Endocytic recycling is central to circadian collagen fibrillogenesis and disrupted in fibrosis

Authors

[1,2]Joan Chang*, [1]Adam Pickard, [1]Jeremy A. Herrera, [2]Sarah O'Keefe, [1]Matthew Hartshorn, [1]Richa Garva, [1]Anna Hoyle, [3]Lewis Dingle, [4]Cédric Zeltz, [3]Jason Wong, [3]Adam Reid, [5]Rajamiyer V. Venkateswaran, [1]Yinhui Lu, [1]Patrick Caswell, [2]Stephen High, [4]Donald Gullberg, and [1]Karl E. Kadler*

Affiliations

[1]Wellcome Centre for Cell-Matrix Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[2]Division of Molecular and Cellular Function, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[3]Blond McIndoe Laboratories, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.

[4]Department of Biomedicine and Centre for Cancer Biomarkers, Norwegian Center of Excellence, University of Bergen, Norway.

[5]Manchester University National Health Service Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK.

* Co-corresponding authors:

joan.chang@manchester.ac.uk (JC)

karl.kadler@manchester.ac.uk (KEK)

Abstract

Collagen-I fibrillogenesis is crucial to health and development, where dysregulation is a hallmark of fibroproliferative diseases. Here, we show that collagen-I fibril assembly required a functional endocytic system that recycles collagen-I to assemble new fibrils. Endogenous collagen production was not required for fibrillogenesis if exogenous collagen was available, but the circadian-regulated vacuolar protein sorting (VPS) 33b and collagen-binding integrin-α11 subunit were crucial to fibrillogenesis. Cells lacking VPS33b secrete soluble collagen-I protomers but were deficient in fibril formation, thus secretion and assembly are separately controlled. Overexpression of VPS33b led to loss of fibril rhythmicity and over-abundance of fibrils, which was mediated through integrin α11β1. Endocytic recycling of collagen-I was enhanced in human fibroblasts isolated from idiopathic pulmonary fibrosis, where VPS33b and integrin-α11 subunit were overexpressed at the fibrogenic front; this correlation between VPS33b, integrin-α11 subunit, and abnormal collagen deposition was also observed in samples from patients with chronic skin wounds. In conclusion, our study showed that circadian-regulated endocytic recycling is central to homeostatic assembly of collagen fibrils and is disrupted in diseases.

Introduction

Collagen fibrils account for ~25% of total body protein mass [1] and are the largest protein polymers in vertebrates [2]. The fibrils can exceed centimeters in length and are organized into elaborate networks to provide structural support for cells. It is unclear how the fibrils are formed and how this process goes awry in collagen pathologies, such as fibrosis. A key mechanism, cell surface mediated fibrillogenesis, has been suggested to occur via indirect binding to fibronectin fibrils, or via direct assembly by collagen-binding integrins [3-5]. Interestingly, fibrils can be reconstituted *in vitro* from purified collagen (reviewed in [6]) but the assembly process is not controlled to the extent seen *in vivo* in terms of number, size and organization; this suggests a much tighter cellular control over the process. Additional support for tight cellular control of collagen fibril formation comes from electron microscope observations of collagen fibrils at the plasma membrane of embryonic avian and rodent tendon fibroblasts [7, 8], where the end (tip) of a fibril is enclosed within a plasma membrane invagination termed a fibripositor [9]. Previously we identified vacuolar protein sorting (VPS) 33b (a regulator of SNARE-dependent membrane fusion in the endocytic pathway) as a circadian clock regulated protein involved in collagen homeostasis [10]. VPS33B forms a protein complex with VIPAS39 (VIPAR) [11], and mutations in the *VPS33B* gene cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome [12]. Here, death usually occurring within the first year of birth with renal insufficiency, jaundice, multiple congenital anomalies, and predisposition to infection [13]. One proposed disease-causing mechanism is abnormal post-Golgi trafficking of lysyl hydroxylase 3 (LH3, PLOD3), which is essential for collagen homeostasis during development [14]. All these observations points to a central role for the endosome, situated in proximity of the plasma membrane and acting as a hub for a complex assortment of vesicles, in sorting collagen molecules to different fates.

Previous studies of collagen endocytosis have focused on collagen degradation and signaling, identifying additional collagen-binding proteins such as Endo180 and MRC1 [15-19], as well as a role for non-integrin collagen-receptors involved in mediating signaling activities such as DDR1/DDR2 (reviewed in [20]). Here, we showed that collagen uptake by fibroblasts is circadian rhythmic. Importantly, instead of degrading endocytosed collagen, cells utilize endocytic recycling of exogenous collagen-I to assemble new fibrils, even in the absence of endogenous collagen production. Further we showed that the secretion of soluble collagen protomers is separate from, and can occur independently to, collagen fibril assembly. We identified VPS33b and integrin-α11 subunit (part of the collagen-binding integrin α11β1 heterodimer [4, 21, 22]) as central molecules specific to fibril formation. Finally, we showed that in idiopathic pulmonary fibrosis (IPF), a life threatening disease with unknown trigger/mechanism where lung tissue is replaced with collagen fibrils, integrin α11 subunit and VPS33B are located at the invasive fibroblastic focus where collagen fibril rapidly accumulates [23]; IPF fibroblasts also have elevated endocytic recycling of exogenous collagen-I, despite having similar collagen-I expression level to normal fibroblasts. Together, these results provide novel insights into the importance of the circadian clock-regulated endosomal system in normal fibrous tissue homeostasis,

where fibrillogenesis occurs with both endogenous collagen and/or recycled scavenged exogenous collagen. This work also highlights that collagen utilization, rather than production (i.e. assembled into fibrils vs. protomeric secretion), is central to the maintenance of homeostasis, and dysregulation leads to disease.

Keywords: collagen-I, endocytic recycling, fibrillogenesis, integrin-α11, circadian, trafficking, VPS33B, fibrosis, chronic wound

Results

Collagen-I is taken up into vesicle-like structures within the cell, and reassembled into fibrils

To study collagen-I endocytosis and the fate of endocytosed collagen, we made Cy3 labeled collagen-I (Cy3-coll) [24] and incubated it with explanted murine tail tendons for 3 days, then imaged the core of the tendon. Cells in tendon (nuclei marked with Hoechst stain) showed a clear uptake of Cy3-coll (**Figure 1A**, yellow box, indicated by yellow arrowheads). Interestingly, fibrillar Cy3 fluorescence signals that extend across cells were also observed (**Figure 1A**, grey box, indicated by white arrows). Pulse-chase experiments were performed where Cy3-coll was added to the tendons for 3 days, followed by 5FAM-coll for a further 2 days (**Figure S1A**). The results showed distinct areas where only Cy3-coll was observed (**Figure S1A**, yellow box); this persistence indicated that not all collagen endocytosed by cells is directed for degradation. 5FAM-positive fibril-like structures were also observed, confirming that collagen-I taken up by tissues is reincorporated into the matrix (**Figure S1A**, grey boxes). To confirm these findings in 2D cultures, we incubated immortalized murine tendon fibroblast cultures (iTTF) with labelled Cy3-coll. Flow cytometry analysis revealed that most cells had taken up Cy3-coll after overnight incubation (**Figure S1B**). Further time course analyses revealed a time-dependent and concentration-dependent increase of collagen-I uptake (**Figure 1B**). Cells incubated with Cy3-Coll for one hour followed by trypsinization were then analyzed using flow cytometry imaging, revealing that collagen-I is endocytosed by the cells into distinct structures (**Figure 1C**). Similar results were obtained using Cy5-labelled collagen-I (**Figure S1C, D, E**). To confirm that the labelled collagen-I is in fact endocytosed into the fibroblasts and not only associated with the cell surface when cells are still attached, we performed live-imaging on cells incubated with Cy3-coll. Time lapse images showed Cy3-coll congregate in structures within the cell, as indicated with white arrows (**Figure 1D**). We then incubated iTTFs with Cy3-coll for one hour, before trypsinizing and replating the cells to ensure any Cy3-coll signal detected hereon originated from the endocytosed pool. Immunofluorescence (IF) staining indicated co-localization of endogenous collagen-I with Cy3-labeled collagen-I (**Figure 1E**). These results demonstrated that exogenous collagen-I can be taken up by cells both *in vitro* and in tissues *ex vivo*, and recycled into fibrils.

We considered the possibility that the fluorescent fibril-like structures were the result of spontaneous cell-free fibrillogenesis of the added Cy3-coll. Thus, we incubated Cy3-coll at concentrations spanning the 0.4 μg/mL critical concentration for cell-free *in vitro* fibril formation [25], with or without cells (**Figure 1F**). In cell-free cultures, Cy3-positive fibrils were observed at 1 μg/mL and 5 μg/mL with no discernible Cy3-positive fibrils at 0.1 μg/mL Cy3-coll (**Figure 1F**, right panel). However, in the presence of cells, Cy3-positive fibrillar structures were observed at the cell surfaces at all concentrations of Cy3-coll examined, including 0.1 μg/mL (**Figure 1F** zoom ins of red box expanded to the left, highlighted by white arrows). These results indicated that the cells actively endocytosed and recycled Cy3-coll into fibrils.

The endocytic pathway controls collagen-I secretion and fibril assembly

Having shown that cells can take up and recycle collagen-I into extracellular fibrils, we then asked if endocytosis is required for collagen deposition. We used Dyngo4a (Dyng) to inhibit clathrin-mediated endocytosis, where Dyng treatment (20 μM) leads to ~ 60% reduction in Cy3-coll uptake relative to control (**Figure S2A**), without affecting cell viability (**Figure S2B**). We then investigated the effects of Dyng treatment on the ability of wild-type fibroblasts to assemble collagen fibrils. Cells were treated with Dyng for 48 h before fixation and IF against collagen-I antibody. A significant reduction (average 60%) in the number of collagen-I fibrils assembled was observed (**Figure 2A**), indicative of endocytosis playing a key role in collagen-I fibrillogenesis. Conditioned media (CM) from these treated cells were also collected and assessed. To our surprise, the amount of soluble collagen-I secreted was also greatly reduced (**Figure 2B**). qPCR analyses revealed that *Col1α1* mRNA was significantly reduced in Dyng-treated cells (**Figure S2C**), suggesting a potential

feedback mechanism between endocytosis and collagen-I synthesis, secretion, and fibrillogenesis. Interestingly, whilst *Fn1* mRNA was significantly lower in Dyng-treated cells (**Figure S2C)**, and the intensity of FN1 signal appears lower, the amount of FN1 fibrils deposited (as determined by the area occupied by fibrils) was not significantly impacted (**Figure 2C**). Taken together, these results showed that inhibiting endocytosis in fibroblasts does not lead to accumulation of soluble collagen-I protomers in the extracellular space. Rather, endocytosis impacts collagen fibrillogenesis and the transcriptional control on collagen-I and fibronectin.

Collagen-I endocytosis is circadian clock regulated, and recycling alone can generate fibrils

Previously, we have shown that clock-synchronized fibroblasts synthesize collagen fibrils in a circadian rhythmic manner [10]; the results here thus far indicate an involvement of endocytosis in collagen fibrillogenesis. Therefore, we hypothesized that collagen-I endocytosis may also be circadian clock regulated. Time-series flow cytometry analyses of fibroblasts incubated with Cy3-coll revealed that the level of Cy3-coll endocytosed by the cells is rhythmic, with a periodicity of 23.8hrs as determined by MetaCycle analysis (**Figure 2D**). When corrected to the running average, the rhythmic nature of Cy3-coll uptake is accentuated (**Figure S2D)**. We noted that, when compared to the number of fibrils produced over time, the peak time of uptake has an inverse correlation with fibril numbers (**Figure 2E**). These data suggest that the cells may be endocytosing exogenous collagen under circadian control and holding it in the endosomal compartment, then trafficking the collagen to the plasma membrane for fibril formation.

To eliminate the possibility that the fluorescent fibril-like structures were due to attachment of fluorescently-labeled collagen protomers to pre-existing fibrils already deposited by cells, we performed an siRNA-mediated knockdown against *Col1a1* (siCol1a1) to target endogenous collagen production. Fibroblasts were then analyzed by IF using anti-collagen-I or anti-fibronectin (FN1) antibodies. Control (scrambled siRNAs, scr) cells synthesized collagen-I and fibronectin (**Figure 3A**, top row) with defined fibrillar structures. In contrast, cells treated with siCol1a1 synthesized lower-levels of collagen-I and fibronectin, with few collagen-I fibrillar structures as well as lower intracellular collagen-I signal (**Figure 3A**, bottom row). A ~ 90% reduction in *Col1a1* mRNA was confirmed using qPCR (**Figure S3A**).

Flow cytometry analysis of cells incubated with Cy3-coll revealed that siCol1a1 fibroblasts retained the ability to endocytose exogenous collagen (**Figure S3B**). To assess if siCol1a1 fibroblasts can assemble a collagen fibril with exogenous collagen, Cy3-coll was incubated with scr or siCol1a1 cells for 1 hour, before trypsinization and replating to ensure no cell surface associated Cy3-coll remains. Within 3 days, Cy3-positive collagen fibrils were observed in both scr (**Figure 3B**, left column) and siCol1a1 cells (**Figure 3B**, right column). The observation of fibrillar Cy3 signals in siCol1a1 cells showed that the cells can repurpose collagen into fibrils without the requirement for intrinsic collagen-I production (red arrow, **Figure 3B**). The cells were also probed for collagen-I using an anti-collagen-I antibody with a secondary antibody conjugated to Alexa-488 (**Figure 3B**, second row). In scr cells we detected multiple fibrillar structures that did not contain Cy3-coll (**Figure 3B,** white arrows), indicating fibrils derived from endogenous collagen-I. In the siCol1a1 cells we detected a more diffuse signal and puncta within the cell, but no discernible Alexa-488-only collagen-I fibrils (**Figure 3B**). These results indicate that fibroblasts can assemble collagen-I fibrils from recycled exogenous collagen-I alone. Importantly, siCol1a1 cells treated with Dyng could not effectively form Cy3-collagen-containing fibrils under these conditions, confirming that siCol1a1 cells, like control cells (see **Figure 1E**), are endocytosing exogenous collagen to be recycled into new fibrils (**Figure 3C**).

VPS33B controls collagen fibril formation but not protomeric secretion

VPS33B is situated in a post-Golgi compartment where it is involved in endosomal trafficking with multiple functions including extracellular vesicle formation [26], modulation of p53 signaling [27], and maintenance of cell polarity [28], dependent on cell type. We previously identified VPS33B to be a circadian-controlled component of collagen-I homeostasis in fibroblasts [10]. As a result, we decided to further investigate its role in endosomal recycling of collagen-I.

We confirmed our previous finding that CRISPR-knockout of VPS33B (VPSko, as verified by western blot analysis and qPCR analysis, **Figure S4A and B**) in tendon fibroblasts led to fewer collagen fibrils without impacting proliferation (**Figure S4C**), as evidenced by both electron microscopy (**Figure 4A**) and IF imaging

(**Figure 4B**). Decellularized matrices showed that VPSko fibroblasts produced less matrix by mass than control, which was mirrored by a reduction in hydroxyproline content (**Figure 4C**). We then stably overexpressed VPS33B in fibroblasts (VPSoe), as confirmed by western blot, qPCR analyses, and flow cytometry of transfected cells (**Figure S4D-G**). IF staining indicated a greater number of collagen-I fibrils in VPSoe cells (**Figure 4D**), although the mean total matrix and mean total hydroxyproline in VPSoe cultures was not significantly higher than control (**Figure 4E**). As VPSoe cells showed equivalent proliferation to controls (**Figure S4F**), this suggests VPSoe specifically enhanced the assembly of collagen fibrils. We then performed time-series IF on synchronized cell cultures to quantify the number of collagen fibrils formed. Tendon fibroblasts exhibited a ~ 24 h rhythmic fluctuation in collagen-I fibril numbers, whereas VPSoe cells continuously deposited collagen-I fibrils over the 55-h period (**Figure 4F**). MetaCycle analyses indicated that fluctuation of fibril numbers in control and VPSoe cultures occurred at a periodicity of 22.7 h and 28.0 h respectively (**Figure S4H**), indicating that continuous expression of VPS33B leads to loss of collagen-I fibril circadian homeostasis. Interestingly, when assessing the levels of secreted soluble collagen-I protomers from control and VPSoe cells, VPSoe CM have lower levels of soluble collagen-I (**Figure 4J**). In addition, when VPS33B is knocked-down using siRNA, there is an elevation of soluble collagen-I secreted; siVPS33B on VPSoe cells also increased the levels of collagen-I in CM (**Figure S4I, J**), whilst VPS33b levels are not correlated with *Col1a1* expression levels (**Figure S4K**). This finding indicates that VPS33B is specifically involved in collagen-I fibril assembly, and not secretion. The reduction of secreted soluble collagen-I in VPSoe cells further supports that VPS33B directs collagen-I towards fibril assembly.

VPS33B-positive vesicles contain collagen-I

We then utilized the split-GFP system [29, 30] to investigate how VPS33B specifically directs collagen-I to fibril assembly. Here, a GFP signal will be present if VPS33B co-traffics with collagen-I. We, and others, have previously demonstrated that insertion of tags at the N-terminus of proα2(I) or proα1(I) chain does not interfere with collagen-I folding or secretion [31, 32]. To determine which terminus of the VPS33B protein GFP1-10 should be added, we performed computational ΔG analysis [33], which predicts that VPS33B contains two regions of extended hydrophobicity towards its C-terminus (**Figure S5A**). The first region (residues 565-587, denoted TMD1) has a ΔG of -0.62 kcal/mol, consistent with a single pass type IV transmembrane domain (TMD), known as a tail-anchor. This arrangement locates the short C-terminus of VPS33B inside the lumen of the endoplasmic reticulum (ER), and subsequently within the endosomal compartment [34]. The second region (residues 591-609, denoted HR1) is significantly less hydrophobic, with a ΔG of +2.61 kcal/mol (**Figure S5B**). Its presence raises the possibility that VPS33B may contain two TMDs that, depending on their relative membrane topologies, result in either a luminal or cytosolic C-terminus (**Figure S5C**).

To define the topology of VPS33B, we used a well-established *in vitro* system where newly synthesized and radiolabeled proteins of interest are inserted into the membrane of ER-derived canine pancreatic microsomes, and created constructs with ER luminal modification of either endogenous N-glycosylation sites (N54, N545 in VPS33b), or artificial sites in an appended OPG2 tag (N2, N15 in residues 1-18 of bovine rhodopsin, Uniprot: P02699) as a robust reporter for membrane protein topology in the ER (**Figure S5D, E**) [35]. Due to the N-terminal region of VPS33b being highly aggregation-prone in our *in vitro* system (**Figure S5F, G**), we created three additional chimeric proteins comprised of different regions of VPS33b and Sec61β to investigate the topology of VPS33B (**Figure S5H**). Whilst a small proportion of each chimera continued to pellet when synthesized in the absence of ER-derived microsomes (**Figure S5I,** lanes 3, 6, 9), N-gvlycosylated species were now clearly identifiable for each of the chimeras tested.

In chimera 1, given the respective efficiency of membrane insertion of the TA region and N-glycosylation of the C-terminal OPG tag of Sec61βOPG2 (**Figure S4I,** lanes 1-2), we attribute the N-glycosylated species to the C-terminal translocation of OPG2 tag, although it is evident that the remaining short stretch of VPS33b (residues 414-564) still impedes efficient membrane insertion, likely due to aggregation (**Figure S4I**, lanes 3-5) [36, 37]. For chimera 2, the efficient modification of its distal N-glycosylation site (N15 of the OPG2 tag) but inefficient use of the proximal site (N2 of the OPG2 tag) reflects the latter residues' close proximity to the ER membrane [38], and supports the *bona fide* membrane insertion of the VPS33b TMD1 and thus translocation of the C-terminal OPG2 tag into the ER lumen (**Figure S4I**, lanes 6-8). Interestingly, the inclusion of HR1 to the C-

terminus of TMD1 in chimera 3 results in a substantial qualitative reduction in the amount of protein that is N-glycosylated (**Figure S4I**, lanes 7 and 10). Further, for the fraction of chimera 3 that is modified, the majority is doubly N-glycosylated; most likely due to the extra length provided by HR1, resulting in both N-glycosylation sites on the OPG2 tag (N2 and N15) now being efficiently modified [38].

The clear reduction in the proportion of N-glycosylated species obtained with chimera 3 compared to chimera 2 (**Figure S4I**) indicates that only a small proportion of HR1 and the appended OPG2 tag are successfully translocated into the ER lumen. This is likely due to a combination of the low proportion of hydrophobic residues and the presence of several charged and polar amino acids in HR1[39]. We thus propose that, for a minority of VPS33b, TMD1 is inserted into the ER membrane as a tail-anchored region with a luminal HR1 that most likely remains associated with the inner leaflet of the bilayer (N-cytosolic, C-luminal; **Figure 5A**, right). In contrast, the majority of VPS33b likely assumes a 'hairpin' like conformation in the ER membrane where both its N- and C-termini remain in the cytosol (N-cytosolic, C-cytosolic) with either: i) a partially membrane inserted TMD1, with HR1 associated with the outer leaflet of the ER membrane (**Figure 5A**, left); or ii) a fully membrane inserted TMD1 followed by the marginally hydrophobic HR1 which may span the membrane through the formation of stabilizing hydrogen bonds with TMD1 (**Figure 5A**, middle) [40].

To test our topology findings, we inserted BFP at either the N- (VPSnBFP) or C-terminus (VPScBFP) of VPS33B. In cells expressing VPSnBFP, the fluorescence signal was diffuse throughout the cell body, and in some cases appeared to be completely excluded from vesicular structures (**Figure 5B**, left), suggestive of protein mistargeting. In contrast, VPScBFP-expressing cells have punctate, vesicular-like blue fluorescence signals and peripherally blue structures (**Figure 5B**, right), supportive of the topology findings. Previous studies have also tagged VPS33B at the C-terminus [11]. Thus, the 214-residue N-terminal fragment (GFP1-10) was cloned onto VPS33B, and the 17-residue C-terminal peptide (GFP11) was cloned onto the proα1(I) chain of collagen-I as previously described [31] (GFP11-proα1(I), **Figure 5C**).

Fibroblasts stably expressing VPS-barrel and GFP11-proα1(I) were imaged to identify any GFP fluorescence. The results showed puncta throughout the cell body (**Figure 5D**). Intriguingly, GFP fluorescence was also observed at the cell periphery (**Figure 5D**, green box zoom). IF staining of endogenous VIPAS39 (a known VPS33B-interacting partner [11]) also revealed co-localization of VIPAS39 with collagen-I in intracellular vesicular structures (**Figure 5E**); in some of the co-localized vesicles, the signal of VIPAS39 is strongest surrounding collagen-I (**Figure 5E**, zoom), suggesting that VIPAS39 is encasing collagen-I, not within the vesicle lumen but present in proximity with the external membrane of the vesicle. VPS33B has been demonstrated to interact with VIPAS in regions before the TMD1 site [41]; thus, in all suggested VPS33B topologies herein, VIPAS39 will still be able to interact with VPS33B, and encase collagen-I-containing vesicles.

Notably, when VPScBFP was overexpressed in Dendra2-collagen-I expressing 3T3 cells [42], the number of Dendra2-positive fibrils significantly increased, verifying the effects of VPS33B overexpression in fibroblast collagen-I deposition; however, the average length of the fibril was not significantly different, suggesting that VPS33B is important in fibril initiation but not elongation (**Figure 5F**). Flow cytometry analyses of fluorescently labelled-coll endocytosed by control, VPSko, and VPSoe fibroblasts revealed no consistent change in uptake by VPSko or VPSoe cells (**Figure 5G**). Taken together, these results suggest that VPS33B is involved in the trafficking of collagen-I to the extracellular space, and not endocytosis. VPSko cells replated after Cy5-coll uptake have conspicuously fewer fibrils when compared to control. In contrast, VPSoe cells have shorter but more Cy5-coll fibrils (**Figure 5H**), highlighting the role of VPS33B in recycling endocytosed collagen-I to fibril assembly sites.

Integrin chain α11 mediates VPS33B-dependent fibrillogenesis

Having identified VPS33B as specifically driving collagen-I fibril formation, we used biotin cell surface labeling coupled with mass spectrometry protein identification to identify other proteins that may be involved in this process at the cell surface. VPSko and VPSoe fibroblasts were analyzed using a "shotgun" approach. In total, 4121 proteins were identified in total lysates (**Table S1**), and 1691 proteins in the enriched-for-surface-protein samples (**Table S2**). Gene Ontology (GO) Functional Annotation analysis [43, 44] identified the top 25 enriched terms (based on p-values) with the top 5 terms all associated with "extracellular" or "cell surface" (**Figure 6A**),

indicative that the biotin-labeling procedure had successfully enriched proteins at the cell surface interacting or associated with the extracellular matrix (ECM).

We then interrogated the differences between VPSko and control cells, visualizing the results in a semi-quantitative manner using spectral counting (**Table S3**). Conspicuously, collagen α1(I) and α2(I) chains were detected at a reduced level at the surface of VPSko cells (**Figure 6B**). The reduction of α1(V) and α2(V) chains (which constitute type V collagen) from the cell surface supports the long-standing view that collagen-V nucleates collagen-I containing fibrils. Gene ontology pointed to integral components of the plasma membrane, which included several integrins. Whilst many integrins were detected in all samples there was an absence of integrin α1 and integrin α11 subunit in VPSko cultures (**Figure 6C**). It is well established that integrins are cell-surface molecules that interact extensively with the ECM [45], with integrin α11β1 functioning as a major collagen receptor on fibroblasts [22]. Interestingly, whilst the level of integrin β1 chain detected was also reduced in VPSko cultures, it was not as drastic as the reduction of integrin α11 chain. This is likely due to the promiscuous nature of integrin β1 subunit, that is being able to partner with other α subunits for functions other than collagen-I interaction. The absence of integrin α11 subunit from VPSko cells suggested a link between VPS33B-mediated collagen fibrillogenesis and integrin α11β1.

VPS33B was conspicuous by its absence from cell-surface labeling studies of control samples (**Figure 6D**). We have previously struggled to detect VPS33B protein using mass spectrometry [10] and postulate that its absence could be due to low abundance. Regardless, VPS33B was detected at the cell surface in VPSoe samples along with PLOD3 (**Figure 6D**), previously identified to be delivered by VPS33B [14]. Complementary western blot analysis of total cell lysates showed that expression of integrin α11 subunit is significantly reduced in VPSko cells and is elevated in VPSoe cells (**Figure 6E**). This correlation between VPS33B and integrin α11 expression levels was confirmed at the mRNA level (**Figure 6F**), inferring a link between VPS33B, collagen fibril, and integrin α11 abundance. We verified this by siRNA knockdown of *Itga11* in control and VPSoe fibroblasts, where knockdown efficiency is confirmed by qPCR (**Figure 6G**). siItga11 in both control and VPSoe cells reduced the number of collagen-I fibrils in culture (**Figure 6H**). Thus, even with elevated levels of VPS33B, integrin α11 subunit is required for collagen-I fibrillogenesis. Interestingly, knocking-down integrin α11 subunit increased exogenous collagen-I uptake as demonstrated by flow cytometry (**Figure 6I**). Thus, we propose that both VPS33B and integrin α11 are involved in directing collagen-I protomers to the formation of collagen-I fibrils, and not mediators of collagen-I endocytosis.

Integrin α11 subunit is localized to the fibroblastic focus of idiopathic pulmonary fibrosis

We next investigated if VPS33B and integrin α11 are involved in human fibrotic diseases, where accumulation of collagen fibrils is a disease hallmark. Lung fibroblasts isolated from control individuals or individuals suffering from IPF were cultured and the mRNA levels of *COL1A1, ITGA11,* and *VPS33B* determined (**Figure 7A**). Although *COL1a1* was similarly expressed, the levels of *VPS33B* and *ITGA11* were significantly increased in IPF fibroblasts compared to control. We next assessed the endocytic capacities of collagen-I, and found that IPF fibroblasts endocytosed significantly more Cy5-labelled exogenous collagen-I when compared to control fibroblasts (**Figure 7B**, left); uptake of Cy3-labelled exogenous collagen-I was also elevated, albeit not significantly (**Figure 7B**, right). Subsequent culture of cells that have taken up exogenous collagen-I revealed that IPF fibroblasts made more fluorescently-labeled collagen fibrils, indicating enhanced recycling of endocytosed collagen-I to fibrillogenic sites (**Figure 7C**). This *in vitro* observation was also represented in IPF pathology, where patient-derived lung samples showed enrichment of collagen-I which overlaps with both integrin α11 subunit and VPS33B within the IPF hallmark lesion (termed the fibroblastic focus [23], encircled by red dotted line, **Figure 7D**). In sites of emerging fibrotic remodeling, integrin α11 subunit and VPS33B are also detected (red asterisks, **Figure S6A**; additional N = 3 IPF specimens, **Figure S6B**), indicating a role for VPS33B/integrin α11 chain in collagen-I deposition in the context of IPF. In control lung samples, collagen-I and VPS33B were present whereas integrin α11 subunit was detected at negligible levels (**Figure 7F, Figure S6C**). These results indicate that the proteins required for assembly of collagen into fibrils (i.e. integrin α11, VPS33B) are present at the fibrotic fronts of IPF, and that enhanced endocytic recycling of collagen-I by fibroblasts may be a disease-potentiating mechanism in IPF.

VPS33b and integrin α11 subunit levels are elevated in chronic skin wounds

Fibrosis has been described as a dysregulation of the normal wound-healing process [46]. Thus, we postulated that chronic skin wounds, a similar chronic inflammation unresolved wound-healing condition, may also share a similar pathological molecular signature as IPF. Immunohistochemical (IHC) staining revealed that in human chronic skin wounds, the expressions of integrin α11 subunit and VPS33B are elevated when compared to normal skin areas taken from the same patient (**Figure 8**), where expression was evident at both the fibrotic wound margins and perivascular tissues, indicative of areas under constant collagen remodeling, suggestive of a similar dysregulation in collagen fibrillogenesis pathway utilizing VPS33B and integrin α11.

Discussion

In this study we have identified the endosome as a major protagonist in controlling the assembly of type I collagen-containing fibrils, through a circadian rhythmic endocytic recycling route. Integral to this process is VPS33B (a circadian clock-regulated endosomal tethering molecule) and integrin α11 subunit (a collagen-binding transmembrane receptor when partnered with integrin β1 subunit). Collagen-I co-traffics with VPS33B in a VIPAS-containing endosomal compartment that directs collagen-I to sites of fibril assembly, which requires integrin α11. These proteins are all enhanced at active sites of collagen pathologies such as fibrosis and chronic skin wounds, suggestive of a common disease mechanism.

Previous research [15, 47-50] has focused on degradation or signaling as the endpoint of collagen-I endocytosis. However, our results show that collagen-I endocytosis contributes towards fibril formation, which is reduced when endocytosis is inhibited. The decision between degrading or recycling endocytosed collagen may depend on cell type, microenvironment, or collagen type itself. In complex environments, for example a wound, degradation of damaged collagen molecules and recycling of structurally sound collagen molecules could be beneficial. The requirement of VPS33B for fibril assembly confirms the previously-observed rhythmicity of collagen fibril formation by fibroblasts, *in vivo* and *in vitro* [10]. Our discovery that the endosomal system is involved in fibril assembly explains how collagen fibrils can be assembled in the absence of endogenous collagen synthesis, namely, if collagen can be retrieved from the extracellular space. We also observed a delay between maximum uptake and maximum fibril numbers. A possible function of this delay may be to increase the concentration of collagen-I in readiness for more efficient fibril initiation. This is supported by our findings that fibroblasts can utilize a solution of Cy3-coll lower than the 0.4 μg/mL critical concentration for fibril formation to assemble fibrils. Surface biotin-labelling in fibroblasts identified two integrin α subunits (α1 and α11) to be absent in VPSko cells which could not make collagen fibrils. There are four known collagen-binding integrins (α1β1, α2β1, α10β1, α11β1), where α2β1 and α11β1 have affinity for fibrillar collagens and are expressed in fibroblasts *in vivo* [22]. α1β1, in contrast, has been demonstrated *in vitro* to display higher affinity for collagen-IV than collagen-I [51]; it has a wide expression pattern *in vivo,* including basement membrane-associated smooth muscle cells, strongly suggesting that collagen-IV is the major ligand for α1β1 integrin instead. Thus, we focused on the effects of integrin α11 subunit. Interestingly, both α1 and α11 subunits were also reduced in VPSoe cells as detected by mass spectrometry; however, validation by western blots indicated that protein levels of integrin α11 follows VPS33B and collagen fibril numbers. We postulate that this discrepancy detected through MS might be due to integrin dynamics at the plasma membrane. Regardless, VPS33B and integrin α11β1 are essential for targeting collagen-I to fibril assembly and not for the secretion of soluble triple helical collagen-I molecules (i.e. protomers); there is also no evidence to support an active role for VPS33B or α11β1 in collagen endocytosis. We have also demonstrated involvement of VPS33B and integrin α11 subunit in IPF, particularly at sites of disease progression (fibroblastic foci). The presence of integrin α11 subunit confirms a recent study utilizing spatial proteomics to decipher regions of IPF tissues, where in addition to integrin α11 the authors also identified proteins involved in collagen biogenesis specific to the fibroblastic foci [52]. Crucially, in our study, levels of *Col1a1* transcript were unchanged between control and IPF fibroblasts, and IHC demonstrated that collagen-I and VPS33B were prevalent in control lungs. This highlights that the production of collagen-I does not equate to fibril formation; indeed, here VPS33B is required for normal collagen homeostasis in the lung, and it is the combination of VPS33B and integrin α11 subunit, coupled with elevated endocytic recycling of exogenous collagen at the fibroblastic focus, that promotes excess collagen-I fibril assemblies, which is a hallmark of

fibrosis. Furthermore, this relationship between excessive collagen-I, VPS33B, and integrin α11 is also observed in chronic skin wounds, where, similar to lung fibrosis, there is a chronic unresolved wound-healing response; this is suggestive of a common pathological molecular pathway between the two organs, based on enhanced endocytic recycling of collagen-I protomers directed to fibrillogenesis.

The discovery that VPS33B is required for collagen fibril formation but not collagen protomer secretion highlights an important distinction between 'collagen secretion' and 'collagen fibrillogenesis', especially in the context of collagen pathologies (e.g. fibrosis, cancer metastasis). In this context, it is crucial to keep in mind that elevated collagen levels (i.e. collagen transcripts or protomer levels) does not equate to the formation of an insoluble fibrillar network. This distinction was also suggested in a recent study on *Pten*-knockout mammary fibroblasts, where SPARC protein acts via fibronectin to affect collagen fibrillogenesis but not secretion [53]. Of note, whilst SPARC and fibronectin were detected in our previously-reported time-series mass spectrometry analyses [10], they were not circadian clock rhythmic in tendon.

The nucleation of a collagen-I fibril is expected to involve collagen-V, a minor fibril-forming collagen that is necessary for the appearance of collagen-I fibrils *in vivo* [54]. Another long-standing view is that fibronectin may tether collagen to a fibronectin-binding integrin and thereby function as a proteolytically cleavable anchor ([55], reviewed by [56]). Whilst the present study did not focus on the role of fibronectin, IF staining suggested that collagen-I and fibronectin do not completely overlap, and that inhibition of endocytosis did not affect fibronectin fibril numbers. Recent data using cells deficient in fibronectin assembly also suggested that cell-mediated collagen fibrillogenesis can be mediated via integrin α11β1 at the cell membrane, independent of FN (Personal correspondence, Donald Gullberg). Further, biotin-surface labeling mass spectrometry analysis in the present study indicated that the levels of the major fibronectin-binding integrins (integrins α5β1, αVβ3) were not drastically altered between control, VPSko, and VPSoe fibroblasts, indicating that VPS33B controls collagen fibrillogenesis via a fibronectin-independent route. Interestingly, in fibronectin knockout liver cells, exogenous collagen-V could be added to initiate collagen-I fibrillogenesis [57], which suggests that other matrix proteins could also be recycled in a similar manner as collagen-I. Whilst collagen α1(V) could only be detected in surface biotin-labeled control but not VPSko or VPSoe cells, collagen α2(V) was only detected in VPSoe cells. Of note, we have previously shown that collagen α2(V) protein was rhythmic and in phase with collagen α1(I) and collagen α2(I). Whether collagen-V co-traffics with collagen-I in VPS33B vesicle-like compartments to fibrillogenic sites at the cell periphery, or collagen-V is directed to the nucleation site by another route, is an important question to address in future studies. Indeed, endocytic recycling of other factors required for collagen fibrillogenesis (e.g. other nucleation factors such as collagen-V, collagen-XI, integrins), in addition to collagen-I protomers, may be required for fibril assembly by fibroblasts.

Previous research has demonstrated that collagen-I can be secreted within minutes [9], whilst other studies have demonstrated a relatively slow emergence of collagen-I fibrils over days in culture [42]. The delay in assembly of fibrils could be due to the requirement to reach the threshold concentration of collagen needed for fibrillogenesis [25]. Procollagen-I can be converted to collagen-I within the secretory pathway [9, 58] in a process that requires giantin for intracellular N-terminal processing of procollagen-I [59]. Thus, it is possible that triple helical collagen-I protomers destined for fibril formation are: 1) directed straight to a VPS33B-positive compartment from the Golgi apparatus, or 2) quickly secreted and then recaptured, via endocytosis, prior to recycling for fibrillogenesis. Such a mechanism would provide the cell with additional opportunities to determine the number of fibrils, the rate at which they are to be initiated, and their required growth in the extracellular matrix (see Figure 9). We have not identified the specific proteins that mediate collagen-I uptake into the cell, however as inhibition of clathrin-mediated endocytosis did not completely inhibit collagen endocytosis, it is likely that collagen-I has multiple routes to enter the cells [17-20]. With regards to the fibrillogenesis route, although a recent model of VPS33B/VIPAS39 protein complexes does not consider the possibility that VPS33B can associate directly with vesicular membranes [60], our *in vitro* studies suggest its hydrophobic C-terminal region can act as a tail-anchor domain. Importantly, the proposed model of the VPS33B/VIPAS39 complex suggests the C-terminus of VPS33B would be available for membrane association, consistent with our *in vitro* findings. Further studies will be required to establish whether differences between our models reflect alternative experimental systems, yeast vs. mammalian, and/or the behavior of VPS33B

alone vs. in complex with VIPAS39. We found that endogenous VIPAS39 encases collagen-I in vesicle-like structures, consistent with a model where VPS33B and VIPAS39 co-traffic with collagen-I at the same endosomal compartment (see also [60]). Our data suggest two membrane bound populations of VPS33B, the majority with a hairpin-like structure and cytosolic C-terminus and a minority that spans the membrane with the C-terminus inside the vesicle lumen. Whether or not alternative topologies of VSP33B are associated with different biological functions remains to be determined. Nonetheless, here we have identified a previously unknown role for endocytosis in collagen fibrillogenesis, which is exploited in collagen pathologies.

Author contributions

JC, AP and KEK conceived the project. KEK supervised the experiments. AP, JC, RG performed experiments and interpreted data. JAH consented and biopsied patients. JAH, LD performed experiments, MH, PC performed super-resolution live imaging, analysis, and data interpretation, YL performed electron microscopy and analysis, CZ produced integrin-α11 antibodies and performed antibody screening, DG, RVV, AH provided reagents and data interpretation, and SH and SO'K performed the VPS33B ΔG analysis, topology analyses, and data interpretation. All authors contributed to writing the manuscript.

Acknowledgements

The proteomics was performed at the Biological Mass Spectrometry Facility with the assistance of Stacey Warwood and Ronan O'Cualain, imaging was performed at the Bioimaging Facility, flow cytometry analyses were performed at the Flow Cytometry Facility with the assistance of Gareth Howell, all in the Faculty of Biology, Medicine and Health (University of Manchester). This report is independent research supported by the North West Lung Centre Charity and National Institute for Health Research Clinical Research Facility at Manchester University NHS Foundation Trust. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the North West Lung Centre Charity, National Institute for Health Research or the Department of Health. The authors would like to acknowledge the Manchester Allergy, Respiratory and Thoracic Surgery Biobank and the North West Lung Centre Charity for supporting this project. We would also like to thank the study participants for their contribution. The authors would like to acknowledge and thank Prof David Stephens for his feedback on this manuscript. Some illustrations were created with Biorender.com.

Funding

Wellcome Senior Investigator Award 110126/Z/15/Z (KEK)

MRC Career Development Award MR/W016796/1 (JC)

Wellcome Centre Award 203128/Z/16/Z (KEK)

Wellcome Investigator Award in Science 204957/Z/16/Z (SH)

Norwegian Centre of Excellence grant 223250 (DG)

Norwegian Cancer Society grant 223052 (DG)

Nasjonalföreningen for folkhelsen grant, project 16216 (DG)

Wellcome Institutional Strategic Support fund 204796/Z/16/Z (AR, JW)

Wellcome Translational Partnership Award 209741/Z/17/Z (AR, JW)

Declaration of interests

The authors declare no competing interests.

Materials and Methods

Procurement of human lung tissue and human lung-derived fibroblasts

The use of human lung tissue was approved by the National Health Research Authority with patient consent (NRES14/NW/0260). The specimens used for this study met the criteria for Idiopathic Pulmonary Fibrosis (IPF) diagnosis [61], as previously described [62]. 4 IPF patient samples and 2 control lung samples were used in this

study. The patient-derived fibroblasts, isolated from 5 IPF patient samples and 5 control samples used here were a kind gift from Professor Peter Bitterman and Professor Craig Henke at the University of Minnesota. Briefly, control and IPF-derived fibroblasts (isolated from were established by allowing fibroblasts to propagate from lung tissue explants as previously described [63]. The fibroblasts used in the described experiments were between passages 4 and 6.

Cell culture

Unless otherwise stated, all cell culture reagents were obtained from Gibco, and all cells were maintained at 37 °C, 5% CO_2 in a humidified incubator. Immortalized mouse tail tendon fibroblasts (iTTF, isolated from 8-10 weeks old C57BL/6 male mice [10]), and NIH3T3 cells with Dendra2-tagged collagen-I expression (Dendra2-3T3 [42]) were cultured at in DMEM/F12 with sodium bicarbonate and L-glutamine (supplemented with 10% fetal calf serum (FCS), 200 µM L-ascorbate-2-phosphate (Sigma), and 10,000 U/mL penicillin/streptomycin), and DMEM with sodium bicarbonate and L-glutamine (supplemented with 10% new born calf serum, 200 µM L-ascorbate-2-phosphate, and 10,000 U/mL penicillin/streptomycin) media respectively. HEK293T cells were cultured in DMEM with sodium bicarbonate and L-glutamine, supplemented with 10% FCS.

CRISPR-Cas9-mediated knockout

iTTFs were treated with CRISPR-Cas9 to delete *VPS33B* gene as previously described [10]. Gene knockout was confirmed by western blotting and qPCR.

Constructs

The cDNA for mouse VPS33b (Uniprot: P59016) was purchased from Genscript (OMu07060D). Sec61β modified with a C-terminal OPG2 tag (residues 1-18 of bovine rhodopsin (Uniprot: P02699) was as previously described [64]. An artificial N-glycosylation site (VPS33b-A606T) and OPG2-tag (VPS33b-OPG2) were incorporated into parent VPS33b by site-directed mutagenesis (Stratagene QuikChange, Agilent Technologies) using the relevant forward and reverse primers (Integrated DNA Technologies; see reagents table) and confirmed by DNA sequencing (GATC, Eurofins Genomics). Linear DNA templates were generated by PCR and mRNA transcribed using T7 polymerase or SP6 RNA polymerase as appropriate (see reagents table). In order to improve the signal intensity of all radiolabeled proteins, an additional five methionine residues (5M) were appended to the C-terminus of all VPS33b linear DNA templates by PCR. All primer combinations used for mutagenesis and PCR are listed in the tables below.

Transfection and stable infection of over-expression vectors in cells

Three different vectors were used to stably over-express VPS33B protein in iTTF and 3T3 cells – untagged VPS33B overexpression (with red fluorescent protein expression as selection), N-terminus BFP-tagged VPS33B overexpression, and C-terminus BFP-tagged VPS33B overexpression. All proteins were cloned into pLV V5-Luciferase expression vector, where luciferase was excised from the backbone prior to gel purification, or into pCMV expression vector, followed by Gibson Assembly® (NEB) [65]. Lentiviral particles were generated In HEK293T cells, and cells were infected in the presence of 8 µg/mL polybrene (Millipore). Cells were then sorted using Flow Cytometry to isolate RFP-positive or BFP-positive cells. Alternatively, cells were treated with puromycin (Sigma) or G418 (Sigma) depending on selection marker on the vectors.

Fluorescence labeling of collagen-I

Cy3 or Cy5 NHS-ester dyes (Sigma) were used to fluorescently label rat tail collagen-I (Corning), using a previously described method [24]. Briefly, 3 mg/mL collagen-I gels were made, incubated with 50 mM borate buffer (pH 9) for 15 min, followed with incubation with either Cy3-ester or Cy5-ester (Sigma) dissolved in borate buffer in dark overnight at 4 °C, gently rocking. Dyes were then aspirated and 50 mM Tris (pH 7.5) were added to quench the dye reaction, and incubated in the dark rocking for 10 min. Gels were washed with PBS (with calcium and magnesium ions) 6x, incubating for at least 30 min each wash. The gels were then resolubilized using 500 mM acetic acid and dialyzed in 20 mM acetic acid.

RNA isolation and quantitative real-time PCR

RNA was isolated using TRIzol Reagent (Thermo Fisher Scientific) following manufacturer's protocol, and concentration was measured using a NanoDrop OneC (Thermo Fisher Scientific). Complementary DNA was

synthesized from 1 μg RNA using TaqMan™ Reverse Transcription kit (Applied Biosystems) according to manufacturer's instructions.

SensiFAST SYBR kit reagents were used in qPCR reactions. Primer sequences can be found in reagents section.

Protein extraction and western blotting

For lysate experiments, proteins were extracted using urea buffer (8 M urea, 50 mM Tris-HCl pH7.5, supplemented with protease inhibitors and 0.1% β-mercaptoethanol). For conditioned media (CM) media, cells were plated out at 200,000 cells per 6-well plate, and left for 48-72 h before 250 μl was sampled. Samples were mixed with 4xSDS loading buffer with 0.1% β-mercaptoethanol and boiled at 95°C for 5 min. The proteins were separated on either NuPAGE Novex 10% polyacrylamide Bis–Tris gels with 1XNuPAGE MOPS SDS buffer or 6% Tris-glycine gels with 1XTris-glycine running buffer (all Thermo Fisher Scientific), and transferred onto polyvinylidene difluoride (PVDF) membranes (GE Healthcare). The membranes were blocked in 5% skimmed milk powder in PBS containing 0.01% Tween 20. Antibodies were diluted in 2.5% skimmed milk powder in PBS containing 0.01% Tween 20. The primary antibodies used were: rabbit polyclonal antibody (pAb) to collagen-I (1:1,000; Gentaur), mouse mAb to vinculin (1:1000; Millipore), rabbit pAb to integrin α11 subunit (1:1000; see [66]) and mouse mAb to VPS33B (1:500; Proteintech). Horseradish-peroxidase-conjugated antibodies and Pierce ECL western blotting substrate (both from Thermo Fisher Scientific) were used and reactivity was detected on GelDoc imager (Biorad). Alternatively, Licor goat-anti-mouse 680, mouse-anti-rabbit 800 were used and reactivity detected on an Odyssey Clx imager.

***In vitro* ER import assays**

Translation and ER import assays were performed in nuclease-treated rabbit reticulocyte lysate (Promega) as previously described [35, 36, 67]: briefly, in the presence of EasyTag EXPRESS ^{35}S Protein Labelling Mix containing [^{35}S] methionine (Perkin Elmer) (0.533 MBq; 30.15 TBq/mmol), 25 μM amino acids minus methionine (Promega), 6.5% (v/v) nuclease-treated ER-derived canine pancreatic microsomes (from stock with OD_{280}=44/mL) or an equivalent volume of water, and 10% (v/v) of *in vitro* transcribed mRNA (~1 μg/μL) encoding relevant precursor proteins. Translation reactions for Sec61βOPG2 (20 μL; 1X) were performed at 30°C for 30 min whereas VPS33b and its variants (20 μL; 2X) were performed at 30°C for 1 h. Irrespective of the precursor protein, all translation reactions were finished by incubating with 0.1 mM puromycin for 10 min at 30°C to ensure translation termination and the ribosomal release of newly synthesized proteins prior to analysis. Microsomal membrane-associated fractions were recovered by centrifugation through an 80 μL high-salt cushion (0.75 M sucrose, 0.5 M KOAc, 5 mM $Mg(OAc)_2$ and 50 mM HEPES-KOH, pH 7.9) at 100,000 g for 10 min at 4°C, the pellet suspended directly in SDS sample buffer and, where indicated, treated with 1000 U of endoglycosidase Hf (New England Biolabs, P0703S). All samples were denatured for 10 min at 70°C and resolved by SDS–PAGE (10% or 16% PAGE, 120 V, 120-150 min). Gels were fixed for 5 min (20% MeOH, 10% AcOH), dried for 2 h at 65°C, and radiolabeled products were visualized using a Typhoon FLA-700 (GE Healthcare) following exposure to a phosphorimaging plate for 24–72 h. Images were opened using Adobe Photoshop and annotated using Adobe Illustrator.

Flow cytometry and imaging

Cy3- or Cy5-tagged collagen was added to cells and incubated at 37°C, 5% CO_2 in a humidified incubator for predetermined lengths of times before washing in PBS, trypsinized, spun down at 2500 rpm for 3 min at 4 °C, and resuspended in PBS on ice. Cells were analyzed for Cy3- or Cy5-collagen uptake using LSRFortessa (BD Biosciences). For imaging, cells were prepared as described, and analyzed on Amnis® ImageStream®^{X}Mk II (Luminex).

Decellularization of cells to obtain extracellular matrix

Cells were seeded out at 50,000 in a 6-well plate and cultured for 7 days before decellularization. Extraction buffer (20mM NH_4OH, 0.5% Triton X-100 in PBS) was gently added to cells and incubated for 2 minutes. Lysates were aspirated and the matrix remaining in the dish were washed gently twice with PBS, before being scrapped off into ddH_2O for further processing.

Hydroxyproline assay

Samples were incubated overnight in 6 M HCl (diluted in ddH_2O (Fluka); approximately 1 mL per 20 mg of sample) in screw-top tubes (StarLab) in a sand-filled heating block at 100 °C covered with aluminum foil. The tubes were then allowed to cool down and centrifuged at 12,000 x*g* for 3 min. Hydroxyproline standards were prepared (starting at 0.25 mg/mL; diluted in ddH2O) and serially diluted with 6 M HCl. Each sample and standard (50 µL) were transferred into fresh Eppendorf tubes, and 450 µL chloramine T reagent (0.127 g chloramine T in 50% N-propanol diluted with ddH2O; brought up to 10 mL with acetate citrate buffer (120 g sodium acetate trihydrate, 46 g citric acid, 12 mL glacial acetic acid, 34 g sodium hydroxide) adjusted to pH 6.5 and then made to 1 litre with dH_2O; all reagents from Sigma) was added to each tube and incubated at room temperature for 25 min. Ehrlich's reagent (500 µL; 1.5 g 4-dimethylaminobenzaldehyde diluted in 10 mL N-propanol:perchloric acid (2:1)) was added to each reaction tube and incubated at 65 °C for 10 min and then the absorbance at 558 nm was measured for 100 µL of each sample in a 96-well plate format.

Electron microscopy

Unless otherwise stated, incubation and washes were done at room temperature. Cells were plated on top of ACLAR films and allows to deposit matrix for 7 days. The ACLAR was then fixed in 2% glutaraldehyde/100mM phosphate buffer (pH 7.2) for at least 2 h and washed in ddH_2O 3 x 5 min. The samples were then transferred to 2% osmium (v/v)/1.5% potassium ferrocyanide (w/v) in cacodylate buffer (100 mM, pH 7.2) and further fixed for 1 h, followed by extensive washing in ddH_2O. This was followed by 40 minutes of incubation in 1% tannic acid (w/v) in 100 mM cacodylate buffer, and then extensive washes in ddH_2O. and placed in 2% osmium tetroxide for 40 min. This was followed by extensive washes in ddH_2O. Samples were incubated with 1% uranyl acetate (aqueous) at 4 °C for at least 16 h, and then washed again in ddH_2O. Samples were then dehydrated in graded ethanol in the following regime: 30%, 50%, 70%, 90% (all v/v in ddH_2O) for 8 min at each step. Samples were then washed 4 x 8 min each in 100% ethanol, and transferred to pure acetone for 10 min. The samples were the infiltrated in graded series of Agar100Hard in acetone (all v/v) in the following regime: 30% for 1 h, 50% for 1 h, 75% for overnight (16 h), 100% for 5 h. Samples were then transferred to fresh 100% Agar100Hard in labeled moulds and allowed to cure at 60 °C for 72 h. Sections (80 nm) were cut and examined using a Tecnai 12 BioTwin electron microscope.

Surface biotinylation-pulldown mass spectrometry

Cells were grown in 6 well plates for 72 h prior to biotinylating. Briefly, cells around 90% confluence were kept on ice and washed in ice-cold PBS, following with incubation with ice-cold biotinylating reagent (prepared fresh, 200 µg/mL in PBS, pH 7.8) for 30 min at 4 °C gently shaking. Cells were then washed twice in ice-cold TBS (50 mM Tris, 100 mM NaCl, pH 7.5), and incubated in TBS for 10 min at 4 °C. Cells were then lysed in ice-cold 1% TritonX (in PBS, with protease inhibitors) and lysates cleared by centrifugation at 13,000 x*g* for 10 min at 4 °C. Supernatant was then transferred to a fresh tube and 1/5 of the lysates were kept as a reference sample. Streptavidin-Sepharose beads are then aliquoted into each sample and incubated for 30 min at 4 °C rotating. Beads were then washed 3x in ice cold PBS supplemented with 1% TritonX, followed by one final wash in ice cold PBS and boiled at 95 °C for 10 min in 2x sample loading buffer, followed by centrifugation at top speed for 5 min follow by sample preparation for Mass Spectrometry analysis. The protocol used for sample preparation was as described previously [62]. Mass spectrometry results files were exported into Proteome Discoverer (PD) for identification and spectral counting. All searches included the fixed modification for carbamidomethylation on cysteine residues resulting from IAA treatment to prevent cysteine bonding. The variable modifications included in the search were oxidized methionine (monoisotopic mass change, +15.955 Da); hydroxylation of asparagine, aspartic acid, proline, or lysine (monoisotopic mass change, +15.955 Da); and phosphorylation of threonine, serine, and tyrosine (79.966 Da). A maximum of 2 missed cleavages per peptide was allowed. The minimum precursor mass was set to 350Da with a maximum of 5000. Precursor mass tolerance was set to 10ppm, fragment mass tolerance was 0.02Da and minimum peptide length was 6. Peptides were searched against the Swissprot database using Sequest HT with a maximum false discovery rate of 1%. Proteins required a minimum FDR of 1% and were filtered to remove known contaminants and to have at least 2 unique peptides. Missing values were assumed to be due to low abundance. For spectral counting, enriched samples were normalized and spectral matches were compared. Peptide spectral matches were only included if they were calculated to have a q-value of less than 0.01.

607 **Imaging and immunofluorescence**

608 For fluorescence live imaging of Cy3-coll internalisation, iTTF cells were seeded for 24 h onto Ibidi μ-plates and
609 stained using CellMask™ Green Plasma Membrane Stain (Invitrogen C37608) according to the manufacturers
610 protocol. Cy3-coll was added to cells and images were collected in a 37°C chamber using a Zeiss 3i spinning
611 disk (CSU-X1, Yokagowa) confocal microscope with a 63x/1.40 Plan-Apochromat oil objective. 7.5 μm z-stacks
612 with a step size of 0.5 μm were captured at 2 minute intervals using 488 nm (100 power, 150 ms exposure)
613 and 561 nm (100 power, 100 ms exposure) lasers with SlideBook 6.0 software (3i) and a front illuminated
614 Prime sCMOS camera. Imaris 9.9.1 software was then used to generate 3D reconstructions.

615 For fixed immunofluorescence imaging, cells plated on coverslips or Ibidi μ-plates were fixed with 100%
616 methanol at -20°C and then permeabilized with 0.2% Triton-X in PBS. Primary antibodies used were as follows:
617 rabbit pAb collagen-I (1:400, Gentaur OARA02579), rabbit pAb VIPAS (1:50, Proteintech 20771-1-AP), mouse
618 mAb FN1 (1:400, Sigma F6140 or Abcam ab6328). Secondary antibodies conjugated to Alexa-488, Alexa-647,
619 Cy3, and Cy5 were used (ThermoScientific), and nucleus were counterstained with DAPI (Sigma). Coverslips
620 were mounted using Fluoromount G (Southern Biotech). Images were collected using a Leica SP8 inverted
621 confocal microscope (Leica) using an x63/0.50 Plan Fluotar objective. The confocal settings were as follows:
622 pinhole, 1 Airy unit; scan speed, 400 Hz bidirectional; format 1024 x 1024 or 512 x 512. Images were collected
623 using photomultiplier tube detectors with the following detection mirror settings: DAPI, 410-483; Alexa-488,
624 498-545nm; Cy3, 565-623 nm; Cy5, 638-750 nm using the 405 nm, 488 nm, 540 nm, 640 nm laser lines.
625 Alternatively, images were collected using EVOS™ M7000 (Thermo Fisher) using 40x/1.3na oil objective.
626 Images were collected in a sequential manner to minimize bleed-through between channels.

627 For fluorescence live-imaging, cells were plated onto Ibidi μ-plates and imaged using Zeiss LSM880 NLO (Zeiss).
628 For split-GFP experiment, cells were seeded for 24 h onto Ibidi μ-plates before imaging with an Olympus
629 IXplore SpinSR (Olympus) with 100x oil magnification. Prior to imaging, media was changed to FluoroBrite
630 media with the appropriate supplements.

631 **Immunohistochemistry**

632 Human lung samples were formalin-fixed and paraffin-embedded (FFPE). 5-μm FFPE sections were obtained
633 and mounted onto Superfrost® Plus (Thermo Scientific™) slides and subjected to antigen heat retrieval using
634 citrate buffer (Abcam, ab208572), in a pre-heated steam bath for 20 minutes, before cooling to room
635 temperature in a water bath for 20 min. Slides were then treated with 3-4% hydrogen peroxide (Leica
636 Biosystems RE7101) for 10 min, blocked in SuperBlock™ (TBS) blocking buffer for a minimum of 1 h (Thermo
637 Scientific™; 37581), and probed with primary antibodies overnight at 4°C in 10% SuperBlock™ solution in Tris
638 Buffered Saline Tween 20 solution (TBS-T, pH 7.6). Primary antibodies used were as follows: Collagen-I A1/A2
639 (Rockland Immunochemicals Inc 600-401-103.0.5), integrin-α11 (integrin α11 mAb 210F4 [68]), VPS33B (Atlas
640 antibodies).

641 After overnight incubation, specimens were subjected to Novolink Polymer Detection Systems (Leica
642 Biosystems RE7270-RE, as per the manufacturer's recommendations), with multiple TBS-T washes. Sections
643 were developed for 5 min with DAB Chromagen (Cell Signaling Technology®, 11724/5) before being
644 counterstained with haematoxylin for 1 minute, followed by acid alcohol and blueing solution application.
645 Slides were dehydrated through sequential ethanol and xylene before being cover-slipped with Permount
646 mounting medium (Thermo Scientific™, SP15).

647 **Histological imaging**

648 Stained slides were imaged using a DMC2900 Leica camera along with Lecia Application Suite X software
649 (Leica).

650 **QUANTIFICATION AND STATISTICAL ANALYSIS**

651 Data are presented as the mean ± s.e.m. unless otherwise indicated in the figure legends. The sample number
652 'N' indicates the number of independent biological samples in each experiment, and 'n' indicates the number
653 of technical repeats, and are indicated in the figure legends. Data were analyzed as described in the legends.
654 The data analysis was not blinded, apart from quantification of fibril numbers over time, and differences were
655 considered statistically significant at $P<0.05$, using Student's t-test or One-way Anova, unless otherwise stated

656 in the figure legends. Analyses were performed using Graphpad Prism 8 software. Significance levels are: *
657 *P*<0.05; ** *P*<0.01; *** *P*<0.005; ****P<0.0001. For periodicity, analysis was performed using the MetaCycle
658 package[69] in the R computing environment [70] with the default parameters.

659 **Reagents**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit polyclonal antibody (pAb) to collagen-I	Aviva Systems Biology	Cat# OARA02579, RRID:AB_10873334
mouse mAb to mouse vinculin	Sigma-Aldrich	Cat# V9131, RRID:AB_477629
rabbit pAb to mouse integrin α11 subunit	This paper	N/A
mouse mAb to mouse VPS33B	Proteintech	Cat# 12195-1-AP, RRID:AB_2215198
Rabbit pAb to mouse VIPAS	Proteintech	Cat# 20771-1-AP, RRID:AB_10695764
Mouse mAb to FN1	Sigma-Aldrich	Cat# F6140, RRID:AB_476981
Mouse mAb to FN1	Abcam	Cat# ab6328, RRID:AB_305428
Rabbit pAb to human VPS33B	Atlas antibodies	Cat# HPA040415, RRID:AB_10795419
Mouse mAb integrin α11 210F4	DOI: 10.1002/cjp2.148	N/A
Rabbit pAb to Collagen-I	Rockland Immunochemicals	600-401-103.0.5, RRID:AB_217595
IRDye® 680RD_Goat anti-Mouse_IgG (H + L)	Li-Cor	Cat# 925-68070, RRID:AB_2651128
IRDye® 800RD_Goat anti-Rabbit_IgG (H + L) Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	Li-Cor Thermo Fisher Scientific	Cat# 925-32211, RRID:AB_2651127 Cat# 32430, RRID:AB_1185566
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	Thermo Fisher Scientific	Cat# 32460, RRID:AB_1185567
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Cyanine3	Thermo Fisher Scientific	Cat# A10520, RRID:AB_2534029
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Cyanine5	Thermo Fisher Scientific	Cat# A10523, RRID:AB_2534032
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Cyanine3	Thermo Fisher Scientific	Cat# A10521, RRID:AB_2534030
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21236, RRID:AB_2535805
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-11029, RRID:AB_2534088
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-11034, RRID:AB_2576217
Bacterial and virus strains		
pLV-V5-VPS33B-GFP1-10	This paper	N/A
pLV-V5-VPS33B-BFP	This paper	N/A
pLV-V5-BFP-VPS33B	This paper	N/A
pCMV-SBP-GFP11-COL1A1	This paper	N/A
pcDNA3.1+/C-(K)-DYK	This paper	N/A
Biological samples		
Human control lung tissues	Manchester University NHS Foundation Trust	NRES14/NW/0260
Human IPF lung tissues	Manchester University NHS Foundation Trust	NRES14/NW/0260

Human chronic skin wound tissue	Manchester University NHS Foundation Trust	NRES 18/NW/0847
Human healthy skin wound tissue	Manchester University NHS Foundation Trust	NRES 18/NW/0847
Chemicals, peptides, and recombinant proteins		
Recombinant S. pyogenes Cas9 nuclease protein	IDT	1081059
Rat tail collagen-I	Corning	354259
Dyngo®4a	Abcam	Ab120689
TRIzol reagent	Invitrogen	15596-018
Cy®3 NHS Ester	Sigma	GEPA13101
Cy®5 NHS Ester	Sigma	GEPA15101
Critical commercial assays		
Gibson Assembly Cloning kit	NEB	E5510S
Site-Directed Mutagenesis QuikChange kit	Agilent	200513
Novolink Polymer Detection Systems	Leica Biosystems	RE7270-RE
Rabbit reticulocyte lysate, nuclease treated	Promega	L4960
TaqMan™ Reverse Transcription kit	Applied Biosystems	N8080234
SensiFAST™SYBR® No-ROX kit	Bioline	BIO-98005
Experimental models: Cell lines		
NIH3T3-Dendra2-Col I	This paper	N/A
Immortalized mouse tail tendon fibroblasts Primary human IPF lung fibroblasts	This paper University of Minnesota	N/A N/A
Primary human control lung fibroblasts	University of Minnesota	N/A
Oligonucleotides		
siCol1a1, Mission® esiRNA	Merck	EMU069551
siItga11, Mission® esiRNA	Merck	EMU042761
Primers for site-directed mutagenesis	See primer list	N/A
Primers for transcription templates Primers for qPCR reactions	See primer list See primer list	N/A N/A
Recombinant DNA		
Vps33b mouse cDNA	Genscript	Omu07060D
Sec61β modified with a C-terminal OPG2 tag	This paper	N/A
VPS33b A-606-T	This paper	N/A
VPS33b-OPG2	This paper	N/A
VPS33b-Sec61βTMD	This paper	N/A
VPS33b-Sec61βTMD-OPG2	This paper	N/A
Software and algorithms		
Fiji ImageJ	doi:10.1038/nmeth.2019	https://imagej.net/software/fiji/
MetaCycle	doi:10.1093/bioinformatics/btw405	https://github.com/gangwug/MetaCycle
Scaffold Proteome Software	doi:10.1002/pmic.200900437	https://www.proteomesoftware.com/products
Graphpad Prism 8	https://www.graphpad.com/scientific-software/prism/	https://www.graphpad.com/scientific-software/prism/
Biorender	https://biorender.com/	https://biorender.com/

Primer list for cDNAs generated by site-directed mutagenesis.

New cDNA	Cloned cDNA	Vector	Forward primer (5'-3')	Reverse primer (5'-3')
VPS33b A-606-T	VPS33b	pcDNA3.1+/C-(K)-DYK	ACTGCTGTTACAAACAGTACCCGCCTCATGGAAGCC	GGCTTCCATGAGGCGGGTACTGTTTGTAACAGCAGT
VPS33b-OPG2	VPS33b	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTATTAGCCTGTTTTGTTGCTGAATGGTACGTAGAAGTTTGGTCCTTCTGTTCCGTTGGATTTCACCTCACTCATGGCTTC
VPS33b-Sec61βTMD	VPS33b	pcDNA3.1+/C-(K)-DYK	GTATTGGTTATGTGTCTTCTGTTCATCGCTTCTGTATTTATGTTGCACATTTGGGGCAAGTACACTCGTTCGTAGCTGCGCCTCATCTTGGTGGTGTTCC	CTACGAACGAGTGTACTTGCCCCAAATGTGCAACATAAATACAGAAGCGATGAACAGAAGACACATAACCAATACTGACTCACTGGAAGCCTTGTCTTCC
Chimera 3 (VPS33b-Sec61βTMD-OPG2)	VPS33b-Sec61βTMD	pcDNA3.1+/C-(K)-DYK	AACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATAGCTGCGCCTCATCTTGGTGGTGTTCCTG	CTATTAGCCTGTTTTGTTGCTGAATGGTACGTAGAAGTTTGGTCCTTCTGTTCCGTTCGAACGAGTGTACTTGCCCCAAATGTGCAAC

Primer list for PCR reactions used to create transcription templates.

Substrate	cDNA	Vector	Forward primer (5'-3')	Reverse primer (5'-3')	Promoter
414-564-5M	VPS33b	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTACATCATCATCATCATTGACTCACTGGAAGCCTTGTC	SP6
414-587-5M	VPS33b	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTACATCATCATCATCATCAGGAAGCGCAGGGCTGATAT	SP6
414-FL-5M	VPS33b	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTACATCATCATCATCATGGATTTCACCTCACTCATGGCTTC	SP6
414-N603-FL-5M	VPS33b-N603	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGATGACGAC	CTACATCATCATCATCATGGATTTCACCTCACTCATGGCTTC	SP6
414-FL-OPG2-5M	VPS33b-OPG2	pcDNA3.1+/C-(K)-DYK	GCCAACGGAACAGAAGGACCAAACTTCTACGTACCATTCAGCAACAAAACAGGCTAATCCGATTACAAGGAT	CTACATCATCATCATCATGCCTGTTTTGTTGCTGAATGGTACGTAGAAGTTTGGTCCTTCTGTTCCGTT	SP6

			GACGAC		
1-564-5M	VPS33b	pcDNA3.1+/C-(K)-DYK	CGCAAATGGGCGGTAGGCGTG	CTACATCATCATCATCATTGACTCACTGGAAGCCTTGTC	T7
1-587-5M	VPS33b	pcDNA3.1+/C-(K)-DYK	CGCAAATGGGCGGTAGGCGTG	CTACATCATCATCATCATCAGGAAGCGCAGGGCTGATAT	T7
1-FL-5M	VPS33b	pcDNA3.1+/C-(K)-DYK	CGCAAATGGGCGGTAGGCGTG	CTACATCATCATCATCATGGATTTCACCTCACTCATGGCTTC	T7
1-N603-5M	VPS33b-N603	pcDNA3.1+/C-(K)-DYK	CGCAAATGGGCGGTAGGCGTG	CTACATCATCATCATCATGGATTTCACCTCACTCATGGCTTC	T7
1-FL-OPG2-5M	VPS33b-OPG2	pcDNA3.1+/C-(K)-DYK	CGCAAATGGGCGGTAGGCGTG	CTACATCATCATCATCATGCCTGTTTTGTTGCTGAATGGTACGTAGAAGTTTGGTCCTTCTGTTCCGTT	T7

Primer list for qPCR reactions.

Gene (ms – mouse, hu – human)	**Forward primer (5'-3')**	**Reverse primer (5'-3')**
msCol1a1	AGAGCATGACCGATGGATTC	AGGCCTCGGTGGACA
msItga11	AGATGTCGCAGACTGGCTTT	CCCTAGGTATGCTGCATGGT
msRplp0	ACTGGTCTAGGACCCGAGAAG	CTCCCACCTTGTCTCCAGTC
msGapdh	CAGCCTCGTCCCGTAGACAA	CAATCTCCACTTTGCCACTGC
msVps33b	GCATTCACAGACACGGCTAAG	ACACCACCAAGATGAGGCG
huCOL1A1	GGGATTCCCTGGACCTAAAG	GGAACACCTCGCTCTCCA
huGAPDH	GAGTCAACGGATTTGGTCGT	GACAAGCTTCCCGTTCTCAG
huACTB	GATCATTGCTCCTCCTGAGC	AAAGCCATGCCAATCTCATC
huITGA11	CACGACATCAGTGGCAATAAG	GACCCTTCCCAGGTTGAGTT
huVPS33B	GAGCTGCCTGACTTCTCCAT	GCTTGTCTACTTCGTGTTGCTG

References

1. Smejkal GB, Fitzgerald C. Revised estimate of total collagen in the human body. International Journal of Proteomics & Bioinformatics. 2017;2:001-2.
2. Craig AS, Birtles MJ, Conway JF, Parry DA. An estimate of the mean length of collagen fibrils in rat tail-tendon as a function of age. Connect Tissue Res. 1989;19(1):51-62. PubMed PMID: 2477190.
3. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol. 2008;20(5):495-501. Epub 2008/07/22. doi: 10.1016/j.ceb.2008.06.008. PubMed PMID: 18640274; PubMed Central PMCID: PMCPMC2577133.
4. Musiime M, Chang J, Hansen U, Kadler KE, Zeltz C, Gullberg D. Collagen Assembly at the Cell Surface: Dogmas Revisited. Cells. 2021;10(3). Epub 2021/04/04. doi: 10.3390/cells10030662. PubMed PMID: 33809734; PubMed Central PMCID: PMCPMC8002325.
5. Zeltz C, Kusche-Gullberg M, Heljasvaara R, Gullberg D. Novel roles for cooperating collagen receptor families in fibrotic niches. Curr Opin Cell Biol. 2023;85:102273. Epub 2023/11/03. doi: 10.1016/j.ceb.2023.102273. PubMed PMID: 37918273.

6. Kadler KE, Holmes DF, Trotter JA, Chapman JA. Collagen fibril formation. Biochem J. 1996;316 (Pt 1):1-11. Epub 1996/05/15. doi: 10.1042/bj3160001. PubMed PMID: 8645190; PubMed Central PMCID: PMCPMC1217307.
7. Trelstad RL, Hayashi K. Tendon collagen fibrillogenesis: intracellular subassemblies and cell surface changes associated with fibril growth. Dev Biol. 1979;71(2):228-42. Epub 1979/08/01. doi: 10.1016/0012-1606(79)90166-0. PubMed PMID: 499658.
8. Birk DE, Trelstad RL. Extracellular compartments in tendon morphogenesis: collagen fibril, bundle, and macroaggregate formation. J Cell Biol. 1986;103(1):231-40. Epub 1986/07/01. doi: 10.1083/jcb.103.1.231. PubMed PMID: 3722266; PubMed Central PMCID: PMCPMC2113791.
9. Canty EG, Lu Y, Meadows RS, Shaw MK, Holmes DF, Kadler KE. Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. J Cell Biol. 2004;165(4):553-63. Epub 2004/05/26. doi: 10.1083/jcb.200312071. PubMed PMID: 15159420; PubMed Central PMCID: PMCPMC2172348.
10. Chang J, Garva R, Pickard A, Yeung CC, Mallikarjun V, Swift J, et al. Circadian control of the secretory pathway maintains collagen homeostasis. Nat Cell Biol. 2020;22(1):74-86. Epub 2020/01/08. doi: 10.1038/s41556-019-0441-z. PubMed PMID: 31907414.
11. Hunter MR, Hesketh GG, Benedyk TH, Gingras AC, Graham SC. Proteomic and Biochemical Comparison of the Cellular Interaction Partners of Human VPS33A and VPS33B. J Mol Biol. 2018;430(14):2153-63. Epub 2018/05/21. doi: 10.1016/j.jmb.2018.05.019. PubMed PMID: 29778605; PubMed Central PMCID: PMCPMC6005816.
12. Gissen P, Johnson CA, Morgan NV, Stapelbroek JM, Forshew T, Cooper WN, et al. Mutations in VPS33B, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis-renal dysfunction-cholestasis (ARC) syndrome. Nat Genet. 2004;36(4):400-4. Epub 2004/03/31. doi: 10.1038/ng1325. PubMed PMID: 15052268.
13. Nezelof C, Dupart MC, Jaubert F, Eliachar E. A lethal familial syndrome associating arthrogryposis multiplex congenita, renal dysfunction, and a cholestatic and pigmentary liver disease. J Pediatr. 1979;94(2):258-60. Epub 1979/02/01. doi: 10.1016/s0022-3476(79)80839-2. PubMed PMID: 762621.
14. Banushi B, Forneris F, Straatman-Iwanowska A, Strange A, Lyne AM, Rogerson C, et al. Regulation of post-Golgi LH3 trafficking is essential for collagen homeostasis. Nat Commun. 2016;7:12111. Epub 2016/07/21. doi: 10.1038/ncomms12111. PubMed PMID: 27435297; PubMed Central PMCID: PMCPMC4961739.
15. Madsen DH, Engelholm LH, Ingvarsen S, Hillig T, Wagenaar-Miller RA, Kjoller L, et al. Extracellular collagenases and the endocytic receptor, urokinase plasminogen activator receptor-associated protein/Endo180, cooperate in fibroblast-mediated collagen degradation. J Biol Chem. 2007;282(37):27037-45. Epub 2007/07/12. doi: 10.1074/jbc.M701088200. PubMed PMID: 17623673.
16. Madsen DH, Leonard D, Masedunskas A, Moyer A, Jurgensen HJ, Peters DE, et al. M2-like macrophages are responsible for collagen degradation through a mannose receptor-mediated pathway. J Cell Biol. 2013;202(6):951-66. Epub 2013/09/11. doi: 10.1083/jcb.201301081. PubMed PMID: 24019537; PubMed Central PMCID: PMCPMC3776354.
17. Madsen DH, Jurgensen HJ, Siersbaek MS, Kuczek DE, Grey Cloud L, Liu S, et al. Tumor-Associated Macrophages Derived from Circulating Inflammatory Monocytes Degrade Collagen through Cellular Uptake. Cell Rep. 2017;21(13):3662-71. Epub 2017/12/28. doi: 10.1016/j.celrep.2017.12.011. PubMed PMID: 29281816; PubMed Central PMCID: PMCPMC5753792.
18. Rainero E. Extracellular matrix endocytosis in controlling matrix turnover and beyond: emerging roles in cancer. Biochem Soc Trans. 2016;44(5):1347-54. Epub 2016/12/03. doi: 10.1042/BST20160159. PubMed PMID: 27911717.
19. Lee TH, McKleroy W, Khalifeh-Soltani A, Sakuma S, Lazarev S, Riento K, et al. Functional genomic screen identifies novel mediators of collagen uptake. Mol Biol Cell. 2014;25(5):583-93. Epub 2014/01/10. doi: 10.1091/mbc.E13-07-0382. PubMed PMID: 24403604; PubMed Central PMCID: PMCPMC3937085.
20. Leitinger B. Discoidin domain receptor functions in physiological and pathological conditions. Int Rev Cell Mol Biol. 2014;310:39-87. Epub 2014/04/15. doi: 10.1016/B978-0-12-800180-6.00002-5. PubMed PMID: 24725424; PubMed Central PMCID: PMCPMC4021107.
21. Tiger CF, Fougerousse F, Grundstrom G, Velling T, Gullberg D. alpha11beta1 integrin is a receptor for interstitial collagens involved in cell migration and collagen reorganization on mesenchymal nonmuscle cells. Dev Biol. 2001;237(1):116-29. Epub 2001/08/24. doi: 10.1006/dbio.2001.0363. PubMed PMID: 11518510.
22. Zeltz C, Gullberg D. The integrin-collagen connection - a glue for tissue repair? J Cell Sci. 2016;129(6):1284. Epub 2016/07/22. doi: 10.1242/jcs.188672. PubMed PMID: 27442113.
23. Herrera J, Forster C, Pengo T, Montero A, Swift J, Schwartz MA, et al. Registration of the extracellular matrix components constituting the fibroblastic focus in idiopathic pulmonary fibrosis. JCI Insight. 2019;4(1).

Epub 2019/01/11. doi: 10.1172/jci.insight.125185. PubMed PMID: 30626754; PubMed Central PMCID: PMCPMC6485370.
24. Doyle AD. Fluorescent Labeling of Rat-tail Collagen for 3D Fluorescence Imaging. Bio Protoc. 2018;8(13). Epub 2018/08/17. doi: 10.21769/BioProtoc.2919. PubMed PMID: 30112424; PubMed Central PMCID: PMCPMC6089535.
25. Kadler KE, Hojima Y, Prockop DJ. Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process. J Biol Chem. 1987;262(32):15696-701. Epub 1987/11/15. PubMed PMID: 3316206.
26. Huang D, Sun G, Hao X, He X, Zheng Z, Chen C, et al. ANGPTL2-containing small extracellular vesicles from vascular endothelial cells accelerate leukemia progression. J Clin Invest. 2021;131(1). Epub 2020/10/28. doi: 10.1172/JCI138986. PubMed PMID: 33108353; PubMed Central PMCID: PMCPMC7773400.
27. Liang Z, Liu Z, Cheng C, Wang H, Deng X, Liu J, et al. VPS33B interacts with NESG1 to modulate EGFR/PI3K/AKT/c-Myc/P53/miR-133a-3p signaling and induce 5-fluorouracil sensitivity in nasopharyngeal carcinoma. Cell Death Dis. 2019;10(4):305. Epub 2019/04/05. doi: 10.1038/s41419-019-1457-9. PubMed PMID: 30944308; PubMed Central PMCID: PMCPMC6447525.
28. Wang C, Cheng Y, Zhang X, Li N, Zhang L, Wang S, et al. Vacuolar Protein Sorting 33B Is a Tumor Suppressor in Hepatocarcinogenesis. Hepatology. 2018;68(6):2239-53. Epub 2018/05/08. doi: 10.1002/hep.30077. PubMed PMID: 29729199.
29. Magliery TJ, Wilson CG, Pan W, Mishler D, Ghosh I, Hamilton AD, et al. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. J Am Chem Soc. 2005;127(1):146-57. Epub 2005/01/06. doi: 10.1021/ja046699g. PubMed PMID: 15631464.
30. Cabantous S, Terwilliger TC, Waldo GS. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat Biotechnol. 2005;23(1):102-7. Epub 2004/12/08. doi: 10.1038/nbt1044. PubMed PMID: 15580262.
31. McCaughey J, Stevenson NL, Cross S, Stephens DJ. ER-to-Golgi trafficking of procollagen in the absence of large carriers. J Cell Biol. 2019;218(3):929-48. Epub 2018/12/28. doi: 10.1083/jcb.201806035. PubMed PMID: 30587510; PubMed Central PMCID: PMCPMC6400576.
32. Calverley BC, Kadler KE, Pickard A. Dynamic High-Sensitivity Quantitation of Procollagen-I by Endogenous CRISPR-Cas9 NanoLuciferase Tagging. Cells. 2020;9(9). Epub 2020/09/16. doi: 10.3390/cells9092070. PubMed PMID: 32927811; PubMed Central PMCID: PMCPMC7564849.
33. Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, et al. Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature. 2007;450(7172):1026-30. Epub 2007/12/14. doi: 10.1038/nature06387. PubMed PMID: 18075582.
34. Rabu C, Schmid V, Schwappach B, High S. Biogenesis of tail-anchored proteins: the beginning for the end? J Cell Sci. 2009;122(Pt 20):3605-12. doi: 10.1242/jcs.041210. PubMed PMID: 19812306; PubMed Central PMCID: PMCPMC2758799.
35. O'Keefe S, High S, Demangel C. Biochemical and Biological Assays of Mycolactone-Mediated Inhibition of Sec61. Methods Mol Biol. 2022;2387:163-81. doi: 10.1007/978-1-0716-1779-3_16. PubMed PMID: 34643911.
36. O'Keefe S, Zong G, Duah KB, Andrews LE, Shi WQ, High S. An alternative pathway for membrane protein biogenesis at the endoplasmic reticulum. Commun Biol. 2021;4(1):828. Epub 20210701. doi: 10.1038/s42003-021-02363-z. PubMed PMID: 34211117; PubMed Central PMCID: PMCPMC8249459.
37. Zong G, Hu Z, O'Keefe S, Tranter D, Iannotti MJ, Baron L, et al. Ipomoeassin F Binds Sec61alpha to Inhibit Protein Translocation. J Am Chem Soc. 2019;141(21):8450-61. Epub 20190515. doi: 10.1021/jacs.8b13506. PubMed PMID: 31059257; PubMed Central PMCID: PMCPMC6627486.
38. Nilsson IM, von Heijne G. Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. J Biol Chem. 1993;268(8):5798-801. PubMed PMID: 8449946.
39. Whitley P, Grau B, Gumbart JC, Martinez-Gil L, Mingarro I. Folding and Insertion of Transmembrane Helices at the ER. Int J Mol Sci. 2021;22(23). Epub 20211126. doi: 10.3390/ijms222312778. PubMed PMID: 34884581; PubMed Central PMCID: PMCPMC8657811.
40. Meindl-Beinker NM, Lundin C, Nilsson I, White SH, von Heijne G. Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly. EMBO Rep. 2006;7(11):1111-6. Epub 20060929. doi: 10.1038/sj.embor.7400818. PubMed PMID: 17008929; PubMed Central PMCID: PMCPMC1679787.
41. Tornieri K, Zlatic SA, Mullin AP, Werner E, Harrison R, L'Hernault S W, et al. Vps33b pathogenic mutations preferentially affect VIPAS39/SPE-39-positive endosomes. Hum Mol Genet. 2013;22(25):5215-28. Epub 2013/08/07. doi: 10.1093/hmg/ddt378. PubMed PMID: 23918659; PubMed Central PMCID: PMCPMC3842178.

42. Pickard A, Adamson A, Lu Y, Chang J, Garva R, Hodson N, et al. Collagen assembly and turnover imaged with a CRISPR-Cas9 engineered Dendra2 tag. bioRxiv. 2018. doi: https://doi.org/10.1101/331496
43. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. Epub 2009/01/10. doi: 10.1038/nprot.2008.211. PubMed PMID: 19131956.
44. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13. Epub 2008/11/27. doi: 10.1093/nar/gkn923. PubMed PMID: 19033363; PubMed Central PMCID: PMCPMC2615629.
45. Humphries JD, Byron A, Humphries MJ. Integrin ligands at a glance. J Cell Sci. 2006;119(Pt 19):3901-3. Epub 2006/09/22. doi: 10.1242/jcs.03098. PubMed PMID: 16988024; PubMed Central PMCID: PMCPMC3380273.
46. Wynn TA. Integrating mechanisms of pulmonary fibrosis. J Exp Med. 2011;208(7):1339-50. doi: 10.1084/jem.20110551. PubMed PMID: 21727191; PubMed Central PMCID: PMCPMC3136685.
47. Knowles GC, McKeown M, Sodek J, McCulloch CA. Mechanism of collagen phagocytosis by human gingival fibroblasts: importance of collagen structure in cell recognition and internalization. J Cell Sci. 1991;98 (Pt 4):551-8. Epub 1991/04/01. PubMed PMID: 1650378.
48. Everts V, Aronson DC, Beertsen W. Phagocytosis of bone collagen by osteoclasts in two cases of pycnodysostosis. Calcif Tissue Int. 1985;37(1):25-31. Epub 1985/01/01. doi: 10.1007/BF02557674. PubMed PMID: 3922593.
49. Arora PD, Wang Y, Bresnick A, Dawson J, Janmey PA, McCulloch CA. Collagen remodeling by phagocytosis is determined by collagen substrate topology and calcium-dependent interactions of gelsolin with nonmuscle myosin IIA in cell adhesions. Mol Biol Cell. 2013;24(6):734-47. Epub 2013/01/18. doi: 10.1091/mbc.E12-10-0754. PubMed PMID: 23325791; PubMed Central PMCID: PMCPMC3596245.
50. Vijayan RS, Arnold E, Das K. Molecular dynamics study of HIV-1 RT-DNA-nevirapine complexes explains NNRTI inhibition and resistance by connection mutations. Proteins. 2014;82(5):815-29. Epub 2013/11/01. doi: 10.1002/prot.24460. PubMed PMID: 24174331; PubMed Central PMCID: PMCPMC4502926.
51. Eble JA, Golbik R, Mann K, Kuhn K. The alpha 1 beta 1 integrin recognition site of the basement membrane collagen molecule [alpha 1(IV)]2 alpha 2(IV). EMBO J. 1993;12(12):4795-802. Epub 1993/12/01. doi: 10.1002/j.1460-2075.1993.tb06168.x. PubMed PMID: 8223488; PubMed Central PMCID: PMCPMC413926.
52. Herrera JA, Dingle L, Montero MA, Venkateswaran RV, Blaikley JF, Lawless C, et al. The UIP/IPF fibroblastic focus is a collagen biosynthesis factory embedded in a distinct extracellular matrix. JCI Insight. 2022;7(16). Epub 2022/07/20. doi: 10.1172/jci.insight.156115. PubMed PMID: 35852874; PubMed Central PMCID: PMCPMC9462507.
53. Jones CE, Sharick JT, Colbert SE, Shukla VC, Zent JM, Ostrowski MC, et al. Pten regulates collagen fibrillogenesis by fibroblasts through SPARC. PLoS One. 2021;16(2):e0245653. Epub 2021/02/04. doi: 10.1371/journal.pone.0245653. PubMed PMID: 33534863; PubMed Central PMCID: PMCPMC7857610.
54. Wenstrup RJ, Smith SM, Florer JB, Zhang G, Beason DP, Seegmiller RE, et al. Regulation of collagen fibril nucleation and initial fibril assembly involves coordinate interactions with collagens V and XI in developing tendon. J Biol Chem. 2011;286(23):20455-65. Epub 2011/04/07. doi: 10.1074/jbc.M111.223693. PubMed PMID: 21467034; PubMed Central PMCID: PMCPMC3121453.
55. Taylor SH, Yeung CY, Kalson NS, Lu Y, Zigrino P, Starborg T, et al. Matrix metalloproteinase 14 is required for fibrous tissue expansion. Elife. 2015;4:e09345. Epub 2015/09/22. doi: 10.7554/eLife.09345. PubMed PMID: 26390284; PubMed Central PMCID: PMCPMC4684142.
56. Moses M, Chang J, Hansen U, Kadler KE, Zeltz C, Gullberg D. Collagen assemble at the cell surface: Dogmas revisited. Cells. 2021;in press.
57. Moriya K, Bae E, Honda K, Sakai K, Sakaguchi T, Tsujimoto I, et al. A fibronectin-independent mechanism of collagen fibrillogenesis in adult liver remodeling. Gastroenterology. 2011;140(5):1653-63. Epub 2011/02/16. doi: 10.1053/j.gastro.2011.02.005. PubMed PMID: 21320502; PubMed Central PMCID: PMCPMC3081910.
58. Canty-Laird EG, Lu Y, Kadler KE. Stepwise proteolytic activation of type I procollagen to collagen within the secretory pathway of tendon fibroblasts in situ. Biochem J. 2012;441(2):707-17. Epub 2011/10/05. doi: 10.1042/BJ20111379. PubMed PMID: 21967573; PubMed Central PMCID: PMCPMC3430002.
59. Stevenson NL, Bergen DJM, Lu Y, Prada-Sanchez ME, Kadler KE, Hammond CL, et al. Giantin is required for intracellular N-terminal processing of type I procollagen. J Cell Biol. 2021;220(6). Epub 2021/05/05. doi: 10.1083/jcb.202005166. PubMed PMID: 33944912; PubMed Central PMCID: PMCPMC8103548.
60. Liu RJY, Al-Molieh Y, Chen SZ, Drobac M, Urban D, Chen CH, et al. The Sec1-Munc18 protein VPS33B forms a uniquely bidirectional complex with VPS16B. J Biol Chem. 2023;299(6):104718. Epub 2023/04/17. doi: 10.1016/j.jbc.2023.104718. PubMed PMID: 37062417; PubMed Central PMCID: PMCPMC10208892.

61. Lynch DA, Sverzellati N, Travis WD, Brown KK, Colby TV, Galvin JR, et al. Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. Lancet Respir Med. 2018;6(2):138-53. Epub 2017/11/21. doi: 10.1016/S2213-2600(17)30433-2. PubMed PMID: 29154106.
62. Herrera JA, Mallikarjun V, Rosini S, Montero MA, Lawless C, Warwood S, et al. Laser capture microdissection coupled mass spectrometry (LCM-MS) for spatially resolved analysis of formalin-fixed and stained human lung tissues. Clin Proteomics. 2020;17:24. Epub 2020/06/23. doi: 10.1186/s12014-020-09287-6. PubMed PMID: 32565759; PubMed Central PMCID: PMCPMC7302139.
63. Herrera J, Beisang DJ, Peterson M, Forster C, Gilbertsen A, Benyumov A, et al. Dicer1 Deficiency in the Idiopathic Pulmonary Fibrosis Fibroblastic Focus Promotes Fibrosis by Suppressing MicroRNA Biogenesis. Am J Respir Crit Care Med. 2018;198(4):486-96. Epub 2018/03/27. doi: 10.1164/rccm.201709-1823OC. PubMed PMID: 29579397; PubMed Central PMCID: PMCPMC6118029.
64. McKenna M, Simmonds RE, High S. Mechanistic insights into the inhibition of Sec61-dependent co- and post-translational translocation by mycolactone. J Cell Sci. 2016;129(7):1404-15. Epub 20160211. doi: 10.1242/jcs.182352. PubMed PMID: 26869228; PubMed Central PMCID: PMCPMC4852723.
65. Campeau E, Ruhl VE, Rodier F, Smith CL, Rahmberg BL, Fuss JO, et al. A versatile viral system for expression and depletion of proteins in mammalian cells. PLoS One. 2009;4(8):e6529. Epub 2009/08/07. doi: 10.1371/journal.pone.0006529. PubMed PMID: 19657394; PubMed Central PMCID: PMCPMC2717805.
66. Popova SN, Rodriguez-Sanchez B, Liden A, Betsholtz C, Van Den Bos T, Gullberg D. The mesenchymal alpha11beta1 integrin attenuates PDGF-BB-stimulated chemotaxis of embryonic fibroblasts on collagens. Dev Biol. 2004;270(2):427-42. Epub 2004/06/09. doi: 10.1016/j.ydbio.2004.03.006. PubMed PMID: 15183724.
67. O'Keefe S, Roboti P, Duah KB, Zong G, Schneider H, Shi WQ, et al. Ipomoeassin-F inhibits the in vitro biogenesis of the SARS-CoV-2 spike protein and its host cell membrane receptor. J Cell Sci. 2021;134(4). Epub 20210219. doi: 10.1242/jcs.257758. PubMed PMID: 33468620; PubMed Central PMCID: PMCPMC7904091.
68. Smeland HY, Askeland C, Wik E, Knutsvik G, Molven A, Edelmann RJ, et al. Integrin alpha11beta1 is expressed in breast cancer stroma and associates with aggressive tumor phenotypes. J Pathol Clin Res. 2020;6(1):69-82. Epub 2019/10/13. doi: 10.1002/cjp2.148. PubMed PMID: 31605508; PubMed Central PMCID: PMCPMC6966706.
69. Wu G, Anafi RC, Hughes ME, Kornacker K, Hogenesch JB. MetaCycle: an integrated R package to evaluate periodicity in large scale data. Bioinformatics. 2016;32(21):3351-3. Epub 2016/10/30. doi: 10.1093/bioinformatics/btw405. PubMed PMID: 27378304; PubMed Central PMCID: PMCPMC5079475.
70. Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics. 1996;5(3):299-314. doi: 10.1080/10618600.1996.10474713.

Figure 1

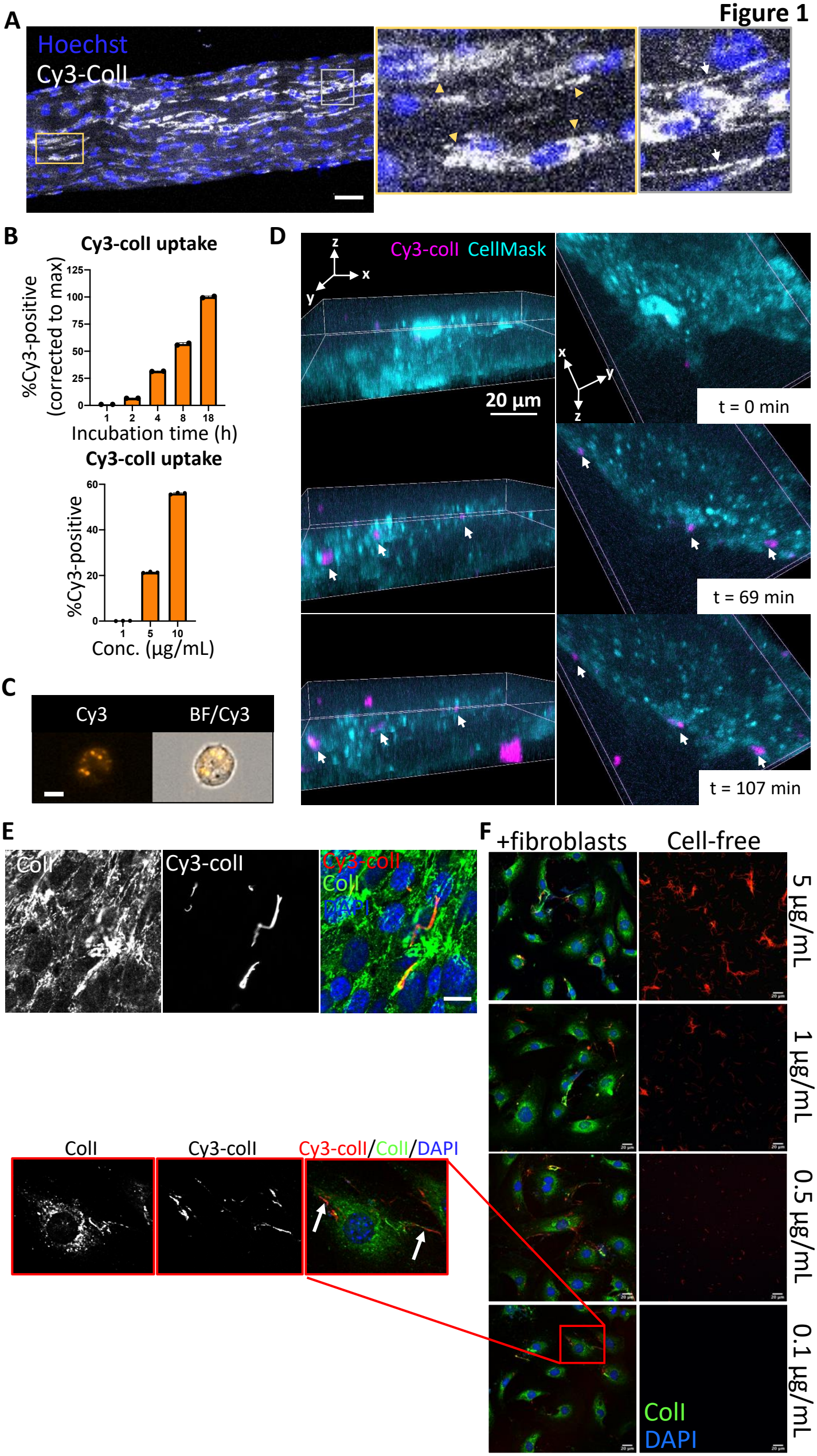

A
Hoechst
Cy3-Coll
B
Cy3-coll uptake
%Cy3-positive (corrected to max)
125
100
75
50
25
0
1
2
4
8
18
Incubation time (h)
Cy3-coll uptake
%Cy3-positive
60
40
20
0
1
5
10
Conc. (µg/mL)
C
Cy3
BF/Cy3
D
Cy3-coll CellMask
z
x
y
20 µm
t = 0 min
t = 69 min
t = 107 min
E
Coll
Cy3-coll
Cy3-coll
Coll
DAPI
Coll
Cy3-coll
Cy3-coll/Coll/DAPI
F
+fibroblasts
Cell-free
5 µg/mL
1 µg/mL
0.5 µg/mL
0.1 µg/mL
Coll
DAPI

Figure 1: Collagen-I is endocytosed and reassembled into fibrils

A. Fluorescent images of tail tendon incubated with Cy3-colI for 5 days, showing presence of collagen-I within the cells, and fibril-like fluorescence signals outside of cells. Hoechst stain was used to locate cells within the tendon. Area surrounded by yellow box expanded on the right, and cells with Cy3-colI present intracellularly pointed out by yellow triangles. Area surrounded by grey box expanded on the right, and fibril-like fluorescence signals indicated with white arrows. Scale bar = 50 µm. Representative of N = 3.

B. Bar chart showing an increase of percentage of fluorescent cells incubated with 1.5 µg/mL Cy3-colI over time (top), and an increase of percentage fluorescent cells incubated with increasing concentration of Cy3-colI for one hour (bottom), suggesting a non-linear time-dependent and dose-dependent uptake pattern. N = 3.

C. Flow cytometry imaging of fibroblasts incubated with 5 µg/mL Cy3-labeled collagen-I for one hour, showing that collagen-I is taken up by cells and held in vesicular structures. Images acquired using ImageStream at 40x magnification. Scale bar = 10 μm. Cy3 – Cy3 channel, BF/Cy3 – merged image of BF and Cy3. Representative of >500 cells images collected per condition.

D. Fluorescent image series of fibroblasts incubated with Cy3-labeled collagen-I at 0 min (t = 0min) 69 min (t = 69 min) and 107 min (t = 107 min) after addition (rows) and viewed from different angles (columns), with cell mask (cyan) to distinguish the cell volume. White arrows point to intracellular collagen in vesicular like structures. Representative of N = 3. Scale bar = 20 µm.

E. Fluorescent image series of fibroblasts incubated with 5 µg/mL Cy3-colI for one hour, trypsinized and replated in fresh media, and cultured for 72 h. Top left labels denotes the fluorescence channel corresponding to proteins detected. Representative of N>3. Scale bar = 20 µm.

F. Fluorescent image series of Cy3-colI incubated at different concentrations for 72 h, either cell-free (right panel), or with fibroblasts (left panel). Representative of N = 3. Scale bar = 20 µm. Red box - zoomed out to the left, and separated according to fluorescence channel. White arrows highlighting Cy3-positive fibrils assembled by fibroblasts when incubated with 0.1 µg/mL Cy3-colI.

Figure 2

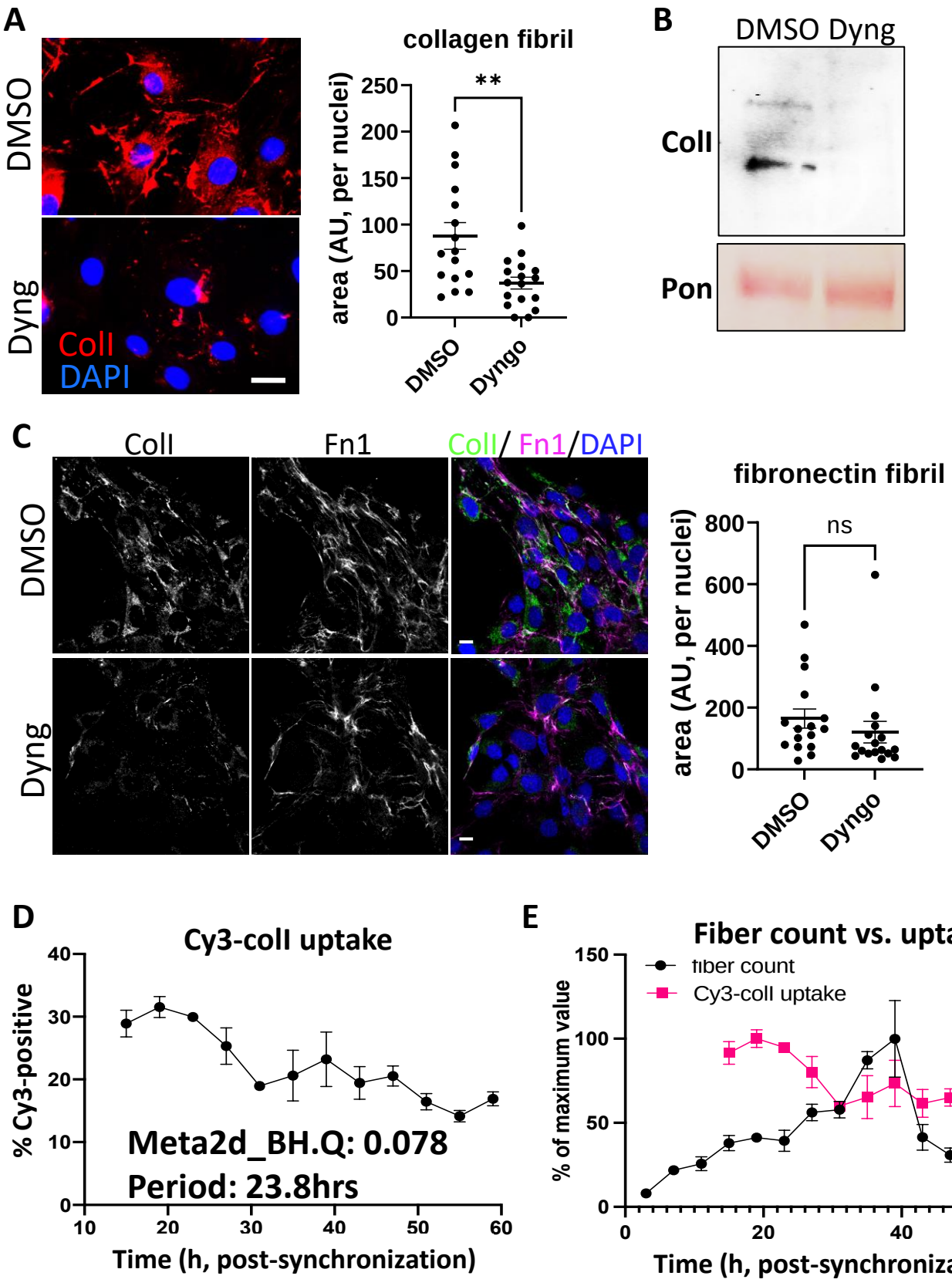

A
DMSO
Dyng
Coll
DAPI
collagen fibril
**
area (AU, per nuclei)
250
200
150
100
50
0
DMSO
Dyngo
B
DMSO Dyng
Coll
Pon
C
Coll
Fn1
Coll/ Fn1/DAPI
DMSO
Dyng
fibronectin fibril
ns
area (AU, per nuclei)
800
600
400
200
0
DMSO
Dyngo
D
Cy3-coll uptake
% Cy3-positive
40
30
20
10
0
Meta2d_BH.Q: 0.078
Period: 23.8hrs
10
20
30
40
50
60
Time (h, post-synchronization)
E
Fiber count vs. uptake
fiber count
Cy3-coll uptake
% of maximum value
150
100
50
0
0
20
40
60
Time (h, post-synchronization)

Figure 2: Inhibition of endocytosis leads to changes in collagen-I homeostasis, and endocytosis is a rhythmic event

A. Left: fluorescent images of collagen-I (red) counterstained with DAPI (blue) in fibroblasts treated with DMSO (top) or Dyng (bottom) for 72 h. Scale bar = 20 µm. Right: quantification of area occupied by collagen-I fibrils, corrected to number of nuclei.

B. Western blot analysis of conditioned media taken from fibroblasts treated with DMSO (ctrl) or Dyng (Dyng) for 72 h, showing a decrease in collagen-I secretion. Top: probed with collagen-I antibody (Col-I), bottom: counterstained with Ponceau (Pon) as control. Representative of N = 3.

C. Left: fluorescent image series of collagen-I (green), fibronectin (magenta) counterstained with DAPI (blue) in fibroblasts treated with DMSO (top) or Dyng (bottom) for 72 h. Scale bar = 20 µm. Right: quantification of area occupied by fibronectin fibrils, corrected to number of nuclei.

D. Percentage Cy3-colI taken up by synchronized fibroblasts over 48 h. Meta2d analysis indicates a circadian rhythm of periodicity of 23.8 h. Bars show mean ± s.e.m. of N = 3 per time point.

E. Percentage of Cy3-colI taken up by synchronized cells, corrected to the maximum percentage uptake of the time course (pink, bars show mean ± s.e.m. of N=3 per time point), compared to the percentage collagen fibril count over time, corrected to the maximum percentage fibril count of the time course (black, fibrils scored by two independent investigators. Bars show mean± s.e.m. of N=2 with n=6 repeats at each time point).

Figure 3

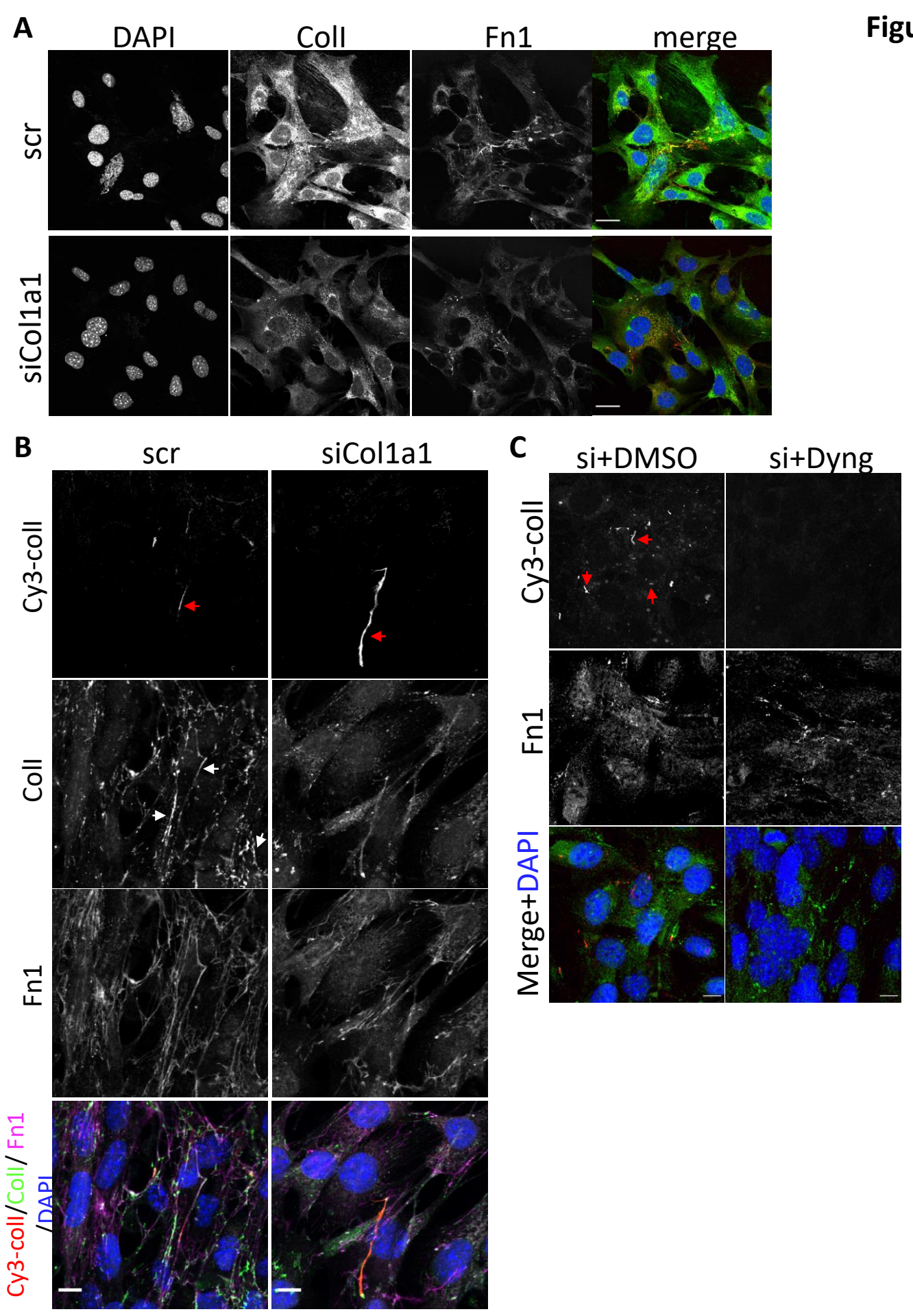

A
DAPI
Coll
Fn1
merge
scr
siCol1a1
B
scr
siCol1a1
Cy3-coll
Coll
Fn1
Cy3-coll/Coll/Fn1
/DAPI
C
si+DMSO
si+Dyng
Cy3-coll
Fn1
Merge+DAPI

Figure 3: Collagen-I recycling can generate fibrils

A. Fluorescent image series of fibroblasts treated with scrambled control (top panel, scr), and siRNA against col1a1 (bottom panel, siCol1a1). Labels on top denotes the fluorescence channel corresponding to proteins detected. Representative of N = 4. Scale bar = 25 µm.

B. Fluorescent image series of scr (left column) and siCol1a1 (right column) fibroblasts incubated with Cy3-colI. Labels on left denotes the fluorescence channel(s) corresponding to proteins detected. Cy3-colI fibrils highlighted by red arrows, and collagen-I fibrils highlighted by white arrows. Both scr cells and siCol1a1 cells can take up exogenous collagen-I and recycle to form collagen-I fibril. Representative of N>3. Scale bar = 10 μm.

C. Fluorescent image series of siCol1a1 cells treated with DMSO control (left) or Dyng (right) during Cy3-colI uptake, followed by further culture for 72 h. Labels on left denotes the fluorescence channel corresponding to proteins detected. Dyngo treatment led to a reduction of Cy3-colI fibrils. Representative of N>3. Scale bar = 25 µm.

Figure 4

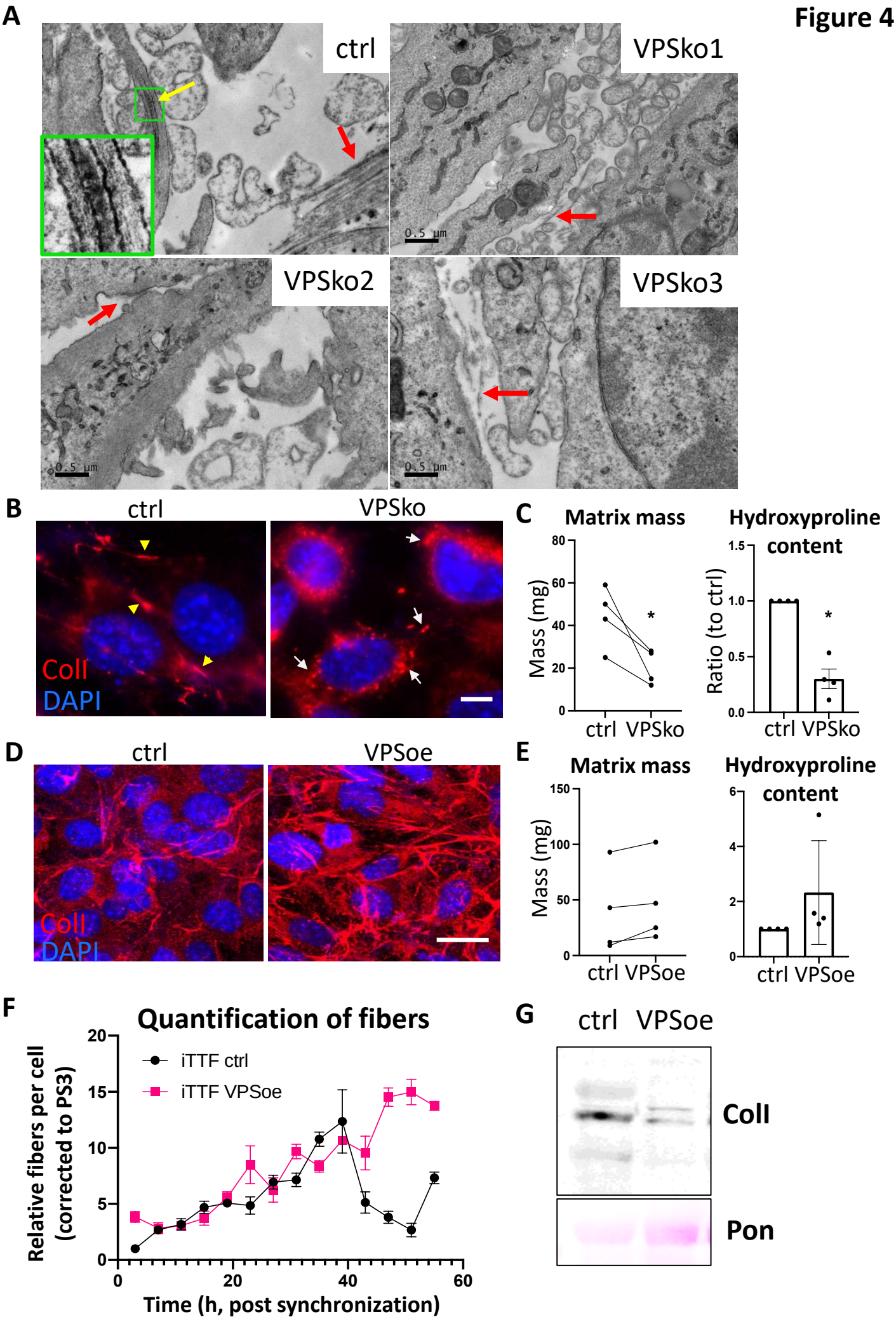

A
ctrl
VPSko1
VPSko2
VPSko3
0.5 µm
B
ctrl
VPSko
Coll
DAPI
C
Matrix mass
Mass (mg)
ctrl VPSko
Hydroxyproline content
Ratio (to ctrl)
ctrl VPSko
D
ctrl
VPSoe
Coll
DAPI
E
Matrix mass
Mass (mg)
ctrl VPSoe
Hydroxyproline content
ctrl VPSoe
F
Quantification of fibers
iTTF ctrl
iTTF VPSoe
Relative fibers per cell (corrected to PS3)
Time (h, post synchronization)
G
ctrl VPSoe
Coll
Pon

Figure 4: VPS33B controls collagen fibril formation at the plasma membrane in a rhythmic manner

A. Electron microcopy images of fibroblasts plated on Aclar and grown for a week before fixation and imaging. Ctrl culture has numerous collagen-I fibrils, as pointed out by arrows. Yellow arrow points to a fibripositor, and green box is expanded to the left bottom corner, showing the distinct D-banding pattern of collagen-I fibril when observed with electron microscopy. VPSko clones all have fewer and thinner fibrils present in the culture (pointed out by red arrows). Representative of N = 3. Scale bar = 0.5 µm.

B. Fluorescence images of collagen-I (red) and DAPI counterstain in ctrl and VPSko fibroblasts. Yellow arrows indicating collagen fibrils, and white arrows pointing to collagen-I presence in intracellular vesicles. Representative of N>6. Scale bar = 25 µm.

C. Matrix deposition by ctrl or VPSko cells, after one week of culture. Left: decellularized matrix mass. N = 4, * p = 0.0299. Right: hydroxyproline content presented as a ratio between ctrl and VPSko cells. N = 4, *p = 0.0254. Ratio-paired t-test used.

D. Fluorescence images of collagen-I (red) and DAPI counterstain in ctrl and VPSoe fibroblasts. Representative of N>6. Scale bar = 20 µm.

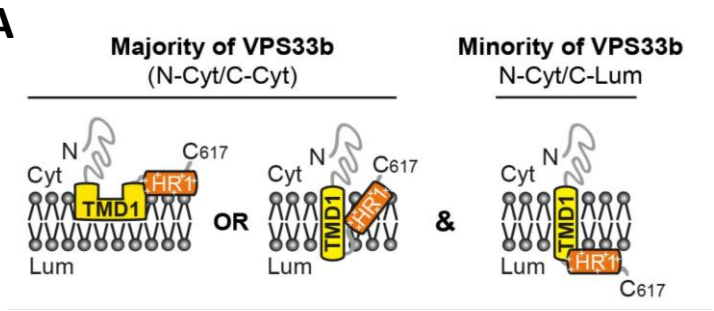
E. Matrix deposition by ctrl or VPSoe cells, after one week of culture. Left: decellularized matrix mass, N = 4. Right: hydroxyproline content presented as a ratio between ctrl and VPSoe cells, N = 4. Ratio-paired t-test used.

F. Relative collagen fibril count in synchronized ctrl (black) and VPSoe (pink) fibroblasts, corrected to the number of fibrils in ctrl cultures at start of time course. Fibrils scored by two independent investigators. Bars show mean± s.e.m. of N = 2 with n=6 at each time point.

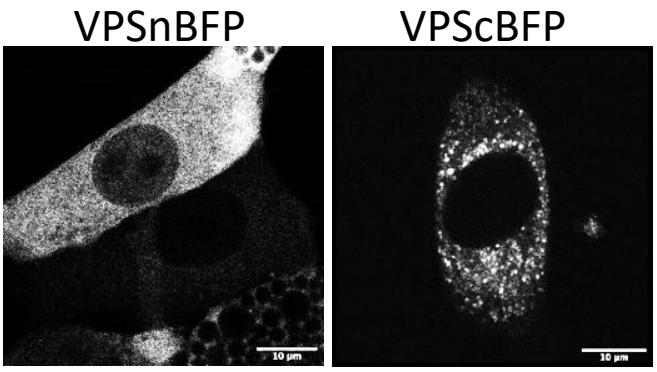
G. Western blot analysis of conditioned media taken from ctrl and VPSoe after 72 h in culture. Top: probed with collage-I antibody (ColI), bottom: counterstained with Ponceau (Pon) as control. Representative of N = 3.

Figure 5

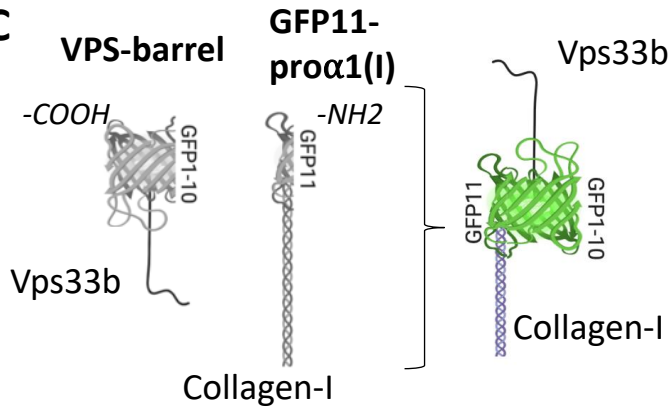
A



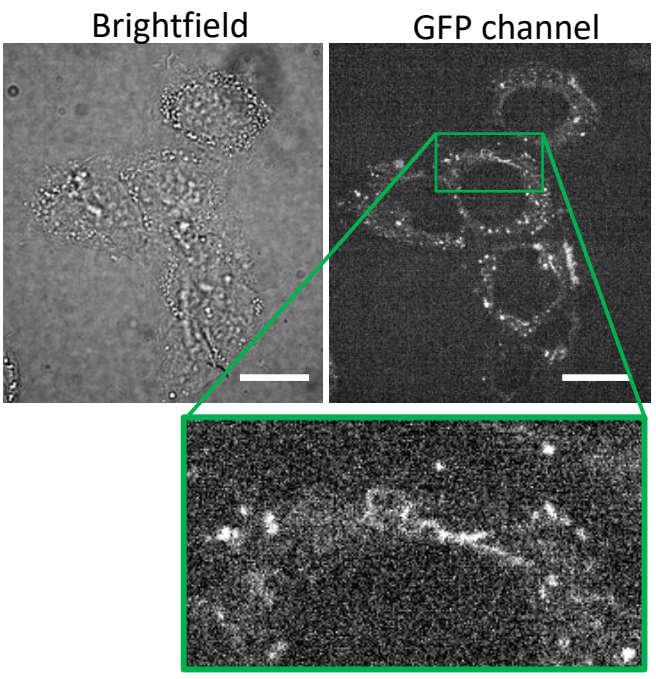
B



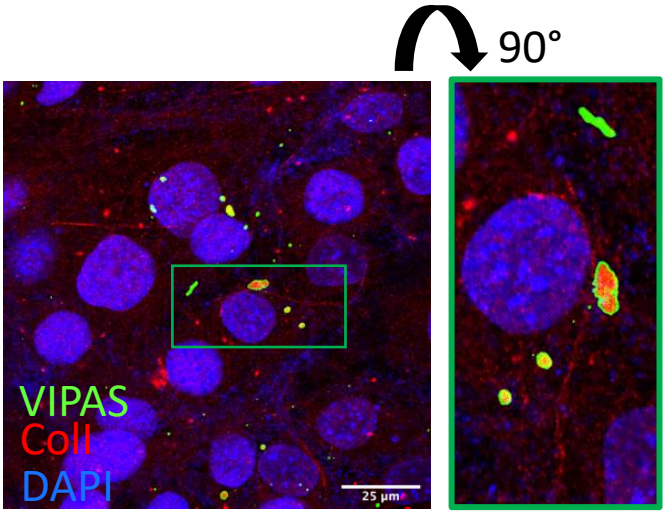
C



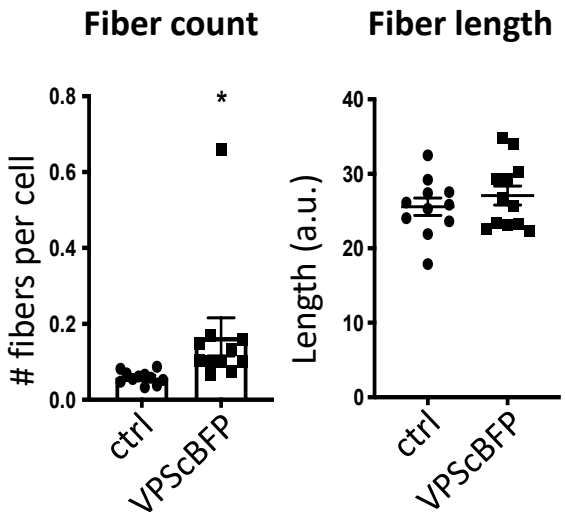
D



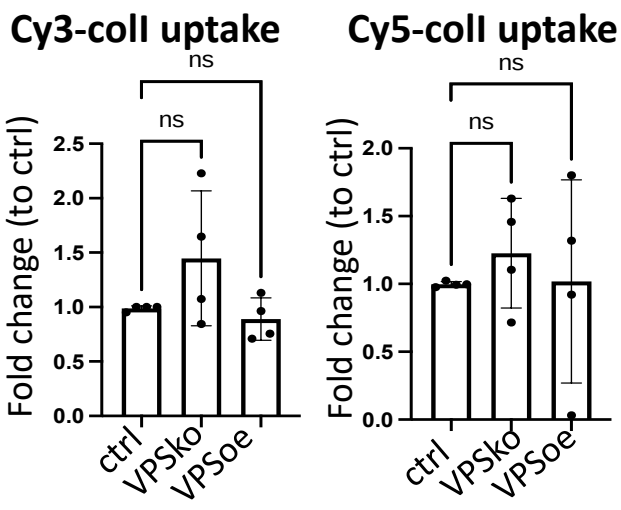
E



F



G



H

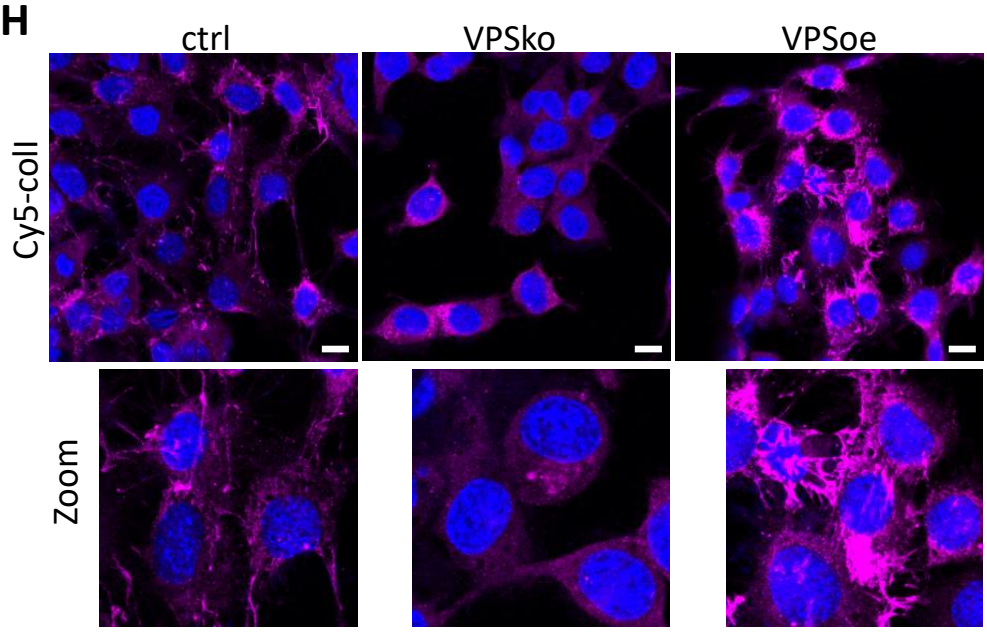


Figure 5: Procollagen-I and VPS33B localize to the same compartments

A. Schematic depicting the proposed membrane topologies of VPS33b.

B. Fibroblasts expressing BFP-tagged VPS33B. Left: BFP tagged on the N-terminal end of VPS33B (VPSnBFP), Right: BFP tagged on the C-terminal end (VPScBFP). Images taken in Airy mode. Representative of N>4. Scale bar = 10 µm.

C. Schematic of the split-GFP system. GFP1-10 barrel is introduced into VPS33B (VPS-barrel), and GFP11 to alpha-1 chain of Collagen-I (GFP11-proα1(I)). If the two tagged proteins co-localize (e.g. in a vesicle), a GFP signal will be emitted.

D. Brightfield (top) and fluorescence (middle) images of fibroblasts expressing both VPS-barrel and GFP11-proα1(I) constructs. Representative of N=5. Green box is expanded to the bottom, to highlight the punctate fluorescence signals within intracellular vesicular structures, as well as fibril-like structures suggestive of fibril assembly sites. Scale bar = 20 µm.

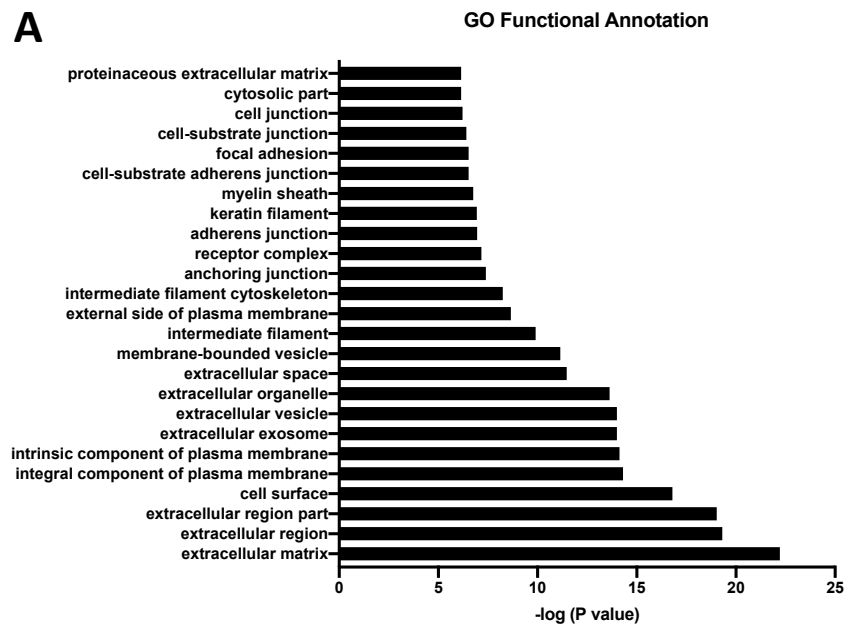
E. Fluorescence images of VIPAS (green), collagen-I (red) and DAPI counterstain in fibroblasts. Representative of N=3. Green box is expanded to the right (flipped 90°) to show strong VIPAS signal encasing collagen-I. Scale bar = 25 µm.

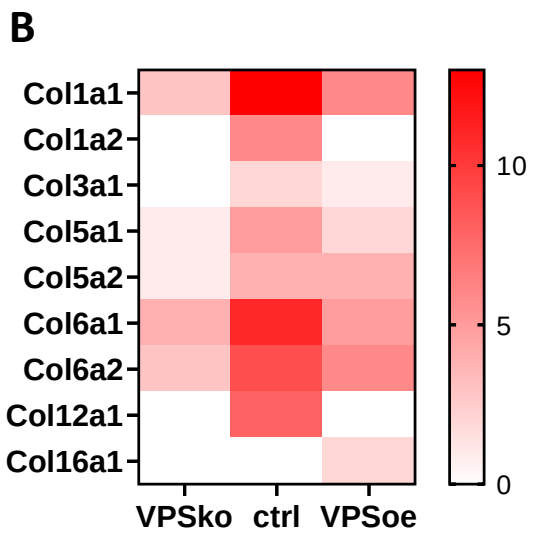
F. Quantification of average number of fibrils per cell (left) and average fibril length (right) in control Dendra-colI expressing 3T3 cells (ctrl) and Dendra-colI expressing 3T3 overexpressing VPScBFP (VPScBFP). >500 cells quantified per condition. N=12. *P=0.048.

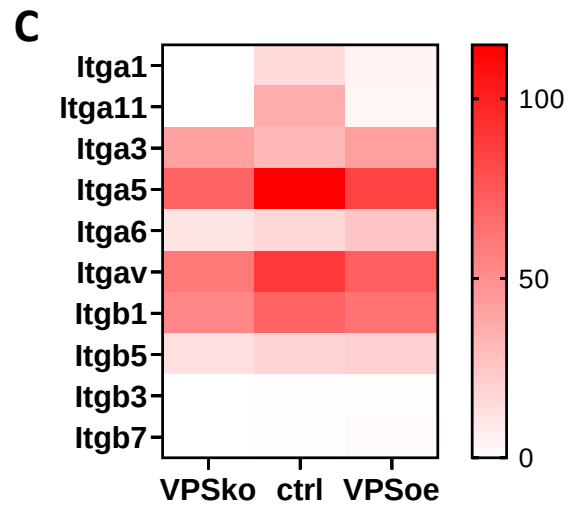
G. Bar charts comparing the percentage of cells that have taken up 5 µg/mL Cy3-colI (left) and Cy5-colI (right) after one hour incubation between control (ctrl), VPS33B-knockout (VPSko) and VPS33B-overexpressing (VPSoe) cells, corrected to control. RM one-way ANOVA analysis was performed. N = 4.

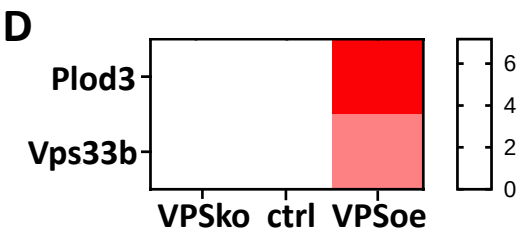
H. Fluorescence images of cells of different levels of VPS33B expression, fed with Cy5-colI and further cultured for 72 h. Cultures were counterstained with DAPI. Bottom panel are zoomed-in images of the fibrils produced by the fibroblasts. Representative of N = 2.

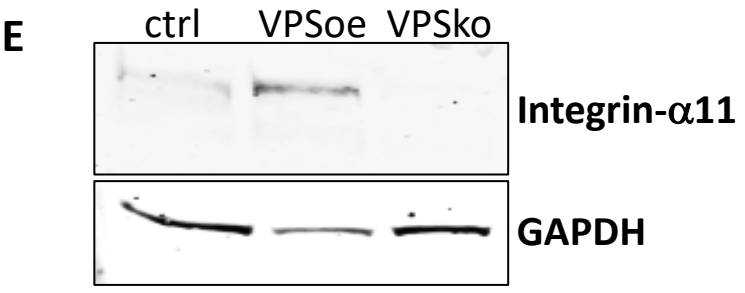
Figure 6

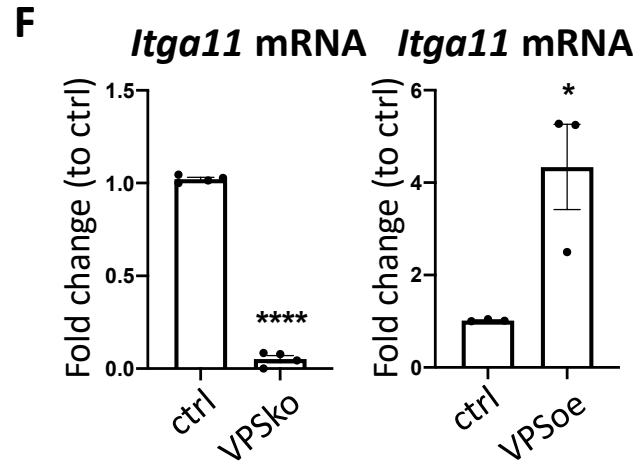

A
GO Functional Annotation
proteinaceous extracellular matrix
cytosolic part
cell junction
cell-substrate junction
focal adhesion
cell-substrate adherens junction
myelin sheath
keratin filament
adherens junction
receptor complex
anchoring junction
intermediate filament cytoskeleton
external side of plasma membrane
intermediate filament
membrane-bounded vesicle
extracellular space
extracellular organelle
extracellular vesicle
extracellular exosome
intrinsic component of plasma membrane
integral component of plasma membrane
cell surface
extracellular region part
extracellular region
extracellular matrix
0
5
10
15
20
25
-log (P value)

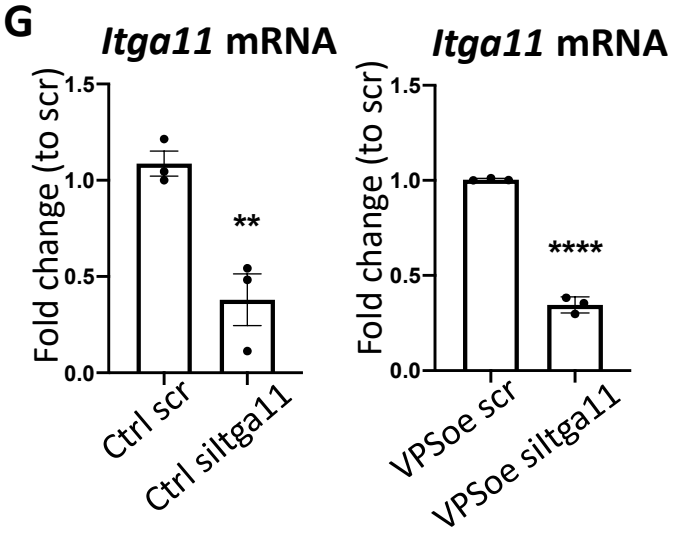

B
Col1a1
Col1a2
Col3a1
Col5a1
Col5a2
Col6a1
Col6a2
Col12a1
Col16a1
VPSko
ctrl
VPSoe
10
5
0

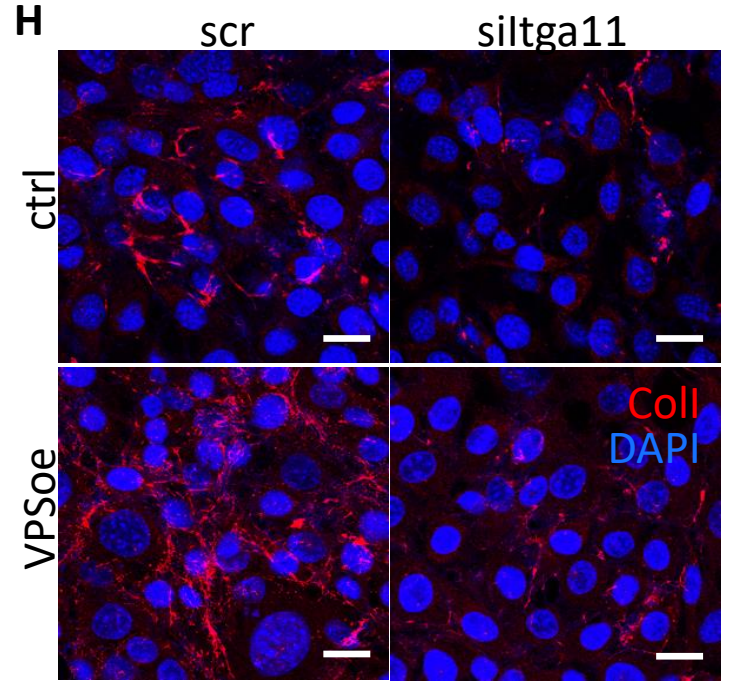

C
Itga1
Itga11
Itga3
Itga5
Itga6
Itgav
Itgb1
Itgb5
Itgb3
Itgb7
VPSko
ctrl
VPSoe
100
50
0


D
Plod3
Vps33b
VPSko
ctrl
VPSoe
6
4
2
0


E
ctrl
VPSoe
VPSko
Integrin-α11
GAPDH


F
Itga11 mRNA
Itga11 mRNA
Fold change (to ctrl)
1.5
1.0
0.5
0.0
****
ctrl
VPSko
Fold change (to ctrl)
6
4
2
0
*
ctrl
VPSoe


G
Itga11 mRNA
Itga11 mRNA
Fold change (to scr)
1.5
1.0
0.5
0.0
**
Ctrl scr
Ctrl siltga11
Fold change (to scr)
1.5
1.0
0.5
0.0
****
VPSoe scr
VPSoe siltga11


H
scr
siltga11
ctrl
VPSoe
Coll
DAPI

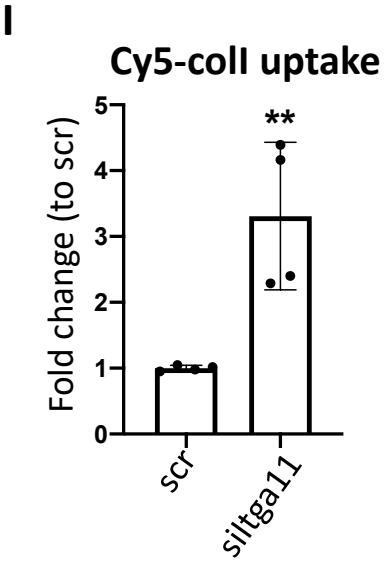

I
Cy5-coll uptake
Fold change (to scr)
5
4
3
2
1
0
**
scr
siltga11

Figure 6: Integrin α11 subunit mediates VPS33B-effects and is required for collagen-I fibrillogenesis

A. Top 25 Functional Annotation of proteins detected in biotin-enriched samples when compared to non-enriched samples based on p-values. Y-axis denotes the GO term, X-axis denotes –log (P value).

B. Heatmap representation of spectral counting of collagens detected in biotin-enriched surface proteins from control (ctrl), VPS33B-knockout (VPSko), and VPS33B-overexpressing (VPSoe) fibroblasts. Scale denotes quantitative value as normalized to total spectra, as determined by Proteome Discoverer.

C. Heatmap representation of spectral counting of integrins detected in biotin-enriched surface proteins from control (ctrl), VPS33B-knockout (VPSko), and VPS33B-overexpressing (VPSoe) fibroblasts. Scale denotes quantitative value as normalized to total spectra, as determined by Proteome Discoverer.

D. Heatmap representation of spectral counting of Plod3 and VPS33B detected in biotin-enriched surface proteins from control (ctrl), VPS33B-knockout (VPSko), and VPS33B-overexpressing (VPSoe) fibroblasts. Scale denotes quantitative value as normalized to total spectra, as determined by Proteome Discoverer.

E. Western blot analysis of integrin α11 subunit levels in control (ctrl), VPS33B-overexpressing (VPSoe), VPS33B-knockout (VPSko) cells. Top: probed with integrin α11 antibody, bottom: reprobed with GAPDH antibody. Representative of N = 3.

F. qPCR analysis of *Itga11* transcript levels in ctrl compared to VPSko fibroblasts (left), and ctrl compared to VPSoe fibroblasts (right). N>3, ****$P<0.0001$, *$P = 0.0226$.

G. qPCR analysis of *Itga11* mRNA expression in ctrl (left) or VPSoe (right) fibroblasts treated either with scrambled control (scr) or siRNA against Itga11 (siItga11), collected after 96 h. N=3, **$P=0.0091$, ****$P<0.0001$.

H. IF images of ctrl and VPSoe fibroblasts treated either with control siRNA (scr) or siRNA again Itga11 (siItga11), after 72 h incubation; collagen-I (red) and DAPI (blue) counterstained. Representative of N = 3. Scale bar = 25 µm.

I. Bar chart comparing the percentage of cells that have taken up 5 µg/mL Cy5-colI after one hour incubation between fibroblasts treated with scrambled control (ctrl) or siRNA against Itga11 (siItga11), corrected to scr. N=3. **$P=0.0062$.

Figure 7

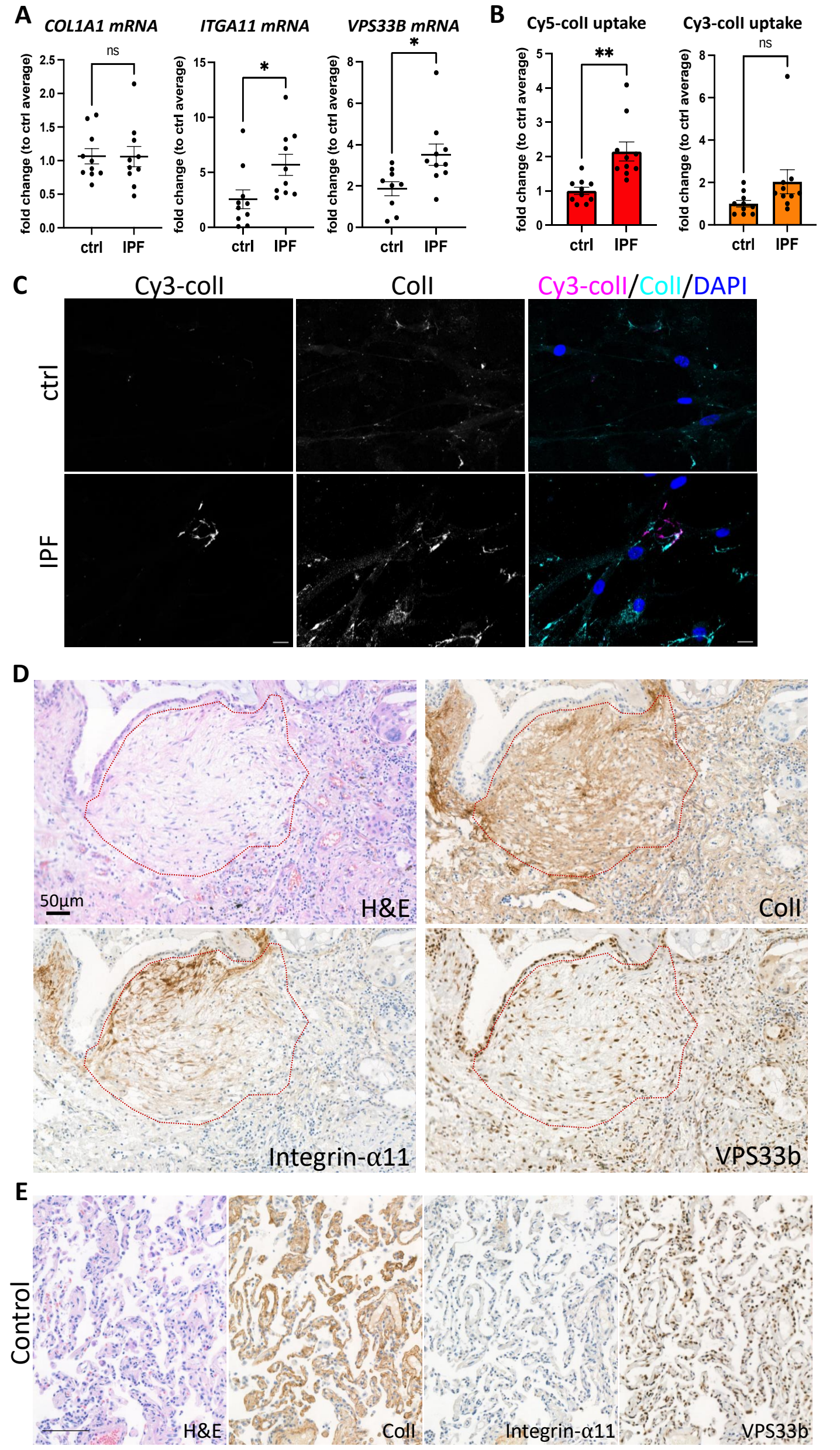

A
COL1A1 mRNA
ITGA11 mRNA
VPS33B mRNA
fold change (to ctrl average)
ns
*
*
ctrl
IPF
B
Cy5-coll uptake
Cy3-coll uptake
**
ns
C
Cy3-coll
Coll
Cy3-coll/Coll/DAPI
ctrl
IPF
D
50µm
H&E
Coll
Integrin-α11
VPS33b
E
Control
H&E
Coll
Integrin-α11
VPS33b

Figure 7: The IPF fibrotic focus is positive for integrin α11 subunit and VPS33B.

A. qPCR analysis of patient-derived fibroblasts isolated from control (ctrl) or IPF lungs. Bars showing mean± s.e.m, 5 patients in each group from 2 independent experiments (technical repeats not shown here). *Itga11, *p* = 0.0259; *VPS33B*, **p* = 0.0183.

B. Fold change of percentage Cy5-colI (left) or Cy3-colI (right) taken up by ctrl or IPF lung fibroblasts, corrected to average of control fibroblasts. Bars showing mean± s.e.m., 5 patients in each group from 2 independent experiments (technical repeats not shown here). ***p* = 0.003.

C. Fluorescent image series of ctrl or IPF lung fibroblasts that have taken up Cy3-colI (magenta), followed by further culture for 48 h, before subjected to collagen-I staining (cyan). Labels on top denotes the fluorescence channel corresponding to proteins detected. IPF fibroblasts produced more Cy3-labelled fibrillar structures. Representative of N = 4. Scale bar = 20 µm.

D. Immunohistochemistry of IPF patient (patient 1) with red dotted line outlining the fibroblastic focus, the hallmark lesion of IPF. Sections were stained with hematoxylin and eosin (H&E), collagen-I (ColI), integrin-α11, VPS33B. Scale bar = 50 µm.

E. Immunohistochemistry of 5 µm thick sequential lung sections taken from lungs classified as control (Control 1). Sections were stained with hematoxylin and eosin (H&E), collagen-I (ColI), integrin-α11, VPS33B. Scale bar = 100 µm.

Figure 8

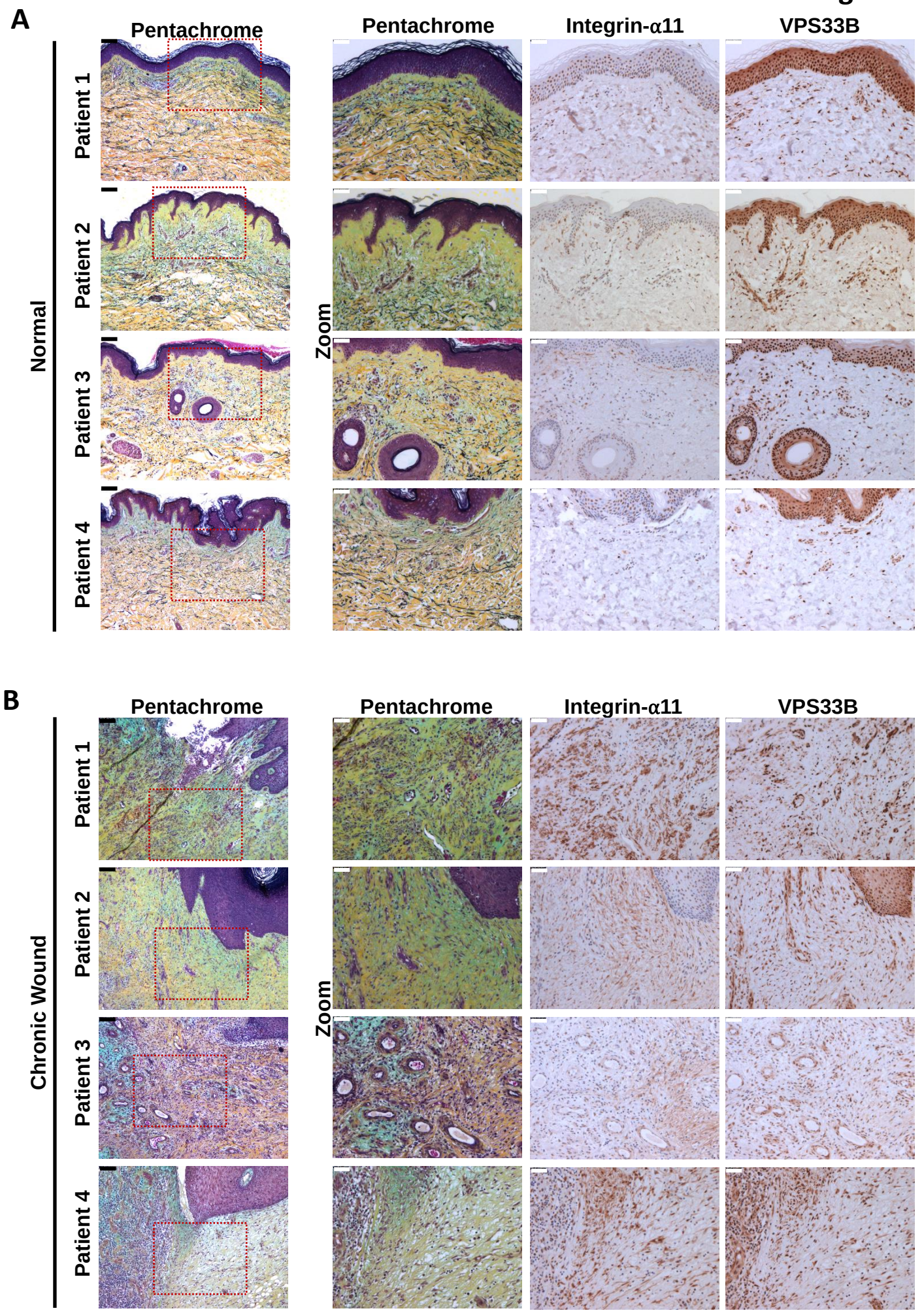

A
Pentachrome
Pentachrome
Integrin-α11
VPS33B
Normal
Patient 1
Patient 2
Patient 3
Patient 4
Zoom
B
Pentachrome
Pentachrome
Integrin-α11
VPS33B
Chronic Wound
Patient 1
Patient 2
Patient 3
Patient 4
Zoom

Figure 8: Proteins responsible for collagen fibrillogenesis are also co-localized to diseased areas of chronic skin wounds.

A. Immunohistochemistry of 5 µm thick sequential skin sections taken from normal skin regions of patients with chronic skin wounds (Patient 1, Patient 2, Patient 3, Patient 4). Sections were stained with pentachrome, integrin-α11, VPS33B. Scale bars positioned in top left corner: black (unzoomed pentachrome) = 100 µm, white (zoomed sections) = 50 µm.

B. Immunohistochemistry of 5 µm thick sequential skin sections taken from the chronic wound areas from patients with chronic skin wounds (Patient 1, Patient 2, Patient 3, Patient 4). Sections were stained with pentachrome, integrin-α11, VPS33B. Scale bars positioned in top left corner: black (unzoomed pentachrome) = 100 µm, white (zoomed sections) = 50 µm.

Figure 9

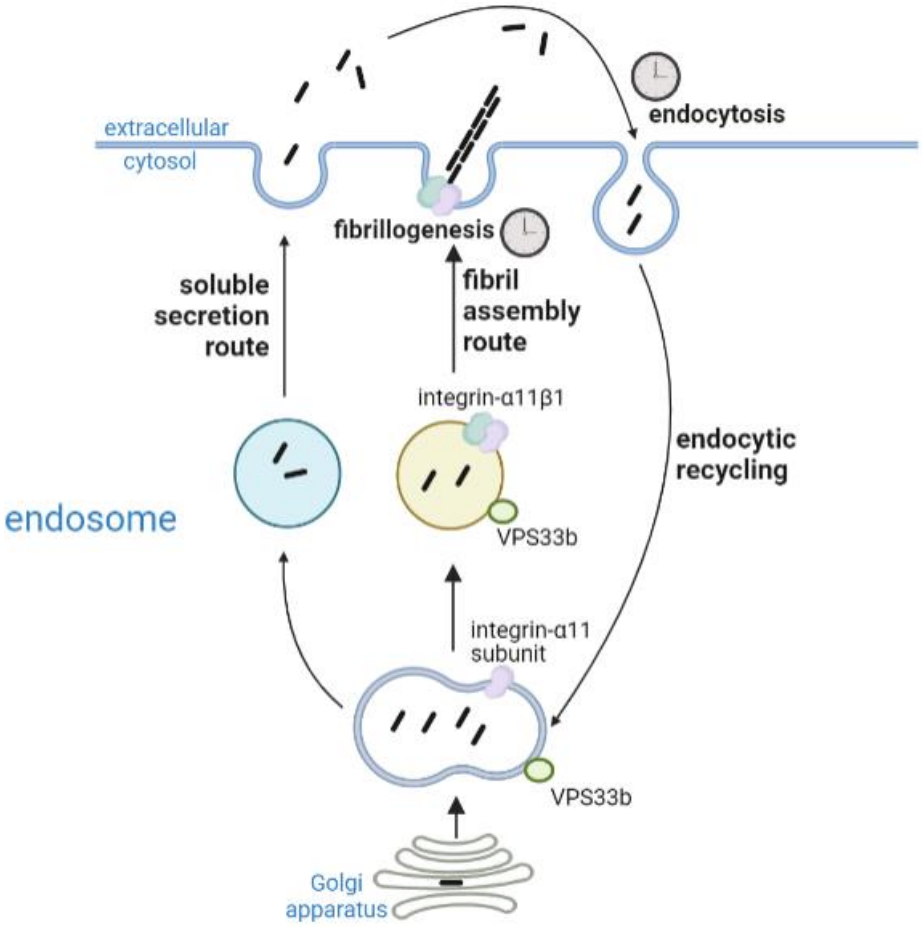

extracellular
cytosol
endocytosis
fibrillogenesis
soluble
secretion
route
fibril
assembly
route
integrin-α11β1
endocytic
recycling
endosome
VPS33b
integrin-α11
subunit
VPS33b
Golgi
apparatus

Figure 9: Proposed working model of collagen homeostasis in fibroblasts.

Endogenous collagen is either secreted as protomers (soluble secretion route, not circadian rhythmic) or made into fibrils (fibril assembly route, circadian rhythmic). Secreted collagen protomers can be captured by cells through endocytosis (circadian rhythmic) and recycled to make new fibrils. Integrin α11 and VPS33b directs collagen to fibril formation.