Identification of the First Prokaryotic Collagen Sequence Motif That Mediates Binding to Human Collagen Receptors, Integrins $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$ *

Received for publication, September 4, 2008, and in revised form, October 15, 2008. Published, JBC Papers in Press, November 5, 2008, DOI 10.1074/jbc.M806865200

Clayton C. Caswell†, Malgorzata Barczyk‡, Douglas R. Keene§, Ewa Lukomska§, and Sławomir Lukomski §†

From the †Department of Microbiology, Immunology, and Cell Biology and the §Mary Babb Randolph Cancer Center, West Virginia University School of Medicine, Morgantown, West Virginia 26506, the ‡Department of Biomedicine, University of Bergen, 5009 Bergen, Norway, and the ¶Shriners Hospital for Children, Portland, Oregon 97201

Many pathogenic bacteria interact with human integrins to enter host cells and to augment host colonization. Group A Streptococcus (GAS) employs molecular mimicry by direct interactions between the cell surface streptococcal collagen-like protein-1 (Scl1) and the human collagen receptor, integrin $\alpha_2\beta_1$. The collagen-like (CL) region of the Scl1 protein mediates integrin-binding, although, the integrin binding motif was not defined. Here, we used molecular cloning and site-directed mutagenesis to identify the GLPGER sequence as the $\alpha_2\beta_1$ and the $\alpha_{11}\beta_1$ binding motif. Electron microscopy experiments mapped binding sites of the recombinant $\alpha_2$-integrin-inserted domain to the GLPGER motif of the recombinant Scl (rScl) protein. rScl proteins and a synthetic peptide harboring the GLPGER motif mediated the attachment of C2C12-$\alpha_2$+ myoblasts expressing the $\alpha_2\beta_1$ integrin as the sole collagen receptor. The C2C12-$\alpha_{11}$+ myoblasts expressing the $\alpha_{11}\beta_1$ integrin also attached to GLPGER-harboring rScl proteins. Furthermore, the C2C12-$\alpha_{11}$+ cells attached to rScl more efficiently than C2C12-$\alpha_2$+ cells, suggesting that the $\alpha_{11}\beta_1$ integrin may have a higher binding affinity for the GLPGER sequence. Human endothelial and dermal fibroblasts adhered to rScl proteins, indicating that multiple cell types may recognize and bind the Scl proteins via their collagen receptors. This work is a stepping stone toward defining the utilization of collagen receptors by microbial collagen-like proteins that are expressed by pathogenic bacteria.

The integrins are a family of $\alpha\beta$ heterodimeric cell surface receptors found in metazoans that bind proteins of the extracellular matrix (ECM) and function in diverse cellular processes, including cell adhesion, spreading, and migration. There are currently 18 $\alpha$ and 8 $\beta$ known subunits that combine to form 24 distinct heterodimers with various ligand binding specificities (1). One group of integrins binds several collagen types with varying affinities and includes the integrins $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_10\beta_1$, and $\alpha_{11}\beta_1$ (2). The $\alpha$ subunit of these integrins contains an inserted (I) domain, and the $\alpha I$ domain confers binding specificity to the integrin heterodimer. Over the past decade much emphasis has been directed toward defining the parameters of collagen recognition by these integrins, as well as the specific amino acid sequences in collagens that are required for binding.

The GFOGER (where O is hydroxyproline) sequence has been identified in some collagens as a major binding motif for the $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins (3, 4). However, not all collagen types harbor the GFOGER motif. Using synthetic collagen peptides, additional sequences with the base sequence GXOGER were proposed to bind the $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$ integrins with lower affinity as compared with the GFOGER sequence (5, 6). The hydroxylated proline residue of GXOGER is required for binding by the $\alpha_2\beta_1$ integrin but not by the $\alpha_{11}\beta_1$ integrin (7). Furthermore, the E residue of the GER triplet is essential for integrin binding because of its function of protruding metal ions into the metal ion-dependent adhesion site of the $\alpha I$ domain (8, 9). Less is known about the binding specificity of the $\alpha_{10}\beta_1$ and $\alpha_{11}\beta_1$ integrins. The recombinant $\alpha_{11}I$ domain binds both fibrillar and non-fibrillar collagens, although the target sequences were not defined (10). The recombinant $\alpha_{11}I$ domain binds the sequence GFOGER, similar to $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$ integrins (5, 11). To date, no information is available regarding the requirement of the hydroxyproline residue for collagen binding to the $\alpha_{10}\beta_1$ or $\alpha_{11}\beta_1$ integrin.

The streptococcal collagen-like proteins, Scl1 and Scl2, are cell surface molecules of the human pathogenic bacterium Streptococcus (GAS) that contain tracks of repeated GXY sequence (12–16). Similar to mammalian collagens, the collagen-like (CL) regions of Scl proteins adopt stable triple helices with midpoint melting temperatures of about 37 °C (17, 18). Because of the lack of prolyl hydroxylase in GAS, Scl proteins are devoid of O residues, which in mammalian collagens are regarded as essential for triple helix stabilization. It was recently reported that electrostatic stabilization by charged residues in an Scl2 protein was similar to the stabilizing contribu-

---

* This work was supported, in whole or in part, by National Institutes of Health Grant AI50666 (to S. L.). 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: West Virginia University, Health Sciences North, Dept. of Microbiology, Immunology, and Cell Biology, Box 9177, Morgantown, WV 26506. Tel.: 304-293-6405; Fax: 304-293-7328; E-mail: stukomska@hsc.wvu.edu.

‡ The abbreviations used are: I, inserted domain; GAS, group A Streptococcus; rScl, (recombinant) streptococcal collagen-like protein; V, variable; CL, collagen-like; Col, collagen; BSA, bovine serum albumin; EM, electron microscopy; HUVEC, human umbilical vein endothelial cells; HDFa, human dermal fibroblasts.
tion of the O residues in mammalian collagens (19). A recombinant Scl1-derived protein binds to the α5β1 integrin on human cells and this interaction resulted in the phosphorylation of integrin pathway-associated cell signaling proteins (20). Thus, Scl proteins are both structural and functional mimics of human collagen. The native Scl1 protein expressed on the GAS cell surface mediates GAS internalization by human cells upon binding to the α5β1 integrin (21). However, the sequence motif within the Scl1-CL region that mediates integrin binding has yet to be identified. The Scl proteins are members of a new and growing family of “prokaryotic collagens” that are predicted to exist in nature. Searches in silico revealed that numerous bacterial and viral genomes harbor ORFs encoding collagen-like sequences (22). Other described examples of prokaryotic collagens include the Scl3/Sc1C protein of S. equi (23, 24), the Bacillus collagen-like proteins BclA and BclB present on anthrax spore surface (25–27), and the pneumococcal collagen-like cell surface protein, PclA (28). The biological function(s) of the prokaryotic collagens has only recently been subjected to investigations, and sequence motifs that mediate their interactions with human receptors have not yet been identified.

In this study we determined that the GLPGER sequence mediates direct binding to the α5β1 and α1β1 integrins. A synthetic collagen peptide (GPP)5GLPGER(GPP)5 and recombinant Scl proteins with single and multiple GLPGER motifs conferred the attachment of human cells expressing these integrins. Taken together, this study identifies for the first time a prokaryotic sequence GLPGER, which is found in the collagenous domain of the Scl1 protein, as a functional integrin binding motif. Identifying the molecular determinants of Scl-integrin interactions may well offer insight into interactions between human collagen receptors and other yet uncharacterized prokaryotic collagen-like proteins.

**EXPERIMENTAL PROCEDURES**

**Recombinant Proteins**—The *Escherichia coli* strain DH5α was used in cloning experiments and *E. coli* BL21 was used for the expression of recombinant proteins. *E. coli* strains harboring plasmid constructs were routinely grown in Luria-Bertani (LB) liquid medium (BD Biosciences) supplemented with ampicillin (100 μg/ml). Recombinant Scl (rScl) proteins (Table 1) were produced using the Strep-tag II cloning, expression, and purification system (IBA-GmbH, Goettingen, Germany), as described previously (17, 18).

To map the region responsible for integrin binding, DNA fragments encoding various portions of the CL region of P176 were amplified by polymerase chain reaction (PCR) and cloned into the construct pSL163 encoding the integrin binding-negative rScl protein P163. DNA fragments were amplified using the primers: GXY 1–30, P176-CL1-F (5'-GAGAGGTCTCACCAGCTTGAGCTCCGG); GXY 15–30, P176-CL15-F (5'-GAGAGGTCTCGCTGAGCGGCACCC-ACAACTGAAAGGCGG) and P176-CL30-R (above). The amplified DNA fragments were treated with BsaI restriction endonuclease and cloned in-frame into a single BsaI site in the pSL163-CL region, resulting in constructs pSL204 and pSL205, respectively. Additionally, GXY repeats 1–15 were also cloned as a single (P214) or multiple insertions (P215, P218, P219, and P220).

Site-directed mutagenesis was performed using the QuikChange® kit (Stratagene). To mutate the glutamic acid residue to alanine (construct P209) or to aspartic acid (construct P210) in the GLPGER motif of P204, a polymerase chain reaction was performed using the primer pairs P204-E2A-F (5’-GGTATCTGGTTA ACCTAGA) and P204-E2A-R (TCTAGGTCTCGTGACCGAGGTAACACGAGGTAAC) or P204-E2D-F (GGTATCTGGTTA CAGGAGGACCTAGA) and P204-E2D-R (TCTAGGTCTCGTGACCGAGG TAACACGAGGTAAC), respectively. Site-directed mutagenesis was also employed to create the GLPGER motif in the CL region of P163 by incorporating two single amino acid substitutions (P→L [CCT→CTT] and A→P [GCT→CTT]) in the GPA triplet of the GPAGER sequence. A polymerase chain reaction with the primers P163-GLPGER-For (5’-GCAGGTCAATGGGTCTTCTCTGGTGAGCAGGTGGTACACGAGGTAAC) and P163-GLPGER-Rev (5’-CCCTTCTACCTCGTCACAGGAGAAGCATTGGACACTG) was used to amplify the mutated plasmid that resulted in pSL222. All constructs were verified by DNA sequencing.

**Rotary Shadowing and Electron Microscopy**—rScl proteins and rScl-rα1 complexes were rotary shadowed and analyzed by EM as described previously (17, 20). Briefly, recombinant proteins were dialyzed against 0.1 M sodium bicarbonate containing 1 mM magnesium chloride. To study interactions, rScl proteins were mixed with rα1 domain at a molar ratio of 1:3 and incubated at room temperature for 3 h. rScl-rα1 domain combinations were mixed with glycerol to a final glycerol concentration of 70%, and 100 μl of each sample was nebulized with an airbrush onto freshly cleaved mica. Samples were then dried in a vacuum and rotary shadowed with carbon/platinum using an electron beam gun tilted at an angle of 6 degrees relative to the mica surface in a Balzers BAE 250 evaporator. The replicas were backed with carbon at 90 degrees, and then were floated in distilled water and picked up onto bare 600-mesh copper grids. Photomicrographs were then taken using a Philips 410 electron microscope operated at 80 kV with a 30-μm objective aperture.

**Cell Attachment Assays**—The ability of cells to use rScl proteins for attachment was tested using a previously described assay (20). BSA or type I collagen (10 μg/well) or rScl proteins (100 nm) in Tris-buffered saline (TBS) were coated overnight onto 24-well tissue culture plates at 4 °C. The following day, wells were blocked with 1% BSA in TBS at room temperature for 2 h. The model eukaryotic C2C12 cells (described below) were detached using a non-enzymatic cell dissociation solution (CellStripper™, Mediatech, Inc., Herndon, VA). Cells were washed twice and suspended in appropriate serum-free cell culture medium containing 2 mM Mg2+. Approximately 5 × 104 cells were added to each well, and the plates were incubated at 37 °C for 60 min. Following washing with phosphate-buffered saline, bound cells were fixed in 3% p-formaldehyde for 15 min and stained with 2% crystal violet for 5 min. Crystal violet was eluted with 100 mM sodium citrate in 50% ethanol, and absorb-

3 M. Pawlowski and J. M. Bujnicki, unpublished data.
ance was recorded at 595 nm. For visualization, digital images of attached cells were captured prior to the elution of crystal violet using a ×50 magnification on a Zeiss Axiovert 40 CFL microscope. Images were processed using AxioVision software (Zeiss).

The mouse myoblast cell lines C2C12, C2C12-α2+, and C2C12-α11+ were used as model cells. The C2C12 cells express no collagen binding integrins, while the C2C12-α2+ and C2C12-α11+ derivatives stably express the corresponding α subunits and, therefore, form αβ1 and αβ1β1 heterodimers, respectively (29). C2C12 and C2C12-α2+/C2C12-α11+ cells were cultured in DMEM/High Glucose medium (Hyclone, Logan, UT) supplemented with 10% FBS in the absence or presence of 10 μg/ml puromycin, respectively.

In some experiments, a GLPGER-containing peptide was employed in cell attachment assays to assess integrin-binding. The (GPP)5GLPGER(GPP)5 peptide, designated GP36, was employed in cell attachment assays to assess integrin-binding. The mouse myoblast cell lines C2C12, C2C12-α2+, and C2C12-α11+ were used as model cells. The C2C12 cells express no collagen binding integrins, while the C2C12-α2+ and C2C12-α11+ derivatives stably express the corresponding α subunits and, therefore, form αβ1 and αβ1β1 heterodimers, respectively (29). C2C12 and C2C12-α2+/C2C12-α11+ cells were cultured in DMEM/High Glucose medium (Hyclone, Logan, UT) supplemented with 10% FBS in the absence or presence of 10 μg/ml puromycin, respectively.

In some experiments, a GLPGER-containing peptide was employed in cell attachment assays to assess integrin-binding. The (GPP)5GLPGER(GPP)5 peptide, designated GP36, was synthesized and HPLC-purified by Chi Scientific (Maynard, MA). The GLPGER guest sequence is embedded in a GPP host sequence, which stabilizes the triple helix (3). Cell attachment was carried out in 24-well tissue culture plates coated with 10 μg/well of GP36 using the C2C12-α2+ myoblasts, the human umbilical vein endothelial cells (HUVEC), and adult human dermal fibroblasts (HDFa; Cascade Biologics, Portland, OR).

**RESULTS**

**Mapping of αβ1 Integrin Binding Region in rScl1**—Our previous study reported binding between the CL region of the recombinant P176 construct, derived from Scl1.41 protein, and the αβ1 integrin on human fibroblasts (20); however, the integrin-binding motif on P176 was not defined. To map the αβ1 integrin binding site of P176 we inserted fragments of the P176-CL region into the integrin binding-negative protein, P163 (Table 1). GXY repeats 1–30 and 15–30 were inserted by cloning to create the constructs P204 and P205, respectively. Mouse myoblasts C2C12 devoid of collagen-binding integrins and the C2C12-α2+ cell line, which stably expresses the αβ1 integrin as the only collagen receptor (29), were used in a cell attachment assay to assess integrin binding to rScl constructs (Fig. 1A and B). As expected, the C2C12 control cells did not attach to any of these rScl constructs nor to control human type I collagen. The C2C12-α2+ cells attached to type I collagen and to rScl proteins P176 and P204. Attachment of C2C12-α2+ was not observed on P205 and P163 negative control. These results indicated that the first 15 GXY repeats in the P176-CL region harbor an integrin binding motif.

**Site-directed Mutagenesis of GLPGER Motif Inhibits Integrin Binding**—The P176-CL region contains the sequence GLPGER at GXY repeats 8–9, and we hypothesized that this motif mediates binding to the αβ1 integrin. The GLPGER sequence is present in the integrin binding positive construct P204 but not in the integrin binding negative P205. Earlier studies determined that the glutamyl residue in the mammalian integrin binding motif GFOGER is essential for binding to the αβ1 integrin. The GLPGER...
sequence of P204 was replaced with an alanine (E → A) or aspartic acid (E → D) residue to create the rScl constructs P209 and P210, respectively (Table 1). Both P209 and P210 failed to support attachment of C2C12-α2+ cells (Fig. 1, A and B), further indicating that GLPGER in the CL region of the P176 protein is involved in integrin binding.

GLPGER Sequence in rScl Is Sufficient for Integrin Binding—We next created the GLPGER motif in the CL region of the P163 protein, which is negative for integrin binding, by site-directed mutagenesis. Two amino acids were replaced (GPA → GLP), and the resulting construct was designated P222. Recombinant P222 supported attachment of C2C12-α2+ cells, but not the C2C12 control cells, indicating that replacement of the two amino acids in P163 rendered its ability to bind integrins (Fig. 1, A and B). Additionally, the synthetic GP36 peptide, which contains a central GLPGER motif flanked on each side by five GPP repeats, supported attachment of C2C12-α2+ cells, but not the C2C12 control cells (Fig. 1A). Altogether, our experiments employing recombinant collagen-like molecules and a synthetic collagen peptide, conclusively demonstrate that the GLPGER sequence found in prokaryotic Scl1 proteins is sufficient to mediate interactions with the α2β1 integrin expressed on the surface of eukaryotic cells.

Visualization of α2 Domain Binding to the GLPGER Motif—The inserted (I) domains of the α subunits of collagen-binding integrins confer collagen binding (2) and recombinant αI domains (rαI) have been used as binding indicators to collagen sequences (10). Here, we used the rαI to map the binding sites present in rScl constructs. The P222 protein containing a single GLPGER motif was incubated with rαI, and the mixture was rotary shadowed and viewed by electron microscopy (EM) (Fig. 2). Our earlier EM studies revealed that rScls share a common “lollipop-like” two-domain organization composed of the globular head formed by the V region and the lollipop stalk made by the fibrous CL region (17, 18). Binding events between the rαI domain and the CL region of P222 were observed in electron micrographs, and the location of bound rαI appeared to be the same in each case. Based on the translation value of 0.286 nm per residue in collagen triple helix (30), we calculated that the GLPGER motif should be located ~31 nm from the base of the globular domain. The measured distance from the base of the P222-V domain to the center of bound rαI was 30.6 ± 0.4 nm, which correlated well with the calculated value. Hence, we physically mapped the rαI binding site on P222 to the GLPGER motif.

Effect of Multiple GLPGER Motifs in the Scl-CL Region on Integrin Binding—To study the effects of increasing numbers of the GLPGER motif in the Scl-CL region, we generated the constructs P214, P215, P218, P219, and P220 (Table 1). The first 15 GXY repeats from the P176-CL region, which contain the sequence GLPGER, were inserted by cloning into the CL region of the integrin binding-negative protein P163 resulting in P214. Multi-insertion constructs were also engendered containing 2, 3, 4, or 5 tandem copies of the 15GXY-GLPGER repeats to produce P215, P218, P219, and P220, respectively. The purity and integrity of the rScl proteins was analyzed by SDS-PAGE and Western immunoblotting (Fig. 3A). The rotary-shadowed samples were viewed by EM to reveal the characteristic structural organization of the rScls with increasing lengths of their CL regions and indicating that they form triple helices (Fig. 3B).

All rScl constructs were tested in a cell attachment assay for mediating the attachment of C2C12-α2+ cells after 20, 40, and 60 min (Fig. 3C). As expected, C2C12 control cells did not adhere to any of the rScl proteins or type I collagen during this time period. C2C12-α2+ cells expressing the α2β1 integrin readily attached to wells coated with type I collagen but not to the negative controls coated either with P163 protein or BSA after the maximum of 60 min of incubation. Binding of C2C12-α2+ cells to P176, P214, P215, P218, P219, and P220 occurred as soon as 20 min into the experiments. We observed that cell attachment was increasing on constructs P176<P214<P215, but remained comparable for constructs P218, P219, and P220. Similar attachment trends were recorded at 40- and 60-min time points.

EM analyses of rαI domain binding to each rScl construct were also carried out (Fig. 4). The distances of bound rαI domains measured from the rScl-V borders were compared with the calculated locations of the GLPGER sequences in the rScl-CL regions (Fig. 4A). For each of the rScl constructs, rαI appeared to bind a site on the rScl collagenous tail corresponding to the location of a GLPGER motif. In constructs containing
multiple GLPGER sequences, single binding events were observed between $\alpha_2\beta_1$ domains and all GLPGER sites. However, despite the presence of multiple GLPGER sequences in constructs P215, P218, P219, and P220, simultaneous binding of multiple $\alpha_2\beta_1$ domains was only observed for P220, where up to three $\alpha_2\beta_1$ molecules bound to a single P220-CL region (Fig. 4B). Thus, our mapping experiments employing $\alpha_2\beta_1$ confirmed the $\alpha_2\beta_1$ integrin binding to multiple GLPGER sequences in rScl proteins, which suggests that prokaryotic molecules containing multiple binding motifs may exhibit increased integrin binding.

**Binding of the $\alpha_1\beta_1$ Integrin Collagen Receptor to rScl Proteins**—The $\alpha_1\beta_1$ integrin has similar binding specificities as $\alpha_2\beta_1$ toward fibrillar collagens (11). However, recognition and binding of the $\alpha_1\beta_1$ to unhydroxylated collagens has not been studied. We carried out cell attachment assays to rScls using C2C12-$\alpha_1\beta_1$ myoblasts, which stably express the $\alpha_1\beta_1$ integrin as the sole collagen receptor (29) (Fig. 5A). The control C2C12 cells did not bind to any rScl construct or to human collagens type I and type IV. On the contrary, the C2C12-$\alpha_1\beta_1$ cells efficiently attached to wells coated with both collagen types and with P176, P204, P214, and P222. The C2C12-$\alpha_1\beta_1$ cells did not attach to P209 or P210 that contain single amino acid substitutions of the glutamyl residue of GLPGER (Table 1).
A

B

FIGURE 5. α11β1 Integrin-mediated binding of myoblasts to rScl constructs. A, cell attachment assay with α1,β1 integrin-expressing C2C12-α11+ myoblasts. 10 μg of BSA, collagens I (ColI), and IV (ColIV), and fibronectin (FN), or 100 nM rScl proteins were coated onto 48-well tissue culture plates, and following blocking, wells were incubated for 60 min with C2C12 myoblasts or C2C12-α11+ myoblasts, which stably express the α11β1 integrin as the sole collagen receptor. Attached cells were fixed and stained with crystal violet, and the optical density of eluted crystal violet was recorded at 595 nm. Each data point represents the average OD value of duplicate wells, and error bars represent the S.D. Concentration-dependent cell attachment of C2C12 cells expressing α11β1 and α11β1 integrins to P176. Increasing concentrations of P176 (0–100 nM) were coated onto tissue culture wells. Equal numbers of C2C12 (○), C2C12-α11+ (●), or C2C12-α11+ (▲) myoblasts were added to the wells and incubated for 60 min. Attached cells were fixed and stained with crystal violet, and the OD of eluted crystal violet was recorded at 595 nm. Each data point represents the average OD value of duplicate wells, and error bars depict the S.D.

we conclude that cell attachment was due to the specific interaction between the α11β1 integrin and the GLPGER sequence of rScls because: (i) the C2C12-α11+ cells did not bind to P163 lacking GLPGER; (ii) P176, P204, P214, and P222, all containing the GLPGER sequence, supported cell attachment; (iii) glutamyl replacement in the GLPGER sequence of P204 (P209 and P210 constructs) disabled cell attachment; and (iv) the C2C12-α11+ cells attached to both eukaryotic and prokaryotic collagens, while the C2C12 cells devoid of collagen receptors did not, although they attached to fibronectin-coated wells. Hence, our data extend the spectrum of human collagen receptors that interact with Scl proteins via the GLPGER motif.

Titration curves for C2C12-α1, and C2C12-α11+ binding to increasing concentrations of P176 were obtained by incubation with equal numbers of both cell types (Fig. 5B). Cells expressing the α11β1 integrin bound better than did the α2β1 integrin-expressing cells, which was more apparent at lower coating concentrations of P176. Importantly, cell adhesion of C2C12-α1, and C2C12-α11+ to type I collagen was comparable (data not shown). These results suggest that the α11β1 integrin binds better to the GLPGER sequence in rScl proteins as compared with the α2β1 integrin.

Multiple Eukaryotic Cell Types Adhere to Scl1 via the GLPGER Motif—Primary adult HDFa and HUVEC were tested for their ability to attach to rScl constructs containing the GLPGER motif (P176 and P222), as well as to a GLPGER-containing synthetic peptide (GP36). The HDFa or HUVEC were incubated in the wells coated with BSA, type I collagen, rScl proteins, and the GP36 peptide for 60 min. (Fig. 6, A and B). Both cell lines attached to type I collagen, as well as to GLPGER-containing rScls (P176 and P222) and the GP36 peptide. Cells did not attach to P163 and BSA negative controls. These results indicate that both HDFa and HUVEC express collagen receptors that bind to the GLPGER motif harbored within the CL region of rScl proteins.

DISCUSSION

The mammalian collagen receptors, the integrins α1β1, α2β1, α10β1, and α11β1 (2), recognize the primary sequence GFOGER (O is hydroxyproline) containing an essential GER motif, which they bind with high affinity (8). In addition, the collagen integrin receptors may recognize and bind derived sequences GXOGEY (X and Y are other amino acids) that they bind with varying specificities and affinities including the GXY triplets GLO, GMO, and GAO, as well as the GEY triplets GEK and GEN (4, 5, 6, 32). Moreover, the ligand recognition depends on integrin activation state (5), collagen hydroxylation (7, 33), and its higher order organization (10, 34). Collagens also contain embedded RGD sequences that are cryptic within the triple helix but become available for binding during tissue remodeling or wound healing (35–37).

The prokaryotic collagens, which include the Scl proteins, are different in many respects since they are unhydroxylated, homotrimeric, and often use GXY repeats that are uncommon in human collagens and vice versa. While some triplets identified in human collagens that are part of the integrin binding sites, such as GFO, GMO, and GEN, are never found within the Scl-CL regions, other triplets may occur frequently, including GER (3.2%), GEK (7.08%), and GLP (4.89%) (18). Interestingly, the RGD cryptic motif is common in the Scl2-CL regions and is also found in recently reported ScIZ protein (38). Prokaryotic collagens may have evolved different integrin binding specificities as compared with their mammalian counterparts, and the threshold that is sufficient to induce integrin-mediated cell signaling is hard to predict based purely on sequence analysis. Therefore, the spectrum of sequence recognition in the prokaryotic collagen-like proteins, including Scl-CL regions, by human collagen receptors is currently unknown. The focus of this work was to define, for the first time, the recognition sequence for the α2β1 and α11β1 integrins in the prokaryotic
collagen-like protein, Sc1, and set up the foundation for future investigations on interactions between prokaryotic collagens and human collagen binding integrins.

In the present study, we demonstrated that the GLPGER sequence in the collagen-like (CL) region of Sc1 proteins is responsible for mediating Sc1 interactions with the $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$ integrins. Sc1 binding to $\alpha_{11}\beta_1$ integrin has not been previously reported. Earlier work concluded that $\alpha_1\beta_1$ integrins bind the GFOGER and GLOGER sequences (11); however, the requirement for the presence of hydroxyprolines by $\alpha_{11}\beta_1$ was not assessed. Our present data show that proline hydroxylation in collagen is not necessary for binding by this integrin. These data also suggest that $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$ integrins have similar binding specificities to Sc1 proteins and, perhaps, to other unhydroxylated prokaryotic collagen-like proteins. In conclusion, the present study has demonstrated that the GLPGER sequence harbored in the CL region of Sc1 proteins is the first defined prokaryotic collagen sequence motif that mediates binding to integrins $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$. Because the number of novel integrin binding motifs continuously increases (6, 32) and because GXY sequences found in prokaryotic collagen-like proteins vary significantly, our determination of the $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$ binding site in Sc1 sheds light on the ability of other pathogenic bacteria to bind human collagen receptors.

REFERENCES

3. Knight, C. G., Morton, L. F., Peachey, A. R., Tuckwell, D. S., Farndale, organ during infections of various anatomical sites in human body. The collagen-like sequence is determined by the presence of the triglycine repeat region, (Gly-Xaa-Yaa)$_n$, or (GXY)$_n$, and tracks of GXY repeats have been identified not only in proteins of higher and lower eukaryotes (40), but also in several bacteria and phages (Ref. 22). It remains unclear whether the collagen-like sequence was transferred to prokaryotes, as it has been postulated for the collagen-like protein (CLP) of the cyanobacterium *Trichodesmium erythraeum* (31), or it evolved independently. Why have pathogenic bacteria adapted CLPs, and what roles do they play in the pathogenesis caused by these organisms? Genes putatively encoding CLPs have been identified in the genomes of several pathogenic streptococci. For example, the collagen-like protein PclA of *S. pneumoniae* is a cell surface protein that was shown to play a role in pneumococcal adherence to and internalization by human cells (28). Most recently, the sequence of the *S. zooepidemicus* genome was reported to harbor 12 distinct genes encoding cell surface CLPs designated Sc1Z.1–12 (38). The authors describe that the Sc1Z proteins contain potential integrin-binding motifs, including RGD, KGD, and GXPGER. This work supports our notion that other prokaryotic CLPs may bind to collagen receptors. The ongoing support of Fran Rubin is greatly appreciated.

 acknowledgments—we thank neung-seon seo and magnus höök for providing the recombinant $\alpha_2$ domain and for helpful discussion. we thank lisa salati for the critical reading of the manuscript. we also thank sarah tufo for assistance with em and fred minnear for providing huvecs. the ongoing support of fran rubin is greatly appreciated.

36174 JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 283 • NUMBER 52 • DECEMBER 26, 2008
Identification of the First Prokaryotic Collagen Sequence Motif That Mediates Binding to Human Collagen Receptors, Integrins $\alpha_2\beta_1$ and $\alpha_{11}\beta_1$

Clayton C. Caswell, Małgorzata Barczyk, Douglas R. Keene, Ewa Łukomska, Donald E. Gullberg and Slawomir Lukomski


Access the most updated version of this article at doi: 10.1074/jbc.M806865200

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

This article cites 39 references, 19 of which can be accessed free at http://www.jbc.org/content/283/52/36168.full.html#ref-list-1