The FEBS Journal

FEBS PRESS science publishing by scientists

Defining the role of αC helix interactions in the activation of the integrin αI domain

Liisa Pösö[1,2], Tomi T. Airenne[1], Jarmo Käpylä[2], Hongmin Tu[3], Eva Bligt-Lindén[1], Vimal Parkash[1], Jyrki Heino[2] and Tiina A. Salminen[1]

1 Structural Bioinformatics Laboratory Biochemistry, Faculty of Science and Engineering, and InFLAMES Research Flagship Center, Åbo Akademi University, Turku, Finland
2 Biochemistry, Department of Life Technologies, Faculty of Technology, and InFLAMES Research Flagship, University of Turku, Finland
3 Biocenter Oulu, University of Oulu, Finland

Keywords
collagen; integrin; integrin activation; signal transduction; VLA-2; X-ray crystallography

Correspondence
T. A. Salminen, Structural Bioinformatics Laboratory and InFLAMES Research Flagship Center, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, Tykistökatu 6A, 20520 Turku, Finland
Tel: +358 405151201
E-mail: tiina.salminen@abo.fi

Liisa Pösö and Tomi T. Airenne contributed equally to this work.

(Received 18 December 2025, revised 12 May 2026, accepted 10 July 2026)

doi:10.1111/febs.70658

Four αI domain-containing integrins (α1β1, α2β1, α10β1, and α11β1) have evolved to recognize various members of the collagen family. A defining structural feature of these receptors is the presence of the αC helix within their αI domains (Glu284–Arg288 in the α2I domain), a structure solely found in collagen-binding integrins, whose functional mechanism has remained unclear. To elucidate the functional role of the αC helix, and to assess the contribution of the mechanistically important Arg288-Glu318 ion pair in α2I domain activation, we created two variants α2IR288A and α2IE318A. Functional solid-phase binding assays and surface plasmon resonance revealed that these variants exhibit strikingly increased avidity for collagens I and IV, whereas affinities for triple-helical GFOGER peptides, representing a single binding motif, were only slightly increased. The variants displayed distinct ligand-binding profiles, including differences in association and dissociation constants. To understand these differences at a structural level, we determined four novel X-ray crystal structures of the variants, including both ligand-free and ligand-bound (succinic acid or malic acid) states. These structures revealed differences in the folding of the αC helix region, suggesting that its conformation regulates α2I domain activation and ligand specificity, consistent with our binding assays. Additionally, the structures captured the movements of the catalytic metal ion in the metal ion dependent adhesion site, providing a detailed view of the sophisticated activation mechanism of the α2I domain.

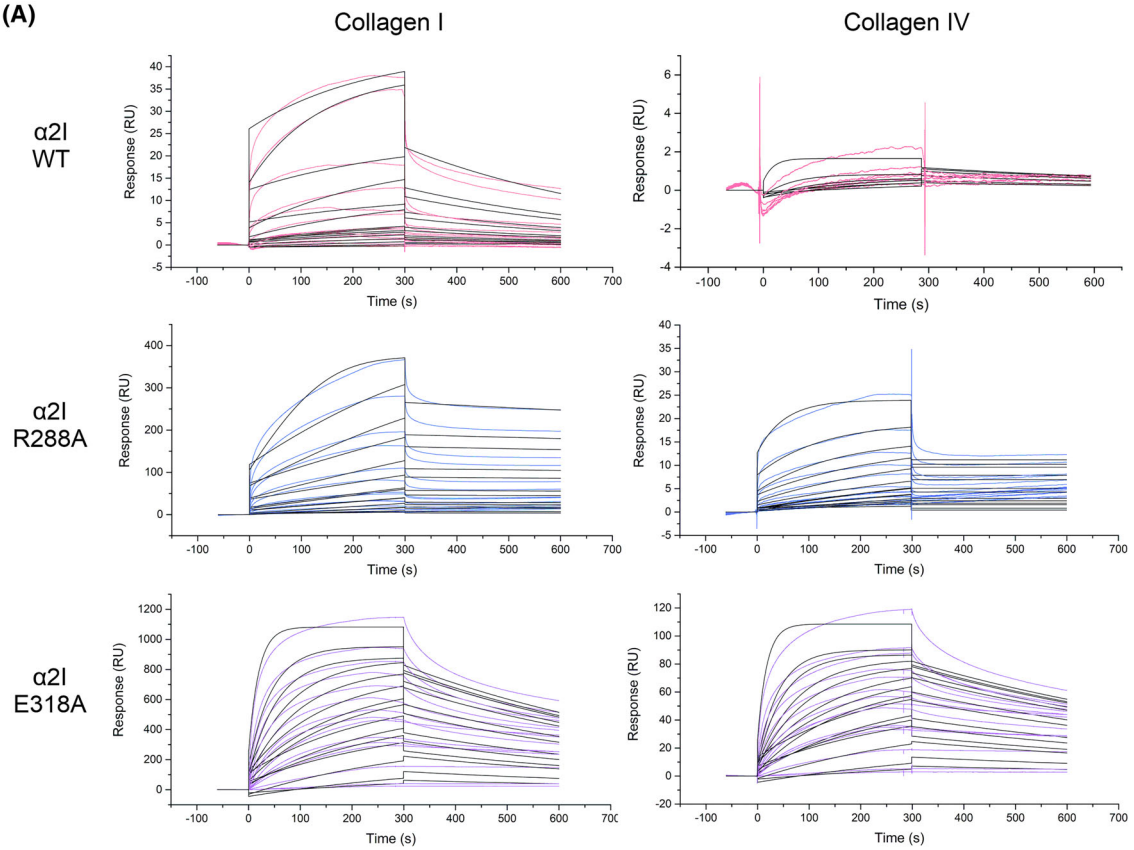
Introduction

Integrin α2β1 is a widely expressed member of the collagen receptor subfamily of the integrins. This group, comprising of four members, is unique to vertebrates and is characterized by the presence of the αI domain (also known as the αA domain) containing an αC helix, a structural element not found in other metazoan integrins [1–4]. The αC helix is therefore presumed to play an important role in high avidity integrin–collagen interactions [3–6].

Integrin α2β1, together with α1β1, has been described as a 'generalist' collagen receptor due to its involvement in a wide range of physiological processes,

Abbreviations
Bindingmax, maximal binding; ka, association rate constant; kd, dissociation rate constant; K_D, equilibrium dissociation constant; $K_{D,app}$, apparent value of K_D; MIDAS, metal ion dependent adhesion site; R_max, maximal binding; RU, resonance units; SPBA, solid-phase binding assay; SPR, surface plasmon resonance; WT, Wild-type; α2I, α2I domain; α2IE318A^{Closed}, PDB ID: 9RHY; α2IE318A^{Open_m}, PDB ID: 9RIE; α2IE318A^{Open_s}, PDB ID: 9RI7; α2IR288A^{Closed}, PDB ID: 9RK7; α2IWT^{Closed}, PDB ID: 1AOX; α2IWT^{Open}, PDB ID: 1DZI; αI, αI domain.

(A)

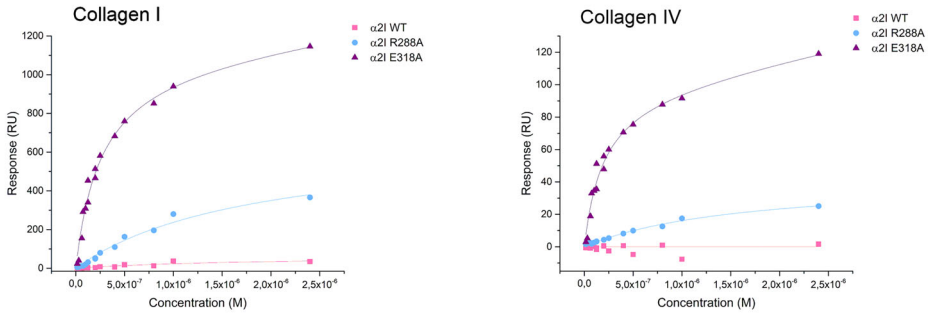


(B)

SPR kinetics analysis of α2I-domain variants binding to Collagen

α2I domain	Ligand	KD (μM)	Rmax (RU)	ka (1/Ms)	kd (1/s)	X^2
α2I WT	Collagen I	0.8	32	2725	0.002	1
α2I R288A	Collagen I	0.09	310	3404	0.0003	79.9
α2I E318A	Collagen I	0.08	859	20690	0.002	1740
α2I WT	Collagen IV	0.06	1	21480	0.001	0.1
α2I R288A	Collagen IV	0.00003	11	8175	0.0000002	1
α2I E318A	Collagen IV	0.06	84	23050	0.001	24

(C)



(D)

SPR affinity analysis of α2I-domain variants binding to Collagen

α2I domain	Ligand	KD1 (μM)	Rmax1 (RU)	KD2 (μM)	Rmax2 (RU)	X^2
α2I WT	Collagen I	1,9	37,4	1,9	30,6	30,6
α2I R288A	Collagen I	1,9	341,6	1,8	326,7	356,0
α2I E318A	Collagen I	0,3	1090	$3{,}8\times10^{5}$	$2{,}5\times10^{7}$	$2{,}2\times10^{3}$
α2I WT	Collagen IV	N/A	N/A	N/A	N/A	N/A
α2I R288A	Collagen IV	1,7	21,5	1,6	20,8	0,8
α2I E318A	Collagen IV	0,2	97,3	$3{,}4\times10^{5}$	4×10^{6}	31,1

Fig. 1. Pre-activated α2IR288A and α2IE318A variants increase avidity to both collagen I and IV in surface plasmon resonance (SPR) analysis. (A) Binding curves for α2IWT, α2IR288A, and α2IE318A binding to collagen I (left) and IV (right). Responses were fitted to a 1 : 1 interaction model. (B) Kinetic values gained with SPR analysis for α2I variants binding to collagen. (C) Biacore affinity analysis performed with the steady state two-site interaction model for collagen I (left) and collagen IV (right). (D) Affinity values gained with SPR affinity analysis for α2I variants binding to collagen. Measurements were performed using 6 different concentrations of each α2I variant, diluted into HBS buffer (pH 7.4) at 10 °C.

albeit without a critically important role [7,8]. They are not essential for organogenesis or survival, as evidenced by the relatively mild phenotypes of α1 and α2 knockout mice [9–11]. Integrin α2β1 is highly expressed on epithelial cells [7,12] and serves as a collagen receptor on platelets [13]. Furthermore, α2β1 is expressed on various immune cells, including T-lymphocytes, macrophages, and neutrophils, where it contributes to immune-related functions [7,12]. In inflammatory cells, α2β1 is classified as one of the 'very late activation antigens' (namely VLA-2), as its surface expression becomes detectable only days after cellular activation [14]. Inhibitors targeting α2 have shown promise as potential therapeutics for thrombotic disorders [15,16], while α2β1-specific antibodies have been shown to inhibit angiogenesis [17]. Together, these diverse roles make α2β1 an attractive therapeutic target in cancer and inflammation.

The activity of integrin α2β1 can be regulated through two primary mechanisms. Like all integrins, α2β1 can be pre-activated by intracellular signaling pathways promoting the binding of proteins, such as talin, to the intracellular tail of the β subunit. The binding of intracellular ligands induces a conformational change within the integrin increasing its affinity to extracellular ligands, a mechanism referred to as 'inside-out signaling' [18]. Outside the cell, the α2I domain (α2I) of α2β1 can bind to specific triple-helical motifs of collagen, such as GFOGER (where O represents hydroxyproline) or other sites containing the consensus sequence GXX'GER (where X and X' can be any amino acid) [19,20]. Upon this interaction, the α2I domain undergoes a conformational change triggering downstream intracellular signaling pathways [21], a process known as 'outside-in signaling' [18].

The extracellular α2I domain adopts a Rossmann-like fold characterized by a central parallel β-sheet flanked by six α-helices [22]. Collagen binding occurs at the metal ion-dependent adhesion site (MIDAS), but only after the domain undergoes substantial conformational rearrangements involving the MIDAS, the αC helix, and the α7 helix. Structural studies have shown that binding a collagenous GFOGER peptide disrupts a key ion pair between Arg288 in the αC helix and Glu318 in the α7 helix [21], inducing subsequent conformational changes and facilitating signal transmission to the I domain of β1 [18,21,23]. Although the detailed function of the αC helix during α2I activation remains unclear, current knowledge suggests it could have the potency to participate in pre-activation, ligand recognition, and ligand-induced activation of α2β1.

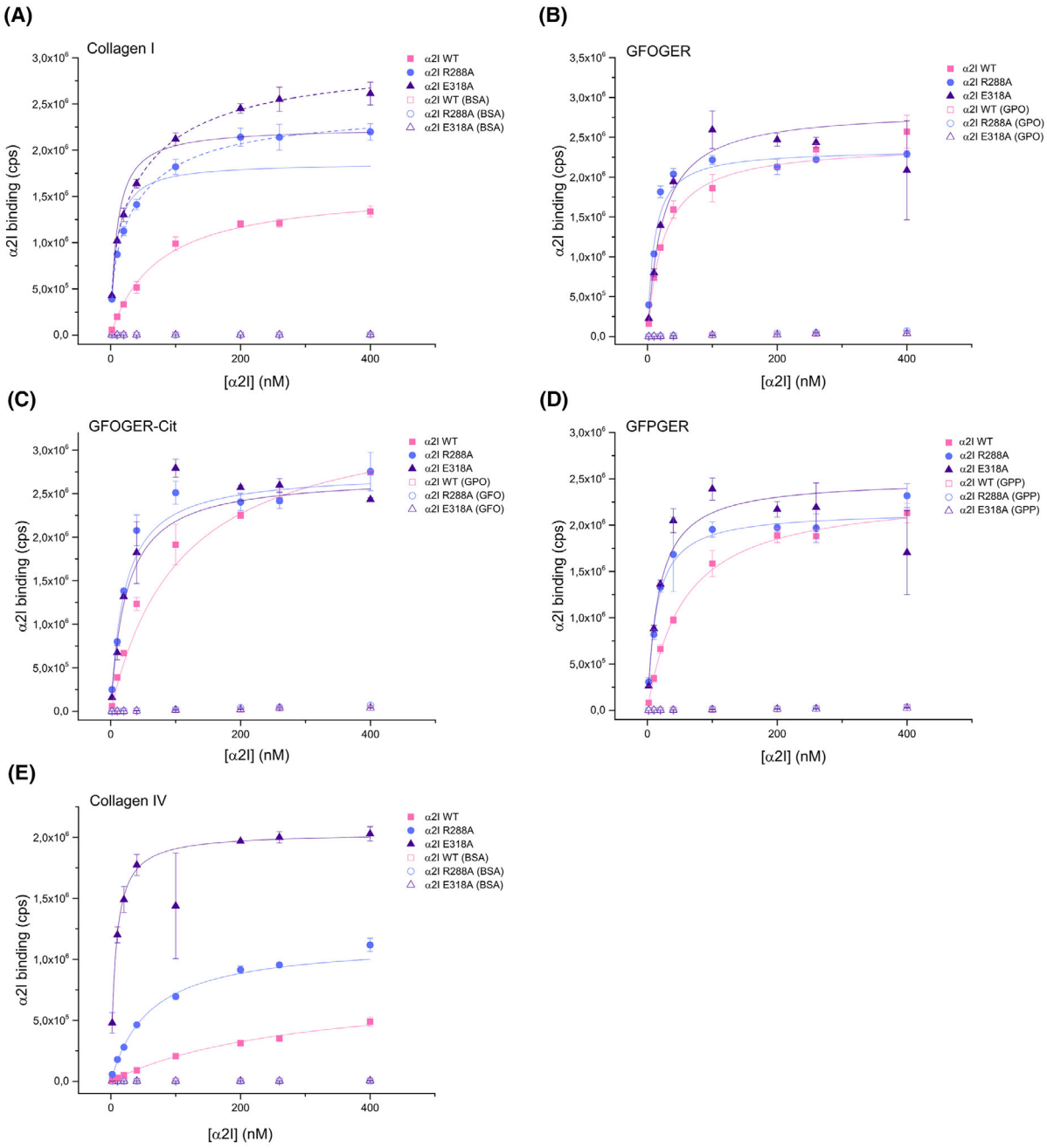
In this study, we investigate α2I domain activation and examine the mechanistic significance of the collagen receptor specific αC helix. We explore the specific role of the conserved Arg288-Glu318 ion pair in α2I activation using two α2I variants: α2IR288A and α2IE318A. Kinetic analyses, together with four novel crystal structures that capture previously unobserved transitional conformations, reveal that Arg288 and Glu318 play distinct roles in the dynamic activation mechanism of the α2I domain. Consistent with the concept of integrin outside-in signaling, our findings suggest a possible activation model in which ligand binding may precede complete α2I activation. Furthermore, the results highlight the evolutionarily unique αC helix as a mechanistic regulator of α2I activation, modulating both collagen type specificity and collagen-binding avidity.

Results

Disruption of the Arg288-Glu318 ion pair enhances α2I avidity toward collagens I and IV

The functional relevance of the ion pair between Arg288 and Glu318 in the α2I was investigated using point mutations designed to disrupt this ionic interaction. Previous studies have shown that breaking the ion pair by mutations of the glutamate residue (α2IE318W, α1IE317A, α2IE318A) can activate the αI domain and increase collagen binding [24–26]. To further probe the contribution of this ion pair to α2I activation, we generated the α2IR288A variant in parallel with the α2IE318A variant and assessed their binding to collagens I and IV using surface plasmon resonance (SPR) and solid-phase binding assays (SPBA).

The SPR binding assays were conducted using immobilized collagens I and IV and varying



(F)

Kinetic analysis of a2I domain variants (SPBA)

	a2I WT		a2I R288A		a2I E318A	
LIGAND	$K_{D,app}$ (nM)	$Binding_{max}$ (cps)	$K_{D,app}$ (nM)	$Binding_{max}$ (cps)	$K_{D,app}$ (nM)	$Binding_{max}$ (cps)
Collagen I	73 ± 6	$1.6x10^6 \pm 0.07x10^6$	9 ± 2	$1.9x10^6 \pm 0.2x10^6$	10 ± 2	$2.2x10^6 \pm 0.2x10^6$
Collagen IV	289 ± 49	$0.8x10^6 \pm 0.1x10^6$	60 ± 9	$1.1x10^6 \pm 0.1x10^6$	7 ± 0,5	$2.0x10^6 \pm 0.007x10^6$
GFOGER	24 ± 1	$2.4x10^6 \pm 0.06x10^6$	10 ± 1	$2.3x10^6 \pm 0.06x10^6$	23 ± 1	$2.9x10^6 \pm 0.1x10^6$
GFOGER-Cit	107 ± 10	$3.5x10^6 \pm 0.1x10^6$	21 ± 2	$2.8x10^6 \pm 0.1x10^6$	24 ± 4	$2.7x10^6 \pm 0.1x10^6$
GFPGER	56 ± 3	$2.4x10^6 \pm 0.05x10^6$	14 ± 1	$2.2x10^6 \pm 0.06x10^6$	17 ± 2	$2.5x10^6 \pm 0.1x10^6$

Fig. 2. Pre-activated variants α2IR288A and α2IE318A show enhanced avidity for collagen I and IV relative to α2IWT in solid-phase binding assay (SPBA) analysis. SPBA results of the pre-activated variants α2IR288A and α2IE318A, and α2IWT binding to: (A) collagen I; (B) GFOGER peptide; (C) citrullinated GFOGER; (D) GFPGER; and (E) collagen IV. (F) Table of kinetic values $K_{D,app}$ and $Binding_{max}$ for all measured ligands. Estimates of the dissociation constants were calculated using the equation: binding = maximal binding/(1 + K_D/[α2I]). All measurements were done at room temperature using DELFIA® Assay Buffer and 2 mM $MgCl_2$ as sample buffer. Results are representative of three independent biological replicates. Error bars represent the standard deviation of three technical replicates.

concentrations of the α2I variants (Fig. 1A,B). Binding kinetics were analyzed using a 1 : 1 binding model to derive the association rate constant (ka), dissociation rate constant (kd), and equilibrium dissociation constant (K_D = kd/ka). Both pre-activated α2IR288A and α2IE318A variants showed approximately 10-fold higher avidity for collagen I in comparison to the wild-type α2I (α2IWT) (K_D ~ 0.08–0.09 μM vs. 0.8 μM; Fig. 1B). More specifically, α2IE318A showed a markedly higher association constant toward collagen I, indicative of an extremely rapid association rate. Similar association was also seen for collagen IV suggesting that association is not selective for specific ligands for the α2IE318A variant.

Consistent with earlier findings [27], α2IWT exhibited low avidity for collagen IV. Interestingly, α2IR288A showed markedly enhanced binding to collagen IV, with a K_D of 0.03 nM, which is over 2000-fold stronger than that of α2IWT or α2IE318A (60 nM; Fig. 1A,B). This high avidity was primarily due to a dramatically low kd, indicating extremely slow dissociation of α2IR288A from collagen IV. A similar slow dissociation was also observed in α2IR288A binding to collagen I.

However, the pre-activated α2I variant binding curves deviated much from the applied 1 : 1 binding mode which displayed poor χ^2 values. To further inspect kinetics results, an additional two-site steady state affinity analysis was performed on the gained SPR data (Fig. 1C). This analysis conformed well with the binding data and confirmed both α2IR288A and α2IE318A variants to exhibit equal or increased binding affinity for collagen I, with the α2IE318A variant showing the strongest interaction (K_D1 0.3 μM vs. ~ 1.9 μM; Fig. 1D). α2IWT did not bind collagen IV; however, both pre-activated variants, especially the α2IE318A, showed increased affinity.

The solid-phase binding (SPBA) experiments supported the SPR findings (Fig. 2). Both α2IR288A and α2IE318A exhibited eight- and sevenfold increased binding to collagen I in comparison to α2IWT, respectively (Fig. 2A). For collagen IV, α2IE318A showed a 41-fold increase, while α2IR288A showed a fivefold increase in binding (Fig. 2E).

In summary, both mutations disrupting the Arg288-Glu318 ion pair significantly enhanced the avidity of the α2I domain for collagen I and IV but clear kinetic differences between the two variants were observed.

Pre-activated α2I variants exhibit changes in ligand selectivity and altered binding kinetics

SPBA analysis (Fig. 2) revealed both α2IR288A and α2IE318A variants to exhibit significantly higher maximal binding values in comparison to α2IWT. For collagen I, the $Binding_{max}$ values were 1.6×10^6 cps for α2IWT, and 1.9×10^6 cps and 2.2×10^6 cps for α2IR288A and α2IE318A, respectively (Fig. 2A,F). However, binding data for collagen I did not conform well to a single-site 1 : 1 binding model. The two-affinity binding model described the data best: binding = maximal $binding_1$/(1 + K_{D1app}/[αI]) + maximal $binding_2$/(1 + K_{D2app}/[αI]) (Fig. 2A). This model reflected more accurately the complex interaction of the pre-activated α2I variants, indicative of them binding to multiple binding sites on collagen I and thus resulting in exceptionally high maximal binding.

These findings were consistent with the conducted SPR kinetics analysis. For collagen I, the R_max for α2IWT was 32 RU, whereas α2IR288A and α2IE318A reached 10-fold and 27-fold higher values, respectively (Fig. 1A,B). A similar pattern was observed for collagen IV: the R_max for α2IWT was minimal, while the R_max for α2IR288A and α2IE318A increased by 11- and 84-fold, respectively (Fig. 1A,B). However, due to the unconformity of the kinetics data, the two-site interaction model was used to further inspect the binding mode of the α2I variants. This model gave a similar fit to the SPBA results where α2I variants were depicted to bind multiple binding sites on collagen I, as further supported by the high R_max values (Fig. 1C). In contrast, binding to collagen IV produced lower R_max and χ^2 values in both SPR analyses for all measured α2I variants, suggesting a simpler binding interaction to less available binding sites on collagen IV (Fig. 1C).

To further probe the binding preferences of the α2I variants, we tested their binding to triple-helical

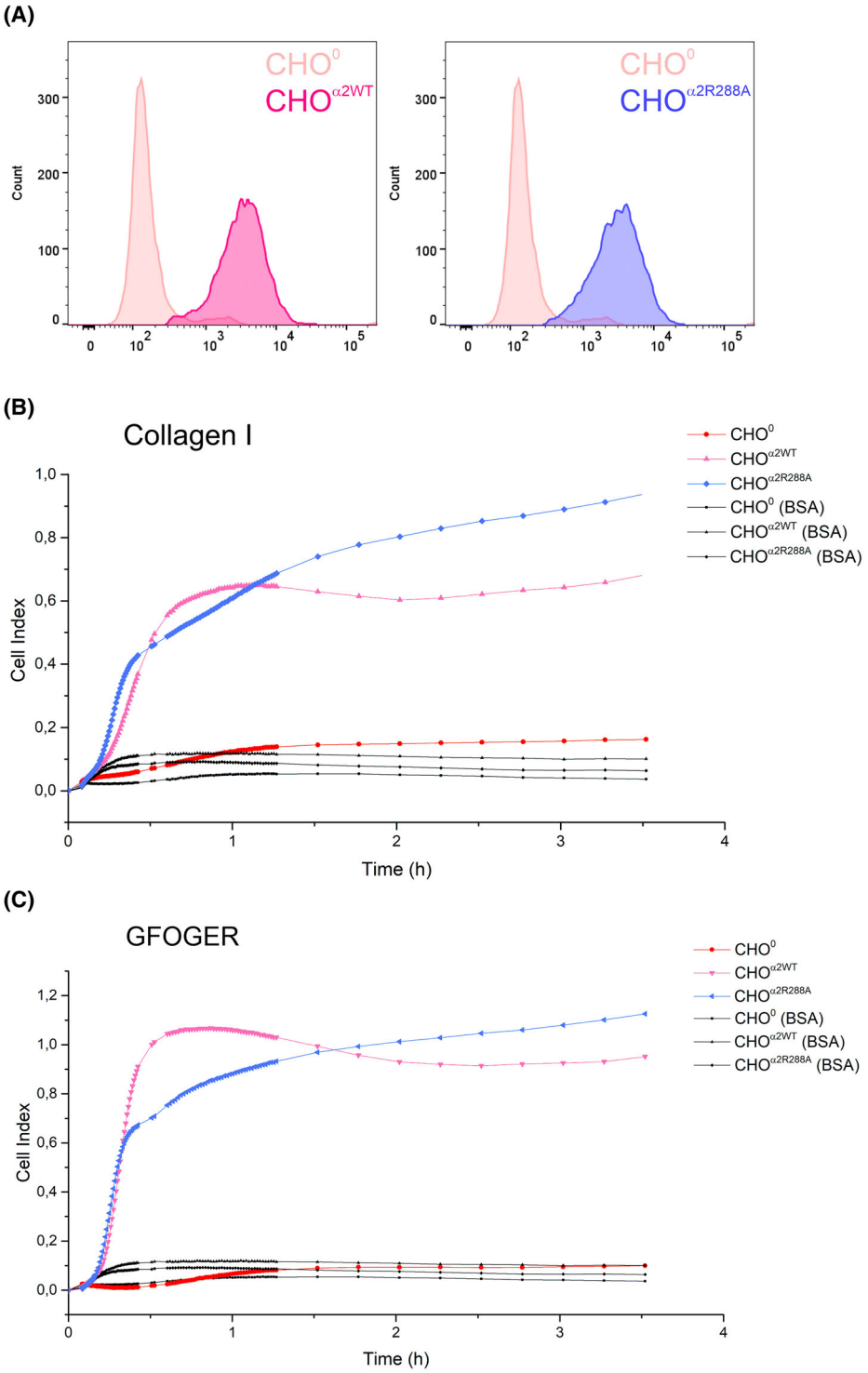


Fig. 3. Effect of R288A mutation on integrin α2–dependent adhesion to collagen I and GFOGER. (A) Flow cytometry analysis (FACS) of α2 transfected CHO cell lines. Cells were sequentially gated to exclude debris, aggregates, and dead cells, followed by selection of α2-positive cells. (B, C) Adhesion of CHO^{0} (red, negative control), $CHO^{\alpha 2WT}$ (pink) and $CHO^{\alpha 2R288A}$ (blue) cells on collagen I and GFOGER. Cells were monitored in real-time with impedance-based xCELLigence technology, where adhesion is expressed as Cell Index. Data shown is the mean value of three technical replicates where BSA was used as negative control. The experiment was performed twice with similar results.

peptides (Fig. 2B–D). GFOGER represents a single high-affinity motif, while collagen I contains at least three high-affinity α2β1-binding sites, most probably along with multiple low-affinity regions [28]. Variants only showed a moderate increase in binding to peptide ligands (Fig. 2B–D,F). α2IR288A bound 2.4- to 5-fold tighter than α2IWT to GFOGER, GFPGER, and GFOGER-Cit. α2IE318A showed slightly improved binding to GFPGER and GFOGER-Cit, but no significant change in GFOGER binding compared to α2IWT (Fig. 2B). Peptides GFPGER and GFOGER-Cit were included in the experiments to test the putative roles of hydroxyproline and arginine residues, respectively. However, they did not significantly differ from the GFOGER peptide.

To conclude, α2IR288A and α2IE318A enhanced maximal binding to both collagen I and IV but, however, not to individual binding motifs. Both variants showed binding to more than one binding site in collagens, revealing separate interactions with different affinities.

The loss of the Arg288-Glu318 ion pair in α2 integrin modifies cell adhesion to collagen I

At cellular level, the functional role of the Arg288-Glu318 ion pair in α2 integrin was investigated by adhesion analyses utilizing real-time impedance-based xCELLigence technology. CHO cell lines expressing either wild-type α2 ($CHO^{\alpha2WT}$) or the α2R288A variant (CHO^{R288A}) were generated, while nontransfected CHO cells (CHO^{0}) were used as a negative control. Before adhesion assays, α2 integrin expression levels were assessed by flow cytometry (Fig. 3A). The results confirmed similar expression levels of α2 wild-type and α2R288A variant (mean fluorescence intensities 3370 and 3077, respectively).

In xCELLigence experiments, cell adhesion to both collagen I and the triple-helical collagenous GFOGER peptide was measured (Fig. 3B,C). CHO^{R288A} showed increased adhesion to collagen I when compared to $CHO^{\alpha2WT}$ (Fig. 3B). Interestingly, the difference was most pronounced after the 2 h time point, indicating that the R228A mutation activated adhesion and adhesion-related signaling events after the initial attachment phase. The differences were less pronounced when adhesion to GFOGER was measured (Fig. 3C). The results are consistent with performed α2I domain binding assays and confirm the activated state of the variant receptor at the cellular level. It has previously been reported that the E318W mutation in the α2 chain of α2β1 integrin also enhances cellular adhesion to collagen [24,27,29].

Overall structures of pre-activated α2IR288A and α2IE318A variants

To understand the different binding mechanisms of the pre-activated α2I variants, several crystal structures of the α2IR288A and α2IE318A variants were solved (For structure determination statistics, see Table S1 and the PDB validation reports). In general, the calculated electron density maps were well interpretable (Fig. S1) and at least 92% of the residues of all structures were in favored regions of their Ramachandran plots.

The solved structures displayed a Rossmann-like fold of secondary elements typical for the αI domains but revealed two new transitional conformations for α2I. The new structures were named based on their similarity to the known closed ($\alpha2IWT^{Closed}$; Fig. 4, orange; PDB ID: 1AOX; [22]) and open ($\alpha2IWT^{Open}$; Fig. 4, gray; PDB ID: 1DZI; [21]) conformations of α2I as follows: $\alpha2IR288A^{Closed}$, the pre-activated α2IR288A variant, (Fig. 5A, raspberry; PDB ID: 9RK7); $\alpha2IE318A^{Closed}$, α2IE318A variant without a ligand at MIDAS, (Fig. 5A, violet; PDB ID: 9RHY); $\alpha2IE318A^{Open\_s}$, α2IE318A variant with succinic acid as ligand (Fig. 5B, cyan; PDB ID: 9RI7); and $\alpha2IE318A^{Open\_m}$ variant with malic acid ligand (Fig. 5B, sky blue; PDB ID: 9RIE) (Table S1). The $\alpha2IE318A^{Open\_s}$ and $\alpha2IE318A^{Open\_m}$ structures co-crystallized with carboxylic group ligands binding directly to the Mg^{2+} ion, unlike $\alpha2IE318A^{Closed}$, which did not have any ligand at MIDAS despite the high malonate concentration in the crystallization setup.

During α2I activation, the ion pair interaction between Arg288 and Glu318 (Fig. 4) is lost and the αC helix unwinds away from the MIDAS region extending the α6 helix. As a result, Glu318 is detached from the αC helix triggering a rearrangement of the C-terminal α7 helix, which is shifted significantly (12 Å) downwards. This movement further opens the MIDAS region for ligand binding ($\alpha2IWT^{Open}$; gray, Fig. 4A–C) and allows signal transmission to the βI domain via the α7 helix [18,21]. In the $\alpha2IR288A^{Closed}$ and $\alpha2IE318A^{Closed}$ structures, both the αC and α7 helix are intact (Fig. 5A) and superimpose well with the wild-type $\alpha2I^{Closed}$ (PDB: 1AOX) with an RMSD value of 0.4 Å (for details, see Materials and Methods).

Like $\alpha2IWT^{Open}$, the $\alpha2IE318A^{Open}$ structures have an unstructured αC helix region and a downwards shifted α7 helix (11 Å, Fig. 5B) like in $\alpha2IWT^{Open}$ (Fig. 4C) even though the shift is smaller due to missing water-mediated Glu318 contacts to the α1 helix

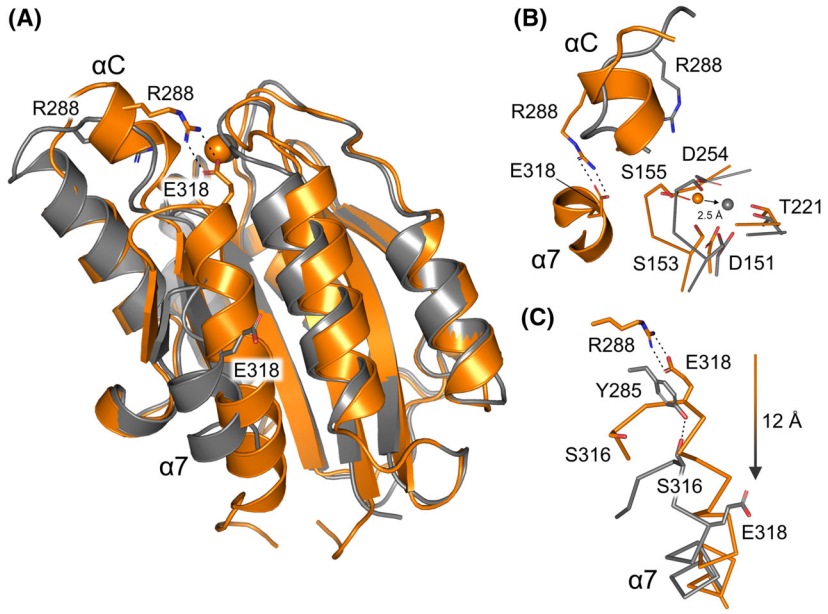


Fig. 4. Known open (α2IWT^{Open}; gray; PDB ID: 1DZI) and closed (α2IWT^{Closed}; orange; PDB ID: 1AOX, chain A) conformations of the α2IWT (Emsley *et al.*, [21,22]). Key amino acid residues (sticks) and helices are labeled, and the catalytic metal ion is shown as a sphere. (A) α2IWT open and closed structures superimposed on each other. (B) Zoomed view of the ionic bond (ion pair; dotted lines) between Arg288 and Glu318 and the metal ion dependent adhesion site (MIDAS) amino acids in α2IWT. In α2I^{Open} the αC helix is unwound and Arg288 has shifted away from Glu318 which is not visible here due to the large conformational shift of the α7 helix. A black arrow pinpoints the 2.5 Å shift in the position of the metal ion. (C) Zoomed view of the conformational change of the α7 helix. In α2IWT^{Open} the α7 helix adapts an alpha helical turn allowing a hydrogen bond (dotted line) to form between Ser316 and Tyr285. The 12 Å movement of the functionally critical Glu318 is indicated with an arrow. Protein structures were visualized using PyMOL (Schrödinger, LLC).

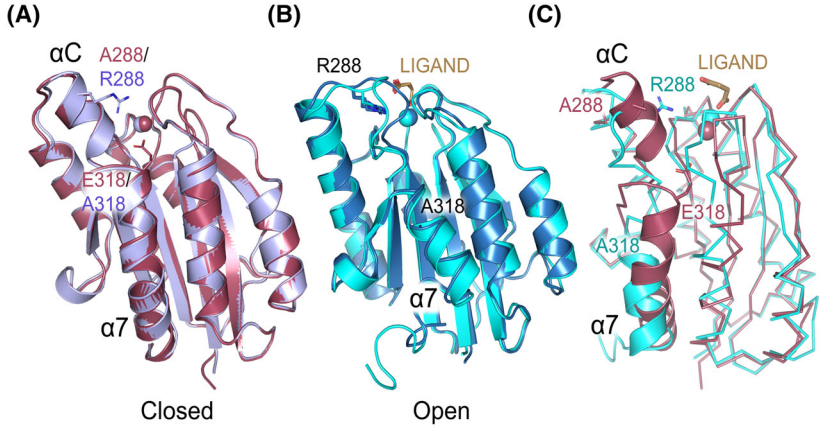


Fig. 5. Overall structures of the pre-activated α2IR288A and α2IE318A variants. Key amino acid residues (sticks) and helices are labeled, and the catalytic Mg^{2+} ion is shown as a sphere. (A) Superimposed α2IR288A^{Closed} (raspberry, PDB ID: 9RK7) and α2IE318A^{Closed} (violet, PDB ID: 9RHY) variants. (B) Superimposed α2IE318A^{Open_s} (cyan, PDB ID: 9RI7) and α2IE318A^{Open_m} (sky blue, PDB ID: 9RIE) variants with a bound succinic acid or malic acid ligand, respectively; the ligands are shown as sticks (sand). (C) Superimposed α2IR288A^{Closed} (raspberry, PDB ID: 9RK7) and α2E318A^{Open_s} (cyan, PDB ID: 9RI7) pre-activated variants. Mobile helices are shown using cartoon and other regions of the structure using ribbon representation for clarity. Succinic acid of α2IE318A^{Open} is shown as sticks (sand). Protein structures were visualized using PyMOL (Schrödinger, LLC).

(Ala161, Asn164, Lys168). When comparing to the α2I^{Open} structure (PDB:1DZI), the superimposition of the α2IE318A^{Open_s} and α2IE318A^{Open_m} structures resulted in RMSD values of 0.5 Å and 0.3 Å (for details, see Materials and Methods), respectively. The pre-activated α2IR288A^{Closed} and α2IE318A^{Open_s}

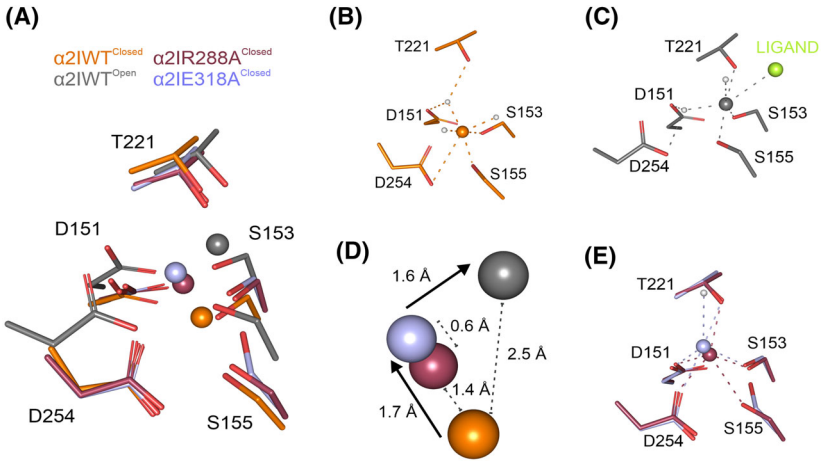


Fig. 6. Conformational changes of the metal ion and its coordinating amino acid residues in the metal ion dependent adhesion site (MIDAS) during α2I activation. In all figures structures are superimposed and the metal ions are shown as large spheres and conserved water molecules as small white spheres. Ionic and hydrogen bonds are shown as dotted lines. (A), Superimposition of α2IWT^{Closed} (orange), α2IWT^{Open} (gray), α2IR288A^{Closed} (raspberry), and α2IE318A^{Closed} (violet) MIDAS. (B), α2IWT^{Closed} MIDAS. The metal ion is hexacoordinated binding to Asp254, Ser155, and Ser153, and three conserved water molecules. (C), α2IWT^{Open} MIDAS. The metal is hexacoordinated binding to Ser155, Ser153, Thr221, two conserved water molecules, and a GFOGER peptide (green sphere). (D), Movement of the α2I metal ion during α2I activation. Distances between metal ions labeled and black arrows indicate movement of the metal ion from closed to open conformation. (E), Transitional conformations of α2IR288A^{Closed} and α2IE318A^{Closed} MIDAS. In the α2IR288A^{Closed} structure, Mg^{2+} is shifted and pentacoordinated to all MIDAS amino acids: Asp254, Ser155, Ser153, Thr221, and Asp151. In the α2IE318A^{Closed} structure, Mg^{2+} is hexacoordinated to all five MIDAS amino acids and to one conserved water molecule. Protein structures were visualized using PyMOL (Schrödinger, LLC).

variants (Fig. 5C) respectively resemble the α2IWT^{Closed} and α2IWT^{Open} WT structures (Fig. 4A).

MIDAS metal coordination of pre-activated α2IR288A domain

Previous studies on the integrin α2I activation have depicted a change in the α2I MIDAS configuration due to ligand binding (Figs 4B, 6A) [21]. In the closed conformation of the α2I (α2IWT^{Closed}; PDB ID: 1AOX), MIDAS is embedded by three surrounding loops and the metal ion is hexacoordinated by three conserved water molecules and side chains of Ser155, Ser153, and Asp254 (Fig. 6B). In the collagen-bound open conformation of the α2I (α2IWT^{Open}; PDB ID: 1DZI), MIDAS is exposed, and the metal ion is hexacoordinated and now binds the negatively charged side chain of the glutamate (E11) from the collagen-peptide in addition to two conserved water molecules and side chains of Ser155, Ser153, and Thr221 (Fig. 6C). Moreover, in both α2IWT^{Open} and α2IWT^{Closed} forms, water-mediated interactions between the metal ion and Asp151 stabilize the coordination. Upon this coordination change due to ligand binding, the metal ion shifts 2.5 Å toward the ligand (Figs 4B and 6A,D) [21,22,30,31].

Here we report that both α2IR288A^{Closed} and α2IE318A^{Closed} pre-activated structures portray a novel metal ion coordination when compared to α2IWT (Fig. 6E). At MIDAS of the solved α2IR288A^{Closed} structure, the metal ion is pentacoordinated and Asp151, Ser153, Ser155, Thr221, and Asp254 directly stabilize the metal, which is shifted 1.4 Å in comparison to the metal ion of α2IWT^{Closed} (Fig. 6D,E). In addition to the same interactions seen in α2IR288A^{Closed}, the MIDAS metal ion of the α2IE318A^{Closed} on the contrary is hexacoordinated since it binds to a water molecule (Fig. 6E, violet). Due to this interaction, the metal ion of α2IE318A^{Closed} is shifted 0.5 Å in comparison to the metal of α2IR288A^{Closed} and 1.6 Å in comparison to the metal of α2IWT^{Closed}. The MIDAS configurations of all the open form α2I structures determined in this study closely resemble that of the previously reported α2IWT^{Open} (PDB: 1DZI) and are therefore not described further.

Interaction between Arg288 and Ser155 is crucial for MIDAS metal stabilization

Despite the very similar overall structure of the α2IR288A^{Closed} with the intact αC helix (Fig. 5A,

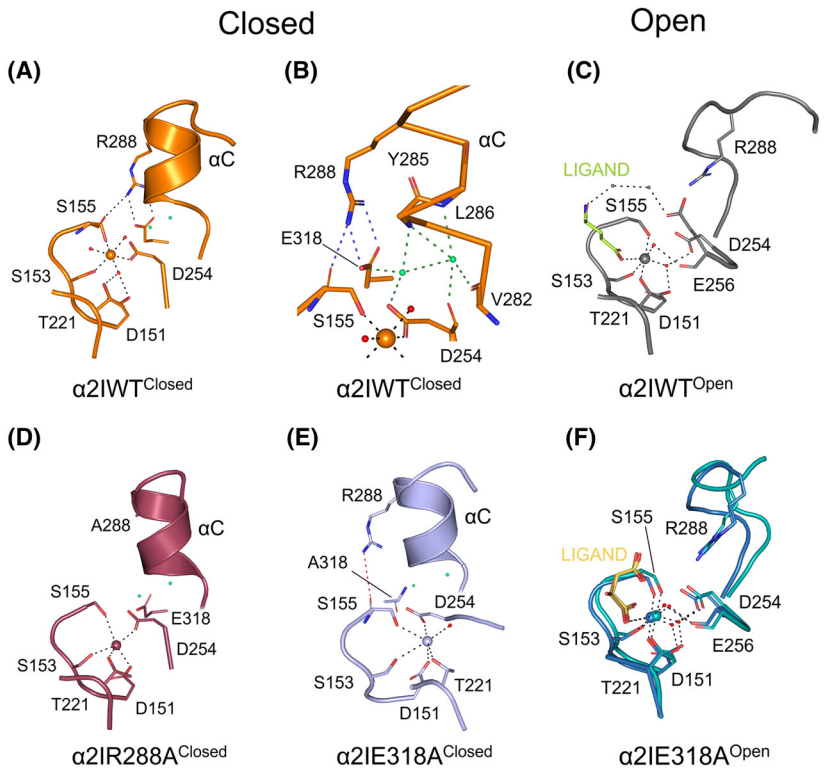


Fig. 7. Comparison of α2I metal ion dependent adhesion site (MIDAS) conformations. Metal ion (large spheres) coordination and position of the αC helix is shown. Conserved water molecules (small spheres) at MIDAS are shown in red and next to the αC helix in green. (A), $\alpha2IWT^{Closed}$ MIDAS (PDB ID: 1AOX) showing a hexacoordinated metal ion and an intact αC helix with conserved waters. (B), Zoomed view of the αC helix of $\alpha2IWT^{Closed}$ in ribbon representation showing Arg288 contacts (blue dotted lines) and Glu318 contacts (green dotted lines). (C), $\alpha2IWT^{Open}$ MIDAS (PDB ID: 1DZI) showing a hexacoordinated metal ion in contact with GFOGER peptide (green) and an unwound αC helix. The metal ion has shifted 2.5 Å and Glu318 is not visible near MIDAS due to movement of the α7 helix. (D), $\alpha2IR288A^{Closed}$ MIDAS (PDB ID: 9RK7) showing a pentacoordinated metal ion and an intact αC helix with conserved waters. The metal ion has shifted 1.4 Å (in comparison to $\alpha2IWT^{Closed}$). (E), $\alpha2IE318A^{Closed}$ MIDAS (PDB ID: 9RHY) showing a hexacoordinated metal ion and an intact αC helix near the E318A mutation. A 1.6 Å shift in metal ion is seen and the Arg288 and Ser155 contact (red dotted line) is missing. (F), Two superimposed $\alpha2IE318A^{Open}$ MIDAS structures with either a bound succinic acid (PDB ID: 9RI7; cyan) or malic acid (PDB ID: 9RIE, blue); the ligands are shown in sand. The metal ion of both structures is hexacoordinated and the αC helix has unwound. E318A is no longer located near the MIDAS and the metal ion has shifted 2.6–2.7 Å, respectively. Protein structures were visualized using PyMOL (Schrödinger, LLC).

raspberry) to that of the $\alpha2IWT^{Closed}$ (Fig. 4A, orange), its metal ion has shifted 1.4 Å toward the open conformation of α2I (Fig. 6D) indicating that the side chain of Arg288 is important for the stabilization of the metal in the closed conformation. Closer inspection of the α2IWT MIDAS revealed that a side chain nitrogen atom of Arg288 is hydrogen bonded to the carboxyl main chain oxygen atom of Ser155, which, in turn, coordinates the metal ion in both the closed and open conformations of the α2IWT (Fig. 7A,C). When Arg288 is mutated to alanine in the $\alpha2IR288A^{Closed}$ variant (Fig. 7D), the interaction between Arg288 and Ser155 is lost and the 1.4 Å shift of the metal ion is seen (Figs 6D and 7D). Losing the contact between Arg288 and Ser155 allows the metal ion to move more freely and approach a negatively charged ligand (e.g., succinic or malic acid or a glutamate residue in collagen) and alter α2I conformation toward the open state.

Similarly, the $\alpha2IE318A^{Closed}$ (Fig. 7E) structure also depicts a shift in the position of the metal ion. The movement results from the E318A mutation, which eliminates the ion pair between Arg288 and Glu318, thus allowing the side chain of Arg288 to move. This movement disrupts the hydrogen bond between Arg288 and Ser155, allowing the metal ion to shift toward the ligand-bound state by 1.6 Å (Figs 6D and 7E). Our results suggest that this Arg288-Ser155 contact is vital for anchoring the metal ion in its closed conformation and once lost, a shift in the metal ion is seen. Binding studies demonstrate that the 1.4 Å

shift of metal enhances avidity to collagen ligands (Figs 1 and 2; see the increased avidity of the α2IR288A variant), yet the changes in MIDAS are insufficient to enable binding of low-affinity ligands as observed in the fully opened, α2IE318A^{Open} structures (Fig. 7F).

Glu318 is important for the stabilization of the αC helix

We have previously discovered that the αC helix (Glu284-Arg288) is needed for the α2I to bind collagen with high avidity and selectivity [6]. However, its specific function in the activation mechanism of the collagen receptor integrin αI domains has remained unknown. Both pre-activated, open α2IE318A structures lacking the Arg288-Glu318 ion pair (α2IE318A^{Open_s} and α2IE318A^{Open_m}) crystallized with an unstructured αC helix (Fig. 7F) bound to low-affinity ligands (succinic or malic acid), highlighting the role of the αC helix in ligand specificity as well as the importance of the C-terminal Glu318 in the stability of the αC helix.

Two water molecules are present in all studied α2I domain structures containing an intact αC helix and allow for water-mediated contacts between the αC helix (Tyr285, Leu286, Val282) and the MIDAS amino acid Asp254 (Fig. 7B, green spheres). In our pre-activated, closed α2I structures, either the main chain nitrogen atom of Ala318 (α2IE318A^{Closed}) or the side chain oxygen atom of Glu318 (α2IR288A^{Closed}) formed hydrogen bonds to one of the two structurally conserved, but previously undescribed, water molecules in the vicinity of the intact αC helix (Fig. 7D,E; green spheres; contacts identical to those in green in Fig. 7B). These contacts anchor the αC helix near the MIDAS in its helical conformation. When the ion pair interaction between Arg288 and Glu318 is lost during α2I domain activation by small acidic ligands in our pre-activated α2IE318A^{Open} variants (Fig. 7F), Glu318 has shifted and can no longer form stabilizing interactions with the αC helix and the MIDAS. This causes the αC helix to unwind and consequently the metal ion to shift toward its ligand-binding position.

Upon the unwinding of the αC helix, Tyr285 is known to form a hydrogen bond with Ser316 of the α7 helix leading to a completely open conformation of the αC helix as reported earlier by Emsley *et al.*, [21] (Fig. 4C). Once this conformation is reached, the restrictive elements governing α2I ligand binding and selectivity are lost, allowing the ligand to bind fully. This is demonstrated by our α2IE318A^{Open} variants (Fig. 5B, Fig. S2) where the unwound αC helix has allowed Tyr285 to bind Ser316, stabilizing the α7 helix in its open conformation and thus allowed for unrestricted binding of low-affinity ligands with a free carboxylic acid side chain.

Discussion

In this study, we have investigated the collagen receptor integrin-specific αC helix and its role in αI domain activation. Despite previous studies on the αI domains, the precise function of the αC helix has remained enigmatic. We have earlier developed small molecule inhibitors that are selective for either nonactivated or pre-activated α2I domains [15] and shown their different efficacy in *ex vivo* and *in vivo* models of thrombosis [15] and inflammation [32]. Thus, the detailed information about the structural basis of the activation mechanism of α2β1 integrin and understanding the role of the αC helix in the activation process is essential in developing new inhibitor molecules and putative drug candidates.

The rise of the αC helix of αI and the specific high-affinity recognition motifs on collagens (e.g. GFOGER) is reported to be the result of coevolution between collagen-binding integrins and their collagen ligands [33]. The αC helix containing, collagen-binding integrins are found from early chordates to primitive vertebrates, including the primitive Agnathostomes (jawless vertebrates) [26,34,35], indicating that the evolvement of the αC helix has been a critical step in the evolution of collagen receptors. We have previously performed functional analysis on the closest living relative of vertebrates, *Ciona intestinalis* (sea squirt), and shown that its αI does not contain an αC helix and, thus, cannot recognize the integrin binding motifs in fibrillar collagens [26] essential for the complexity of vertebrate tissue [36]. Indeed, human α2I lacks specificity in the absence of the αC helix, allowing both rapid binding and releasing of its collagen ligand [6].

To examine the function of the αC helix we created two pre-activated α2I variants by mutating the Arg288 and Glu318 residues, which form the mechanistically conserved ion pair that breaks prior to the αC helix unwinding upon ligand binding [21]. The four new crystal structures of the pre-activated α2I variants, along with vast kinetic analysis on their collagen binding, indicated the αC helix to be vital for the controlled regulation of α2I activation. Our kinetic studies with the pre-activated variants revealed reduced ligand specificity, which was explained by the α2I variant crystal structures in which the forced destruction of the stabilizing ion pair resulted in the unwinding of

the αC helix. However, structural and functional differences between the studied variants indicated that additional factors beyond ion pair disruption contributed to α2I activation, pointing to a more nuanced mechanism of activation, potentially initiated by ligand engagement.

The generally accepted model of integrin function stresses the role of integrin activation before high-affinity ligand binding and ligand-induced conformational changes leading finally to activation of cellular signaling pathways [18,37]. On blood platelets integrin α2β1 represents one of the three principal molecular machineries mediating platelet binding to collagen, the two others being glycoprotein VI (GPVI) and von Willebrand factor (VWF) together with its receptor GPIb [38]. It has been shown that VWF–GPIb interaction can activate α2β1, inducing a conformational change in the α2I [38]. Interestingly, recently published comprehensive analyses of α4β1 and α5β1 integrin function have found evidence that ligand binding may precede, rather than follow, activation by inside-out signaling [39]. Regulation of collagen receptor function is known to be controlled by the strength of blood flow-rate dependent shear stress [40], and under these conditions, selective small molecule inhibitors have been reported to bind to nonactivated α2β1 WT or variants suggesting that extracellular ligands can activate the α2I domain as well [15]. Thus, both cellular signals and ligand-induced activation of α2β1 may break the Arg288-Glu318 ion pair. Our results propose that αC helix and the Arg288-Glu318 ion pair regulate the number of potential binding sites recognized in collagen.

Surprisingly, the forced destruction of the Arg288-Glu318 ion pair by the R288A mutation did not open the α2I fold, but depicted a new transitional closed conformation without ligand, (α2IR288A^{Closed}; Fig. 5A), in which the pentacoordinated Mg^{2+} ion had shifted toward its collagen bound position (Fig. 6D,E). We suggest that α2IR288A^{Closed} structure mimics an early transitional state in the α2I domain activation following initial collagen recognition indicating that the early MIDAS rearrangements likely constitute the first step of α2I conformational activation. Previous studies have shown Asp219 and His258 in α2I to play a crucial role in collagen binding near the MIDAS [21,41,42]. These contacts could be responsible for the initial collagen recognition, which allows the negatively charged E11 of the GFOGER middle chain to come closer to the Mg^{2+} ion. The intermolecular contacts between E11 in collagen and the metal ion could cause the MIDAS to rearrange disrupting the bond between Arg288 and Ser155 as seen in the α2IR288A^{Closed}, α2IE318A^{Closed} structures (Fig. 7B,E) and consequently leading to the breakage of the Arg288-Glu318 ion pair. Thereafter the αC helix can unwind, the metal ion moves closer to E11 and Glu318 shifts down to stabilize the α7 helix in its open conformation (e.g., α2IE318A^{Open}; Figs 5B, 7F and Fig. S2A).

Results from our previous study indicate that breaking of the corresponding ion pair in integrin α1I domain also activates ligand binding but leads to different conformational changes [25] when compared to α2I domain structures in this paper. Based on the existing knowledge, it is not possible to speculate potential functional similarity or differences when α1I and α2I are compared to the two other collagen receptor integrins.

To conclude, our results illuminate initial α2I MIDAS rearrangements that promote unwinding of the αC helix, leading to the open, ligand-bound α2I conformation. We propose that outside-in activation can be initiated by introduction of an extracellular ligand, which is followed by MIDAS rearrangement, αC helix unwinding, and finally the downward displacement of the α7 helix, ultimately transmitting conformational changes to the β1 I domain. This suggested activation mechanism secures α2I to maintain its high ligand specificity and avidity in collagen binding, thus preventing redundant activation of the generalist α2β1 integrin and allowing it to participate dynamically in the many cellular functions it mediates.

Methods

Mutagenesis of human integrin α2I domain

The human α2IWT (^{139}PCP – IEG^{335}) carrying the point mutation C140S was previously cloned into the pGEX-2T GST vector (Cytiva Lifesciences, Uppsala, Sweden) [6]. Point mutations R288A and E318A were created using the QuikChange site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA), and all created constructs were sequenced to confirm success of the mutations.

Protein expression and purification

α2I domains for solid-phase binding assays were expressed and purified as GST fusion proteins in *Escherichia coli* BL21 as described previously [27,43]. For crystallization of α2IR288A and α2IE318A, GST was cleaved away by overnight thrombin digestion (GE Healthcare, Uppsala, Sweden) conducted at room temperature. About 100 NIH units of thrombin were directly added to the purified α2I and left to sit in a disposable GST-affinity chromatography column (Bio-Rad, Hercules, CA, USA). The purified α2I was eluted into Tris/HCl buffer (0.4 M, 2 mM $MgCl_2$) and lastly

concentrated with a centrifugal filter device (Merck Millipore, Darmstadt, Germany) to 13–16 mg·mL^{-1}. Purification of the cleaved α2I domains was assessed with electrophoresis using 8–25% gradient SDS/PAGE gels and the Phast System (Amersham Biosciences, Uppsala, Sweden).

Solid-phase binding assays

Solid-phase binding assays for α2IWT, α2IR288A, and α2IE318A were carried out as previously described [6]. 96-well plates (DELFIA® Microtitration Plate, PerkinElmer Inc., Wallac, Turku, Finland) were coated with PBS diluted collagens I and IV (5 μg·cm^{-2}/well) (Corning®, Corning, NY, USA) or with BSA (3.75%) (Diluent II, DELFIA®/AutoDELFIA®, PerkinElmer Inc., Wallac) or GFOGER peptides (1 μg·cm^{-2}/well) (AUSPEP, Tullamarine, VIC, Australia). Plates were incubated overnight in + 4C. The following day all wells were washed with PBS + 2 mM $MgCl_2$ and blocked with BSA (Diluent II : PBS, 1 : 1), left to incubate for 1 h in room temperature. After, a 10–2000 ng/100 μL concentration series of all studied α2I were formulated using DELFIA® Assay Buffer (PerkinElmer Inc., Wallac) and 2 mM $MgCl_2$ as sample buffer. The blocking solution was washed as previously mentioned, and α2I samples were added on wells and left to bind for 1 h. After, wells were washed three times, and DELFIA® europium-labeled anti-GST antibody (Eu-N1-anti-GST; 1 : 1000) (PerkinElmer Inc., Wallac) was added and left to incubate for 1 h. Lastly, the wells were washed three times, DELFIA® Enhancement solution was added (PerkinElmer Inc., Wallac), and the signals were measured with time-resolved fluorescence spectrophotometry (Victor3 multilabel counter, PerkinElmer Inc., Wallac). Estimates of dissociation constants for α2I were calculated using the equation: measured binding = maximal binding/(1 + Kd/[α2I]). All measurements included BSA as a negative control, and all points were replicated three times. Data were analyzed and graphs prepared using OriginLab 2016 (OriginLab Corporation, Northampton, MA, USA), and final figures were made with MS PowerPoint and CorelDRAW 2025.

Cell culture and transfection of CHO-α2 cells

Chinese hamster ovary (CHO^0) (CHO-K1; ATCC CCL-61; RRID:CVCL_0214) cells were transfected with the human α2 integrin expression construct (CHO$^{\alpha2WT}$) as described previously [44]. Additionally, the point mutation R288A was introduced into the human α2 (CHO$^{\alpha2R288A}$) using the HilyMax (Dojindo Laboratories, Kumamoto, Japan) transfection reagent. Cells were cultured in α-minimum essential medium (α-MEM; EuroClone, Milan, Italy) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Gibco, Thermo Fisher Scientific, Waltham, MA, USA), 2 mM glutamine (Gibco, Thermo Fisher Scientific), 100 IU·mL^{-1} of penicillin, and 100 μg·mL^{-1} streptomycin (EuroClone). Since the focus of this work was on integrin α2 expression and function in an integrin α2-null background, authentication of the CHO cell line was not performed by genotyping. Instead, the cells were characterized based on their distinctive CHO phenotype, specifically by confirming the absence of α1β1 and α2β1 integrin expression using flow cytometry and the lack of functional collagen receptor activity using xCELLigence technology. All experiments were performed using cells that tested negative for mycoplasma contamination.

Flow cytometry

Integrin α2 expression on transfected CHO cells was assessed by flow cytometry. Cells were grown to ~ 75% confluency and detached with trypsin. Approximately 3 × 10^5 cells per sample were incubated with mouse anti-human α2 primary antibody (3 : 200, CD49b, 12F1; BD Pharmingen, San Diego, CA, USA) for 1 h at 4 °C with continuous mixing, followed by Alexa Fluor 488–conjugated goat anti-mouse secondary antibody (1 : 100, IgG (H + L), Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, USA) for 30 min at 4 °C. Cells were washed and analyzed on an LSRFortessa (BD Biosciences, San Jose, CA, USA). Unstained samples and secondary-only controls were included as negative controls. Data were analyzed with FlowJo™ software (BD Biosciences), and cell sorting was performed using a Sony SH800S benchtop sorter (Sony Biotechnology Inc., San Jose, CA, USA).

Real time cell adhesion analysis

Real-time cell adhesion was assessed using the xCELLigence RTCA eSight (Agilent Technologies) system. E-plates (Agilent Technologies) were coated overnight at 4 °C with Collagen I (5 μg·cm^{-2}, Corning®) or the GFOFER peptide (5 μg·cm^{-2}, AUSPEP) diluted in PBS, with BSA-coated wells serving as negative controls. After coating, unspecific binding sites were blocked with 100 μL·well^{-1} of 0.1% BSA in PBS for 1 h at 37 °C. Wells were then washed twice with PBS, and 50 μL of serum-free media was added to establish the background measurement. Subsequently, 50 000 cells/well were seeded. The cell lines used were CHO, CHO$^{\alpha2WT}$, and CHO$^{\alpha2R288A}$. Impedance was monitored for 3.5 h, focusing on early adhesion events. All measurements were replicated twice. Data analysis and graphing were carried out in OriginLab 2016 (OriginLab Corporation), and final figures were assembled using CorelDRAW Graphics Suite 2025 (Corel Corporation, Ottawa, Canada).

Surface plasmon resonance preparation

Biacore T200 (GE healthcare) was used to immobilize both collagen I and IV on a CM5 chip. A noncoated reference

channel was included on the chip. Collagen was immobilized on the CM5 gold surfaces using the amine coupling kit provided. Collagen I was diluted to the immobilization buffer (0.01 mg·mL^{-1} in 10 mM NaAc, pH 5.5) and injected onto the chip at 30 µL·min^{-1} (HBS running buffer: 10 mM HEPES, 1.15 NaCl, 2 mM $MgCl_2$, pH 7.4). Collagen IV (0.2 mg·mL^{-1} in 19 mM NaAc, pH 7) was immobilized similarly and all channels were cleaned using a cleaning buffer (50 mM NaOh). All buffers were filtered before use. CM5 surfaces were coupled to RU values of ~ 500 and ~ 2000, respectively.

Surface plasmon resonance data acquisition

A concentration series (2400, 800, 400, 200, 100, 75, and o nM) of all α2I (α2IWT, α2IR288A, α2IE318A) diluted in HBS was injected to the channels using 300 s as the association time. Additionally, to the concentration series measured, each run included a technical duplicate and all samples were run at 10 °C. Afterwards, HBS was injected for 300 s as the dissociation phase. The same flow rate was used as previously mentioned. After dissociation, the CM5 chip was regenerated with 1 mM NaOH as the regeneration buffer. Regeneration was repeated three times for 30 s and after each analyte concentration. Gained kinetic data was analyzed with BIAevaluation software (GE Healthcare Life Sciences, Uppsala, Sweden). All binding curves were fitted using 1 : 1 binding, and data were corrected by subtraction of reference channel values. Data were analyzed and graphs prepared using OriginLab 2016 (OriginLab Corporation), and final figures were made with MS PowerPoint (Microsoft Corporation, Redmond, WA, USA) and CorelDRAW 2025 (Corel Corporation).

Crystallization, data collection, and structure determination

Crystallization of the α2IR288A and α2IE318A variants was done using the sitting drop vapor diffusion method and the Mosquito® (SPT Labtech Ltd., Melbourn, UK) or NT8 (Formulatrix, Bedford, MA, USA) robotics with a typical drop size of 200–400 nL. Initial crystallization screens were performed with the commercial screening kits PACT premier™ HT-96/FX-96 (MD1-36/MD1-36-FX; Molecular Dimensions Ltd., Newmarket, UK) and JCSG-plus™ (MD1-37; Molecular Dimensions Ltd.), both in room temperature and + 10°C, and using SPT Labtech's 96-well plates. All αI domains were in 40 mM Tris/HCl, pH 7.4, and 2 mM $MgCl_2$, and specific crystallization conditions were as follows: (a) α2IR288A^{Closed} (PDB ID: 9RK5, 9RK7), 2 : 1 ratio of protein (16.5 mg·mL^{-1}), and well solution (1.92 ammonium sulphate, 0.096 M Bis-Tris, pH 5.5); (b) α2IR288A^{Closed} (PDB ID: 9RKA), 1 : 1 ratio of protein (16.5 mg·mL^{-1}), and well solution (1 M tri-Na citrate, 0.1 M Na cacodylate); (c) α2IE318A^{Closed} (PDB ID: 9RHY), 1 : 1 ratio of protein (70 mg·mL^{-1}), and well solution (2.4 M sodium malonate, pH 7.0); (d) α2IE318A^{Open_s} (PDB ID: 9RI7), 2 : 1 ratio of protein (55 mg·mL^{-1}), and well solution (0.1 M succinic acid, pH 7.0, 15% PEG 3350); (e) α2IE318A^{Open_m} (PDB ID: 9RIE), 1 : 1 ratio of protein (100 mg·mL^{-1}), and well solution (0.15 M DL-malic acid, pH 7.0, 20% PEG 3350). For the α2IR288A^{Closed} variants, glycerol (typically 13–15% v/v) was used as a cryoprotectant, which was added to the crystallization drops just prior freezing them in liquid nitrogen. Data were collected at the European Synchrotron radiation facility (ESRF, Grenoble France) beam lines ID23-1 (α2IR288A^{Closed}; PDB IDs 9RKA and 9RK7; 2022), BM14 (α2IE318A^{Closed}; 2014) or ID29 (α2IE318A^{Open_s} and α2IE318A^{Open_m}; 2013), or for α2IR288A^{Closed} with PDB ID 9RK5 at the local X-ray source (MicroMax-007 HF X-ray generator/VariMax mirrors/hybrid photon counting X-ray area detector; Turku Bioscience, Finland; 2022).

XDS/XSALE [45] was used for data processing and data were merged with Aimless [46] of the CCP4i2 program suite [47]. The phase problem was solved using the molecular replacement program Phaser [48] and refinement was done using Refmac 5 [49], both within CCP4i2. The published α2I structure (PDB IDs: 1DZI), or structures solved here (PDB IDs: 9RK5 and 9RI7), were used as search templates in the replacements. coot [50] was used for manual building of the structures, and adding water molecules and ligands to the models, and the final models were validated with the inbuilt tools of Coot including Molprobity [51]. Since the α2IR288A^{Closed} structures (PDB IDs 9RK7, 9RK5, and 9RKA) were highly similar to each other, we chose the highest resolution 1.6 Å α2IR288A^{Closed} structure (PDB ID: 9RK7) as the representative model for further analysis. Structural analysis was performed with PyMOL and RMSD values were calculated with the super command of s (A chains were used). Figures were created with PyMOL [52] and labels added with MS PowerPoint (Microsoft Corporation) and CorelDRAW 2025 (Corel Corporation).

Acknowledgements

We thank Johanna Jokinen for the valuable help in establishing the cell lines and Maria Tuominen for technical support. Additionally, we thank the bioinformatics (J.V. Lehtonen) and FINStruct structural biology infrastructure facilities supported by Åbo Akademi University and Biocenter Finland. CSC-IT Center for Science for computational infrastructure support at the Structural Bioinformatics Laboratory (SBL), Åbo Akademi University. SBL is part of the Nordic Research Infrastructure Hub for Structural Biology (NORDSTRUCT), the NordForsk Nordic POP (Patient Oriented Products), NordicPharmaTrain (Nordic Pharmaceutical Translation and Innovation), the

Solutions for Health strategic profile area of Åbo Akademi University. We are grateful for the Sigrid Jusélius Foundation and the InFLAMES Flagship Programme of the Research Council of Finland for support in this research. Biocenter Oulu Infrastructure for Structural Biology, an Instruct-ERIC center, Turku Protein Core and Turku Bioscience are thanked for providing instrumentation. Artificial intelligence was used for language editing when needed. All outputs were carefully reviewed and verified by the authors.

This study has been financially supported by grants from the Sigrid Jusélius Foundation (J.H., T.A.S.), the Research Council of Finland's Flagship InFLAMES (337530 and 357910, J.H.; 337531, 357911 and 359346, T.A.S.), the European Union – NextGenerationEU instrument, FIRI2019 and FIRI2021, and is, respectively, funded by the Research Council of Finland (352823 T.T.A. and T.A.S.; 328119 and 345514 T.A.S.), the Cancer Foundation Finland (J.H.), The Research Council of Finland (329743 J.H.), Turku University Foundation (L.P.), the University of Turku Doctoral programme in Technology (L.P.), and Instruct-ERIC (PID 29721) (L.P.). Open access publishing facilitated by Abo Akademi, as part of the Wiley - FinELib agreement.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

LP, JH, TTA, and TAS contributed to conceptualization; LP, EB-L, VP, HT, TTA, and JK contributed to data curation; LP, TTA, HT, EB-L, and VP contributed to formal analysis; LP, JH, and TAS contributed to funding acquisition; LP, EB, VP, TTA, HT, JK, JH, and TAS contributed to investigation; JK, HT, TTA, and TAS contributed to methodology; JH, and TAS contributed to project administration; JH and TAS contributed to resources; TTA, JK, JH, and TAS contributed to supervision; LP, HT, TTA, JK, EB-L, VP, JH, and TAS contributed to validation; LP contributed to visualization; LP and TTA contributed to writing—original draft; LP, TTA, JK, JH, and TAS contributed to writing—review and editing.

Peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.70658.

Data availability statement

The original contributions presented in this study are included in the article; further inquiries can be directed to the corresponding author. The atomic coordinates and structure factors for the crystal structures published in this article have been deposited in the Protein Data Bank (PDB) [53] with the following accession IDs: α2IR288A^{Closed} (9RK5, 9RKA and 9RK7), α2IE318A^{Closed} (9RHY), α2IE318A^{Open_s} (9RI7), and α2IE318A^{Open_m} (9RIE).

References

1 Farndale R, Sonnenberg A, DiPersio CM, Eble JA, Heino J & Gullberg D (2023) What does it take to be a collagen receptor? *Matrix Biol* **115**, 128–132.
2 Whittaker CA & Hynes RO (2002) Distribution and evolution of von Willebrand/integrin a domains: widely dispersed domains with roles in cell adhesion and elsewhere. *MBoC* **13**, 3369–3387.
3 Johnson MS, Lu N, Denessiouk K, Heino J & Gullberg D (2009) Integrins during evolution: evolutionary trees and model organisms. *Biochim Biophys Acta Biomembr* **1788**, 779–789.
4 Johnson MS, Käpylä J, Denessiouk K, Airenne T, Chouhan BS & Heino J (2013) Evolution of cell adhesion to extracellular matrix. In Evolution of Extracellular Matrix (Keeley FW & Mecham RP, eds), pp. 243–283. Springer Berlin Heidelberg, Berlin, Heidelberg.
5 Heino J, Huhtala M, Käpylä J & Johnson MS (2009) Evolution of collagen-based adhesion systems. *Int J Biochem Cell Biol* **41**, 341–348.
6 Käpylä J, Ivaska J, Riikonen R, Nykvist P, Pentikäinen O, Johnson M & Heino J (2000) Integrin α2I domain recognizes type I and type IV collagens by different mechanisms. *J Biol Chem* **275**, 3348–3354.
7 Heino J & Siljamäki E (2023) Integrins α1β1 and α2β1: the generalist collagen receptors. In Integrins in Health and Disease (Gullberg D & Eble JA, eds), pp. 3–25. Springer International Publishing, Cham.
8 Zeltz C, Kusche-Gullberg M, Heljasvaara R & Gullberg D (2023) Novel roles for cooperating collagen receptor families in fibrotic niches. *Curr Opin Cell Biol* **85**, 102273.
9 Holtkötter O, Nieswandt B, Smyth N, Müller W, Hafner M, Schulte V, Krieg T & Eckes B (2002) Integrin α2-deficient mice develop normally, are fertile, but display partially defective platelet interaction with collagen. *J Biol Chem* **277**, 10789–10794.
10 Chen J, Diacovo TG, Grenache DG, Santoro SA & Zutter MM (2002) The α2 integrin subunit-deficient mouse. *Am J Pathol* **161**, 337–344.

11 Pozzi A, Wary KK, Giancotti FG & Gardner HA (1998) Integrin α1β1 mediates a unique collagen-dependent proliferation pathway in vivo. *J Cell Biol* **142**, 587–594.
12 The Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, Cell type annotation, Writing group, Supplemental text writing group & Principal investigators (2018) Single-cell transcriptomics of 20 mouse organs creates a tabula Muris. *Nature* **562**, 367–372.
13 Sarratt KL, Chen H, Zutter MM, Santoro SA, Hammer DA & Kahn ML (2005) GPVI and α2β1 play independent critical roles during platelet adhesion and aggregate formation to collagen under flow. *Blood* **106**, 1268–1277.
14 Hemler ME, Jacobson JG & Strominger JL (1985) Biochemical characterization of VLA-1 and VLA-2. Cell surface heterodimers on activated T cells. *J Biol Chem* **260**, 15246–15252.
15 Nissinen L, Koivunen J, Käpylä J, Salmela M, Nieminen J, Jokinen J, Sipilä K, Pihlavisto M, Pentikäinen OT, Marjamäki A *et al.* (2012) Novel α2β1 integrin inhibitors reveal that integrin binding to collagen under shear stress conditions does not require receptor Preactivation. *J Biol Chem* **287**, 44694–44702.
16 Huang K, Li Z, He X, Dai J, Huang B, Shi Y, Fan D, Zhang Z, Liu Y, Li N *et al.* (2024) Gut microbial co-metabolite 2-methylbutyrylcarnitine exacerbates thrombosis via binding to and activating integrin α2β1. *Cell Metab* **36**, 598–616.e9.
17 Senger DR, Claffey KP, Benes JE, Perruzzi CA, Sergiou AP & Detmar M (1997) Angiogenesis promoted by vascular endothelial growth factor: regulation through $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins. *Proc Natl Acad Sci USA* **94**, 13612–13617.
18 Luo B-H, Carman CV & Springer TA (2007) Structural basis of integrin regulation and signaling. *Annu Rev Immunol* **25**, 619–647.
19 Knight CG, Morton LF, Peachey AR, Tuckwell DS, Farndale RW & Barnes MJ (2000) The collagen-binding A-domains of integrins α1β1 and α2β1Recognize the same specific amino acid sequence, GFOGER, in native (triple-helical) collagens. *J Biol Chem* **275**, 35–40.
20 Hamaia S & Farndale RW (2014) Integrin recognition motifs in the human collagens. In I Domain Integrins (Gullberg D, ed.), pp. 127–142. Springer Netherlands, Dordrecht.
21 Emsley J, Knight CG, Farndale RW, Barnes MJ & Liddington RC (2000) Structural basis of collagen recognition by integrin α2β1. *Cell* **101**, 47–56.
22 Emsley J, King SL, Bergelson JM & Liddington RC (1997) Crystal structure of the I domain from integrin α2β1. *J Biol Chem* **272**, 28512–28517.
23 Connors WL, Jokinen J, White DJ, Puranen JS, Kankaanpää P, Upla P, Tulla M, Johnson MS & Heino J (2007) Two synergistic activation mechanisms of α2β1 integrin-mediated collagen binding. *J Biol Chem* **282**, 14675–14683.
24 Aquilina A, Korda M, Bergelson JM, Humphries MJ, Farndale RW & Tuckwell D (2002) A novel gain-of-function mutation of the integrin α2 VWFA domain. *Eur J Biochem* **269**, 1136–1144.
25 Lahti M, Bligt E, Niskanen H, Parkash V, Brandt A-M, Jokinen J, Patrikainen P, Käpylä J, Heino J & Salminen TA (2011) Structure of collagen receptor integrin α1I domain carrying the activating mutation E317A. *J Biol Chem* **286**, 43343–43351.
26 Tulla M, Huhtala M, Jäälinoja J, Käpylä J, Farndale RW, Ala-Kokko L, Johnson MS & Heino J (2007) Analysis of an ascidian integrin provides new insight into early evolution of collagen recognition. *FEBS Lett* **581**, 2434–2440.
27 Tulla M, Pentikäinen OT, Viitasalo T, Käpylä J, Impola U, Nykvist P, Nissinen L, Johnson MS & Heino J (2001) Selective binding of collagen subtypes by integrin α1I, α2I, and α10I domains. *J Biol Chem* **276**, 48206–48212.
28 Xu Y, Gurusiddappa S, Rich RL, Owens RT, Keene DR, Mayne R, Höök A & Höök M (2000) Multiple binding sites in collagen type I for the integrins α1β1 and α2β1. *J Biol Chem* **275**, 38981–38989.
29 Jokinen J, White DJ, Salmela M, Huhtala M, Käpylä J, Sipilä K, Puranen JS, Nissinen L, Kankaanpää P, Marjomäki V *et al.* (2010) Molecular mechanism of α2β1 integrin interaction with human echovirus 1. *EMBO J* **29**, 196–208.
30 Huizinga EG, Tsuji S, Romijn RAP, Schiphorst ME, De Groot PG, Sixma JJ & Gros P (2002) Structures of glycoprotein Ibα and its complex with von Willebrand factor A1 domain. *Science* **297**, 1176–1179.
31 Lee J-O, Bankston LA, Arnaout MA & Liddington RC (1995) Two conformations of the integrin A-domain (I-domain): a pathway for activation? *Structure* **3**, 1333–1340.
32 Nissinen L, Ojala M, Langen B, Dost R, Pihlavisto M, Käpylä J, Marjamäki A & Heino J (2015) Sulfonamide inhibitors of $\alpha\ 2\ \beta\ 1$ integrin reveal the essential role of collagen receptors in in vivo models of inflammation. *Pharmacol Res Perspect* **3**, e00146.
33 Huhtala M, Heino J, Casciari D, De Luise A & Johnson MS (2005) Integrin evolution: insights from ascidian and teleost fish genomes. *Matrix Biol* **24**, 83–95.
34 Chouhan BS, Käpylä J, Denessiouk K, Denesyuk A, Heino J & Johnson MS (2014) Early chordate origin of

the vertebrate integrin αI domains. *PLoS One* **9**, e112064.
35 Ewan R, Huxley-Jones J, Mould AP, Humphries MJ, Robertson DL & Boot-Handford RP (2005) The integrins of the urochordate Ciona intestinalis provide novel insights into the molecular evolution of the vertebrate integrin family. *BMC Evol Biol* **5**, 31.
36 Boot-Handford RP & Tuckwell DS (2003) Fibrillar collagen: the key to vertebrate evolution? A tale of molecular incest. *BioEssays* **25**, 142–151.
37 Kim C, Ye F & Ginsberg MH (2011) Regulation of integrin activation. *Annu Rev Cell Dev Biol* **27**, 321–345.
38 Cruz MA, Chen J, Whitelock JL, Morales LD & López JA (2005) The platelet glycoprotein Ib–von Willebrand factor interaction activates the collagen receptor α2β1 to bind collagen: activation-dependent conformational change of the α2-I domain. *Blood* **105**, 1986–1991.
39 Li J, Jo MH, Yan J, Hall T, Lee J, López-Sánchez U, Yan S, Ha T & Springer TA (2024) Ligand binding initiates single-molecule integrin conformational activation. *Cell* **187**, 2990–3005.
40 Nuyttens BP, Thijs T, Deckmyn H & Broos K (2011) Platelet adhesion to collagen. *Thromb Res* **127**, S26–S29.
41 Kamata T, Liddington RC & Takada Y (1999) Interaction between collagen and the α2 I-domain of integrin α2β1. *J Biol Chem* **274**, 32108–32111.
42 Smith C, Estavillo D, Emsley J, Bankston LA, Liddington RC & Cruz MA (2000) Mapping the collagen-binding site in the I domain of the glycoprotein Ia/IIa (integrin α2β1). *J Biol Chem* **275**, 4205–4209.
43 Käpylä J, Jäälinoja J, Tulla M, Ylöstalo J, Nissinen L, Viitasalo T, Vehviläinen P, Marjomäki V, Nykvist P, Säämänen A-M *et al.* (2004) The fibril-associated collagen IX provides a novel mechanism for cell adhesion to cartilaginous matrix. *J Biol Chem* **279**, 51677–51687.
44 Ivaska J, Reunanen H, Westermarck J, Koivisto L, Kähäri V-M & Heino J (1999) Integrin α2β1 mediates isoform-specific activation of p38 and upregulation of collagen gene transcription by a mechanism involving the α2 cytoplasmic tail. *J Cell Biol* **147**, 401–416.
45 Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
46 Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **67**, 282–292.
47 Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, Jenkins HT, Keegan R, Krissinel E, Stevenson K *et al.* (2018) *CCP4i2*: the new graphical user interface to the*CCP4* program suite. *Acta Crystallogr D Struct Biol* **74**, 68–84.
48 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
49 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
50 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of *Coot*. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
51 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21.
52 Schrodinger L (2010) The PyMOL molecular graphics system.
53 Berman HM (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Structure determination statistics for α2I domain constructs.

Figure S1. Electron density maps (2FoFc) around the Mg^{2+} ion of MIDAS and amino acids and water molecules coordinating the metal binding.

Figure S2. Stabilization of the α7 helix by Tyr285 in the α2IE318AOpen variants.