Integrin Activation State Determines Selectivity for Novel Recognition Sites in Fibrillar Collagens

Received for publication, April 27, 2004, and in revised form, August 13, 2004
Published, JBC Papers in Press, September 2, 2004, DOI 10.1074/jbc.M404685200

Pia R.-M. Siljander‡§, Samir Hamai‡, Anthony R. Peache yay‡, David A. Slatter‡,
Peter A. Smethurst‡§, Willem H. Ouwehand‡, C. Graham Knight‡, and Richard W. Farndale‡

From the Departments of ‡Biochemistry and §Haematology, University of Cambridge, Downing Site,
Cambridge CB2 1QW, United Kingdom

Only three recognition motifs, GFOGER, GLOGER, and GASGER, all present in type I collagen, have been identified to date for collagen-binding integrins, such as αβ1. Sequence alignment was used to investigate the occurrence of related motifs in other human fibrillar collagens, and located a conserved array of novel GER motifs within their triple helical domains. We compared the integrin binding properties of synthetic triple helical peptides containing examples of such sequences (GLSGER, GMOGER, GAOGGER, and GQRGER) or the previously identified motifs. Recombinant inserted (I) domains of integrin subunits α1, α2, and α11 bound poorly to all motifs other than GFOGER and GLOGER. Similarly, αβ1-containing resting platelets adhered well only to GFOGER and GLOGER, while ADP-activated platelets, HT1080 cells and two active αII domain mutants (E318W, locked open) bound all motifs well, indicating that affinity modulation determines the sequence selectivity of integrins. GxO/SGER peptides inhibited platelet adhesion to collagen monomers with order of potency F ≥ L ≥ M > A. These results establish GFOGER as a high affinity sequence, which can interact with the αI domain in the absence of activation and suggest that integrin reactivity of collagens may be predicted from their GER content.

Collagen, the most abundant structural protein of the vertebrate organism, currently has 27 reported family members (1). As either a mechanical support or as a bioactive surface, collagen plays a crucial role in processes as diverse as morphogenesis, wound repair, inflammation, tumor metastasis, hemostasis, and thrombosis. The tensile strength of the fibrillar collagens, types I-III, V, XI, and XXVII, is crucial to the function of connective tissues including the blood vessel wall. The non-fibrillar collagens, such as types IV and VI, provide a flexible support for endothelial and epithelial cell attachment and development. Integrins, heterodimeric adhesion molecules, form an important subgroup of receptors, which mediate interaction between collagen and cells. Four α-subunits, α1, α2, α10, and α11, which associate non-covalently with β1, constitute the native collagen-binding integrin family (2).

Integrin αβ1 is a well characterized and widespread receptor for collagen, laminin, and other non-matrix ligands among nucleated cells including epithelial and endothelial cells, smooth muscle cells, fibroblasts, leukocytes, and mast cells (3, 4), mediating a wide range of cellular activities. αβ1 is the only collagen binding integrin in platelets, and is crucial for deposition on collagen exposed in damaged arterial walls (5, 6). Accordingly, the expression levels of αβ1 have been associated with myocardial infarction and stroke (7). Recently, knockout studies suggested that the platelet activatory collagen receptor glycoprotein VI (GPVI) is mandatory for the initiation of integrin-mediated adhesion (8), but this concept was challenged by further mouse studies (9). Moreover, in vitro studies suggest that soluble collagen binding to αβ1 requires a change in the conformation of the integrin (10). Therefore the mechanism by which αβ1 acts as a primary adhesive receptor and yet is subject to affinity modulation remains to be reconciled.

The native collagen binding integrins contain within their α-subunit an inserted (I) domain (αI), which binds collagen through its metal ion-dependent adhesion site (MIDAS)1, perhaps the only site of interaction, although in integrins lacking an αI domain, the I-like domain of the β-subunit constitutes the ligand binding site (11). Use of collagen-derived triple helical peptides identified the sequence GFOGER (where O is hydroxyproline) as the minimal recognition motif for α1, αI, and α11 (12–14), and yielded the first integrin-ligand co-crystal (15), which revealed the crucial interaction of the MIDAS and the GER sequence. The peptide glutamate residue is directly coordinated to the divalent cation bound to the MIDAS, whereas the arginine forms a salt bridge to the αI asparagine, 219. Changing Glu to Asp abolishes binding, while Arg to Lys reduces the binding by 50% (13). In contrast, the phenylalanine may be less important, since two other collagen type I α chain sequences, GLOGER and GASGER, were found to bind αI and αII, when I-domain-interacting areas were mapped within collagen by rotary shadowing (16). No other recognition sequences have been unequivocally identified hitherto, although collagens I and III contain 11 and 14 GER triple helical motifs respectively, some of which occur within cyanogen bromide-cleaved peptides, which support integrin binding (17–21). One conserved sequence, GLOGER, is cleaved at methionine by cyanogen bromide, and so is not present in such peptides.

We used bioinformatics to locate and identify novel GXXGER sequences within the collagen triple helices, and examined the affinity of synthetic peptides representing examples from them for αβ1 in different cell types, and for recombinant α1, α2, and

1 The abbreviations used are: MIDAS, metal ion-dependent adhesion site; GST, glutathione S-transferase; PBS, phosphate-buffered saline; ANOVA, analysis of variance; mAb, monoclonal antibody; O, hydroxyproline; LO, locked open; Fmoc, N-(9-fluorenyl)methoxycarbonyl.
α₁I I domains. All the double triplets were synthesized in two different flanking hosts (22, 23), (GPP)₅ and (GPO)₅, which provide the minimal structure to maintain triple helix stability at 20 °C (24). We identified new sequences analogous to GFOGER at conserved loci within the collagen helix that can bind integrin, and confirmed the reactivity of GLOGER and GASGER. Competition experiments established an affinity series for the recognition motifs: GFOGER > GLOGER > GLSGER > GAOGER > GAOGER ≈ GASGER ≈ GGQGER. The cellular context was found to be crucial in determining the binding to these sequences, since, in contrast to adhesion of αβ₁-containing HT1080 cells (a constitutively adherent cell line), αβ₁-mediated platelet adhesion was poor to GER sequences other than GFOGER unless the peptide also contained GPVI recognition motifs (GPO₅) (25, 26), or the platelets were otherwise activated (27) in line with their hemostatic surveillance function. This concept was confirmed by non-selective binding of active α₁I mutants to all sequences.

**EXPERIMENTAL PROCEDURES**

**Materials**—Human platelets were isolated from citrate-anticoagulated whole blood, provided by the National Blood Service (Cambridge, UK). Pepsin-digested monomeric collagen type I from bovine skin has been previously described (19). Monoclonal anti-α₁ antibody 6F1 was a kind gift from Dr. Barry Coller (Mount Sinai Hospital, New York). GR144505F was a gift from Glaxo Wellcome (Stevenage, UK). Anti-GR144505 Ab 10B12 was generated as previously described (28). Horseradish peroxidase-conjugated anti-GST antibody was from Amersham Biosciences UK Limited ( Bucksheath, UK). Unless otherwise stated, other reagents were from Sigma.

**Sequence Alignment**—Fibrillar collagen sequences CA11, CA21, CA12, CA13, CA15, CA25, CA1B, CA2B, and CA1R were downloaded from the Swiss-Prot database (ca.expasy.org/sprot/), and non-helical parts of the sequences were removed before alignment using ClustalX. Small modifications to the alignment were made manually before assigning loci of GER sequences. Residue 1 of the type 1 collagen helix was denoted residue 1 of the alignment and of the first D-period of 234 residues. Subsequently, all 90 non-cuticle collagen sequence from the Swiss-Prot database with the “CA” prefix encompassing all collagen types across a number of species were downloaded using the sequence retrieval system (hrs.embl-heidelberg.de:8000/hrs/S). These were analyzed for the presence of GXXGER sequences (see Supplementary Data).

**Peptide Synthesis and Purification**—The peptides synthesized for this study are shown in Table I. Peptides were made as C-terminal amidases on Tentagel R RAM resin (0.1 mmol) in an Applied Biosystems Pioneer peptide synthesizer using standard Fmoc protocols (29). Briefly, Fmoc amino acids (0.4 mmol) were activated with O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) or O-(6-chlorobenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) (0.4 mmol) in the presence of diisopropylamine (0.8 mmol). Fmoc deprotection was with a mixture of 2% (v/v) piperidine and 2% (v/v) 1,8-diazabicyclo[5.4.0]undec-7-ene. Peptides were released from the resin by treatment with a mixture of trifluoroacetic acid (9.5 ml), thioisobutyric acid (0.5 ml) and triisopropylsilane (2.5 ml). Ammonium iodide (0.15 g) was added when cleaving methionine peptides to suppress oxidation (0.3%). Crude peptides were purified by reverse phase HPLC (PerkinElmer Life Sciences LC200) using ACE diethylamino columns (Hichrom Ltd, Theale, Berkshire, UK). The identity of the purified peptides was confirmed by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry at the Protein and Nucleic Acid Chemistry Facility (Cambridge Centre for Molecular Recognition).

**Cell Adhesion and Spreading**—Human fibrosarcoma cells, HT1080, obtained from the European Collection of Animal Cell Cultures (Porton Down, UK), were maintained in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum, 2 mM glutamine, 100 international units/ml penicillin, 100 µg/ml streptomycin, and 2.5 µg/ml amphotericin B. Cells were harvested with trypsin/EDTA, washed, and suspended at 0.3 × 10⁶/ml in platelet adhesion buffer, supplemented with glucose at 1 g/l and 2 mM EDTA or MgCl₂. 100 µl of cells were allowed to adhere at 20 °C for 30 min for standard adhesions and 20 min in the presence of GEP. Adhesion (% total number of added cells) was determined as for platelets. Cells were allowed to spread on GER substrates for 90 min, then fixed and stained with 2% Crystal Violet in PBS, then ×200 bright-field images from a Nikon TMS microscope were analyzed using a Leica Q550C image analyser and QWIN software (Leica, Cambridge UK). The area of at least 100 cells was measured for each peptide.

**Platelet Adhesion**—Platelet (1.25 × 10⁶/ml) adhesion was determined colorimetrically (absorbance 405 nm), as described (31). Peptides or collagen were coated at 10 µg/ml on Immulon-2 HB 96-well plates (Thermo Life Sciences, Basingstoke, UK), at which concentration platelet binding reached a maximum. When peptides were used to inhibit platelet adhesion to collagen, they were added to the platelets prior to plating with 2 µM GR14505F, which prevents platelet aggregation. Stimulation by 30 µM ADP was carried out in the presence of 2 µM GR144505F 5 min prior to the addition of platelets to peptide-coated wells.

**Data Analysis**—Data from each donor platelet batch used for competitive inhibition of platelet binding to collagen (shown in Figs. 4 and 8) were fitted to the ligand binding Equation 1 (32),

\[ P = \frac{P_{\text{max}} - \frac{P_{\text{min}}}{A_{\text{Hill}}}}{1 + \left(\frac{A}{A_{\text{Hill}}}ight)^n} + P_{\text{min}} \]

where \( P \) = response from platelet binding, \( P_{\text{max}} \) = response at zero peptide concentration, \( P_{\text{min}} \) = response at infinite peptide concentration, \( A_{\text{Hill}} \) = hill coefficient.

The maximal inhibition of platelet binding, \( I_{\text{max}} \), for each peptide was expressed as a percentage defined in Equation 2.

\[ I_{\text{max}} = \left(\frac{P_{\text{max}} - P}{P_{\text{max}}}\right) \times 100 \]

**As \( P_{\text{max}} \) varied between patients, datasets for any one peptide were normalized such that variance was evenly distributed between peptide concentration points and that the average \( P_{\text{max}} \) value across them was 1. Mean and S.E. values for \( P_{\text{50}} \), \( H \), and \( I_{\text{max}} \) were determined. Values for \( P_{\text{50}} \) were converted to \( A_{\text{Hill}} \) values in micromolar units, and divided by 3 to give the concentration of triple helix.

**Plasmids**—The recombinant I-domain encoding plasmids of α₁ (33), α₂ (34), α₂ mutant (E112W) (35), and αM (36) were a generous gift from Danny Newsholt, FKF, University of Kentucky. For integrin α₁I expression plasmid, a DNA fragment corresponding to the integrin α₁I gene was generated by PCR using the truncated human α₁I-fos HMT vector (a kind gift from D. Gullberg, Uppsala, Sweden) as template. The forward and reverse primers with an additional BamHI and EcoRI sites were (5’-ATACATCAATTTGATCCCTGGATGCTGCAACAGCAT-3’) and (5’-TCGATATGGAATCCAGGGACCTGACGAATCTCCTG-3’), respectively. The DNA sequence was confirmed. Data from transformants was sequenced and compared with the published sequence (37).

**Mutagenesis of α₁I—**P2EG-2T containing the α₁I DNA was used to generate the I-domain mutant “locked open” (LO) containing cysteine at 172 and 322 positions (G172C and L322C) by site-directed mutagenesis, essentially as describe previously for α₁I (38). The underlying rationale was to secure helix 7 in the position defined in the I-domain:peptide co-crystal, so that the MIDAS remained open. The primers used were: GAGAAATTGTGACAACTTGCACGATGCTGAGCCCGACACCACTAGG, CCTGTGTTTGGGCTGGCATATCAAGGCGCAACTTGAAGTCCT, GCTGATTGAAACAGCCTCATCTGCGAGAAAGCTGCGCATTAAGAG, and TCTGAAATGTCGCCCTTTTTCGTGCAAGACATCACTAC.

**Expression of I Domains**—For protein expression and purification of α₁I, α₂I, α₁I, α₂I, LO, and E318W I domains, a 40-ml overnight culture of transformants was used to inoculate 400 ml of Luria Broth, 50 mg/ml ampicillin. The culture was grown for 1 h at 37 °C and then induced for 4 h with 0.1 mM isopropyl-β-D-thiogalactoside. Cells were harvested by centrifugation (4,500 × g, 10 min), and pellets resuspended in PBS without divalent cations (PBS-). Suspensions were sonicated and centrifuged (2,500 × g, 10 min), and the supernatants were adjusted to 1% Triton X-100. Pellets were resuspended in PBS–, sonicated, and centrifuged twice more, and the supernatants pooled. The lysate was passed down a glutathione-agarose column equilibrated in 150 mM NaCl, 2 mM Tris-HCl, pH 8.0 (TBS), the column washed with 10 volumes of TBS and the glutathione S-transferase-I domain fusion proteins eluted with 50 mM glutathione in 50 mM Tris-HCl, pH 8.0. The proteins were then dialyzed against TBS and concentrated using a Microcon 3 microconcentrator (Amicon, Stonehouse, Gloucester, UK). The I domains were checked for purity and degradation by 10% SDS-gel and by Western blotting. Nitrocellulose blot were probed with HRP-conjugated polyclonal anti-GST antibody.
I-domain Binding—Immuno-2 96-well plates were coated as described for platelet adhesion, and blocked for 2 h with 200 μl of TBS with 50 mg/ml bovine serum albumin. Wells were washed four times with 200 μl of the adhesion buffer (TBS with 1 mg/ml bovine serum albumin) before adding 100 μl of adhesion buffer containing 5 μg/ml of recombinant GST I domains in the presence of either 2 mM MgCl₂ or EDTA for 1 h at room temperature. This concentration of I domains provides optimal detection at a submaximal level of binding (39). Wells were washed five times with 200 μl of adhesion buffer containing MgCl₂ or EDTA, before adding 100 μl of adhesion buffer containing the anti-GST horseradish peroxidase-conjugate (1:5000) for 1 h at room temperature. After washing, color was developed using an ImmunoPure TMB Substrate Kit (Pierce) according to the manufacturer’s instructions.

Statistics—The absorbance values for the adhesion of cells or I domains to peptide substrates were compared using either 1- or 2-way ANOVA. Specific comparisons were performed using the Neumann-Keuls (1-way) or Bonferroni (2-way) post-tests. Student’s ANOVA. Specific comparisons were performed using the Neumann-Mins to peptide substrates were compared using either 1- or 2-way ANOVA. Specific comparisons were performed using the Neumann-Keuls (1-way) or Bonferroni (2-way) post-tests. Student’s ANOVA.

RESULTS
Collagen primary amino acid sequences GFOGER, GLOGER, and GASGER were previously shown to bind α₅β₁-containing cells, the purified integrin and recombinant α₁, α₂, and α₁₁ I domains (13, 14, 16). A bioinformatic approach was taken to analyze the presence of GER sequences in human fibrillar collagens. An alignment of individual α-chains from collagens I, II, III, V, and VI suggested that GXXGER sequences occur at conserved loci within triple helical collagen domains, and these sites contain novel GER sequences (Fig. 1). Three conserved loci for GFOGER and GLOGER in collagen types I, II, and VI, started at residues 127, 502, and 550 in the alignment. When other sequences occur in these loci, they are usually GMOGER or GAOGER, a previously proposed recognition sequence (25). Because the C-terminal GASGER, at residue 811, occupied a less conserved site and was shown to bind of poor affinity at best (16), we also identified another C-terminal GXXGER locus nearby at 787, which contains several GXXGER sequences each having a polar rather than a hydrophobic residue within the first triplet. From this locus we chose to test GQGER. Peptides containing these novel sequences were synthesized in addition to the other previously identified integrin recognition sequences using two different hosts, (GPP)₁₀ and (GPO)₁₀ (Table I). As controls, (GPP)₁₀ or the GPPGPQ sequence in a (GPO)₁₀ host were used. All the peptides had stable triple helices at room temperature with GER peptides lacking hydroxyproline as their third amino acid exhibiting lower melting points, demonstrating its important role in stabilizing the peptide helix.

To assess cellular function of these recognition sequences, two different cell types, human fibrosarcoma cell line HT1080 and human platelets, each of which express no collagen-binding integrin other than α₅β₁, were assayed to determine for their capacity to adhere to the peptides in the (GPP)₁₀ host. HT1080 cells bound well to all GER peptides, whereas the binding was poor to (GPP)₁₀ (Fig. 2). The HT1080 expressed a distinct binding profile: while GFO did not differ from GLO, GLS, or GMO, nor GAS from GAO, GOM, and GLS, significantly greater binding occurred to GFO or GLO than to GAO and GAS (Fig. 2A). Adhesion to (GPP)₁₀ was significantly lower than to each GXXGER sequence. In separate experiments, the binding of HT1080 to the polar GQGER motif was found to be within the same range as to GAO/GAS (data not shown). Integrin specificity of the adhesion to these sequences was demonstrated by their MgCl₂ dependence (Fig. 2B), and by blocking with an α₅ inhibitory mAb 6F1 (Fig. 2C). To detail the differential binding of HT1080 cells, we also measured their spreading. HT1080 cells expressed peptide preference for spreading, which paralleled their adhesion levels and which was clearly observed as a reduction in number of fully spread cells on GAO, GAS, and GQR and (GPP)₁₀ control (Fig. 2D). GQGER supported greater spreading, cell area being 150% of that obtained on GQR (p < 0.001), with cell area on other substrates decreasing in order GPO, GLO > GLS, GMO > GAO, GAS, GQR (data not shown). Finally, we assayed adhesion of another cell type, platelets, to the GXXGER peptide set. In striking contrast to the HT1080 cells, divergent cation-dependent binding of platelets to all sequences except GFOGER was reduced, and platelet binding to GAS, GAO, and GQR was almost negligible, paralleling the lower affinity observed with HT1080 cells (Fig. 3).

To test whether the binding capacity of the GXXGER sequences for platelets could be improved by allowing their interaction in solution rather than as an immobilized ligand, the peptides were tested as inhibitors of platelet adhesion to bovine type I collagen monomers. Inhibition experiments indicated three affinity classes of peptides: GFO had a 45-fold higher affinity over the next best sequence, GLO, which had medium affinity together with GLS and GMO (Fig. 4A), whereas GAO/ GAS/GQR exhibited no additional inhibition when compared with the (GPP)₁₀ control (Fig. 4B and data not shown). This qualitative order of affinity was always preserved irrespective of the absolute binding levels, which varied between platelets from different donors. A* values and statistical data derived from the curves modeled for Figs. 4 and 8 as described under “Experimental Procedures” are given for the different ligands.
in Table II. Inhibition of platelet adhesion to human collagens types I and III displayed the same affinity order as bovine monomers, as expected (data not shown).

To analyze the direct capacity of isolated αI to bind the GER peptides, we purified recombinant I domains as glutathione S-transferase (GST) fusion proteins (Fig. 5). Binding to GFOGER was used as an inter-assay reference to normalize the results with the three I domains. Mg²⁺-dependent adhesion of all domains was significant (p < 0.05) to all peptides except (GP₅)₁₀ (all αI) and GQR (to which α₁I bound significantly). The binding profile of wild-type α₁I was more similar to that of α₂β₁ in situ in platelets than in HT1080 cells. Effective binding was observed to GFO, GLO, and GLS. α₁I bound less to GMO than platelets did, and less well to GLO and GLS than HT1080 cells. Comparison of the I domains demonstrated both qualitative and quantitative differences (Fig. 5). α₁I strongly preferred GLO (−90% binding of that of GFOGER) in comparison to α₂I (−40% p < 0.001) and α₁₁I (−70% p < 0.05). In contrast to α₂I and α₁₁I, α₁I also had better affinity to GLS than GMO, the two former having GLS and GMO binding comparable to GAS and GAO. As a control, no α₃I binding to the sequences was observed (data not shown).

The difference between the adhesion profiles of HT1080 cells and platelets to GXXGER peptides suggested that α₂β₁ was displayed in different conformation in the two cell types. Platelet binding to collagen monomers in suspension has previously been shown to require affinity modulation of the integrin α₂β₁ through activation by agonists such as ADP or via collagen’s main signaling receptor in platelets, GPVI (10). GPVI can be specifically activated with a cross-linked (GPO)₁₀ polymer, collagen-related peptide (CRP), and peptides containing (GPO)₉ can bind GPVI when immobilized to a surface (26). To allow platelet interaction or activation by GPVI, we synthesized the same set of GXXGER peptides in (GPO)₉ hosts (Table I). In contrast to data shown above, platelets bound to all immobilized GXXGER sequences in which the (GP₅)₉ host was replaced by (GPO)₉ (Fig. 6A). An α₂β₁-dependent element of adhesion could be resolved for GPO peptides containing GFO, GLO, GMO, and GLS as well as for the controls, GFOGER in (GPO)₉ and monomeric collagen, since a significant decrease in binding was observed in the presence of either EDTA or 6F1 (p < 0.001 ANOVA). Adhesion to GAS and GAO within (GPO)₉ host, in contrast, was almost completely dependent on GPVI, judged from the marked attenuation of adhesion in the presence of an anti-GPVI antibody, 10B12 (28). The effect of 10B12 on adhesion to all these peptides was uniformly significant (p < 0.001, ANOVA). Thus, for most GER peptides, GAO and GAS being exceptions, platelet binding was dependent on both α₂β₁ and GpVI. Platelets from 30 donors bound GFOGER in (GPO)₉ host about 40% better than in (GPP)₉ host, revealing that α₂β₁ interaction with GFOGER predominates over the GPVI involvement (Fig. 6B), whereas HT1080 cells, which lack GPVI, bound to peptides in (GPP)₉ and (GPO)₉ hosts equally well (Fig. 6C). Together, these data suggested that GFOGER is a high affinity sequence not requiring affinity modulation of the integrin for its interaction.

To determine whether platelet activation improves the capacity of α₂β₁ to interact with low affinity sequences through affinity modulation, we stimulated platelets prior to adhesion to the peptides in (GPO)₉ host with 30 μM ADP, providing a GPVI-independent activation. ADP treatment markedly increased platelet adhesion to all substrates except for GFOGER and collagen (Fig. 7), and now GAO, GAS, and GQR demonstrated clear binding. (GPP)₁₀ controls showed little binding activity even after preactivation of the integrin. Although ADP stimulation improved platelet adhesion to the weaker affinity peptides more than introduction of (GPO)₉ host (likely to provide only weak activation of GPVI, Ref. 40), replacement of (GPP)₉ with (GPO)₉ host dramatically improved the capacity of all peptides to competitively inhibit platelet adhesion to monoclonic collagen. The most improvement was achieved for the medium affinity sequences GMO/GLS, which became almost as effective as GLO (Fig. 8A). Also, the lower affinity sequences GAS and GAO differed from their respective GPP-GPO control (Fig. 8B), but GQRGER (not shown) remained similar to control. A₅₀ was >200 μM; nH and Iₘₐₓ were indeterminate (n = 2). Relative A₅₀ values were determined where possible from pooled experiments (Figs. 4 and 8) for the set of peptides, which demonstrates an overall decrease in A₅₀ with all the peptides in a (GPO)₉ host in comparison to (GPP)₉ host, particularly for GLS and GMO (Table II). However, GFOGER was a marked exception and conversely, it actually had a significantly lower A₅₀ in the (GPP)₉ host (p = 0.008). These data confirm that GFOGER supports I-domain interaction without integrin affinity modulation.

Final proof that the activation state of the integrin underlies the observed adhesion profiles, was obtained using two different active mutants of α₁I, E318W (35) and a newly generated

| Name         | Sequence                                                                 | Mₑ  | Tₑ   |
|--------------|--------------------------------------------------------------------------|-----|------|
| GFOGER in GPP* | H-GPC(GPP)₅GFOGER(GPP)₅GPC-NH₂                                           | 3704.2 | 45 |
| GFOGER in GPO | H-GPC(GPO)₅GFOGER(GPO)₅GPC-NH₂                                           | 2795.1 | 43 |
| GLOGER in GPO | H-GPC(GPO)₅GLOGER(GPO)₅GPC-NH₂                                           | 3670.2 | 50 |
| GASGER in GPO | H-GPC(GPO)₅GASGER(GPO)₅GPC-NH₂                                           | 2761.1 | 45 |
| GAGGER in GPO | H-GPC(GPO)₅GAGGER(GPO)₅GPC-NH₂                                           | 3602.1 | 42 |
| GAOGGER in GPO| H-GPC(GPO)₅GAOGER(GPO)₅GPC-NH₂                                           | 2692.9 | 34 |
| GAOGGER in GPO| H-GPC(GPO)₅GAOGER(GPO)₅GPC-NH₂                                           | 3628.1 | 50 |
| GLSGER in GPP | H-GPC(GPO)₅GLSGER(GPO)₅GPC-NH₂                                           | 2719.0 | 48 |
| GLSGER in GPO | H-GPC(GPO)₅GLSGER(GPO)₅GPC-NH₂                                           | 3644.2 | 39 |
| GMGER in GPO  | H-GPC(GPO)₅GMOGER(GPO)₅GPC-NH₂                                           | 2735.0 | 33 |
| GMGER in GPO  | H-GPC(GPO)₅GMOGER(GPO)₅GPC-NH₂                                           | 3688.2 | 56 |
| GMGER in GPO  | H-GPC(GPO)₅GMOGER(GPO)₅GPC-NH₂                                           | 2779.1 | 44 |
| GMGER in GPO  | H-GPC(GPO)₅GMOGER(GPO)₅GPC-NH₂                                           | 3738.2 | 46 |
| GMGER in GPO  | H-GPC(GPO)₅GMOGER(GPO)₅GPC-NH₂                                           | 2819.1 | 41 |
| GDPGER in GPP | H-GPC(GPO)₅GDPGER(GPO)₅GPC-NH₂                                           | 3044.5 | 51 |
| GDPGER in GPP | H-GPC(GPO)₅GDPGER(GPO)₅GPC-NH₂                                           | 2637.9 | 55 |

* Bold text indicates amino acid changes in the peptide set.
disulfide LO form, which was produced by substituting cysteines at residues 172 and 322, in which helix 7 is predicted from I-domain crystal to be constrained at the activated down position (15). Both active mutants were found to effectively adhere to all peptides (Fig. 9). The overall absorbance levels of both LO and E318W binding to GFOGER were very similar to that of wild-type. Most interestingly, the binding pattern of both E318W and LO mutants to lower affinity sequences resembled that of HT1080 cells, except that E318W binding to GQR binding was poor.

In summary, our data show that integrin recognition sequences of different affinities exist in conserved loci within fibrillar collagens, and that the activation state of the integrin depends on the cellular milieu and determines its affinity for the collagen species. The first identified integrin recognition sequence, GFOGER, is a high affinity motif, which can bind the integrin without its prior activation. In contrast, binding to most other GER sequences, including the newly found GMOGER, GLSGER, GAOGER, and GQRGER, is sequence-dependent, exhibiting a spectrum of affinity, both in the resting state and after integrin activation.

DISCUSSION

Human fibrillar collagen α-chains contain multiple two triplet integrin binding motifs, which have GER as their second triplet. The conserved spacing of these various sequences within collagens I, II, III, V, and XI, reflects their close evolutionary origins. The final collagen XXVII has a large fibrillar domain, which contains only three GXXGER sequences of unknown affinity. When 90 available collagen sequences were investigated, a total of 452 GXXGER sequences were found, of which just 70 are predicted according to our results to have significant binding to integrins based on their first triplet (GFO, GLO, GMO, GAO, GAS, GVR, and GMS; see also Supplementary Data). Of the sequences studied here, there exist 21 GFOGER, 23 GLOGER, 15 GMOGER, 3 GAOGER, 19 GASGER sequences, whereas GLSGER is unique to human type III collagen. The first three have very high occurrence given the relatively low amino acid frequency of F, L, and M in collagen and thus are likely to have important function. GFOGER was found in collagen chains of Iα1, II, IVα1, IVα3, and IIα1.
IVα4, IVα5, VII, XIα1, and XIα2, but not in any other collagens. In eel and Drosophila, GYoger might substitute for the absence of FOGGER based on amino acid similarity. Moreover, a multitude of GXXGEK sequences in collagens coincide with the proposed GXXGER loci, suggesting that they too, are integrin binding sites with a lower affinity than GER, as the Arg to Lys change attenuates affinity by 50% (13). Nine GFOGER sequences that may bind integrin were located, including in human collagens Xα1 and Vα3, followed by a further 26 GLOGEK sequences in collagen types III, IV, V, VIII, and XI. This remarkable sequence conservation suggests that these motifs serve an important role, such as the ability to support receptor avidity and subsequent signaling as receptors align with the higher order structure of the fibril. The GXXGER sites may also have other functions: it has been proposed that dimerization of collagen monomers depends on the juxtaposition of GER with the A-domain of collagen VI (41).

The affinity series of α2β1 for GER sequences was remarkably well conserved between HT1080 cells, platelets, and recombinant I domains, whether direct adhesion to peptides or displacement from collagen was measured: GFO > GLO > GLS > GMO > GAO = GAS = GQR. Minor exceptions to this rule were observed for the constitutively active mutants, where GLO, GLS, and GMO did not differ significantly from FOGGER. The elucidation of different preferences of binding requires structural characterization of these mutants. Relatively low binding was found for α1I and resting platelets to GLS and GMO in comparison to HT1080 cells and activated platelets. Elegant studies by Chen et al. (42) suggest a structural basis for intermediate affinity forms of integrins regulated by changes within the β-subunit. Here we present complementary data, which define the requirements within collagen for binding to the integrins, and indicates a role for activation. Such activation, possibly involving both integrin subunits, could determine the affinity for the intermediate sequences GLS and GMO.

Table II
Parameters describing GXXGER peptide inhibition of platelet binding to collagen type I monomers

| Peptide | \( A_{50} \) | \( nH \) | \( I_{\text{max}} \) | \( r^2 \) | \( n \) |
|---------|-----------|------|--------|------|-----|
| GFOGER-GPP | 0.14, 0.16, 0.26 | 0.74 ± 0.05 | 99.5 ± 0.4 | 0.984 | 4 |
| GFOGER-GPO | 0.28, 1.03, 1.60 | 1.14 ± 0.11 | 99.7 ± 0.2 | 0.980 | 4 |
| \( p \) values | 0.008 | NS | NS |
| GLOGER-GPP | 6.24, 5.40, 7.22 | 1.18 ± 0.09 | 95.8 ± 0.23 | 0.963 | 5 |
| GLOGER-GPO | 3.91, 3.41, 4.47 | 1.24 ± 0.05 | 100 ± 0.00 | 0.991 | 5 |
| \( p \) values | 0.009 | NS | NS |
| GLSGER-GPP | 33.2, 23.3, 47.2 | 1.23 ± 0.21 | 92.6 ± 5.0 | 0.962 | 5 |
| GLSGER-GPO | 4.68, 3.44, 6.37 | 1.37 ± 0.05 | 99.9 ± 0.04 | 0.974 | 5 |
| \( p \) values | 0.0005 | NS | NS |
| GMOGER-GPP | 95.0, 81.4, 111 | 1.19 ± 0.19 | 86.1 ± 5.8 | 0.898 | 5 |
| GMOGER-GPO | 16.0, 11.7, 21.9 | 0.97 ± 0.10 | 100 ± 0.00 | 0.986 | 5 |
| \( p \) values | 0.008 | NS | NS |
| GAGGER-GPP | >200 | ND | ND | ND | 5 |
| GAGGER-GPO | 151, 124, 184 | 1.06 ± 0.13 | 100 | 0.875 | 5 |
| GASGER-GPP | >200 | ND | ND | ND | 4 |
| GASGER-GPO | 108, 90.1, 130 | 1.10 ± 0.33 | 100 | 0.857 | 4 |
| GPGP-GPP | >200 | ND | ND | ND | 4 |
| GPGP-GPO | >200 | ND | ND | ND | 4 |
| GPP order | GFO, GLO, GLS, GMO | GAS = GAO = GPP |
| GPP order | GFO, GLO, GLS, GMO | GAS > GAO > GPP |

a Concentrations are given for the triple helix.

b NS, no significant difference.
c ND, effectively no or very little inhibition (<10%) detected at maximal concentration.
d Inhibition curve was incomplete. As a result, \( I_{\text{max}} \) was defaulted to 100% in order to get a converging solution for the equation.

e \( p \) values for the difference between the series of GXXGER guest peptides with either (GPP)₅ or (GPO)₃ hosts are summarized in “GPP order” and “GPP order” at the foot of the table where *, \( p < 0.05 \); **, \( p < 0.01 \); and ***, \( p < 0.005 \).
The affinity series illustrates how sequence variation influences I-domain-ligand interaction. Although the amino acids E and R are crucial for the GFOGER:I-domain complex, the aromatic ring of the F residue contributes to the binding by associating with the hydrophobic pocket on the surface of the I-domain (15). The first triplet may diminish binding if it lacks a sufficiently long hydrophobic chain in the second position.

**Fig. 5. Recombinant wild-type α1, α2, and α11 I-domain binding to GXXGER peptides.** A, anti-GST probed Western blot of the purified I domains used in Figs. 5 and 9, run on a 7% SDS-polyacrylamide gel. B, 96-well plates were coated with GXXGER peptides and recombinant I-domain binding was performed as described under “Experimental Procedures.” The residual binding in the presence of EDTA was subtracted, and the level of binding was normalized to GFOGER-binding to allow interassay comparisons. Each point represents the mean ± S.E. of the number of determinations indicated in the graph. Binding to all sequences significantly differed from zero (*p < 0.05), except for (GPP)10 and α11 binding to GQRGER. The statistical difference from GFOGER is indicated in the graph, or text. A450 values for GFOGER binding were 0.98 (α21), 1.07 (α21), and 0.65 (α11).

**Fig. 6. GXXGER peptides in (GPO)3 host.** A, platelet adhesion to GXXGER peptides in the presence of 2 mM Mg2+, 4 mM EDTA, 10 μg/ml of 6F1 or 10B12. Each point represents the mean ± S.E. of three determinations. Missing error bars were too small to show. Statistics are detailed under “Results.” B, comparison of platelet adhesion (n = 30) to GFOGER in (GPP)3 or (GPO)3 host. C, comparison of HT1080 adhesion (n = 15) to peptides in (GPP)3 or (GPO)3 host. Results are expressed as a mean ± S.E. *p < 0.0001, Student’s paired t test.

**Fig. 7. Effect of platelet activation on adhesion to GXXGER peptides in (GPP)3 host.** Platelet adhesion in the presence of 2 mM Mg2+ is shown with and without treatment with 30 μM ADP for 5 min prior to plating. 2-way ANOVA (Bonferroni’s post-test) showed the effect of ADP to be significant for adhesion to all peptides other than GFOGER and for collagen. *, p < 0.05; **, p < 0.01; ***, p < 0.001.
position. Here we show that L, and to a lesser extent M and A, may substitute for F in this context, but the efficacy of the binding depends on the activation state of the integrin. GAOGGER has previously been proposed, but not proven, to mediate \( \alpha_5 \beta_1 \)-dependent binding of collagen III, which does not contain GFOGER (25), while GMOGER and GLSGER sequences identified here are completely novel. The final hydrophobic amino acid-containing sequences, GIOGER, which occurs in collagens types I (chicken, frog) and III (chicken), and GVOGER, which occurs in collagens types VIII and X (human) would also be predicted to be a higher affinity integrin binding sequences as isoleucine and valine also have a large non-polar side chains.

No side chain contacts were observed between the hydroxyproline in the third position of the first triplet and the \( \alpha_2 \)I in the GFOGER co-crystal. Previously GFPGER has been shown to bind platelet \( \alpha_5 \beta_1 \) almost as well as GFOGER. Thus, other residues (Ser, Pro) may be substituted without wholly compromising binding to \( \alpha_2 \)I. However, while a recombinant type I collagen devoid of hydroxyproline bound \( \alpha_5 \beta_1 \) well, it had a poor interaction with both \( \alpha_2 \)I and \( \alpha_2 \)-expressing cells (43), suggesting selectivity among I domains. GLS was a substantially weaker binder than GLO for all wild-type \( \alpha_2 \)I domains, but GAS was similar to GAO, the latter presumably being a poor substrate compared with GFO by virtue of the short alanine side chain. We also identified 8 GXGER loci in the collagens containing several alternative polar residue combinations within the first triplet. GQRGER, present at locus 787 in the consensus alignment in Fig. 1 of I domain, II, and Va2 chains, was tested as an example of them. Integrin and the I domains had similar affinity to GQR as to GAS and GAO, and GQR also required integrin affinity modulation for its binding. Thus, the GXGER sequences provide a versatile means of interaction with the collagen binding integrins.

Several lines of evidence presented here demonstrate that GFOGER is a unique high affinity recognition motif. First, of all the sequences tested, GFO had the highest affinity for I domains, and no difference was observed in the level of binding between the mutants and the wild-type \( \alpha_2 \)I. Second, resting platelets were fully capable of binding GFO: inclusion of a GPVI interacting/activating host sequence did not change the binding or inhibitory capacity of GFO, neither did prior activation of platelets with ADP. As HT1080 cells had no significantly higher affinity for GFO than GLO/GLS/GMO, and their adhesion profile resembled that of a gain-of-function mutant E318W, these cells are likely to express the integrin in an active conformation.

In platelets \( \alpha_5 \beta_1 \) is thought to act as a primary adhesive receptor upon vascular trauma, detailed in the “two-site two-step” process of platelet–collagen interaction (44). Paradoxically however, in circulation \( \alpha_5 \beta_1 \) is in a resting state regulated by the divalent cation concentration in the blood. How can \( \alpha_5 \beta_1 \) operate? Platelet \( \alpha_5 \beta_1 \) has been suggested to be a shear-dependent mechanoreceptor (45), a concept which has been strengthened by studies of I-domain-containing integrins in...
leukocytes (46). It also becomes activated through inside-out signaling by soluble agonists and by GPVI (27). The concept of a need for prior activation of αβ₃ for hemostatic function was furthered by murine studies, in which platelet adhesion was markedly attenuated by the absence of GPVI (8). Here, we show that although platelet binding to the lower affinity sequences does indeed require prior activation, binding to GFOGER (and to a lesser extent GLOGER) does not, which may allow platelets to interact with such sequence-containing collagens without a preceding switch in the integrin conformation. Thus, the collagen-integrin interaction represents a continuum, defined by the spectrum of affinities of the different GEP sequences and the activation state of the integrin itself.

These results are supported by the previous observations that under flow, primary platelet adhesion is impaired by blocking the function of human α₂ (47, 48), while the case is less clear with thrombus formation in mice (49–51). Comparison of the structures of the free α₂ I-domain alone and in complex with GFOGER reveals the existence of two distinct conformations, open and closed. Freed from the constraints imposed by the remainder of the integrin, the I-domain most likely exists in equilibrium between the two conformations, with GFOGER peptide most readily able to select and stabilize the open state, i.e. activate the integrin. Whether the GFOGER-ligated and unliganded structures represent the extremes of movement within the I-domain, and whether they are also sufficient to fully describe also the inside-out activation process is currently under investigation. The concept of ligand binding-induced activation is supported by intracellular platelet signaling by GFOGER (52).³

The GXXGER sequences of various affinities in conserved loci within the collagen fibril may underpin the basis of differential cellular fine-tuning by collagens, a topic which needs further study. A simple example is that of collagen types I and III: while type I monomer has 2 GFO, 4 GLO, 2 GMO, 2 GAS, and 2 GQRGER among several other GXXGER sequences, type III has only 3 GMO, 3 GLS, and 6 GAOGER, all of relatively low affinity based on this study. This correlates well with the observed differences in studies using these collagens (5, 28, 53) and with the different αβ₃ reactivity of different collagen types (54). Thus the requirement for integrin activation in thrombus deposition, by GPVI or other agonists such as ADP, may be determined by the relative expression of collagen types in the subendothelium, which changes upon e.g. development of an atheromatous plaque (55). Further, the influence of the quaternary fiber structure on platelet function has been a long standing issue, and these results may begin to offer insights into the differential αβ₃ dependence of platelet or cell function on collagen monomers compared with fibrils (48, 56, 57). The possible modulation of integrin binding provided by the different chains in heterotrimeric collagens has not yet been fully elucidated, and it is essential to identify whether these sequences are accessible for cell interaction within an intact fibril (58), or whether they only become exposed upon degrada-

Acknowledgments—We thank Dr. Gavin Jarvis for helping with the pharmacological equation. Stephan Luis is acknowledged for preliminary testing of the GMO peptide.

REFERENCES

¹ Boot-Handford, R. P., Tuckwell, D. S., Plumb, D. A., Rock, C. F., and Poulsom, R. (2003) J. Biol. Chem. 278, 31067–31077

² Guilberg, D. E., and Lundgren-Akerlund, E. (2002) Prog. Histochem. Cyto-

³ P. Sundaesans, unpublished results.
45. Nishiya, T., Kainoh, M., Murata, M., Handa, M., and Ikeda, Y. (2002) *Blood* **100**, 136–142
46. Salas, A., Shimaoka, M., Chen, S., Carman, C. V., and Springer, T. (2002) *J. Biol. Chem.* **277**, 50255–50262
47. Verkleij, M. W., Morton, L. F., Knight, C. G., de Groot, P. G., Barnes, M. J., and Sixma, J. J. (1998) *Blood* **91**, 3808–3816
48. Vajander, P., and Lassila, R. (1999) *Arterioscler. Thromb. Vasc. Biol.* **19**, 3033–3043
49. Chen, J., Diacovo, T. G., Grenache, D. G., Santoro, S. A., and Zutter, M. M. (2002) *Am. J. Pathol.* **161**, 337–344
50. Holtkotter, O., Niewandt, B., Smyth, N., Muller, W., Hafner, M., Schulte, V., Krieg, T., and Eckes, B. (2002) *J. Biol. Chem.* **277**, 10769–10794
51. Knijpers, M. J. E., Schulte, V., Bergmeier, W., Lindhout, T., Brakebusch, C., Offermanns, S., Fassler, R., Heemskerk, J. W., and Niewandt, B. (2003) *FASEB J.* **17**, 685–687
52. Inoue, O., Suzuki-Inoue, K., Dean, W. L., Frampton, J., and Watson, S. P. (2003) *J. Cell Biol.* **160**, 769–789
53. Mori, M., Jung, S. M., Shinmyozu, K., Tomiyama, Y., Ordinas, A., and Diaz-Ricart, M. (1996) *Blood* **88**, 2081–2092
54. Saelman, E. U. M., Nieuwenhuis, H. K., Hese, K. M., de Groot, P. G., Heijnen, H. F. G., Sapa, E. H., Williams, S., McKeown, L., Grahnick, H. R., and Sixma, J. J. (1994) *Blood* **83**, 1244–1250
55. Fitzsimmons, C., and Shanahan, C. (2002) in *Pan Vascular Medicine* (Lanzer, P., and Topol, E. J., eds) pp. 217–231, Springer-Verlag, Berlin, Heidelberg
56. Savage, B., Ginsberg, M. H., and Ruggeri, Z. M. (1999) *Blood* **94**, 2704–2715
57. Koyama, H., Raines, E. W., Bornfeldt, K. E., Roberts, J. M., and Ross, R. (1996) *Cell* **87**, 1069–1076
58. Sweeney, S. M., DiLallo, G., Slater, S. J., Martinez, J., Iozzo, R. V., Lauer-Fields, J. L., Fields, G. B., and San Antonio, J. D. (2003) *J. Biol. Chem.* **278**, 30516–30524
Integrin Activation State Determines Selectivity for Novel Recognition Sites in Fibrillar Collagens
Pia R.-M. Siljander, Samir Hamaia, Anthony R. Peachey, David A. Slatter, Peter A. Smethurst, Willem H. Ouwehand, C. Graham Knight and Richard W. Farndale

J. Biol. Chem. 2004, 279:47763-47772.
doi: 10.1074/jbc.M404685200 originally published online September 2, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M404685200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/09/16/M404685200.DC1

This article cites 55 references, 35 of which can be accessed free at
http://www.jbc.org/content/279/46/47763.full.html#ref-list-1