Proline hydroxylation in collagen supports integrin binding by two distinct mechanisms

Received for publication, January 31, 2018, and in revised form, March 23, 2018 Published, Papers in Press, April 3, 2018, DOI 10.1074/jbc.RA118.002200

Kalle H. Sipilä[‡1], Kati Drushinin[§], Pekka Rappu[‡], Johanna Jokinen[‡], Tiina A. Salminen[¶], Antti M. Salo[§], Jarmo Käpylä[‡], Johanna Myllyharju[§], and Jyrki Heino[‡2]

From the [‡]Department of Biochemistry, University of Turku, FI-20014 Turku, Finland, the [§]Biocenter Oulu and Faculty of Biochemistry and Molecular Medicine, University of Oulu, FI-90014 Oulu, Finland, and the [¶]Structural Bioinformatics Laboratory, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, FI-20520 Turku, Finland

Edited by Amanda J. Fosang

Collagens are the most abundant extracellular matrix proteins in vertebrates and have a characteristic triple-helix structure. Hydroxylation of proline residues is critical for helix stability, and diminished prolyl hydroxylase activity causes wide-spread defects in connective tissues. Still, the role of proline hydroxylation in the binding of collagen receptors such as integrins is unclear. Here, we isolated skin collagen from genetically modified mice having reduced prolyl 4-hydroxylase activity. At room temperature, the reduced proline hydroxylation did not affect interactions with the recombinant integrin α2I domain, but at 37 °C, collagen hydroxylation correlated with the avidity of α2I domain binding. Of note, LC–MS/MS analysis of isolated skin collagens revealed no major changes in the hydroxyproline content of the main integrin-binding sites. Thus, the disrupted α2I domain binding at physiological temperatures was most likely due to structural destabilization of the collagenous helix. Integrin α2I binding to the triple-helical GFPGER motif was slightly weaker than to GFOGER (O = hydroxyproline). This phenomenon was more prominent when α1 integrin was tested. Integrin α1β1 expressed on CHO cells and recombinant α1I domain showed remarkably slower binding velocity and weaker avidity to GFPGER when compared with GFOGER. Structural modeling revealed the critical interaction between Arg-218 in α1I and the hydroxyproline residue in the integrin-binding motif. The role of Arg-218 was further validated by testing a variant R218D α1I domain in solid-phase binding assays. Thus, our results show that the lack of proline hydroxylation in collagen can affect integrin binding by a direct mechanism and via structural destabilization of the triple helix.

This work was supported by Academy of Finland Project Grants 259769 (to J. H.) and 296498 (to J. M.), Academy of Finland Center of Excellence 2012–2017 Grant 251314 (to J. M.), grants from the Sigrid Jusélius Foundation (to J. H., J. M., and T. A. S.), the Jane and Aatos Erkko Foundation (to J. M.), Tor, Joe, and Pentti Borg's Foundation (to T. A. S.), and the Turku Doctoral Programme of Biomedical Science. J. M. owns equity in FibroGen Inc., which develops hypoxia-inducible factor P4H inhibitors as potential therapeutics. This company supports hypoxia response research in the J. M. group.
The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD008802.

[1] Present address: King's College London, London WC2R 2LS, United Kingdom.
[2] To whom correspondence should be addressed. Tel.: 358505238351; E-mail: jyrki.heino@utu.fi.

Collagens are the most abundant extracellular matrix proteins in vertebrates. They have a characteristic triple-helical structure that is composed of three collagen α chains. The primary structure of the collagen α chain enables the formation of the collagenous triple-helix. Especially the location of glycine as every third residue and the presence of posttranslationally hydroxylated proline and lysine residues are critical factors (1). Lack of hydroxyproline dramatically decreases the melting temperature (T_m) of collagen and the T_m of the triple-helical collagen molecule is directly proportional to the hydroxyproline content (2, 3). Collagen prolyl 4-hydroxylases (C-P4Hs)[3] catalyze the formation of hydroxyprolines in collagens in the *Y* position prolines of the *-X-Y-*Gly- repeats through a mechanism that requires 2-oxoglutarate, Fe^{2+}, molecular oxygen, and vitamin C (1). Scurvy, a disease caused by vitamin C deficiency, leads to broad connective tissue destruction and problems in the healing of skin wounds (4). In addition to vitamin C deficiency, C-P4H activity can be disrupted by mutations. Vertebrates have three C-P4H isoenzymes, of which C-P4H-I is expressed in all tissues studied, C-P4H-II is expressed particularly in chondrocytes, osteoblasts, and endothelial cells, and C-P4H-III has a wide expression pattern, but the expression is minor when compared with C-P4H-I and -II (1). C-P4H-I and C-P4H-II are likely to have distinct differences in their substrate-binding sites because the K_m values of C-P4H-II for synthetic collagen-like $(Pro\text{-}Pro\text{-}Gly)_{10}$ and Gly-Val-Pro-Gly-Val peptides and a full-length procollagen chain are 3–6-fold higher than in the case of C-P4H-I (5, 6). Furthermore, a marked difference exists between the two enzymes in the inhibition by a competitive inhibitor poly(L-proline), the K_i of C-P4H-II being 200–1000 times higher than that of C-P4H-I, depending on the length of poly(L-proline) (5, 6). Studies with synthetic peptide substrates have shown that the rate of hydroxylation of a proline residue by C-P4H-I is affected by the nature of amino acids in the *X* position and also in other parts of the peptide substrate (3). Systematic comparison between C-P4H-I and C-P4H-II on their hydroxylation rates of such peptides is yet lacking, however. Studies with gene-modified mice have shown that homozygous inactivation of the P4ha1 gene, which encodes the catalytic α(I) subunit of the main isoenzyme, C-P4H-I, leads to an 80% decrease in total C-P4H activity and

[3] The abbreviations used are: C-P4H, collagen prolyl 4-hydroxylase; PDB, Protein Data Bank; CHO, Chinese hamster ovary.

embryonic death at E10.5, whereas P4ha2$^{-/-}$ mice lacking C-P4H-II have no obvious phenotypic abnormalities (7, 8). Compound P4ha1$^{+/-}$;P4ha2$^{-/-}$ mice with 65% reduction of total C-P4H activity in chondrocytes have moderate chondrodysplasia caused by defective growth plate development associated with soft extracellular matrix and reduced thermal stability of cartilage collagen (8). In humans, bi-allelic mutations of the P4HA1 gene have recently been shown to lead to a congenital connective tissue disorder with joint hypermobility, contractures, mild skeletal dysplasia, and high myopia in a compound heterozygous patient with a 50% reduction in total C-P4H activity in fibroblasts (9).

The members of the large integrin family mediate cell adhesion and are critical for survival, proliferation, differentiation, and migration, as well as tissue homeostasis. There are four integrin-type collagen receptors in mammals, namely $\alpha1\beta1$, $\alpha2\beta1$, $\alpha10\beta1$, and $\alpha11\beta1$ heterodimers (10). They bind to ligands in a metal-dependent manner via the inserted domain (I domain) that is structurally similar to von Willebrand factor type A domain (11). Consensus integrin-binding motif in collagens can be formulated as G*XX*GER and the best characterized high avidity site is triple-helical GFOGER (O = 4-hydroxyproline) (12). Other similar motifs include GROGER (13), GLOGER (14), GMOGER (15), GLOGEN (14), GAOGER (13), GVOGEA (16), GFKGER (17), and GLQGER (17). Importantly, most of these sequences contain a hydroxyproline residue. Previously, it has been reported that the $\alpha2$I domain can bind to GFPGER, but more weakly than GFOGER peptides (18). Furthermore, earlier papers show that the integrin $\alpha2\beta1$, but not $\alpha1\beta1$, can bind to unhydroxylated collagen I that has been produced in transgenic plants (19). However, both integrin $\alpha1$I and integrin $\alpha2$I domains can bind to GFPGER and GLPGER motifs when they have been inserted into Streptococcal collagen-like protein Scl2 (20). Thus, it is not clear whether prolyl hydroxylation plays a direct or indirect role in collagen–integrin interaction. Additionally, the structural basis of the putative difference between $\alpha1$ and $\alpha2$ integrin action is unknown.

To examine the role of collagen hydroxylation in the collagen–integrin interaction, we extracted skin collagen from P4ha1$^{+/+}$;P4ha2$^{+/-}$, P4ha1$^{+/-}$;P4ha2$^{+/-}$, P4ha1$^{+/+}$;P4ha2$^{-/-}$, and P4ha1$^{+/-}$;P4ha2$^{-/-}$ mutant mice. The collagen preparations as well as triple-helical collagen mimicking peptides were tested with soluble recombinant integrin αI domains and CHO cells transfected to overexpress specific integrin–collagen receptors. We report that hydroxyprolines generated by C-P4Hs can affect integrin binding by two distinct mechanisms: stabilization of the collagenous helix at body temperature and by strengthening the direct interaction between specific integrin αI domains.

Results

Hydroxylation of proline residues in skin collagens is essential for the intact binding function of integrin α2I domain at physiological temperatures

To study if reduction in C-P4H activity affects the collagen–integrin interaction, we extracted bulk collagen from mouse skin using a acetic acid/pepsin digestion method followed by NaCl precipitation. The binding of soluble human recombinant integrin $\alpha2$I domain to isolated collagen was measured by a solid-phase binding assay. We could not detect any significant differences between binding to skin collagen derived from the WT mice ($n = 5$) and collagen from P4ha1$^{+/-}$;P4ha2$^{-/-}$ mice ($n = 6$), when the experiments were performed at room temperature (Fig. 1*A*). However, when the binding measurements were repeated at 37 °C, the $\alpha2$I domain showed ~30% weaker binding to collagen derived from the mice with deficient C-P4H activity (Experiment 1, $p = 0.004$; Experiment 2, $p = 0.0004$; Fig. 1*B*).

Next, we tested binding of the integrin $\alpha2$I domain to collagenous triple-helical peptides, which harbor the well-described hydroxyproline containing an integrin-binding motif (GFOGER) or the same site with proline (GFPGER) at room temperature and at 37 °C. The $\alpha2$I domain binding was slightly weaker to the proline containing peptide, but the difference was the same in both temperatures (Fig. 1, *C* and *D*). These results indicate that there are temperature-dependent and temperature-independent mechanisms by which the hydroxylation state of proline residues in collagen can affect the function of collagen receptor integrin.

To confirm our findings, we isolated skin collagen from WT mice ($n = 5$) and mutant mice with different genetic combinations leading to C-P4H deficiency: P4ha1$^{+/+}$;P4ha2$^{+/-}$ ($n = 4$), P4ha1$^{+/-}$;P4ha2$^{+/-}$ ($n = 5$), P4ha1$^{+/+}$;P4ha2$^{-/-}$ ($n = 7$), and P4ha1$^{+/-}$;P4ha2$^{-/-}$ ($n = 6$). The LC–MS/MS shotgun proteomics analysis revealed that the skin collagen preparations almost exclusively contained collagen I and collagen III. On average, 72.3% of the detected peptide spectrum matches belonged to the Colα1(I) or Colα2(I) chains and 26.5% to the Colα1(III) chain. The average fraction of spectra matched to the other 29 proteins was only 1.2% showing the high purity level of our collagen preparations. The hydroxylation frequency of prolines in -*X*-P-G- sequences in collagen was progressively and significantly diminished when the number of gene alleles coding for C-P4H-I and -II was reduced (Fig. 2*A*). At room temperature binding of the soluble $\alpha2$I domain did not correlate with the hydroxyproline content in collagen (Fig. 2*B*), whereas there was a strong correlation at 37 °C ($r = 0.647$, $p = 0.0005$) (Fig. 2*C*). We used circular dichroism (CD) and trypsin/chymotrypsin digestion analysis to study the thermal stability of the triple-helical conformation of collagen isolated from WT and P4ha1$^{+/-}$;P4ha2$^{-/-}$ mice. CD analysis of two individual skin collagen samples from both genotypes gave melting temperature (T_m) values of 39.3 and 39.9 °C for the WT collagen samples, whereas the T_m of P4ha1$^{+/-}$;P4ha2$^{-/-}$ collagen samples was markedly lower, 36.3 and 36.4 °C. Similarly, at 37 °C collagen derived from P4ha1$^{+/-}$;P4ha2$^{-/-}$ mice was remarkably less trypsin/chymotrypsin resistant, which indicates opening of the triple-helical conformation, than collagen derived from WT animals (Fig. 2*D*). In conclusion, the temperature-dependent changes in integrin binding seem to be due to the altered hydroxylation state of the proline residues in collagen.

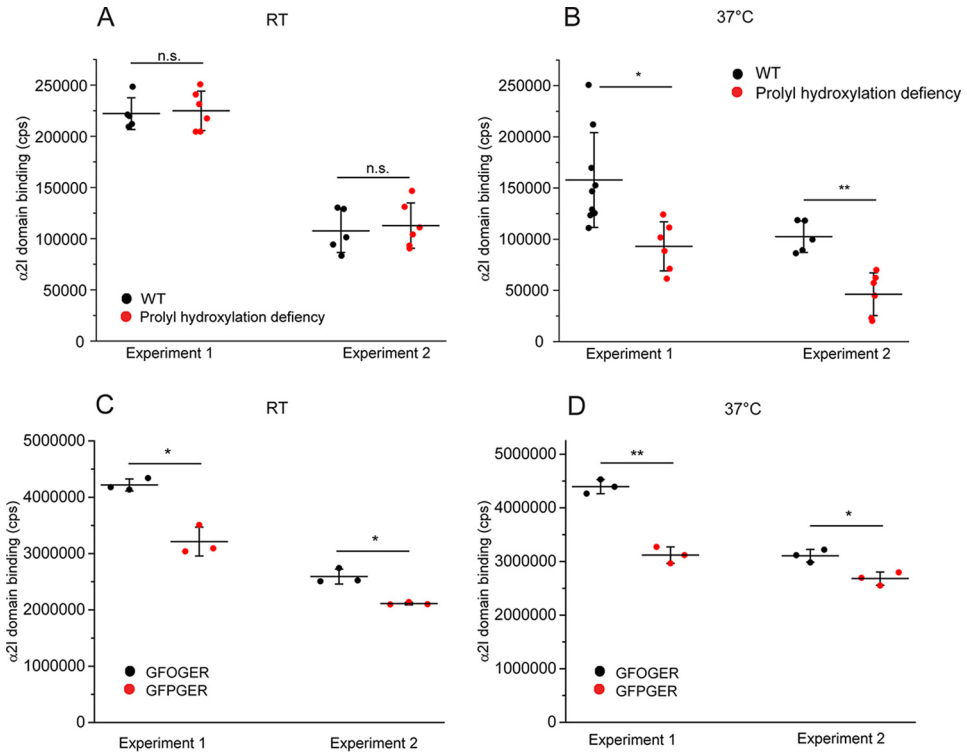


Figure 1. At 37 °C, but not at room temperature, the binding of integrin α2I domain is weaker to collagen isolated from P4ha1$^{+/-}$;P4ha2$^{-/-}$ mice than to collagen from WT mice. *A,* binding of α2I domain to collagen isolated from 6 P4ha1$^{+/-}$;P4ha2$^{-/-}$ and 5 WT mice at room temperature (*RT*). Two independent experiments are shown. *B,* binding of α2I domain to collagen isolated from 6 P4ha1$^{+/-}$;P4ha2$^{-/-}$ and 5 WT mice at 37 °C. Two independent experiments are shown. *C,* binding of α2I domain to triple-helical GFOGER and GFPGER peptides at room temperature. Two independent experiments are shown. *D,* binding of α2I domain to triple-helical GFOGER and GFPGER peptides at 37 °C. Two independent experiments are shown. *p* values from Student's *t* test and mean ± S.D. are shown. *, $p < 0.01$; **, $p < 0.001$. Integrin αI domain concentrations used in the experiments were 200 (*A* and *B*) or 400 nM (*C* and *D*).

Partial genetic inactivation of C-P4Hs in vivo does not affect the main integrin-binding motifs

To study further the hypohydroxylation of proline residues in collagen and its temperature-dependent consequences, we analyzed the LC–MS/MS data for the hydroxylation status of prolines in the known integrin α2β1-binding motifs.

The proline residues in the GLOGER, GMOGER, GROGER, and GFOGER motifs of the Colα1(I) chain and the GROGER motif of the Colα2(I) chain were always (100%) hydroxylated both in the WT and mutant mice (Fig. 3*A*). We detected both proline and hydroxyproline containing forms of the GLOGER motif in the Colα2(I) chain and the GAOGER motif in the Colα1(III) chain (Fig. 3*B*). The hydroxyproline content of GAOGER decreased from 100% in WT mice to 85% in P4ha1$^{+/-}$; P4ha2$^{-/-}$ mice. In the case of GLOGER of the Colα2(I) chain, the hydroxyproline content decreased from 70 to 40%, correspondingly. The relative abundance of the detected peptides containing the first triplet of individual hydroxylated integrin-binding motifs (GRO, GAO, GLO, GFO, and GMO) was equal in collagen samples derived from the different mouse lines (Fig. 3*C*).

The changes detected in the hydroxyproline content of GLOGER of the Colα2(I) chain and GAOGER of the Colα1(III) chain are not sufficient to explain the affected binding of the integrin α2I domain. Giving also the fact that the α2I domain bound only slightly weaker to GFPGER than to GFOGER containing triple-helical peptides and that the difference was not temperature sensitive (Fig. 1, *C* and *D*), the most likely explanation for the decreased α2I binding to hypohydroxylated skin collagen at 37 °C is a partial or local destabilization of the triple helix.

Replacement of hydroxyproline with proline in the GFOGER integrin-binding motif leads to affected binding by integrins α1β1 and α11β1

To investigate the role of the direct interaction between integrin αI domain and the hydroxyproline residue in the integrin-binding motif, we analyzed, in addition to the α2I domain, the binding of the soluble recombinant integrin α1I domain to GFOGER and GFPGER containing peptides in solid-phase binding assays. Furthermore, we also used CHO cells transfected to express α1, α2, or α11 integrins. The adhesion of these cells to the collagenous peptides was measured by impedance-based xCELLigence technology that monitors cell–substrate interaction at real-time. WT CHO cells do not express collagen receptor integrins on their cell surface and the adhesion of the CHO transfectants to collagen is solely mediated by the over-expressed integrin α subunit that has formed a heterodimer with CHO cell's own β1 subunit. The protein-binding assays were performed at room temperature and the cell-adhesion assays at 37 °C. The CD-spectra of GFOGER and GFPGER indicated a small difference in melting temperature (48.9 *versus* 46.9 °C) (Fig. 4*A*). However, at room temperature and 37 °C it was possible to study the direct interaction between

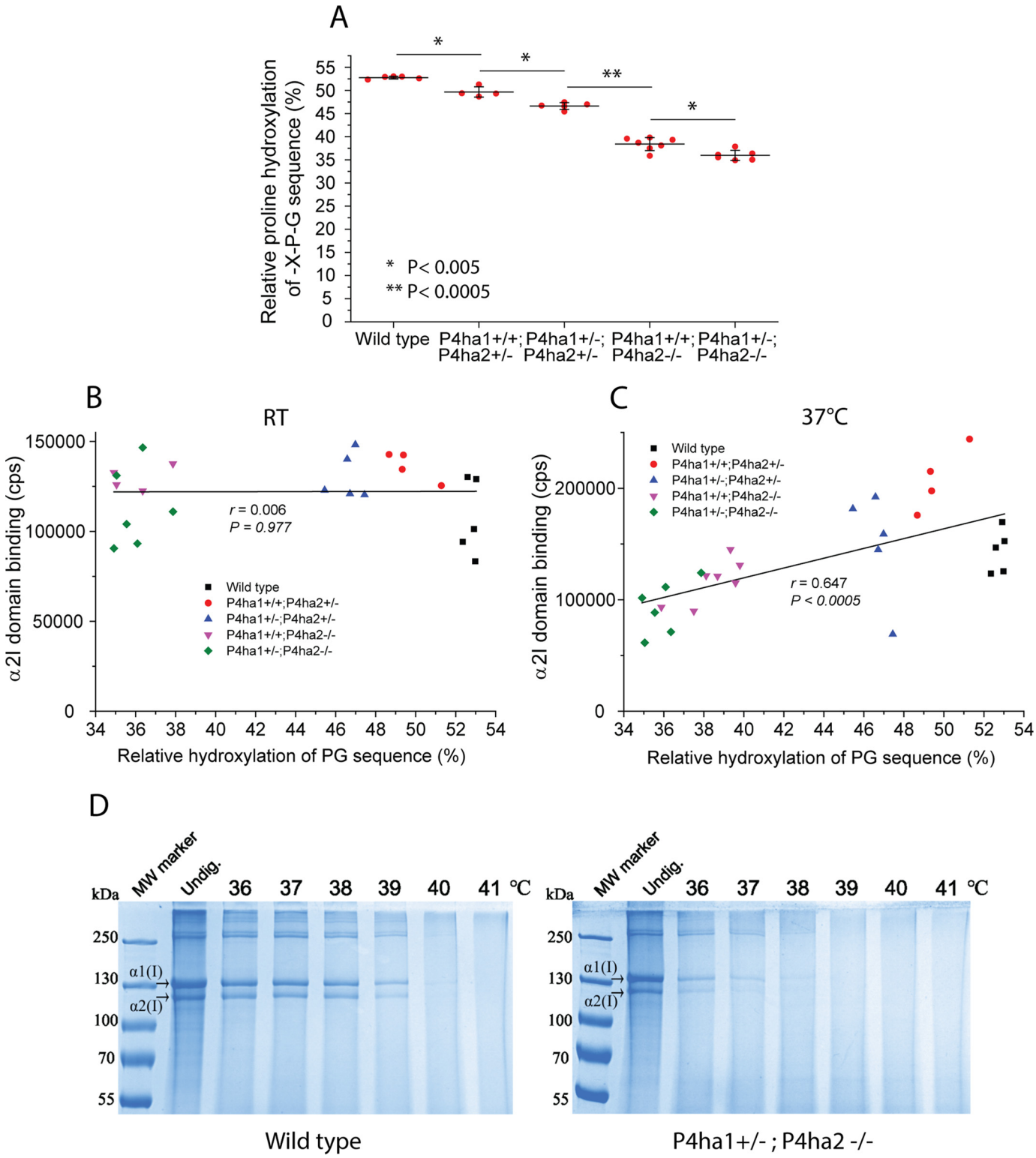


Figure 2. Integrin binding correlates with the proline hydroxylation levels at 37 °C but not at room temperature. *A,* relative hydroxylation of the -*X*-P-G-sequence in skin collagen of C-P4H mutant mice calculated by dividing the number of unique tandem mass spectra matching to a peptide containing -*X*-O-G-sequences (but no -*X*-P-G- sequences) with the number of unique spectra matching to the peptides containing -*X*-O-G or -*X*-P-G- sequences. The data points represent the average of two LC–MS/MS runs of collagen preparations from individual mice. *p* values from Tukey HSD test and mean ± S.D. are shown. *B* and *C,* Pearson correlation analyses between integrin binding (α2I domain) and relative hydroxylation of collagen (isolated separately from 4 P4ha1$^{+/+}$;P4ha2$^{+/-}$, 5 P4ha1$^{+/-}$;P4ha2$^{+/-}$, 7 P4ha1$^{+/+}$;P4ha2$^{-/-}$, 6 P4ha1$^{+/-}$;P4ha2$^{-/-}$, and 5 WT mice) at room temperature (*RT*) (*B*) or 37 °C (*C*). Integrin αI domain concentration used in the experiments was 200 nM. *D,* analysis of thermal stability of fibrillar (mainly type I) collagen isolated from the skin of the WT and P4ha1$^{+/-}$; P4ha2$^{-/-}$ mice by trypsin/chymotrypsin digestion. The collagen samples were treated with a mixture of trypsin and chymotrypsin at temperatures between 36 and 41 °C and analyzed by 8% SDS-PAGE under reducing conditions followed by Coomassie Blue staining, and representative images are shown. An undigested sample without trypsin/chymotrypsin is shown as a control. Molecular weight markers are shown on the *left* of the gels and the *arrows* show the positions of the α1(I) and α2(I) collagen chains.

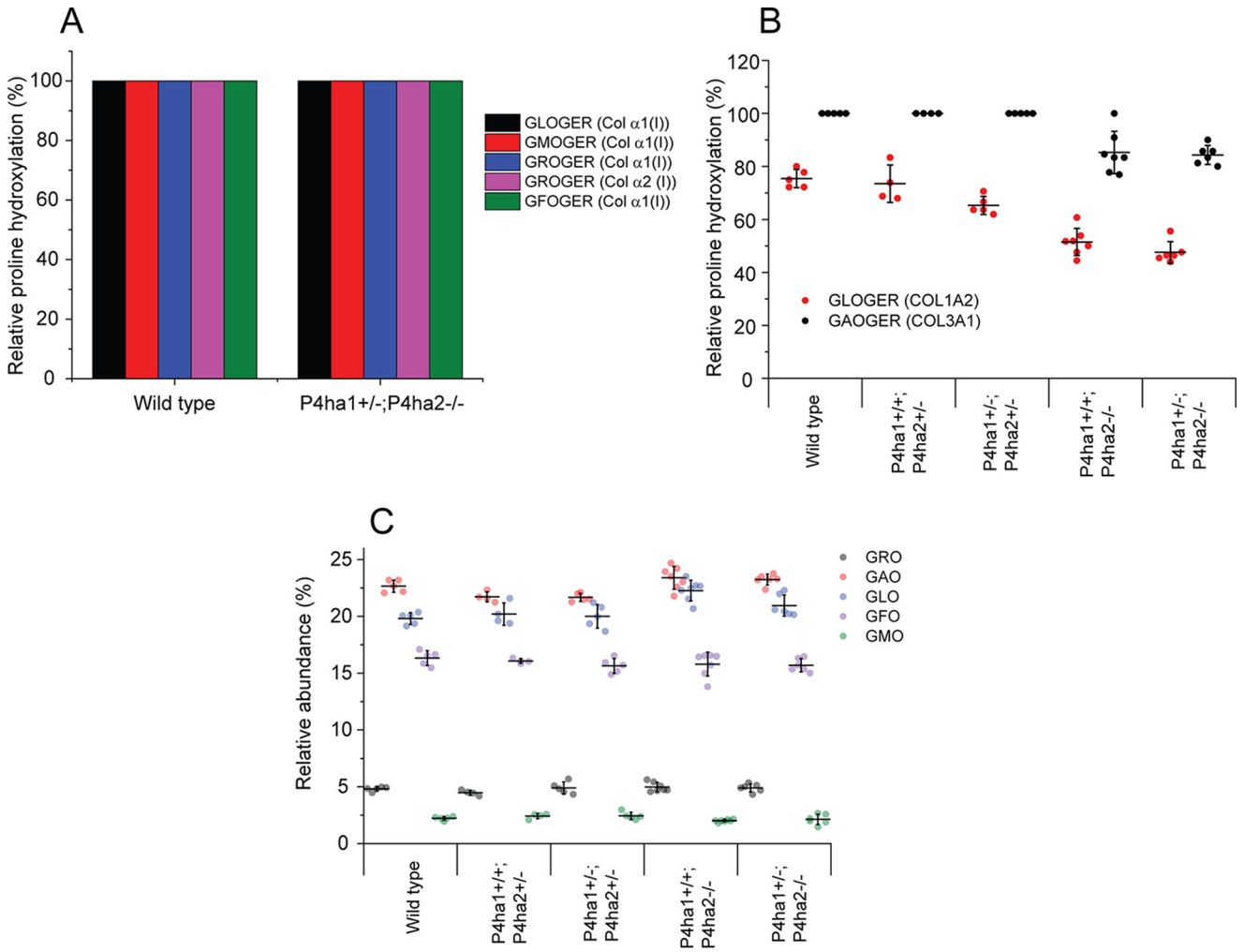


Figure 3. The hydroxylation of infrequent integrin-binding sites in collagen is changed due to genetic ablation of C-P4Hs. *A* and *B*, relative hydroxylation, calculated by dividing the number of unique spectra matching to a peptide containing a hydroxylated motif with the number of unique spectra matching a peptide containing hydroxylated or nonhydroxylated motifs, of unaffected sites (*A*) and affected sites (*B*). *C*, relative abundance of sequences containing the first triplet of the hydroxylated integrin binding sequence calculated as number of spectra matching to a peptide containing the hydroxylated triplet divided by number of all spectra matching to collagen peptides in the sample. The data points represent the average of two LC–MS/MS runs of collagen preparations from individual mice. Mean ± S.D. are shown.

integrin and hydroxyproline independently of triple-helix stability.

As reported earlier, the effect of prolyl hydroxylation on α1I binding was substantial (19). K_d of α1I domain binding to GFOGER was about 125 ± 4 nM, when it was 670 ± 130 nM in binding to GFPGER (Fig. 4*B*). In accordance with the data in Fig. 1*C*, α2I binding was only slightly weaker to GFPGER (K_d 30 ± 2 nM) than to GFOGER (K_d 23 ± 1 nM) (Fig. 4*C*). Batch-to-batch variation in the recombinant integrin α11I domain prevented valid interpretations concerning the binding of this αI domain. In accordance with the αI domain data, CHO-α1 integrin cells spread remarkably slower on GFPGER peptide when compared with GFOGER peptide (Fig. 4, *D* and *F*), whereas in the case of CHO-α2 integrin cells only a minor difference was detected (Fig. 4, *E* and *F*). Interestingly, CHO-α11 integrin cells strongly favored the adhesion to GFOGER peptide when compared with GFPGER (Fig. 4, *G* and *H*).

Further analyses of the real-time cell spreading revealed that CHO-α1 integrin cells adhered to nonhydroxylated peptide with slower velocity than to GFOGER peptide (Fig. 5*A*), whereas in the case of CHO-α2 integrin cells no difference was detected (Fig. 5*B*). Similar analysis of the integrin αI domain binding supported these observations (Fig. 5, *C* and *D*). Binding of the α2I domain to GFOGER and GFPGER similarly reached a steady binding state within the first 10 min (Fig. 5*D*), whereas α1I domain reached a steady binding state very slowly and there was a remarkable difference between GFOGER and GFPGER (Fig. 5*C*). Based on the observations that indicated the slow binding velocity, we hypothesized that hydroxylation of the proline residue in the GFOGER motif may be needed for proper activation of the integrin α1I domain. To test this possibility, we measured the binding velocity using a recombinant α1I harboring gain-of-function mutation E317A, which is known to pre-activate the integrin (21). The mutation did not abolish the difference between GFOGER and GFPGER (Fig. 5*E*). The observation suggested that the hydroxyproline–integrin interaction is essential during the initial interaction rather than at the later activation stages.

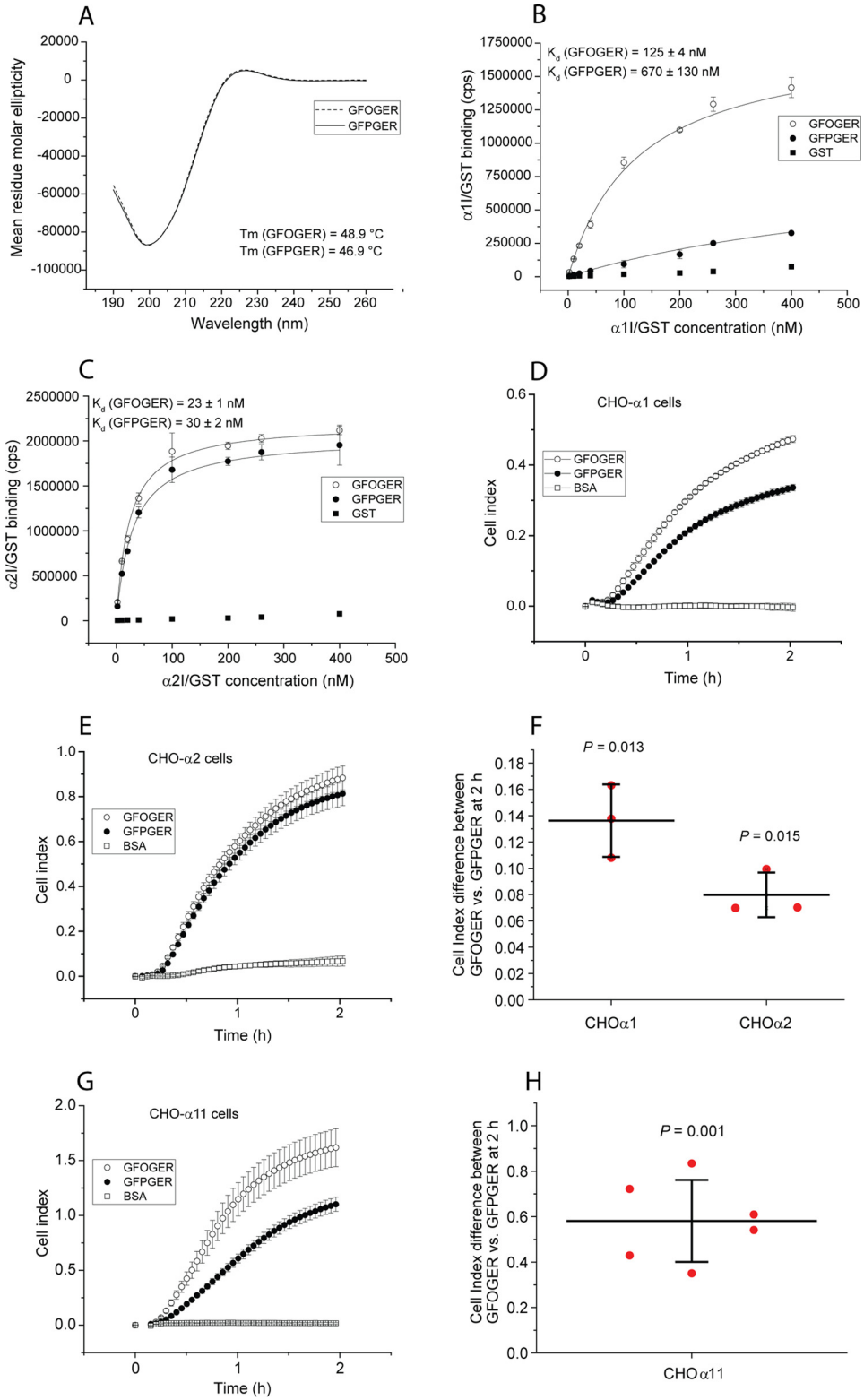


Figure 4. The lack of hydroxyproline in the integrin-binding motif GFOGER reduces the avidity of integrins α1 and α11 but only slightly the binding of integrin α2. *A*, CD-spectra for GFOGER and GFPGER peptides. The data were collected between 260 and 190 nm 22 °C using a 0.1-cm path length quartz cuvette at 0.1 mg/ml of protein concentration using triplicates. Melting temperatures for the peptides are shown. *B* and *C*, binding of α1I (*B*) and α2I (*C*) domains to GFPGER or GFOGER in a solid-phase binding assay (BSA as a background control). *D*, *E*, and *G*, spread of CHO cells transfected with α1 (*D*), α2 (*E*), or α11 (*G*) integrin on GFOGER or GFPGER analyzed by impedance-based xCELLigence technology (BSA as a background control). Mean ± S.D. is shown. Each sample was measured with three parallel wells. The difference in peptide binding was confirmed in 4 (α2I) or 3 (α1I) independent experiments. Estimates for the dissociation constants were obtained using an equation: measured binding = maximal binding/(1 + K_d/[α]). *F*, difference of cell adhesion between GFPGER and GFOGER in 2 h in the case of CHO-α1 and CHO-α2. *p* values from one-sample *t* test and mean ± S.D. of three independent experiments are shown. *H*, difference of cell adhesion between GFPGER and GFOGER in 2 h in the case of CHO-α11. *p* values from one-sample *t* test and mean ± S.D. of 7 independent experiments are shown.

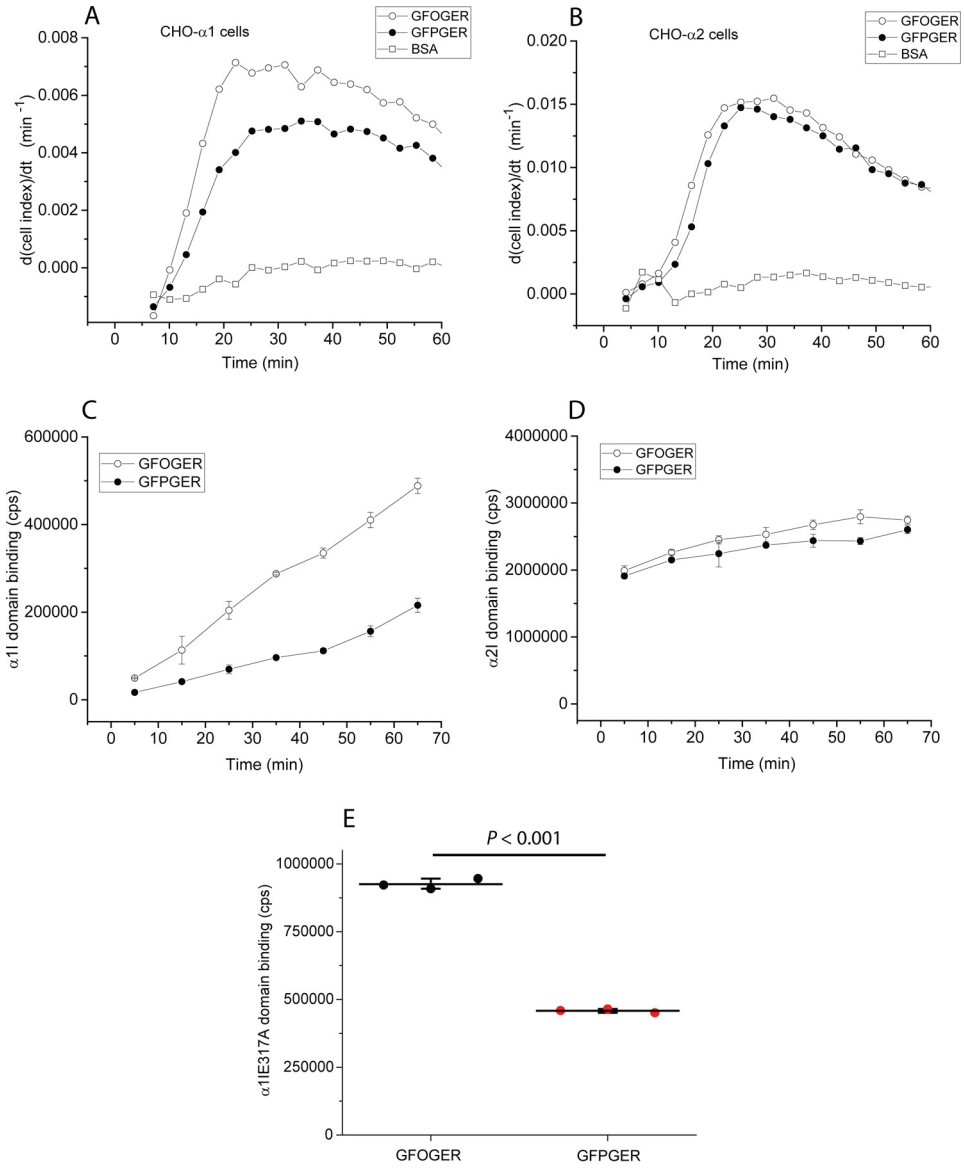
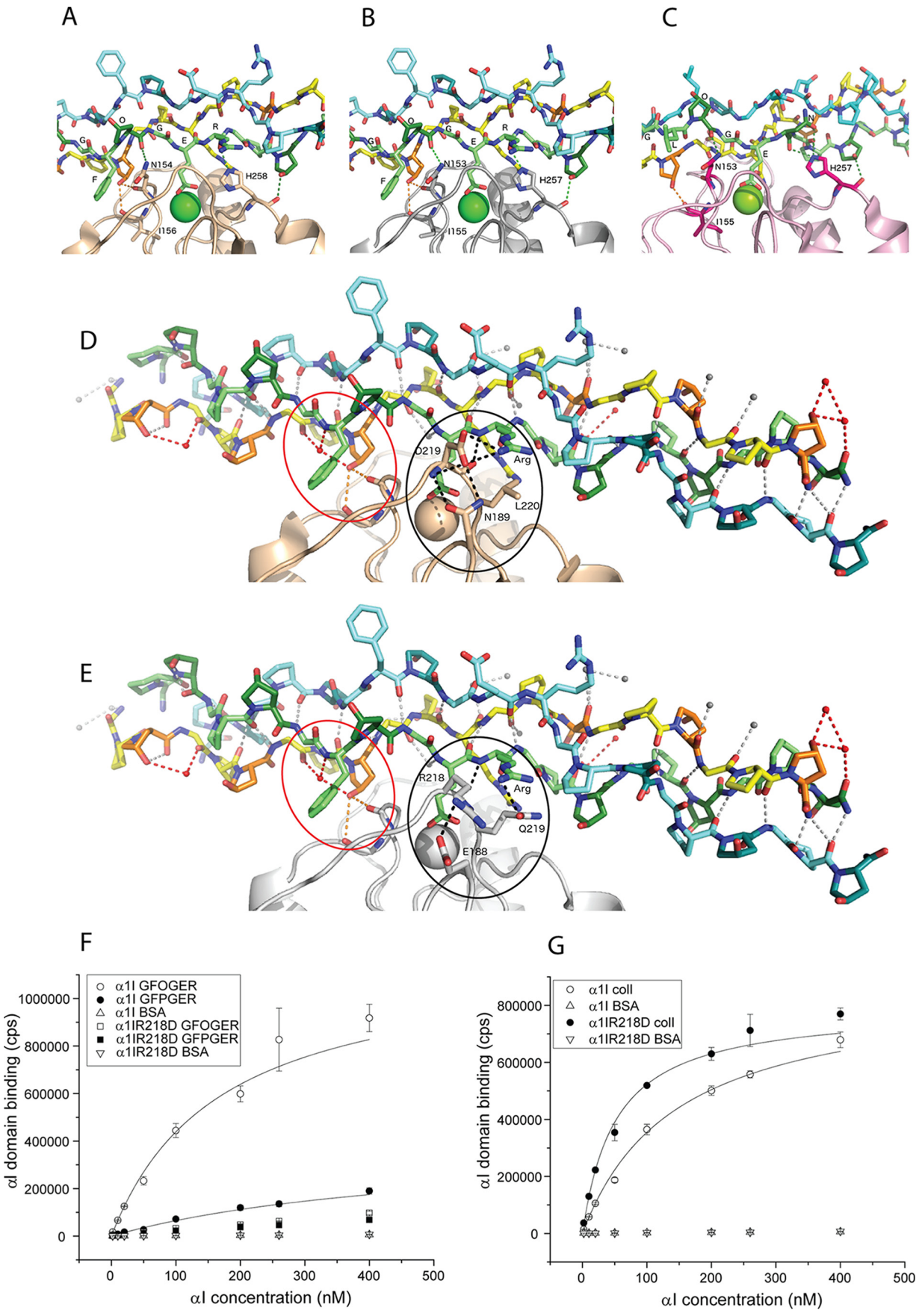


Figure 5. Lack of hydroxylation affects the binding of integrin $\alpha1\beta1$ both in nonactivated and preactivated conformation. *A* and *B,* spreading velocity of CHO-α1 (*A*) or CHO-α2 (*B*) cells on GFOGER and GFPGER as a function of time measured by xCELLigence technology. The experiments have been repeated three times with similar results. *C* and *D,* binding of recombinant α1I domain (*C*) or α2I domain (*D*), to GFPGER and GFOGER in a solid-phase binding assay as a function of time. The integrin αI domain concentration used in the experiments was 200 nM. The experiments have been repeated twice with similar results. *E,* binding of recombinant α1 E317I domain (α1 domain containing preactivation mutation) to GFPGER and GFOGER in a solid-phase binding assay. *p* values from Student's *t* test and mean ± S.D. are shown. The integrin αI domain concentration used in the experiment was 400 nM. The experiment has been repeated twice with similar results.

Arginine 218 in α1I domain is critical for the interaction between integrin and hydroxyproline residue in the integrin-binding site

To analyze the role of the hydroxyprolines in the integrin αI–peptide interactions, a structural model for integrin α1I–GFOFER was created and compared with the X-ray structures of integrin α2I–GFOGER (PDB code 1DZI) (11) and the α1I–GLOGEN complex (PDB code 2M32) (22) (Fig. 6, *A–C*). Interestingly, the interactions between the hydroxyprolines and αI are similar in the α2I–GFOGER (Fig. 6*A*) and α1I–GLOGEN (Fig. 6*C*) complexes and these interactions are conserved also in the α1I–GFOGER model (Fig. 6*B*), which was created assuming that α1I interacts with the GFOGER peptide similarly as α2I. The conserved residues that form hydrogen bonds with the hydroxyprolines are Ile-156, Ans-154, and His-258 in the integrin α2I domain and corresponding Ile-156, Asn-153, and His-257 in the integrin α1I domain (Fig. 6, *A–C*). The hydroxyproline in the trailing strand of the collagen (Fig. 6*A*, *orange*) forms hydrogen bonds with the carbonyl oxygen of Ile-156 and Asn-154, whose side chain nitrogen in turn makes a hydrogen bond with the carbonyl of a middle strand hydroxyproline (Fig. 6*A*, *green*). On the other side of the metal ion-dependent adhesion site (MIDAS), the carbonyl group of His-258 forms a hydrogen bond with the hydroxyl group of the middle strand hydroxyproline and the side chain of His-258 with the carbonyl of the middle strand Arg of GFOGER. All the integrin α2I domain residues (Asn-154, Tyr-157, and His-258) that form direct hydrogen bonds with the collagenous GFOGER peptide (11)


A
B
C
N154
H258
I156
N153
H257
I155
N153
H257
I155
D
O219
Arg
L220
N189
E
R218
Arg
Q219
E188
F
α1I GFOGER
α1I GFPGER
α1I BSA
α1IR218D GFOGER
α1IR218D GFPGER
α1IR218D BSA
αI domain binding (cps)
αI concentration (nM)
G
α1I coll
α1I BSA
α1IR218D coll
α1IR218D BSA
αI domain binding (cps)
αI concentration (nM)

are very conserved in the integrin α1I domain. However, the salt bridge between the middle strand Arg of the peptide and Asp-219 in α2I domain (Fig. 6*D*) is not conserved, as the corresponding residue in the α1I domain is a positively charged Arg-218, which cannot form a similar interaction. Based on the α1I–GFOGER model, Arg-218 presumably stacks with the middle strand Arg of GFOGER, which forms a hydrogen bond with Gln-219 (Fig. 6*E*). The position of Arg-218 in the α1I domain is stabilized by a salt bridge with Glu-188, whereas Asp-219 in the α2I domain is positioned by a hydrogen bond with Asn-189 (Fig. 6, *D* and *E*).

As the differing interactions did not directly involve collagen hydroxyprolines and, thus, did not explain why the integrin α1I domain binds more weakly to GFPGER than GFOGER, whereas the α2I domain binds them equally well, we next analyzed the water-mediated hydrogen bonds formed by the hydroxyprolines within GFOGER (Fig. 6, *D* and *E*, waters are shown as *red spheres* and hydrogen bonds with *red dashes*). Based on the analysis, all these bonds are formed by the hydroxyprolines in the trailing strand and mainly stabilize its structure, but one of the waters connects the trailing and middle strands (Fig. 6, *D* and *E, circles in red*). Intriguingly, the same hydroxyproline makes conserved interactions with the αI domain (Fig. 6, *A–C*). As the GFPGER peptide lacks the hydroxyprolines involved in the stabilizing water-mediated hydrogen bond network, its triple-helical structure is presumably more flexible compared with GFOGER. The lack of the water-mediated bonds is likely to weaken the positioning of the Arg in the middle strand. Thus, Arg might not be optimally oriented for the hydrogen bonding and stacking interactions in the α1I–GFPGER complex (Fig. 6*C*), which require more precise positioning than the stronger Arg–Asp-219 salt bridge within the α2I–GFPGER complex (Fig. 6*A*). Additionally, the differences in the residues in the vicinity of Asp-219 of the α2I domain (Leu-220; Asn-189) and Arg-218 of the α1I domain (Gln-219; Glu-188) and the formed interaction networks likely contribute to αI–peptide interactions (Fig. 6, *D* and *E*, *circles in black*).

The structural model was validated by solid-phase binding assays using variant domain α1I R218D. Domain α1I R218D bound very weakly to GFOGER peptide confirming the central role of Arg-218 in the interaction between integrin and hydroxyproline (Fig. 6*F*). In addition, α1I R218D binding to GFOGER and GFPGER was equally poor supporting the idea (Fig. 6*F*). In accordance with own previous data (23), binding of α1I R218D domain to collagen I was not affected, but actually slightly improved (Fig. 6*G*). Here, this observation serves as an important control showing that the α1I R218D domain is folded properly and is fully functional. In general, the result indicates that in collagen I there must be other high avidity-binding motifs than GFOGER for α1β1 integrin.

Discussion

Mammalian cells can directly recognize collagens by using a specific subgroup of integrins that, in complex with the β1 subunit, contain one of the four collagen-binding α subunits. These αI domain-harboring collagen receptor integrins can be found in vertebrates and lampreys but not in other multicellular animals (24). The co-evolution of collagen receptor integrins and collagens has led to the development of high affinity integrin recognition motifs in collagens. These sites, including the prototype motif GFOGER, typically contain one hydroxyproline residue. The biological role of direct cell adhesion to collagen is unclear. In some *in vitro* models integrin α2β1 and α2I domains can bind to collagen I fibrils (25), whereas other lines of experimentation suggest that cartilage type, collagen II containing fibrils may not be direct targets of the collagen receptors or their αI domains (26). Some researchers have suggested that the collagen receptors participate in the organization of the collagen fibrils (27, 28), but so far the described phenotypes of the α1, α2, and α11 integrin-deficient mice have not supported this idea. In general the phenotypes of collagen receptor knockouts are mild and show relatively small defects in immunoresponse, platelet function, wound healing, angiogenesis, bone fracture healing, or tooth development (29–34). Phenotypes of double knockouts have not revealed significant overlap or compensation in the functions of these receptors (35, 36). Studies on integrin α10-deficient mice have, however, revealed a reduced density of the collagen fibrillar network matrix in the growth plates (37). In general, the alterations in cartilage of the α10 knockout mice are similar as the defects in the P4ha1$^{+/-}$; P4ha2$^{-/-}$ mice (8), including abnormal arrangement of growth plate chondrocytes and an elevated number of apoptotic chondrocytes (8, 37).

Figure 6. Interaction of the GFOGER and GFPGER peptides with the αI domains. Despite the differences in the peptides and their binding modes to αI domain, the interactions with the αI domain including the key interaction between the metal ion and the Glu in peptide are similar in (*A*) the α2I-GFOGER (PDB ID 1DZI), (*B*) the modeled α1I-GFOGER, and (*C*) the α1I-GLOGEN (PDB ID 2M32) complexes. The leading, middle, and trailing strands are colored in *cyan*, *green*, and *orange*, respectively, and the hydroxyprolines within each strand in a darker color. The hydrogen bonds between the αI domain and the hydroxyprolines are shown as a *dashed line* and colored according to the color of the connected strands. In the GFOGER–α2I complex (*A*), one of the hydroxyprolines interacts with the main chain of His-528 and the other one with the main chain oxygens of Ile-156 and Asn-154. Furthermore, the side chain of His-528 interacts with the main chain oxygen of Arg in the peptide and the side chain nitrogen of Asn-154 with the main oxygen of the hydroxyproline in the peptide. These interactions are conserved in the modeled α1I–GFOGER complex (*B*) and in the X-ray structure of the GLOGEN–α1I complex (*C*). The salt bridge between the middle strand Arg of the peptide and Asp-219 in α2I (*D*; *circled in black*) is not conserved, as the corresponding residue Arg-218 in α1I cannot form a similar interaction (*E*; *circled in black*). Based on the α1I–GFOGER model (*E*), the middle strand Arg of GFOGER stacks with Arg-218 and forms a hydrogen bond with Gln-219. Glu-188 stabilizes the orientation of Arg-218 in α1I by a salt bridge (*E*) and, similarly, Asn-189 in α2I positions Asp-219 by a hydrogen bond (*D*). The waters and hydrogen bonds involving hydroxyprolines in the αI domain–peptide complexes are shown in *red* and the rest of in *gray*. One of the hydroxyprolines in the trailing strand (*circled in red*) makes a connection to the middle strand via a water-mediated hydrogen bond and the same hydroxyproline is also involved in the conserved interactions with the αI domain. *F*, binding of α1I domain and α1I R218D domain to GFPGER or GFOGER in a solid-phase binding assay (BSA as a background control). Each sample was measured with three parallel wells. Estimates for the dissociation constants were obtained using equation: measured binding = maximal binding/(1 + K_d/[αI]). *G*, binding of the α1I domain and α1I R218D domain to collagen I in a solid-phase binding assay (BSA as a background control). Each sample was measured with three parallel wells. Integrin αI domain concentration used in the experiments was 400 nM. Estimates for the dissociation constants were obtained using equation: measured binding = maximal binding/(1 + K_d/[αI]). The experiments in *F* and *G* have been repeated twice with similar results.

Here, we made an interesting observation that after concomitant inactivation of both alleles of P4ha2 and one allele of the P4ha1 gene all the major integrin-binding motifs were completely hydroxylated, despite the fact that hydroxylation of proline residues in -*X*-P-G- sequences dropped significantly (from 53 to 36%). There are two possible explanations for the result. It can be hypothesized that the site specificity of C-P4H-I has evolved to favor the integrin-recognition motifs over other sites in collagen or that only the tropocollagens that have intact integrin-binding sites end up to skin collagen fibrils. We cannot exclude either one of the two hypotheses, but they both stress the importance of direct integrin– collagen interaction.

Atomic structures of integrin α1I–GLOGEN and α2I–GFOGER complexes suggest that there is a direct interaction between the integrin and hydroxyproline residue (11, 22). Previous studies have, however, indicated that recognition of the GFOGER motif by the major epithelial cell collagen receptor, $\alpha2\beta1$ integrin, is not dependent on the hydroxylated proline residue (19). Our observations are in accordance with this result. Still, *in vivo* hypohydroxylation of proline residues results in collagen that cannot effectively interact with the α2I domain, when the test is performed in a physiological temperature. Thus, the structural stability of the collagen triple-helix may be more important regulator of integrin binding than the proline hydroxylation in the individual recognition motifs.

In agreement with a previous paper (19) we report here that the high avidity binding of α1I domain, and similarly $\alpha1\beta1$ integrin, requires the presence of hydroxyproline as part of the GFOGER motif. To study further the mechanism of this interaction we used molecular modeling and binding assay utilizing variant recombinant α1I domains. Based on the structural analysis, we were able to identify differences in the αI domain–peptide interactions that may result in the observed differences in the binding assay, despite the interactions that hydroxyprolines form with the αI domains being totally conserved in the X-ray structures of α2I–GFOGER (PDB code 1DZI) (11) and the α1I–GLOGEN complex (PDB code 2M32) (22), as well as in the created structural model of the α1I–GFOFER complex. Intriguingly, the water-mediated hydrogen bonds formed by one of the hydroxyprolines seem to be the key element for the different binding characteristics of α1I and α2I domains. Due to the lack of these bonds and the increased flexibility of the GFPGER triple-helix, the resulting decreased binding of α1I compared with α2I was predicted to originate from the Asp-219$^{\alpha2I}$/Arg-218$^{\alpha1I}$ and Leu-220$^{\alpha2I}$/Gln-219$^{\alpha1I}$ differences, which allows the aspartate in the α2I domain to form a salt bridge with the Arg of GFPGER, whereas the interactions formed by the α1I domain are weaker hydrogen bond and stacking interactions. In general, the interactions formed by these αI domain residues to the Arg and their positioning by the interacting residues (Asn-189$^{\alpha2I}$/Glu-188$^{\alpha1I}$) seems to be crucial for the collagen peptide binding properties.

The experiments with the variant α1I R218D domain revealed the importance of Arg-218 in the interaction between the α1I domain and GFOGER. Importantly, here and also in our previous paper (23) we have shown that the same mutation actually improves the interaction between the α1I domain and collagen I. Thus, $\alpha1\beta1$ integrin must also have other binding sites in collagen I. Collagen I is not a homotrimer, but composed of two α1(I) chains and one α2(I) chain. The presence of multiple α chains has complicated the search for integrin-binding sites in collagen I and also in other heterotrimeric collagens (23), whereas integrin recognition motifs have systematically been searched in homotrimeric collagens II and III, only. Indeed, there is evidence that in collagen IV all three α chains may participate in the formation of an $\alpha1\beta1$ integrin-binding site, in which the critical residues include one arginine and two aspartic acid residues (38–40). The presence of similar motifs in collagen I would also explain our results.

To conclude, our results show that intact collagen receptor function requires proper hydroxylation of collagen, which can partially explain the phenotypes of C-P4H mutant mice. We also provide two distinct molecular mechanisms, an indirect and a direct one, by which proline hydroxylation modulates integrin function.

Experimental procedures

Mouse lines with inactivated P4ha1 and P4ha2 genes

Mice with different combinations of inactivated P4ha1 and P4ha2 genes were generated and genotyped as described earlier (8). Briefly, P4ha1$^{+/-}$ mice were cross-bred with P4ha2$^{-/-}$ mice to first obtain double heterozygous P4ha1$^{+/-}$;P4ha2$^{+/-}$ mice, which were then intercrossed with P4ha2$^{-/-}$ mice. WT, P4ha1$^{+/+}$;P4ha2$^{+/-}$, P4ha1$^{+/-}$;P4ha2$^{+/-}$, P4ha1$^{+/+}$;P4ha2$^{-/-}$, and P4ha1$^{+/-}$;P4ha2$^{-/-}$ mice were used in the studies. Animal experiments were approved by the Animal Experiment Board of Finland, following the regulations of the EU Directive 86/609/EEC, the European Convention ETS123, and the national legislation of Finland. The recommendations given by the Federation of European Laboratory Animal Science Associations (FELASA) and the Finnish and EU legislations concerning laboratory animal experiments and handling were: approvals ESAVI/6154/04.10.07/2014, ESAVI/10390/04.10.07/2015, and ESAVI/8179/04.10.07/2017.

Collagen extraction and precipitation from the mouse skin

Fibrillar collagen was purified as described earlier (41), with slight modifications. Briefly, to extract collagen, skin punches were minced and incubated in 0.5 M acetic acid overnight at 4 °C. The tissue was minced further by vigorous shaking with stainless steel beads using TissueLyser LT (Qiagen) for 3 × 30 s, 50 Hz. After this, a pepsin-soluble collagen fraction was obtained by adding 1 mg/ml of pepsin for 1 week with constant mixing at 4 °C and collecting the supernatant after centrifugation at 13,000 × *g* for 30 min. Type I collagen was precipitated by adding 5 M NaCl in 0.05 M Tris-HCl, pH 7.4, to a final concentration of 0.7 M. Collagen containing pellet was obtained after overnight incubation at 4 °C followed by 21,000 × *g* centrifugation for 20 min. The pellet was washed by resuspension in 4 M NaCl and overnight incubation at 4 °C, followed by 21,000 × *g* centrifugation for 20 min. The wash was repeated once and the final collagen pellet was suspended in 0.1 M acetic acid.

Thermal stability analysis of collagen

The native triple-helical conformation of a collagen molecule is resistant to most proteinases and becomes susceptible for digestion only after unfolding of the triple helix, for example, by thermal denaturation, and proteolytic enzymes are used as probes for quantitative assay of fully aligned triple-helical collagen molecules (52). The thermal stability of the precipitated skin collagen was analyzed by trypsin/chymotrypsin digestion performed at various temperatures (52). Collagen concentration was measured with Direct Detect® IR spectrometer (Millipore Corporation) and 25-μl aliquots were neutralized and preincubated at 35 °C for 5 min. 2.5 μl of trypsin/chymotrypsin solution (0.1 mg/ml of trypsin (Sigma) and 0.25 mg/ml of chymotrypsin (Sigma) in 0.2 M NaCl, 50 mM Tris-HCl buffer, pH 7.4) was added to the samples, which were then incubated at 36, 37, 38, 39, 40, or 41 °C for 2 min. The reaction was stopped with 2.5 μl of trypsin inhibitor solution (0.5 mg/ml of trypsin inhibitor (Honeywell Fluka, Morris Plains, NJ), in 0.2 M NaCl, 50 mM Tris-HCl buffer, pH 7.4). Trypsin/chymotrypsin was omitted from the undigested control sample. Samples were analyzed by 8% SDS-PAGE under reducing conditions, stained with Coomassie Brilliant Blue R-250, destained with 10% EtOH and 10% CH_3COOH in H_2O, and imaged with Gel Doc™ RX+ Molecular Imager (Bio-Rad).

CD

Circular dichroism (CD) spectroscopy was performed using Chirasca CD Spectrometer (Applied Photophysics, Leatherhead, UK). The data were collected between 190 and 260 nm at 22 °C using a 0.1-cm path length quartz cuvette triplicate samples with 0.1 mg/ml of protein concentration. CD measurements were acquired every 1 nm with 1 s as an integration time and repeated three times with baseline correction. The data were analyzed with Pro-Data Viewer (Applied Photophysics). Thermal unfolding was recorded between 190 and 260 nm with a 2 °C step size at 1 °C/min ramp rate with ± 0.2 °C tolerance. The melting temperature was analyzed with Global3 (Applied Photophysics).

Mass spectrometry

The collagen samples were vacuum-dried and dissolved in 6 M urea in 25 mM ammonium bicarbonate. The cysteines were reduced by 10 mM DTT at 37 °C for 1 h and alkylated by 40 mM iodoacetamide at room temperature for 30 min. The solution was diluted to 1:10 with 25 mM ammonium bicarbonate, and the proteins were digested with 0.5 μg of trypsin (Promega, Madison, WI) per 10 μg of protein at 37 °C overnight. The peptides were desalted by StageTips (42), vacuum dried, dissolved in 1% formic acid, and loaded on a nanoflow HPLC system (Easy-nLC1000, Thermo Fisher Scientific) coupled to the Q Exactive mass spectrometer (Thermo Fisher Scientific) equipped with a nano-electrospray ionization source. The peptides were first loaded on a trapping column and subsequently separated in-line on a 15-cm C18 column (75 μm × 15 cm, Magic 5 μm 200 Å C18, Michrom BioResources Inc., Sacramento, CA). The mobile phase consisted of water/acetonitrile (98:2 (v/v)) with 0.2% formic acid (solvent A) and acetonitrile/water (95:5 (v/v)) with 0.2% formic acid (solvent B). The peptides were separated with a 10-min gradient from 5 to 35% of solvent B. Before the end of the run, the percentage of solvent B was raised to 100% in 2 min and kept there for 8 min. Data-dependent acquisition was enabled and higher energy collisional dissociation of the top 5 ions from the survey scan over the mass-to-charge (m/z) range 300–2000 with dynamic exclusion time of 10 s was performed in each cycle. Two repeated runs per sample were performed. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (43) partner repository with the dataset identifier PXD008802.

Data analysis

Tandem mass spectra were searched against mouse Swissprot sequences (release 2014_08) with Proteome Discoverer software, version 1.4 (Thermo Fischer Scientific) using the Mascot search engine (Matrix Science, version 2.4) allowing for 5 ppm precursor mass tolerance and 0.02 Da fragment mass tolerance. Carbamidomethyl (C) as a fixed modification, and oxidation (M, K, P) as dynamic modifications were included. Maximum of two missed cleavages were allowed. Decoy database search using reversed mouse SwissProt sequences was used to assess the false discovery rate. Only the peptides with false discovery rate <0.05 and determined as "rank 1" by Proteome Discoverer software were accepted for further analysis. All statistical tests were done by IBM SPSS Statistics, version 23.

Integrin αI domains and solid-phase binding assays

The cDNA of integrin α1, α2, and α11 I domains or mutants α1I E317A and α1I R218D were transformed to *Escherichia coli* BL21 Tuner (Novagen-Merck-EMD-Millipore, Darmstadt, Germany) in GEX-2T vector (21, 44–47). The αI domains were produced as glutathione *S*-transferase fusions and purified by GSH-agarose affinity chromatography as described earlier (46). Solid-phase binding assays of the proteins were performed at room temperature or at 37 °C using a published method (23). Briefly, 96-well plates (DNA bind; Corning Costar, Corning, NY) were coated with 1.64 μg/cm^2 of the isolated mouse type I collagen or GPC-$(GPP)_5$-GFOGER-$(GPP)_5$-GPC-NH_2 (O = hydroxyproline) or GPC-$(GPP)_5$-GFPGER-$(GPP)_5$-GPC-NH_2 (Auspep, Tullamarine, VIC, Australia) in PBS, pH 8.5, for 1 h. The wells were blocked by using Delfia diluent II (PerkinElmer Life Sciences) in PBS (1:1 dilution, 1 h). Recombinant αI domains were diluted (400 nM if not otherwise mentioned) in Delfia assay buffer (2 mM $MgCl_2$, PerkinElmer Life Sciences) and were added into the blocked wells. After the incubation (1 h, if not otherwise indicated) wells were washed three times and Europium-labeled anti-glutathione *S*-transferase antibody (PerkinElmer) was added (1:1000 dilution, 1 h). Finally, the wells were washed three times and Enhancement solution (PerkinElmer Life Sciences) was added followed by the measurement of signal with time resolved fluorescence spectrophotometry (Victor^3 multilabel counter; PerkinElmer Life Sciences).

Cell cultures

Chinese hamster ovary cells (CHO, ATCC) were stably transfected with human integrin α1 (21), α2 (45), or α11 (48)

subunits as described. Cells were cultured in α minimum essential medium (Gibco) supplemented with 10% heat-inactivated fetal calf serum (PromoCell), 2 mM glutamine, 100 IU/ml of penicillin G, and 100 μg/ml of streptomycin. CHO-α1 cells were cultured in the presence of 150 μg/ml of Zeocin (Invitrogen), with other cell lines, 500 μg/ml of G418 (Invitrogen) was used.

Real-time cell adhesion

Plates (96-well E-Plate 96; ACEA Biosciences Inc.) were coated with synthesized triple-helical collagen mimetic peptides GFOGER or GFPGER (see above, 10 μg/ml in PBS) or BSA (1 mg/ml in PBS) overnight at +4 °C. The nonspecific binding sites were blocked with BSA (1 mg/ml, 37 °C, 1 h). 30,000 CHO cells/well were seeded in serum-free α-minimum Eagle's medium. Cell adhesion and spreading was followed for at least 24 h with the xCELLigence RTCA (ACEA Biosciences Inc.). The velocity of cell spreading was calculated by the following equation: $(C2 - C1)/(t2 - t1)$, where C is a cell index value at a specific time point and t is time.

3D modeling and structural analysis

A 3D model for the integrin α1I domain (UniProt Knowledgebase (UniProtKB) P56199) in complex with the GFOGER peptide was constructed using the α2I–GFOGER peptide complex as a structural template (PDB code 1DZI) (11), as the binding studies suggest that the binding mode of the α1I domain is similar to that of α2I (18). The model with the lowest value of MODELLER objective function (49) of the 10 created models was chosen for structural analysis and visualization. The intermolecular interactions were studied by the "Protein interfaces, surfaces and assemblies" service (PISA) at the European Bioinformatics Institute (50). The figures and structural analysis were done with PyMOL (51).

Author contributions—K. H. S., J. M., and J. H. conceptualization; K. H. S., K. D., P. R., J. J., T. A. S., A. M. S., and J. K. investigation; K. H. S. and J. H. writing-original draft; K. H. S., K. D., P. R., J. J., T. A. S., A. M. S., J. K., J. M., and J. H. writing-review and editing; P. R. and T. A. S. methodology; T. A. S., J. M., and J. H. funding acquisition; J. M. and J. H. resources; J. M. and J. H. supervision; J. M. and J. H. project administration.

Acknowledgments—We thank Anne Kokko and Maria Tuominen for excellent technical assistance and Dr. Hongmin Tu from the Biocenter Oulu core facility for protein analysis (co-funded by the University of Oulu and Biocenter Finland) for help in CD analyses. We thank the bioinformatics (J.V. Lehtonen), translational activities, and structural biology infrastructure support from the Biocenter Finland, and the CSC IT Center for Science for computational infrastructure support at the Structural Bioinformatics Laboratory, Åbo Akademi University. We also thank the University of Oulu Laboratory Animal Center.

References

1. Myllyharju, J., and Kivirikko, K. I. (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* **20,** 33–43 CrossRef Medline
2. Rosenbloom, J., Harsch, M., and Jimenez, S. (1973) Hydroxyproline content determines the denaturation temperature of chick tendon collagen. *Arch. Biochem. Biophys.* **158,** 478–484 CrossRef Medline
3. Kivirikko, K. I., Myllylä, R., and Pihlajaniemi, T. (1992) Prolyl hydroxylase, protein disulfide isomerase, and other structurally related proteins. in *Post-Translational Modifications of Proteins* (Harding, J. J., and Crabbe, M. J. C., eds) pp. 1–51, CRC Press, Boca Raton, FL
4. Grosso, G., Bei, R., Mistretta, A., Marventano, S., Calabrese, G., Masuelli, L., Giganti, M. G., Modesti, A., Galvano, F., and Gazzolo, D. (2013) Effects of vitamin C on health: a review of evidence. *Front. Biosci.* **18,** 1017–1029 CrossRef
5. Helaakoski, T., Annunen, P., Vuori, K., MacNeil, I. A., Pihlajaniemi, T., and Kivirikko, K. I. (1995) Cloning, baculovirus expression, and characterization of a second mouse prolyl 4-hydroxylase α-subunit isoform: formation of an α2β2 tetramer with the protein disulfide-isomerase/β subunit. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 4427–4431 CrossRef Medline
6. Annunen, P., Helaakoski, T., Myllyharju, J., Veijola, J., Pihlajaniemi, T., and Kivirikko, K. I. (1997) Cloning of the human prolyl 4-hydroxylase α subunit isoform α(II) and characterization of the type II enzyme tetramer: the α(I) and α(II) subunits do not form a mixed α(I)α(II)β2 tetramer. *J. Biol. Chem.* **272,** 17342–17348 CrossRef Medline
7. Holster, T., Pakkanen, O., Soininen, R., Sormunen, R., Nokelainen, M., Kivirikko, K. I., and Myllyharju, J. (2007) Loss of assembly of the main basement membrane collagen, type IV, but not fibril-forming collagens and embryonic death in collagen prolyl 4-hydroxylase I null mice. *J. Biol. Chem.* **282,** 2512–2519 CrossRef Medline
8. Aro, E., Salo, A. M., Khatri, R., Finnilä, M., Miinalainen, I., Sormunen, R., Pakkanen, O., Holster, T., Soininen, R., Prein, C., Clausen-Schaumann, H., Aszódi, A., Tuukkanen, J., Kivirikko, K. I., Schipani, E., and Myllyharju, J. (2015) Severe extracellular matrix abnormalities and chondrodysplasia in mice lacking collagen prolyl 4-hydroxylase isoenzyme II in combination with a reduced amount of isoenzyme I. *J. Biol. Chem.* **290,** 16964–16978 CrossRef Medline
9. Zou, Y., Donkervoort, S., Salo, A. M., Foley, A. R., Barnes, A. M., Hu, Y., Makareeva, E., Leach, M. E., Mohassel, P., Dastgir, J., Deardorff, M. A., Cohn, R. D., DiNonno, W. O., Malfait, F., Lek, M., Leikin, S., Marini, J. C., Myllyharju, J., and Bönnemann, C. G. (2017) P4HA1 mutations cause a unique congenital disorder of connective tissue involving tendon, bone, muscle and the eye. *Hum. Mol. Genet.* **26,** 2207–2217 CrossRef Medline
10. Zeltz, C., and Gullberg, D. (2016) The integrin-collagen connection: a glue for tissue repair? *J. Cell Sci.* **129,** 653–664 CrossRef Medline
11. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000) Structural basis of collagen recognition by integrin α2β1. *Cell* **101,** 47–56 CrossRef Medline
12. Knight, C. G., Morton, L. F., Onley, D. J., Peachey, A. R., Messent, A. J., Smethurst, P. A., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (1998) Identification in collagen type I of an integrin α2β1-binding site containing an essential GER sequence. *J. Biol. Chem.* **273,** 33287–33294 CrossRef Medline
13. Kim, J. K., Xu, Y., Xu, X., Keene, D. R., Gurusiddappa, S., Liang, X., Wary, K. K., and Höök, M. (2005) A novel binding site in collagen type III for integrins α1β1 and α2β1. *J. Biol. Chem.* **280,** 32512–32520 CrossRef Medline
14. 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 α2β1-binding motifs in human collagen III. *J. Biol. Chem.* **281,** 3821–3831 CrossRef Medline
15. Siljander, P. R., Hamaia, S., Peachey, A. R., Slatter, D. A., Smethurst, P. A., Ouwehand, W. H., Knight, C. G., and Farndale, R. W. (2004) Integrin activation state determines selectivity for novel recognition sites in fibrillar collagens. *J. Biol. Chem.* **279,** 47763–47772 CrossRef Medline
16. Hamaia, S. W., Pugh, N., Raynal, N., Némoz, B., Stone, R., Gullberg, D., Bihan, D., and Farndale, R. W. (2012) Mapping of potent and specific binding motifs, GLOGEN and GVOGEA, for integrin α1β1 using collagen toolkits II and III. *J. Biol. Chem.* **287,** 26019–26028 CrossRef Medline

17. Zwolanek, D., Veit, G., Eble, J. A., Gullberg, D., Ruggiero, F., Heino, J., Meier, M., Stetefeld, J., and Koch, M. (2014) Collagen XXII binds to collagen-binding integrins via the novel motifs GLQGER and GFKGER. *Biochem. J.* **459,** 217–227 CrossRef Medline
18. 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 CrossRef Medline
19. Perret, S., Eble, J. A., Siljander, P. R., Merle, C., Farndale, R. W., Theisen, M., and Ruggiero, F. (2003) Prolyl hydroxylation of collagen type I is required for efficient binding to integrin $\alpha1\beta1$ and platelet glycoprotein VI but not to $\alpha2\beta1$. *J. Biol. Chem.* **278,** 29873–29879 CrossRef Medline
20. Seo, N., Russell, B. H., Rivera, J. J., Liang, X., Xu, X., Afshar-Kharghan, V., and Höök, M. (2010) An engineered $\alpha1$ integrin-binding collagenous sequence. *J. Biol. Chem.* **285,** 31046–31054 CrossRef Medline
21. Lahti, M., Bligt, E., Niskanen, H., Parkash, V., Brandt, A.-M., Jokinen, J., Patrikainen, P., Käpylä, J., Heino, J., and Salminen, T. A. (2011) Structure of collagen receptor integrin $\alpha(1)$I domain carrying the activating mutation E317A. *J. Biol. Chem.* **286,** 43343–43351 CrossRef Medline
22. Chin, Y. K., Headey, S. J., Mohanty, B., Patil, R., McEwan, P. A., Swarbrick, J. D., Mulhern, T. D., Emsley, J., Simpson, J. S., and Scanlon, M. J. (2013) The structure of integrin $\alpha1$I domain in complex with a collagen-mimetic peptide. *J. Biol. Chem.* **288,** 36796–36809 CrossRef Medline
23. Käpylä, J., Jäälinoja, J., Tulla, M., Ylöstalo, J., Nissinen, L., Viitasalo, T., Vehviläinen, P., Marjomäki, V., Nykvist, P., Säämänen, A.-M., Farndale, R. W., Birk, D. E., Ala-Kokko, L., and Heino, J. (2004) The fibril-associated collagen IX provides a novel mechanism for cell adhesion to cartilaginous matrix. *J. Biol. Chem.* **279,** 51677–51687 CrossRef Medline
24. Johnson, M. S., Käpylä, J., Denessiouk, K., Airenne, T., Chouhan, B., and Heino, J. (2013) Evolution of cell adhesion to extracellular matrix. in *Evolution of Extracellular Matrix* (Keeley, F., and Mecham, R., eds) pp. 243–283, Springer, New York
25. Jokinen, J., Dadu, E., Nykvist, P., Käpylä, J., White, D. J., Ivaska, J., Vehviläinen, P., Reunanen, H., Larjava, H., Häkkinen, L., and Heino, J. (2004) Integrin-mediated cell adhesion to type I collagen fibrils. *J. Biol. Chem.* **279,** 31956–31963 CrossRef Medline
26. Woltersdorf, C., Bonk, M., Leitinger, B., Huhtala, M., Käpylä, J., Heino, J., Gil Girol, C., Niland, S., Eble, J. A., Bruckner, P., Dreier, R., and Hansen, U. (2017) The binding capacity of $\alpha1\beta1$-, $\alpha2\beta1$- and $\alpha10\beta1$-integrins depends on non-collagenous surface macromolecules rather than the collagens in cartilage fibrils. *Matrix Biol.* **63,** 91–105 CrossRef Medline
27. Velling, T., Risteli, J., Wennerberg, K., Mosher, D. F., and Johansson, S. (2002) Polymerization of type I and III collagens is dependent on fibronectin and enhanced by integrins $\alpha11\beta1$ and $\alpha2\beta1$. *J. Biol. Chem.* **277,** 37377–37381 CrossRef Medline
28. Li, S., Van Den Diepstraten, C., D'Souza, S. J., Chan, B. M. C., and Pickering, J. G. (2003) Vascular smooth muscle cells orchestrate the assembly of type I collagen via $\alpha2\beta1$ integrin, RhoA, and fibronectin polymerization. *Am. J. Pathol.* **163,** 1045–1056 CrossRef Medline
29. Gardner, H., Broberg, A., Pozzi, A., Laato, M., and Heino, J. (1999) Absence of integrin $\alpha1\beta1$ in the mouse causes loss of feedback regulation of collagen synthesis in normal and wounded dermis. *J. Cell Sci.* **112,** 263–272 Medline
30. Ekholm, E., Hankenson, K. D., Uusitalo, H., Hiltunen, A., Gardner, H., Heino, J., and Penttinen, R. (2002) Diminished callus size and cartilage synthesis in $\alpha1\beta1$ integrin-deficient mice during bone fracture healing. *Am. J. Pathol.* **160,** 1779–1785 CrossRef Medline
31. Holtkötter, O., Nieswandt, B., Smyth, N., Müller, W., Hafner, M., Schulte, V., Krieg, T., and Eckes, B. (2002) Integrin $\alpha2$-deficient mice develop normally, are fertile, but display partially defective platelet interaction with collagen. *J. Biol. Chem.* **277,** 10789–10794 CrossRef Medline
32. Chen, J., Diacovo, T. G., Grenache, D. G., Santoro, S. A., and Zutter, M. M. (2002) The $\alpha(2)$ integrin subunit-deficient mouse: a multifaceted phenotype including defects of branching morphogenesis and hemostasis. *Am. J. Pathol.* **161,** 337–344 CrossRef Medline
33. Popova, S. N., Barczyk, M., Tiger, C.-F., Beertsen, W., Zigrino, P., Aszodi, A., Miosge, N., Forsberg, E., and Gullberg, D. (2007) $\alpha11\beta1$ integrin-dependent regulation of periodontal ligament function in the erupting mouse incisor. *Mol. Cell. Biol.* **27,** 4306–4316 CrossRef Medline
34. Zweers, M. C., Davidson, J. M., Pozzi, A., Hallinger, R., Janz, K., Quondamatteo, F., Leutgeb, B., Krieg, T., and Eckes, B. (2007) Integrin $\alpha2\beta1$ is required for regulation of murine wound angiogenesis but is dispensable for reepithelialization. *J. Invest. Dermatol.* **127,** 467–478 CrossRef Medline
35. Blumbach, K., Niehoff, A., Belgardt, B. F., Ehlen, H. W., Schmitz, M., Hallinger, R., Schulz, J.-N., Brüning, J. C., Krieg, T., Schubert, M., Gullberg, D., and Eckes, B. (2012) Dwarfism in mice lacking collagen-binding integrins $\alpha2\beta1$ and $\alpha11\beta1$ is caused by severely diminished IGF-1 levels. *J. Biol. Chem.* **287,** 6431–6440 CrossRef Medline
36. Ghatak, S., Niland, S., Schulz, J.-N., Wang, F., Eble, J. A., Leitges, M., Mauch, C., Krieg, T., Zigrino, P., and Eckes, B. (2016) Role of integrins $\alpha1\beta1$ and $\alpha2\beta1$ in wound and tumor angiogenesis in mice. *Am. J. Pathol.* **186,** 3011–3027 CrossRef Medline
37. Bengtsson, T., Aszodi, A., Nicolae, C., Hunziker, E. B., Lundgren-Akerlund, E., and Fässler, R. (2005) Loss of $\alpha10\beta1$ integrin expression leads to moderate dysfunction of growth plate chondrocytes. *J. Cell Sci.* **118,** 929–936 CrossRef Medline
38. Eble, J. A., Golbik, R., Mann, K., and Kühn, K. (1993) The $\alpha1\beta1$ integrin recognition site of the basement membrane collagen molecule $[\alpha1(IV)]2$ $\alpha2(IV)$. *EMBO J.* **12,** 4795–4802 Medline
39. Golbik, R., Eble, J. A., Ries, A., and Kühn, K. (2000) The spatial orientation of the essential amino acid residues arginine and aspartate within the $\alpha1\beta1$ integrin recognition site of collagen IV has been resolved using fluorescence resonance energy transfer. *J. Mol. Biol.* **297,** 501–509 CrossRef Medline
40. Saccà, B., Fiori, S., and Moroder, L. (2003) Studies of the local conformational properties of the cell-adhesion domain of collagen type IV in synthetic heterotrimeric peptides. *Biochemistry* **42,** 3429–3436 CrossRef Medline
41. Miller, E. J., and Rhodes, R. K. (1982) Preparation and characterization of the different types of collagen. *Methods Enzymol.* **82,** 33–64 CrossRef Medline
42. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670 CrossRef Medline
43. Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44,** D447–D456 CrossRef Medline
44. Käpylä, J., Ivaska, J., Riikonen, R., Nykvist, P., Pentikäinen, O., Johnson, M., and Heino, J. (2000) Integrin $\alpha(2)$I domain recognizes type I and type IV collagens by different mechanisms. *J. Biol. Chem.* **275,** 3348–3354 CrossRef Medline
45. Nykvist, P., Tu, H., Ivaska, J., Käpylä, J., Pihlajaniemi, T., and Heino, J. (2000) Distinct recognition of collagen subtypes by $\alpha(1)\beta(1)$ and $\alpha(2)\beta(1)$ integrins: $\alpha(1)\beta(1)$ mediates cell adhesion to type XIII collagen. *J. Biol. Chem.* **275,** 8255–8261 CrossRef Medline
46. Tulla, M., Pentikäinen, O. T., Viitasalo, T., Käpylä, J., Impola, U., Nykvist, P., Nissinen, L., Johnson, M. S., and Heino, J. (2001) Selective binding of collagen subtypes by integrin $\alpha1$I, $\alpha2$I and $\alpha10$I domains. *J. Biol. Chem.* **276,** 48206–48212 CrossRef Medline
47. Zhang, W.-M., Kapyla, J., Puranen, J. S., Knight, C. G., Tiger, C.-F., Pentikainen, O. T., Johnson, M. S., Farndale, R. W., Heino, J., and Gullberg, D. (2003) $\alpha11\beta1$ integrin recognizes the GFOGER sequence in interstitial collagens. *J. Biol. Chem.* **278,** 7270–7277 CrossRef Medline
48. Sipilä, K., Haag, S., Denessiouk, K., Käpylä, J., Peters, E. C., Denesyuk, A., Hansen, U., Konttinen, Y., Johnson, M. S., Holmdahl, R., and Heino, J. (2014) Citrullination of collagen II affects integrin-mediated cell adhesion in a receptor-specific manner. *FASEB J.* **28,** 3758–3768 CrossRef Medline

49. Šali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815 CrossRef Medline
50. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797 CrossRef Medline
51. DeLano, W. L. (2002) *The PyMOL molecular graphics system,* DeLano Scientific, San Carlos, CA
52. Bruckner, P., and Prockop, D. J. (1981) Proteolytic enzymes as probes for the triple-helical conformation of procollagen. *Anal. Biochem.* **110,** 360–368 Medline