integrin also serves as a cell surface receptor for collagens and laminins (for review see Ref. 4), and the adhesion of cells via the \(\alpha_2\beta_1\) integrin to these ligands also depends upon the presence of divalent cations (5). However, the substrate specificity of the \(\alpha_2\beta_1\) integrin depends upon the cell type on which it is expressed. The \(\alpha_2\beta_1\) integrin on platelets and fibroblasts is a collagen receptor (5); on endothelial and epithelial cells it is a receptor for both collagens and laminin-1 (6, 7). The \(\alpha_2\beta_1\) integrin also mediates echovirus 1 attachment and infection (8).

Near their amino termini, the \(\alpha_1\) and \(\alpha_2\) integrin subunits share (along with approximately half of the integrin \(\alpha\) subunits) an autonomously folding domain of approximately 220 amino acids known as the I (inserted) domain (for review see Ref. 9). The \(\alpha\) subunit I domains are critical determinants for ligand recognition and binding of both the \(\alpha_1\beta_1\) and the \(\alpha_2\beta_1\) integrins. Function blocking antibodies directed against both of these integrins map to their \(\alpha\) subunit I domain, and several mutations in the I domains adversely affect ligand binding activity (10, 11). Bacterially expressed I domains from both of these integrins bind collagens and laminin-1 in a manner that is saturable, blocked by function blocking antibodies against the parent integrin, and, like the parent integrins, dependent upon the presence of divalent cations (12–14).

The crystal structures of several different integrin I domains have been solved, providing much insight into I domain structure. I domain structures that have been solved include those of the \(\alpha_1\), \(\alpha_1\), and \(\alpha_2\) subunits (15–17). The three-dimensional structures of these three I domains are very similar. Each has a core that consists of a largely parallel \(\beta\) sheet structure, and the core is surrounded by several \(\alpha\) helices. All three contain a single divalent cation-binding site in a crevice near the top of the \(\beta\) sheet. Unique to the \(\alpha_2\) I domain, however, is an additional short \(\alpha\) helix at the top of the domain, in close proximity to the divalent cation-binding site (17).

Although cells expressing either the \(\alpha_1\beta_1\) or the \(\alpha_2\beta_1\) integrins adhere to several different collagens, they exhibit different relative affinities depending on the identity of the substrate. For instance, cells expressing the \(\alpha_1\beta_1\) integrin preferentially adhere to collagen IV, whereas cells expressing the \(\alpha_2\beta_1\) integrin preferentially bind to collagen I (18). Likewise, the binding of purified recombinant \(\alpha_1\) and \(\alpha_2\) integrin I domains to collagen types I and IV reflects the same relative affinities to these ligands as do the parent integrins (13). To identify the structural components of these I domains responsible for conferring the observed substrate specificities, we

* This work was supported in part by National Institutes of Health Grants AI35667, HL40506, HL63446, and DK45181. 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.

‖ Established Investigator of the American Heart Association.

¶ To whom correspondence should be addressed.

This paper is available on line at http://www.jbc.org
have engineered a series of α1β1/α2 integrin I domain chimeras and examined their ligand binding activities in collagen I and IV and echovirus 1 binding assays. Initially, with the aid of the crystal structure data mentioned above, relatively large structural components of the α2 integrin I domain were replaced with the corresponding region from the α1 integrin I domain. The chimeras were examined for both a gain of function (enhanced collagen IV binding activity) and a loss of function (echovirus 1 binding). Regions that satisfied one or both of these criteria were subjected to subsequent rounds of chimera construction, purification, and analysis to define further the critical structural and functional determinants.

**EXPERIMENTAL PROCEDURES**

Cloning and Mutagenesis of Integrin I Domain cDNAs—The cloning and expression of the α1 integrin I domain has been previously described (12). Briefly, cDNA encoding the I domain was amplified using the polymerase chain reaction (PCR) with the full-length α1 integrin cDNA as the template. This cDNA encodes Ser155–Met349 of the published α1 integrin sequence (19). In addition, a cDNA encoding a shorter I domain protein lacking the 35 amino-terminal amino acids was also prepared. The shorter protein, referred to as Δ1, contains Trp159–Met349. This protein lacks the DXXXS portion of the metal ion dependent adhesion site (MIDAS) motif (15). The PCR primers were designed such that both of the amplification products would contain a BglII restriction site at their 5′ ends and a stop codon followed by an XhoI restriction site at their 3′ ends. The PCR products were digested with BglII and XhoI, purified in agarose gels, and cloned into BamHI- and XhoI-digested glutathione-S-transferase (GST) fusion protein expression vector pGEX-5X-1 (Amersham Pharmacia Biotech). cDNA encoding the α1 integrin I domain was amplified by PCR using the full-length human α1 cDNA (Dr. Eugene E. Marcantonio, Columbia University) as the template. Analogous to the α2 integrin I domain construct, this cDNA encodes Ser155–Met349 of the published α1 integrin sequence (20). As with the α2 integrin I domain, the PCR primers were designed to create a product that contained a BglII site at the 5′ end and a stop codon followed by a XhoI site at the 3′ end. The PCR product was digested, purified, and cloned into pGEX-5X-1 as described above.

To facilitate the construction of chimeric α1/α2 integrin I domain cDNAs encoding the α2 integrin I domain with an internal region replaced with α1 integrin sequence, the cDNAs encoding both the I domains were transferred into pBluescript KS+ (Stratagene). A 681-bp EcoRI-XhoI fragment of pGEX-5X-1/α1, encoding the entire α1 I domain (amino acids 1–349) of the first seven amino acids, was isolated and ligated into pBluescript KS+ digested with EcoRI and XhoI. For the α2 I domain, a 1157-bp BglII-XhoI fragment of pGEX-5X-1/α2, encoding all of the α2 I domain and containing approximately two-thirds of the GST coding sequence, was isolated and ligated into pBluescript KS+ digested with SmaI and XhoI. Chimeric α1/α2 integrin I domain cDNAs were prepared by one of two methods, depending upon the length of the internal portion of α2 to be replaced with α1 sequence. An example of the construction of one of each type of chimera will be described. Chimera A, an example of one of the chimeras with a longer replacement, consists of the α1 I domain with an internal region from the α2 I domain consisting of the α2 helix and the α2β2 loop. First, pBluescript/α1 was mutated using Kunkel’s method (21) and oligo 1 as the mutagenic oligo (Table I). Oligo 1 introduced an Eco47III site into the α1 sequence at the amino-terminal end of the α1 helix to create p4. Simultaneously, pBluescript/α2 was mutated using oligo 2 as the mutagenic oligo. This introduced an EcoRV site at the amino-terminal end of the α2 helix (position analogous to the position of the Eco47III site in p4) to create p5. Eco47III and EcoRV were chosen so that the proper amino acids, serine and isoleucine, would be encoded by the 3′ end of the α1 fragment and the 5′ end of the α2 fragment, respectively, upon restriction enzyme digestion. The 3482-bp XhoI-Eco47III fragment of p4 and the 591-bp EcoRV-XhoI fragment of p5 were isolated and ligated to create p6, encoding amino acids 124–155 of α2 followed by amino acids 155–182 of α1, which is in turn followed by amino acids 184–349 of α1. It is important to note that although amino acids 155–158 of α1 encoded in this chimeric I domain, only amino acids 160–178 differ in the two I domains. This is due to the fact that amino acids 155–158 of α2 are identical to amino acids 156–160 of α1 and that amino acids 179–182 of α1 are identical to amino acids 180–183 of α2 (Figs. 1 and 3). Finally, the 966-bp Bst BI-XhoI fragment of p8 and the 4669-bp Bst BI-XhoI fragment of pGEX-5X-1/α1 were isolated and ligated to create the expression plasmid pGEX-5X-1/Ch A, capable of directing the expression of the GST-chimera A fusion protein.

Chimera B2, an example of one of the chimeras with a shorter internal portion of the α1 integrin I domain, consists of the α1 I domain with an internal region from the α2 I domain consisting of only the α3 helix, amino acids 205–210 (VLVAAK). Amino acids 206–211 (MIVATS) of the α2 I domain sequence were replaced with amino acids 205–210 (VLVAAK) of the α1 I domain sequence in a single mutagenesis reaction using pBluescript/α1 and oligo 5 (Table I) to create pBluescript/Ch B2. The 966-bp Bst BI-XhoI fragment of pGEX-5X-1/α1 and the 4669-bp Bst BI-XhoI fragment of pGEX-5X-1/α2 were isolated and ligated to create pGEX-5X-1/Ch B2. The sequences of the cDNAs used in this study, including all of the chimeras, were determined using the BigDye terminator cycle sequencing method (PE Applied Biosystems), and were compared with the published α1 and α2 integrin sequences (19, 20).

Expression and Purification of GST-I Domain Fusion Proteins—Trial inductions were performed to determine whether the selected clones could direct the expression of appropriately sized GST fusion proteins. Escherichia coli DH5α containing each of the plasmid constructs was grown at 37 °C in 5 ml of 2× YT medium supplemented with 0.2% glucose and 100 μg/ml ampicillin. Uninduced samples were removed from each culture after 1 h. Isopropyl-β-D-thiogalactoside was then added to a final concentration of 1 μM, and the cultures were returned to the incubator for 3 h to permit for accumulation of the expressed proteins. Cell lysates from the uninduced and induced samples were analyzed by SDS-polyacrylamide gel electrophoresis (22) followed by Coomassie Blue staining. Constructs that directed the expression of recombinant proteins of the expected size were used for large scale induction of protein expression.

For the purification of the fusion proteins, the inductions were performed as above except that the culture volume was increased to 500 ml and the expression was induced with 1 μM isopropyl-β-D-thiogalactoside when the A500 reached 0.3–0.4. At the end of the induction period, the cells were recovered by centrifugation at 2800 × g for 10 min, washed twice with 10 ml of ice-cold phosphate-buffered saline (10 mM Na2HPO4, 1.8 μM KH2PO4, 140 μM NaCl, 2.7 mM KCl, pH 7.4), and stored at −70 °C until needed. The GST-I domain fusion proteins were expressed, purified, and characterized as recently described (12).

**Collagen Binding Assays**—The wells of a 96-well microtiter plate (Immulon 2, Dynatech) were coated overnight at 4 °C with 0.1 ml/well of 30 μg/ml collagen I from calf skin (Sigma) or collagen IV from human placenta (Sigma) in 0.09% acetic acid. The wells were washed twice
a dihydrochloride (Sigma) was prepared according to the manufacturer’s instructions. 1:20,000 dilution of pig anti-goat IgG secondary antibody horseradish peroxidase conjugate (Roche Molecular Biochemicals) in the appropriate wash buffer and 0.1 ml of the appropriate wash buffer was added per well for 1 h at room temperature. The wells were again washed three times, and 0.1 ml of tetramethylbenzidine dihydrochloride (Sigma) prepared according to the manufacturer’s directions was added per well. After 15 min of substrate conversion, reactions were stopped with 0.025 ml of 4 N H2SO4 and the plates read at 450 nm using a Molecular Devices ELx808 microplate reader. Each protein was expressed and purified twice, and both preparations were tested in collagen I and IV binding assays. All data represent means of triplicate determinations.

**Echovirus 1 Binding Assays**—[35S]Methionine-labeled echovirus 1, prepared as described (23), was added to purified fusion proteins immobilized on glutathione-Sepharose (10,000 cpm, 2 µg of protein/µl) and then incubated for 1 h at room temperature with rocking. Beads were washed to remove unbound virus and then dissolved for scintillation counting. Additional aliquots of immobilized fusion proteins were boiled in Laemmli sample buffer (22) and analyzed by SDS-polyacrylamide gel electrophoresis to confirm equal protein loading.

ELISA—I domain-containing proteins were diluted to 10 µg/ml in TBS containing 2 mM MgCl2, and used to coat the wells of a 96-well microtiter plate (Immulon 2, Dynatech). Coating was carried out overnight at 4 °C with 0.1 ml of solution/well. The wells were washed twice with 0.15 ml of TBS containing 2 mM MgCl2 and then blocked for 1 h at room temperature with 0.15 ml of TBS containing 2 mM MgCl2. Following blocking the wells were washed once with 0.15 ml of wash buffer, and then 0.1 ml of primary antibody was added and allowed to interact for 1 h at room temperature. The wells were washed three times with 0.15 ml of wash buffer, then 0.1 ml of a 1:10,000 dilution of goat anti-mouse IgG secondary antibody horseradish peroxidase conjugate (Roche Molecular Biochemicals) in the appropriate wash buffer was added per well for 1 h at room temperature. The wells were again washed three times, and 0.1 ml of tetramethylbenzidine dihydrochloride (Sigma) prepared according to the manufacturer’s directions was added per well. After 10 min of substrate conversion, reactions were stopped with 0.025 ml of 4 N H2SO4 and the plates read at 450 nm using a Molecular Devices ELx808 microplate reader. Each protein was expressed and purified twice, and both preparations were tested in the ELISA assays. All data represent means of triplicate determinations.

**Results**

Binding of the α1β1 and α2β1 integrins to a variety of extracellular matrix molecules has been examined previously. The α1β1 integrin preferentially binds to collagen IV, whereas the α2β1 integrin preferentially binds collagen I (18). Analysis of the binding of the α1 and α2 integrin I domains to collagens I and IV using surface plasmon resonance and solid phase binding assays indicate that the substrate preferences of the α1 and α2 I domains reflect those of the parent integrins (13). We also have examined the binding of these I domains to collagens I and IV in solid phase substrate binding assays. In agreement with the above, we observed that the α1 I domain preferentially bound collagen IV, and the α2 I domain preferentially bound collagen I (Fig. 2A). The collagen type preferences of the α1 and α2 I domains were confirmed by more extensive studies over a wide range of I domain concentrations (see Fig. 7, left and middle panels).

The α2β1 integrin is also a cell surface receptor for echovirus 1 (8). Binding of echovirus 1 to the integrin is mediated by the α1 I domain (24). We tested the binding of echovirus 1 to both the α1 and α2 integrin I domains. As expected, the virus bound to the α2 I domain. However, the virus failed to bind to the α1 I domain (Fig. 2B). To determine which regions of the α1 I domain were responsible for enhanced collagen IV binding activity and which regions of the α2 I domain were required for echovirus binding, we designed a series of α1/α2 I domain chimeras and examined their collagen I and IV and echovirus 1 binding activities. As a guide for designing the chimeras, we used the recently published crystal structure of the α2 integrin I domain (17). Initially, relatively large portions of the α2 I domain were replaced with the analogous regions of the α1 I domain. Because loops and helices are exposed on the I domain surface, the initial round of chimeras contained loops...
and α helices from the α1 I domain. Fig. 3 (A–F) shows the initial round of six α1/α2 I domain chimeras and indicates the regions of the α2 I domain that were replaced with α1 I domain sequence.

Each of the chimeras was tested in collagen I and IV binding assays and the echovirus I binding assay. In addition, to assess the folded state of the I domain chimeras, each was tested by ELISA using a panel of five α2 I domain monoclonal antibodies. Because any given epitope could be lost because of replacement of the epitope with protein sequence from the α1 I domain, the chimeras were considered to be folded properly if a majority of the monoclonal antibodies bound. The results of these analyses are shown in Fig. 3. The collagen binding data are given as percentages ± S.E. of control (binding of the α2 I domain to collagen I); the echovirus binding data are also expressed as the percentages ± S.E. of control (binding of echovirus I to α2 I domain). Of the six chimeras, A and C failed to bind to collagen I, collagen IV, and echovirus I. Furthermore, neither of these chimeras was effectively recognized by any of the α2 I domain monoclonal antibodies. Therefore, we assume chimeras A and C are incorrectly folded, and no conclusions can be drawn from their lack of ligand binding. The remaining four chimeras, B, D, E, and F, were recognized by a majority of the antibodies and bound at least one of either collagen I, collagen IV, or echovirus I at a level comparable with or greater than that of the α2 I domain. Of these four chimeras, B, D, and E showed enhanced collagen IV binding activity. This represents a gain of function; these chimeras have become more α1 I domain-like with respect to their collagen binding activity. Also, of the four chimeras showing positive antibody reactivity, only chimera E lost the ability to bind echovirus I. This finding indicated an important role in echovirus I binding for the region of the α2 I domain consisting of the αC helix, the αC-α6 loop, and the α6 helix.

To further localize regions of the α1 I domain responsible for enhanced collagen IV binding activity and regions of the α2 I domain necessary for echovirus I binding, iterative rounds of chimera design, expression, purification, and analysis were undertaken. The design of each round was based on the results of the analysis of the previous round. Fig. 4 shows the resultant B series of chimeras. Because chimera B possessed enhanced collagen IV binding activity and contained both the βC-α3 loop and the α3 helix from the α1 I domain, chimeras B1 and B2 contained just the βC-α3 loop (and one residue from the amino-terminal end of the α3 helix), or the remainder of the α3 helix from the α1 I domain, respectively. The enhanced collagen IV binding activity of chimera B segregated completely into chimera B2 containing only the α3 helix from the α1 I domain. Chimera B1 reverted to the parental α2 I domain phenotype of preferential binding to collagen I. An attempt to further map the region of the α3 helix necessary for the enhancement (Fig. 4, chimeras B2a and B2b) resulted in reversion to the α2 I domain collagen binding properties, indicating the requirement of the entire α3 helix from the α1 I domain for the effect.

Chimera D, containing the α5 helix from the α1 I domain, as well as three amino-terminal flanking residues and one carboxy-terminal flanking residue, had also shown enhanced collagen IV binding activity. To determine which region was required for the effect, chimeras D1 and D2 were prepared. These chimeras contained the mutations GSMLKA → NHRLKK (three residues from the βD-α5 loop and the amino-terminal half of the α5 helix) and DQCNHD → QDCEDE (carboxy-terminal half of the α5 helix and one residue from the α5-βE loop). The results of collagen I and IV binding assays using these chimeras are shown in Fig. 5. Chimera D1, containing the amino-terminal half of the α5 helix of the α1 I domain, bound to both collagens I and IV considerably less effectively than did chimera D. Chimera D2, containing the carboxy-terminal half of the α5 helix of the α1 I domain, bound to collagen I as well as chimera D, but again the binding to collagen IV was diminished with respect to that of chimera D. Thus it appears that the entire α5 helix from the α1 I domain is required to observe the enhanced collagen IV binding activity.

Both enhanced collagen IV binding activity as well as complete loss of echovirus binding activity was observed for chimera E, containing the αC helix, the αC-α6 loop, and the α6...
The enhanced collagen IV binding activity segregated to the amino-terminal end of the α3 helix, or the remainder α3 helix, respectively. Chimeras B2a and B2b contained the mutations MI → VL and TS → AK within the α3 helix. Each of these chimeras was tested in the collagen I and IV binding assays. The binding data are given as percentages ± S.E. of control (α2 I domain binding to collagen I).

Chimera E2 showed both enhanced collagen IV binding activity as well as significantly diminished echovirus binding activity. Chimera E1, which contained a two-residue mutation in the αC helix (YL → SY), bound to both collagens I and IV. However, as opposed to the binding of both the α1 and α2 I domains as well as chimera E, the binding of chimera E1 to both collagens I and IV was independent of divalent cations. The binding of chimera E1 to collagen IV was not as pronounced as the binding of chimera E to collagen IV. Chimera E1α, which contained a single amino acid replacement (Y → S) within the αC helix, also bound both collagens I and IV in a divalent cation-independent manner. However, chimera E1R, in which the residues YL in the αC helix of the α2 I domain were reversed to LY, showed enhanced collagen IV binding activity and bound collagens I and IV in a divalent cation-dependent manner.

To localize the regions of E2 required for enhanced collagen IV binding activity and loss of echovirus binding, E2a, E2b, and E2c were prepared. (Fig. 6). The amino acid replacements in these chimeras are NA → GN, D → S, and KN → EK, respectively. The enhanced collagen IV binding activity segregated to chimera E2c, containing residues EK from the α1 I domain at the amino-terminal end of the α6 helix. Chimera E2c1, containing only the point mutation K → E at the amino-terminal end of the α6 helix, reverted to the collagen binding properties of the α2 I domain, indicating the importance of K294 for the enhanced collagen IV binding activity of chimera E2c. The loss of echovirus binding activity segregated completely to E2a, in which NA in the αC-α6 loop had been replaced. Chimeras E2a1 and E2a2, containing point mutations N → G and A → N in the αC-α6 loop, were prepared to determine which of the two residues from the α1 I domain present in E2a was responsible for the loss of echovirus 1 binding activity. This analysis revealed that Asn289 of the α2 I domain, within the αC-α6 loop, is required for echovirus binding.

Because chimeras E2 and E2a containing amino acids 288–294 and 288–289 of the α1 I domain had severely diminished echovirus 1 binding activity (E2) or exhibited in a complete loss (E2a) of echovirus 1 binding activity, reciprocal I domain chimeras were constructed in which amino acids 288–294 and 288–289 of the α1 I domain were replaced with the corresponding residues from the α2 I domain to determine whether either of these replacements could confer echovirus 1 binding activity upon the α1 I domain. Neither of these two chimeras, E2' or E2a', gained echovirus 1 binding activity (data not shown). However, these two chimeras also failed to bind collagens I and IV in a divalent cation-dependent manner and therefore may represent incorrectly folded species.

To gain a more complete understanding of the reciprocal collagen type specificities of the α1 I domain and the α2 I domain, as well as the cation-independent collagen binding activity of chimera E1, the binding of each of these proteins to collagens I and IV was measured over a wide range of I domain concentrations and in the presence of 1 mM EDTA or 2 mM Mg²⁺ (Fig. 7). The preferential binding of the α1 I domain to collagen IV and of the α2 I domain to collagen I was evident over the entire range of I domain concentrations. In addition, the divalent cation-independent binding of chimera E1 to collagens I and IV was confirmed over the entire range of protein concentrations tested.

Chimera F, containing the α7 helix from the α1 I domain, was recognized by a majority of the α2 integrin I domain monoclonal antibodies and echovirus 1 but showed significantly impaired collagen (both I and IV) binding activity. Three more chimeras, F1, F2, and F3, were prepared to ascertain which region of the α7 helix from the α1 I domain was necessary for the loss of collagen binding activity. Chimeras F1, F2, and F3 contained

---

**Fig. 4. Structure of B series of α1/α2 integrin I domain chimeras and ligand binding results.** Chimera B contained the βC-α3 loop and the α3 helix from the α1 I domain. Chimeras B1 and B2 contained just the loop and one residue from the amino-terminal end of the α3 helix, or the remainder α3 helix, respectively. Chimeras B2a and B2b contained the mutations MI → VL and TS → AK within the α3 helix. Each of these chimeras was tested in the collagen I and IV binding assays. The binding data are given as percentages ± S.E. of control (α2 I domain binding to collagen I).

**Fig. 5. Structure of D series of α1/α2 integrin I domain chimeras and ligand binding results.** Chimera D contained the α5 helix as well as three amino-terminal flanking residues and one carboxyl-terminal flanking residue from the α3 I domain. Chimeras D1 and D2 contained the amino-terminal mutations (GSMLKA → NHRLLK) and the carboxyl-terminal mutations (DQCNEF → QDCEDE) found in chimera D, respectively. Each of these chimeras was tested in the collagen I and IV binding assays. The binding data are given as percentages ± S.E. of control (α2 I domain binding to collagen I).
the mutations A → L, LEKAG → VTIVK, and Q → R within the α7 helix, respectively. Chimera F1 reverted to the collagen binding activity profile of the α2 I domain (Fig. 8). Chimera F2 showed somewhat diminished collagen I binding activity (in comparison to the α2 I domain) and severely impaired collagen IV binding activity. Of the chimeras F1–F3, F3 showed the

FIG. 6. Structure of E series of α1/α2 integrin I domain chimeras and ligand binding results. Chimera E contained the αC helix, the αC-α6 loop, and the α6 helix from the α1 I domain. Chimeras E1, E2, and E3 contained just the αC helix, the αC-α6 loop, or the α6 helix from the α1 I domain, respectively. Chimera E1a contained a point mutation (Y → S) within the αC helix; chimera E1R contained a reversal of two residues in the αC helix of the α2 I domain (YL → LY). Chimeras E2a, E2b, and E2c contain mutations in the αC-α6 loop (E2a and E2b) or at the amino-terminal end of the α6 helix (E2c). These mutations are NA → GN, D → S, and KN → EK, respectively. Chimeras E2a1 and E2a2 contain mutations in the αC-α6 loop, N → G and A → N, respectively. Chimera E2c1 contains a mutation in the amino-terminal end of the α6 helix, K → E. Each of these chimeras were tested in the collagen I and IV binding assays. The collagen binding data are given as percentages ± S.E. of control (α2 I domain binding to collagen I). Data marked with a lowercase letter reflect divalent cation-independent binding to collagen. Many of these chimeras were also tested for echovirus 1 binding. Echovirus binding data are given as percentages ± S.E. of control (echovirus 1 binding to the α5 integrin I domain).

FIG. 7. Binding of integrin I domain proteins to collagens I and IV. The binding of the α1 I domain, the α2 I domain, and chimeric I domain E1 to collagens I and IV was measured in a solid phase binding assay over a range of I domain concentrations from 1.6 to 400 nm. Binding was determined in the presence of either 1 mM EDTA or 2 mM Mg²⁺.
most impaired collagen I binding activity, only marginally higher than chimera F. Thus it appears that the loss of collagen binding activity seen in chimera F is a result of a combination of effects from the F2 region (LEKAG) of the α7 helix as well as the Q → R mutation near the carboxyl terminus of the α7 helix.

As mentioned earlier, five α5 integrin I domain monoclonal antibodies recognizing complex conformational epitopes were used as probes in ELISAs to assess the folded state of chimeric I domain molecules. The chimeras were considered to be properly folded if a majority of the antibodies were able to bind. All of the chimeras presented in this study were deemed properly folded with the exception of chimeras A and C in the first round and E2′ and E2a′.

New epitope mapping information for three of the five α2 integrin I domain monoclonal antibodies was obtained in this study from the analysis of their binding to the α1α2 integrin chimeras (Table II). These antibodies were 6F1, 12F1, and 5E8. These antibodies had previously been mapped to the same relatively large region (residues 173–259) of the I domain using human/bovine chimeric α5 I domain proteins (11). In our experiments, 6F1 binding was abrogated by mutation of the sequence TYKTK (residues 199–203) within the βC-strand (chimera B1). P1E6 recognition required DQCNHD (residues 268–273), the carboxyl-terminal half of the α3 helix (chimera B2b). The binding of 12F1 required the sequence QTIVK (residues 199–203) within the βC-strand (chimera B1). P1E6 recognition required DQCNHD (residues 268–273), the carboxyl-terminal half of the α3 helix (chimera B2b). With the exception of the P1E6-binding site, these regions are all within the range of amino acids previously specified as required for binding using the human/bovine chimeras. Amino acids 268–273, required for P1E6 binding in this study, are close in the primary structure to the region previously shown to be important for P1E6 recognition (11); it appears that P1E6 binding may require amino acids within both regions.

### DISCUSSION

Although purified α1β1 and α5β1 integrins each bind both collagen I and collagen IV, the α1β1 integrin exhibits a greater affinity for collagen IV than for collagen I, whereas the α5β1 integrin preferentially binds collagen I (18). BIAcore and solid phase binding studies using isolated α subunit I domains from these two integrins demonstrate that the collagen type specificities of the I domains are those of the parent integrins (13). The α5β1 integrin has been shown to be a cell surface receptor for echovirus 1 (8). The α2 I domain was shown to contain the echovirus 1-binding site (24). Murine α2 I domain fails to bind the virus (24) and human/murine α2 I domain chimeras have been used to identify some regions of the human α2 I domain required for echovirus binding (25).

To determine which regions of the α1 integrin I domain are responsible for enhanced collagen IV binding activity, and to identify additional regions of the α5 integrin I domain required for echovirus 1 binding, we undertook a systematic study involving the preparation and analysis of human α1/α2 integrin I domain chimeras. Initially, a series of six chimeras were constructed that contained primarily α2 sequence, each with a single internal region that had been replaced with the analogous residues from the α1 I domain. The design of the initial round of chimeras was undertaken with the aid of the recently published crystal structure of the α2 integrin I domain (17). For the most part, chimeras in the initial round (Fig. 3, A–F) contained a combination of secondary structural elements from the α1 I domain, for instance a helix and a loop, as in chimera A. Helices and loops were chosen, as opposed to strands, because of their surface exposure. Two of the chimeras (D and F) contained only a helix and chimera E contained two helices and a loop. Following the identification of chimeras with enhanced collagen IV binding activity or impaired echovirus binding activity, additional chimeras were constructed that had progressively smaller regions of replacement by α1 sequence. The pattern of chimera design, expression, purification, and analysis was repeated until the chimeras reverted to the α2 phenotype with regard to ligand binding or until the effect was isolated to a single amino acid.

The only chimera in the first round to lose echovirus 1 binding activity was chimera E. This chimera contained the αC helix, the αC-α6 loop and the α6 helix from the α1 I domain. Through iterative rounds of expression and analysis, the loss or severe impairment of echovirus binding was traced through chimera E2 (αC-α6 loop) and chimera E2a (residues NA in the loop) and finally to chimera E2a1 (Asn 289) (25). The fact that the loss of echovirus 1 binding can be traced through these four chimeras in the E series provides compelling evidence for the involvement of this region and Asn 289 in the interaction of echovirus 1 with the α1 I domain.

A similar approach has been previously used to map the echovirus 1-binding site on the α2 I domain (25). However, chimeras in the earlier study were human/murine α2 integrin I domain chimeras. Murine α2 integrin I domain fails to support echovirus 1 binding (24). Two regions of the I domain were

### TABLE II

| Protein | 6F1 | 5E8 | P1E6 |
|---------|-----|-----|------|
| GST     | 6.81±0.29 | 4.29±0.06 | 3.99±0.05 |
| α1 I    | 7.79±0.29 | 5.76±0.15 | 5.81±0.14 |
| α2 I    | 100±2   | 100±1  | 100±0  |
| B       | 7.52±0.07 | 7.21±0.08 | 88.7±1.1 |
| B1      | 78.1±1.9 | 10.5±0.3 | 77.4±2.5 |
| B2      | 8.40±0.02 | 98.5±0.22 | 99.8±1.8 |
| B2a     | 87.7±1.6 | 85.7±2.4 | 88.7±1.6 |
| B2b     | 5.28±0.07 | 98.6±0.5 | 97.1±0.3 |
| D       | 98.6±2.2 | 98.8±1.4 | 70.7±0.18 |
| D1      | 82.1±1.0 | 78.5±1.1 | 64.7±2.0 |
| D2      | 101±1   | 104±2  | 16.6±0.4 |

The data are given as percentages ± S.E. of the absorbance obtained with the α2 I domain.

**FIG. 8. Structure of F series of α1α2 integrin I domain chimeras, and ligand binding results.** Chima F contained the α7 helix from the α1 I domain. Chimeras F1, F2, and F3 contained residues from the amino-terminal third (A → L), the middle third (LEKAG → VTVIK), and the carboxyl-terminal third (Q → R) of the α7 helix of the α1 I domain, respectively. Each of these chimeras was tested in the collagen I and IV binding assays. The binding data are given as percentages ± S.E. of control (α2 I domain binding to collagen I).
shown to be required for maximal echovirus 1 binding activity. These regions consisted of amino acids 199–201 in the βC-α3 loop and amino acids 212, 214, and 216 in the α3-α4 loop. In a reciprocal experiment, replacement of the analogous residues within the mouse α2 I domain with human residues 199–216 conferred echovirus 1 binding upon the mouse α2 I domain. This provided unambiguous evidence for the role of this region of the α2 I domain in echovirus binding. A similar attempt was undertaken in this study involving the αC-α6 loop. Two chimeras were prepared in which portions of the α1 I domain results in chimeras with enhanced collagen IV binding activity, an α1 I domain-like quality. Tyr285, in the αC helix, when mutated to serine resulted in enhanced collagen IV binding activity and divalent cation-independent binding to both collagens I and IV. Residues 199–201 and 212–216 have previously been shown to be required for echovirus 1 binding using human/mouse domain chimeras (25). Based on this study, Asn289 is also required for echovirus 1 binding. In the figure on the right, the secondary structural elements are labeled (α helices 1, 3–7, and αC are shown in red, and β strands A–F are shown in blue).

Three of the chimeric α1/α2 integrin I domains (B, D, and E) in the first round of chimeras showed enhanced collagen IV binding activity, i.e., they exhibited a gain of function, becoming α1 I domain-like with respect to collagen IV binding. Chimera B contained the βC-α3 loop and the α3 helix from the α1 I domain. Chimera D contained the α5 helix, and chimera E contained the αC helix, the αC-α6 loop, and the α6 helix from the α1 I domain. As described above for the echovirus 1 binding studies, iterations of chimera design and analysis were undertaken to determine the minimal amount of α1 I domain sequence that would still confer the gain of function upon the α2 I domain, for each of these three regions.

For chimera B, the minimal sequence replacement that still conferred enhanced collagen IV binding activity was residues 205–210 of the α1 I domain sequence. This region contains the α3 helix and is the sequence present in chimera B2. Attempts
to further localize the portion of the a5 helix necessary for the enhanced collagen IV binding activity in chimera D led to the construction of chimeras D1 and D2. However, neither of these chimeras possessed collagen IV binding activity comparable with that of chimera D, and D1 also had significantly diminished collagen I binding activity. Thus, the enhanced collagen IV binding activity effect of chimera D required the entire a5 helix. Significant progress was made, however, at localizing the region within chimera E that was responsible for the enhanced collagen IV binding activity of this protein. Enhanced cation-dependent collagen IV binding activity segregated to chimera E2, containing only the aC-a6 loop from the a1 I domain (residues 288-294). The effect was further localized in chimera E2c, in which the sequence KN at the amino terminus of the a6 helix was mutated to EK, the analogous sequence from the a1 I domain. Somewhat surprisingly, chimera E2a1 (Asn289 → G), also exhibited increased collagen IV binding activity, although its immediate parent in the chimera series E2a (NA → GN) did not. These findings clearly identify an important role for residues of the aC-a6 loop and the amino-terminal portion of the a6 helix in establishing collagen type specificity.

Chimera F, containing the a7 helix from the a1 I domain, failed to bind either collagen I or IV. It was, however, recognized by all five of the a2 I domain monoclonal antibodies and was as effective as the wild-type a2 I domain at binding echovirus. Dissection of F revealed that two regions within the a7 helix contributed to the loss of collagen I binding activity. These regions, replaced in chimeras F2 and F3, consist of the amino acids LEKAG (residues 322-326) and Glu331 of the a2 I domain. Glu331 is most critical for collagen I binding by the a2 I domain. Mutation of this glutamate to arginine (chimera F3) resulted in a 50% reduction in collagen I binding activity.

Chimeras B2, D, and E2c each exhibited enhanced collagen IV binding activity. It is unlikely that amino acid replacements in these regions (a5 helix, a6 helix, and EK in the a6 helix) could each exert their effect through a direct contact with a rod-shaped molecule like collagen. It is more likely that the amino acid replacements within these regions induce the same or a highly similar conformational change at the site of collagen binding, leading to the enhanced collagen IV binding activity seen in chimeras B2, D, and E2c.

There are several lines of evidence supporting the existence of multiple I domain conformations revealing significant structural differences, some of which may lead to differences in ligand binding activities. For example, the a5 I domain has been crystallized in the presence of both Mg2+ and Mn2+ (28). These divalent cation-dependent differences in conformation of the a5 I domain may be responsible for observed differences in the in vitro ligand binding activity of this I domain in the presence of Mg2+ or Mn2+. Although the identity of the divalent cation has no effect on the binding of the a5 I domain to neutrophil inhibitory factor (29), binding to iC3b in the presence of Mn2+ is twice that observed in the presence of physiologic levels of Mg2+ and Ca2+ (30). Similarly, the extent of binding of the a5 I domain to collagen is identical in the presence of Mg2+ or Mn2+ (12), yet laminin binding is enhanced in the presence of Mn2+ as opposed to Mg2+ (14). Thus it is possible, perhaps even likely, that the enhancement of collagen IV binding activity observed with chimeras B2, D, and E2c is due to a conformational change in the I domain that is a direct result of the replacement of residues in the a3, a5, and a6 helices rather than to the alteration of critical residues that directly interact with collagen.

It is, however, also possible that some of the amino acids that were replaced in these chimeras do come into contact with collagen. In the model of the collagen-I domain complex (17), the collagen molecule is very near the carboxyl terminus of the a6 helix. It is possible, therefore, that residues 294 and 295 (Fig. 9), those replaced in chimera E2c, exert their effect through a direct contact with collagen.

In addition to the amino acid replacements in the a3, a5, and a6 helices, mutation of Tyr285 to either a serine (chimeras E1 and E1a) or to leucine (chimera E1R) also resulted in enhanced collagen IV binding activity of the a2 I domain. The Tyr → Ser mutation was the only change in chimera E1a. The presence or absence of this tyrosine in the a1 I domain, the a2 I domain, and chimeras E1, E1a and E1R correlates perfectly either with or enhanced collagen IV binding activity. This tyrosine is present within the aC helix of the a2 I domain and is positioned near the groove at the top of the I domain predicted to be the collagen-binding site (17). Fig. 10 shows a schematic of the model of the a2 I domain-collagen complex, looking directly down the collagen molecule. Side chains are included, and Tyr285 is shown in a space-filling representation (yellow). The tyrosine side chain and in particular the hydroxyl group come into extremely close contact with collagen. Our data predict that Tyr285 performs a “gatekeeper” function, composing part of the collagen-binding site on the a2 I domain that leads to preferential binding of collagen I over collagen IV. Additional evidence for such a role comes from analysis of chimera E1R, in which residues YL in the a6 helix were reversed to LY. Chimera E1R showed enhanced binding to collagen IV, while retaining the requirement for divalent cations. Thus, it is possible that replacement of amino acids in chimeras B2, D, and E2c leads to enhanced collagen IV binding activity by causing an alteration in the positioning of the side chain of Tyr285.

In summary, we have identified critical amino acid residues and structural motifs that determine the collagen type specificity of the a1 and a2 integrin I domains. It appears likely that...
collagen type specificity is established not only by residues that interact directly with collagen but also by more distant residues and motifs that interact with other structural elements of the I domain to indirectly, presumably by subtle alteration of the topography of the collagen-binding site, control selectivity. We have furthermore identified determinants that are essential for the echovirus 1 binding activity of the $\alpha_2$ integrin I domain. It is noteworthy that collagen type specificity and echovirus binding activity are determined by distinctly different structural features. It is also notable that several of the motifs that determine collagen type specificity presumably represent relatively large interactive surfaces or faces because they cannot be successfully reduced to single amino acid residues prior to reversion by the chimera to the parental I domain phenotype. This finding indicates the importance, indeed the necessity, of approaches such as that described here rather than the more traditional alanine mutagenesis approach.

Acknowledgments—We are grateful to Dr. Eugene E. Marcantonio for the gift of the $\alpha_1$ integrin cDNA. We thank Drs. Barry S. Coller, Virgil L. Woods, Jr., and Richard B. Bankert for the gifts of the monoclonal antibodies 6F1, 12F1, and 5E8, respectively. We are also grateful to Dr. Robert C. Liddington for providing the coordinates for the model of the I domain-collagen complex.

REFERENCES

1. Hynes, R. O. (1992) Cell 69, 11–25
2. Turner, D. C., Flier, L. A., and Carbonetto, S. (1989) J. Neurosci. 9, 3287–3296
3. Turner, D. C., Flier, L. A., and Carbonetto, S. (1987) Dev. Biol. 121, 519–525
4. Santoro, S. A., and Zutter, M. M. (1995) Thromb. Haemostasis 74, 813–821
5. Santoro, S. A. (1986) Cell 46, 913–920
6. Elices, M. J., and Hemler, M. E. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9906–9910
7. Languino, L. R., Gehlein, K. R., Wayner, E., Carter, W. G., Engvall, E., and Ruoslahti, E. (1989) J. Cell Biol. 109, 2455–2462
8. Bergelson, J. M., Shepley, M. P., Chan, B. M. C., Hemler, M. E., and Finberg, R. W. (1992) Science 255, 1718–1720
9. Dickeson, S. K., and Santoro, S. A. (1998) Cell. Mol. Life Sci. 54, 556–566
10. Kern, A., Briesewitz, R., Bank, I., and Marcantonio, E. E. (1994) J. Biol. Chem. 269, 22911–22916
11. Kamata, T., Puzon, W., and Takada, Y. (1994) J. Biol. Chem. 269, 9659–9663
12. Dickeson, S. K., Walsh, J. J., and Santoro, S. A. (1997) J. Biol. Chem. 272, 7661–7668
13. Calderwood, D. A., Tuckwell, D. S., Eble, J., Kuhn, K., and Humphries, M. J. (1997) J. Biol. Chem. 272, 12311–12317
14. Dickeson, S. K., Walsh, J. J., and Santoro, S. A. (1998) Cell Adh. Commun. 3, 273–281
15. Lee, J. O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995) Cell 80, 631–638
16. Qu, A., and Leahy, D. J. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 10277–10281
17. Emsley, J., King, S. L., Bergelson, J. M., and Liddington, R. C. (1997) J. Biol. Chem. 272, 28512–28517
18. Kern, A., Eble, J., Gobik, R., and Kuhn, K. (1993) Eur. J. Biochem. 215, 151–159
19. Takada, Y., and Hemler, M. E. (1989) J. Cell Biol. 109, 397–407
20. Briesewitz, R., Epstein, M., and Marcantonio, E. E. (1993) J. Biol. Chem. 268, 2989–2996
21. Kunkel, T. A. (1987) Methods Enzymol. 154, 367–382
22. Laemmli, U. K. (1970) Nature 227, 680–685
23. Bergelson, J. M., St. John, N. F., Kawaguchi, S., Chan, M., Stubdal, H., Modlin, J., and Finberg, R. W. (1993) J. Virol. 67, 6847–6852
24. Bergelson, J. M., St. John, N. F., Kawaguchi, S., Pasqualini, R., Berdichevsky, F., Hemler, M. E., and Finberg, R. W. (1994) Cell Adh. Commun. 2, 455–464
25. King, S. L., Kamata, T., Cunningham, J. A., Emsley, J., Liddington, R. C., Takada, Y., and Bergelson, J. M. (1997) J. Biol. Chem. 272, 28518–28522
26. Filman, D. J., Wien, M. W., Cunningham, J. A., Bergelson, J. M., and Hogle, J. M. (1998) Acta Crystallograph. 54, 1261–1272
27. Dickeson, S. K., Bhattacharyya-Pakrasi, M., Mathis, N. L., Schlesinger, P. H., and Santoro, S. A. (1998) Biochemistry 37, 11280–11288
28. Lee, J. O., Bankston, L. A., Arnaout, M. A., and Liddington, R. C. (1995) Structure 3, 1333–1340
29. Rieu, P., Ueda, T., Haruta, I., Sharma, C. P., and Arnaout, M. A. (1994) J. Cell Biol. 127, 2081–2091
30. Michishita, M., Videm, V., and Arnaout, M. A. (1993) Cell 72, 857–867