α11β1 Integrin Recognizes the GFOGER Sequence in Interstitial Collagens*

Received for publication, October 8, 2002, and in revised form, December 20, 2002
Published, JBC Papers in Press, December 20, 2002, DOI 10.1074/jbc.M210313200

The integrins αβ1, αβ2, αβ3, αβ4, and αβ5 are referred to as a collagen receptor subgroup of the integrin family. Recently, both αβ1 and αβ2 integrins have been shown to recognize triple-helical GFOGER (where single letter amino acid nomenclature is used, O = hydroxyproline) or GFOGER-like motifs found in collagens, despite their distinct binding specificity for various collagen subtypes. In the present study we have investigated the mechanism whereby the latest member in the integrin family, αβ1, recognizes collagens using C2C12 cells transfected with α11 cDNA and the bacterially expressed recombinant α11 I domain. The ligand binding properties of α11β1 were compared with those of αβ1, Mg2+-dependent α11β1 binding to type I collagen required micromolar Ca2+ but was inhibited by 1 mM Ca2+, whereas αβ1-mediated binding was refractory to millimolar concentrations of Ca2+. The bacterially expressed recombinant α11 I domain preference for fibrillar collagens over collagens IV and VI was the same as the α2β1 domain. Despite the difference in Ca2+ sensitivity, α11β1-expressing cells and the α11 I domain bound to helical GFOGER sequences in a manner similar to αβ1-expressing cells and the α2β1 domain. Modeling of the α1 I domain-collagen peptide complexes could partially explain the observed preference of different I domains for certain GFOGER sequence variations. In summary, our data indicate that the GFOGER sequence in fibrillar collagens is a common recognition motif used by αβ1, αβ2, and also α11β1 integrins. Although α10 and α11 chains show the highest sequence identity, α10 and α11 are more similar with regard to collagen specificity. Future studies will reveal whether αβ1 and α11β1 integrins also show overlapping biological functions.

The collagen family currently includes at least 24 members (1, 2), and four different collagen-binding integrins αβ1, αβ2, αβ3, and αβ4 are known. The αβ1 integrin does not interact directly with collagen, but it does act as a laminin receptor (5) that can affect the activity of the collagen receptor αβ1 through receptor cross-talk (6). αβ1, αβ2, αβ3, and αβ4 possess an inserted, or I domain† closely related to the von Willebrand factor A domain, which mediates binding to native collagens. In different in vitro assays, αβ2 also appears to be able to interact with type I collagen (7–9). However, this interaction most likely involves RGD motifs in denatured or partially unfolded collagen chains.

Studies of collagen-binding integrins in various in vitro assays show that they take part in cell adhesion, cell migration, control of collagen synthesis, matrix metalloproteinase synthesis, remodeling of collagen matrices, and influence such complex processes as cell proliferation, cell differentiation, angiogenesis, platelet adhesion/aggregation, and epithelial tubulogenesis (10, 11).

Using transfected cells and recombinant I domain, αβ1 has been shown to bind collagens, with a preference for collagens IV and VI over collagen I and II. Collagen XIII in vitro is also a ligand for αβ1 (12). Other identified ligands include laminin-1/2 (13, 14), the cartilage protein matrilin-1 (15), and the C-propeptide of collagen I (16). The affinity of αβ1 for laminin-1 has been reported to be about 10-fold lower than for collagen IV (17). In accordance with this, when the α1 integrin chain is expressed in K562 cells, αβ1 will bind type IV collagen, but it requires activation to bind laminin-1. In Chinese hamster ovary cells, αβ2 integrin does not mediate spreading on collagen II (12). It has been suggested that in these cells a coreceptor is needed for αβ2-mediated spreading on collagen II.

Integrin αβ1 and its α I domain have been shown to bind a variety of collagens (19–21), the C-propeptide of collagen I (16, 22), laminin-1 (23), laminin-2 (14), decorin (24), and the cartilage protein chondroadherin (25). Unlike the other ligands for αβ1 and αβ2, chondroadherin does not support cell spreading. Early studies using antibodies showed that αβ1 on some cells (melanoma LOX cells), but not others (fibroblasts, platelets), mediated the binding to laminin-1 (26).

The integrin subunit α10 was originally identified by affinity purification of collagen type II-binding integrins from adult chondrocytes (3). The rather restricted expression of α10β2 to cartilage indicates that the ligands are to be found in the cartilage extracellular matrix. Intriguingly, using recombinant protein, the collagen binding preference of the α10 I domain is

* This work was supported by grants from the Swedish Research Council (to D. G.), the Medical Research Council (to R. W. F.), the British Heart Foundation (to R. W. F.), the Wenner-Gren Foundation (to W.-M. Z.), and the Kung Gustaf V:s 80-års Fond (to D. G.). 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. Tel.: 46-18-471-4175, Fax: 46-18-471-4673; E-mail: donald.gullberg@imbim.uu.se.

© 2003 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in U.S.A.

1 Integrin Recognizes the GFOGER Sequence in Interstitial Collagens*

From the †Department of Medical Biochemistry and Microbiology, Biomedical Center, Box 582, Uppsala University, Uppsala S-751 23, Sweden, the ‡Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä FIN-40351, Finland, and the §Department of Biochemistry and Pharmacy, Åbo Akademi University, Turku FIN-20520, Finland, and the ¶Department of Biochemistry, University of Cambridge, Bldg. O, Downing Site, Tennis Court Rd., Cambridge CB2 1QW, United Kingdom.
most similar to that of the α1 I domain, so that the α11 I domain prefers the basement membrane collagen IV and the beaded filament-forming collagen VI over the interstitial collagens I and II (27). In the same study, mutational analysis of the I domains showed that the amino acid residues Arg-218 in α1 and α11 and Asp-219 in α2 are involved in determining this collagen preference.

α11 was initially detected in differentiating human fetal muscle cells (28). α11 protein and mRNA expression analysis in human embryos, however, revealed that expression is localized to mesenchymal non-muscle cells in areas of highly organized interstitial collagen networks. No expression was seen in muscle cells in vivo (29). In the developing skeletal system, α1β1 and α11β1 thus show nonoverlapping, complementary expression patterns (11). In accordance with the expression of α11β1 in areas rich in interstitial collagens, α11β1 binds more efficiently to collagen I than to collagen IV (29).

Cyanogen bromide cleavage of collagen chains identified the non-RGD-containing helical CB3 fragment of collagen α1 I as a cell-binding fragment that could be used to purify α1β1 (30). The α1 I and α2 I integrin binding site located within triple-helical α1 I CB3 has been identified as GFOGER (31, 32). Two related sequences, GLOGER2 and GASGER, were identified elsewhere in collagen I (33), and other GER-containing sequences in the collagen chains can also mediate cell adhesion through α1β1.3 The GER motif thus appears to be a major cell adhesion motif used by collagen-binding integrins.

Examination of the crystal structure of an α2 I domain-GFOGER complex suggested that other hydrophobic residues might replace phenylalanine, which together with the glutamate residue of the hydroxyproline group of the hydroxyproline residue, suggesting that hydroxyproline itself may not be required specifically for collagen-integrin interaction.

Arginine interacts with negatively charged Asp-219 on the surface of the α1 I domain, and although this appears a relatively nonspecific interaction, GEK will not substitute fully in human platelets. The ligand binding groove of the integrin α2 I domain-ligand complexes could in part explain the observed differences in α2 I and α11 I domain binding to different collagen peptides. The results are potentially promising for future attempts to generate reagents effective in blocking multiple collagen-binding integrins simultaneously.

MATERIALS AND METHODS

Production of Human Recombinant Integrin α1, I, α2, I, and α11, I Domains as Fusion Proteins

cDNAs encoding α1 I and α2 I domains were generated by PCR as described earlier (27) using human integrin α1 and α2 cDNAs as templates. Vectors pGEX-4T-3 and pGEX-2T (Amersham Biosciences) were used to generate recombinant glutathione S-transferase (GST) fusion proteins of human α1 I and α2 I domains, respectively. Human integrin α11 I cDNA (4) was used as a template when the α11 I domain was generated by PCR. The PCR product having BamHI and EcoRI sites was cloned to pGEX-KT, and the DNA sequence was checked by sequencing the whole insert. The same vector was used for expression of recombinant GST fusion proteins of the human α11 I domain. Competent Escherichia coli BL21 cells were transformed with the plasmids for protein production. 500 ml of LB medium (Biokar) containing 100 μg/ml ampicillin was inoculated with a 50-ml overnight culture of BL21/pOα1 I, BL21/pOα2 I, or BL21/pOα11 I, and the cultures were grown at 37 °C until the A600 of the suspension reached 0.6–1.0. Cells were induced with isopropyl-1-thio-β-d-galactopyranoside and allowed to grow for an additional 4–6 h before harvesting by centrifugation. Pelleted cells were resuspended in PBS (pH 7.4) and then lysed by sonication followed by the addition of Triton X-100 to a final concentration of 2%. After incubation for 30 min on ice, suspensions were centrifuged, and supernatants were pooled. Glutathione-Sepharose (Amersham Biosciences) was added to the lysate, which was incubated at room temperature for 30 min with gentle agitation. The lysate was then centrifuged, the supernatant was removed, and glutathione-Sepharose with bound fusion protein was transferred into disposable chromatography columns (Bio-Rad). The columns were washed with PBS, and fusion proteins were eluted using 30 mm glutathione. Purified recombinant and glutathione-tagged α1 I, α2 I, and α11 I domains were analyzed by SDS and native PAGE. The recombinant α1 I domain produced was 227 amino acids in length, corresponding to amino acids 129–358 of the whole α1 integrin, whereas the α2 I domain was 223 amino acids long, which

2 Where single letter amino acid nomenclature is used, O = hydroxyproline.
3 R. W. Farndale, P. R.-M. Siljander, and C. G. Knight, in preparation.
4 R. W. Farndale, P. R.-M. Siljander, and C. G. Knight, unpublished observation.
corresponded to amino acids 124–329 of the whole α2β1 integrin. The carboxyl termini of the α1 and α2 domains contained 10 and 6 non-integrin amino acids, respectively. Recombinant α1, I domain contains a total of 204 amino acids: at the amino terminus there are 2 extra residues (GS) before the α1, I domain, which starts from CQTY and ends with SLEG (residues 159–354); at the carboxyl terminus there are 6 extra amino acids (EFIVTD). The recombinant α1, I domain contains some residues (GS) before the α2 domain as an impurity caused by endogenous proteolytic activity during expression and purification. Recombinant I domains were used as GST fusion proteins for collagen binding experiments.

**Synthesis of Peptides**

Peptides were synthesized as carboxyl-terminal amides on TentaGel R RAM resin in a PerSeptive Biosystems 9050 Plus PepSynthesizer express machine and purified by reverse phase high performance liquid chromatography (HPLC) on a column of Ydc 219TP101522 using a linear gradient of 5–45% acetonitrile in water containing 0.1% trifluoroacetic acid. Fractions containing homogeneous product were identified by analytical HPLC on a column of Ydc 219TTP54, pooled, and freeze dried. All peptides were found to be of the correct theoretical mass by mass spectrometry. The triple-helical stability of each peptide was assessed by polarimetry as described previously.

**Solid Phase Binding Assay for α1, I, α2, I, and α11, I Domains**

The coating of a 96-well high binding microtiter plate (Nunc) was done by exposure to 0.1 ml of PBS containing 5 μg/ml (15 μg/ml) collagens or 20 μg/ml synthetic triple-helical collagen peptides overnight at 4°C. Rabbit (rat tail) collagen, type IV mouse (membrane of Engelbreth-Holm-Swarm mouse sarcoma) collagen, and type IV human "cut" (human placenta) collagen were purchased from Sigma. Type IV human collagen and type II bovine collagen were purchased from Biodiagnostics International and Chemicon, respectively. Type I bovine (bovine dermal) collagen was from Cellon S. A. Blank wells were coated with a 1:1 solution of 0.1 ml Delfia® Dilent II (Wallac) and PBS. Residual protein absorption sites on all wells were blocked with a 1:1 solution of 0.1 ml of Delfia® Dilent II and PBS. Recombinant proteins α1, I-GST, α2, I-GST, and α11, I-GST were added to the coated wells at the desired concentration in Delfia® assay buffer and incubated for 1 h at room temperature. Europium-labeled anti-GST antibody (Wallac) was then added (typically 1:1,000), and the mixtures were incubated for 1 h at room temperature. All incubations mentioned above were done in the presence of 2 mM MgCl2. Delfia® enhancement solution (Wallac) was added to each well, and the europium signal was measured by time-resolved fluorometry (Victor2 multilabel counter, Wallac). In every case, at least three parallel wells were analyzed.

**Cells**

Murine C2C12 myoblasts from the American Type Culture Collection were provided by A. Starzinski-Powitz. The generation of C2C12 cells stably transfected with integrin α2, I DNA or integrin α11, I DNA has been described previously (29). Cells were cultured at 37°C in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum and antibiotics (Statens veterinärmedicinska anstalt, Uppsala). The cells were grown to subconfluence and passed every 2–3 days.

**Antibodies**

Rabbit antibodies to the cytoplasmic tail of α11 integrin have been described previously (4). To immunoprecipitate β1 integrins, a polyclonal antibody to rat integrin β1 chain was used (38).

**Immunoprecipitation and Electrophoresis**

Cell cultures were washed three times in Dulbecco’s modified Eagle’s medium devoid of cysteine and methionine and metabolically labeled overnight in the presence of 25 μCi/ml [35S]methionine/cysteine (pre-Mix®7S cell labeling mix; Amersham Biosciences). Proteins were extracted from the culture dish in lysis buffer (1 ml of solubilization buffer (1% Triton X-100, 0.15 M NaCl, 20 mM Tris-HCl pH 7.4, 1 mM MgCl2, 1 mM CaCl2) containing protease inhibitors (1 mM Pefabloc SC (Roche Molecular Diagnostics), 1% aprotinin, 1 μg/ml pepstatin, 1 μg/ml leupeptin). Solubilized proteins were centrifuged for 10 min at 15,000 g. The centrifuged supernatant was precleared by incubating with 100 μg/ml preimmune IgG and protein A-Sepharose CL-4B (Amersham Biosciences) for 2 h. After centrifugation, immune IgG was incubated with the extract for 2 h. Specifically bound proteins were recovered with protein A-Sepharose. The precipitate was washed three times with buffer A (1% Triton X-100, 0.5 M NaCl, 20 mM Tris-HCl pH 7.4, 1 mM MgCl2, 1 mM CaCl2) and three times with buffer B (0.1% Triton X-100, 0.15 M NaCl, 20 mM Tris-HCl pH 7.4, 1 mM MgCl2, 1 mM CaCl2) prior to solubilization in electrophoresis sample buffer. Proteins were separated on 6% SDS-polyacrylamide gels and processed for fluorography.

**Cell Attachment Assay**

**General Setup**—Ligands and metal ions are described separately for the two experiments. 24-well culture plates (Nunc) were coated with ligands (500 μl to a 2-cm2 well) diluted in PBS overnight at 4°C, followed by blocking with 2% BSA in PBS for 2 h at room temperature and then washed in Puck’s saline (137 mM NaCl, 5 mM KCl, 4 mM Na2CO3, 5.5 mM D-glucose, pH 7.0). Transfected cells were trypsinized, washed three times in Puck’s saline, deeded into the wells at a concentration of 250,000 cells/well, and were allowed to attach for 45 min at 37°C and 5% CO2. Wells were washed three times in Puck’s saline, and plates were rapidly frozen at −20°C for later assay using the hexose-aminidase test as described previously (29). For each cell line used, a cell number standard was made. Each experiment was performed in triplicate. To minimize errors from unequal trypsinization stress between cell lines and handling of plates, for example, data were normalized as follows. For each plate the adhesion to 10 μg/ml fibronectin (provided by S. Johanson, Upsala University) was used as the 100% reference level, and the background found on BSA-only coated wells was used as the base-line (0%) reference level.

**C2** over-7272

**Inhibition Set-Up**—Wells were coated with 10 μg/ml bovine collagen type I (Vitrogen®100, Cohesion) or 10 μg/ml human plasma fibronectin. Wells were filled with Puck’s saline, and MgCl2 + EGTA was added to obtain a final concentration (after addition of cells) of 2 mM MgCl2 and 0.01 mM EGTA. CaCl2 was added according to Fig. 1.

**Cell Attachment to Synthetic Peptides**—24-well plates were coated with 10 μg/ml synthetic triple-helical collagen peptides at 4°C overnight according to Fig. 4, or 10 μg/ml bovine collagen type I (Vitrogen®100), or 10 μg/ml fibronectin. MgCl2 and CaCl2 were added to a final concentration of 2 mM MgCl2 and 0.01 mM CaCl2.

**Homology Modeling**

Sequences of integrin α1 (accession code PS56199 (39)) and α11, (Q9UKX5 (40)) were obtained from SWISS-PROT (41). The crystal structure of the α1, I domain in complex with the triple-helical collagen mimetic peptide (PDB code 1dzi (34)) was obtained from the Protein Data Bank (42).

The sequence alignment was made using the program MALIGN (43) in the BODIL modeling environment (www.abo.fi/fak/mnf/bkf/search/johnson/bodil.html) using a structure-based sequence comparison matrix (44) with a gap penalty of 40.

Homology models were built with HOMODEG in BODIL. The amino acid side chain rotamer library incorporated into BODIL was used to evaluate alternative possibilities for side chain conformations for sequence differences in the alignment of α1 and α11, I domains sequences with the template structure.

In the α1, I domain-peptide complex structure, 1dzi, and the three peptide chains (identical in sequence but having different interactions with the α1, I domain) of the collagen mimetic tripeptide are labeled B, C, and D. The corresponding chain labels are used for the tripeptides docked to the model structures built for the α1, I and α11, I domains.

**RESULTS**

**Influence of Ca2+ on Cell Attachment to Collagen**—To compare the mechanism whereby α1β1 and α11β1 recognize collagens, we used the satellite cell line C2C12 transfected with α2 or α11 cDNAs (C2C12 α2β1 and C2C12 α11β1, respectively). The parental cell line C2C12 expresses members of the β1 subfamily, such as α2β1 and α11β1 (45), but does not adhere to collagen (29). C2C12 cells transfected with either α2 or α11 acquire the ability to interact with collagens I and IV, with a preference for collagen I (29). The α2β1-mediated binding of platelets to collagen has been reported to require micromolar Ca2+ but to be inhibited by millimolar Ca2+ in the presence of Mg2+ (37). To test the effect of Ca2+ on cell adhesion to collagen I, the transfected C2C12 cells were plated on collagen I in the presence of Mg2+ and EGTA with increasing concentrations of Ca2+ added.

5 J. V. Lehtonen, V. V. Rantanen, D. J. Still, M. Gyllenberg, and M. S. Johnson, unpublished observation.
In the absence of Ca$^{2+}$, adhesion of cells expressing $\alpha_{11}\beta_1$ was virtually absent, as reported for human platelets (45), and a biphasic response to added Ca$^{2+}$ was observed with peak adhesion occurring at a free Ca$^{2+}$ of around 30 $\mu$M and adhesion being substantially abolished at 4 mM, with an IC$_{50}$ of about 1 mM. In marked contrast, C2C12 $\alpha_2\beta_1$ cell adhesion to collagen I was largely refractory to both the removal of Ca$^{2+}$ using EGTA or the subsequent addition of millimolar Ca$^{2+}$ (Fig. 1).

$\alpha_1\beta_1$ and $\alpha_2\beta_1$ I Domains Differ in Their Affinity for Collagen I—To estimate the $K_d$ values for $\alpha_{11}$ for collagen I, we produced an $\alpha_{11}$ I domain in *E. coli*. Initial attempts to express the $\alpha_{11}$ I domain as a bacterial GST fusion protein yielded low amounts of protein. We therefore expressed the $\alpha_{11}$ I domain as a His-tagged fusion protein with a Missouri active site upstream of the GFOGER-like peptide (33) in *Pichia pastoris*. Production of this protein was subsequently optimized. Large scale expression of the $\alpha_{11}$ I domain yielded enough protein to perform binding studies. Approximated $K_d$ values based on solid state binding assays (26) and the use of the Michaelis-Menten equation suggested a relatively low $\alpha_{11}$ I domain avidity to collagen I (750 ± 50 nM) when compared with the $\alpha_2$ I domain binding to collagen I (20 ± 5 nM) (Fig. 2).

Collagen Preference of $\alpha_{11}$ I Domain—Previous studies from several laboratories have shown that $\alpha_\beta_1$ integrins prefer fibril-forming collagens over network-forming type IV and beaded filament-forming type VI collagen. The same pattern can be seen in the binding of the $\alpha_2$ I domain. Here $\alpha_2$ and $\alpha_{11}$-mediated binding to different collagens was compared. The $\alpha_{11}$ I domain was shown to prefer the fibril-forming collagen types I and II (Fig. 3), whereas its binding was weaker to type III (data not shown), a member of the same collagen subgroup. Collagens IV and VI were poor ligands for the $\alpha_{11}$ I domain. Thus, in the terms of its binding pattern the $\alpha_{11}$ I domain was closer to the $\alpha_2$ I domain than to either the $\alpha_1$ or $\alpha_{10}$ I domain.

Binding to GER-containing Peptides—Helical GFOGER and GFOGER-like peptides have recently been shown to represent high affinity integrin recognition motifs in collagens (32, 33). To determine whether $\alpha_{11}\beta_1$ also differed from $\alpha_\beta_1$, with regard to its recognition sequences, C2C12 $\alpha_2\beta_1$ and C2C12 $\alpha_{11}\beta_1$ cells were tested for their ability to attach to different collagen-like peptides. C2C12 $\alpha_2\beta_1$ cells adhered to GFOGER and GFO-
GPOGES, from the collagen I α2 chain where it corresponds to GFOGER in the α1 chain, was similarly without significant binding activity. Comparing the overall peptide binding patterns, α2-I and α11-I domains appear most similar in their peptide binding preferences.

Modeling of Collagen Peptide Binding to the α11-I Domain—The basic assumption in modeling was that the collagen mimetic tripeptide GFOGER and its mutants bind to all of the integrin α1 domains in a way similar to that seen in the crystal structure of the complex between α2-I domain and the GFOGER triple-helical peptide (34). The sequence identities of the α1 and α11-I domains to the α2-I domain are 51 and 45% respectively, thus experience dictates that high quality models will be produced. Only one region of the α11-I domain model is uncertain, where Pro-310 (threonine in the α1 and α2-I domains) is located within a region that corresponds to helix 6 of the open fold of the α1 and α2-I domains. Proline generally does not promote helix stability, so it is very likely that the helix begins at or after position 310 in the α11-I domain. In addition, the local alignment of residues 179 and 180 seems peculiar because the charged residue Glu-180 would be buried, and the hydrophobic residue Val-179 would be exposed toward the solvent. If Glu-180 is buried, then the conserved residue Tyr-157 may change its conformation in the α11-I domain and affect the binding of the collagen mimetic tripeptide. Thus, it is possible that the binding conformation seen in α2-I domain-tripeptide

FIG. 4. α11β1 binds the helical GFOGER sequence. C2C12, C2C12 α2−, and C2C12 α11− cells were allowed to adhere to synthetic collagen peptides in the presence of 1 mM Mg2+ and 10 μM Ca2+, and cell adhesion (triplicate wells) was evaluated (±S.D.).

FIG. 5. Binding of α1, α2, I, and α11-I domains to synthetic triple-helical collagen peptides. Microtiter plates were coated with 20 μg/ml collagen peptides. Diluent II containing BSA was used as a background control and to block the wells. GST-fusion α1-I (A), α2-I (B), and α11-I domains (C) were allowed to bind for 1 h in the presence of 2 mM MgCl2. The wells were washed three times. Bound α-I domains were detected with europium-labeled anti-GST antibody. Time-resolved fluorescence measurements were used. The data are the means of three parallel determinations (±S.D.).


**Arginine in the Collagen Mimetic Tripeptide—**In the crystal structure of the $\alpha_2$ I domain model, Arg-218 (aspartate in $\alpha_2$ and threonine in $\alpha_{11}$) may form an additional interaction with the glutamate of chain D of the collagen mimetic tripeptide. This interaction would still be possible in the tripeptide containing the aspartate mutation.

In the $\alpha_1$ I domain model, Arg-218 (aspartate in $\alpha_1$ and threonine in $\alpha_{11}$) may form an additional interaction with the glutamate of chain D of the collagen mimetic tripeptide. This interaction would still be possible in the tripeptide containing the aspartate mutation.

**Glutamate in the Collagen Mimetic Tripeptide—**In the crystal structure of the integrin $\alpha_1$ I domain in complex with the collagen mimetic peptide (3 × GFOGER), the side chain of only one of the glutamate residues in the collagen mimetic tripeptide, that of the middle strand, chain C, interacts with the I domain. This glutamate is coordinated to the metal ion of the MIDAS motif and thus, represents a key interaction in tripeptide binding (Fig. 6). Moreover, even a conservative change, mutation to aspartate, lowers the binding dramatically for the peptide binding (Fig. 6A).

In both $\alpha_1$ and $\alpha_{11}$, the arginine of the tripeptide could form a salt bridge with glutamate at the position equivalent to Asp-189 in $\alpha_2$ (Fig. 6, B and C). Furthermore, in $\alpha_{11}$, there is a threonine equivalent to Asp-219 in $\alpha_2$ whose side chain hydroxyl group can accept a hydrogen bond from the Ne of arginine from the tripeptide (Fig. 6C). For $\alpha_2$, the Arg → Lys mutation in the collagen mimetic tripeptide (GFOGEC) does not affect binding as dramatically as seen for $\alpha_1$ and $\alpha_{11}$ (Fig. 5). In $\alpha_2$, the repulsion resulting from the charged amino group of lysine positioned near the Leu-220 side chain would be offset by the formation of a somewhat more optimal hydrogen bond/salt bridge between lysine and Asp-219 (Fig. 6D). The mutation of arginine to lysine reduces the binding affinity of collagen mimetic tripeptide to the $\alpha_1$ and $\alpha_{11}$ I domains because the salt bridge to glutamate, at the position equivalent to Asp-189 in $\alpha_2$ cannot be maintained. When arginine of the collagen mimetic tripeptide is mutated to lysine, lysine cannot reach glutamate because a lysine residue is shorter than an arginine residue. Moreover, in $\alpha_{11}$ the lysine residue can form a hydrogen bond with the threonine equivalent to Asp-219 in $\alpha_2$, and thus, the effect of the mutation is not as dramatic as for $\alpha_1$.

In the $\alpha_2$ I domain, the arginine from chain B of the collagen mimetic tripeptide interacts mainly with other parts of the tripeptide and not with the I domain. Ne is hydrogen-bonded to the main chain oxygen of arginine in chain C, and the planar end of the arginine side chain has a hydrophobic interaction with proline in peptide chain C. In addition, hydrophobic interactions with the hydrophobic part of the Glu-256 side chain and weak electrostatic interactions with the main chain oxygen atom of Ser-257 can be seen. These interactions should be present and identical in each of the I domains in this study. Thus, the effect of the mutation of arginine to lysine, caused by chain C, should be same for all I domains.

In the $\alpha_2$ structure, the arginine from chain D of the collagen mimetic tripeptide is exposed to the solvent, and thus, the mutation can only have an indirect influence on I domain binding.

**Phenylalanine in the Collagen Mimetic Tripeptide—**In the $\alpha_3$ I domain structure, the phenyl ring of phenylalanine in chain B of the collagen mimetic tripeptide is stacked with the phenol ring of the conserved tyrosine (position 157 in $\alpha_3$). This phenylalanine also has hydrophobic interactions with Leu-286 in $\alpha_3$.

In addition, the phenylalanine of chain B forms an unfavorable interaction with the main chain oxygen atom of Tyr-285 (Fig.

---

**Fig. 6. Stereo view of the interactions between the arginine from chain C of the collagen mimetic tripeptide and integrin $\alpha_2$ (1dzi) (A), $\alpha_1$ (B), and $\alpha_{11}$ (C) I domains.** Detailed interactions of Arg → Lys mutation of the collagen mimetic tripeptide with the integrin $\alpha_1$ I domain in the same region are shown in D. The backbone of the collagen mimetic tripeptide is shown as yellow dotted lines, and the I domain backbone is shown in blue. Strong electrostatic interactions are shown as black dotted lines, and weak electrostatic interactions are shown as yellow dotted lines.
ligand binding.

When phenylalanine of the collagen mimetic tripeptide is mutated to leucine, some favorable interactions would be lost, but this loss is offset by the removal of unfavorable interactions with the main chain oxygen atom of the residue at position 285 (tyrosine in α2 and α11; serine in α1). Thus, there is a small change in the binding affinity of α1 and α2 when the Phe → Leu mutant is compared with the “wild type” tripeptide (Fig. 5). The effects seen for α11 are difficult to predict because the model is inaccurate in this region. The binding affinity is lowered dramatically when phenylalanine is replaced with alanine, resulting in the loss of all favorable interactions (Fig. 5).

In the α1 I domain structure, the phenylalanine in chain C of the collagen mimetic tripeptide leans against the side chain of Asn-154, which is conserved in the α1, α9, and α11 I domains. This interaction is not very critical, and thus the mutation of phenylalanine to leucine or alanine would not affect the binding affinity by much. The phenylalanine in chain D is exposed to the solvent, and thus it has practically no role in binding.

**DISCUSSION**

In recent years an increasing effort has been spent trying to understand the mechanism whereby cells bind collagen. In vertebrates more than 24 different collagens exist, and the role of some of these is yet unclear. Integrins are major receptors for collagens. A common feature of the collagen-binding integrins is the presence of an α I domain that is directly involved in ligand binding.

The I domain is not found in integrin α chains from the invertebrate *Drosophila melanogaster* but it is present in 9 of the 18 currently known vertebrate integrin α chains (11) including αL, αM, αX, αD, and αE, which are all involved in different aspects of leukocyte functions and pair exclusively with the β2 subunit (10). The overall importance of integrin-mediated cell-collagen interactions involving the α11, α22, α10, and α11 integrin chains is largely unknown because of the limited information available for α10β1 and α11β1. Based on the appearance of I domain integrin α chains during vertebrate evolution it is possible that these integrin chains play important roles in vertebrate-specific structures of the musculoskeletal system.

Gene knockout experiments and recombinant expression of the α I domains have yielded considerable information about the characteristics and functions of collagen-binding integrins α1β1 and α2β1 (11, 46, 47). Phylogenetically, α1 and α2 form a subfamily distinct from α10 and α11, which most likely have formed through two distinct gene duplication events.

Both α1 and α2 integrin chains are fairly widely expressed throughout the body. Gene knockout experiments of α1 and α2 chains have shown that inactivation of the individual collagen-binding integrins does not seem to impair embryonic development (46, 47). Rather, mild phenotypes are observed where either fibroblast and leukocyte interactions with collagens or platelet interactions with collagens are affected. Recent analysis of α11 and α1 expression (3, 29) reveals a restricted embryonic expression pattern, which is not overlapping but complementary. In the near future it will be important to determine to what extent the collagen-binding integrins show overlapping functions and to what extent different collagen-binding integrins can functionally compensate for each other’s absence. Crossing different mice strains lacking certain collagen receptors will shed light on these issues.

As a part of understanding the biological function of collagen-binding integrins it is important to characterize all of the different collagen-binding integrins with regard to collagen affinity, collagen specificity, divalent ion requirements, and ligand recognition motifs. Studies of α1 I, α2 I, and α10 I domains have shown that they bind collagens with different specificity (27). This specificity seems in part to be determined by residues located outside the MIDAS motif in the α I domain. Data from several groups have convincingly shown that α1 prefers collagens IV and VI over collagen I and that the preferences of α2 are opposite. More recently the α10 I domain was shown to display a collagen binding specificity similar to that of α1 (27).

Prior to this study no binding studies had been performed with the α11 I domain. It thus appears that although α11 is, in terms of evolution, more similar to α2 and α10, it is more closely related to α11, another grouping can be made based on their collagen specificity. The finding that α11 prefers interstitial collagens over nonfibrillar collagens supports our previous cell binding data (29), but the difference is even more pronounced at the α1 I domain level. A candidate amino acid that might play a role in determining this preference is Thr-238 found in a position corresponding to Arg-218 in α1.

The relatively low avidity for collagens estimated for both α10 I and α11 I domains is intriguing. Our experience is that as recombinant GST fusion proteins, these I domains are less soluble than α1 I and α2 I domains, and they might have a tendency to form aggregates. This may affect the Kd estimates. Furthermore, we have shown that in the length of the produced protein a difference of one amino acid residue might lead to changes in the avidity of collagen binding (48). Thus, the approximated Kd values should be used for comparing the binding of a recombinant α I domain with different collagens rather than for comparing the α I domains with each other. Low avidity may indicate that the major function for α10 and α11 is not that of firm adhesion but that these integrins engage in dynamic interactions with collagen during events such as cell migration. It is also possible that the true ligands have not yet been identified. For example, for α10 I, a cartilage ligand, possibly a collagen other than collagen II, might be the preferred ligand. In the case of α11 I, a perichondrium ligand other than
collagen I might bind this integrin with higher affinity.

The $\alpha_\beta_1$-mediated binding of platelets to collagen I is inhibited by mM concentrations of Ca$^{2+}$ (37). The finding that $\alpha_\beta_1$ is not inhibited by Ca$^{2+}$ when expressed in C2C12 cells is intriguing. In the case of platelets and C2C12 cells this difference might be related to the activation status of the integrin. Whereas platelets and leukocytes have a more elaborate system for regulating integrin activation status, integrins in C2C12 cells are expected to be mainly in the activated state, displaying a higher affinity. $\alpha_\beta_1$ is not expressed on platelets, so a direct comparison with $\alpha_\beta_1$ is not possible. However, when expressed in C2C12 cells, $\alpha_\beta_1$ binding to collagen I requires $\mu$M Ca$^{2+}$ and is sensitive to mM concentrations of Ca$^{2+}$, as $\alpha_\beta_1$ when expressed on the platelet surface. The recent crystallization of soluble $\alpha_\beta_1$ supports a role of Ca$^{2+}$ ions in allosteric modulation of integrin conformation (49). It is possible that a high affinity interaction with collagen I, such as that mediated by $\alpha_\beta_1$, is less affected by Ca$^{2+}$-induced allosteric conformational changes. Conversely, a lower affinity interaction with collagen I, such as that mediated via $\alpha_\beta_2$, might be more sensitive to allosteric changes in other regions of the receptor. This differential sensitivity to Ca$^{2+}$ might be physiologically important in the formation and turnover of the musculoskeletal system, where the local concentration of Ca$^{2+}$ varies.

Despite the differences in collagen specificity, helical GFOGER-like sequences are recognized by $\alpha_1\beta_1$, $\alpha_\beta_1$, and as shown in this study, also by $\alpha_1\beta_1$. Careful analysis of the occurrence of GFOGER-like peptides has shown that in addition to the CB3-derived sequences GFOGER and GFOGEK, which are present in the central part of the collagen chain, an amino-terminal $\alpha_2$ I and $\alpha_1$ I domain binding site overlaps with the peptide GLOG3 (33). As shown in this study, the binding of the differen I domains to different collagen peptides varied somewhat. The glutamate in GFOGER was central for the binding of all I domains, whereas the phenylalanine seemed to be more important for $\alpha_1$ binding, and the arginine was especially important for $\alpha_1$ binding. It is possible that in vivo collagen-binding integrins prefer certain sites on the collagen molecules. In a particular cell expressing multiple collagen receptors, a number of factors might determine which region in collagen is bound by a particular integrin. Some of the factors that might affect ligand binding include local Ca$^{2+}$ concentrations, expression levels of the different receptors, and subcellular localization within the cell. Collagen receptors have been shown to affect collagen and matrix metalloproteinase synthesis. Already now it is possible to envisage how GFOGER peptides have the potential to become universal reagents blocking cell-collagen interactions. It will be important to determine whether $\alpha_\beta_1$ also binds GFOGER peptides. Triple-helical collagen peptides might be of use in conditions characterized by excessive collagen production such as various fibrotic conditions.

REFERENCES

1. Mlynyiharju, J., and Kivirikko, K. I. (2001) Ann. Med. 33, 7–21.
2. Hashimoto, T., Watanabe, T., Kato, A., Sawai, H., and Ohtsuki, R. (2000) J. Biol. Chem. 275, 20380–20389.
3. Velting, T., Kusche-Gullberg, M., Sejersen, T., and Gullberg, D. (1999) J. Biol. Chem. 274, 26735–26743.
4. Carter, W. G., Ryan, M. C., and Gahr, P. J. (1991) Cell 65, 599–610.
5. R. W. Farndale, unpublished observation.