

# Recombinant Collagen Engineered to Bind to Discoidin Domain Receptor Functions as a Receptor Inhibitor\*

Received for publication, June 23, 2015, and in revised form, December 15, 2015. Published, JBC Papers in Press, December 23, 2015, DOI 10.1074/jbc.M115.674507

Bo An<sup>‡1</sup>, Vittorio Abbonante<sup>§1</sup>, Huifang Xu<sup>¶2</sup>, Despoina Gavriilidou<sup>¶</sup>, Ayumi Yoshizumi<sup>||</sup>, Dominique Bihan<sup>\*\*</sup>, Richard W. Farndale<sup>\*\*</sup>, David L. Kaplan<sup>‡</sup>, Alessandra Balduini<sup>‡5</sup>, Birgit Leitinger<sup>¶3</sup>, and Barbara Brodsky<sup>‡4</sup>

From the <sup>‡</sup>Department of Biomedical Engineering, Tufts University, Medford, Massachusetts 02155, the <sup>§</sup>Department of Molecular Medicine, Istituto di Ricerca e Cura a Carattere Scientifico San Matteo Foundation, University of Pavia, 27100 Pavia, Italy, the <sup>¶</sup>Molecular Medicine Section, National Heart and Lung Institute, Imperial College London, London SW7 2AZ, United Kingdom, the <sup>||</sup>Department of Microbiology and Infectious Diseases, Faculty of Medicine, Toho University School of Medicine, Tokyo 143-8540, Japan, and the <sup>\*\*</sup>Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, United Kingdom

A bacterial collagen-like protein Scl2 has been developed as a recombinant collagen model system to host human collagen ligand-binding sequences, with the goal of generating biomaterials with selective collagen bioactivities. Defined binding sites in human collagen for integrins, fibronectin, heparin, and MMP-1 have been introduced into the triple-helical domain of the bacterial collagen and led to the expected biological activities. The modular insertion of activities is extended here to the discoidin domain receptors (DDR), which are collagen-activated receptor tyrosine kinases. Insertion of the DDR-binding sequence from human collagen III into bacterial collagen led to specific receptor binding. However, even at the highest testable concentrations, the construct was unable to stimulate DDR autophosphorylation. The recombinant collagen expressed in *Escherichia coli* does not contain hydroxyproline (Hyp), and complementary synthetic peptide studies showed that replacement of Hyp by Pro at the critical Gly-Val-Met-Gly-Phe-Hyp position decreased the DDR-binding affinity and consequently required a higher concentration for the induction of receptor activation. The ability of the recombinant bacterial collagen to bind the DDRs without inducing kinase activation suggested it could interfere with the interactions between animal collagen and the DDRs, and such an inhibitory role was confirmed *in vitro* and with a cell migration assay. This study illustrates that

recombinant collagen can complement synthetic peptides in investigating structure-activity relationships, and this system has the potential for the introduction or inhibition of specific biological activities.

Genomic studies of prokaryotic organisms identified more than 100 genes that encode proteins with collagen-like (Gly-Xaa-Yaa)<sub>n</sub> repeating sequences (1). A number of these bacterial collagen-like proteins have been expressed in recombinant systems, and all formed triple-helical structures with stability close to the  $T_m = 37^\circ\text{C}$  found for mammalian collagens (2). These proteins lack the post-translational modification of Pro to 4-hydroxyproline (Hyp),<sup>5</sup> which is essential for triple helix stabilization in animal collagens, and alternative stabilization strategies are utilized (3, 4). Pure bacterial collagen-like proteins can be produced in high yield in a recombinant *Escherichia coli* system where their sequences can be easily modified, making them an attractive source of recombinant collagenous material for bioengineering and biomedical applications (5–7). One collagen-like protein, Scl2 (*Streptococcus* collagen-like protein 2) from the Gram-positive bacterium *Streptococcus pyogenes*, has been extensively characterized (8, 9). The Scl2 collagen domain appears to be biologically inert, making it an excellent “blank slate” framework for the introduction of specific activities found in animal collagens (5, 10, 11). Human collagens are known to interact with cell surface receptors, extracellular matrix proteins, proteoglycans, glycosaminoglycans, and enzymes, and the specific (Gly-Xaa-Yaa)<sub>n</sub> sequences responsible for an increasing number of interactions have been determined through protein and synthetic peptide studies (12–17). Identification of specific ligand binding sequences in human collagen presents an opportunity to insert defined biological activities in a stable triple-helical scaffold provided by recombinant bacterial collagens such as Scl2.

The development of a recombinant bacterial collagen system with inserted human bioactivities depends on a modular model

\* This work was supported by National Institutes of Health Grants EB011620 (to B. B. and D. L. K.) and GM60048 (to B. B.), CARIPLO13-ICH (Cariplo Foundation) Contract 2013.0717 (to A. B.), and in part by Medical Research Council United Kingdom Grant G0701121 (to B. L. and R. W. F.) and Biotechnology and Biological Sciences Research Council United Kingdom Grant BB/I011226/1 (to B. L.). Peptide synthesis was supported by Biomedical Resource Grant 094470/Z/10/Z (to R. W. F.) from the Wellcome Trust. The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

✂ Author's Choice—Final version free via Creative Commons CC-BY license.

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> Present address: Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China.

<sup>3</sup> To whom correspondence may be addressed: Molecular Medicine Section, NHLI, Imperial College London, Sir Alexander Fleming Bldg., London SW7 2AZ, United Kingdom. Tel.: 44-20-7594-1591; Fax: 44-20-7594-3100; E-mail: b.leitinger@imperial.ac.uk.

<sup>4</sup> To whom correspondence may be addressed: Dept. of Biomedical Engineering, Tufts University, Medford, MA 02155. Tel.: 617-627-3064; Fax: 617-627-3231; E-mail: barbara.brodsky@tufts.edu.

<sup>5</sup> The abbreviations used are: Hyp, 4-hydroxyproline; DDR, discoidin domain receptor; ECM, extracellular matrix; VWF, von Willebrand factor; Mk, megakaryocyte; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol. Molar concentrations of DDR-Fc listed in this work were single chain concentrations, calculated based on single chain molecular mass of DDR-Fc at ~90 kDa.

## DDR Binding Collagen-like Protein with Inhibitory Function

of collagen activity, where a specific (Gly-Xaa-Yaa)<sub>n</sub>-binding motif is responsible for a given bioactivity. Typically, short sequences of tripeptides (Gly-Xaa-Yaa)<sub>n</sub>, where  $n = 2$  to 6, define a ligand-binding site, and experiments introducing the sequences for recognizing integrins, fibronectin, and heparin into the Scl2 triple-helical domain have demonstrated the expected biological activity, both in solid state binding assays and cell culture studies (11, 18–20). In addition, insertion of the sequence for the unique human matrix metalloproteinase cleavage site of type III collagen led to specific digestion of the recombinant collagen protein at this site (21). Here, we extend the concept of designing modular recombinant collagen with separate collagen functionalities by introducing a high affinity binding site for the discoidin domain receptors (DDRs), which are collagen-activated receptor tyrosine kinases (22).

The cell surface DDR receptors are widely expressed in human tissues and play key roles in the communication of cells with the extracellular matrix. The DDRs regulate fundamental cellular functions, including cell adhesion, proliferation, and migration (22, 23). The DDR family consists of two closely related receptors, DDR1 and DDR2, that are both activated by a number of different collagen types, in particular fibrillar collagens (24, 25). Both receptors play important roles in embryo development, and alterations in DDR function have been related to organ fibrosis, osteoarthritis, and tumor progression (22, 26, 27). Collagen binding to their discoidin homology domain induces receptor autophosphorylation with slow kinetics (24, 25, 28). DDR binding to collagen requires its presentation as a native triple-helical structure. Studies using synthetic triple-helical collagen-mimetic peptides demonstrated that the major binding site for DDR1 and DDR2 in the interstitial fibrillar collagen types I–III contains an essential GVMGFO motif (where O = Hyp) (29, 30). A crystal structure of the discoidin domain of human DDR2 bound to a triple-helical peptide revealed the DDR-collagen (GVMGFO) interface at atomic level resolution (31). Peptide studies indicated additional binding sites, with sequences other than GVMGFO, in the fibrillar collagens for DDR2 but not DDR1 (29, 30).

The goal of this work was to exploit modular designs of bacterial collagens to study interactions with the DDRs. A human type III collagen sequence containing the GVMGFO-based DDR-binding site, conserved in collagen II and the  $\alpha 1$  chain of collagen I, was inserted between two triple-helical domains of the bacterial Scl2 sequence. Solid-phase binding assays demonstrated that this recombinant collagen protein bound to recombinant DDR ectodomains, as expected. However, the engineered bacterial collagen did not induce DDR activation, as assessed by receptor autophosphorylation, suggesting a more complicated story than simple modular activity. The production of a recombinant collagen, which could bind to but not activate DDR receptors, raised the possible utility of such constructs as DDR inhibitors, and this inhibitory effect was demonstrated in both a competitive binding assay as well as in megakaryocyte (Mk) migration assays.

### Experimental Procedures

Chemicals used in all experiments were purchased from Sigma unless otherwise indicated. Collagens type I and type III

preparations, provided by Dr. John Ramshaw, were pepsin-extracted from bovine skin as described (32). Hemate P (Aventis-Behring, Milan, Italy), which contains high purity human von Willebrand factor (VWF) and FVIII, was used as-is in the VWF binding assay.

**Production of Recombinant Bacterial Collagen Proteins**—The protein sequence for the bacterial collagen constructs was based on the original Scl2.28 sequence from *S. pyogenes* (9). The original protein construct with the N-terminal globular domain (V) was modified to include the following: two tandem repeats of the triple-helical domain (CL); an N-terminal hexahistidine tag for purification; and a protease-susceptible sequence (LVPRGSP) between the V and the first CL domains for V domain removal, as described previously (19, 33). Oligonucleotides encoding the type III collagen DDR-binding sequence were designed and synthesized (Invitrogen; sequences provided upon request). These were inserted as annealed dsDNA between the two CL domains of the bacterial collagen constructs through restriction sites XmaI and ApaI included in the oligonucleotide sequences. The final constructs containing the DDR-binding sites were cloned into the pColdIII vector (Takara Bio Inc.) through NdeI and BamHI restriction sites. All enzymes for cloning were purchased from New England Biolabs. DNA sequencing to confirm fidelity was carried out at the Tufts Core Facility. The DNA constructs and their subsequent proteins were denoted as SCI, for *Streptococcus* collagen-like proteins. The recombinant SCI proteins containing DDR-binding sequences were denoted as SCI-DDR<sub>4</sub> (inserted sequence, GPRGQPGVMGFP), SCI-DDR<sub>6</sub> (inserted sequence, GPSGPRGQPGVMGFPGPK), and SCI-DDR<sub>6F-A</sub> (inserted sequence, GPSGPRGQPGVMGAPGPK), respectively. The bacterial collagen sequences immediately flanking the inserted human DDR-binding sequences were...GKDGQPGPK (Insertion) GPRGEQGPT....

All SCI constructs in the pColdIII vector were expressed in *E. coli* BL21 strain, grown in 20 ml of LB medium with 100  $\mu$ g/ml ampicillin overnight at 37 °C. This starting culture was used to inoculate 500 ml of LB/ampicillin media in a shaking flask and grown at 37 °C to an  $A_{600\text{ nm}} = 0.8$ . To induce protein expression, 1 mM isopropyl  $\beta$ -D-thiogalactopyranoside (Thermo Scientific) was added to the culture, and the temperature was lowered to 22 °C. After 16 h of induction, cells were harvested by centrifugation and resuspended in His tag column binding buffer (20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl, 10 mM imidazole) containing 0.25 mg/ml lysozyme and frozen at  $-80$  °C until purification. Purification of the His-tagged SCI proteins was carried out by affinity chromatography with nickel ion binding and imidazole gradient elution on an ÄKTA pure 25L FPLC (GE Healthcare). Frozen cells were thawed and further lysed by sonication. Cellular debris was removed by centrifugation at  $8000 \times g$  at 4 °C. The supernatant containing the soluble target protein was injected onto binding buffer equilibrated 5-ml HisTrap HP columns (GE Healthcare). The column was then washed sequentially with 3 column volumes of binding buffer, binding buffer plus 50 mM imidazole, and binding buffer plus 100 mM of imidazole. The His-tagged protein was eluted by elution buffer (binding buffer plus 400 mM imidazole). Protein purity was checked by SDS-PAGE

(NuPAGE® BisTris 4–12%, Invitrogen). The protein concentration was determined using an extinction coefficient of  $\epsilon_{280} = 9970 \text{ M}^{-1} \text{ cm}^{-1}$  after dialysis into phosphate-buffered saline (PBS, pH 7.4).

**Production of Recombinant DDR Ectodomain Proteins**—The production and purification of recombinant human DDR proteins were as described previously (30). The Fc-tagged DDR1 (DDR1-Fc, with IgG2 sequence) and DDR2 (DDR2-Fc with IgG2 sequence) were isolated from episomally transfected HEK293-EBNA cells. Proteins were purified by affinity chromatography on HiTrap rProtein A column using an ÄKTA™ purifier (GE Healthcare). For some experiments, human DDR1-Fc (with IgG1 sequence) was purchased from R&D Systems (Minneapolis, MN).

**Peptide Synthesis**—The sequences of the peptides used in this study were as follows: GPC(GPP)<sub>5</sub>-GPRGQOGVMGFO-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; GPC(GPP)<sub>5</sub>-GPRGQPGVMGFO-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; GPC(GPP)<sub>5</sub>-GPRGQOGVMGFP-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; GPC(GPP)<sub>5</sub>-GPSGPRGQOGVMGFO-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; GPC(GPP)<sub>5</sub>-GPSGARGQOGVMGFO-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; GPC(GPP)<sub>5</sub>-GPSGPRGQAGVMGFO-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; GPC(GPP)<sub>5</sub>-GPSGPRGQOGVMGFA-(GPP)<sub>5</sub>GPC-NH<sub>2</sub>; and GPC(GPP)<sub>10</sub>-GPC-NH<sub>2</sub>. Peptides were synthesized by Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry as C-terminal amides on TentaGel R RAM resin in either an Applied Biosystems Pioneer or a microwave-assisted CEM Liberty Discover automated synthesizer, and purified as described (34). Peptides were verified by mass spectrometry and shown to adopt triple-helical conformation by polarimetry.

**Mass Spectrometry**—Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) was performed on a Microflex LT system (Bruker Corp., Billerica, MA) with 50% laser intensity using standard LP (linear positive) 60-kDa mode provided by the software. MALDI matrix was prepared by making a saturated sinapinic acid solution in 50% (v/v) acetonitrile and 0.3% (v/v) trifluoroacetic acid. A 6- $\mu\text{l}$  aliquot of 1  $\mu\text{g}/\text{ml}$  sample was mixed with 24  $\mu\text{l}$  of matrix, and 1  $\mu\text{l}$  of this solution was plated onto a 96 spot target plate and allowed to dry before spectra acquisition.

**Trypsin Digestion**—Purified SCL proteins in PBS buffer were incubated with 0.01 mg/ml (430 nM) trypsin at 25 °C for 120 min. The reaction was stopped by addition of phenylmethylsulfonyl fluoride (PMSF) to 1 mM. Cleavage was visualized by SDS-PAGE.

**Circular Dichroism**—CD spectra were obtained on AVIV model 420 CD spectrometer (AVIV Biomedical, Lakewood, NJ) using glass cuvettes with a 1-mm path length. Protein solutions were equilibrated for at least 24 h at 4 °C before measurement. Wavelength scans were collected from 190 to 260 nm in 0.5-nm steps with a 4-s averaging time, 1.0 nm bandwidth, and repeated three times. Temperature scans were monitored from 15 to 70 °C at 225 nm with a 10-s averaging time and a 1.5-nm bandwidth. Samples were equilibrated for 2 min at each temperature, and the temperature was increased at an average rate of 0.1 °C/min.

**Differential Scanning Calorimetry**—Differential scanning calorimetry was performed on a NANO DSC II model 6100 (Calorimetry Sciences Corp, Lindon, UT). Each sample was re-

dialyzed against PBS overnight before measurement to collect the dialyzed buffer as reference in the experiment. Sample solutions were loaded at 0 °C into the cell and heated at a rate of 1 °C/min until 100 °C.

**Dynamic Light Scattering**—DLS measurements were performed using a DynaPro Titan instrument (Wyatt Technology Corp., Santa Barbara, CA) equipped with a temperature controller using Eppendorf UVette cuvettes with 10-mm path length. Protein concentrations of the samples (in PBS) were adjusted to 1 mg/ml. Samples were centrifuged at 14,000  $\times g$  for 10 min and filtered through 0.2- $\mu\text{m}$  Whatman Anotop 10 syringe filters before measurement. Samples were measured at 80% laser intensity. Twenty acquisitions were taken for every sample with each acquisition lasting 60 s. To obtain the hydrodynamic radius ( $R_h$ ), the intensity autocorrelation functions were analyzed by Dynamic software.

**Solid-phase Binding Assays**—The binding of recombinant DDR-Fc proteins to immobilized collagen proteins was measured using a solid-phase binding assay. 50  $\mu\text{l}$  of 200  $\mu\text{g}/\text{ml}$  collagen samples in PBS were coated onto Immulon 2HB 96-well assay plates (ThermoFisher Scientific, Waltham, MA) overnight at 4 °C. Denatured samples were first incubated in a 90 °C water bath for 30 min before coating. Plates were washed and blocked by 1 mg/ml BSA in PBS with 0.05% Tween 20 (PBS-T) for 1 h at room temperature. 50  $\mu\text{l}$  of recombinant DDR1-Fc or DDR2-Fc was subsequently added to each collagen-coated well at a concentration of 20  $\mu\text{g}/\text{ml}$  (~220 nM) in PBS and incubated for 3 h at room temperature. For dose-response assays, serial concentrations of DDR-Fc at 60, 20, 6, 2, 0.6, 0.3, 0.15, 0.06, 0.03, 0.015, and 0  $\mu\text{g}/\text{ml}$  were used, which correspond to ~667, 222, 67, 22, 6.7, 3.3, 1.7, 0.7, 0.3, and 0.17 nM DDR-Fc in molar concentration. For competitive binding assays, the plates were coated with 50  $\mu\text{l}$  of 10  $\mu\text{g}/\text{ml}$  bovine collagen type III overnight at 4 °C, washed, and blocked as described above. A final concentration of 3  $\mu\text{g}/\text{ml}$  (43 nM) DDR-Fc was mixed and incubated with each collagen sample at a final molar concentration of 0, 0.3, 1.3, 5, and 10  $\mu\text{M}$  for 1 h at room temperature. 50  $\mu\text{l}$  of the collagen/DDR-Fc mixtures was subsequently added onto the type III collagen-coated plates and incubated at room temperature for 3 h. Bound DDR-Fc on the plates was detected by incubating first with primary antibody as follows: mouse anti-human IgG1 Fc mAb (Clone 97924, R&D Systems, Minneapolis, MN) at 1  $\mu\text{g}/\text{ml}$ ; and then secondary antibody as follows: anti-mouse HRP antibody (Santa Cruz Biotechnology, Dallas, TX) at 1:2,000 dilution, each for 1 h. Finally, 50  $\mu\text{l}$  of 3,3',5,5'-tetramethylbenzidine solution (Invitrogen) was added for the colorimetric reaction. Plates were washed with 200  $\mu\text{l}$  of PBS-T 6 times between every step. Color was allowed to develop at room temperature for 5 min, and 50  $\mu\text{l}$  of 1 M HCl was added to stop the reaction.  $A_{450 \text{ nm}}$  was recorded from the 96-well plate using a Spectra Max M2 plate reader (Molecular Devices, Sunnyvale, CA) for data analysis. Appropriate controls were included using the same setup. The VWF binding assay was carried out using the same mass concentrations as DDR-Fc binding assays. For VWF detection, rabbit anti-VWF N-terminal antibodies (1:1000) and anti-rabbit HRP antibodies (1:5000) (Abcam) were used. Binding of DDR1-Fc and DDR2-Fc to chemically synthesized triple-helical peptides

## DDR Binding Collagen-like Protein with Inhibitory Function

was performed using a similar protocol, as described previously, with detection of bound proteins using anti-human Fc coupled to horseradish peroxidase and *o*-phenylenediamine dihydrochloride as a substrate (29, 30).

**Collagen-induced DDR Activation Assays**—Collagen-induced DDR autophosphorylation in HEK293 cells was performed as described (28–30). Transiently transfected cells were stimulated with collagen I or peptides for the indicated times and temperature, and cell lysates were analyzed by Western blotting with mouse anti-phosphotyrosine monoclonal antibody (clone 4G10 from Upstate Biotechnology), goat anti-DDR2 (AF2538 from R&D Systems), or rabbit anti-DDR1 (SC-532 from Santa Cruz Biotechnology) and appropriate HRP-conjugated secondary antibodies. The DDR activation assay on Mks was performed as follows. Mks were differentiated from cord blood-derived CD34<sup>+</sup> cells for 13 days in StemSpan medium (Stem Cell, Vancouver, Canada) supplemented with 10 ng/ml thrombopoietin, 10 ng/ml IL-11 (Pepro-Tech, London, UK), 1% L-glutamine, and 1% penicillin/streptomycin as described previously (35). Mks ( $7 \times 10^5$ ) at day 13 of differentiation were plated on 25  $\mu$ g/ml type III collagen-coated 6-well plates upon 30 min of incubation with 100  $\mu$ g/ml (580 nm) of the indicated bacterial collagen. In parallel, unstimulated Mks were plated on uncoated wells as control. After 8 h of incubation at 37 °C in a 5% CO<sub>2</sub> fully humidified atmosphere, Mks were lysed by scraping the well as described previously (36). Lysates were used for Western blotting analysis of DDR1 activation. Membranes were stained with anti-phospho DDR1 (Tyr-792; Cell Signaling, Danvers, MA) and anti-actin (Sigma, Milan, Italy) to show equal loading.

**Megakaryocytes Migration Assay**—Mk migration and invasion assays were performed as described previously (36). Briefly, 8- $\mu$ m trans-well migration inserts (Millipore) were coated with 25  $\mu$ g/ml type III collagen overnight at 4 °C. Mks ( $25 \times 10^3$ ) were seeded on the upper well in 100  $\mu$ l of StemSpan and incubated at 37 °C and 5% CO<sub>2</sub>. After 16 h, Mks that had passed through the trans-well to the other side of the filters and in the outer wells, which contained StemSpan medium with 100 ng/ml SDF1- $\alpha$  (PeproTech, London, UK), were recovered and counted under an inverted microscope. Thereafter, the upper side of the filters was carefully washed with cold PBS, and cells remaining on the upper face of the filters were removed with a cotton wool swab. Trans-well filters were fixed in 4% paraformaldehyde for 20 min at room temperature, stained using a monoclonal antibody against CD61 (Santa Cruz Biotechnology) and with Hoechst 33258, cut out with a scalpel, and mounted onto glass slides, putting the lower face on the top. Each experiment was performed in at least triplicates. Data are expressed as numbers of total migrated cells per insert or as percentages of cells related to that of the control. Images were acquired using an Olympus BX51 using  $\times 20/0.5$  UPlanF1 objective.

**Data Analysis**—Quantitative analyses were performed in triplicate, with results based on the averages of data points and standard deviations presented as error bars. However, the experiments for the dose-response curves consisted of duplicate or single measurements, to measure all samples on the same 96-well plate. In this case, three independent repeat

experiments were performed to confirm the consistency of results. The significance level was determined by *p* value using paired-sample Student's *t* test between the means of two samples. *p* < 0.05 is considered significant. Raw data were processed and plotted in either Origin 6.0 (MicroCal Inc.), Excel 2013 (Microsoft Corp.) or Prism (GraphPad Software, Inc).

## Results

**Design and Expression of Recombinant Bacterial Collagen with a DDR-binding Motif**—A major DDR-binding site contained in fibrillar collagens I–III is the six amino acid motif GVMGFO, which was identified using triple-helical synthetic collagen-mimetic peptides (29, 30). Although these six amino acids are sufficient for binding recombinant DDR ectodomains, as assessed by solid-phase binding assays, the activation of full-length receptors on the surface of cells requires the presence of additional amino acids N-terminal to the GVMGFO motif; peptides presenting GPRGQOGVMGFO (human type III collagen sequence) induce DDR autophosphorylation with similar kinetics to full-length collagen (29). For this study, we modified streptococcal bacterial collagen Scl2.28, by recombinant DNA technology to insert the DDR-binding sequences from human type III collagen. Sequences from the homotrimeric human  $\alpha 1$ (III) chain were used because the bacterial system produces homotrimers. The GVMGFO motif is completely conserved among the fibrillar types II and III collagens and the  $\alpha 1$  chains of type I, with the surrounding sequences highly conserved (type III sequence, GPRGQOGVMGFO; type II sequence, GARGQOGVMGFO, and sequence of the  $\alpha 1$  chain of type I collagen GARGQAGVMGFO). Note that formation of Hyp (O) from Pro in the Yaa position of the Gly-Xaa-Yaa triplets is a post-translational modification in animal collagens that does not take place in the bacterial expression system; therefore, GVMGFP rather than GVMGFO is present in the recombinant bacterial protein.

Two constructs were made containing either four or six triplets of human type III collagen sequence inserted between two tandem CL (Gly-Xaa-Yaa)<sub>79</sub> domains of bacterial collagen, and these are denoted as SCI-DDR<sub>4</sub> (GPRGQPGVMGFP) or SCI-DDR<sub>6</sub> (GPSGPRGQPGVMGFP<sub>6</sub>GPK) (Fig. 1A). A specific negative control with a single Phe to Ala replacement was also constructed. Previous peptide studies indicated that replacement of the critical Phe in the GVMGFP motif by Ala eliminated DDR binding and receptor activation (29, 30); therefore, a bacterial protein containing six triplets, including this replacement GPSGPRGQPGVMGFP<sub>6</sub>GPK was also expressed and denoted as SCI-DDR<sub>6</sub><sub>F-A</sub>. All recombinant proteins were expressed in *E. coli* at 22 °C, and the N-terminal His-tagged proteins were purified through a nickel-nitrilotriacetic acid column (IMAC), yielding ~50 mg/liter of collagen proteins. The purity and identity of the proteins were confirmed by SDS-PAGE and MALDI-TOF (Fig. 1, B and C; Table 1).

**Structural Characterization of Recombinant Bacterial Collagens**—Experiments were carried out to confirm that the insertion of the human collagen sequences did not disrupt the triple-helical conformation or stability of the chimeric collagens. The CD spectra of all recombinant proteins showed a characteristic maximum near 220 nm and a minimum near 198

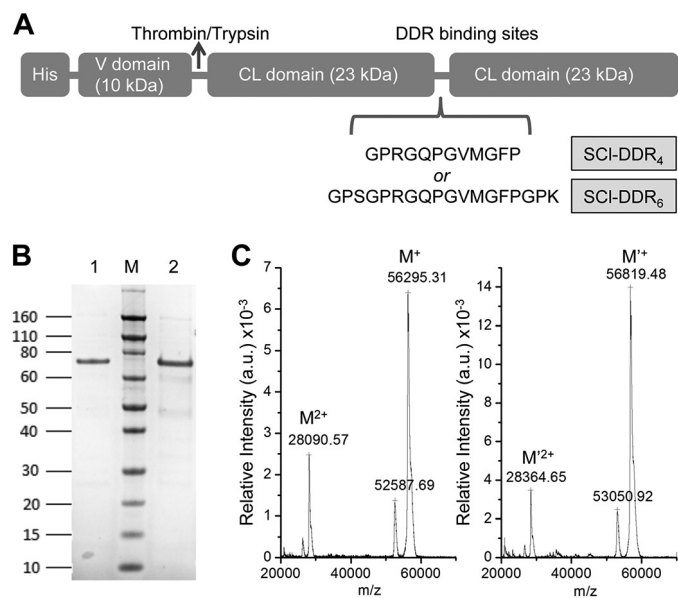
nm, confirming that they adopted a triple-helical conformation (Fig. 2A). Monitoring the  $MRE_{220\text{ nm}}$  versus temperature indicated a sharp thermal transition close to 37 °C for all SCI variants (Fig. 2B and Table 1), with a very small decrease in  $T_m^{CD}$  of  $\sim 0.5$  °C for the proteins with insertions versus the control. Differential scanning calorimetric data of the samples showed similar pattern of thermal stabilities (Fig. 2C and Table 1), with higher  $T_m^{CAL}$  observed compared with CD likely due to the 10 times faster heating rate (37). The calorimetry enthalpy of SCI-DDRs also decreased slightly versus control but is comparable with other SCI variants with 4–6 triplet human collagen sequence insertions (19).

Trypsin digestion was performed to investigate whether a tight triple helix was maintained at the inserted human sequences. After a 2-h digestion at 25 °C, the globular V domain of SCI was cleaved off, yielding bands with decreased mobility on SDS-PAGE, corresponding to the size of the CL-CL unit (Fig. 2D). This indicates that insertions of up to six triplets of

DDR-binding site sequences into the CL-CL protein did not alter the triple-helical conformation, as it remained resistant to trypsin digestion. These results confirm that the SCI-DDR<sub>4</sub>, SCI-DDR<sub>6</sub>, and SCI-DDR<sub>6 F-A</sub> proteins form a typical triple-helical structure.

Dynamic light scattering was used to analyze the particle sizes of the bacterial collagens and to detect potential soluble aggregation of the SCI-DDR solution. The hydrodynamic radii ( $R_h$ ) of all SCI proteins were  $\sim 44 \pm 1.2$  nm (Table 1), which is consistent with the previously reported values. These results indicate that proteins with the DDR-binding site insertions did not form aggregates under these conditions.

**DDR1 and DDR2 Bind to SCI-DDR Proteins**—Recombinant DDR ectodomain constructs, DDR1-Fc and DDR2-Fc, were



**FIGURE 1. Production of recombinant collagen-like protein with DDR-binding sequences (SCI-DDR).** A, schematic diagram of SCI construct with four or six triplet DDR-binding sequence insertions. B, SDS-PAGE showing the purity of the recombinant collagens after His tag column purification. Lane 1, SCI-DDR<sub>4</sub>; lane 2, SCI-DDR<sub>6</sub>; lane M, Novex® Sharp protein standard (Invitrogen); molecular mass is indicated in kDa. Note that collagen proteins always migrate slower in SDS-PAGE than expected on the basis of protein standards. C, MALDI-TOF mass spectrometry of SCI-DDR<sub>4</sub> (left) and SCI-DDR<sub>6</sub> (right), showing the molecular mass of the proteins is in good agreement with that expected (Table 1). a.u., arbitrary units.

**TABLE 1**  
Characterization of the recombinant bacterial collagen constructs

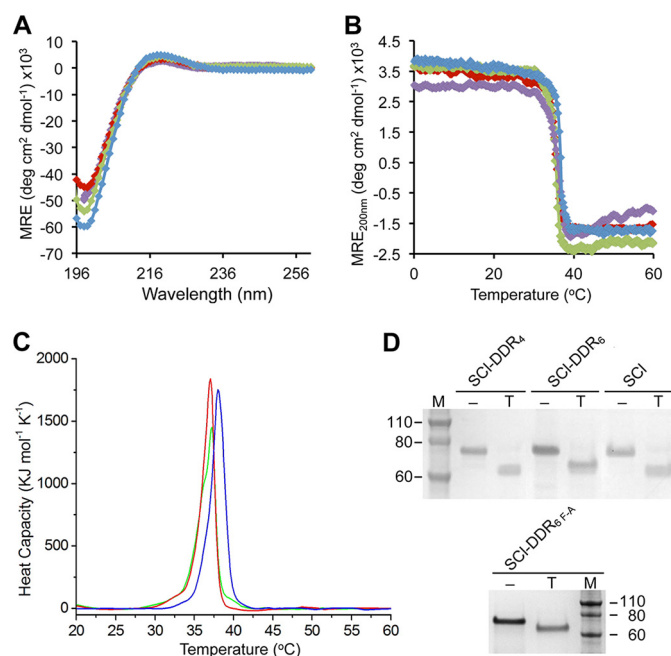
	Molecular mass <sup>a</sup>		CD		DSC			Binding affinity (apparent $K_d$ ) <sup>d</sup>		
	Calculated	Observed	$MRE_{200\text{ nm}}$ <sup>b</sup>	$T_m^{CD}$	Enthalpy <sup>c</sup>	$T_m^{CAL}$	DLS, $R_h$	DDR1	DDR2	VWF
	kDa	kDa		°C		°C	nm	nm	nm	μg/ml
SCI	55734.39	ND	4677	37.1	4217	38.0	45.1	—	—	—
SCI-DDR <sub>4</sub>	56429.18	56295	3456	36.4	3993	37.0	44.9	53.3 ± 7.1	81.2 ± 17.2	1.2 ± 0.12
SCI-DDR <sub>6</sub>	56952.77	56819	4307	36.6	3818	37.2	42.4	66.0 ± 8.8	69.7 ± 7.4	1.2 ± 0.09
SCI-DDR <sub>6 F-A</sub>	56876.67	ND	2913	36.5	ND	ND	ND	—	—	—
Bovine Collage III	ND	ND	ND	ND	ND	ND	ND	15.2 ± 0.9	20.0 ± 1.9	1.0 ± 0.09

<sup>a</sup> Calculated mass is based on the amino acid sequence in the open reading frame of each construct. Observed mass is acquired from MALDI-TOF results.

<sup>b</sup> Unit for the mean residue ellipticity of the CD is degrees·cm<sup>2</sup>·dmol<sup>-1</sup>.

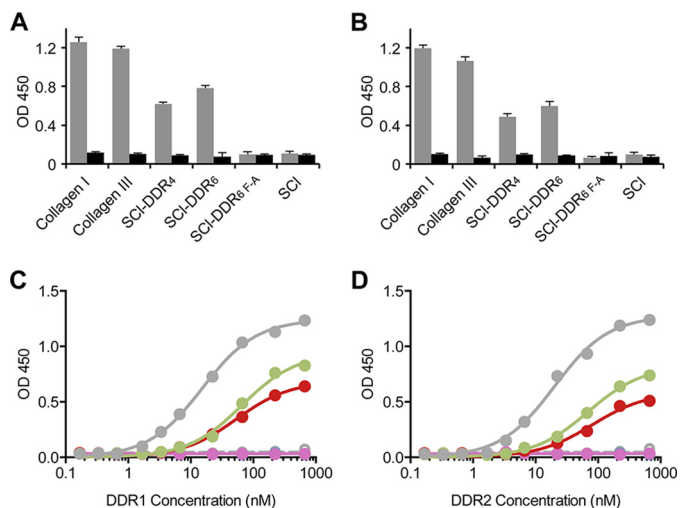
<sup>c</sup> Unit for the calorimetric enthalpy of the DSC is kJ·mol<sup>-1</sup>·K<sup>-1</sup>.

<sup>d</sup> The apparent  $K_d$  values are calculated based on binding curves obtained from the solid-state binding assays; they are good for internal comparison but do not represent the accurate binding constant between the ligand and receptor. Cells with dash symbols indicate no binding. Cells with ND indicate value not determined for that sample.



**FIGURE 2. Structural and thermal analysis of SCI-DDR recombinant collagens.** A, circular dichroism wavelength spectra of the recombinant collagens showing triple helix formation. Color code is as follows: SCI (blue); SCI-DDR<sub>4</sub> (red); SCI-DDR<sub>6</sub> (green); SCI-DDR<sub>6 F-A</sub> (purple). B, circular dichroism temperature scans showing the unfolding profiles of all recombinant collagens. Color code is the same as in A. C, differential scanning calorimetry of SCI (blue), SCI-DDR<sub>4</sub> (red), and SCI-DDR<sub>6</sub> (green) showing heat capacity as a function of temperature. Areas under the curve indicate the calorimetric enthalpy. D, 2-h trypsin digestion (T) showing resistance of the triple-helical domain against enzyme digestion. Proteins were analyzed by SDS-PAGE. Lane M, Novex® Sharp protein standard (Invitrogen); molecular mass is indicated in kDa.  $MRE$ , mean residue ellipticity.

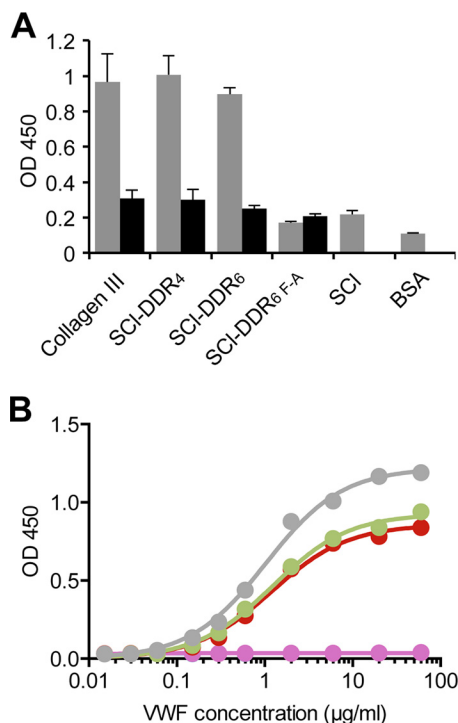
## DDR Binding Collagen-like Protein with Inhibitory Function



**FIGURE 3. Solid-phase binding assay of recombinant DDR1-Fc or DDR2-Fc to recombinant SCI-DDR collagens.** A and B, SCI-DDR constructs in both triple-helical (gray) or denatured (black) states were immobilized on 96-well plates and incubated with 20  $\mu\text{g/ml}$  ( $\sim 220$  nM), of DDR1-Fc (A) or DDR2-Fc (B). Bound proteins were detected with anti-Fc antibodies and measured as absorbance at 450 nm. C and D, dose response of DDR1-Fc (C) and DDR2-Fc (D) to collagens adsorbed onto plastic plates. Color code is as follows: collagen type III (gray); SCI-DDR<sub>4</sub> (red); SCI-DDR<sub>6</sub> (green); SCI (blue); SCI-DDR<sub>6 F-A</sub> (purple); and denatured collagen type III (open gray symbols). Data obtained with denatured SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub> overlap with those of SCI and SCI-DDR<sub>6 F-A</sub>. Data were fitted in GraphPad Prism<sup>®</sup> using a non-linear fit (one site, specific binding).

used to measure binding to immobilized collagens and SCI-DDR proteins (Fig. 3, A and B). As expected, the original SCI construct without an insertion did not bind DDR1 or DDR2, whereas the insertion of the four or six tripeptide sequence containing the GVMGFP motif led to DDR binding for both SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub>. The binding signal for these constructs was about 30–50% lower than those seen for the control native type I and type III collagens. DDR binding was seen only when the recombinant proteins were in their native triple-helical state; no binding was seen following denaturation, in agreement with the receptors' binding specificities that are strictly dependent on a native triple-helical conformation (24, 25, 28). The SCI-DDR<sub>6 F-A</sub> construct showed no binding, consistent with the peptide studies showing Phe is required at the GFP site (29, 30). Initial binding assays were carried out at a DDR concentration of 20  $\mu\text{g/ml}$  ( $\sim 220$  nM), and this was expanded to obtain dose-response curves using receptor concentrations from 0 to 60  $\mu\text{g/ml}$  (0–667 nM) for DDR1-Fc (Fig. 3C) and DDR2-Fc (Fig. 3D). SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub> had similar apparent  $K_d$  values for both DDR1-Fc and DDR2-Fc, which were about 3–4 times higher than those for collagen III binding to DDR1-Fc or DDR2-Fc (Table 1), and there was a modest reduction in  $B_{\text{max}}$  (maximal binding) relative to collagen III.

**SCI-DDR Proteins Are Ligands for von Willebrand Factor**—The DDR-binding motif overlaps with the high affinity-binding site for VWF on collagen type III (RGQOGVMGFO) (38). As this sequence (*i.e.* RGQPGVMGFP, without Hyp) is contained in both SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub>, we tested the proteins for their ability to bind to VWF. Fig. 4 shows dose-dependent binding of VWF to both SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub>. In this case, collagen III, SCI-DDR<sub>4</sub>, and SCI-DDR<sub>6</sub> did not differ in their affinity to VWF (Table 1). In agreement with the collagen bind-

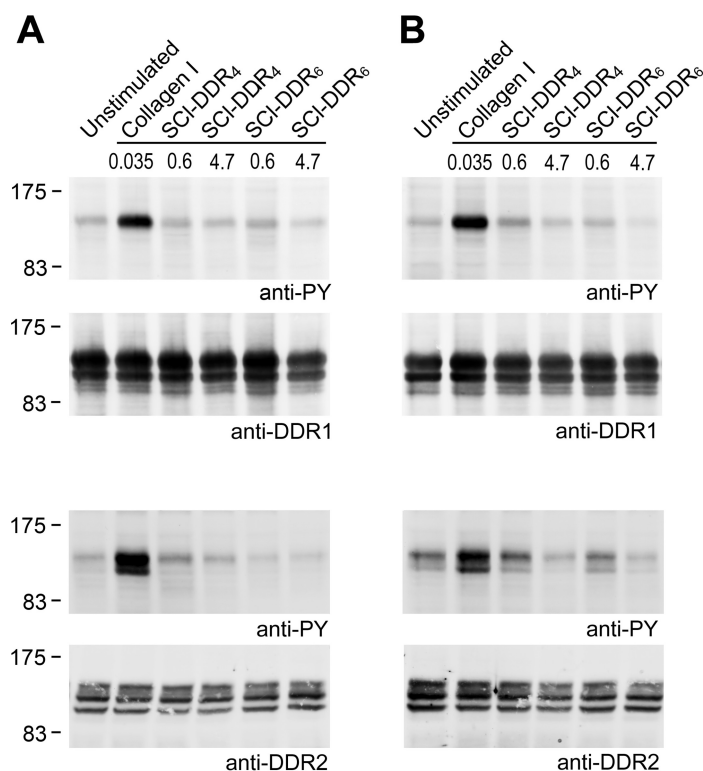


**FIGURE 4. Solid-phase binding assay of VWF to recombinant SCI-DDR collagens.** A, SCI-DDR constructs in both triple-helical (gray) or denatured (black) states were immobilized on 96-well plates and incubated with 20  $\mu\text{g/ml}$  VWF. B, dose response of VWF to collagens adsorbed onto 96-well plates. Bound protein was detected with anti-VWF antibodies and measured as absorbance at 450 nm. Color code is as follows: collagen type III (gray); SCI-DDR<sub>4</sub> (red); SCI-DDR<sub>6</sub> (green); SCI (blue); SCI-DDR<sub>6 F-A</sub> (purple). Data obtained with denatured collagens are obscured by those obtained with SCI-DDR<sub>6 F-A</sub>. Data were fitted in GraphPad Prism<sup>®</sup> using a non-linear fit (one site, specific binding).

ing specificity of VWF (38), VWF did not bind to denatured collagens or the SCI-DDR<sub>6 F-A</sub> construct (Fig. 4). This shows that the same sequence GPRGQPVMGFP in a bacterial collagen context enables binding to VWF as well as to the DDRs, consistent with the peptide findings. These data demonstrate that our engineered SCI proteins have the expected binding characteristics for the inserted binding module.

**SCI-DDR Proteins Do Not Induce DDR Activation**—Because SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub> bound to DDR1 and DDR2 in the solid-phase binding assays, experiments were carried out to determine whether these chimeric recombinant constructs could promote receptor activation in HEK293 cells that transiently express full-length DDR1 or DDR2. Collagen binding to the extracellular discoidin homology domain induces activation of the intracellular DDR tyrosine kinase domain, which is manifested by autophosphorylation of cytoplasmic Tyr residues (24, 25). The presence of Tyr phosphorylation was assayed by Western blotting of cell lysates, as described previously (28–30). Collagen type I (at 10  $\mu\text{g/ml}$  or 35 nM)<sup>6</sup> induced a clear phosphorylation signal for both DDR1 and DDR2, as expected (Fig. 5A). In sharp contrast, no DDR activation above background was observed when cells were incubated with either

<sup>6</sup> The concentrations of animal collagens, the SCI-DDR collagen constructs, and the triple-helical peptides are expressed in terms of molarity of trimer (based on single chain molecular mass of animal collagen at  $\sim 95$  kDa, SCI-DDR at  $\sim 57$  kDa, and peptide molecular masses depending on their sequence).



**FIGURE 5. Recombinant collagens do not induce DDR1 or DDR2 autophosphorylation.** Full-length DDR1 (DDR1a isoform) or DDR2 was transiently expressed in HEK293 cells. After stimulation for 90 min with the indicated collagen samples at 37 °C (A) or at 23 °C (B), cell lysates were analyzed by SDS-PAGE and Western blotting. Collagen I was used at 10 µg/ml (0.035 µM), which is the concentration previously established to trigger autophosphorylation; SCI-DDRs were assayed at 100 µg/ml (0.6 µM) initially, which is the concentration previously shown to trigger autophosphorylation for collagen peptides, and later at a much higher concentration 800 µg/ml (4.7 µM). For DDR1 samples, the blot was probed with anti-phosphotyrosine mAb 4G10 (upper panel), followed by stripping and reprobing with anti-DDR1 (lower blot). For DDR2 samples, samples were resolved on two gels. The corresponding blots were probed with anti-phosphotyrosine mAb 4G10 (upper panel) or anti-DDR2 (lower panel). The positions of molecular mass markers are indicated (in kDa).

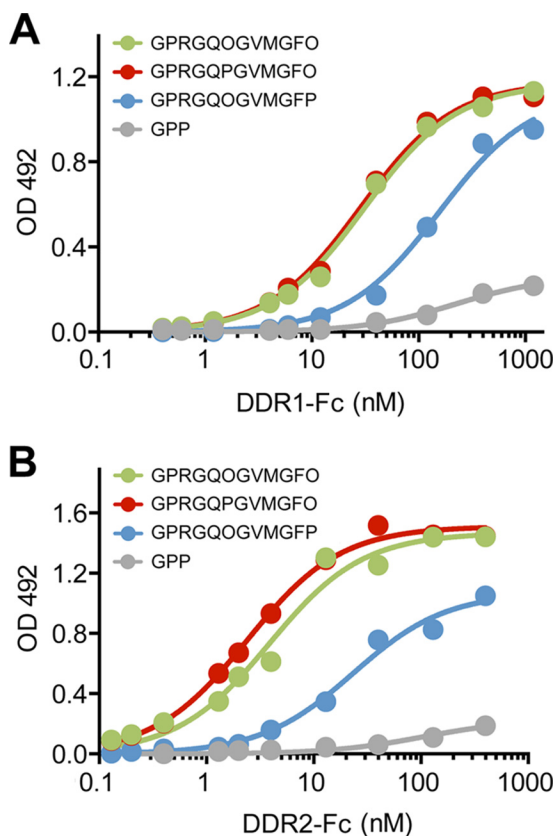
SCI-DDR<sub>4</sub> or SCI-DDR<sub>6</sub> using standard incubation conditions at 37 °C. When the initial concentration of 100 µg/ml (0.6 µM) recombinant collagen failed to activate the DDRs, the concentration was increased to the highest testable concentration of 800 µg/ml (~4.7 µM), but no phosphorylation was observed. Because the triple-helical conformation is required for DDR binding, and activation at 37 °C is close to the melting temperature of the bacterial collagen constructs, it is possible that lack of DDR activation by the bacterial collagen constructs was due to denaturation of the collagens during the 90-min incubation at 37 °C. The activation assays were therefore also carried out at lower temperatures, to avoid possible unfolding of triple helices. Type I collagen induced strong DDR autophosphorylation when cells were incubated at 23 °C, but the bacterial collagen constructs did not induce phosphorylation above background of either DDR1 or DDR2 (Fig. 5B). These data suggest that although the SCI-DDR proteins can bind to the DDRs, they do not act as receptor agonists.

**Effect of Proline Hydroxylation on DDR Binding and Activation**—Because the recombinant collagen expressed in *E. coli* is not hydroxylated on proline, it is possible that the absence of Hyp affects DDR binding and receptor activation. Introducing post-translational hydroxylation of Pro in this recombinant bacterial system is complicated, because under aerobic conditions *E. coli* cells do not biosynthesize or transport L-ascorbate into the cytosol (39), which would be required to activate human prolyl-4-hydroxylase introduced through plasmids.

Experiments based on a previously published strategy to bypass this requirement (39) have thus far only yielded bacterial collagen with low Hyp levels in our hands. Therefore, experiments to investigate the effect of Hyp on DDR binding and activation were carried out on synthetic triple-helical peptides. Two Hyp are present in the GPRGQOGVMGFO DDR-binding sequence of collagen type III (underlined), and it was previously shown that mutation of the Hyp to Ala in the GQO triplet did not affect the binding affinity to DDR1 or DDR2, although a Hyp to Ala change in the GFO triplet caused a substantial decrease in DDR binding (29, 30). To further investigate whether it is the Pro residue itself or the hydroxylation of the Pro that is critical in the GVMGFO sequence, homologous peptides with GPRGQOGVMGFP and GPRGQOGVMGFO were synthesized and tested. A control peptide with the sequence GPRGQPGVMGFO was also synthesized, which was expected to bind to the DDRs with similar affinity to the native sequence. Solid-phase binding assays showed that the peptide with GPRGQPGVMGFO did indeed bind to DDR1-Fc and DDR2-Fc to similar extents as a peptide with GPRGQOGVMGFO (Fig. 6). In contrast, GPRGQOGVMGFP resulted in weaker binding to DDR1-Fc and DDR2-Fc, compared with the native sequence.

The peptides were tested for their ability to induce DDR phosphorylation (Fig. 7). Although the GPRGQPGVMGFO peptide could effectively induce DDR1 or DDR2 phosphorylation when tested at 50 µg/ml (3.9 µM) or 100 µg/ml (7.9 µM), the peptide with GPRGQOGVMGFP required higher concentrations

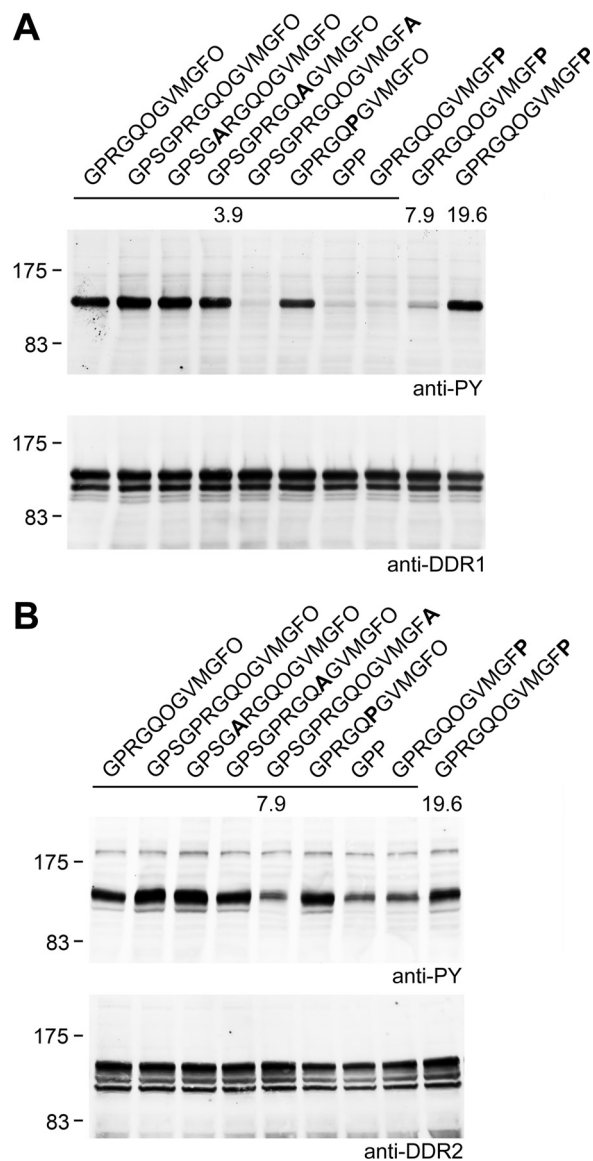
## DDR Binding Collagen-like Protein with Inhibitory Function



**FIGURE 6. Solid-phase binding assays of DDR1-Fc or DDR2-Fc to triple-helical collagen-mimetic peptides.** DDR1-Fc (A) or DDR2-Fc (B) was added for 3 h to 96-well plates coated with peptides at 10  $\mu\text{g}/\text{ml}$ . Bound proteins were detected with anti-Fc antibodies and measured as absorbance at 492 nm. Data were fitted in GraphPad Prism<sup>®</sup> using a non-linear fit (one site, specific binding).

( $\sim 250 \mu\text{g}/\text{ml}$  or  $19.6 \mu\text{M}$ ) to stimulate similar levels of DDR1 or DDR2 autophosphorylation. These results are in keeping with the reduced binding affinity of GPRGQOGVMGFP compared with GPRGQOGVMGFO and suggest that the presence of Hyp in GVMGFO, although important for higher affinity DDR binding, is not essential to induce DDR phosphorylation. Additional peptides contained Pro to Ala substitutions or Hyp to Ala substitutions in the sequence GPSGPRGQOGVMGFO. Activation of the DDRs by these peptides was as expected; peptides with GPSGARGQOGVMGFO or GPSGPRGQAGVMGFO induced DDR phosphorylation to the same extent as GPSGPRGQOGVMGFO, in line with their ability to bind DDR1 and DDR2 to similar extents as peptides with the native sequence (29, 30). In contrast, a peptide with GPSGPRGQOGVMGFA drastically reduced DDR binding (29, 30) and receptor activation (Fig. 7). Compared with collagen, a much higher molar concentration of these synthetic peptides is required for DDR activation ( $\sim 4 \mu\text{M}$  peptide *versus*  $0.03 \mu\text{M}$  collagen), but the reasons for this are not known (29, 30).

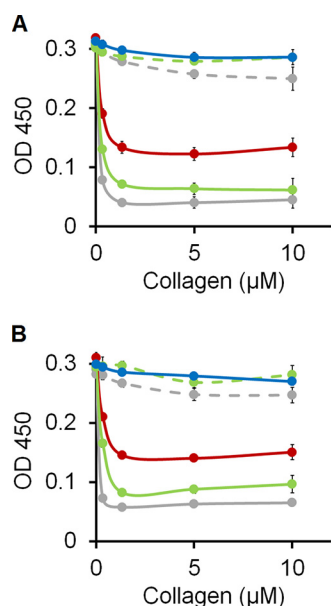
**Recombinant SCI-DDR as an Inhibitor of DDR Binding and Activation**—Given that the recombinant chimeric collagens could bind to DDR1 and DDR2, but not activate them, their role as a potential inhibitor was explored. First, a competitive solid-phase binding assay was performed to test whether the recombinant collagens could inhibit DDR1 or DDR2 binding to immobilized type III collagen (Fig. 8). Denatured collagens as



**FIGURE 7. DDR activation by collagen-mimetic peptides.** Hyp in GVMGFO is important but not essential for DDR1 or DDR2 activation. Full-length DDR1 or DDR2 was transiently expressed in HEK293 cells. After stimulation for 90 min at  $37^\circ\text{C}$  with the indicated peptide samples, cell lysates were analyzed by SDS-PAGE and Western blotting. Peptides were used at  $50 \mu\text{g}/\text{ml}$  ( $3.9 \mu\text{M}$ ) or  $100 \mu\text{g}/\text{ml}$  ( $7.9 \mu\text{M}$ ); the peptide without Hyp was also assayed at  $250 \mu\text{g}/\text{ml}$  ( $19.6 \mu\text{M}$ ). A, DDR1 samples, blot was probed with anti-phosphotyrosine mAb 4G10 (upper panel), followed by stripping and reprobing with anti-DDR1 (lower blot). B, DDR2 samples, samples were resolved on two gels. The corresponding blots were probed with anti-phosphotyrosine mAb 4G10 (upper panel) or anti-DDR2 (lower panel). The positions of molecular mass markers are indicated (in kDa).

well as the original SCI, with no insertion, showed only negligible inhibition. Efficient inhibition was observed when excessive amounts of native SCI-DDR recombinant collagens were pre-incubated with DDR1-Fc or DDR2-Fc before addition onto type III collagen-coated wells. Maximal inhibition was achieved at  $\sim 1.3 \mu\text{M}$  ( $228 \mu\text{g}/\text{ml}$ ) of recombinant collagens, giving a molar ratio between dimeric DDR receptors to trimeric recombinant collagen molecules of  $\sim 1:60$ . Under the same condition, type I collagen acted as a somewhat more potent inhibitor, showing inhibitory effects at lower concentrations. The ability of the recombinant SCI-DDR collagens to inhibit animal colla-





**FIGURE 8. Competitive inhibition by SCI-DDR of the interactions between collagen type III and recombinant DDR2-Fc.** Collagen type I, SCI-DDR<sub>4</sub>, or SCI-DDR<sub>6</sub> were incubated with DDR1-Fc (A) or DDR2-Fc (B) for 1 h at room temperature before addition to plate-immobilized collagen type III. Bound DDR-Fc was detected with anti-Fc antibodies and measured as absorbance at 450 nm. Color code is as follows: collagen type I (gray); SCI-DDR<sub>4</sub> (red); SCI-DDR<sub>6</sub> (green); SCI (blue). Dashed lines show data for the denatured samples.

gen binding to DDR suggests potential for inhibition in a biological setting.

DDR1 is known to support the migration of several types of cells, and the ability of recombinant SCI-DDR<sub>6</sub> protein to inhibit such migration was investigated using a previously established Mk migration system where DDR1 could be activated by fibrillar collagen, which subsequently promoted cell migration and invasion (36). SCI-DDR was added, at increasing concentrations, into the cell culture media to evaluate whether it could block the binding between Mk surface DDR1 receptors and type III collagen, and thus inhibit cell migration or cell invasion. When SCI-DDR<sub>4</sub> or SCI-DDR<sub>6</sub> was added at 10, 50, and 100  $\mu\text{g}/\text{ml}$  (equals 58, 290, and 580 nM), fewer cells migrated through the Transwell<sup>®</sup> membrane, compared with control conditions without adding recombinant collagen (Fig. 9, A–C). When the same concentration of SCI without insertion or SCI-DDR<sub>6 F-A</sub> was added, there was no obvious change in the number of cells migrating through the membrane. Of note, 1  $\mu\text{g}/\text{ml}$  (6 nM) of all the bacterial collagens tested did not result in any reduction of Mk migration. The level of phosphorylated DDR1 in the Mks plated on collagen type III was also analyzed by Western blot. In the presence of SCI-DDR<sub>4</sub> and of SCI-DDR<sub>6</sub>, there was a reduction in the level of phosphorylated DDR1 to a level comparable with unstimulated Mks as shown by the level of phosphorylated DDR1 normalized to actin (Fig. 9, D and E). SCI or SCI-DDR<sub>6 F-A</sub> did not affect the level of DDR1 phosphorylation. These results suggest that SCI-DDR<sub>6</sub> and SCI-DDR<sub>4</sub> compete with coated type III collagen for binding DDR1 on the surface of Mks, thereby preventing the activation of DDR1 and delaying cell migration in this cell culture system.

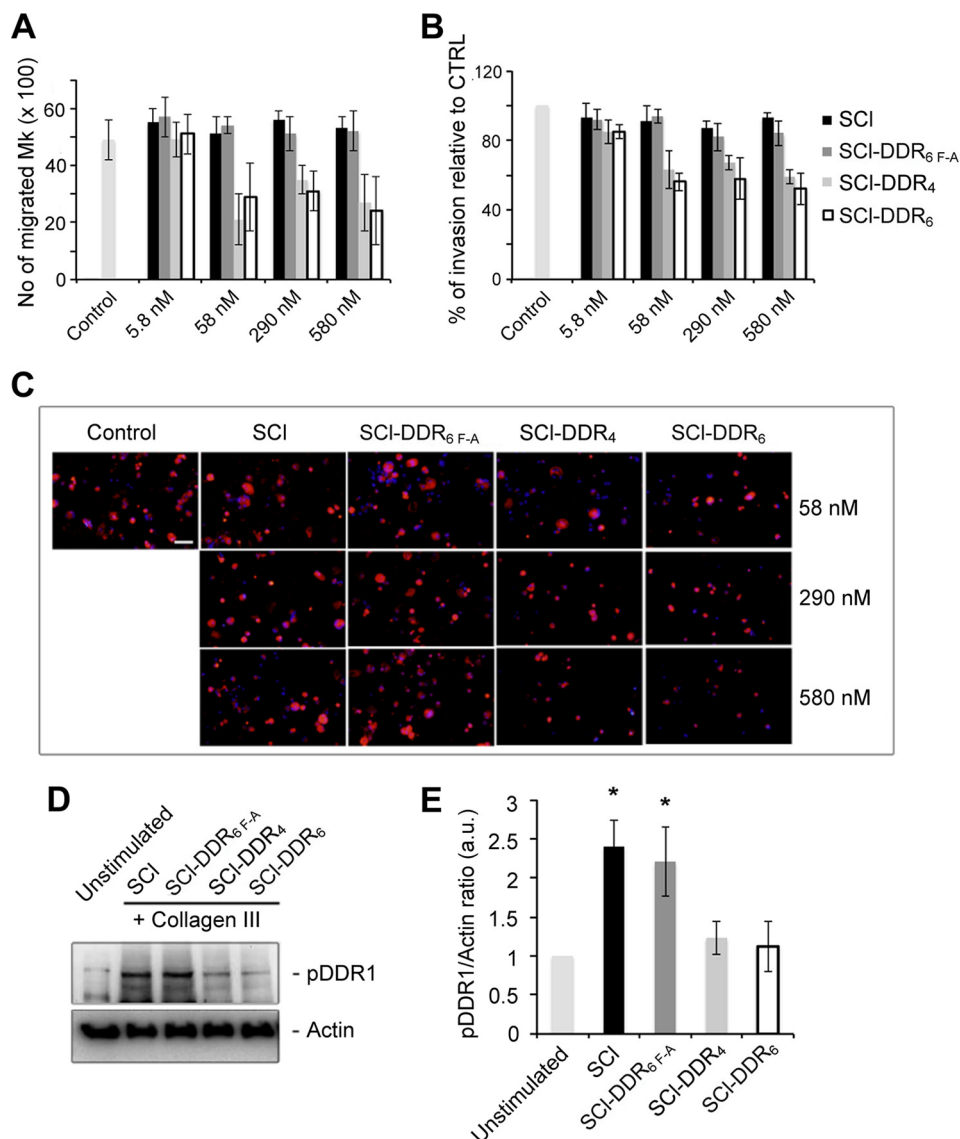
## Discussion

We report on a recombinant protein with sequences from the DDR binding region of human collagen type III inserted within a triple-helical bacterial collagen context. Collagen, as the major component of the extracellular matrix, contains more than 50 biologically active sites that interact with other extracellular matrix proteins or cell surface receptors and regulates normal cellular and tissue functions (40). The stringent requirement of forming triple-helical structures complicates the analysis of collagen ligand-binding sites. Early studies used collagen fragments from cyanogen bromide digestion and rotary shadow electron microscopy to identify the rough location of some bioactive sites within collagen. The development of libraries of overlapping triple-helical peptides, the so-called collagen Toolkit peptide libraries, allowed a comprehensive analysis of binding sites within the fibrillar collagens type II and type III (12) and formed the basis for successful definition of the minimal binding sites of an increasing number of key collagen activities, including the DDR-binding sequence used in this study (29, 30).

Although the development of synthetic triple-helical peptides led to major advances in defining collagen binding activities, alternative recombinant strategies are valuable for confirming the peptide results and allowing their application to biomaterials. The use of engineered bacterial collagen-like proteins as a complementary approach allows large scale production of the recombinant protein by standard protein production methods and the creation of bioactive hydrogels (5, 6, 41). To date, several human collagen interaction regions have been inserted within a recombinant bacterial collagen-like protein, including the binding sites for different integrins, the binding site for fibronectin and heparin, and the cleavage site for MMP1 and MMP-13 (11, 18–21). Varying the length of inserted sequences also allowed definition of the minimally required collagen sequence of a given activity in this system. Using the bacterial collagen system raises the possibility of efficient production of “plug and play” recombinant collagen-like proteins containing selected and multiple collagen functions. Here, we extend the concept of designing modular collagen with separate collagen functionalities by creating a group of recombinant collagen-like molecules to target and regulate DDR signaling. Recombinant collagens containing DDR-binding sequences did bind to DDR receptors, as expected, but they differed from mammalian collagens in their weaker binding affinity and their inability to stimulate DDR autophosphorylation.

Non-modified SCI did not bind to the DDRs; binding was dependent on the presence of a DDR-binding sequence and an intact triple helix, as expected from the strict requirement for native triple-helical conformation for the DDR-collagen interaction. Both SCI-DDR<sub>4</sub> and SCI-DDR<sub>6</sub> bound DDR1 and DDR2, but there are small differences in the amount of maximal binding and in the level of inhibition. It is not clear whether these differences are significant, but we cannot exclude the possibility that low inhibition of SCI-DDR<sub>4</sub> points to a secondary binding site in DDR. The dose-response curves of DDR-Fc proteins to SCI constructs showed reduced binding affinity of recombinant SCI-DDR compared with collagens type I and III

## DDR Binding Collagen-like Protein with Inhibitory Function



**FIGURE 9. SCI-DDRs reduce DDR1 activation on type III collagen and result in decreased megakaryocyte migration and invasion.** *A*, Mks were treated with the indicated concentrations of the SCI-DDR recombinant collagens and allowed to migrate in a trans-well migration chamber coated with type III collagen for 16 h. Mks that passed through the filter in the lower chamber were counted. *B*, Mks adhering on the lower side of the Transwell filter were fixed and stained with anti-CD61 antibody and then counted by fluorescence microscopy. *C*, representative images showing Mks that invaded the lower side of the Transwell filter. Cells were stained with anti-CD61 antibody (red). Nuclei were counterstained with Hoechst 33288 (blue). Scale bar, 50  $\mu$ m. *D*, DDR1 activation assay on Mks stimulated with 100  $\mu$ g/ml (580 nM) of the indicated SCI-DDR recombinant collagen and plated on type III collagen-coated wells. Unstimulated Mks plated on uncoated wells were used as control. Membranes were probed with the indicated antibodies. *E*, densitometry analysis of pDDR1 Western blots. pDDR1/actin ratio was calculated and shown as relative to unstimulated Mk. Data are presented as means  $\pm$  S.D. of at least three independent experiments. Asterisks indicate  $p < 0.05$ .

(Fig. 3). The most obvious difference between animal collagen and recombinant collagen is the presence of hydroxyproline in the former. Our data with synthetic peptides show that Hyp in the six amino acid-binding motif GVMGFO is required for high affinity DDR binding but that substitution of Hyp with Pro in the preceding triplet (GPRGQPGVMGFO context) does not alter binding affinity (Fig. 6). A previous crystal structure of the DDR2 discoidin domain complexed to a collagen-mimetic peptide showed the Hyp in GVMGFO in the leading chain of the triple helix participates in a hydrogen bonding network involving a salt bridge between Arg-105 and Glu-113 of DDR2, whereas the preceding Hyp makes no contact with DDR2 (31). It is therefore not surprising that mutation to GVMGFP reduces DDR binding, although the GPRGQPGVMGFP substi-

tion has no effect on the DDR-SCI interaction. The presence of Hyp in the GVMGFO motif is thus critical for high affinity DDR-collagen interactions.

In contrast, for VWF, it was previously shown that substitution of Hyp with Ala (GVMGFA context) in a synthetic triple-helical peptide did not affect VWF binding (38), indicating that Hyp in the GVMGFO motif is not involved in binding to VWF. Consistent with this notion, our SCI-DDR constructs bound to VWF with similar affinity to that seen for type III collagen (Fig. 4), which likely reflects the lack of involvement of Hyp in VWF binding.

All previous studies found a correlation between the ability of different collagen types to bind the DDRs and their ability to induce receptor activation, as assessed by autophosphorylation

of cytoplasmic tyrosine residues (24, 25, 28, 42, 43). This correlation is extended to collagen-mimetic peptides (Fig. 7) (present study and Refs. 29, 30). It was therefore surprising to see no receptor activation by the recombinant SCI-DDR proteins, despite their ability to bind to the DDRs in a specific manner. The inability to induce transmembrane signaling cannot be explained by denaturation of the SCI-DDR triple helix, because no phosphorylation was observed at 23 °C as well as at 37 °C, although control animal collagens led to a strong phosphorylation signal at both temperatures. Although a lack of hydroxylation in GPRGQOGVMGFO reduced binding affinity, peptide studies showed that hydroxylation was not essential for inducing receptor activation. Higher concentrations of GPRGQOGVMGFP peptide were required for DDR activation, compared with the GPRGQOGVMGFO peptide, but the relative peptide concentrations required for activation were in line with their relative affinities.

Because of the reduced binding affinity of the recombinant collagens compared with animal collagens, higher concentrations of SCI-DDR proteins were tested for DDR activation, but no significant autophosphorylation was observed, even at 800  $\mu\text{g/ml}$  (4.7  $\mu\text{M}$ ), the highest concentration that was practical in the cell-based assay following the dilution of the protein stock into cell culture medium. Synthetic peptides require much higher concentrations than native collagens to induce DDR phosphorylation (3–4  $\mu\text{M}$  versus 0.02–0.03  $\mu\text{M}$ ), especially in the absence of Hyp in the GVMGFO essential sequence, but the underlying reasons are not clear (29, 30). We cannot exclude that the failure of our recombinant collagens to induce DDR phosphorylation is due to an inability to test them at high enough concentrations, but the unexpected uncoupling observed between DDR binding and DDR activation may be related to the unusual nature of DDR activation compared with other receptor tyrosine kinases. Typical receptor tyrosine kinases are activated with fast kinetics (seconds to minutes after ligand binding), whereas the DDRs have protracted activation kinetics and require prolonged incubation with collagen for kinase activation (24, 25). Despite a detailed understanding of the DDR1 ectodomain structure and the DDR2-collagen interaction (31, 44), very little is known about the molecular mechanism of how collagen binding to the DDR discoidin domains results in activation of the cytoplasmic kinase domain. In particular, the cellular mechanisms behind the slow kinetics are unclear. DDR1 binding to animal collagen has been shown to cluster the receptor on the cell surface with fast kinetics (45). It is likely that SCI-DDR proteins bind to the DDRs on the cell surface, but perhaps the bacterial collagen cannot cluster the receptors in the required fashion or cannot mediate a “second step” in activation such as the transition from clustered receptors to phosphorylated receptors. Future experiments are required to address whether SCI-DDR proteins lead to DDR1 clustering on the cell surface. It is currently not established whether DDR phosphorylation results from interactions with collagen that has aggregated during the prolonged incubation times necessary to observe receptor phosphorylation. Possible reasons why the SCI-DDR proteins cannot activate DDRs could include the absence of aggregation for the recombinant SCI-DDR protein, which may be required, or the large number of

charges in the SCI triple helix, which could disrupt or prevent interactions of the DDRs with molecules on the cell surface, which may be required for a transition to the phosphorylated state.

The observation that SCI-DDR proteins bind to the DDR ectodomain without stimulating receptor transmembrane signaling led to the prediction that they could inhibit collagen-DDR interactions, which was demonstrated *in vitro* and in a cell migration system. The ability of SCI-DDR molecules to interfere with DDR-collagen interactions provides a basis for the generation of recombinant collagen-based biomaterials with DDR inhibitory properties. Recombinant collagen presents an attractive alternative to animal collagens for collagen biomaterial production, allowing standardized, high yield production with easy sequence modification. The recombinant collagen approach has been hindered by the requirement for post-translational modification of Pro to Hyp, an issue that is largely avoided in bacterial collagen SCI which forms a stable triple helix in the absence of Hyp. Biological activities such as integrin, fibronectin, and heparin binding have been conferred on bacterial collagen through the insertion of modular human collagen sequences. This study thus presents the first example where a modular insertion of the defined DDR-binding sequence did not fully capture biological functionality. However, it showed the potential of inserting a human ligand binding module to inhibit a natural biological activity or process. The generation of recombinant collagen-based biomaterials with DDR inhibitory properties could be useful in the study of the pathological roles of DDR signaling in a number of disease models where DDR signaling is believed to contribute to disease progression (22, 26, 27).

*Author Contributions*—B. A., B. L., A. B., and B. B. designed the study and wrote the paper. B. A., V. A., D. L. K., and B. L. analyzed and critically evaluated the results. H. X. and D. G. performed experiments with triple-helical peptides and DDR activation assays in HEK293 cells. V. A. performed the experiments with megakaryocytes. D. B. synthesized the collagen-mimetic peptides. A. Y. performed early cloning experiments. R. W. F. oversaw the peptide design and experiments. All authors read and approved the final version of the manuscript.

*Acknowledgments*—We thank Dr. John Ramshaw from the Commonwealth Scientific and Industrial Research Organization of Australia for bovine type I and III collagen preparations and helpful discussion; we also thank Dr. David Wilbur from Tufts University Chemistry Department for allowing us to access the MALDI-TOF MS equipment.

## References

1. Rasmussen, M., and Björck, L. (2001) Unique regulation of SclB—a novel collagen-like surface protein of *Streptococcus pyogenes*. *Mol. Microbiol.* **40**, 1427–1438
2. Yu, Z., An, B., Ramshaw, J. A., and Brodsky, B. (2014) Bacterial collagen-like proteins that form triple-helical structures. *J. Struct. Biol.* **186**, 451–461
3. Mohs, A., Silva, T., Yoshida, T., Amin, R., Lukomski, S., Inouye, M., and Brodsky, B. (2007) Mechanism of stabilization of a bacterial collagen triple helix in the absence of hydroxyproline. *J. Biol. Chem.* **282**, 29757–29765

4. Xu, C., Yu, Z., Inouye, M., Brodsky, B., and Mirochnitchenko, O. (2010) Expanding the family of collagen proteins: recombinant bacterial collagens of varying composition form triple-helices of similar stability. *Bio-macromolecules* **11**, 348–356
5. Cosgriff-Hernandez, E., Hahn, M. S., Russell, B., Wilems, T., Munoz-Pinto, D., Browning, M. B., Rivera, J., and Höök, M. (2010) Bioactive hydrogels based on designer collagens. *Acta Biomaterialia* **6**, 3969–3977
6. Peng, Y. Y., Howell, L., Stoichevska, V., Werkmeister, J. A., Dumsday, G. J., and Ramshaw, J. A. (2012) Towards scalable production of a collagen-like protein from *Streptococcus pyogenes* for biomedical applications. *Microb. Cell Fact.* **11**, 146
7. An, B., Kaplan, D. L., and Brodsky, B. (2014) Engineered recombinant bacterial collagen as an alternative collagen-based biomaterial for tissue engineering. *Front. Chem.* **2**, 40
8. Lukomski, S., Nakashima, K., Abdi, I., Cipriano, V. J., Shelvin, B. J., Graviss, E. A., and Musser, J. M. (2001) Identification and characterization of a second extracellular collagen-like protein made by group A *Streptococcus*: control of production at the level of translation. *Infect. Immun.* **69**, 1729–1738
9. Xu, Y., Keene, D. R., Bujnicki, J. M., Höök, M., and Lukomski, S. (2002) Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices. *J. Biol. Chem.* **277**, 27312–27318
10. Humtsoe, J. O., Kim, J. K., Xu, Y., Keene, D. R., Höök, M., Lukomski, S., and Wary, K. K. (2005) A streptococcal collagen-like protein interacts with the  $\alpha 2\beta 1$  integrin and induces intracellular signaling. *J. Biol. Chem.* **280**, 13848–13857
11. Seo, N., Russell, B. H., Rivera, J. J., Liang, X., Xu, X., Afshar-Kharghan, V., and Höök, M. (2010) An engineered  $\alpha 1$  integrin-binding collagenous sequence. *J. Biol. Chem.* **285**, 31046–31054
12. Farndale, R. W., Lisman, T., Bihan, D., Hamaia, S., Smerling, C. S., Pugh, N., Konitsiotis, A., Leitinger, B., de Groot, P. G., Jarvis, G. E., and Raynal, N. (2008) Cell-collagen interactions: the use of peptide Toolkits to investigate collagen-receptor interactions. *Biochem. Soc. Trans.* **36**, 241–250
13. Manka, S. W., Carafoli, F., Visse, R., Bihan, D., Raynal, N., Farndale, R. W., Murphy, G., Enghild, J. J., Hohenester, E., and Nagase, H. (2012) Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 12461–12466
14. Erat, M. C., Slatter, D. A., Lowe, E. D., Millard, C. J., Farndale, R. W., Campbell, I. D., and Vakonakis, I. (2009) Identification and structural analysis of type I collagen sites in complex with fibronectin fragments. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4195–4200
15. Sekiya, A., Okano-Kosugi, H., Yamazaki, C. M., and Koide, T. (2011) Pigment epithelium-derived factor (PEDF) shares binding sites in collagen with heparin/heparan sulfate proteoglycans. *J. Biol. Chem.* **286**, 26364–26374
16. Knight, C. G., Morton, L. F., Peachey, A. R., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (2000) The collagen-binding A-domains of integrins  $\alpha 1\beta 1$  and  $\alpha 2\beta 1$  recognize the same specific amino acid sequence, GFOGER, in native (triple-helical) collagens. *J. Biol. Chem.* **275**, 35–40
17. Jarvis, G. E., Raynal, N., Langford, J. P., Onley, D. J., Andrews, A., Smethurst, P. A., and Farndale, R. W. (2008) Identification of a major GpVI-binding locus in human type III collagen. *Blood* **111**, 4986–4996
18. An, B., DesRochers, T. M., Qin, G., Xia, X., Thiagarajan, G., Brodsky, B., and Kaplan, D. L. (2013) The influence of specific binding of collagen-silk chimeras to silk biomaterials on hMSC behavior. *Biomaterials* **34**, 402–412
19. An, B., Abbonante, V., Yigit, S., Balduini, A., Kaplan, D. L., and Brodsky, B. (2014) Definition of the native and denatured type II collagen binding site for fibronectin using a recombinant collagen system. *J. Biol. Chem.* **289**, 4941–4951
20. Peng, Y. Y., Stoichevska, V., Schacht, K., Werkmeister, J. A., and Ramshaw, J. A. (2014) Engineering multiple biological functional motifs into a blank collagen-like protein template from *Streptococcus pyogenes*. *J. Biomed. Mater. Res. A* **102**, 2189–2196
21. Yu, Z., Visse, R., Inouye, M., Nagase, H., and Brodsky, B. (2012) Defining requirements for collagenase cleavage in collagen type III using a bacterial collagen system. *J. Biol. Chem.* **287**, 22988–22997
22. Leitinger, B. (2014) Discoidin domain receptor functions in physiological and pathological conditions. *Int. Rev. Cell Mol. Biol.* **310**, 39–87
23. Fu, H. L., Valiathan, R. R., Arkwright, R., Sohail, A., Mihai, C., Kumarasiri, M., Mahasenan, K. V., Mobashery, S., Huang, P., Agarwal, G., and Fridman, R. (2013) Discoidin domain receptors: unique receptor tyrosine kinases in collagen-mediated signaling. *J. Biol. Chem.* **288**, 7430–7437
24. Vogel, W., Gish, G. D., Alves, F., and Pawson, T. (1997) The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol. Cell* **1**, 13–23
25. Shrivastava, A., Radziejewski, C., Campbell, E., Kovac, L., McGlynn, M., Ryan, T. E., Davis, S., Goldfarb, M. P., Glass, D. J., Lemke, G., and Yancopoulos, G. D. (1997) An orphan receptor tyrosine kinase family whose members serve as nonintegrin collagen receptors. *Mol. Cell* **1**, 25–34
26. Valiathan, R. R., Marco, M., Leitinger, B., Kleer, C. G., and Fridman, R. (2012) Discoidin domain receptor tyrosine kinases: new players in cancer progression. *Cancer Metastasis Rev.* **31**, 295–321
27. Borza, C. M., and Pozzi, A. (2014) Discoidin domain receptors in disease. *Matrix Biol.* **34**, 185–192
28. Leitinger, B. (2003) Molecular analysis of collagen binding by the human discoidin domain receptors, DDR1 and DDR2. Identification of collagen binding sites in DDR2. *J. Biol. Chem.* **278**, 16761–16769
29. Konitsiotis, A. D., Raynal, N., Bihan, D., Hohenester, E., Farndale, R. W., and Leitinger, B. (2008) Characterization of high affinity binding motifs for the discoidin domain receptor DDR2 in collagen. *J. Biol. Chem.* **283**, 6861–6868
30. Xu, H., Raynal, N., Stathopoulos, S., Myllyharju, J., Farndale, R. W., and Leitinger, B. (2011) Collagen binding specificity of the discoidin domain receptors: binding sites on collagens II and III and molecular determinants for collagen IV recognition by DDR1. *Matrix Biol.* **30**, 16–26
31. Carafoli, F., Bihan, D., Stathopoulos, S., Konitsiotis, A. D., Kvensakul, M., Farndale, R. W., Leitinger, B., and Hohenester, E. (2009) Crystallographic insight into collagen recognition by discoidin domain receptor 2. *Structure* **17**, 1573–1581
32. Silva, T., Kirkpatrick, A., Brodsky, B., and Ramshaw, J. A. (2005) Effect of deamidation on stability for the collagen to gelatin transition. *J. Agric. Food Chem.* **53**, 7802–7806
33. Yoshizumi, A., Yu, Z., Silva, T., Thiagarajan, G., Ramshaw, J. A., Inouye, M., and Brodsky, B. (2009) Self-association of *Streptococcus pyogenes* collagen-like constructs into higher order structures. *Protein Sci.* **18**, 1241–1251
34. Raynal, N., Hamaia, S. W., Siljander, P. R., Maddox, B., Peachey, A. R., Fernandez, R., Foley, L. J., Slatter, D. A., Jarvis, G. E., and Farndale, R. W. (2006) Use of synthetic peptides to locate novel integrin  $\alpha 2\beta 1$ -binding motifs in human collagen III. *J. Biol. Chem.* **281**, 3821–3831
35. Balduini, A., Pallotta, I., Malara, A., Lova, P., Pecci, A., Viarengo, G., Balduini, C. L., and Torti, M. (2008) Adhesive receptors, extracellular proteins and myosin IIA orchestrate proplatelet formation by human megakaryocytes. *J. Thromb. Haemost.* **6**, 1900–1907
36. Abbonante, V., Gruppi, C., Rubel, D., Gross, O., Moratti, R., and Balduini, A. (2013) Discoidin domain receptor 1 protein is a novel modulator of megakaryocyte-collagen interactions. *J. Biol. Chem.* **288**, 16738–16746
37. Persikov, A. V., Xu, Y., and Brodsky, B. (2004) Equilibrium thermal transitions of collagen model peptides. *Protein Sci.* **13**, 893–902
38. Lisman, T., Raynal, N., Groeneveld, D., Maddox, B., Peachey, A. R., Hui-zinga, E. G., de Groot, P. G., and Farndale, R. W. (2006) A single high-affinity binding site for von Willebrand factor in collagen III, identified using synthetic triple-helical peptides. *Blood* **108**, 3753–3756
39. Pinkas, D. M., Ding, S., Raines, R. T., and Barron, A. E. (2011) Tunable, post-translational hydroxylation of collagen domains in *Escherichia coli*. *ACS Chem. Biol.* **6**, 320–324
40. Kadler, K. (1994) Extracellular matrix. 1: fibril-forming collagens. *Protein Profile* **1**, 519–638
41. Browning, M. B., Guiza, V., Russell, B., Rivera, J., Cereceres, S., Höök, M., Hahn, M. S., and Cosgriff-Hernandez, E. M. (2014) Endothelial cell response to chemical, biological, and physical cues in bioactive hydrogels. *Tissue Eng. Part A* **20**, 3130–3141

## DDR Binding Collagen-like Protein with Inhibitory Function

42. Leitinger, B., Steplewski, A., and Fertala, A. (2004) The D2 period of collagen II contains a specific binding site for the human discoidin domain receptor, DDR2. *J. Mol. Biol.* **344**, 993–1003
43. Leitinger, B., and Kwan, A. P. (2006) The discoidin domain receptor DDR2 is a receptor for type X collagen. *Matrix Biol.* **25**, 355–364
44. Carafoli, F., Mayer, M. C., Shiraishi, K., Pecheva, M. A., Chan, L. Y., Nan, R., Leitinger, B., and Hohenester, E. (2012) Structure of the discoidin domain receptor 1 extracellular region bound to an inhibitory Fab fragment reveals features important for signaling. *Structure* **20**, 688–697
45. Mihai, C., Chotani, M., Elton, T. S., and Agarwal, G. (2009) Mapping of DDR1 distribution and oligomerization on the cell surface by FRET microscopy. *J. Mol. Biol.* **385**, 432–445