



**HAL**  
open science

# Identification of extracellular matrix biomarkers and targets in Dystrophic epidermolysis bullosa

Mélissa Dussoyer

► **To cite this version:**

Mélissa Dussoyer. Identification of extracellular matrix biomarkers and targets in Dystrophic epidermolysis bullosa. Biochemistry, Molecular Biology. Université de Lyon, 2022. English. NNT : 2022LYSE1124 . tel-04748683

**HAL Id: tel-04748683**

**<https://theses.hal.science/tel-04748683v1>**

Submitted on 22 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : 2022LYSE1124

**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale Interdisciplinaire Sciences-Santé (205)**

**Spécialité de doctorat : Biochimie**

Soutenue publiquement le 08/07/2022, par :  
**Mélissa Dussoyer**

---

**Identification of extracellular matrix biomarkers and  
targets in Dystrophic epidermolysis bullosa**

---

Devant le jury composé de :

Ulrich Valcourt, Professeur, Université Claude Bernard Lyon 1, France  
Marie-Pierre Bousquet, Maîtresse de conférences, Université Toulouse III Paul Sabatier, France  
Alain Colige, Maître de recherche, Université de Liège, Belgique  
Matthias Titeux, Chargé de recherche, INSERM, France

Président  
Rapporteure  
Rapporteur  
Examineur

Catherine Moali, Directrice de recherche, CNRS, France  
Patricia Rousselle, Directrice de recherche, CNRS, France

Directrice de thèse  
Co-directrice de thèse

# UNIVERSITE CLAUDE BERNARD – LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	Mme Céline BROCHIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAND

## COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud – Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directeur : M. Claude DUSSART

## COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Directeur : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : M. Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLE

## ACKNOWLEDGEMENTS

---

Je tiens tout d'abord à m'adresser à ma directrice de thèse, Catherine Moali. Un grand merci pour ce travail de recherche que nous avons partagé et pour la confiance que tu m'as accordée tout au long de ce projet. Ton accompagnement et les enseignements que tu as partagés avec moi m'ont permis de gagner en assurance et en maturité scientifique. Tu m'as toujours encouragée à me poser plus de questions, à explorer mes hypothèses et mes résultats plus en profondeur. Tu as aussi su te rendre disponible (malgré un emploi du temps qu'on peut qualifier de très chargé !), m'écouter et m'encourager de sorte à ce que ces trois années soient riches et stimulantes sur tous les plans. Je suis fière d'avoir été ton étudiante et d'avoir appris à tes côtés !

Mes pensées vont ensuite vers Sandrine, ma colocataire au laboratoire. Merci d'avoir partagé avec moi ton bureau, mais aussi ta riche expérience scientifique et tes idées sur tous les sujets que l'on a pu aborder ensemble pendant cette thèse. J'ai, je pense, beaucoup de chance d'avoir pu ponctuer la quasi-totalité de mes journées de travail par des discussions prenantes sur l'actualité ou sur nos vies respectives (à part les jours où j'étais un peu « chonchon », et j'en suis désolée). J'ai beaucoup appris à ton contact, tant sur le plan scientifique qu'humain, et je peux d'ores et déjà annoncer que tu as placé la barre très haut pour mon prochain voisin de bureau !

Merci également à Patricia, pour l'expertise scientifique que tu as apportée tout au long de cette thèse, pour tes encouragements et pour nos riches discussions à chacune de mes visites dans ton bureau. La peau, les lames basales et leurs occupants ont aujourd'hui un peu moins de secrets pour moi grâce à tes nombreux éclairages !

Vient ensuite Priscillia, ma *partner in crime* sans qui cette thèse n'aurait pas été aussi pailletée ! Je te remercie pour tes conseils, ton soutien, tes blagues, ton sourire au quotidien, et ton énergie débordante qui m'a entraînée durant ces trois années. Merci pour les bêtises que l'on a faites, celles que l'on a évitées, les fous-rires au labo et ailleurs, les midis running (faute d'avoir réussi à me lever tôt pour des sessions matinales !). Merci pour ta générosité, ton écoute, et aujourd'hui ton amitié. Tu es une de ces très belles rencontres que la vie nous offre de temps à autre, et toujours pour le meilleur.

Un grand merci au reste de l'équipe Métalloprotéases et Remodelage tissulaire, Natacha, Cindy et Manon, pour votre bonne humeur, votre sourire, pour les manip sous-traitées (vive le N-TAILS, l'intéressée se reconnaîtra...) et ou terminées lorsque je partais en enseignement. La vie au laboratoire

## Acknowledgements

n'aurait pas été aussi agréable et motivante sans les nombreuses discussions partagées à la paillasse, au bureau ou en pauses déjeuner et goûter. J'espère que j'aurais la chance de participer à la partie de Molky annuelle de l'équipe encore cette année !

Une mention spéciale pour Adeline et Frédéric, qui ont grandement contribué à ma formation en protéomique et en spectrométrie de masse en général. Malgré mes incessantes demandes de traitement, de retraitement, de re-retraitement des données (et aussi un petit incident matériel, je vais mettre du temps à m'en remettre), vous m'avez toujours reçue avec le sourire et accompagnée au mieux avec vos conseils avisés. Votre bureau a même fini par devenir mon deuxième spot de travail ces derniers mois ! Un grand merci pour votre douceur, votre bienveillance et vos encouragements tout au long de ces trois années de travail.

Je pense ensuite à Laura, Shérine et Lisa, des doctorantes d'énergie et de caractère présentes au quotidien, pour papoter de science, ou de tout et de rien au gré de nos rencontres dans l'institut ! Je remercie aussi Christelle, Nicolas, Naïma, Marielle et Virginie pour leur aide dans mes différentes expérimentations et les conseils prodigués sur mes travaux.

I also thank our German collaborators for the joint work we conducted these last three years. Many thanks to Alexander for his optimism, his availability and the expertise generously shared with me during my thesis. Thanks to Svenja for her support with samples, parcels and her welcoming when I came to Freiburg. To Dimitra too, thank you for the clinical coloration you gave to the project.

J'ai aussi eu la chance de partager un peu du quotidien des enseignants de l'Université Claude Bernard-Lyon 1. Je remercie les membres des équipes de Biochimie, notamment Leyre et Sébastien, de Biologie Cellulaire, en particulier Christelle, Sandrine et Ulrich, et du SOIE, pour m'avoir intégrée au corps enseignant et aidée à mieux cerner les besoins et les attentes des étudiants. J'espère avoir gagné en pédagogie à votre contact !

Je remercie également les membres de mon comité de suivi de thèse, Virginie Brun et Ulrich Valcourt, pour avoir suivi mes travaux pendant trois ans et pour les conseils éclairés qu'ils m'ont prodigués à chacune de nos réunions. Merci aux membres de mon jury, Marie-Pierre Bousquet, Alain Colige, Matthias Titeux, et à nouveau Ulrich Valcourt, pour l'évaluation de cette thèse et pour les échanges riches et constructifs qui ont rythmé ma soutenance.

Je remercie enfin mes proches, en premier lieu mes parents. Merci de m'avoir toujours soutenue dans mes projets, et donné la chance de chercher le domaine dans lequel je pouvais m'épanouir. Vous m'avez aussi aidée à comprendre quelles étaient les qualités les plus indispensables en tant qu'être

## Acknowledgements

humain, quelle femme je souhaitais devenir, et quelle empreinte je voulais laisser de mon passage, en me donnant l'envie de faire bouger les choses plutôt que de simplement les contempler. Un grand merci enfin pour votre support logistique pendant mes années d'études, notamment les visites d'appartements, les emménagements/déménagements ou encore les nombreux aller-retour à la gare les soirs.

A ma sœur Marjorie, un grand merci pour tes encouragements et ta présence, durant cette thèse certes, mais surtout depuis notre enfance. Ton soutien indéfectible et notre complicité ont été et demeurent un des piliers de mon petit monde, certainement le plus solide mais aussi le plus précieux. Sans lui, sans ta générosité ni ta bienveillance, je ne serais pas devenue la femme que je suis aujourd'hui (sans les chansons Disney aussi, mais bon, elles passent après quand même !).

A Gabriel également, un peu plus récemment arrivé dans la famille mais tout aussi important. Merci pour ta bienveillance aussi débordante que celle de ma sœur, pour l'intérêt que tu as toujours porté à mes travaux et pour ton grain de bêtise qui rend la vie incontestablement plus légère.

A Hélène, Mélanie et Mathilde, des super-nanas avec un grain de folie que j'ai eu la chance de croiser au gré de ma vie étudiante pour ne plus les lâcher par la suite. Vous êtes de ces personnes lumineuses et bienveillantes, qui égayent l'existence des autres avec un naturel déconcertant, tout simplement. Merci pour votre présence et votre soutien, pour les journées, soirées, week-ends partagés ces huit dernières années, qui ont dynamisé ma vie tout en y apportant une touche de douceur. Le présent de votre amitié est un des plus beaux que j'ai eu la chance de recevoir, et je fais le vœu que cette belle route que nous avons pris ensemble soit encore très longue et heureuse.

Et pour terminer, à Simon, merci pour tout ce que tu sais déjà. Merci de traverser les moments les plus heureux comme les plus pénibles avec moi, merci pour l'équilibre et le soutien que tu m'apportes au quotidien, et pour les rires que tu arrives toujours à faire naître même dans les situations délicates. Merci de me faire partager ta vie, d'illuminer la mienne et de faire avancer un peu plus chaque jour la nôtre.

# ABSTRACT

---

Dystrophic epidermolysis bullosa (DEB) is an inherited disease caused by mutations in the *COL7A1* gene encoding collagen VII. This protein assembles into anchoring fibrils and is key for the stability of basement membranes, especially in the dermo-epidermal junction (DEJ) of the skin. In patients suffering from the most severe forms of the disease (recessive DEB or RDEB), the deficiency in collagen VII is responsible for epidermal blistering, progressively evolving into slow-healing wounds, soft-tissue fibrosis and aggressive cancers. To monitor disease progression and develop novel therapies, there is a crucial need to identify new biomarkers and targets. The main goal of this study was to meet these needs using an extracellular matrix (ECM)-centered approach. Thus, we aimed at describing ECM alterations in RDEB skin to propose new therapeutic tools in clinic.

We first defined the ECM enrichment protocols which were most well-suited to the characterization of mouse skin matrisome by mass spectrometry-based proteomics. Using a RDEB mouse model, we next applied the developed methods to quantitatively compare the skin matrisome of control and RDEB mice. Imaging by transmission electron microscopy was also performed to evaluate ECM damages at the tissue level. Finally, the biosynthesis of procollagen  $\alpha 1(I)$ , a major ECM constituent, was analyzed in more detail.

The skin of RDEB mice displayed strong dysregulations at the molecular and tissue levels. On the DEJ epidermal side, a decrease in the size of hemidesmosomes and impairments in their distribution along the basement membrane were detected. In the dermal ECM, decreases in the abundance of some proteoglycans and fibrillar collagens, especially collagen I, were surprisingly detected, combined with a reduction in the diameter of collagen fibrils and a less compact organization of collagen bundles. When the biosynthesis of procollagen  $\alpha 1(I)$  was investigated more thoroughly, diminutions in the expression of *Colla1* and in the abundance of the collagen-specific chaperone heat shock protein-47 (HSP-47) were observed. In the ECM, a global decrease in the amount of secreted procollagen  $\alpha 1(I)$  was also evidenced. In addition, the abundance of bone morphogenetic protein-1 (BMP-1), a protease involved in the elaboration of the collagen network, and of its regulator procollagen C-proteinase enhancer-1 (PCPE-1) were decreased at the RNA and protein levels. If defects in the C-terminal processing of procollagen  $\alpha 1(I)$  by the protease could not be conclusively demonstrated due to technical limitations, the impairment in BMP-1 activity is likely to affect additional molecules involved in collagen fibrillogenesis such as proteoglycans and cross-linking enzymes.

Taken together, our data underlined strong damages in the composition and in the organization of the DEJ and dermal ECM in RDEB skin and demonstrated the relevance of our ECM-centered strategy. The fibrillar collagen network, which acts as a scaffolding component for the whole skin ECM, was massively affected. Even if additional analyses are now needed to understand these dysregulations, the reported alterations already indicate some collagen-related proteins and proteolytic fragments with biomarker or therapeutic target potential.

## RESUME SUBSTANTIEL

---

L'épidermolyse bulleuse dystrophique (DEB en anglais) est une maladie génétique affectant le gène codant pour le collagène VII (*COL7A1*), une protéine matricielle stabilisant les lames basales par la formation de fibrilles d'ancrage. Au sein de la peau, le collagène VII est notamment un composant clé de la jonction dermo-épidermique (DEJ), un assemblage de protéines membranaires et matricielles assurant la cohésion entre l'épiderme et le derme sous-jacent. Les formes les plus sévères de la maladie, observées dans le cas de l'épidermolyse bulleuse dystrophique récessive (RDEB), se traduisent entre autres par des décollements chroniques de l'épiderme, conduisant à des plaies non cicatrisantes qui évoluent progressivement vers une fibrose du derme. L'environnement hautement inflammatoire présent au niveau de ces lésions chroniques favorise également le développement de carcinomes à cellules squameuses (SCC) de très mauvais pronostic. Si plusieurs essais cliniques visant à restaurer un gène *COL7A1* fonctionnel ou à pallier le manque de collagène VII par des apports exogènes sont en cours, ce sont les thérapies destinées à limiter l'apparition et la progression des symptômes qui semblent, à l'heure actuelle, les plus promptes à améliorer la qualité et l'espérance de vie des patients. Développer de telles approches nécessite l'identification préalable de cibles thérapeutiques, mais aussi de biomarqueurs qui permettraient de suivre l'efficacité d'un traitement ou encore l'évolution de la maladie, afin d'adapter la prise en charge clinique des patients.

L'objectif de ma thèse a été défini de sorte à répondre à ces besoins, en adoptant une approche centrée sur la matrice extracellulaire (ECM) de la peau, qui est le siège de nombreuses modifications au cours de la progression de la maladie. Mes travaux ont consisté à décrire et comprendre les altérations de l'ECM de la peau RDEB, de sorte à proposer de potentiels nouveaux biomarqueurs et/ou cibles thérapeutiques issus de ce compartiment tissulaire pour la maladie.

Pour ce faire, un modèle murin de RDEB (souris hypomorphiques pour le collagène VII) a été utilisé et la matrice de la peau entière a été caractérisée par différentes approches protéomiques (spectrométrie de masse de découverte ou ciblée, Western blot) et d'imagerie (microscopie électronique à transmission TEM, immunofluorescence). Dans un premier temps, des protocoles d'enrichissement des protéines matricielles de la peau de souris ont été mis au point et comparés afin de caractériser au mieux le matrisome de l'organe. Les résultats obtenus ont ensuite été appliqués à la description des dérégulations de l'ECM de la peau de souris RDEB. En parallèle, l'ultrastructure de la peau malade a été caractérisée à l'échelle tissulaire, avec un intérêt particulier pour la DEJ et les fibres de collagène du derme. Enfin, certaines altérations mises à jour par ces différentes approches ont été étudiées plus en profondeur de sorte à expliquer leur développement dans le contexte de la pathologie.

Nos résultats ont tout d'abord permis de déterminer les stratégies d'enrichissement de l'ECM les plus adaptées à l'étude du matrisome de la peau de souris entière par spectrométrie de masse. Afin d'éliminer les protéines cellulaires du tissu cutané, des protocoles utilisant des sels et détergents en concentrations connues (décellularisation chimique, *quantitative detergent solubility profiling* dit « QDSP ») ou non (enrichissement compartimental, protocole de référence dans la littérature pour les études de matrisomes), mais aussi la phospholipase A2 (décellularisation enzymatique), ont été évalués. De façon surprenante, le protocole de référence (compartimental) s'est révélé le moins adapté à l'étude de la peau, avec une identification des composés du matrisome inférieure de 20 % à celle obtenue *via* les trois autres approches. Les trois autres protocoles ont conduit à des résultats similaires mais la décellularisation chimique semble la plus compétitive en termes de coût de mise en œuvre (réactifs et analyses). Ces analyses ont aussi révélé une très large gamme de solubilité pour les protéines de l'ECM de la peau, soulignant ainsi la nécessité d'étudier plusieurs extraits d'enrichissement en plus de la fraction insoluble afin d'obtenir une vision aussi exhaustive que possible du matrisome. Au total, ces études de mise au point ont permis l'identification de 236 protéines de l'ECM et ont conduit à la première description du matrisome de la peau de souris saine dans la littérature.

Le protocole d'enrichissement de l'ECM sélectionné ci-dessus (décellularisation chimique) a ensuite permis la comparaison quantitative du matrisome de la peau de souris adultes saines et RDEB (stade pré-lésion). La fraction insoluble du protocole compartimental a également été intégrée dans l'analyse, malgré les limites relevées ci-dessus, dans le but de faciliter la comparaison avec la littérature abondante qui y fait référence. Cent quatre-vingt-cinq composés du matrisome ont ainsi été identifiés, 101 ont pu être quantifiés dans tous les réplicats et 22 présentaient une abondance significativement modifiée en condition pathologique. De façon très inattendue, une forte sous-représentation des chaînes des principaux collagènes fibrillaires du derme, en particulier du collagène I (variations significatives), a été observée dans la peau RDEB. Une très forte diminution de plusieurs inhibiteurs de protéase extracellulaires a aussi été relevée, ainsi qu'une tendance à la baisse pour les protéoglycanes tels que la décorine, le biglycan et le lumican qui régulent les propriétés des fibres de collagène et l'inflammation. Ces résultats sont cohérents avec les altérations du réseau des collagènes fibrillaires observées en TEM à l'échelle tissulaire dans la matrice du derme (diamètres des fibrilles de collagène et densité globale du réseau de collagène diminués). De façon surprenante, ces altérations tissulaires ont été détectées dans la peau du dos des souris adultes, dans les pattes fibrotiques mais également très précocement dans la peau de souriceaux RDEB nouveaux-nés. De plus, il est apparu que l'absence de fibrilles d'ancrage fonctionnelles avait des répercussions sur la composition de la lame basale (abondance en chaînes de laminines et de collagène IV, en nidogène-1 tendant à la hausse) mais aussi sur l'organisation épidermique de la DEJ, avec une réduction de la taille et de la distribution des hémidesmosomes le long de la lame basale.

Par différentes méthodes, nous avons enfin montré que les dérégulations affectant la chaîne de collagène  $\alpha 1(I)$ , principal constituant du collagène I sans lequel les trimères de collagène I ne peuvent se former, résultaient d'un défaut dans l'expression du gène correspondant *Coll1a1* ayant pour conséquence une diminution du taux de procollagène  $\alpha 1(I)$  dans la peau de souris RDEB. Une baisse de la quantité de HSP-47 (heat shock protein-47), une chaperone nécessaire à la structuration des trimères de collagène fibrillaire au niveau intracellulaire, a renforcé la piste d'un défaut de biosynthèse par les cellules. Ceci a pour conséquence de modifier le rapport entre les différents collagènes I/III/V dans les fibres et pourrait expliquer certaines altérations observées à l'échelle tissulaire. En complément, une baisse dans l'expression de BMP-1 (bone morphogenetic protein-1), une protéase matricielle impliquée dans l'élaboration du réseau des collagènes fibrillaires, et de son régulateur PCPE-1 (procollagen C-proteinase enhancer-1) a été mise en évidence, à l'échelle des transcrits mais aussi des protéines. Si aucun défaut dans le processing C-terminal du procollagène  $\alpha 1(I)$  par la protéase n'a pour le moment pu être validé du fait de la faible teneur en procollagène I dans la peau RDEB, une dérégulation de BMP-1 demeure susceptible d'affecter d'autres molécules impliquées dans la fibrillogénèse des collagènes (protéoglycanes, enzymes de réticulation, thrombospondine-1).

L'ensemble de nos résultats suggère une altération majeure dans l'élaboration de la DEJ mais aussi du réseau des collagènes fibrillaires dans la peau de souris RDEB. De façon marquante, cette dernière dérégulation semble résulter directement de la déficience en collagène VII et non d'un processus développé progressivement au cours de la maladie, tel que l'inflammation. Seuls les collagènes fibrillaires et leur maturation C-terminale ont été étudiés en détails au cours de ce travail, mais nos données suggèrent que de nombreux autres composés matriciels impliqués dans la fibrillogénèse des collagènes pourraient être affectés au cours de la pathologie RDEB. C'est par exemple le cas pour la décorine, le biglycan et le lumican, connus pour moduler la croissance latérale des fibres, ou encore pour le collagène XIV, un autre régulateur des propriétés de surface des fibres. D'autre part, nos travaux soulignent que des processus intracellulaires de biosynthèse du collagène semblent aussi affectés par la maladie. Des études supplémentaires seront donc certainement nécessaires pour comprendre tous les mécanismes à l'origine des dérégulations rapportées ici. Par exemple, des analyses de la protéolyse matricielle ont été initiées afin de mieux décrire la maturation et la dégradation des composés de l'ECM, dont les collagènes. De façon générale, l'ensemble des dérégulations détectées au cours de ce travail sont susceptibles d'entraîner des variations dans les molécules circulantes dérivées des collagènes, qui pourraient alors être utilisées comme biomarqueurs. Le développement d'une méthode PRM (parallel reaction monitoring) de quantification des biomarqueurs potentiels liés à la chaîne  $\alpha 1(I)$  du collagène a été initié dans ce sens. De plus, si les modifications dans les acteurs et processus associés à la fibrillogénèse étaient confirmées dans la pathologie humaine (HSP-47, BMP-1, PCPE-1), ces derniers pourraient constituer de nouvelles cibles thérapeutiques.

En conclusion, ce travail démontre la pertinence de l'étude de la matrice extracellulaire pour identifier des biomarqueurs et cibles thérapeutiques de la pathologie RDEB. Il ouvre de nombreuses perspectives très prometteuses pour améliorer la prise en charge des symptômes cutanés les plus invalidants.

# TABLE OF CONTENTS

---

<b>Acknowledgements</b>	<b>3</b>
<b>Abstract</b>	<b>6</b>
<b>Résumé substantiel</b>	<b>8</b>
<b>Table of contents</b>	<b>12</b>
<b>Abbreviations</b>	<b>17</b>
<b>List of figures</b>	<b>20</b>
<b>List of tables</b>	<b>23</b>
<b>Chapter I: Introduction</b>	<b>24</b>
<b>I. Skin: structure and functions</b>	<b>25</b>
1. Epidermis	26
2. Dermo-epidermal junction	28
3. Dermis	29
i. Organization of the dermal extracellular matrix	29
ii. Mesenchymal resident cells	31
4. Hypodermis	31
5. Vasculature and innervation	32
6. Skin immunity and healing	32
<b>II. The extracellular matrix, a complex and dynamic network</b>	<b>36</b>
1. Collagens: the basis of skin ECM	36
i. Collagens: structure and expression in skin	36
ii. Biosynthesis of fibrillar collagens	37
iii. Biosynthesis of collagen VII	38
2. Glycosaminoglycans and proteoglycans in skin matrix	39
3. Metzincins and proteolysis-driven dynamics of the matrix	40
i. Matrixins	41
ii. Astacins	42
iii. Adamalysins	43
4. The Matrisome project	45
i. Matrisome definition	45
ii. <i>In silico</i> build-up of the matrisome	45

## Table of contents

iii.	Tools for the analysis of the matrisome	48
<b>III.</b>	<b>Dystrophic epidermolysis bullosa</b>	<b>49</b>
1.	Etiology, classification and symptoms	49
2.	Research on DEB	51
i.	Overview of medical research	51
ii.	RDEB mouse model	53
3.	ECM dysregulations in DEB	54
i.	Modifications in the basement membrane	54
ii.	Dermal fibrosis	55
<b>IV.</b>	<b>Proteomics and mass spectrometry</b>	<b>58</b>
1.	Introduction to mass spectrometry	58
i.	Principle of mass spectrometry	58
ii.	MS to study proteins	59
2.	Approaches in MS-based proteomics	60
i.	Top-down and bottom-up strategies	61
ii.	Data acquisition	61
iii.	Discovery and targeted proteomics	63
3.	Quantification in MS-based proteomics	64
i.	Overview of absolute quantification	65
ii.	Relative quantification methods	65
4.	Studying the ECM by MS-based proteomics	67
	<b>Objectives</b>	<b>69</b>
	<b>Chapter II: Materials and methods</b>	<b>71</b>
<b>I.</b>	<b>General procedures and reagents</b>	<b>72</b>
1.	Recovery of biological samples	72
i.	Biological samples for the comparison of enrichment protocols	72
ii.	RDEB mouse model and recovery of samples	72
2.	Chemicals	72
<b>II.</b>	<b>Matrisome enrichment protocols</b>	<b>74</b>
1.	Chemical decellularization	74
2.	Compartmental enrichment	76
<b>III.</b>	<b>MS-based discovery proteomics</b>	<b>78</b>
1.	Mass spectrometry for the comparison of enrichment protocols	78
2.	Quantitative MS-based proteomics	78
i.	Sample preparation	78

## Table of contents

ii.	TMT labeling of peptides	79
iii.	HPLC-MS/MS analyses	79
iv.	MS/MS data processing	80
v.	Analysis of quantitative data	80
vi.	Data processing using SwissProt database (first matrisome analysis test, abundance ratios of collagen chains and peptide mapping)	81
<b>IV.</b>	<b>Transmission electron microscopy</b>	<b>83</b>
1.	Sample preparation	83
2.	Image acquisition	83
3.	Hemidesmosome and collagen fibrils analyses, statistics	83
<b>V.</b>	<b>Other protein analyses</b>	<b>85</b>
1.	Immunoblotting	85
i.	Sample preparation	85
	Skin lysates	85
	Enrichment fractions	85
	Fibroblast lysates	86
ii.	Western blotting	86
iii.	Quantification and statistics	87
2.	Immunofluorescence on tissue sections	88
<b>VI.</b>	<b>RNA analyses</b>	<b>89</b>
1.	RNA extraction	89
2.	Reverse transcription	89
3.	Quantitative real time PCR	89
4.	Computing of gene expression and statistics	90
<b>VII.</b>	<b>MS-based targeted proteomics (parallel reaction monitoring)</b>	<b>92</b>
1.	Sample preparation	92
i.	Synthetic peptides	92
ii.	NaCl fractions	92
iii.	Skin lysates	93
2.	PRM analysis	93
3.	Data analysis	93
<b>Chapter III: Results</b>		<b>94</b>
<b>I.</b>	<b>Development of protocols for the analysis of mouse skin matrisome</b>	<b>95</b>
1.	Preliminary ECM enrichment tests	95
2.	Comparison of different ECM enrichment protocols for mouse skin	98

## Table of contents

3.	Final selection of ECM-enrichment protocols	118
<b>II.</b>	<b>Global characterization of the extracellular matrix in the skin of RDEB mice</b>	<b>120</b>
1.	Selected mouse samples	120
2.	Global description of the RDEB skin matrisome by MS-based proteomics	120
i.	Qualitative overview of the MS/MS analyses	120
ii.	Analysis of quantitative data	125
iii.	Impaired biological processes	130
3.	Analysis of skin ultrastructure by transmission electron microscopy	131
i.	Characterization of the dermo-epidermal junction in RDEB skin	131
DEJ in the back skin	131	
Analysis of the hemidesmosomes in paws	136	
ii.	Characterization of the collagen network in RDEB skin	137
Characterization of the collagen network in the back skin	137	
Characterization of the collagen network in the paw skin	140	
4.	Discussion	142
<b>III.</b>	<b>Preliminary evaluation of the mechanisms involved in the collagen phenotype in RDEB skin</b>	<b>146</b>
1.	Analyses of mature collagen forms	146
2.	Expression of procollagens I and III in RDEB skin	151
i.	Expression of the genes encoding collagens I and III	151
ii.	Expression of procollagen I at the protein level	152
3.	Evaluation of the protease-dependent maturations of fibrillar collagens	156
i.	MS/MS and Western blot data for the analysis of proteolysis	156
ii.	Expression of procollagen C-proteinases and related partners in RDEB skin	159
iii.	Preliminary evaluation of the expression of HSP-47	164
4.	Proof-of-concept: development of a PRM assay to monitor procollagen-derived biomarkers in skin	166
i.	PRM analysis of unlabeled samples	166
ii.	PRM-analysis of TMT-labeled samples	168
5.	Discussion	171
	<b>General discussion and conclusion</b>	<b>178</b>
	<b>Bibliography</b>	<b>185</b>
	<b>Appendices</b>	<b>199</b>
	<b>Publications</b>	<b>216</b>
	<b>Oral and poster communications, prizes</b>	<b>296</b>

Table of contents

<b>Training and complementary activities</b>	<b>298</b>
<b>Funding</b>	<b>299</b>

# ABBREVIATIONS

---

AA: amino-acid  
ACN: acetonitrile  
ADAM: a disintegrin and metalloproteinase  
ADAMTS: a disintegrin and metalloproteinase with thrombospondine motifs  
ANR: agence nationale de la recherche (French national agency for research)  
BMP-1: bone morphogenetic protein-1  
BTP: BMP-1/tolloid-like proteinase  
CAPS: N-cyclohexyl-3-aminopropanesulfonic acid  
CMT: combinatorial isobaric mass tag  
CP: procollagen C-propeptide  
CP  $\alpha$ 1(I): procollagen  $\alpha$ 1(I) C-propeptide  
DAPI: 4',6-diamidino-2-phenylindole  
DDA: data-dependent acquisition  
DEB: dystrophic epidermolysis bullosa  
DDEB: dominant dystrophic epidermolysis bullosa  
DEJ: dermo-epidermal junction  
DFG: deutsche forschungsgemeinschaft (German research foundation)  
DIA: data-independent acquisition  
DiLeu: N-N-dimethyl leucine  
DNA: deoxyribonucleic acid  
DNase I: deoxyribonuclease I  
DTT: dithiothreitol  
EB: epidermolysis bullosa  
ECM: extracellular matrix  
EDTA: ethylenediaminetetraacetic acid  
Ef: efficiency  
EGF: epidermal growth factor  
ESI: electrospray ionization  
ER: endoplasmic reticulum  
FA: formic acid  
FACIT: fibril-associated collagen with interrupted triple helices  
FGF: fibroblast growth factor  
GAG: glycosaminoglycan

## Abbreviations

GE: gene expression  
GLT25D1: procollagen galactosyltransferase-1  
GLT25D2: procollagen galactosyltransferase-2  
GOI: gene of interest  
GuHCl: guanidinium hydrochloride  
HCD: higher-energy collisional dissociation  
HD: hemidesmosome  
HEPES: N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid  
HPLC: high performance liquid chromatography  
HPLC/MS/MS: liquid chromatography/tandem mass spectrometry  
HSP-47: heat shock protein-47  
HTRA-1: high-temperature requirement protein-A1  
IAA: iodoacetamide  
IL: interleukin  
iTRAQ: isobaric tag for relative and absolute quantitation  
KO: knock-out  
LBTI: laboratory of tissue biology and therapeutic engineering  
LM-332: laminin-332  
Lys-C: endoproteinase Lys-C  
MMP: matrix metalloproteinase  
MS: mass spectrometry  
MS/MS: tandem mass spectrometry  
mTLD: mammalian tolloid protein  
mTLL-1: mammalian tolloid-like 1  
mTLL-2: mammalian tolloid-like 2  
MT-MMP: membrane-type matrix metalloproteinase  
N-TAILS: N-terminal amine isotopic labeling of substrates  
LOX: lysyl oxidase  
LOXL: lysyl oxidase homolog-1  
PBS: phosphate-buffered saline  
PCPE-1: procollagen C-proteinases enhancer-1  
PCPE-2: procollagen C-proteinases enhancer-2  
pN-proteinase: procollagen N-proteinase  
PCR: polymerase chain reaction  
PDGF: platelet-derived growth factor  
PI: protease inhibitors  
PLA2: phospholipase-A2

## Abbreviations

PLOD: procollagen-lysine-1, 2-oxoglutarate 5-dioxygenase  
PNGase F: N-glycosidase F  
PTC: premature termination codon  
PTM: post translational modification  
PVDF: polyvinylidene difluoride  
QDSP: quantitative detergent solubility profiling  
RG: reference gene  
RDEB: recessive dystrophic epidermolysis bullosa  
RG: reference gene  
RIPA (buffer): radioimmunoprecipitation assay (buffer)  
RNA: ribonucleic acid  
RT: reverse transcription  
SCC: squamous cell carcinoma  
SD: sodium deoxycholate  
SDS: sodium dodecyl sulfate  
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis  
SILAC: stable isotope labeling of amino acids in cell culture  
SLRP: small leucine-rich proteoglycan  
SOT: sulfoxide tag  
TCE: 2,2,2-trichloroethanol  
TEM: transmission electron microscopy  
TFA: trifluoroacetic acid  
TGF- $\beta$ : transforming growth factor  $\beta$   
TGFBI: transforming growth factor  $\beta$ -induced protein  
TIMP: tissue inhibitor of metalloproteinase  
TMT: tandem mass tag  
TSP-1: thrombospondin-1  
WB: western-blot  
WT: wild-type

# LIST OF FIGURES

---

Figure 1. General structure of the skin. ....	25
Figure 2. Structuration and fate of keratinocytes in the epidermis. ....	26
Figure 3. Simplified structure of the DEJ. ....	29
Figure 4. Human skin observed by photonic microscopy. ....	31
Figure 5. Summary of the healing process in skin. ....	34
Figure 6. Schematic representation of supramolecular assemblies formed by collagens in the skin. ...	36
Figure 7. Schematic representation of the production and processing of fibrillar collagens. ....	37
Figure 8. Extracellular processing and supramolecular assembly of collagen VII in the basement membrane of the skin. ....	39
Figure 9. Main proteases of metzincin family. ....	41
Figure 10. Methodology for the definition of the matrisome. ....	46
Figure 11. Schematic view of human and mouse matrisomes. ....	47
Figure 12. RDEB micro- and macroscopic phenotypes. ....	50
Figure 13. Strategies for the development of new therapies for DEB. ....	51
Figure 14. Phenotype of the RDEB mouse. ....	54
Figure 15. Summary of the processes involved in fibrosis (general processes currently described in literature). ....	56
Figure 16. Schematic representation of the key steps and equipment composition in mass spectrometry. ....	58
Figure 17. Nomenclature of the MS <sup>2</sup> ions that can be generated during fragmentation by mass spectrometry. ....	59
Figure 18. Strategies for the acquisition of data in MS-based proteomics. ....	62
Figure 19. Schematic view of discovery and targeted approaches in MS-based proteomics. ....	63
Figure 20. Schematic representation of strategies for relative quantification and their main steps. ....	65
Figure 21. Structure of TMT 11-plex tag. ....	67
Figure 22. Schematic view of the statistical approach applied to the MS/MS quantitative data from the two independent experiments (n = 5 mice/genotype in total). ....	81
Figure 23. Protein extraction during compartmental enrichment on mouse skin. ....	98
Figure 24. Overview of the matrisome content in the enrichment fractions analyzed by MS/MS. ....	122
Figure 25. Heatmap of the abundance ratio of core matrisome proteins from WT and RDEB mouse skin. ....	127
Figure 26. Heatmap of the abundance ratio of associated matrisome proteins from WT and RDEB mouse skin. ....	129
Figure 27. Altered biological processes in RDEB mouse skin. ....	130

## List of figures

Figure 28. TEM observations of the DEJ in the back skin of WT and RDEB mice aged 9-12 weeks. .....	133
Figure 29. Analysis of the hemidesmosomes of basal keratinocytes in the back skin of WT and RDEB mice. .....	135
Figure 30. Analysis of the hemidesmosomes of basal keratinocytes in the forepaw skin of WT and RDEB mice aged 9-12 weeks. ....	136
Figure 31. TEM observations of the collagen network in the back skin of WT and RDEB mice aged 9-12 weeks. ....	138
Figure 32. Analysis of collagen fibrils by TEM in the back skin of WT and collagen VII-hypomorphic mice. ....	139
Figure 33. Analysis of collagen fibrils in the skin of the forepaw of WT and RDEB mice aged 9-12 weeks. ....	140
Figure 34. Immuno- and MS/MS-detection and quantification of collagen I in the back skin of WT and RDEB mice aged 9-12 weeks. ....	149
Figure 35. Abundance ratios between the different chains of fibrillar collagens in WT and RDEB insoluble enrichment fractions. ....	151
Figure 36. Expression of the genes encoding fibrillar $\alpha 1(I)$ , $\alpha 2(I)$ and $\alpha 1(III)$ collagen chains in the back skin of WT and RDEB mice aged 9-12 weeks. ....	151
Figure 37. Abundance ratio between collagen $\alpha 1(I)$ , $\alpha 2(I)$ and $\alpha 1(III)$ at transcript and protein levels in the back skin of WT and RDEB mice aged 9-12 weeks. ....	152
Figure 38. Immuno-detection of procollagen $\alpha 1(I)$ and of other C-propeptide $\alpha 1(I)$ -containing forms in the NaCl fraction obtained from the back skin of WT and RDEB mice aged 9-12 weeks. ....	153
Figure 39. Immuno-detection of procollagen $\alpha 1(I)$ and of other C-propeptide $\alpha 1(I)$ -containing forms in the back skin of WT and RDEB mice aged 9-12 weeks. ....	156
Figure 40. Peptide mapping for collagens $\alpha 1(I)$ and $\alpha 2(I)$ in the enrichment fractions obtained from the back skin of WT and RDEB mice aged 9-12 weeks. ....	157
Figure 41. Peptide mapping for collagens $\alpha 1(III)$ , $\alpha 1(V)$ and $\alpha 2(V)$ in the insoluble fractions obtained from the back skin of WT and RDEB mice aged 9-12 weeks. ....	158
Figure 42. Expression of the genes encoding BMP-1/mTLD, mTLL-1 and PCPE-1 in the back skin of WT and RDEB mice aged 9-12 weeks. ....	160
Figure 43. Immunofluorescent labeling of the main BMP-1/tolloid-like proteinases (BMP-1/mTLD and mTLL-1) in the back skin of WT and RDEB mice aged 9-12 weeks. ....	161
Figure 44. Immuno-detection of PCPE-1 in skin lysates from WT and RDEB mice aged 9-12 weeks. .....	162
Figure 45. Immunofluorescent labeling of PCPE-1 in the skin of WT and RDEB mice aged 9-12 weeks. .....	163

## List of figures

Figure 46. Expression of HSP-47 at transcript and protein levels in the back skin of WT and RDEB mice aged 9-12 weeks. ....	166
Figure 47. Schematic representation of the procollagen $\alpha 1(I)$ cleavage products generated by BMP-1. ....	167
Figure 48. Detection of the ADDANVVR peptide by PRM. ....	168
Figure 49. Detection of TMT-labeled ADDANVVR and DDANNVR peptides by PRM. ....	169
Figure 50. Detection of the SGEYWIDPNQGCNLDAIK peptide by PRM. ....	170
Figure 51. Summary of collagen biosynthesis related-processes dysregulated in RDEB mouse skin	181

# LIST OF TABLES

---

Table 1. Composition of the principal transmembrane and matrix proteoglycans contained in the skin. .....	40
Table 2. Examples of skin ECM substrates for the main metzincins expressed in normal skin. ....	44
Table 3. Examples of current clinical trials on DEB. ....	53
Table 4. Characteristics and quick comparison of top-down and bottom-up approaches in MS-based proteomics. ....	61
Table 5. Details of samples and processed images for hemidesmosome (HD) analyses in WT and RDEB mice. ....	83
Table 6. Details of samples and processed images for the measurements of the diameters of dermal collagen fibrils in WT and RDEB mice. ....	84
Table 7. Amounts of biological material and gel percentages used for immunoblotting in relation with target protein.....	86
Table 8. Antibodies used for immunoblotting and immunofluorescence. ....	87
Table 9. Composition of the mix used for the reverse transcription of 200 ng of RNA. ....	89
Table 10. Sequences and efficiency of primer pairs used for RT q-PCR. ....	90
Table 11. Composition of the mix used for qPCR. ....	90
Table 12. Matrisome proteins detected by MS/MS in the insoluble pellet after compartmental enrichment from mouse skin. ....	97
Table 13. Summary of total and matrisome proteins identified by MS/MS in mouse skin.....	121
Table 14. Matrisome proteins identified in the skin of WT and RDEB mice by MS-based quantitative proteomics. ....	125
Table 15. Possible hypotheses to explain the alterations of the collagen network in RDEB mouse skin. .....	146
Table 16. Summary of the conducted experiments and their results in relation with the collagen phenotype in RDEB skin.....	172

---

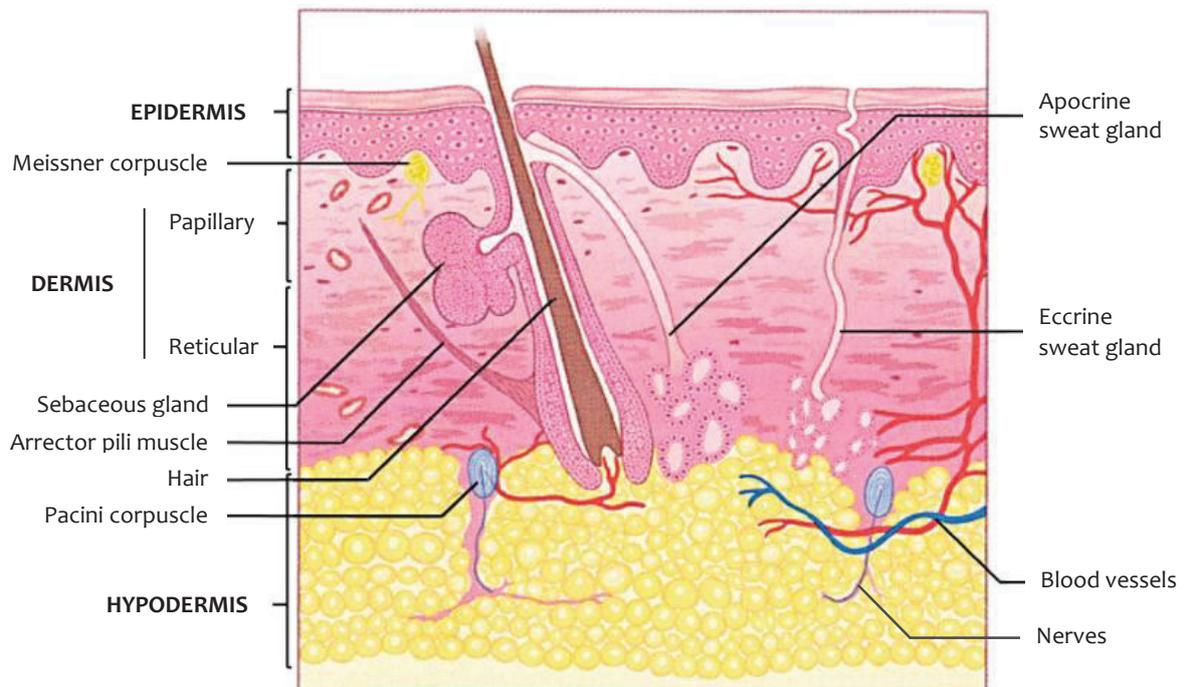
## CHAPTER I: INTRODUCTION

---

## I. SKIN: STRUCTURE AND FUNCTIONS

The skin is one of the largest organs in human, with the intestine and lungs, and covers almost all the surface of the organism in contact with the external environment. It accounts for 15 % of total body weight in adults, and exhibits an approximate surface between 1.5 and 2 square metre<sup>1</sup>. As the direct interface with the outer world, the skin ensures the homeostasis of the whole body by protecting the underlying organs against aggressions<sup>2</sup> (biotic, physical, chemical) and by allowing the sense of external stimuli. It is also involved in thermoregulation and performs endocrine activities. All these vital functions depend on the structural cohesion and functional coordination of skin components and activities.

The skin is composed of three superposed tissues (**Figure 1**). The epidermis is the outermost layer and consists of a protective multi-layered epithelium controlling exchanges of fluids and molecules between the body and the environment. The underlying dermis is a connective layer rich in extracellular matrix (ECM) and provides pliability to skin as well as metabolic support to the epidermis. Finally, the hypodermis is the deepest layer, interfaced with internal organs, which is mainly involved in lipid metabolism. The skin is highly innervated and vascularized by blood and lymph vessels. It exhibits a very dynamic homeostasis based on the permanent production, degradation and renewing of cells and ECM components to maintain its structural and functional properties.

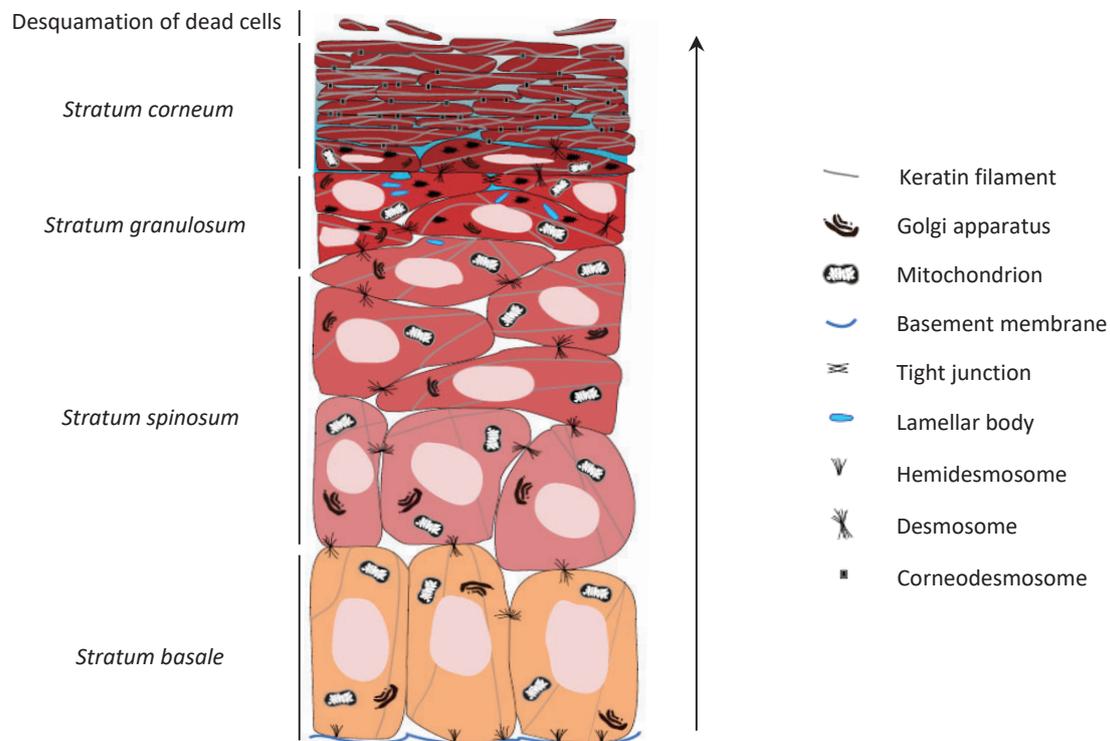


**Figure 1. General structure of the skin.**

The organ is made of three superposed tissues, each specialized in specific functions. Skin appendages (hair, sebaceous and sweat glands) result from profound invaginations of the epithelial tissue in the dermis; their growth/activities depend on nutrients supplied by blood vessels, and nervous stimulations. Adapted from<sup>3</sup>.

## 1. Epidermis

The epidermis is the most superficial compartment of the skin, at the direct interface between the environment and the body<sup>3</sup>. This protective epithelium contains sensitive nervous structures but is avascular: nutrients and gases diffuse from the external environment and from dermal blood vessels to allow the nutrition of epidermal cells. Its thickness ranges from 0.1 to 1.0 mm in accordance with the level of exposure to mechanical stresses of the different anatomical regions: the eyelid possesses the thinnest epidermis of the body, while palms and feet soles, perpetually submitted to mechanical stresses, exhibit the thickest one<sup>4</sup>. The main function of the epidermis is to prevent skin from dehydration, infections by pathogens and shear stress damages<sup>2</sup>.



**Figure 2. Structuration and fate of keratinocytes in the epidermis.**

Within 40 days, keratinocytes undergo a differentiation process to maintain the protective properties of the skin. During the whole process, keratinocytes remain tightly bound together thanks to desmosomes and tight junctions. The diffusion of gases and nutrients from dermal blood vessels support the active proliferation of basal keratinocytes as well as the intense metabolic activity of the whole layer. The direction of the arrow indicates the movement of cells during differentiation. Adapted from<sup>5</sup>.

The “barrier” function of the epidermis is performed by a squamous, stratified and keratinized structure involving dense layers of keratinocytes (the main cells of the epidermis) that are tightly bound together<sup>3,5</sup>, controlling exchanges of water, large molecules or bacteria between the environment and the underlying tissues. Keratinocytes represent about 80 % of epidermis total cell content and are responsible for the synthesis of the main components contributing to its protective properties: keratins, lipids, antimicrobial protein factors. The others cell types hosted in the epithelium<sup>3</sup> are melanocytes,

Merkel cells and Langerhans cells (**Chapter I, I-6**). The epidermis is in perpetual renewal through a permanent proliferation of the basal layer followed by a “bottom-to-top” differentiation of keratinocytes, during which cells move towards the skin surface and change their morphology and biochemical properties<sup>6</sup>. Based on these progressive evolutions, four layers of keratinocytes are commonly described in the epidermis (**Figure 2**), from the depth to the surface:

- The *stratum basale* is the deepest layer, composed of a single layer of columnar proliferating keratinocytes. These cells are strongly attached together by desmosomes and adherent junctions, and to the underlying basement membrane by hemidesmosomes. Their asymmetric divisions permanently produce keratinocytes that detach from the basement membrane to enter the upper layer. The detachment is associated with changes in gene expression that initiate cell differentiation while the cells lose their ability to proliferate<sup>7</sup>.

The *stratum basale* also hosts melanocytes, which are responsible for skin pigmentation through melanin synthesis. After production in melanosome specialized organelles, melanin is transferred to keratinocytes by an exo/endocytosis process. This layer last ensures mechanosensing: membrane deformations of basal keratinocytes are transmitted to Merkel cells thanks to desmosomes, leading to the emission of chemical signals able to stimulate dermal afferent neurons.

- The *stratum spinosum* is composed of five to ten layers of keratinocytes which can be very heterogeneous in terms of shape and composition, depending on their differentiation level. While moving upward, cells flatten, accumulate keratin filaments and cytosolic lamellar granules containing lipids (sterols, phospholipids), glycoproteins and hydrolases. Cell-to-cell adherence is promoted by desmosomes while connections by gap junctions facilitate the intercellular communication and the synchronization of differentiation steps.
- The last epidermal layer containing living cells is the *stratum granulosum*. During their progress in this layer, cells accumulate keratohyaline granules and filaggrin in their cytoplasm while flattening goes on. Keratohyaline granules contain enzymes that modify the structure of keratin filaments, in association with filaggrin which can bind and aggregate filaments together. Just before cells leave the *stratum granulosum*, the initiation of organelle degradation and cell membrane destabilization marks the beginning of “cornification”, the last step of keratinocyte differentiation leading to cell death. Tight junctions are found in this layer.
- The *stratum corneum* is the outermost layer of the epidermis. It hosts the main steps of cornification, a process by which keratinocytes die and are incorporated in a dense supracellular structure protecting underlying tissues<sup>5</sup>. Cell organelles are fully degraded; membrane and intracellular contents are released in the extracellular compartment and integrated in the surrounding environment. The *stratum corneum* is thus composed of dead cells called “corneocytes” embedded in a matrix made of free sterols, cholesterol, phospholipids, ceramids and various proteases, lipases and nucleases derived from lamellar granules<sup>8</sup>. Lipids protect against dehydration<sup>6</sup>, and exert anti-microbial properties as a complement to enzymes. Even if cells are dead, their “skeletons” remain

strongly bound together thanks to numerous desmosomes called “corneodesmosomes”. These junctions and the various components of the matrix are progressively degraded by proteases and other enzymes while corneocytes progress outward, leading to the progressive desquamation of cells.

The transit of keratinocytes in the epidermis lasts for about 40 days from production to desquamation. Finally, the strong intracellular keratin networks, the tight cell-to-cell cohesion in all layers<sup>2</sup> and the dense matrix of the *stratum corneum* all contribute to the exceptional resistance of the epidermis to shear stress.

## 2. Dermo-epidermal junction

In skin, the connection between the epithelium and the underlying tissue is performed by a thick structure including keratinocyte transmembrane proteins and extracellular components produced by both epidermal and dermal cells<sup>9,10</sup>. This interface called “dermo-epidermal junction” (DEJ) is characterized by a very tight assembly of its components, which guarantees the structural and functional association between the epidermis and dermis (**Figure 3**).

The DEJ can be split into 3 compartments<sup>11</sup>:

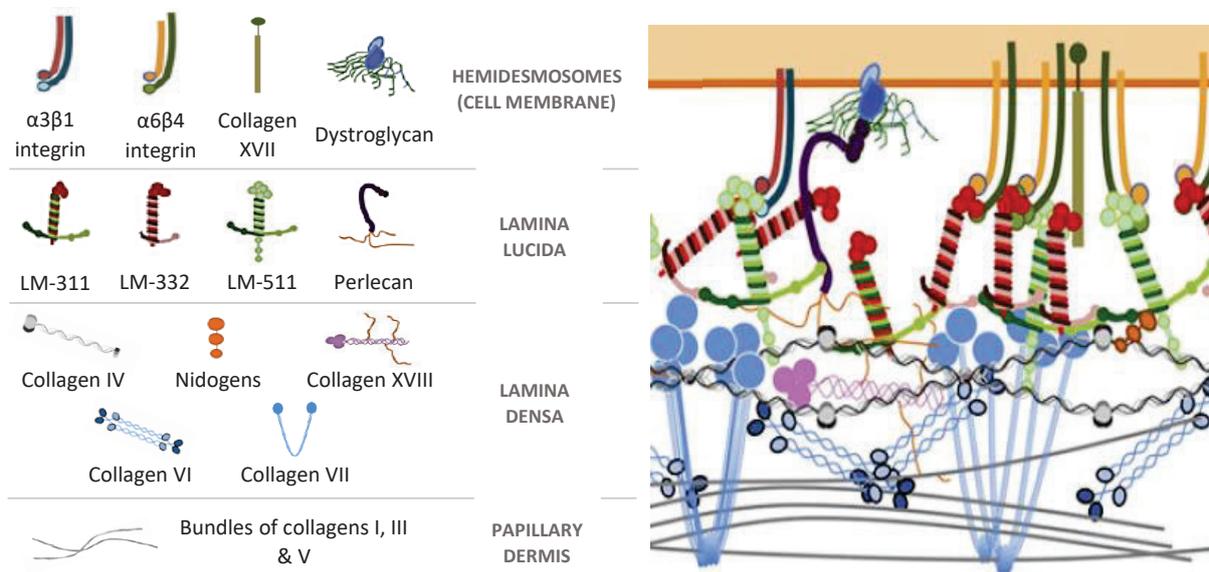
- The basal membrane of keratinocytes from the *stratum basale* directly facing the dermis is the epidermal and most external part of the DEJ. Membrane complexes called “hemidesmosomes” link intracellular keratin filaments with transmembrane integrins and collagen XVII<sup>12</sup>.
- The basement membrane interfacing epidermis and dermis is a complex network structured in two sublayers: superficial *lamina lucida* and deep *lamina densa*.

In the *lamina lucida*, the  $\alpha 6\beta 4$  integrin and the collagen XVII from the membrane of keratinocytes interact with matrix laminin-332 (LM-332), with LM-332<sup>13,14</sup>. Together, LM-332 and collagen XVII form “anchoring filaments” and in these LM-332 and collagen XVII strongly interact with collagens IV and VII of the underlying *lamina densa*<sup>11</sup>, respectively. Collagen IV is distributed as a dense network parallel to skin surface and structures the *lamina densa*, while collagen VII molecules constitute dense “anchoring fibrils” which project deeply to reach the dermis<sup>15</sup>. The proteoglycan perlecan and glycoproteins nidogen-1 and -2 enhance DEJ cohesion and stability<sup>10,16</sup>.

- In the papillary dermis (**Chapter I, I-3-i**), anchoring fibrils from the *lamina densa* exhibit a “V” shape and strongly grasp dermal collagen fibers<sup>17</sup>. Collagen VII is, as a consequence, the key element for basement membrane/dermis connection, as will be further discussed below. In some nomenclatures, this area of the papillary dermis where anchoring fibrils and dermal ECM interact together is also called “*sublamina densa*”.

In addition to mechanical cohesion, the DEJ performs a gate-keeping activity by finely controlling the exchanges of molecules and cells from a layer to another<sup>12</sup>. It especially allows the diffusion of gases

and nutrients from dermal blood vessels to support the growth of keratinocytes, as well as immune cells trafficking. Finally, the DEJ plays a key role in wound healing: the neosynthesis of the basement membrane contribute to drive skin wound healing<sup>15,18,19</sup>, and some proteoglycans of the network act as reservoirs of growth factors in healthy skin and during wound healing<sup>20</sup>.



**Figure 3. Simplified structure of the DEJ.**

Keratinocytes and dermal ECM are tightly associated through complex protein interactions involving *lamina lucida* and *lamina densa* components. The integrins of keratinocytes form bridges connecting cell cytoskeleton with the anchoring filaments of the *lamina lucida*. The latter bind the *lamina densa*, especially collagen VII anchoring fibrils that project deeply and grasp the fibrillar collagen network of the dermis. LM: laminin. Adapted from<sup>21</sup>.

### 3. Dermis

The dermis is a matrix-rich, fibrous connective tissue<sup>22</sup> covering the hypodermis and interfaced with the overlying epidermis by the DEJ. It hosts a dense vascular network as well as several lymphatic vessels, and is innervated. Its thickness varies from 0.1 to 4 mm depending on body site<sup>8</sup>. The dermis performs numerous key functions in skin: it provides pliability and elasticity, contributes to water retention, supports epidermis nutrition, allows the perception of various stimuli and supports skin wound healing, among others. It also contains skin appendages (hair follicles, sebaceous glands, sweat glands) that form deep epidermis-covered invaginations within the dermal layer. Dermis functions mainly depend on the extracellular components, explaining why the abundance in resident cells is relatively limited. The latter are however fundamental for skin homeostasis because they allow the production and renewal of the ECM, and drive skin repair after injury<sup>23</sup>.

#### i. Organization of the dermal extracellular matrix

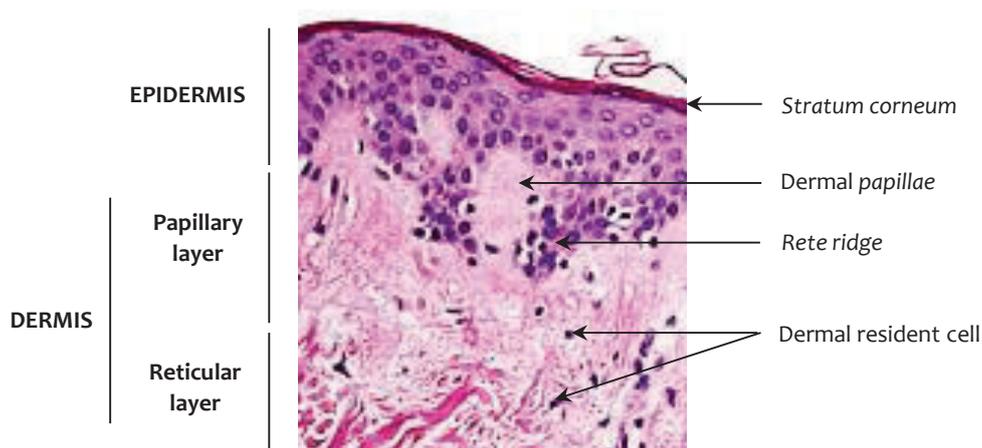
The ECM is defined as the acellular compartment surrounding cells. In the dermis, it mainly consists of a 3-dimensional scaffold made of collagens and elastic fibers embedded in a hydrogel of

proteoglycans, free glycosaminoglycans (GAGs) and glycoproteins<sup>10,15,24</sup>. This complex network does not only provide the architectural basis necessary for cell development but also supports several events of tissue biology such as mechanical stimuli-sensing, cell migration or the storage and release of active molecules<sup>15</sup>. As the hypodermis and epidermis contain smaller amounts of ECM, the term “skin ECM” often found in the literature commonly refers to “dermal ECM”, the basement membrane being usually considered separately.

Collagens are the most abundant proteins in animal tissues; in human, they account for about 70 % of the dry weight of dermal ECM<sup>6,25</sup>. Among the 28 types reported in human (**Chapter I, II-1-i**), the fibrillar collagen I is the most represented in the dermis (85 % of dermal collagens). The collagen I network forms a dense, fibrous, porous structure to which collagens III (10 % of dermal collagens) and V (< 3 %) are associated<sup>6</sup>. It serves as a support for the organization of the whole ECM and is responsible for skin tensile strength. In addition, elastic fibers provides elasticity to the tissue<sup>24</sup>. They result from the association of elastin to protein microfibrils (fibrillins, fibulins, EMILIN-1), and form a second protein network able to stretch under tension and snap back when the mechanical stress is lifted<sup>26</sup>. These protein networks are structurally and functionally associated to free GAGs and proteoglycans<sup>27</sup>. Finally, the ECM contains several glycoproteins<sup>24</sup> that contribute to structurally and functionally connect skin cells with the surrounding ECM. Matrilins, tenascins, thrombospondins, fibronectin and vitronectin especially regulate the balance between adhesion and migration as well as intercellular signaling. In addition to structural properties, numerous proteins and proteoglycans of the ECM are also involved in the modulation of cell behavior through the binding and release of various growth factors<sup>22</sup>.

Superficial and deep parts of the dermis exhibit variations in ECM composition and structure which are responsible for the division of this compartment into two sublayers (**Figure 4**):

- The papillary dermis is the outer part of the compartment connected to the epidermis by the DEJ. It contains a loose and deformable collagen network made of thin fibers, randomly distributed and connected to DEJ components. Regarding composition, the papillary sublayer contains more collagen III, decorin and biglycan than the deep sublayer, and hosts most of the dermal fibroblasts. Blood capillaries that progress through the papillary dermis to the DEJ support feeding of the epidermis. In human skin, the papillary dermis emits protrusions called “*papillae*” in the epidermis, separated by epidermal invaginations named “*rete ridge*”. This structure greatly improves the contact surface between epidermis and dermis, to reinforce their cohesion and promote exchanges.
- The reticular dermis is the deepest part of the dermis and hosts a dense and organized collagen network composed of thick bundles of fibers parallel to skin surface. It also contains thicker elastic fibers, and more versican. If the papillary dermis is responsible for the interactions with epidermis, the main function of the reticular compartment is to confer tensile strength to the skin.



**Figure 4. Human skin observed by photonic microscopy.**

The papillary dermis emits *papillae* in the overlying epidermis, and contains an ECM network that is less dense than the reticular dermis one. Back skin, hematoxylin/eosin coloration, hypodermis not shown. Adapted from<sup>28</sup>.

## ii. Mesenchymal resident cells

In addition to endothelial and nerve cells, fibroblasts are the only type of non-immune dermal resident cells. These cells are scattered in the dermis with few direct cell-cell contacts and perform various functions.

Fibroblasts guarantee the maintenance of the dermal ECM, and therefore the ECM-based homeostasis of this skin layer. Most of the components of the ECM, including collagens and other fibrous proteins, proteoglycans, GAGs and glycoproteins, are synthesized by fibroblasts. These cells also secrete enzymes, especially proteases, which orchestrate the maturation, degradation and cross-linking of these components to allow ECM remodeling and renewal. Furthermore, fibroblasts cooperate with keratinocytes to produce and organize the DEJ. Collagens IV and VII, two major components of the basement membrane, are produced by both cell types.

The homeostasis-maintaining function of fibroblasts also relies on the secretion of a variety of growth factors such as fibroblast-, epidermal- and platelet-derived- growth factors (FGF, EGF and PDGF respectively) to modulate cell growth and the organization of the vascular network and to communicate with resident and infiltrated immune cells<sup>29,30</sup>. Finally, fibroblasts are involved in skin repair through their differentiation into myofibroblasts, a contractile subtype of fibroblasts that exert strong traction forces on the ECM to close wounds.

## 4. Hypodermis

The hypodermis, or *panniculus adiposus*, is the deepest layer of the skin. It is innervated, vascularized and directly in contact with underlying tissues. Its structure consists of projections of fibrous ECM *septa*, rich in collagens, that divide the layer in adipocyte lobules. Bundles of collagen fibers originating from the overlying dermis also penetrate the ECM of the hypodermis to create

cohesion between the two layers. *Septa* contains nerves, large blood and lymph vessels, that progress toward the surface to irrigate the dermis. Lobules host adipocytes which are large mesenchymal cells specialized in lipid storage and release: their cytosol contains few organelles and a large unique (in adult) or multiple (in child) lipid droplet(s). In non-obese adults, the hypodermis concentrates 80 % of the body fat. The hypodermis is an energy-storage tissue and it protects the organism from environmental temperature variations<sup>8</sup>. It also performs endocrine functions: adipocytes produce leptin, a hormone involved in the regulation of body weight, appetite, energy metabolism, and perform a part of the aromatase-based conversion of androgens in estrogens.

### 5. Vasculature and innervation

The irrigation and nervous functions of the skin are mainly ensured by the dermis and the hypodermis.

Blood and lymph vessels originate from the hypodermis and progress to the papillary dermis. They form a dense network in the extracellular matrix and both vessel types contribute to tissue homeostasis<sup>8</sup>. Blood supplies hypodermal, dermal and epidermal cells with nutrients, O<sub>2</sub>, growth factors and endocrine hormones, contributes to metabolic waste elimination and supports the immune response to injury<sup>31</sup>. The lymphatic system acts as a complementary system to the vascular one: it contributes to the transport of lipids, the exchanges of fluids and the elimination of debris and adverse substances that accumulate in the extracellular environment. The lymphatic system is also a key player of the immune response by supporting immune cell proliferation and migration<sup>32</sup>.

As the direct interface with the environment, the skin is a privileged area to sense various external stimuli<sup>3,33</sup>. Nervous activities principally depend on the highly innervated dermis. If epidermal Merkel cells also contribute to sense pressure stimuli, most of the latter are detected by specialized dermal structures called Meissner (papillary dermis), Vater-Pacini and Ruffini (superficial and profound reticular dermis, respectively) corpuscles. Free unmyelinated nerve fibers, especially located in the papillary dermis, and around hair follicles allow the sense of temperature, pain and itch. The hypodermis and dermis also contain efferent nerve fibers modulating vasoconstriction and sweat gland secretions.

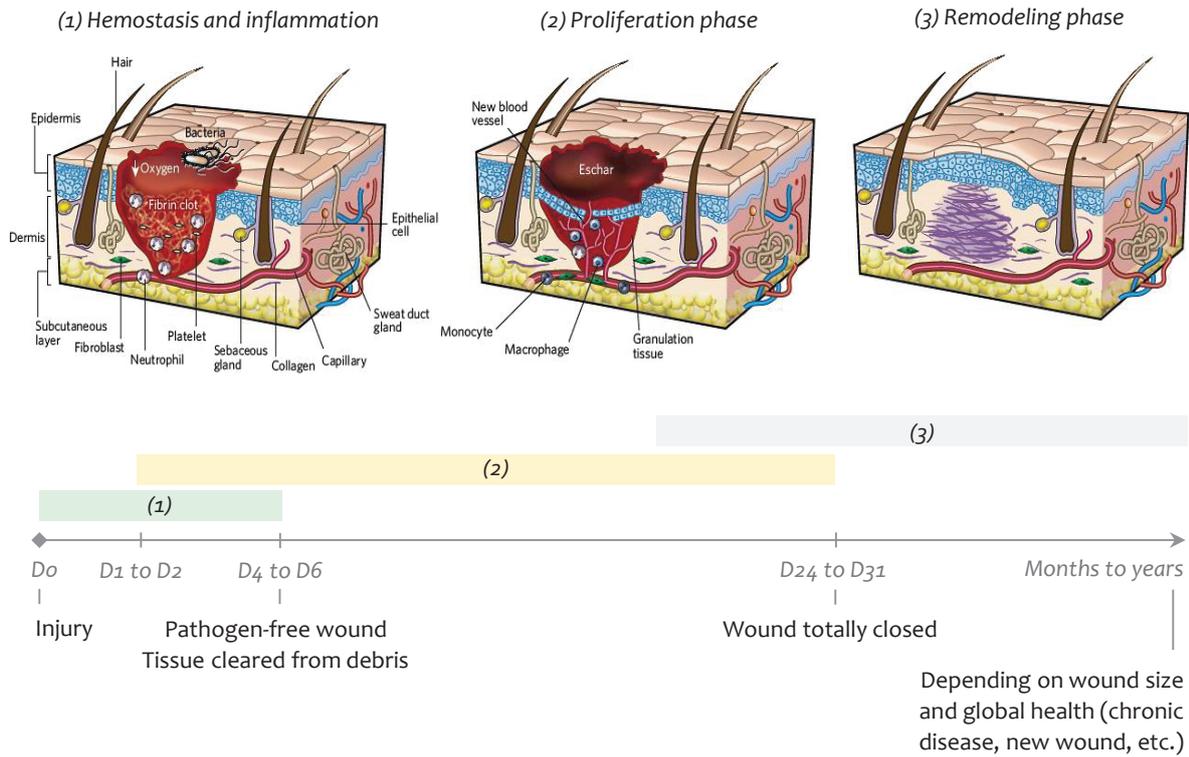
### 6. Skin immunity and healing

As directly exposed to the external environment, skin is a preferential entrance way for pathogens. Numerous actors and mechanisms contribute to the immunity of the organ to prevent infections. Several reviews have proposed extensive descriptions of skin immunity<sup>29,30</sup>.

Briefly, in the epidermis, molecules of the *stratum corneum* such as lipids, small peptides (called antimicrobial peptides) and enzymes can act as an anti-microbial barrier. To reinforce this first but

limited protection, innate immune cells including resident Langerhans cells in the epidermis (5 % of total cells) and dendritic cells in the dermis patrol to detect pathogens and activate adaptive immune response. Resident macrophages are also distributed in the dermis to perform the phagocytosis of pathogens, while resident mast cells drive the allergic response. In the skin, the immune response mainly takes place in the dermis due to its high blood and lymph perfusion. All the cells involved in this response, including monocytes-derived macrophages and neutrophils (innate immunity), CD4+ and CD8+ T-cells and B-lymphocytes (adaptive immunity), reach the skin to fight the infection.

Immune and resident cells are also involved in the wound healing process occurring after both biotic and abiotic (burn, cut, etc.) injuries. Wound healing depends on the inflammatory response that is triggered after injury and aims at restoring the structure and thus the functions of the skin after a trauma<sup>34,35</sup>. It is divided into three phases that come one after another, in parallel with the fight against pathogens, with some overlap: a hemostasis and inflammation step, a proliferation step and a final remodeling phase (**Figure 5**)<sup>36</sup>. All along this process, injured and dead cells are eliminated and a progressive reconstruction of the three layers of the skin occurs through the promotion of angiogenesis, the proliferation and migration of resident cells and the deposition of ECM. The tissue produced during wound healing is named granulation tissue and it differs from the native one. A higher proportion in cells, especially contractile myofibroblasts, is particularly observed, as well as a composition and an organization of the matrix different from the physiological one<sup>37</sup>. The collagens especially exhibit a lower solubility and a higher sensibility to protease-driven maturation in granulation tissue, while the proportion of collagen III is increased<sup>38,39</sup>. This temporary structure is finally remodeled from weeks to months after injury (years for large wounds) to restore a physiologic extracellular environment fulfilling all the functions of the skin<sup>40</sup>.



**Figure 5. Summary of the healing process in skin.**

The three healing phases overlap in time to progressively clean and close the wound. The remodeling phase lasts for weeks to months, possibly years, in physiological healing. Adapted from<sup>36</sup>.

The skin healing process is efficient for small and acute wounds. For some injuries such as large burns, only a limited recovery of the initial structure, and thus functions, of the skin is possible, probably because of the definitive loss of structures involved in tissue repair such as skin appendages. Furthermore, dysregulations in the wound healing process can occur spontaneously and lead to pathologies such as chronic wounds<sup>34</sup>, keloids and hypertrophic scars<sup>41,42</sup>. Fibrosis is a form of pathological wound healing which is also observed in skin. It is characterized by an excessive and aberrant deposition of ECM in the dermis, probably due to a dysregulation in the transition from proliferative to remodeling phases<sup>43,44</sup>. The mechanisms leading to fibrosis are still only partially understood at present day (**Chapter I, III-3-ii**); in addition to dysregulations in the wound healing process itself, environmental factors such as chronic mechanical stimulations of wounds or hypoxia have been suggested to contribute to fibrosis<sup>45</sup>. Whatever its origin and type, fibrosis is an inflammation-driven process. It leads to modifications in the structure of the skin that consequently disturb the functions and the mechanics of the skin and progressively become debilitating for patients.

As a conclusion, the skin is a multi-layered organ responsible for the protection and the adaptation of the body to its environment. The cohesion of the cells in the epidermis and the matrix networks in the dermis are mainly responsible for the properties of the skin and the upkeep of its homeostasis. Cells and

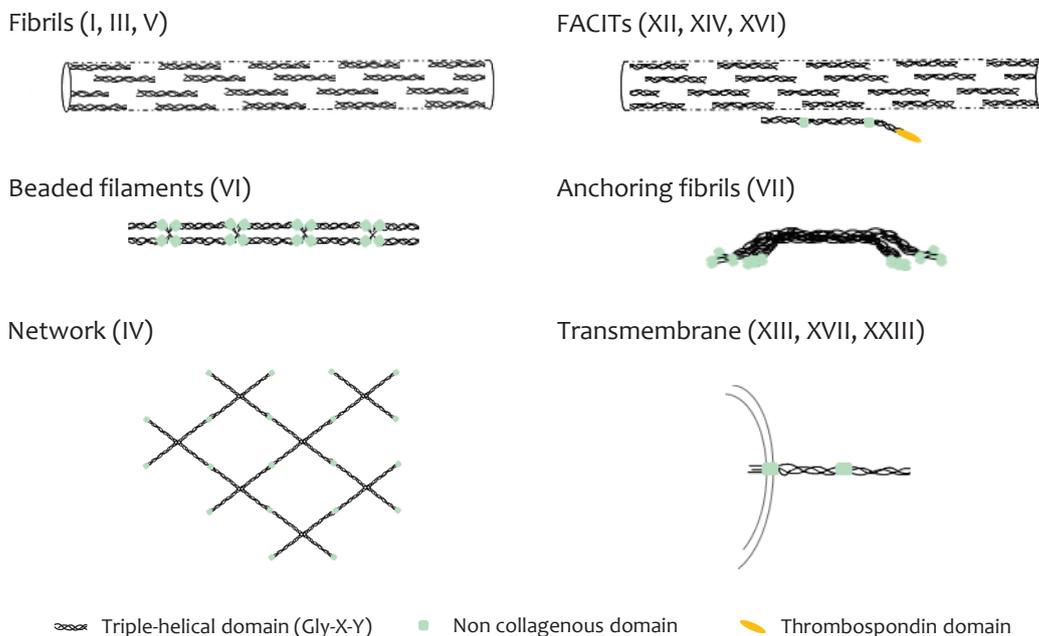
ECM also contribute, with the immune system, to the wound healing process that restores the structure and functions of the organ after injuries.

## II. THE EXTRACELLULAR MATRIX, A COMPLEX AND DYNAMIC NETWORK

### 1. Collagens: the basis of skin ECM

#### i. Collagens: structure and expression in skin

Collagens are the most abundant proteins found in mammalian tissues (30 % of total proteins mass) and constitute the basis of skin ECM, especially in dermis. 28 collagens have been identified and described in human, the composition in collagens varying from one tissue to another<sup>6,25</sup>. A collagen is defined as a protein composed of three identical (homotrimer) or different (heterotrimer) polypeptides named “ $\alpha$ -chains” assembled at least one time into a triple-helical domain<sup>46</sup>. The proportion of triple-helical domains drastically varies from a collagen to another (> 90 % of the mature sequence of collagen I, < 10 % in collagen XII), but the repetition of the common amino acids sequence Gly-X-Y motif is responsible for the folding into helix in all collagens<sup>47</sup>. X and Y amino-acids frequently correspond to proline and 4-hydroxyproline respectively, but may be occupied by any other amino-acids except tryptophan. 3-hydroxyproline is also present at these sites in fibrillar collagens I, II, III and V. Hydroxyprolines act as stabilizers of the triple helix through the creation of hydrogen bonds between the  $\alpha$ -chains of the trimer<sup>46,47</sup>. The interruptions of the Gly-X-Y repeats are responsible for the flexibility of the trimer.



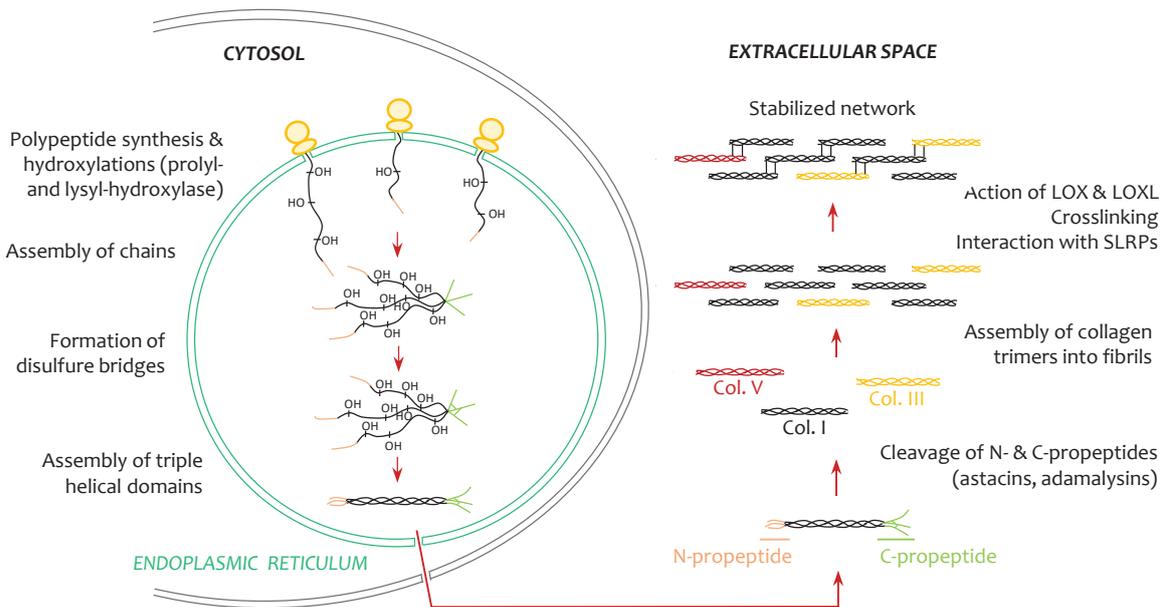
**Figure 6. Schematic representation of supramolecular assemblies formed by collagens in the skin.** FACIT: fibril-associated collagen with interrupted triple helices. Adapted from<sup>25,48</sup>.

All collagens are deposited in the extracellular matrix, except for collagens XIII, XVII, XXIII and XXV that are also transmembrane proteins. For the extracellular ones, supramolecular assemblies of

trimers are very common and a variety of collagen architectures are known (**Figure 6**). In the skin, especially in dermis, the most abundant collagens (I, III, V) assemble into fibrils<sup>6</sup>, which associate together into bundles responsible for skin tensile strength. The dermis also contains FACITs (fibril-associated collagen with interrupted triple helices; XII, XIV, XVI), associated to fibrillar collagens. In the basement membrane of the DEJ, collagen IV forms a mesh that supports the epidermis layer while collagens VI and VII respectively assemble into “beaded” filaments and anchoring fibrils<sup>17</sup>. Epidermal cells exhibit transmembrane collagens (XIII, XVII, XXIII) involved in cell-to-cell and cell to basement membrane adhesion<sup>25</sup>.

**ii. Biosynthesis of fibrillar collagens**

Collagens I, III and V form a network of dense fibrils and bundles in the dermis. Their assembly results from sequential steps of production and maturation involving intracellular and extracellular mechanisms<sup>49</sup>. Fibrillar collagens are synthesized by resident fibroblasts and secreted in the form of precursors; in the extracellular matrix, various proteolytic enzymes can process collagen trimers, thereby triggering fibrils assembly (**Figure 7**)<sup>50,51</sup>.



**Figure 7. Schematic representation of the production and processing of fibrillar collagens.** The steps from translation by ER-bound ribosomes to final integration into the mature collagen network through cross-links are represented. In the mature fibrillar network, black, red and yellow molecules represent collagen I, III and V respectively.

The  $\alpha$ -chains of fibrillar collagens are produced by the ribosomes associated to the endoplasmic reticulum (ER). They are composed of a central “collagenous” part containing repeated Gly-X-Y sequences and flanked by N- and C-propeptides. In the endoplasmic reticulum, they also acquire key post-translational modifications. First, prolines of the Gly-X-Y motif are grafted with 4- (major) or 3-

hydroxylation (minor) through the action of prolyl-3- or prolyl-4-hydroxylases<sup>48</sup>. Hydroxylation also occurs on lysine residues thanks to PLOD (procollagen-lysine-1, 2-oxoglutarate 5-dioxygenase) -1, -2 and -3 exhibiting lysyl hydroxylase activities. Some hydroxylysines are finally glycosylated by the glycosyltransferases GLT25D1 or GLT25D2. The three  $\alpha$ -chains are then covalently linked together through the formation of inter-chain disulphide bridges in the C-propeptides<sup>50</sup>. Finally, the collagenous parts of the chains associate together to create a triple helix. Fibrillar collagens are secreted in this trimerized form called “procollagen” containing a central triple helical domain and N- and C-propeptides responsible for the solubility of the precursor<sup>51</sup>.

The maturation process of procollagens I, III and V occurs in the extracellular matrix under enzymatic actions<sup>46</sup>. First, N- and C-propeptides are cleaved by metalloproteases from the metzincin family (**Chapter I, II-3**). The N-propeptide of fibrillar collagens are mainly processed by ADAMTS-2, -3 and -14 (a disintegrin and metalloproteinase with thrombospondine motifs), and by BMP-1 for collagens  $\alpha 1(V)$  and  $\alpha 3(V)$ . BMP-1, mTLD, mTLL-1 and -2 (for bone morphogenetic protein-1, mammalian tolloid protein, mammalian tolloid-like 1 and 2, respectively), meprins  $\alpha$  and  $\beta$  and furin can potentially remove the C-propeptide. The release of the propeptides from the collagen trimer leads to the formation of insoluble collagen molecules<sup>51</sup> that spontaneously assemble to form collagen fibrils. LOX (lysyl-oxidase) and LOXL (lysyl-oxidase like) catalyze the deamination of lysines and hydroxylysines, a mandatory step for the formation of cross-links that occurs spontaneously and stabilize the final assembly of collagens<sup>25,52</sup>. In addition, small leucine-rich proteoglycans (SLRPs), such as decorin and biglycan, control the spacing and organization of collagen fibrils<sup>27,53</sup>. Finally, fibrils assemble to create bigger collagen structures such as fibers and bundles that further increase skin tensile strength.

In normal adult dermis, the composition of fibrillar collagen trimers is constant (**Appendix 1**). Collagen III always consists in a  $\alpha 1(III)$ -homotrimer annotated  $[\alpha 1(III)]_3$ <sup>25,54</sup>. Heterotrimers are observed for Collagen I ( $[\alpha 1(I)_2, \alpha 2(I)]$ ) and collagen V ( $[\alpha 1(V)_2, \alpha 2(V)]$ ,  $[\alpha 1(V), \alpha 2(V), \alpha 3(V)]$ )<sup>55</sup>;  $\alpha 1(V)_3$  is also present in normal adult skin<sup>56</sup>. Homotrimers of  $\alpha 1(I)$  can also be found in diseases such as fibrosis<sup>57</sup> or cancer<sup>58</sup>.

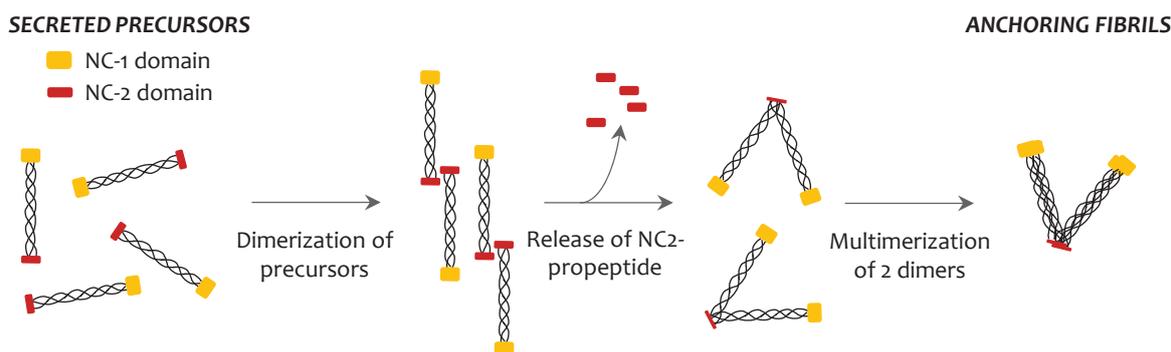
### iii. Biosynthesis of collagen VII

Collagen VII is a key component in DEJ. Basal keratinocytes and dermis fibroblasts both contribute to its synthesis<sup>12</sup>. Like fibrillar collagens, collagen VII is secreted as a hydroxylated and glycosylated precursor which is then matured in the ECM by proteolysis.

A single gene (*COL7A1*) encodes collagen VII homotrimer. The  $\alpha 1(VII)$  chain is produced in the form of a central collagenous part flanked by large non-collagenous NC-1 (N-term) and NC-2 (C-term),

containing a C-propeptide) domains. The same enzymes as those described for fibrillar collagen biosynthesis perform 4-hydroxylation of prolines and 5-hydroxylation of lysines in the collagenous domain. The polypeptide is N-glycosylated at least once in the NC-1 domain; based on the amino-acid sequence, other putative O- and N-glycosylation sites have been suggested in both the NC-1 and collagenous domains but not validated experimentally<sup>59</sup>. Finally, several inter-chain disulphide bonds in the NC-2 domain favour, in combination with the assembly of collagenous domains into a triple helix, the trimerization of the  $\alpha$ -chains. The collagen VII precursor is 2944 amino-acid long and is characterized by numerous interruptions of the triple helix structure (19 in total)<sup>25</sup>.

After secretion, collagen VII aligns as anti-parallel dimers with overlapping C-terminal ends, which are proteolytically matured by proteases including astacin-like proteases<sup>60</sup> that remove the major part of the NC-2 domain. Dimers of mature collagen VII (2822 amino acids each in final form) laterally aggregate to form the anchoring fibrils found in the DEJ (**Figure 8**). Interactions of NC-1 domains with LM-332 and collagen IV insure a tight and strong integration in the basement membrane while the basis of the “V” is projected into the dermis to grasp the network of fibrillar collagens<sup>16</sup>.



**Figure 8. Extracellular processing and supramolecular assembly of collagen VII in the basement membrane of the skin.**

Secreted precursors dimerize and lose their NC-2 propeptide to assemble into anchoring fibrils.

## 2. Glycosaminoglycans and proteoglycans in skin matrix

Glycosaminoglycans consist in long carbohydrates chains which can be found free in the ECM (hyaluronic acid) or bound to a core protein to form matrix-bound, transmembrane or intracellular proteoglycans<sup>15</sup>. The most frequent GAGs found in the matrix of the dermis are keratan sulphate, chondroitin sulphate and heparan sulphate (**Table 1**)<sup>27</sup>. Depending on their nature, GAGs exhibit a broad variety of properties and perform various functions in skin<sup>53</sup>. Versican is the most abundant proteoglycan in the matrix of the dermis and acts as a stabilizer of collagen and elastic fibers, as well as a modulator of cell migration. Small leucine-rich proteoglycans (SLRPs) such as biglycan, decorin, lumican and mimecan are also found in the matrix of the dermis. Biglycan and decorin are especially important for matrix homeostasis: they regulate collagen fibrillogenesis and fiber spacing, by direct binding to

collagen I<sup>22</sup>. They also bind transforming growth factor  $\beta$  (TGF- $\beta$ ) and modulate its bioavailability in the ECM<sup>61,62</sup>. Regarding betaglycan (or type III TGF- $\beta$  receptor), a proteoglycan either found at the cell surface or soluble after ectodomain shedding, it regulates TGF- $\beta$  bioavailability and interaction with signaling receptors. If all these proteoglycans contribute to fill some of the empty space between protein fibers, the main space fillers in dermal ECM are free hyaluronic acid and hyaluronan (polymer of hyaluronic acid residues). These molecules exhibit a high water-retention property that is essential to skin tautness and hydration; furthermore, the interaction of hyaluronan with the transmembrane receptor CD44 regulates skin regeneration<sup>27</sup>.

Localization	Proteoglycan	SLRP	GAG content
DEJ	Perlecan		HS
Dermis	Biglycan	X	CS, DS
	Decorin	X	CS, DS
	Fibromodulin	X	KS
	Lumican	X	KS
	Mimecan	X	KS
	Versican		CS
Transmembrane	Betaglycan		CS, HS
	Glypicans		HS
	Syndecans		CS, HS

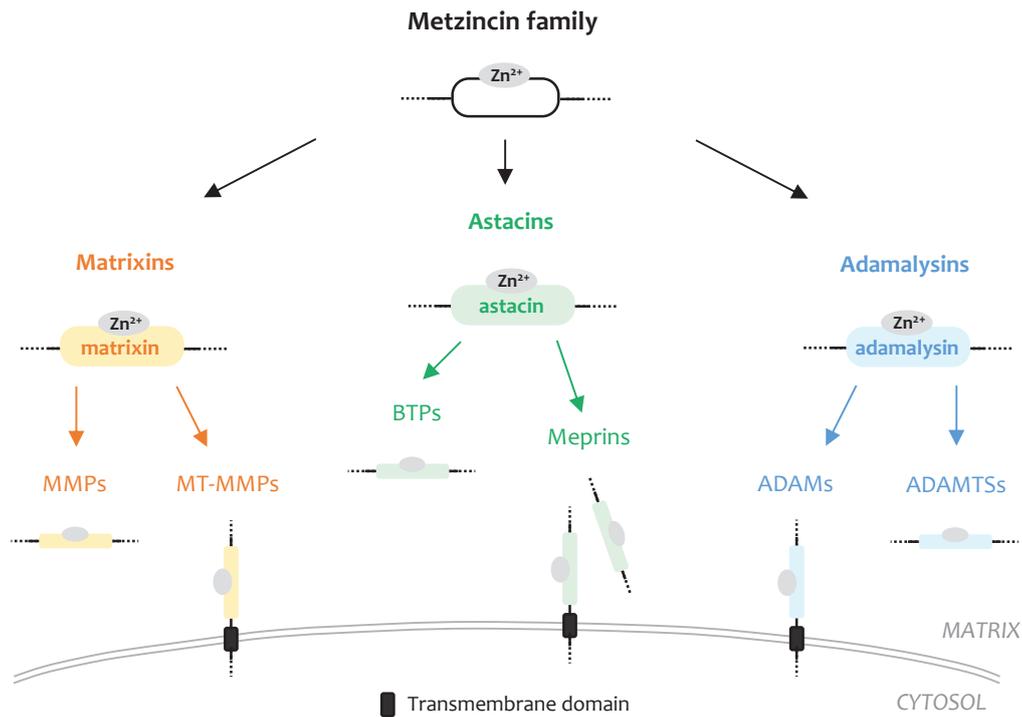
**Table 1. Composition of the principal transmembrane and matrix proteoglycans contained in the skin.**

All these components contribute to matrix structure, functional properties and modulate growth factor signaling. Hyaluronan is the only GAG-polymer found in the free form in the ECM. CS: chondroitin sulfate, DS: dermatan sulfate, HA: hyaluronic acid, HS: heparan sulfate, KS: keratan sulfate.

### 3. Metzincins and proteolysis-driven dynamics of the matrix

Collagens are key components for the structure of skin ECM. The continuous production, maturation and remodeling of matrix proteins, controlled by matrix proteases, are responsible for a complex and dynamic homeostasis of the ECM, especially in the dermis.

Among matrix proteases, metzincins are key actors in ECM homeostasis. The metzincin family is a group of zinc metalloproteinases characterized by the presence of a methionine turn close to the active site and a conserved zinc-binding sequence<sup>63</sup>. Metzincins are involved in numerous activities linked with ECM remodeling and signaling. The metzincin family in human is composed of 3 main groups based on structural features<sup>63</sup>: matrixins, astacins and adamalysins (**Figure 9**). Two other subfamilies, pappalysins and serralysins, also belong to metzincins but they will not be detailed here.



**Figure 9. Main proteases of metzincin family.**

The metzincins are zinc-binding proteases and are distributed into three main subfamilies depending on the nature of the domain binding Zn<sup>2+</sup>. Matrixins comprise soluble MMPs and membrane-type MT-MMPs, astacins include BMP-1/tolloid-like proteases (BTPs) and meprins (sometimes released from the membrane), and adamalysins regroup soluble ADAMTSs and membrane-bound ADAMs. Ovastacin, and serralysins and pappalysins subfamilies are not represented here.

### i. Matrixins

Matrixins are also named “matrix metalloproteinases” (MMPs). To date, 23 MMPs have been described in human and in mouse, including 17 soluble forms and 6 GPI-anchored or transmembrane forms named “membrane-type matrix metalloproteinases” (MT-MMPs): MMP-14 to -17 (or MT1- to MT4-MMP), -24 (or MT5-MMP) and -25 (or MT6-MMP)<sup>64,65</sup>. MMPs are produced in the form of inactive precursors and are most often activated by furin or other active metzincins. In the ECM, soluble MMPs mainly perform catalytic activities in the matrix while the membrane-bound ones mainly modulate cell adhesion through modifications of the structure of several receptors. In normal skin, MMP-1, -2, -3, -7, -8, -9, -13 and -14 – have especially been reported to degrade both protein and proteoglycan targets. They contribute to maintain the homeostasis of fibrillar and non-fibrillar skin collagens by degrading collagen fibers or isolated chains prior to their elimination from the matrix. Others MMPs can also be found in the skin when processes such as inflammation or wound healing occur (e.g. MMP-10, MMP-12).

MMPs activities are modulated by tissue inhibitors of metalloproteinases (TIMPs) and circulating  $\alpha$ 2-macroglobulin<sup>66,67</sup>. Proteoglycans such as lumican and versican also regulate active MMPs as well as their maturation from inactive precursors<sup>68</sup>. Regulation of MMP-based proteolysis is however very complex, as illustrated by the dual role of some regulators depending on their interaction with other

ECM components. For instance, syndecan-2 inhibits pro-MMP2 activation when it is free but activates the same process when it is bound to chondroitine-sulfate proteoglycan 4<sup>69,70</sup>.

## ii. Astacins

Astacins form another subgroup of metzincins performing a wide variety of functions in the ECM. Similarly to MMPs, they contribute to the maturation or degradation of numerous matrix components (structural and enzymatic ones) and to signaling.

Meprins are the first subtype of astacins. These proteases are composed of  $\alpha$  and/or  $\beta$  chains associated in homo- or hetero-multimers of variable size<sup>71</sup>. Meprins can be secreted in the matrix or bound to cell membranes in the form of homo- or hetero-dimers. They are involved in the maturation of fibrillar procollagens I and III through the removal of C- and/or N-propeptides prior to fibril formation<sup>72,73</sup>. Meprins also activate other metzincins such as ADAM10<sup>74</sup>, metzincins regulators such as PCPE-1<sup>72</sup> and they are regulated by fetuin B which inhibits their activity<sup>75</sup>.

BTPs are the second subtype of astacins and include 4 proteases: BMP-1, mTLD, mTLL-1 and -2 (BMP-1 and mTLD are two splicing variants produced from *BMP1* gene)<sup>76</sup>. In skin, BTPs are essential to the organization of the dermis ECM, especially for the collagen network. They mature all fibrillar procollagens through the removal of the C- and/or the N-propeptides<sup>46</sup>, modulate the structure of some SLRPs involved in the regulation of the lateral growth and spacing of collagen fibrils (decorin, osteoglycin/mimecan, biglycan), and release the mature form of elastin and collagen cross-linking enzymes LOX and LOXL<sup>76</sup>. BTPs also contribute to the assembly of the DEJ through the maturation of procollagen VII and LM-332<sup>13</sup>.

BTPs, especially BMP-1, are also involved in matrix-dependent signaling events, in particular those involving TGF- $\beta$ . They modulate the bioavailability of the growth factor through the cleavage of proteins involved in its sequestration in the matrix, such as latent TGF- $\beta$ -binding protein 1 (LTBP1)<sup>77</sup>. They also shed betaglycan, a co-receptor for TGF- $\beta$ <sup>78</sup>. The action of BTPs leads to the release of the active TGF- $\beta$  growth factor and promotes signaling. Recently, the cleavage of thrombospondin-1 (TSP-1) by BMP-1 was also shown to increase TGF- $\beta$  signaling *in vitro*<sup>79</sup>. BTPs participate to other processes such as the modulation of angiogenesis through the release of the LG3 domain from the C-terminal part of perlecan core protein, or the mineralization of soft tissues by action on dentin matrix protein-1<sup>76</sup>.

Several regulators of BTPs have also been identified. Some affect the activity of the proteases on all their substrates, like fibronectin and serum  $\alpha$ 2-macroglobulin that respectively exert activating and inhibiting effects. Other regulators named “procollagen C-proteinases enhancers” (PCPEs) regulate the activity of BMP-1 on some of its substrates. PCPE-1 stimulates the processing of fibrillar procollagens, while PCPE-2 seems to intriguingly affect this processing in a dose-dependent manner, as we recently reported<sup>80</sup>: the cleavage of procollagens by BMP-1 is activated in the presence of low concentrations of

PCPE-2, but tended to decrease at higher concentrations of the regulator. I contributed to this work on PCPE-2 during my PhD, especially for the characterization at tissue level of the effect of knocking down PCPE-2 (article submitted, see Publications section at the end of this manuscript)

### iii. Adamalysins

The adamalysins include a disintegrin and metalloproteinase (ADAM) and a disintegrin and metalloproteinase with thrombospondine motifs (ADAMTS) enzymes. 13 ADAMs among the 22 reported in human exhibit proteolytic activity<sup>81</sup>. They are membrane-bound proteases and mainly involved in the regulation of the interactions of cells with the matrix, and in signaling through the processing of membrane-bound receptors and precursors of growth factors<sup>82</sup>. Conversely, ADAMTSs are secreted proteases and their main targets are ECM components. 19 ADAMTSs have been reported in human, from ADAMTS-1 to -20, ADAMTS-5 and -11 being exactly the same molecule<sup>83,84</sup>. In normal skin, ADAM-10 and 17 are found in the epidermis<sup>85</sup>. In the dermis, ADAMTS-2, -3, -14 are responsible for the maturation of fibrillar procollagens in addition to astacins: they can release the N-propeptide of most fibrillar collagen precursors (I, III,  $\alpha$ 1(V), and the C-propeptide of procollagen III<sup>86</sup>. Proteoglycan (e.g. betaglycan, decorin) and glycoprotein (e.g. fibronectin) substrates have also been suggested by proteomic analyses but the biological significance of these cleavages is not fully understood<sup>86,87</sup>. ADAMTS-17 could also take part in the homeostasis of dermal ECM since the absence of this protease in mouse leads to the thickening of the dermis and to enhanced deposition of fibrillar collagens<sup>88</sup>.

Examples of substrates of matrixins, astacins and adamalysins are given in **Table 2**.

Protease	Collagen substrates	Proteoglycan substrates	Other substrates
<i>Matrixins</i>			
MMP-1	Collagens I, III, VII	Perlecan	
MMP-2	Collagens I, III, IV, V, VII	Perlecan	Elastin Laminins
MMP-3	Collagens III and IV	Decorin Perlecan	Elastin Fibronectin Laminins
MMP-7	Collagens IV, V	Decorin Versican	Elastin Fibronectin Laminins Nidogen-1
MMP-8	Collagens I, III, V, VII		Laminins
MMP-9	Collagens V, VI, VII		Fibronectin
MMP-13	Collagens I, III, V	Biglycan Fibromodulin Perlecan	Fibronectin Laminins
MMP-14 (MT1-MMP)	Collagens I, III, IV	Perlecan	Fibronectin Laminins
<i>Astacins</i>			
BTPs	Procollagens I, III, V ( $\alpha 1$ , $\alpha 2$ ): C-propeptide	Betaglycan Biglycan	LM-332 ( $\alpha 3$ , $\gamma 2$ ) DMP-1
	Procollagen V ( $\alpha 1$ , $\alpha 3$ ): N-propeptide	Decorin Mimecan	DSPP Pro-LOX
	Procollagen VII: NC-2 domain	Perlecan	Pro-LOXL
Meprin $\alpha$ , $\beta$	Procollagens I, III: N- & C-propeptides		BMP-1 MMP-1, -7 PCPE-1
<i>Adamalysins</i>			
ADAMTS-2, -3, -14	Procollagens I and III: N- & C-propeptides Procollagen V ( $\alpha 1$ ): N-propeptide	Betaglycan	Fibronectin

**Table 2. Examples of skin ECM substrates for the main metzincins expressed in normal skin.**

The main substrates involved in the structuration of the ECM are listed. DMP-1: dentin matrix protein-1, DSPP; dentin sialophospho-protein

#### 4. The Matrisome project

These last 20 years, numerous studies aiming at describing the biology of the ECM and its involvement in pathology, especially in cancers, have been conducted. The recovery of the ECM from native tissues combined with MS-based proteomics is now frequently used to study matrix proteins (**Chapter I, IV-4**). The extracellular matrix is however made of a very large variety of molecules, the abundance of which can vary by several orders of magnitude, making the identification of all ECM components in large omics datasets particularly challenging. Before the launch of the matrisome concept, only a few ECM target molecules could usually be characterized in each study.

##### i. Matrisome definition

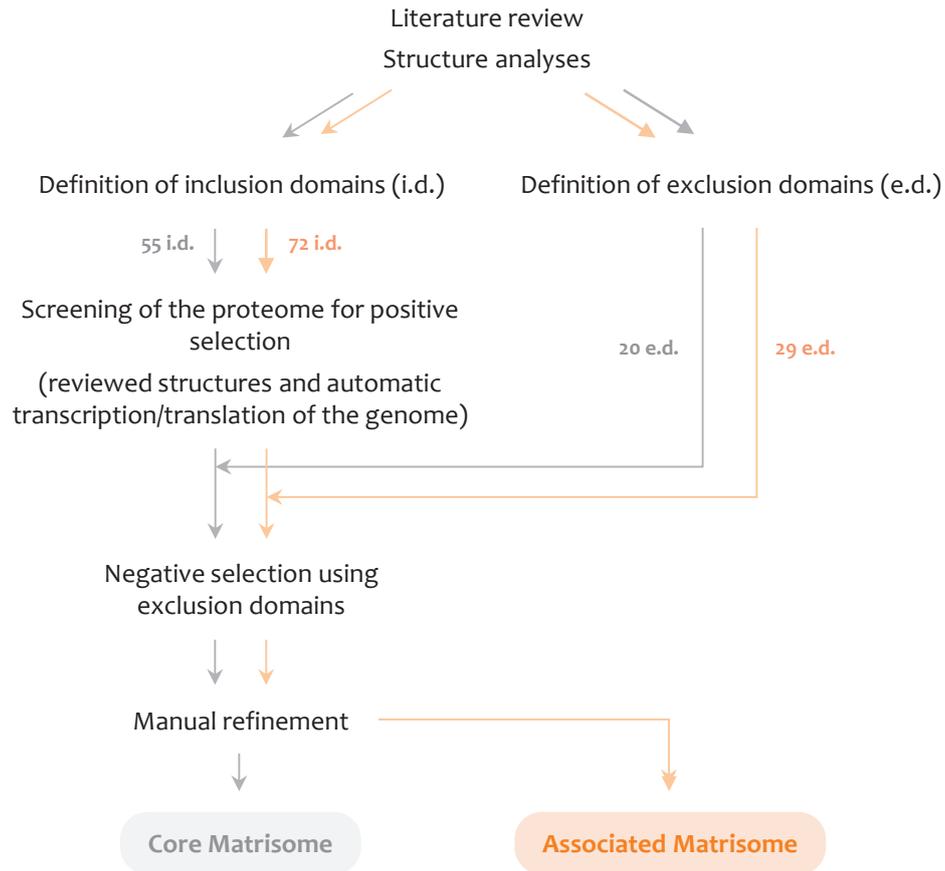
To promote more standardized evaluations of ECM-centered analyses and combine all previously-generated data in a single database, the Matrisome project has been initiated in 2011 by the lab of Richard Hynes (Koch Institute for Integrative Cancer Research, MIT, USA) under the supervision of Alexandra Naba<sup>89</sup>. The aim of the project was the elaboration of a list, made as exhaustive as possible, of all the ECM proteins contained in the proteome to help scientists identify them in their omics datasets. The matrisome was thus defined as “all the protein components constituting the extracellular matrix, and components associated with it” and bioinformatics screening of human and murine proteomes were performed to identify all the proteins corresponding to this definition.

##### ii. *In silico* build-up of the matrisome

To identify all the proteins matching the definition of the matrisome for each species, *in silico* screenings of both genome and proteome databases were performed. Prior to bioinformatic analyses, the matrisome was split into two categories:

- the “core matrisome” that includes proteins mostly fulfilling structural functions (including collagens),
- the “associated matrisome” that gathers proteins interacting with the core matrisome, secreted factors and components sharing structural domains with core proteins or regulating the architecture of the matrix.

For human and then mouse, matrisome lists have been established based on structural features. For core and associated matrisomes, separated lists of domains commonly found (“inclusion” domains) and, conversely, not found (“exclusion” domains) in matrix proteins have been defined. The inclusion/exclusion criteria were based on an extensive literature review, structural annotations of proteins (UniProt database) and structural predictions obtained from the analysis of genes sequences (Interpro database)<sup>89</sup>. The proteins defined as belonging to the matrisome have thus successfully passed positive and negative selection steps using these lists as well as a final manual refinement (**Figure 10**). The same inclusion and exclusion domains were used to screen both human and mouse proteomes.



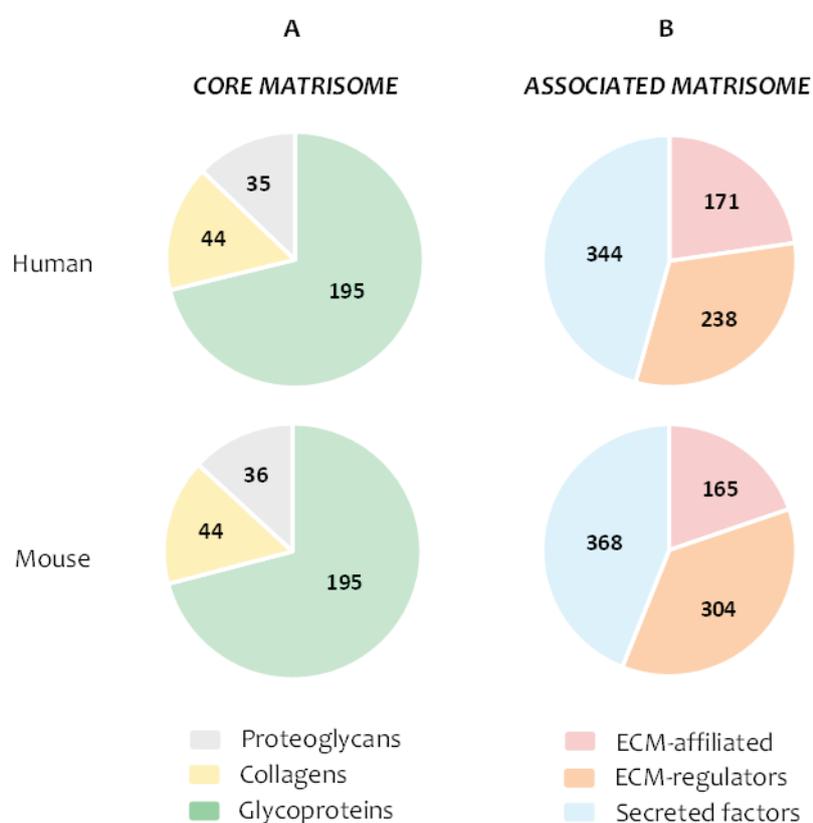
**Figure 10. Methodology for the definition of the matrisome.**

Data from literature and structural analyses were used to define inclusion and exclusion domains for the core and the associated matrisome. These domains were used to screen the real and virtual proteomes in human and mouse, and define the proteins forming the core and the associated categories of the matrisome. Inclusion and exclusion domains are presented in appendices 2 and 3, and more detailed can be found in<sup>89</sup>.

For the core matrisome, 55 inclusion (e.g. collagen triple helical, matrix fibril-associated, laminin domains, fibronectin domains) and 20 exclusion domains (e.g. specific catalytic domains, ion channel domains) have been defined (**Appendix 2**). Final steps of exclusion of proteins containing transmembrane domains and of inclusion of proteins not contained in the bioinformatically-generated panels but known to be part of the ECM based on literature reports were also performed. The final core matrisome list is divided into 3 subgroups depending on the biochemical nature of the proteins: collagens (all the verified and putative collagen chains), glycoproteins and proteoglycans (core proteins). The human core matrisome is composed of 274 proteins, including 44 collagens, 195 glycoproteins and 35 proteoglycans. The mouse one contains 275 proteins, with one additional proteoglycan (nephrocan) and a few differences in the nature of collagen chains (**Figure 11A**).

72 inclusion domains (e.g. catalytic domains, domains belonging to protease inhibitors or cytokine) and 29 exclusion domains (e.g. domains belonging to membrane-bound proteases or integrins) were used to define the associated matrisome (**Appendix 3**). Final refinement steps identical to those

described for the core matrisome were also applied. A division of the final associated matrisome list into 3 subgroups was finally proposed, based on the biological role of the proteins: structural components not involved in the network of core proteins but interacting with it were named “ECM-affiliated”, enzymes and their regulators were grouped as “ECM regulators” and signaling molecules as “Secreted factors”. The associated matrisome is the category gathering the largest number of proteins. In human, it includes 753 components: 171 ECM-affiliated proteins, 238 ECM-regulators and 344 secreted factors. The mouse one contains even more components, 837 in total: 165 ECM-affiliated proteins, 304 ECM-regulators and 368 secreted factors (**Figure 11B**). The high difference in the number of ECM-regulators between mouse and human associated matrisomes results from numerous gene duplications in the mouse genome. For instance, 36 genes encode serine protease inhibitors (serpins) in human, compared to 60 in mouse<sup>90</sup>.



**Figure 11. Schematic view of human and mouse matrisomes.**

**A.** Core matrisome proteins. **B.** Associated matrisome proteins (last update: August 2014<sup>91</sup>).

A total of 1027 and 1111 genes respectively encode the matrisome in human and mouse. They account for about 4 % of the total genome of each species.

### iii. Tools for the analysis of the matrisome

All the databases generated *in silico* are available on the website of the Matrisome project<sup>91</sup> (last update: 2014 for human and mouse lists) and can be used to identify ECM proteins in large proteomic datasets. Since 2017, the *in silico* matrisomes of *Danio rerio* (zebrafish)<sup>92</sup>, *Drosophila melanogaster*<sup>93</sup> and *Caenorhabditis elegans*<sup>94</sup> were also generated and deposited on the platform.

To facilitate the annotation and the exploitation of datasets, the Matrisome platform also proposes tools based on the matrisome databases. First, an on-line annotator can be run automatically to detect and annotate ECM proteins in raw omics datasets. In addition, a new platform has been launched in 2019 to facilitate the retrospective cross-study of the previously-reported proteomic studies of the matrix<sup>95</sup>. Based on a compilation of various proteomics datasets, this tool proposes, for each matrix protein from both human and mouse organisms, a collection of metadata that have been reported in the literature (tissues where it was detected, expression in health and disease, other ECM-related proteins, etc.). For some older works, the lists of peptides detected by MS/MS are also available *via* the on-line platform PRIDE<sup>96</sup>, this information being especially useful for the development of targeted proteomics studies.

The Matrisome approach and databases were used to study several tissues by transcriptomics and MS-based proteomics, especially in diseases such as cancer or fibrosis. For instance, the matrisome of mammary glands<sup>97</sup>, pancreatic islets<sup>98</sup>, lung<sup>99,100</sup>, liver and colon<sup>101</sup>, and various tumors affecting these tissues have been described.

To sum up, the ECM is a complex and dynamic environment made of structural and regulatory proteins. Proteases are key actors in the homeostasis of the skin, through the formation/upkeep of the matrix structure and the modulation of matrix-associated signaling events. To facilitate the study of the very wide variety of ECM proteins, the matrisome concept was introduced 10 years ago and specific tools were progressively developed to facilitate data exploitation.

### III. DYSTROPHIC EPIDERMOLYSIS BULLOSA

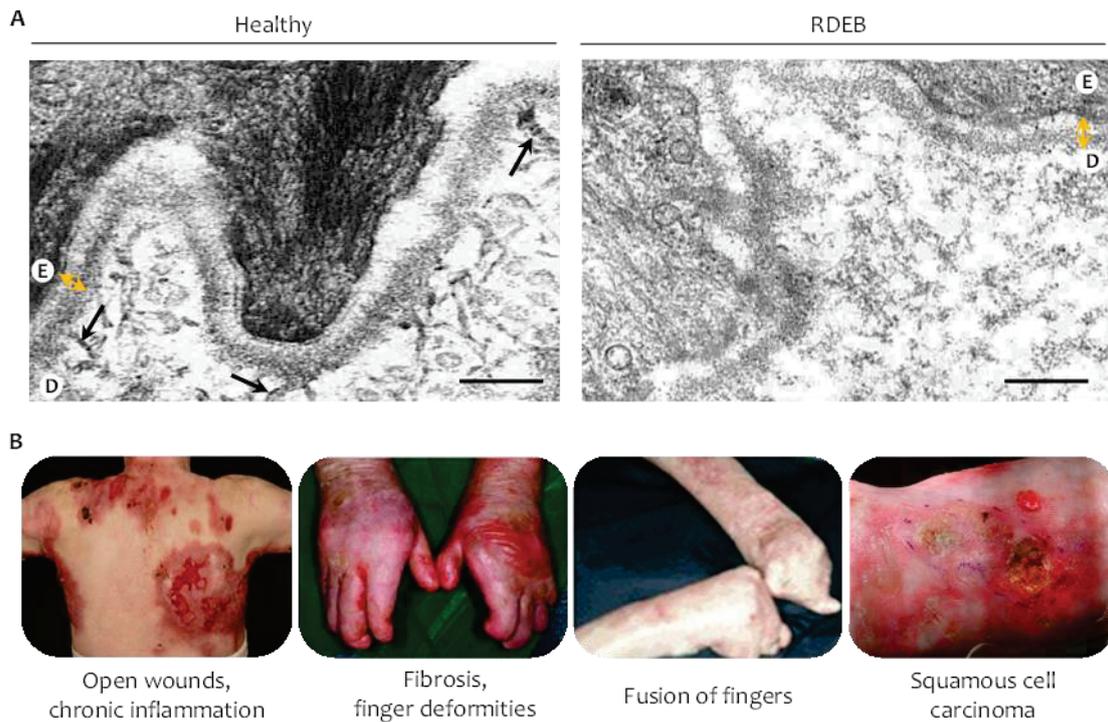
Epidermolysis bullosa (EB) is a group of inherited disorders caused by mutation(s) in an autosomal gene encoding one of the components of basement membranes and DEJ. To date, mutations in at least 18 genes have been identified as responsible for EB. These disorders are characterized by skin fragility and blistering in response to even mild shear stress, and their severity depends on the EB form and gene mutation.

#### 1. Etiology, classification and symptoms

DEB is a rare disease affecting 1 birth out of 49,000 to 420,000, depending on the regions of the world. This pathology is the consequence of mutations in *COL7A1*, the gene encoding collagen VII<sup>17</sup>. To date, about 1000 different dominant (DDEB) or recessive (RDEB) mutations have been identified in DEB patients<sup>102</sup>: they range from missense mutations to premature termination codons (PTC). They are responsible for partial to total loss in collagen VII functionality and cause basement membranes instability and epithelia detachment, especially in DEJ where collagen VII greatly contributes to the cohesion of epidermis with the underlying dermis. The severity of the pathology and its symptoms to some extent depend on the mutation. The most serious cases of DEB are those associated with recessive PTC mutations or recessive combinations of mutations, and result in very low levels of functional collagen VII. In skin for example, the absence of anchoring fibrils can be observed by transmission electron microscopy (TEM, **Figure 12A**).

Chronic skin blistering due to the detachment of the epidermis is an early manifestation of DEB. It occurs in body areas submitted to repeated mechanical stress (mainly hands, feet, elbows and knees). Chronic slow-healing wounds, scarring and permanent inflammation of the underlying tissues finally result from epithelium detachment (**Figure 12B**)<sup>103</sup>.

In the most serious DEB forms<sup>104-106</sup>, especially RDEB, the manifestations of the pathology are diverse and progressive in terms of nature. Blistering occurs on all the surface of the body and on oral mucosae. Soft tissue fibrosis progressively develops and leads to deformations in limb extremities (finger deformities and fusion) and to esophagus narrowing. As feeding becomes harder, patients commonly suffer from nutrient deficiencies such as anemia, and underweight. Alterations of corneal epithelium affecting vision are also frequent in the eye<sup>107</sup>. Finally, the inflammatory and fibrotic environment of the skin favors the development of aggressive squamous cell carcinomas (SCC) responsible for the high morbidity of the disease<sup>104,108</sup>. For people with severe RDEB, the cumulative risk of SCC is higher than 90 % at the age of 55 years, and life expectancy after the first tumor occurrence does not exceed 5 years<sup>109</sup>.



**Figure 12. RDEB micro- and macroscopic phenotypes.**

**A.** TEM observation of DEJ in normal (left) and RDEB (right) human skins. Collagen VII anchoring fibrils are visible in normal skin (black arrows) but absent in RDEB skin. Bar: 200 nm. Orange double arrows: basement membrane with *laminae lucida* and *densa*, E: epidermis, D: dermis. **B.** Progressive evolution of skin symptoms in RDEB. Adapted from<sup>110–113</sup>.

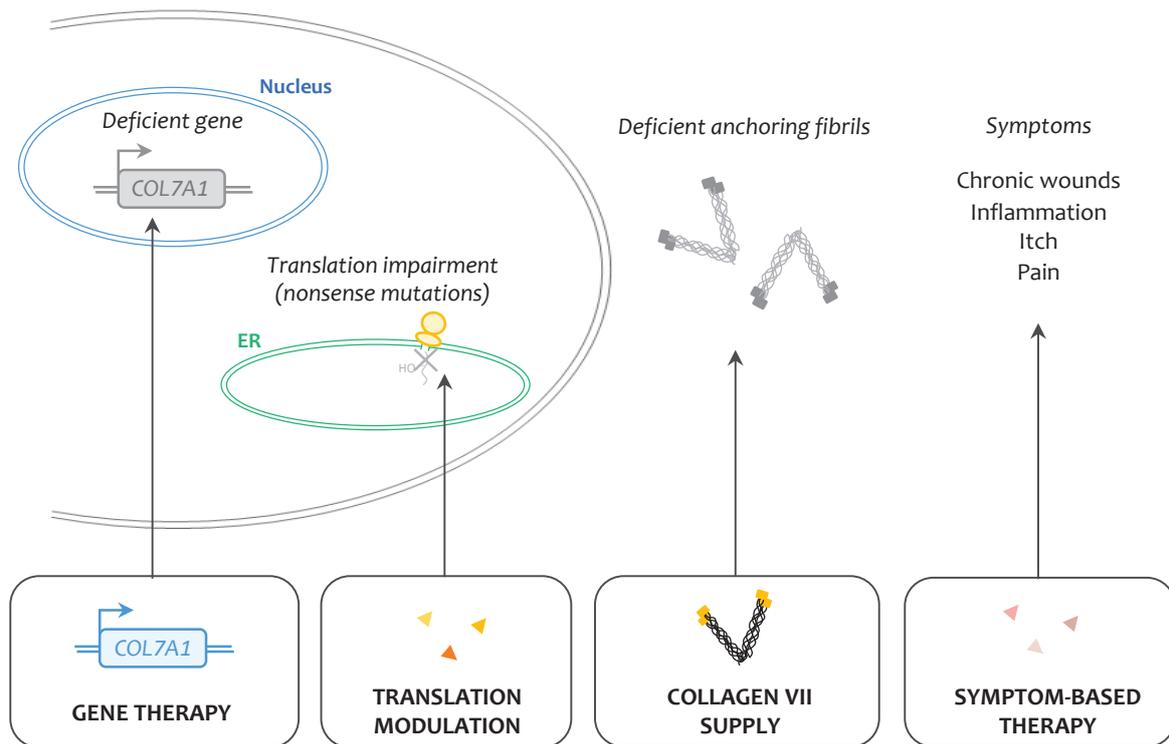
Although a link between the level of functional collagen VII and the severity of the disease exists, some observations suggest the involvement of additional mechanisms to explain the progression of symptoms (epigenetics, compensation of collagen VII deficiency, etc.)<sup>114</sup>. For instance, the case of two homozygous DEB twins exhibiting different disease severities questions the “purely” genetic cause of the disease<sup>115</sup>. In addition, studies on patients presenting the same *COL7A1* mutations from the same or different families also revealed substantial variations in disease severity<sup>116</sup>. Mutations in the promoter of *MMP1* (encoding matrix metalloprotease-1) responsible for an increased expression of the gene were also reported to exacerbate the symptoms in RDEB patients<sup>117</sup>.

To date, no curative treatment is available for DEB: medical care consists in the treatment of symptoms to limit their extension and improve the quality of life of patients<sup>118,119</sup>. Daily management of DEB involves careful wound asepsis and dressing to prevent infections and try to favor healing. Skin substitutes can be grafted to protect and help the closure of large wounds<sup>120,121</sup>. In addition, drugs are commonly administered to limit inflammation and reduce wound-associated pain and itch. Wrapping of limb extremities (mainly fingers) is also recommended for severe DEB cases to prevent fusion events. Nutrition needs to be adapted in terms of consistency and composition to limit the alterations of the oral mucosae and prevent deficiencies in nutrients. Hyper-protein and hyper-caloric diets are especially recommended for children.

## 2. Research on DEB

### i. Overview of medical research

Different strategies have been proposed and tested to treat DEB (**Figure 13**). They mainly consist of collagen VII restoration through gene therapy, modulation of the translation or supply in exogenous collagen VII. Therapies to limit the manifestation and the progression of symptoms have also been proposed.



**Figure 13. Strategies for the development of new therapies for DEB.**

Approaches to restore functional anchoring fibrils extend from the replacement of the deficient *COL7A1* gene to the supply of exogenous collagen VII. Symptom-based therapies are also essential to limit the progression of the disease and improve the quality of life of patients. Small colored triangles represent drugs.

Briefly, gene therapy aims at restoring the expression of native collagen VII through the correct reprogramming of the deficient *COL7A1* gene<sup>122,123</sup>. This can be achieved through the *ex vivo* reprogramming of cells retrieved from patient skin biopsy before re-administration in skin. Currently, autologous corrected keratinocytes seem to be preferred to fibroblasts for gene therapy, probably because they can easily be applied on wounds after engineering through topical application or grafting<sup>124,125</sup>. Promising results were recently reported using these cells: the administration of corrected keratinocytes improved the closure of more than 90 % of the lesions treated in 7 RDEB patients<sup>126</sup>. Engineered autologous fibroblasts were also reported to restore the expression of collagen VII in non-wounded skin after dermal injection, even if the the treatment appeared less efficient in some patients

than in others<sup>127</sup>. Clinical trials using both types of cells are currently on going<sup>128</sup> to evaluate the effect of treatments on larger patients cohorts (**Table 3**). One remaining limitation to these strategies is the recovery of skin cells from DEB patients presenting healing defects and sometimes generalized open wounds. In addition, the use of grafts appears limited to the cure of small-size wounds. To overcome these limitations, *in vivo* gene therapy was proposed, in the form of a topical application on wounds of a viral transduction-vector containing the functional *COL7A1* gene. Very recently, promising results from the first phases of a clinical trial on a cohort of 12 RDEB patients were reported in literature and the speed of wound closure seemed to be significantly improved by the treatment<sup>129</sup>. If no early adverse effects were reported a few weeks after the administration of the cure, the study now needs to be conducted on a longer duration to evaluate their possible occurrence later in time. More recently, approaches aiming at modulate protein translation to skip deleterious mutations were proposed. PTC read-through using gentamycin (pilot study on a small number of RDEB patients<sup>130</sup>) and exon skipping induced by an antisens oligonucleotide (proof of concept using DDEB and/or RDEB cells<sup>131,132</sup>) were shown to restore the production of functional collagen VII. Nevertheless, these approaches are only adapted to some of the mutations responsible for DEB.

Regarding the exogenous supply in collagen VII, promising results were reported: injections of recombinant collagen VII indeed successfully reversed diseased phenotype in RDEB mouse model<sup>133-135</sup>. A clinical trial has been initiated in 2020 to evaluate the efficiency of the treatment in RDEB patients.

Contrary to the strategies aiming at restoring functional collagen VII, symptom-relief therapies do not target the cause of the disease but aim at limiting its manifestations. One example of such approach is losartan, a drug already used to treat hypertension (angiotensin II receptor type 1 (AT<sub>1</sub>) receptor antagonist) and shown to reduce inflammation, TGF- $\beta$  level and signaling, as well as fibrosis in a RDEB mouse model. Tissue-derived acellular matrices<sup>136</sup> were also tested to cover open wounds. In particular, amnion-based dressings have been shown to be efficient in the treatment of eye-located symptoms in RDEB patients<sup>137,138</sup>.

Whatever the type of therapy, there is a need for a better characterization of molecular and cellular mechanisms involved in DEB pathogenesis<sup>139</sup>. Such data are necessary to better understand symptom development and disease evolution and to provide new targets for symptom-based treatments. They could also help to understand the long-term evolution of reprogrammed cells after grafting and improve gene therapy approaches. A lack of biomarkers of disease severity and progression has finally been highlighted; they are needed to assist the monitoring of the disease and the evaluation of new therapies<sup>106</sup>.

For all these strategies, the main on-going clinical trials are summarized in **Table 3**.

Approach	Clinical trial number	
Recombinant collagen VII	NCT 03752905	
<i>Ex vivo</i> gene therapy		
Keratinocytes	NCT 01263379	NCT 02984085
	NCT 02493816	NCT 04186650
	NCT04227106	
Fibroblasts	NCT 04213261	NCT02810951
<i>In vivo</i> gene therapy	NCT 03536143	
	NCT 02698735	
ARNm reading correction	NCT 03012191	
	NCT 03605069	
	NCT 03928093	
Symptom-based	ECT 2015-003670-32	
	NCT 01716169	
Wound management	NCT 02286427	

**Table 3. Examples of current clinical trials on DEB.**

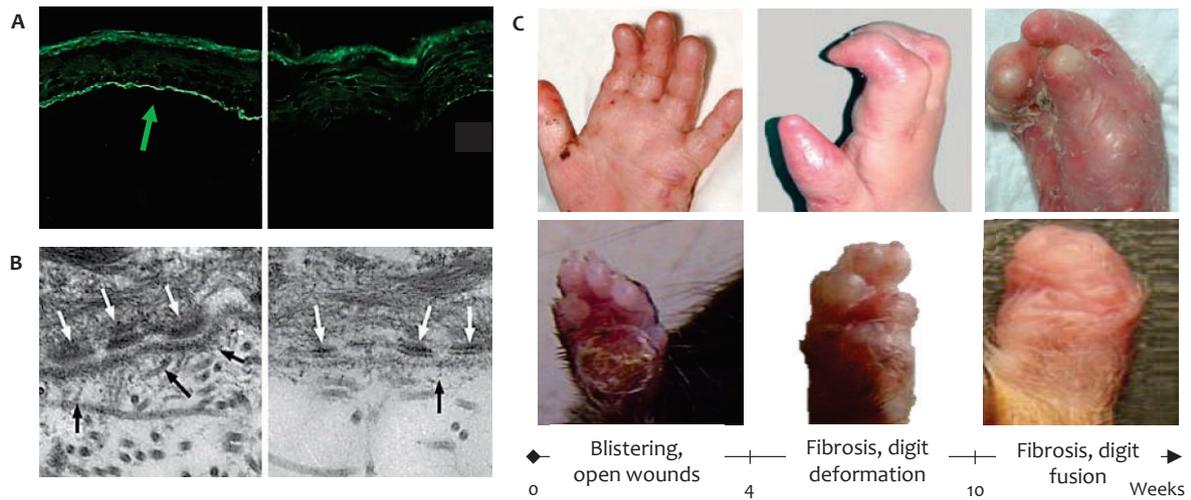
NCT: trials indexed in U.S.A National Trial Library (available on-line on <https://clinicaltrials.gov/>). ECT: trial indexed in European Trial Library (available on-line on <https://www.clinicaltrialsregister.eu/>).

## ii. RDEB mouse model

To provide fundamental and medical studies with a suitable model to analyze the pathology, an animal model of DEB was generated in 2008 at the University of Freiburg, by genetic engineering of mice<sup>140</sup>.

The model is characterized by a homozygous partial knock-out (KO) of the *Col7a1* gene induced by the introduction of the PGK-Neo (Phosphoglycerin kinase-neomycin) cassette in exon 2, leading to aberrant splicing of the mRNA transcript. The mutation approximatively leads to an 85-percent reduction of functional collagen VII approximatively<sup>140</sup>. In contrast to complete *Col7a1* knockout mice which are early postnatal lethal, a proportion of these mice hypomorphic for collagen VII survive into adulthood and develop all signs associated with severe generalized RDEB; consequently they can be referred to as RDEB mice<sup>141,142</sup>. As described in the complete characterization of the model published in 2008 (**Figure 14**), KO mice exhibit epithelium detachment and blistering in tongue and skin, as early as just after birth for some mice. Back of the skin can be used to study the early stages of the disease before the apparition of open wounds. The evolution of forepaws is for its part representative of the evolution of RDEB in human, from blistering to fibrosis. Epidermis blistering and inflammation are observed in young mice, and significant underweight due to malnutrition can be detected early, from 14

to 20 days post-birth. With time, inflammation and fibrosis develop; digit fusion and deformation of the forepaws progressively occur. Around week 10 post-birth, fibrosis is well established in paws and all digits are generally fused: this stage is therefore used to study late stages of the pathology, before cancer development. Because of the pain and of the low quality of life of RDEB mice (malnutrition, multiple wounds and damaged forepaws), animals are sacrificed at the latest at week 12 post-birth, for ethical considerations.



**Figure 14. Phenotype of the RDEB mouse.**

**A.** Collagen VII immunostaining in normal (left) and RDEB (right) mouse skin. The green arrow indicates the detection of collagen VII, which is aligned parallel to the skin surface. **B.** TEM observation of normal (left) and RDEB (right) mouse DEJ. White arrows indicate hemidesmosomes, black arrows anchoring fibrils (also found in hypomorphic skin because of partial deficiency). **C.** Comparison of RDEB mouse forepaws and RDEB human hand. Mouse symptoms mimic the human ones, in terms of nature and evolution (timeline indicates mouse age). Adapted from<sup>140</sup>.

To sum up, the collagen VII-hypomorphic mouse model provides a strong support to the study of RDEB and mimics the human pathology from birth to late fibrosis stage. For the study of SCC, the treatment of RDEB mice with a carcinogen substance was reported to induce the development of cancers very similar to those described in RDEB patients<sup>143</sup>.

### 3. ECM dysregulations in DEB

#### i. Modifications in the basement membrane

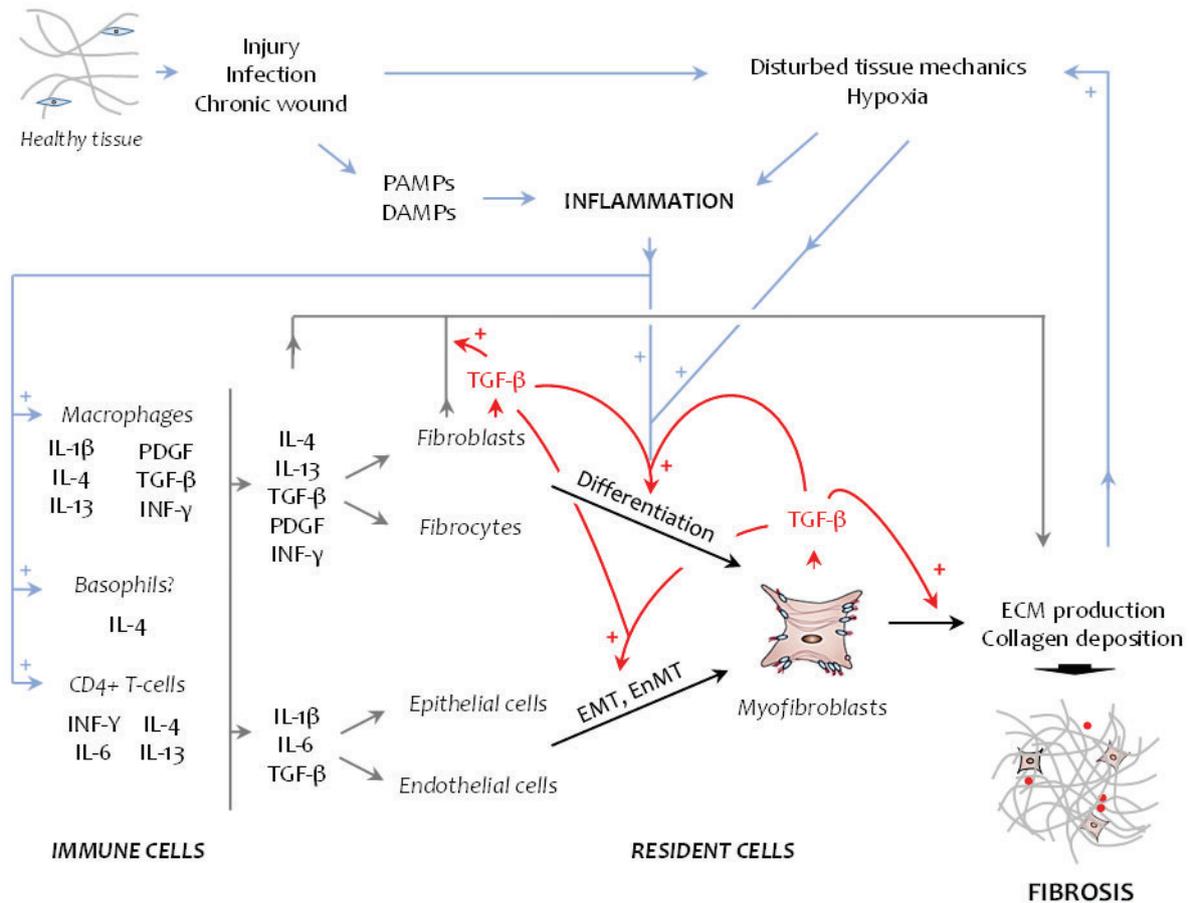
As described above (**Chapter I, I-2**), basement membranes result from the tight association of membrane and matrix components, and are connected with the ECM of underlying tissues. Given the fundamental role of collagen VII in basement membrane homeostasis, efforts have been focused on the description of the dysregulations that occur in this compartment subsequent to collagen VII loss, especially in skin. These dysregulations may help to understand the pathogenesis and the evolution of DEB.

The organization of LM-332, one prominent partner of collagen VII, is especially altered in DEB. In normal skin, LM-332 is distributed linearly in the DEJ, parallel to the skin surface. In RDEB chronic skin wounds, the LM-332 alignment was shown to be disturbed and *in vitro* studies on primary keratinocytes from RDEB mice have shown that the protein aggregates around keratinocytes instead of adopting the physiological patterning<sup>144</sup>. In addition, transcriptomic analyses on biopsies of the unwounded skin from three RDEB patients have shown a decreased expression of the genes encoding chains of collagen IV and VI<sup>145</sup>. Proteomic analyses of the ECM produced by human RDEB fibroblasts from four patients grown *in vitro* have confirmed the diminution in collagen IV deposition (also observed in fibrotic skin from five RDEB patients<sup>146</sup>) and revealed a reduced deposition of nidogen-1 too<sup>146</sup>. Potential variations in the distribution of these molecules *in vivo* remain to be studied.

## ii. Dermal fibrosis

In RDEB, fibrosis affects the dermis in the late stages of the disease. Fibrosis is defined as a chronic and pathological accumulation of extracellular matrix in a tissue resulting from alterations of the wound healing process<sup>103</sup>.

During physiological skin wound healing, the beginning of the proliferative phase is characterized by an inflammatory process that progressively resolves while the wound evolves towards the remodeling phase (**Chapter I, I-6**). In a fibrotic environment, the proliferative phase fails to resolve and dermal cells (fibroblasts, myofibroblasts) keep their over-activated state. The process of fibrosis is complex, not completely understood and exhibits substantial variations in drivers and molecular manifestations between different tissues and contexts<sup>43,147</sup>. However, it generally involves the development of a highly inflammatory environment driven first by immune cells (macrophages, T-cells) and then maintained through the action of both resident cells (including fibroblasts, endothelial and epithelial cells) and immune cells. The secretion of various cytokines, especially interleukins (ILs)<sup>148,149</sup> and growth factors by all these cell types leads to an excessive and aberrant deposition of extracellular matrix, especially fibrillar collagens<sup>44,149</sup>. Whatever the tissue affected, TGF- $\beta$  is considered as the key promoter of fibrosis that mainly drives the fibrotic process<sup>150</sup>. A summary of the cross-talk between immune cells and other skin cells is proposed in **Figure 15**.



**Figure 15. Summary of the processes involved in fibrosis (general processes currently described in literature).**

Immune cells release growth factors and cytokines that activate resident skin cells and circulating cells and drive their differentiation into myofibroblasts (grey arrows); basophils are suspected to participate to this activation process. In addition to myofibroblasts, macrophages and fibroblasts contribute to ECM excessive deposition. The modifications of matrix structure due to resident cell activity and the adverse effects of open wound contribute to maintain fibrosis through direct actions on fibroblast differentiation and promotion of inflammation (blue arrows). Finally, fibroblasts and myofibroblasts secrete TGF- $\beta$  in response to immune cell action and amplify the mechanisms leading to ECM production (red arrows) and inflammation upholding. DAMP: damage-associated molecular pattern, EMT: epithelial-mesenchymal transition, EnMT: endothelial-mesenchymal transition, IL: interleukin, INF- $\gamma$ : interferon  $\gamma$ , PAMP: pathogen-associated molecular pattern, PDGF: platelet derived growth factor, TGF- $\beta$ : transforming growth factor  $\beta$ .

Some aspects of RDEB-associated fibrosis have already been described in RDEB mouse (fibrotic paw skin used for analyses), including dysregulations of the ECM. Increased abundance of fibronectin and altered arrangement of dermal collagen fibrils were especially reported<sup>141</sup>, as well as an improved abundance of tenascin-C, a protein reported to exert pro-fibrotic activity<sup>151</sup>. This latter observation was also made in fibrotic skin biopsies from RDEB patients<sup>146</sup>. In addition, the increase in the abundance of periostin, a protein modulating the cross-linking of fibrillar collagens<sup>152</sup> and reported to contribute to the development of idiopathic pulmonary fibrosis<sup>153</sup> and myocardial fibrosis<sup>154</sup>, was recently reported in the secretome of RDEB patients<sup>155</sup>. The protein was suggested as a potential biomarker for RDEB monitoring too. Modifications in TGF- $\beta$  signaling activity were also described in the fibrotic skin of RDEB mouse. The abundance of the growth factor at the protein level and its canonical downstream

signaling were increased. In addition, TSP-1, an activator of the signaling of TGF- $\beta$ , was reported to be up-regulated<sup>156</sup>. Finally, recurrent bacterial infections affecting open wounds may exacerbate inflammation and contribute to the enhancement of TGF- $\beta$  signaling and promote inflammation and fibrosis in DEB<sup>142,157</sup>.

DEB can finally be described as a progressive disease characterized by an evolution of symptoms from skin blistering to a debilitating fibrosis preceding the development of aggressive cancers. Some ECM alterations occurring before (basement membrane) and upon fibrosis (dermal compartment) have already been reported. Nowadays, current research focuses on cures aiming at correcting the cause of the disease, but the development of new symptom-relief therapies remains necessary to improve the quality of life of patients.

## IV. PROTEOMICS AND MASS SPECTROMETRY

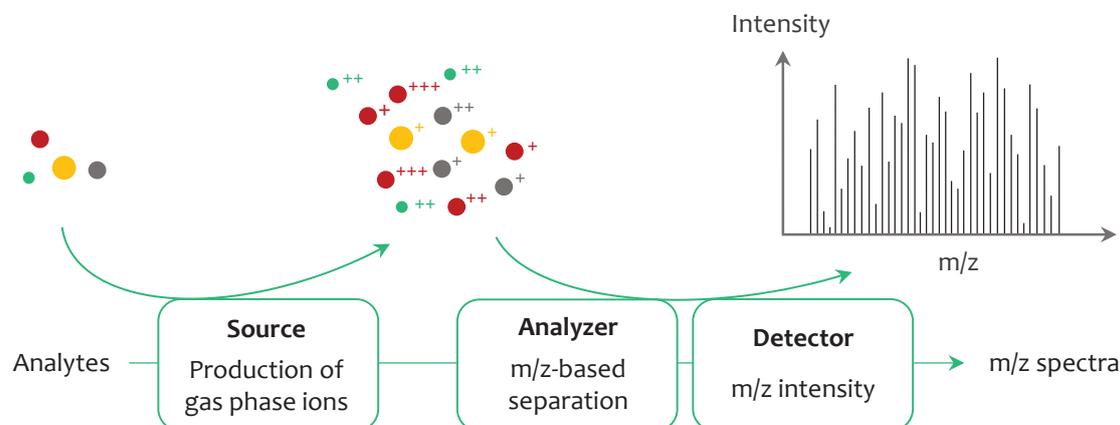
The protein composition of a living entity at a specific time point is called “proteome”. In this way, the proteome reveals the biological activities taking place in an organelle, a cell or a tissue and reflects the “global state” of the structure under study. Proteomic approaches aim at describing proteomes to better characterize physiological mechanisms and understand the dysregulations responsible for and/or associated to pathology.

### 1. Introduction to mass spectrometry

Mass spectrometry (MS) approaches are adapted to the characterization of the structure of both organic and inorganic molecules. Initially developed to work on single or limited numbers of analytes, mass spectrometry is now one of the top technologies for broad-range descriptive or targeted studies of proteomes.

#### i. Principle of mass spectrometry

Mass spectrometry involves the separation of ionized and gaseous molecules to determine their respective mass/charge ratio ( $m/z$ ) used to deduce their structure.



**Figure 16. Schematic representation of the key steps and equipment composition in mass spectrometry.** Analytes are transformed into gas phase ions in the source and analyzed in terms of chemical composition (analyzer) and amount (detector) to get final  $m/z$  spectra.

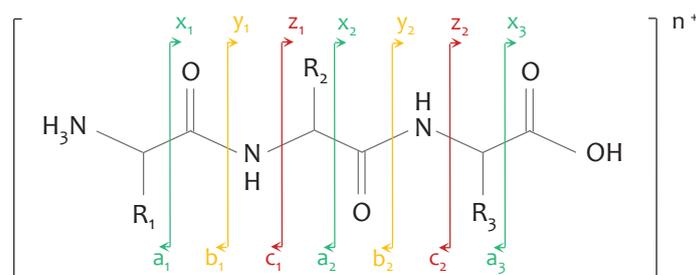
The first mass spectrometry analysis was reported in 1912 by Joseph John Thomson<sup>158</sup> and led to the discovery of isotopes. Successive improvements of the initial Thompson’s equipment were proposed and modern mass spectrometers are now composed of<sup>159</sup>: (1) a “source” to produce gas phase ions from the initial analytes, (2) a “mass analyzer” to separate ions according to their  $m/z$  ratios, (3) a “detector” to count the ions and finally determine the abundance of each  $m/z$  signal (**Figure 16**). For each ion, these two values ( $m/z$  and abundance) are collected in the form of a spectrum that constitutes the output

of mass spectrometry analysis. At the end of the MS scan, a list of signals (= spectra) is generated and finally processed to determine the composition and organization in atoms for each analyte.

## ii. MS to study proteins

MS-based proteomics aim at identifying proteins in complex samples. The chemical structures detected by MS are used to determine the amino-acid (AA) sequences of protein-derived analytes and some of their post-translational modifications (PTMs), to finally identify the proteins contained in the sample and eventually quantify them<sup>160,161</sup>. MS processes and equipment are continuously improved to meet these specific needs, mainly in terms of resolution, accuracy in mass and sensitivity.

For small molecules with chemical formula of low complexity (chemicals), full-scan mass spectra (MS1) can be sufficient to identify their composition in atoms. For bigger molecules with complex molecular formula, including proteins made of tens to thousands AAs, MS1 signals are not sufficient to precisely infer a composition in atoms. Tandem mass spectrometry (MS/MS) has been developed to meet this specific need by coupling two analyzers with a fragmentation cell<sup>159,160</sup>. Initial ionized analytes named “MS1 precursors” are selected in a first analyzer, then fragmented by collision against an inert gas into smaller product ions (“MS2 ions”) following a fixed reaction. The fragment ions are finally separated in a second analyzer and a MS2 spectrum is obtained. Different types of fragmentation exist for proteins and peptides, each inducing a specific cut along the chain of amino acids depends in; a-, b-, c-, x-, y- and z-ions can be produced (**Figure 17**). Collision-induced dissociation mainly generate b- and y-ions by colliding biomolecules with an inert gas such as argon or helium to break MS1 ions up<sup>162</sup>. The method was adapted into higher-energy collisional dissociation (HCD) that allows the additional study of a- and z-ions. These two fragmentation methods are the main used in MS-based proteomics, especially for bottom-up studies (**Chapter I, IV-2-i**)<sup>163</sup>. Electron-transfer dissociation, involving an electron donor such as fluoranthene to fragment MS1 precursors and generate c- and z-ions, was also used to analyze whole proteins and post-translational modifications by top-down proteomics<sup>164</sup>. Whatever the fragmentation type applied, MS2 spectra are used to determine the structure of MS2 ions and finally the structure of each MS1 generating precursor is obtained from its MS2-derived ions<sup>165</sup>. In addition, the sensitivity of the MS/MS scan clearly outperforms that of the MS scan.



**Figure 17. Nomenclature of the MS2 ions that can be generated during fragmentation by mass spectrometry.**

For MS-based analysis of biomolecules, three main characteristics are particularly important for the mass spectrometer<sup>159</sup>:

- the sensitivity, corresponding to the lowest amount of analyte that can be detected,
- the resolution, corresponding to the lowest difference between two  $m/z$  values that can be detected by the mass analyzer (the higher the resolution, the better the ability to discriminate two different analytes with very close  $m/z$ ),
- the dynamic range, in other words the range from minimal to maximal amount that can be analyzed for all ions; if the dynamic range is too low, the analysis of proteins with low abundance is impossible and the detection is limited to the most abundant ones (for instance in plasma, cytokines exhibit concentration  $10^{12}$  times lower than immunoglobulins<sup>160</sup>).

Mass spectrometers and their components have been specifically developed/optimized by suppliers to improve these parameters and propose the best combinations of sensitivity, resolution and dynamic range.

All the MS-based proteomics experiments of this project were conducted on a Q Exactive HF mass spectrometer (Thermo Scientific), which was specifically developed for proteomics. Briefly, the Q Exactive HF is equipped with an electro-spray source to ionize molecules: the gaseous ions are produced by successive fragmentation/desolvation of droplets from a liquid sample. Electro-spray especially allows a soft molecule-preservative ionization of analytes present in a wide molecular mass range in the sample, from a limited mass of biomolecules (one to a few hundreds of nanograms used for analysis)<sup>159,166</sup>. Ions separation and analysis are performed by MS/MS through the combination of two different mass analyzers. Prior to fragmentation, a quadrupole<sup>159</sup> allows a fast separation of MS1 precursors<sup>163,166</sup>. For the analysis of MS2 ions, an Orbitrap<sup>159</sup> exhibiting a very high resolution (140,000) and compatible with the characterization of ions exhibiting a wide range of abundance is used. Between the two analyzers, the fragmentation is ran by HCD. This equipment is consequently adapted to the analysis of limited amounts of complex mixtures of biomolecules and fits the specific needs described above for proteomic studies<sup>167</sup>. In addition, both discovery and targeted experiments (**Chapter I, IV-2-iii**) can be performed using the Q Exactive HF.

## 2. Approaches in MS-based proteomics

MS-based proteomics can be applied to the analysis of the protein content of any biological material. Different types of samples have been characterized by MS-based proteomics: lysates or supernatants from cells cultured *in vitro*, tissue lysates<sup>97,98</sup>, biological fluids such as plasma<sup>168,169</sup> and wound exudates<sup>170</sup>, formalin-fixed paraffin-embedded samples<sup>171</sup> and even 1.5 million year-old fossil bones<sup>172</sup>.

To characterize the protein content of biological samples using MS, various strategies have been proposed in the literature. They can be classified depending on the state of the protein during MS injection, the method to acquire MS signals and the diversity of the targets to analyze.

### i. Top-down and bottom-up strategies

In MS-based proteomics, the sequence in AA of proteins can be obtained from the study of native proteins (top-down approach). Another method, especially well-suited for complex samples, involves the analysis of peptides, generated by the protease-driven digestion of proteins, to identify the proteins originally contained in the sample (bottom-up proteomics). The main characteristics of these two strategies are summarized in **Table 4**<sup>173,174</sup>.

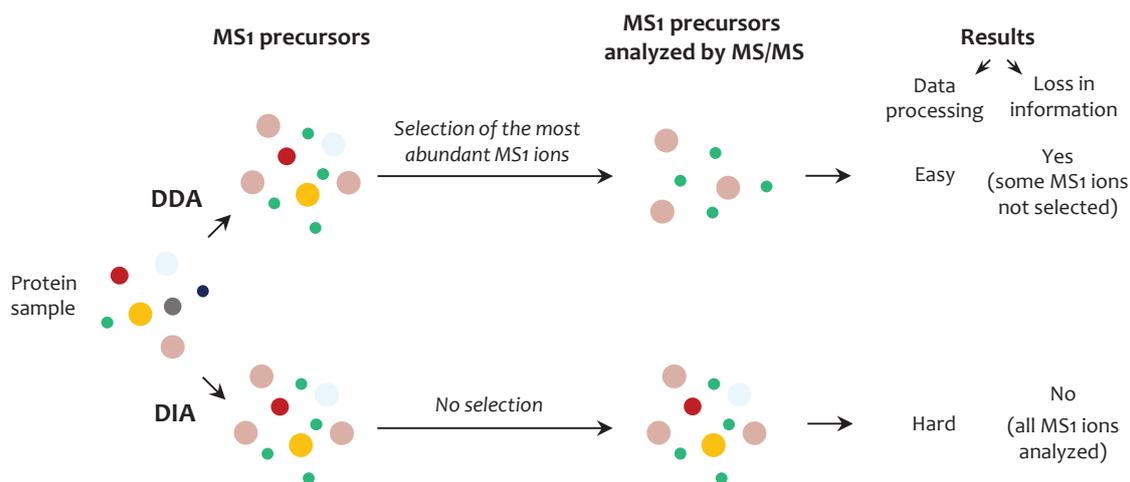
	Top-down	Bottom-up
<b>Nature of the analytes</b>	Denatured proteins, purified from the sample of origin	Proteins extracted from the sample of origin and digested into peptides by a protease (generally trypsin)
<b>Size range of MS1 precursors obtained</b>	A few to 10 <sup>6</sup> Da (a few to hundreds of amino acids)	50 to 3,500 Da
<b>Ionization and fragmentation for MS/MS</b>	Dependent on the mass of the analyte (difficult when the analyte's molecular weight exceeds 50 kDa <sup>173,174</sup> ) Analysis of peptides consequently easier than proteins	
<b>Sensitivity</b>	Top-down 100-fold less sensitive than bottom-up <sup>173</sup>	
<b>Applications</b>	<ul style="list-style-type: none"> <li>- Proteoforms<sup>175</sup> and phosphorylation identification</li> <li>- Description of low-complexity samples</li> <li>- Exact elucidation or preliminary study of the sequence of pure proteins</li> </ul>	<ul style="list-style-type: none"> <li>- PTM identification</li> <li>- Identification and quantification of tenths to thousands of analytes from complex protein samples</li> <li>- Targeted proteomics</li> </ul>

**Table 4. Characteristics and quick comparison of top-down and bottom-up approaches in MS-based proteomics.**

At present, most of the proteomic studies reported in the literature aim at describing the complete proteome of biological samples and favors bottom-up approaches.

### ii. Data acquisition

In MS-based proteomics, two acquisition methods can be used, depending if a selection of MS1 precursors prior to MS/MS analyses (= use of the data acquired during the MS1 scan) is made or not (**Figure 18**).



**Figure 18. Strategies for the acquisition of data in MS-based proteomics.**

DDA: data-dependent acquisition, DIA: data-independent acquisition.

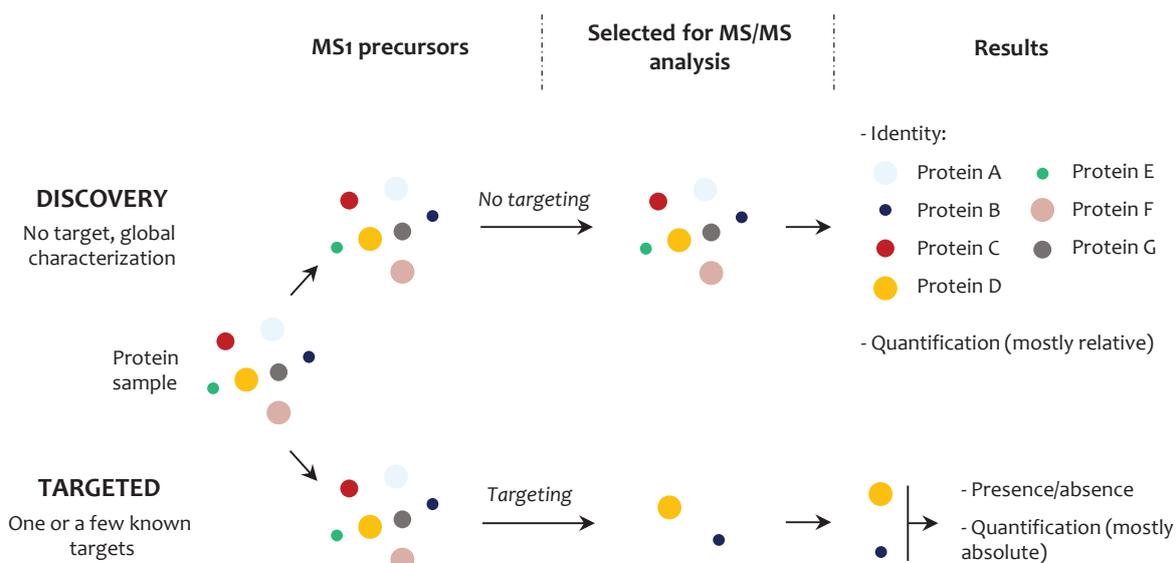
Data-dependent acquisition (DDA) is the reference method in the literature and involves a selection of MS1 precursors prior to fragmentation, based on parameters defined by the operator (intensity,  $m/z$ , charge, etc.). The MS/MS analysis is limited to some MS1 ions matching the parameters, and thus “depends” on MS1 data. This results in an incomplete characterization of the sample since the information brought by unselected MS1 precursors is lost. The use of fractionation strategies such as HPLC separation prior to MS/MS analysis limits the loss of information in DDA by decreasing the complexity of MS1 precursors analyzed at a given point during the scan. Another possibility is the configuration of the equipment to avoid the repeated analysis of the MS1 precursors that have already been characterized by MS/MS during the run, to favour the selection of additional precursors of interest. In spite of this main drawback, DDA permits a fast acquisition of data and generates result files which can be easily processed by bioinformatics<sup>159,160</sup>.

In the data-independent acquisition (DIA) mode, no selection of MS1 precursors occurs: all MS1 ions are accumulated and specific windows of  $m/z$  ranges are successively selected for MS/MS analysis. For each window, the total MS1 content is submitted to MS/MS, preventing any information loss. The MS scan lasts until all the windows of the MS1  $m/z$  values of interest have been analyzed. DIA however generates very complex results due to the computing of every MS2 signal and data are very difficult to process compared to DDA. In addition, the building of a reference database containing reference spectra to match the signals acquired in DIA with the corresponding peptides is mandatory, but very difficult to achieve especially for the analysis of complex samples. Finally, DIA suffers from the dynamic range problem more than DDA does because of the simultaneous analysis of hundreds of MS1 precursors contained in each  $m/z$  window. If DDA remains the reference strategy, studies using DIA have strongly increased during the last 5 years<sup>176,177</sup>.

### iii. Discovery and targeted proteomics

When running MS-based proteomics, the objective can be the description of the whole protein content (discovery proteomics) or the specific analysis of one or several targets among all the molecules contained in a mixture (targeted proteomics; **Figure 19**).

Discovery or shotgun proteomics aim at characterizing as many proteins as possible in a complex sample, mainly by tandem MS using a bottom-up strategy (**Chapter I, IV-2-i**)<sup>174</sup>. In discovery proteomics, all the MS/MS signals acquired by DDA or DIA are processed by bioinformatics pipelines to match as many spectra as possible with peptides derived from protease digestion. Peptides are then used to identify the proteins from which they originate using large proteome databases. Databases are species-specific, and they contain all the protein sequences that can be encoded by the genome (obtained through a virtual transcription/translation process). Depending on the protease used to digest proteins, bioinformatic algorithms predict the digestion of this whole virtual proteome into peptides *in silico* and generate a virtual “peptidome” database. Each  $m/z$  detected by MS/MS is compared to this peptidome to be matched with a corresponding peptide and used to identify proteins<sup>165</sup>. Each protein identification is associated with a score reflecting the confidence that can be given to the result. To date, discovery proteomics has especially been used for quantitative studies aiming at describing the global changes occurring in the proteomes of cells or tissues in disease. Such studies are useful to reveal novel potential biomarkers of disease progression. In addition, bioinformatic tools have been developed to exploit the raw data obtained at the protein level and facilitate subsequent biological interpretation with functional annotation, over- or under-representation analyses of GO (Gene Ontology) terms, heat maps, etc. These tools are mainly open-source and provided as on-line services (e.g. DAVID<sup>178</sup>, Panther<sup>179</sup>, ProteoRE<sup>180</sup>).



**Figure 19. Schematic view of discovery and targeted approaches in MS-based proteomics.**

Discovery proteomics aim at identifying as many proteins as possible in a complex sample, while only a few known and specific proteins are searched in targeted proteomics. The two strategies are compatible with quantitative analyses.

In targeted proteomics, only one or a few biomolecules in a sample are specifically monitored by MS/MS, with the main objective being to compare their abundance in different conditions (health versus disease, treatment *vs* no treatment, treatment 1 versus treatment 2 versus treatment 3, etc.)<sup>181</sup>. Almost all targeted proteomics studies are based on bottom-up approaches, because peptides are easier to analyze by MS/MS compared to proteins (higher ionization and fragmentation efficiencies, **Chapter I, IV-2-i**) and compatible with the use of heavy-labeled standards for absolute quantification (**Chapter I, IV-3-i**). Different technologies have been developed to analyze specific peptides using both DIA and DDA approaches. They all rely on the targeted selection of a few specific MS1 ions from the same or different proteins, followed by a global (parallel reaction monitoring PRM) or targeted (single reaction monitoring SRM, multiple reaction monitoring MRM) monitoring of their MS2-derived ions (“MS2 transitions”). This selection of ions accounts for the high specificity and sensitivity of targeted proteomics methods, especially in SRM and MRM where two selections are successively performed (MS1 and MS2 ions). Most targeted proteomics studies reported in literature have been conducted using MRM and PRM because these approaches allow the simultaneous monitoring of several MS2 transitions<sup>159</sup>. Adaptations of MRM into MRM<sup>3</sup>, a targeted method involving an additional fragmentation of MS2 transitions for the monitoring of MS3 transitions, have also been proposed<sup>182</sup>. Targeted proteomics can be used in line with discovery proteomics for the study of biological samples: after the initial step of proteome screening by discovery proteomics, potential protein biomarkers are selected and then validated by targeted proteomics. Numerous studies are currently on-going to transfer these technologies from research to clinic for diagnostic or treatment-efficacy evaluation purposes<sup>168-170</sup>.

### 3. Quantification in MS-based proteomics

Almost all the studies involving MS-based proteomics aim at comparing the amount of proteins between different experimental conditions or phenotypes. Two main strategies are available to quantify molecules in MS-based proteomics:

- absolute quantification, to determine the exact amount of analytes in a classical quantity unit (mass, moles),
- relative quantification involving the comparison of the amounts of analytes between various conditions without determination of the exact quantities (mainly used for discovery proteomics)<sup>165,177</sup>.

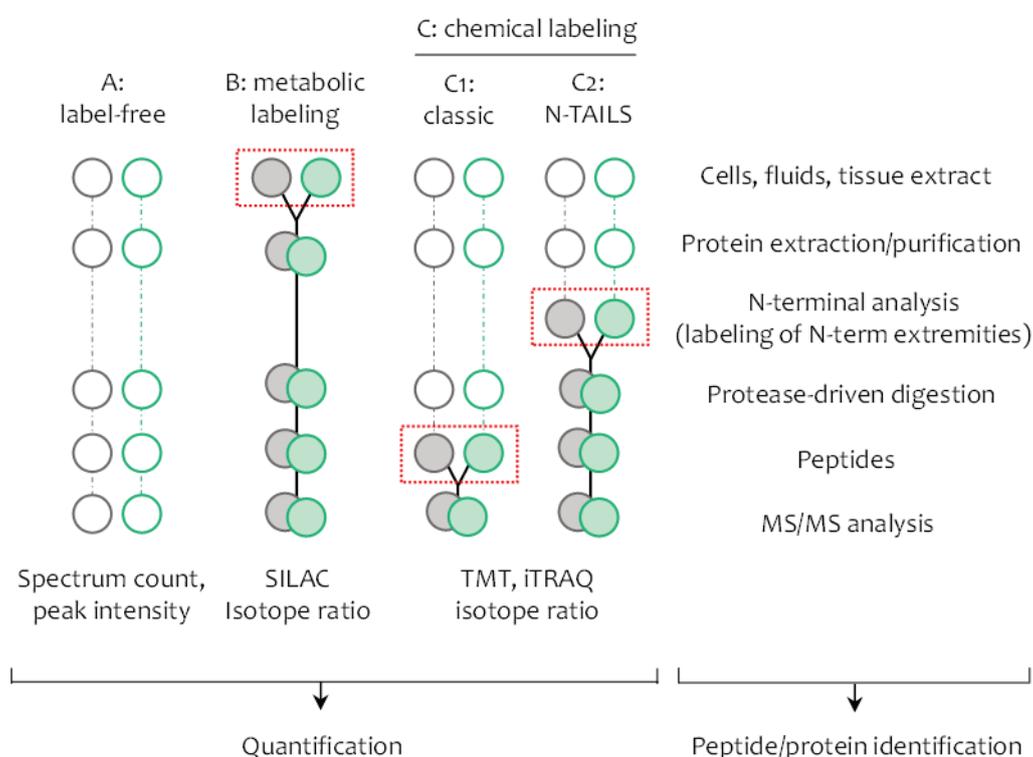
A concise description of the principle and the tools available for each of these two strategies is proposed below. For extensive descriptions, all label-free and label-based quantification methods have been extensively reviewed and compared in the literature<sup>177,183</sup>.

### i. Overview of absolute quantification

Absolute quantification requires a standard for each molecule of interest to compare the MS signal intensity with the abundance of the analyte<sup>184,185</sup>. Standard can be internal (directly mixed with the sample to analyze, isotope-labeled) or external (used in parallel to the sample to analyze) and consists in a synthetic peptide/protein exhibiting the same sequence in AAs than the natural one. Because each protein/peptide behaves differently during sample preparation and MS analysis, each natural analyte studied requires its own standard for correct quantification. As a consequence, this quantification strategy is limited to targeted proteomics where only a few known peptides are monitored. In discovery proteomics, peptide identities are not known at the time of the experiment and this approach is not compatible with absolute quantification. The different absolute quantification methods (label-free, label-based using internal or external isotope-containing calibrators) will not be detailed in this manuscript but can be found in recent literature reviews<sup>184,186</sup>.

### ii. Relative quantification methods

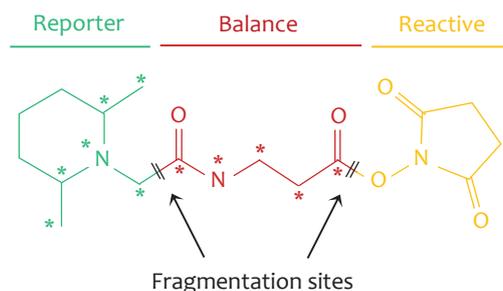
Numerous strategies have been developed in the last 20 years to propose fast and accurate relative quantification methods applicable to discovery proteomics. They are divided into label-free and label-based methods (**Figure 20**).



**Figure 20. Schematic representation of strategies for relative quantification and their main steps.** Label-free and label-based methods can be applied to compare protein abundance between different samples. Red rectangles indicate labeling steps; white and colored circles respectively indicate unlabeled and labeled samples. Adapted from Dussoyer et al<sup>185</sup>.

Label-free approaches do not involve any modification, any “labeling” of proteins or peptides prior to MS/MS analysis: every sample is injected separately from the others to differentiate the proteins/peptides from one condition to another. The number of detected spectra (“spectral counting” method) or the intensity of the peak (“area under the curve” or “extracted ion chromatogram” method) for each MS1 (DDA mostly) or MS2 (DIA mostly) ions is used to compare conditions<sup>165,184</sup>. These strategies do not involve any additional processing of samples before injection in the mass spectrometer. They however remain time-consuming and expensive when studying large cohorts (injection of each sample individually). The comparability between samples is also limited because signals are not acquired during the same MS run. The implementation of methods to normalize quantities between different samples/injections is mandatory when working with label-free strategies, as well as the preliminary demonstration of the reproducibility of sample processing and data acquisition<sup>187</sup>.

To improve the comparability of results, label-based strategies have been developed to allow the co-injection of samples. They involve the labeling of the proteins/peptides of each condition using specific isotope combinations to differentiate the biomolecules from different samples<sup>165</sup>. Carbon and nitrogen isotopes are favored for labeling, hydrogen ones being known to induce shifts in retention times when HPLC is used prior to MS/MS<sup>177</sup>. Isotope(s) can be integrated to biomolecules *in vitro*<sup>186</sup> or *in vivo*<sup>188</sup> by feeding cells in culture or animals with labeled essential amino-acids; these strategies are grouped under the SILAC (stable isotope labeling of amino acid in cell culture) and *in-vivo* SILAC appellations. They are however expensive and time-consuming to implement (integration of labeled amino-acid to all proteins can be long to achieve). Chemical labeling is the top-used strategy for MS-based quantitative proteomics because easy and quick to implement. It consists in the covalent linkage of isobaric and isotope-labeled tags to biomolecules and allows the multiplexed analysis of up to 18 conditions (11 at the beginning of my PhD)<sup>184</sup>. In regular quantitative proteomics, tags are bound to peptides after protease-driven digestion. Adaptations to analyze specific subproteomes, such as the N-terminome which can be analyzed by the N-TAILS approach (N-terminal amine isotopic labeling of substrates), were also reported: tags are then bound to the N-terminal extremity of native proteins before digestion and used as flags to identify native N-termini<sup>86,87</sup>. All these tags are divided into 3 parts (**Figure 21**): (1) a reporter group, made of a defined combination of isotopes which is released during fragmentation and is used for relative quantitation, (2) a balance group to maintain a constant overall mass for all tags and insure identical physico-chemical properties of labeled peptide ions, and (3) a reactive group to create the covalent peptide/tag link<sup>177</sup>.



**Figure 21. Structure of 11-plex Tandem Mass Tag.**

The reactive group binds reactive amine functions carried by proteins/peptides, the reporter part constitutes the quantitative information recorded during analysis and the balance part is used to conserve a constant mass, and thus identical physico-chemical properties, between the different tags used during an experiment. \*: position of isotopes.

Different types of tags have been developed<sup>177,184,189–191</sup> but, in all cases, molecules from all samples can be pooled after labeling and analyzed simultaneously during a single MS run. The reporter ion released during the MS/MS experiment provides the quantitative information: each type of tag, characterized by a specific  $m/z$  in relation with its isotope content, corresponds to one condition and is quantified after fragmentation. The relative abundance of peptides or proteins is finally obtained by computing the relative abundance of tags for all ions belonging to a specific peptide or protein. If all these strategies are compatible with the quantification of MS1 and MS2 ions (except for SILAC which allows a MS1 quantification only), MS2-quantification is favored: exploring biological hypotheses indeed involves the comparison of several conditions, and MS1-based quantification is more limited in terms of multiplexing<sup>177</sup>. At present day, Tandem Mass Tag (TMT, **Figure 21**) is the most widely used chemical tag for label-based quantitative proteomics. It allows the combination of 18 different samples in a single run (highest multiplexing capacity currently available) and benefits from a specific software, optimized to process the quantitative data obtained with this tag.

#### 4. Studying the ECM by MS-based proteomics

Discovery and targeted proteomics now allow the global or specific analyses of proteins in complex fractions in terms of quality and quantity and numerous strategies were proposed to characterize the protein content of biological samples. Almost all of them focus on the analysis of healthy or diseased tissues from human or mouse origin. To determine the contribution of the ECM to disease processes, the matrisome databases are now frequently used to identify matrix protein from MS/MS raw datasets (**Chapter I, II-4**).

For the specific analysis of the ECM from native tissues by MS-based proteomics, sample preparation is more difficult compared to what is usually performed for studies on whole tissues.

It is indeed necessary to eliminate as many cell proteins as possible while avoiding the loss of matrix components during sample processing. The reference protocol for the recovery of ECM components from native tissues is compartmental enrichment. It was proposed by Naba and colleagues<sup>192</sup> in line with the publication of matrisome lists. The strategy aims at eliminating proteins from the different compartments of the cell (cytosol, nucleus, membrane, and cytoskeleton) through the treatment of the tissue with commercial buffers with increasing concentrations of salts and detergents. At the end of the procedure, the remaining insoluble pellet of tissue is considered as the ECM and is analyzed by MS/MS. This strategy was used several times, especially for the characterization of cancer samples and tumors (e.g. breast<sup>97</sup>, lung<sup>100</sup>, pancreas<sup>98</sup>). Another approach, using salts and detergents to eliminate all cell proteins independently from their compartment of origin, was applied several times to describe the matrisome of tissue from the cardiovascular system<sup>193-195</sup>. More recently, a strategy named QDSP (quantitative detergent solubility profiling) was proposed to fractionate proteins depending in relation with their solubility<sup>99</sup>. At the beginning of my PhD, no description of the matrisome of human or mouse skin by MS-based proteomics was available in the literature.

# OBJECTIVES

---

The main objective of my PhD work was to characterize the dysregulations of the skin extracellular matrix in RDEB in order to identify new potential biomarkers and/or targets for therapy. As collagen VII is a key molecule for the connection of the two ECM-rich compartments of the skin (basement membrane and dermis), we hypothesized that major impairments in the skin ECM composition and organization would occur when it is present in abnormally low amounts. As described above, the fragile skin of RDEB patients becomes severely affected by recurrent wounding and, with time, displays signs of scarring and fibrosis which can lead to finger fusion. Furthermore, cutaneous squamous cell carcinoma driven by the injured and fibrotic dermal microenvironment is the major cause of death in RDEB patients. Therefore, in addition to causal therapies aiming at restoring a normal level of collagen VII, it is also very valuable to propose new options to treat the pathological manifestations occurring subsequent to skin fragility. For such efforts to be sustainably successful, an increased knowledge of the mechanisms involved in the establishment and progression of skin symptoms is needed. Some ECM alterations have already been reported but we suspected that they were incompletely described and that a better understanding of these alterations would provide new directions to develop symptom-relief therapies. A more modest but equally valuable objective would be to identify new biomarkers that can be used to evaluate disease severity and/or assist clinical trials. This work was conducted in the context of the DEBtarget project funded by ANR and DFG and also involved the group of Alexander Nyström (group leader) and Dimitra Kiritsi (dermatologist) at the Medical Center of the University of Freiburg, in Germany.

As a first approach to identify dysregulated ECM remodeling pathways, patient samples being limited and precious, we decided to work with a RDEB mouse model, the well-characterized collagen VII-hypomorphic mouse. In addition, we used ECM-centered proteomics as a primary technique to get a broad and hypothesis-free overview of the changes occurring in the matrix.

Thus, the first part of the present work was dedicated to the development of a protocol adapted to the efficient recovery and MS-based identification of extracellular matrix proteins contained in whole skin samples. This also allowed me to get familiar with the matrisome concepts and annotations which were developed these last years and are now largely used in the ECM field.

The second objective of my PhD work was to use the previously-optimized approach to describe ECM dysregulations in the skin of RDEB mice. To do so, the skin matrisome was screened by

## Objectives

quantitative MS-based proteomics to evidence potential changes in protein abundance at the molecular level. To complement proteomic analyses, the ultrastructure of the skin was also analyzed by transmission electron microscopy with a thorough evaluation of the consequences of the loss of anchoring fibrils on matrix organization at the tissue level.

Finally, some of the most striking observations made in the previous part were investigated further to try to identify the mechanisms involved in these alterations. This was done through targeted methods such as RT-qPCR, Western blot or immunohistostaining. At the end of this part, we also proposed a preliminary proof-of-concept that targeted proteomics (PRM) can be a suitable approach to monitor ECM-derived biomarkers in RDEB skin.

---

## CHAPTER II: MATERIALS AND METHODS

---

## I. GENERAL PROCEDURES AND REAGENTS

### 1. Recovery of biological samples

#### i. Biological samples for the comparison of enrichment protocols

Animal experiments followed governmental and international guidelines. Mouse back skin biopsies were collected from pathogen- and disease-free C57BL/6J female animals aged 22 weeks obtained from the Jackson Laboratory and hosted in the animal facility of the LBTI until sacrifice. After sacrifice, the back skin was shaved to remove as much hair as possible and depilated twice using depilatory cream (3059944026734, Veet). The skin was dissected using scissors and separated from the underlying tissues, taking care of removing as much subcutaneous fat as possible, and stored in 2-mL Eppendorf protein Lobind tubes. Tissues were flash-frozen in liquid nitrogen and stored at - 80 °C until use.

#### ii. RDEB mouse model and recovery of samples

The mice used all along the present work (WT and RDEB, **Chapter I, III-2-ii**) were hosted by our German collaborators. The studies on animals were approved by the regional review board (Regierungspräsidium Freiburg, Freiburg, Germany; approval no. 35/9185.81/G15-140) and followed governmental and international guidelines. For skin biopsies, animals were sacrificed and the back skin was recovered and stored as described above. Only the depilatory cream step was skipped, in order to prevent unwanted epidermis blistering. Paws were also recovered immediately after sacrifice (cut with scissors), stored in 2-mL Eppendorf protein Lobind tubes and flash-frozen in liquid nitrogen. All samples were recovered in Freiburg and kept at - 80 °C during all storage and transport steps until their use in Lyon. Samples from control animals will be named WT and samples from collagen VII-hypomorphic animals will be named RDEB in the following parts of the manuscript.

### 2. Chemicals

Ethanol (4146012) was purchased from Carlo Ebra Reagents.

4',6-diamidino-2-phenylindole (DAPI, 10-50A), N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES, 10-110), tris(hydroxymethyl)aminomethane (Tris, 26-128-3094-B) and NaCl (1112-A) were purchased from Euromedex.

2,2,2-trichloroethanol (TCE, T54801), acetonitrile (ACN; 34998), ammonium bicarbonate (A6141), calcium chloride (CaCl<sub>2</sub>, 2233506), chloroform (34854), N-cyclohexyl-3-aminopropanesulfonic acid (CAPS, 2632), dimethyl sulfoxide (DMSO, dithiothreitol (DTT; D0632), DNase I (04536282001), ethylenediaminetetraacetic acid (EDTA; E5134), formic acid (FA, 33015), glycerol (G9012), guanidinium hydrochloride (GuHCl; G4630), NP-40 (I3021), iodoacetamide (I1149), hydroxylamine (43-822-7), MgCl<sub>2</sub> (M1028), β-mercaptoethanol (M7154), methanol (34860), phospholipase A2 (PLA2; P9279), cOmplete™ protease inhibitor cocktail (PI, 11836153001, prepared according to the

manufacturer's instructions, stock solution to be dissolved 7-fold before use), cOmplete™ EDTA-free protease inhibitor cocktail (11836170001, prepared according to manufacturer's instructions, stock solution to be dissolved 7-fold before use), sterile phosphate buffered saline solution (PBS; D8537), sodium acetate (S2889), sodium deoxycholate (SD; D6750), sodium dodecyl sulfate (SDS; L3771), NaOH (S5881), trifluoroacetic acid (TFA; 91707), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl, T3253), Triton X-100 (T8787), Tween-20 (P1379) and urea (33247) were from Merck.

10-plex TMT reagents (90110) were purchased from Thermo Fisher Scientific.

For MS-based proteomics, N-glycosidase-F (PNGase-F, P0709S), endoproteinase Lys-C (Lys-C, 12505061) and trypsin (V5111) were purchased from New-England Biolabs, Promega and Wako respectively.

All solutions were prepared in ultra-pure water.

## II. MATRISOME ENRICHMENT PROTOCOLS

For sample preparation, five different protocols have been used: skin lysis, quantitative detergent-solubility profiling (QDSP), enzymatic-based decellularization, chemical decellularization and compartmental enrichment. All of them were used for the comparison of enrichment protocols. They were conducted as described in the Materials and Methods section of the *J Proteomics* publication presented in **Chapter III, I-2**.

Chemical decellularization and compartmental enrichment were also used for the experiments on WT and RDEB back skin samples. More details are given below for each of them. Two matrisome enrichment experiments were successively conducted with  $n = 3$  mice/genotype or  $n = 2$  mice/genotype. The schematic representation of tissue processing during chemical decellularization and compartmental enrichment is available in **Chapter III, I-2** (Figure 1 of the *J Proteomics* publication).

All the procedures presented in this section were conducted using protein LoBind tubes (Eppendorf).

### 1. Chemical decellularization

All solutions were prepared in ultra-pure water and sterilized by 0.22  $\mu\text{m}$  filtration prior to use. The following stock solutions were prepared the day preceding the experiment:

- NaCl stock solution: 1 M NaCl, 200 mM HEPES pH 7.4. Storage: 4 °C.
- SDS/SD stock solution: 0.6 % SDS (w/v), 1.5 % SD (w/v). Dissolution at 40 °C in a water bath and storage at room temperature. When precipitates were present before the preparation of working buffers, the stock solution was heated at 40 °C in a water bath to obtain a clear and homogeneous solution. Storage: room temperature.
- GuHCl stock solution: 8 M GuHCl, 100 mM sodium acetate. Storage: room temperature.
- EDTA stock solution: 250 mM EDTA. Dissolution at 40 °C in a water bath and storage at room temperature. When precipitates were present before the preparation of working buffers, the stock solution was heated at 40 °C in a water bath to obtain a clear and homogeneous solution. Storage: room temperature.
- Protease inhibitors (PI, EDTA-containing, **Chapter II, I-2**) stock solution: 7x solution prepared according to the recommendations of the manufacturer (1 tablet in 1.5 mL of ultra-pure water). Storage: - 20 °C.

Working solutions with the following compositions were prepared extemporaneously using the stock solutions:

- NaCl buffer: 0.5 M NaCl, 100 mM HEPES, 25 mM EDTA, 1x PI.

- SDS/SD: 0.3 % SDS (w/v), 0.75 % SD (w/v), 25 mM EDTA, 1x PI.
- Buffer Triton X-100: 1 % Triton X-100 (v/v), 25 mM EDTA, 1x PI.
- Buffer GuHCl: 4 M GuHCl, 50 mM sodium acetate, 25 mM EDTA, 1x PI.
- Buffer EDTA: 25 mM EDTA, 1x PI.

Incubations and rinsing steps were conducted at 23 °C, 600 rpm (Thermomixer, Thermo Fisher Scientific) and centrifugation steps for 10 min, 23 °C, 16,000 rpm. After each centrifugation step and before adding the next buffer, tissue fragments were detached from the bottom of the tube and carefully disrupted using micropipette tips.

About 35 mg of frozen back skin biopsy were weighed in a 2-mL tube, cut into small pieces with a scalpel and processed as follows:

- (1) Skin samples were immersed in 600  $\mu$ L of NaCl buffer, gently vortexed during 10 s to homogenize the solution and incubated for 2 h 30.
- (2) Samples were briefly centrifuged during 20 s using a microcentrifuge to pellet tissue fragments. The supernatant (NaCl fraction) was recovered and split into 2 fractions of 200  $\mu$ L (for MS/MS analysis, storage in a 2-mL tube) and about 400  $\mu$ L (for Western blot analyses).
- (3) Samples were rinsed with 700  $\mu$ L of EDTA buffer for 10 min, centrifuged and the supernatant was discarded. This step will be referred as the “rinsing step” for the rest of the protocol.
- (4) 600  $\mu$ L of SDS/SD buffer were added, samples were briefly vortexed for 10 s to resuspend tissue pieces in the buffer and incubated for 18 h.
- (5) Samples were centrifuged to pellet skin fragments. The supernatant (SDS/SD fraction) was recovered and split into 2 fractions as described above for the NaCl fraction.
- (6) A rinsing step was applied.
- (7) 600  $\mu$ L of Triton-X100 buffer were added, samples were briefly vortexed for 10 s to resuspend skin pieces in the buffer and incubated for 1 h.
- (8) Samples were centrifuged to pellet skin fragments. The supernatant was recovered in a single tube (Triton X-100 fraction, no splitting).
- (9) Two rinsing steps of 1 h 30 were next applied.
- (10) Samples were finally incubated in 600  $\mu$ L of GuHCl buffer for 72 h.
- (11) A final centrifugation step was performed for 20 min. The supernatant (GuHCl fraction) was recovered and split into 2 fractions as previously described.
- (12) Insoluble fragments of pelleted tissues obtained after centrifugation were also recovered (Insoluble fraction, abbreviated Ins.).
- (13) All fractions (liquid and pellet) were flash-frozen in liquid nitrogen immediately after recovery and stored at – 80 °C until use.

## 2. Compartmental enrichment

Compartmental enrichment was performed with a commercial CNMCS kit (2145, Merck). All buffers (C, N, M, CS, W and PBS) were prepared extemporaneously and kept on ice at least 30 min before use, except for Buffer CS kept at room temperature. Protease inhibitors from the kit were added to C, N, M and CS buffers to reach a final concentration of 1x, according to the instructions of the manufacturer. Buffer W and PBS were supplemented with 1x PI (EDTA-containing) prepared as described above in the Chemical decellularization section. Tissue lysis was performed using a Tissue Lyser II (Qiagen). Beads adapted to protein extraction were cleaned with detergent, extensively washed with tap water, ultra-pure water and finally disinfected with 70 % ethanol (v/v). Beads were stored in 5-mL Eppendorf tubes on ice until use.

All steps were conducted using protein Lobind tubes (Eppendorf). All incubation and rinsing steps were conducted using a rotational shaker at speed 3 and centrifugations were performed at 16,000 g. After each centrifugation step and before adding the next buffer, the tissue pellet was detached from the bottom of the tube and carefully disrupted using micropipette tips. Samples were never vortexed to avoid buffer loss on tube walls.

About 35 mg of frozen back skin biopsy were weighed in a 1.5-mL tube in the form of 1 or 2 big pieces and processed as follows:

- (1) 2 beads were added per skin sample together with 420  $\mu$ L of buffer C. 2 cycles (2 min, 4 °C, 25 Hz) were successively applied in the Tissue Lyser II to homogenize the tissue.
- (2) Beads were removed using pliers and micropipette tips and samples were incubated in buffer C for 30 min, 4 °C.
- (3) Samples were then centrifuged for 30 min, 4°C and buffer C was recovered.
- (4) Samples were successively incubated in buffers W, N and M as follows, at 4 °C (incubation and centrifugation):
  - i. 330  $\mu$ L of buffer W for 20 min, and centrifugation for 20 min,
  - ii. 125  $\mu$ L of buffer N for 30 min, and centrifugation for 30 min,
  - iii. Rinsing as described above in step i.,
  - iv. 90  $\mu$ L of buffer M for 30 min, and centrifugation for 30 min,Buffers N et M were recovered after centrifugation while buffer W was discarded.
- (5) Samples were next incubated in 120  $\mu$ L of buffer CS for 30 min, room temperature (20 to 23 °C)
- (6) A centrifugation step was applied for 30 min, room temperature and buffer CS was recovered.
- (7) Insoluble pellets were finally rinsed with 350  $\mu$ L of PBS for 5 min, 4 °C, then centrifuged for 5 min, 4 °C and the supernatant was eliminated. All these steps were performed 3 times.

- (8) The centrifugation following the last rinsing with PBS was extended to 10 min to completely eliminate liquid. The final insoluble pellet was recovered (Compartmental Insoluble fraction, abbreviated Comp. Ins.)
- (9) All the fractions (liquid and pellet) were flash-frozen immediately after recovery and stored at - 80 °C until use.

### III. MS-BASED DISCOVERY PROTEOMICS

All the procedures presented in this section were conducted using protein LoBind tubes (Eppendorf).

#### 1. Mass spectrometry for the comparison of enrichment protocols

Sample preparation, HPLC-MS/MS analysis and protein identification were performed as described in the Materials and Methods section of the paper presented *J Proteomics* publication presented in **Chapter III, I-2**.

#### 2. Quantitative MS-based proteomics

##### i. Sample preparation

For liquid fractions, 200  $\mu\text{L}$  of sample were reduced with 10 mM DTT for 40 min, 60  $^{\circ}\text{C}$  and alkylated with 25 mM iodoacetamide for 40 min, 23  $^{\circ}\text{C}$ , in the dark. Proteins were next precipitated at room temperature using the methanol/chloroform method. 800  $\mu\text{L}$  of methanol, 200  $\mu\text{L}$  of chloroform and 600  $\mu\text{L}$  of ultra-pure water were successively added and samples were strongly vortexed after each addition. The mixture was centrifuged for 10 min, room temperature, 16,000 g. The upper liquid phase was carefully discarded and 800  $\mu\text{L}$  of methanol were added to wash the proteins. Samples were strongly vortexed, centrifuged for 10 min, room temperature, 16,000 g and methanol was discarded. Protein pellets were vacuum-dried in a vacuum centrifuge for 15 min at room temperature and solubilized in 45  $\mu\text{L}$  of 100 mM NaOH. Protein concentration was determined (2 dilutions analyzed: 1/40 and 1/60) using a commercial BCA kit (23225, Thermo Fisher Scientific). 20  $\mu\text{g}$  of proteins were further processed for HPLC-MS/MS analysis. 1.8 M HEPES pH 8.0 was added to adjust the pH to 8.0. Proteins were deglycosylated using PNGaseF (10 U) for 2 h and digested into peptides by successive treatments with Lys-C (1:150 enzyme to protein mass ratio) for 2 h and trypsin (1:50) overnight. The following day, trypsin was added again (1:100) and the incubation continued for 2 h. All enzyme treatments were performed at 37  $^{\circ}\text{C}$ , 700 rpm. Samples were centrifuged for 10 min, room temperature, 16,000 g to pellet tissue debris and the supernatants were recovered for TMT-labeling.

For insoluble pellets, 5 mg were dissolved in 120  $\mu\text{L}$  of 100 mM NaOH, strongly vortexed and incubated for 2 h, 23  $^{\circ}\text{C}$ , 600 rpm (Thermomixer). Proteins were reduced, alkylated, diluted four-fold in 100 mM ammonium bicarbonate and digested as described for liquid samples. Amounts of enzymes were increased as follows to allow the complete digestion of proteins: 50 U PNGaseF, 2  $\mu\text{g}$  Lys-C and 2  $\mu\text{g}$  or 1  $\mu\text{g}$  trypsin for the overnight incubation or the 2 hour treatment, respectively. Samples were acidified to pH 2.0 using 50 % TFA in ultra-pure water and desalted on C18 spin columns (89870, Thermo Fisher Scientific) according to the recommendations of the manufacturer. Peptides were eluted in 60  $\mu\text{L}$  of elution solvent (30/70  $\text{H}_2\text{O}$ /acetonitrile (ACN)); 50  $\mu\text{L}$  were vacuum-dried and the remaining 10  $\mu\text{L}$  were used to determine peptide concentration (tested dilutions: 1/4 and 1/10) using the Pierce<sup>TM</sup>

quantitative fluorimetric peptide assay (23290, Thermo Fisher Scientific). Dried peptides were then resuspended in 15  $\mu$ L of NaOH 100 mM. For TMT labeling, 20  $\mu$ g were taken and the pH was adjusted to 8.0 by addition of HEPES 1.8 M, pH 8.0 as described above for proteins from soluble fractions. For chemical decellularization, the peptides from the insoluble pellet were pooled with the peptides from the GuHCl extract in equal amounts prior to labeling (GuHCl + Ins. fraction generated).

### ii. TMT labeling of peptides

TMT reagents were equilibrated for 15 min at room temperature and dissolved in DMSO. The volume of added DMSO was equal to the volume of samples to be labeled in order to reach 50 % DMSO in the reaction medium. For each sample, the corresponding TMT reagent was added to peptides (0.4 mg TMT / 20  $\mu$ g peptides), the mixture was briefly vortexed and incubated for exactly 1 h ( $t_0$  = addition of the TMT reagent), 23 °C, 300 rpm (Thermomixer). The labeling reaction was stopped by the addition of 50 % hydroxylamine to a final concentration of 5 %. Samples were briefly vortexed and incubated for 30 min, 23 °C, 300 rpm. For each type of fraction (NaCl, SDS/SD, GuHCl + Ins., Comp. Ins.), equal amounts of each sample (4 to 5  $\mu$ g) were pooled together. Pools were acidified to pH 2 by addition of TFA 50 % in ultra-pure water, desalted on C18 spin columns and vacuum-dried at room temperature. Dried samples were stored at – 20 °C prior to HPLC-MS/MS analysis.

### iii. HPLC-MS/MS analyses

HPLC-MS/MS analyses were performed by Protein Science Facility (PSF), SFR BioSciences (CNRS UMS3444, Inserm US8, UCBL, ENS de Lyon, Lyon, France). Samples were analyzed in triplicates using an Ultimate 3000 nano-RSLC (Thermo Fisher Scientific) coupled on line with a Q Exactive™ HF mass spectrometer *via* a nano-electrospray ionization source (Thermo Fisher Scientific). Signals were acquired in data-dependent acquisition (DDA).

Peptides were diluted in 20  $\mu$ L of 0.1 % formic acid (FA) in water. 1  $\mu$ L was injected and loaded on a C18 Acclaim PepMap100 trap-column 75  $\mu$ m ID x 2 cm, 3  $\mu$ m, 100 Å, (Thermo Fisher Scientific) for 3 min at 5  $\mu$ L/min with 2 % ACN, 0.05 % TFA in water. Peptides were then separated on a C18 Acclaim Pepmap100 nano-column, 50 cm x 75  $\mu$ m i.d, 2  $\mu$ m, 100 Å (Thermo Fisher Scientific) with a 60 min linear gradient from 4 % to 50 % buffer B in buffer A (A: 0.1% FA in water, B: 80 % ACN, 0.1 % FA in water). The proportion of buffer B was increased from 50 to 95 % in 2 min, hold for 10 min, returned to the initial conditions in 1 min and then kept constant for 14 min. The total duration was set to 90 min at a flow rate of 300 nL/min. The oven temperature was kept constant at 40 °C. Fractions were analyzed with the TOP15 HCD method: MS data were acquired in a data dependent strategy selecting the fragmentation events based on the 15 most abundant precursor ions in the survey scan (375-1800 Th). The resolution of the survey scan was 120,000 at  $m/z$  200 Th and the resolution of the MS/MS scan was to 45,000 at  $m/z$  200 Th. The Ion Target Value for the survey scans in the Orbitrap and the MS/MS scan were set to 3E6 and 1E5 respectively and the maximum injection time was set to 50 ms for MS

scan and 100 ms for MS/MS scan. Parameters for acquiring HCD MS/MS spectra were as follows: collision energy = 32, isolation width = 1.2 m/z. The precursors with unknown charge state, charge state of 1 and 8 or greater than 8 were excluded. Peptides selected for MS/MS acquisition were then placed on an exclusion list for 30 s using the dynamic exclusion mode to limit duplicate spectra.

#### iv. MS/MS data processing

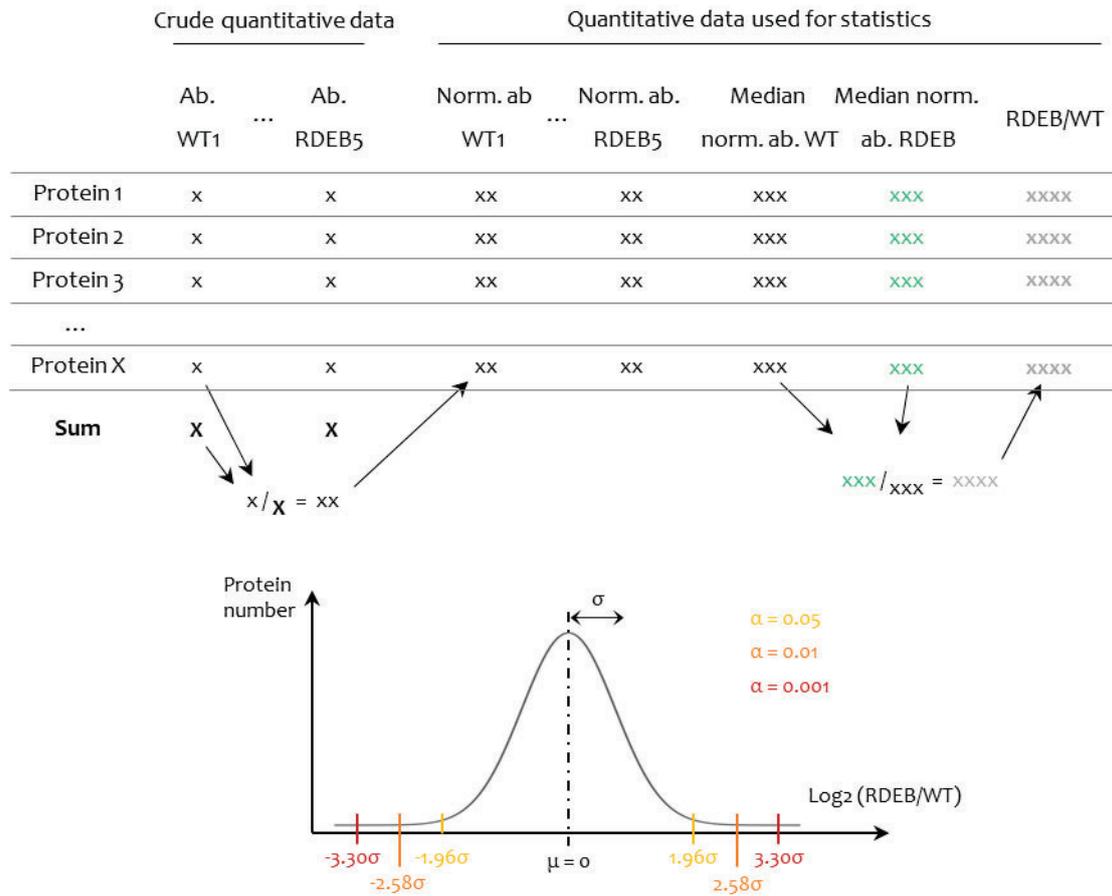
Proteins were identified through database searching using SequestHT with Proteome Discoverer 2.4 software (Thermo Fisher Scientific) against the UniProt mouse database (release: December 2020, 101 412 entries). Trypsin was set as digestion enzyme. Precursor mass tolerance was set to 10 ppm and fragment mass tolerance was set to 0.02 Da; up to 2 missed cleavages were allowed. Minimum peptide length was set to 6 amino acids. Oxidation (M, P), acetylation (protein N-terminus) and TMT labeled peptides in primary amino groups (+229.163 Da, K and peptide N-terminus) were set as variable modification, and carbamidomethylation (C) as fixed modification. Peptides and proteins were filtered with a false discovery rate (FDR) of 1 % using Percolator and proteins were identified with at least 1 unique peptide in rank 1. Protein quantitation was performed with reporter ion quantifier node in Proteome Discoverer 2.4 software with an integration tolerance of 20 ppm. Peptide and protein quantitations were based on summed abundance ratios. The purity correction factor of TMT labels was applied according to the manufacturer's instructions.

To identify matrisome proteins, the IDs of the genes encoding all the proteins identified by MS/MS were searched against matrisome lists. To do so, the mouse matrisome database<sup>91</sup> (last update: August 2014) and the tools of the on-line platform ProteoRE (<http://www.proteore.org/>)<sup>180</sup> were used.

#### v. Analysis of quantitative data

Quantitative data from two independent HPLC-MS/MS experiments (n = 2 mice/genotype and n = 3 mice/genotype) were pooled to reach 5 biological replicates per genotype. Only proteins quantified in the 10 analyzed animals (5 WT and 5 RDEB) were kept to run statistics. To do so, the data from each fraction type (NaCl, SDS/SD, GuHCl + Ins., Comp. Ins.) were processed as follows (**Figure 22**):

- (1) For each sample (WT1 to WT5 and RDEB1 to RDEB5), the total protein abundance was computed by summing the abundances of all the proteins detected in the sample.
- (2) This total abundance was used to normalize the abundance of each protein in the sample.
- (3) To run statistics, the median normalized abundance in WT and in RDEB samples was used to compute the RDEB/WT (RDEB/WT) ratio and then the Log<sub>2</sub>(RDEB/WT) value for each protein.
- (4) All Log<sub>2</sub>(RDEB/WT) ratios were used to build the distribution graph and the parameters of the corresponding normal distribution were used for the determination of the cutoffs of the experiment ( $\mu \pm 1.96\sigma$ ,  $2.58\sigma$  and  $3.30\sigma$  corresponding to p-values of 0.05, 0.01 and 0.001 respectively).



**Figure 22. Schematic view of the statistical approach applied to the MS/MS quantitative data from the two independent experiments (n = 5 mice/genotype in total).**

The individual abundance of each protein was normalized by the total protein abundance in the sample and the Log<sub>2</sub>(RDEB/WT) distribution was plotted. Ab.: abundance, norm. Ab.: normalized abundance.

Heatmaps were generated with Microsoft Excel (text hidden and conditional shaping depending on the Log<sub>2</sub> (RDEB/WT) value), from deep blue for the lowest value to deep red for the highest one; Microsoft Inc.). For the analysis of biological processes, the tools from the online platform Panther v17.0 (<http://www.pantherdb.org/>) were used. The gene IDs encoding all the significantly dysregulated proteins in RDEB skin were submitted to a statistical over-representation test, conducted using the “GO biological process complete” mode. The organism of reference used for the analysis was *Mus musculus*.

**vi. Data processing using SwissProt database (first matrisome analysis test, abundance ratios of collagen chains and peptide mapping)**

The identification of proteins using SwissProt database (reviewed sequences only) was applied for the first matrisome qualitative analysis (Chapter III, I-1). For quantitative analyses on WT and RDEB samples, SwissProt database was used for the study of the ratios between the different chains of collagens and for peptide mapping (to avoid the redundancy in proteoforms for a single gene, especially observed for collagens). Proteins were identified by database searching using SequestHT with Proteome

Discoverer 2.4 software (Thermo Fisher Scientific) against the SwissProt mouse database (release: May 2020, 17418 entries, for the first matrisome analyses, and August 2021, 17015 entries, for the quantitative analyses on WT and RDEB samples). For quantitative analyses on WT and RDEB samples, the quantification of peptides and proteins was performed as described above for quantification using the UniProt database (**Chapter II, III-2-iv**). The heatmap of collagen chain abundances was constructed as described above (**Chapter II, III-2-v**). For peptide mapping, data from the triplicate experiment (n = 3 mice/genotype) and RDEB/WT ratios computed by Proteome Discoverer 2.4 were used. To compare the abundance ratio of peptides from the propeptides and peptides from the triple helix, Mann-Whitney tests were applied (at least 2 peptides from propeptides detected).

## IV. TRANSMISSION ELECTRON MICROSCOPY

Sample processing for transmission electron microscopy (TEM) and image acquisition were performed by the Center of quantitative imaging Lyon-Est (CIQLE, SFR Santé-Lyon Est, CNRS UAR3453, INSERM US7, UCBL, Lyon, France).

### 1. Sample preparation

Back skin biopsies from mice aged 7 days old, and back skin or forepaw skin from mice aged 9-12 weeks were processed for TEM. Small tissue pieces (around 1 mm<sup>3</sup> in size) were fixed in one volume of 4 % glutaraldehyde and one volume of 0.2 M sodium cacodylate at 4 °C. Samples were carefully rinsed three times at 4 °C and post-fixed with 2 % OsO<sub>4</sub> for 1 h, 4 °C, before being dehydrated in graded series of ethanol and transferred to propylene oxide. Impregnation was performed with Epon epoxy resin (Electron Microscopy Sciences). Inclusion was obtained by polymerization at 60 °C for 72 h. Ultrathin sections (approximately 70-nm thick) were cut on a UC7 ultramicrotome (Leica), mounted on 200 mesh copper grids coated with poly-L-lysine (Electron Microscopy Sciences), stabilized for 1 day, room temperature, and contrasted with uranyl acetate.

### 2. Image acquisition

Sections were examined with a 1400JEM (Jeol), 120kV transmission electron microscope equipped with an Orius 600 camera (Gatan) on wide field position and Digital Micrograph software (Gatan Inc.).

### 3. Hemidesmosome and collagen fibrils analyses, statistics

The length of hemidesmosomes (HD) and the proportion of the basement membrane covered with HD were analyzed with the tools available in Fiji v1.53 software<sup>196</sup>, both in the back skin of WT and RDEB mice aged 7 days, and in the back skin and forepaw skin of WT and RDEB mice aged 9-12 weeks. The number of biological replicates, processed images and analyzed hemidesmosomes are given in **Table 5**.

Age	n	Anatomic location	Number of images processed	Total number of WT HD analyzed	Total number of RDEB HD analyzed
7 day-old	2 WT 2 RDEB	Back	5 to 7	124	202
9-12 week-old	4 WT 4 RDEB	Back	3 to 5	181	330
9-12 week-old	4 WT 4 RDEB	Forepaw	4 to 6	211	153

**Table 5. Details of samples and processed images for hemidesmosome (HD) analyses in WT and RDEB mice.**

Collagen fibril diameters were analyzed with Fiji v1.53 software<sup>196</sup> in the back skin of WT and RDEB mice aged 7 days, and in the back skin and forepaw skin of WT and RDEB mice aged 9-12 weeks. The number of biological replicates, processed images and analyzed hemidesmosomes are given in **Table 6**. Fibril diameters were measured using a macro (**script available in Appendix 4**) as the diameters of the smallest inscribed circles in transverse fibril sections.

Age	n	Anatomic location	Number of processed images per mouse	Total number of analyzed fibrils per genotype
7 day-old	2 WT 2 RDEB	Back	5	1,288
9-12 week-old	4 WT 4 RDEB	Back	3	4,000
9-12 week-old	4 WT 4 RDEB	Forepaw	4 to 6	2,400

**Table 6. Details of samples and processed images for the measurements of the diameters of dermal collagen fibrils in WT and RDEB mice.**

Statistics on the various analyzed parameters (length of hemidesmosomes, proportion of the basement membrane covered with hemidesmosomes, collagen fibrils diameter) were conducted using RStudio software (Rstudio Inc.). For each of these 3 variables, WT and RDEB values were compared by non-parametric Mann Whitney test.

## V. OTHER PROTEIN ANALYSES

### 1. Immunoblotting

#### i. Sample preparation

##### *Skin lysates*

Back skin samples from control and RDEB mice aged 10-12 weeks were used to prepare lysates (samples from mice different from those used for MS-based quantitative proteomics). 25 mg of frozen skin were weighed in a 2-mL protein Lobind tube (Eppendorf) and rapidly cut into pieces with a scalpel. 200  $\mu$ L of lysis buffer (8.6 % SDS (w/v), 110 mM Tris pH 6.8, 43 % glycerol (w/v), 1x PI with EDTA) were added and samples were strongly vortexed for 10 s and boiled for 20 min, 95 °C. Samples were strongly vortexed again for 3 minutes and centrifuged for 20 min, room temperature, 16,000 g to pellet tissue debris. Supernatants were then transferred in a new protein Lobind tube, without taking the white fat “snowflakes” sometimes formed at the supernatant surface. A new centrifugation step was performed for 20 min, room temperature, 16,000 g, and supernatants were pooled with those previously recovered. Samples were aliquoted and stored at – 20 °C until use.

Prior to SDS-PAGE migration, protein concentration was measured (applied dilutions: 1/20 and 1/30) using a commercial BCA kit (23225, Thermo Fisher Scientific). Prior to SDS-PAGE, the volume of sample containing the required amount of proteins was diluted to 1/5 by addition of DTT to a final concentration 50 mM, bromophenol blue to a final concentration of 0.01 % (w/v) and water. Samples were vortexed, boiled for 5 min, 95 °C and loaded in gel wells for SDS-PAGE migration.

##### *Enrichment fractions*

NaCl and SDS/SD fractions were assayed using BCA (2 analyzed dilutions, ranging from 1/30 to 1/60). Prior to SDS-PAGE migration, samples were mixed with reducing buffer (final concentrations in the sample: 10 % SDS (w/v), 125 mM Tris pH 6.8, 12.5 % glycerol (w/v), 0.01 % bromophenol blue (w/v), 50 mM DTT), briefly vortexed and boiled for 5 min, 95 °C.

4 mg of the insoluble pellets obtained through compartmental enrichment (fraction Comp. Ins.) were weighed in a 1.5-mL protein Lobind tube (Eppendorf). 50  $\mu$ L of 100 mM NaOH, 250 mM DTT were added and samples were boiled for 15 min, 95 °C. Samples were briefly vortexed and centrifuged for 10 min, room temperature, 16,000 g. Supernatants were recovered and stored at – 20 °C until use. Prior to SDS-PAGE, 15  $\mu$ L of samples were mixed with sample buffer (final concentrations in the sample: 10 % SDS (w/v), 125 mM Tris pH 6.8, 12.5 % glycerol (w/v), 0.01 % bromophenol blue (w/v)), briefly vortexed and boiled for 5 min, 95 °C.

*Fibroblast lysates*

Lysates of human dermal fibroblasts were previously prepared by Agnès Tessier during her PhD as follows: cells were washed twice with PBS and proteins from the cell layer were extracted with RIPA buffer (50 mM Tris, 0.15 M NaCl, 1% Np-40, 1% SDS, 0.5% sodium deoxycholate, pH 8 with protease inhibitors) incubated at 4°C for 30 min under agitation<sup>197</sup>. Samples were stored at – 20 °C until use. Prior to SDS-PAGE migration, 25 µL of cell lysate were mixed with reducing buffer (final concentrations in the sample: 10 % SDS (w/v), 125 mM Tris pH 6.8, 12.5 % glycerol (w/v), 0.01 % bromophenol blue (w/v), 50 mM DTT), briefly vortexed and boiled for 5 min, 95 °C.

For all sample types, protein amounts/lane and acrylamide percentages are detailed below in **Table 6**.

Target protein	Skin lysate	NaCl, SDS fraction	Comp. Ins.	Fibroblast lysate	% SDS-PAGE gel
Collagen I (triple helix)	15 µg	15 µg	N/A	N/A	Home-made, 8 % (+ TCE)
Procollagen α1(I) C-propeptide	50 µg	50 µg	15 µL	N/A	4-20 % gradient gel, Stainfree (4568094, BioRad)
Procollagen α1(I) N-propeptide	50 µg	50 µg	15 µL	N/A	4-20 % gradient gel, Stainfree (4568094, BioRad)
HSP-47	30 µg	30 µg	N/A	25 µL	Home-made, 10 % (+ TCE)
PCPE-1	30 µg	30 µg	N/A	N/A	Home-made, 10 % (+ TCE)

**Table 7. Amounts of biological material and gel percentages used for immunoblotting in relation with target protein.**

TCE: 2,2,2-trichloroethanol, allowing Stainfree detection.

## ii. Western blotting

Proteins were first separated by SDS-PAGE in 25 mM Tris pH 8.4, 0.2 M Glycine, 0.1% SDS buffer. After electrophoresis, gels were exposed to ultraviolet rays (two activation cycles of 30 s each) using the Fusion FX camera system (Vilber Lourmat) to perform the Stainfree reaction (reaction between TCE and Tryptophan residues). Then, proteins were electroblotted for 2 h onto PVDF membranes (polyvinylidene difluoride, 831.1, Merck) in 10 mM CAPS pH 11.0, 10 % ethanol. After transfer, the Stainfree signal was detected using the Fusion FX camera system. The membranes were then blocked in 10 % skim milk (w/v, Regilait) diluted into PBS for 2 h and incubated overnight with primary antibodies diluted into 5 % skim milk in PBS with 0.05 % Tween-20 (v/v). Membranes were then washed three times with PBS, 0.05 % Tween-20 and incubated with the secondary antibody for 1 h, room temperature. Primary and secondary antibodies used for Western blotting are detailed in **Table 8**. All target proteins were detected by chemiluminescence with the Amersham™ ECL™ Select Western blotting detection reagent (RPN2235, Cytiva) using the Fusion FX camera system.

Antibody	Reference	Concentration for immunoblotting	Concentration for immunofluorescence
$\alpha$ - BMP-1	Home-made	N/A	1/100
$\alpha$ - collagen I (triple helix)	20151, Novotec	0.25 $\mu$ g/mL (1/8000)	N/A
$\alpha$ - procollagen $\alpha$ 1(I) C-propeptide	LF41 <sup>198</sup> , Gift from Larry Fisher	12.5 $\mu$ g/mL (1/200)	N/A
$\alpha$ - procollagen $\alpha$ 1(I) N-propeptide	LF39 <sup>198</sup> , Gift from Larry Fisher	12.5 $\mu$ g/mL (1/200)	N/A
$\alpha$ - HSP-47	10875-1-AP, Proteintech	0.67 $\mu$ g/mL (1/750)	N/A
$\alpha$ - PCPE-1	AF2239, R&D Systems	0.2 $\mu$ g/mL (1/1000)	2 $\mu$ g/mL (1/100)
$\alpha$ - goat IgG, HRP conjugated	DkxGt-003-DHRPX, Immuno Reagents	0.1 $\mu$ g/mL (1/10000)	N/A
$\alpha$ - rabbit IgG, HRP conjugated	7074, Cell Signaling	1/10000 (concentration of the stock solution unknown)	N/A
$\alpha$ - goat IgG, Alexa Fluor 647 conjugated	705-605-003, Jackson Immuno Research	N/A	3.75 $\mu$ g/mL (1/400)
$\alpha$ - rabbit IgG, Cyanine3 conjugated	711-165-152, Jackson Immuno Research	N/A	3.75 $\mu$ g/mL (1/400)

**Table 8. Antibodies used for immunoblotting and immunofluorescence.**

### iii. Quantification and statistics

For each target, immunoblotting was performed in three technical replicates (3 membranes). The intensity of the chemiluminescent and Stainfree signals were quantified using ImageQuant TL software (Cytiva Inc.). For each membrane, a double normalization was conducted as follows:

- (1) The Stainfree signal was first used to normalize the chemiluminescent one in each lane (= each biological replicate), to obtain the “normalized signal” to total protein amount,
- (2) In order to compare the signals obtained on independent membranes, a second normalization to one randomly-chosen lane (usually corresponding to WT1) was performed. A “doubly-normalized” signal was thus obtained for each biological replicate (with a value of 1 for the reference, i.e. WT1).
- (3) Mean signal intensities were computed for WT and RDEB genotypes using “doubly-normalized” signals. Normality and homoscedasticity were checked using Shapiro-Wilk and Fisher tests respectively. WT and RDEB were compared by Student t-test when normality and homoscedasticity were respected, or Mann-Whitney test when the t-test was not applicable.

In addition, the three technical replicates were pooled to run an additional statistical test. To do so, for each biological replicate, the mean signal intensity was computed from the 3 doubly-normalized values of the 3 technical replicates. The mean WT and RDEB signal intensities computed in step (3) for each membrane were analyzed as three replicates. All statistics were run on RStudio software (Rstudio Inc.)

## 2. Immunofluorescence on tissue sections

Paraffin blocks of back skin biopsies (fixed in formalin, dehydrated and embedded in paraffin) were obtained from our collaborator in Freiburg and sliced into 5 µm-thick sections. The latter were deparaffinized through three successive 100 % methycyclohexane baths (5 min each), rehydrated into ethanol solutions of decreasing concentrations (100 %, 100 %, 95 %, 70 %, successive 5 min baths) and finally immersed in tap water for 15 min. Antigens were unmasked using 0.05 % pronase (10165921001, Roche) in PBS (45 min, room temperature) for BMP-1 or 0.1 % trypsin (T0303, Merck), 0.1 % CaCl<sub>2</sub> in PBS (1 h, 37 °C) for PCPE-1. Sections were washed three times with PBS (5 min, room temperature) and blocked in 4 % bovine serum albumin (A3059, Sigma) in PBS for 1 h, room temperature. Samples were stained overnight, 4 °C with primary antibody prepared in the blocking buffer, rinsed three times in PBS as described above and incubated in the dark for 1 h, room temperature, with a solution of secondary antibody prepared in the blocking solution (**Table 9**) and supplemented with DAPI (1 µg/µL). Sections were finally rinsed three times with PBS in the dark, mounted in Fluorescent mounting medium (F6182, Sigma) and dried for at least 4 h at room temperature, in the dark, prior to observations. Analyses were conducted with an Eclipse Ti-E fluorescence microscope (Nikon) and pictures were processed with Fiji v1.53 software<sup>196</sup>.

## VI. RNA ANALYSES

### 1. RNA extraction

RNA was extracted from frozen back skin biopsies using RNeasy Micro Kit (74004, Qiagen). Briefly, 4 mg of tissue were weighed in a 1.5-mL Eppendorf tube and disrupted with Tissue Lyser II (Qiagen). Beads adapted to RNA extraction were cleaned with detergent, abundantly rinsed with tap water, ultra-pure water and disinfected with 70 % ethanol (v/v). RNA extraction was performed following the recommendations of the manufacturer (protocol “Purification of total RNA from fibrous tissues”). Two steps were adapted: treatment with proteinase K was extended to 15 min, 55 °C (instead of 10 min, 55 °) and RNA was eluted in 20 µL of RNase-free water (instead of 14 µL). Immediately after recovery, RNA samples were flash-frozen in liquid nitrogen and stored at – 80 °C until analysis by RT-qPCR.

### 2. Reverse transcription

RNA concentration was determined using a Nanodrop instrument (Thermo Fisher Scientific). 200 ng of RNA were next reverse-transcribed (RT) into cDNA using the PrimeScript™ RT kit (RR037A, TakaraBio) as described in **Table 9**, and a MJ Mini PCR instrument (Biorad). The reaction was conducted for 15 min, 37 °C and followed by a 5 s step at 85 °C to degrade remaining RNA and inactivate reverse transcriptase. cDNA was finally diluted to 1/5 prior to qPCR analysis.

Reagent	Volume
5 X PrimeScript Buffer	4 µL
Oligo dT primers	1 µL
Random 6-mers	1 µL
PrimeScript RT Enzyme Mix	1 µL
RNase free water	Adjustment to 20 µL
<b>Total volume</b>	<b>20 µL</b>

**Table 9. Composition of the mix used for the reverse transcription of 200 ng of RNA.**

Reagents from the PrimeScript™ RT kit (TakaraBio) were used to reverse transcribe RNA into cDNA prior to qPCR.

### 3. Quantitative real time PCR

Fluorescence-based quantitative real-time PCR (qPCR) of cDNA was performed using the FastStart Universal SYBR Green Master mix (4913850001, Roche) and a Rotor-Gene Q instrument (Invitrogen). All primers were obtained from Sigma (purity level: dessaler, dissolved in water, 100 µM, **Table 10**). The composition of the reaction mix is detailed in **Table 11**. The qPCR conditions included an initial denaturation (10 min, 95 °C) followed by 50 amplification cycles, each comprising denaturation (10 s, 95 °C) and annealing (20 s, 60 °C) steps. Fluorescence was measured at the end of the annealing step of

each cycle and the quantification cycle (Ct) was determined for each gene. At the end of amplification, a melting curve was recorded by heating slowly at 0.5 °C/sec from 70 °C until 94 °C.

Gene	Primers (5' to 3')	Efficiency
<i>Bmp1</i>	Forward: TGCCTGTGGAGTTCAAGTCT Reverse: CCGAGGTGAGCTTCCTGAGT	0.975
<i>Colla1</i>	Forward: AACGAGATCGAGCTCAGAGG Reverse: GACTGTCTTGCCCCAAGTTC	0.925
<i>Colla2</i>	Forward: CACCCAGCGAAGAACTCAT Reverse: TCTCCTCATCCAGGTACGCA	1.000
<i>Col3a1</i>	Forward: TGACTGTCCCACGTAAGCAC Reverse: GAGGGCCATAGCTGAACTGAA	0.955
<i>Pcolce</i>	Forward: CTTCGGGCACTGAGCACCA Reverse: ACCTGGTTTGAGGGTGCAAT	0.952
<i>Rpl13a</i>	Forward: ATCCCTCCACCCTATGACAA Reverse: GCCCAGGTAAGCAAACCTT	0.964
<i>Serpinh1</i>	Forward: CCAAGGGGGAACGGATCG Reverse: AAGAGGCATAAGGTGCCAG	1.000
<i>Tll1</i>	Forward: GGCTGGAGTTCTTACATCTACG Reverse: CTTATCTCCCCTCCACAAATCG	0.823

**Table 10. Sequences and efficiency of primer pairs used for RT q-PCR.**

Reagent	Volume
cDNA diluted to the 1/5	4 µL
FastStart Universal SYBR Green Master	10 µL
Forward primer	1 µL
Reverse primer	1 µL
RNase-free water	4 µL
<b>Total volume</b>	<b>20 µL</b>

**Table 11. Composition of the mix used for qPCR.**

FastStart Universal SYBR Green Master mix (Roche) was used to run qPCR on WT and RDEB cDNA.

#### 4. Computing of gene expression and statistics

All experiments were performed in three technical replicates. *Rpl13a* was used as reference gene (RG) and the sample WT2 was chosen as reference sample. For each gene of interest (GOI), the gene expression (GE) was computed as follows (Ef = efficiency of primer pair):

$$\Delta Ct_{GOI \text{ sample } X} = Ct_{GOI \text{ sample } X} - Ct_{RG \text{ sample } X}$$

$$\Delta\Delta Ct_{GOI \text{ sample } X} = \Delta Ct_{GOI \text{ sample } X} - \Delta Ct_{GOI \text{ sample } WT2}$$

$$GE_{GOI \text{ sample } X} = (1+Ef)^{-\Delta\Delta Ct_{GOI \text{ sample } X}}$$

For each sample, mean GE was computed from the three values obtained from the three technical replicates. Mean WT and RDEB GEs were finally compared by non-parametric Mann Whitney test using RStudio software (Rstudio Inc).

## VII. MS-BASED TARGETED PROTEOMICS (PARALLEL REACTION MONITORING)

All the procedures presented in this section were conducted using protein LoBind tubes (Eppendorf).

### 1. Sample preparation

#### i. Synthetic peptides

Synthetic peptides (Thermo Fisher Scientific) were ordered in liquid form: DDANVVR and SGEYWIDPNQGCVNLDAIK were received dissolved in DMSO (PEPotech Grade 2 quality) while ADDANVVR was received in 5 % ACN in ultra-pure water (AQUA QuantProLight quality). 100 ng of each peptide were pooled together, vacuum-dried and stored at – 20 °C until preparation for PRM analysis (non-labeled synthetic peptides).

To generate TMT-labeled synthetic peptides, 100 ng of ADDANVVR and DDANVVR were pooled together, vacuum-dried and labeled with TMT as previously described (**Chapter II, III-2-ii**). Briefly, peptides were dissolved in 100 mM NaOH and 1.8 M HEPES pH 8.0, labeled with TMT reagents and the reaction was stopped with hydroxylamine. Peptides were desalted on C18 spin columns, vacuum-dried and stored at – 20 °C until analysis.

#### ii. NaCl fractions

15 µg of proteins from WT NaCl fractions were reduced and alkylated as described above. Samples were diluted to 1/5 by addition of ultra-pure water. The proteins were next purified using the SP3 protocol (“single-pot, solid-phase-enhanced sample preparation”<sup>199</sup>, sera-mag magnetic beads used: 65152105050250 and 45152105050250, Cytiva). 75 µg of each type of beads were added to the samples (final protein/beads ratio: 1/10, w/w) and the mixture was homogenized by gentle pipetting. 100 % ethanol was added (1:1, v/v) to initiate the fixation of proteins onto the beads and the samples were incubated for 5 min, 24 °C, 300 rpm (Thermomixer). Tubes were deposited on a magnetic rack for 5 min at room temperature to immobilize beads and the supernatant was discarded. The proteins were rinsed with 500 µL of 80 % ethanol, the solution was homogenized by gentle pipetting and the supernatant was next eliminated after bead immobilization (see above). This rinsing step with 80 % ethanol was applied 3 times in total. Proteins were eluted from the beads in 20 µL of 100 mM HEPES, pH 8.0. Proteins were digested with trypsin (1:50 enzyme to protein mass ratio) overnight, 37 °C, 700 rpm (Thermomixer). Samples were divided into 2 fractions containing the same amount of material (5 µg of proteins). One fraction was acidified to pH 2.0 with 50 % TFA, desalted on C18 spin columns, vacuum-dried and stored at – 20 °C until PRM analysis (non-labeled NaCl fraction). The other fraction was first labeled with TMT as previously described (**Chapter II, III-2-ii**). Peptides were also acidified

to pH 2.0 with 50 % TFA, desalted on C18 spin columns, vacuum-dried and then treated as above (TMT-labeled NaCl fraction).

### iii. Skin lysates

15 µg of proteins from WT skin lysates prepared for immunoblotting (**Chapter II, V-1-i**) were treated as described above for the NaCl fraction (from reduction to sample storage at – 20 °C). Two fractions of protein from skin lysates, one labeled with TMT, the other unlabeled, were finally generated for PRM.

## 2. PRM analysis

Prior to PRM scan, libraries were generated to analyze non-labeled and TMT-labeled samples by DDA scans of non-labeled and TMT-labeled synthetic peptides respectively. These libraries therefore contain the MS/MS fragmentation profile for each of the analyzed peptides (list of transitions and their relative abundance). These reference signals were used to compare the signal acquired by PRM and to evaluate the quality of peptide detection in biological samples. To generate libraries, synthetic peptides were analyzed by DDA applying the same HPLC-MS/MS parameters as those described for the analyses of WT samples during ECM-enrichment protocol comparison (**Chapter II, V-1-iii**). To run PRM analyses, the Q Exactive HF mass spectrometer (Thermo Fisher Scientific) was used in PRM mode with the following parameters: precursors were targeted in a 1.0 m/z isolation window around the m/z of interest, and fragmented in HCD mode with an NCE energy of 27. MS/MS was performed at 60,000 resolution (at 200 m/z), AGC target was set to 10,000 and maximum injection time to 120 ms. Loop count and MSX count were respectively set to 4 and 1.

## 3. Data analysis

PRM data were processed with Skyline v19.1 software (open-source, developed by the Mac Coss Lab, University of Washington, USA). Skyline settings used to analyze the data are detailed in Appendices 5 and 6.

---

## CHAPTER III: RESULTS

---

## I. DEVELOPMENT OF PROTOCOLS FOR THE ANALYSIS OF MOUSE SKIN MATRISOME

All previous studies reporting MS-based matrisome analyses were conducted on highly cellularized tissues (tumors, lung, endocrine and exocrine glands). Skin also displays a high cell density in the epidermis but this is not the case for the dermis which rather contains sparse cells scattered in a very abundant ECM. Furthermore, the dermal matrix is mainly composed of collagen and elastic fibers; it contains as a consequence both soluble (ECM protein precursors, molecules involved in protein processing such as enzymes and their regulators) and insoluble (mature core matrisome components) proteins. Thus, the first challenge of this project was the recovery of all these skin matrisome components exhibiting different solubility profiles while eliminating as many cellular proteins as possible, especially the epidermal cell proteins that may disturb MS/MS analyses because of their high abundance.

### 1. Preliminary ECM enrichment tests

The most-widely used ECM enrichment protocol is the “compartmental” strategy proposed by Naba *et al*<sup>192</sup> a few years after the first definition of the matrisome. It consists in the processing of the native tissue with different solutions from the CNMCS commercial kit (Merck, detailed composition unknown) to generate a final insoluble pellet, supposedly containing matrix proteins. Compartmental enrichment was thus the first strategy that we tested to recover the ECM from mouse back skin. Because DEB is primarily caused by a deficiency in the DEJ, where the expression of many proteins is altered due to collagen VII deficiency (**Chapter I, III-1**), it seemed fundamental to include basement membrane components to the global characterization of the skin matrisome of diseased mice. For this reason, we decided to enrich the matrix from the whole native skin instead of focusing on the matrix-rich dermis only. No dermis/epidermis separation was therefore performed in this study; only a soft manual elimination of the subcutaneous fat was applied to limit the lipid content of the samples.

The first test of the protocol led to the identification of 357 proteins by HPLC-MS/MS in the insoluble pellet (no quantification). The identity of proteins was obtained through the search of the SwissProt database during data processing. Matrix proteins were identified by comparing the IDs of the genes encoding these 357 proteins with the matrisome genes obtained from the mouse matrisome database<sup>91</sup>. Sixty-one matrisome components were detected, including 43 core and 18 associated proteins (**Table 12**). This number was surprisingly low compared to those obtained with other tissues, for which the ECM content was in addition lower, with the same protocol: 113<sup>100</sup> and 185<sup>98</sup> matrisome components were respectively detected in mouse lung and pancreatic islets for instance.

Furthermore, a lot of proteins already reported in the extracellular matrix of the skin, including insoluble ones, were missing in our data.

CORE MATRISOME				ASSOCIATED MATRISOME			
Gene	Protein name	Urea	NaOH	Gene	Protein name	Urea	NaOH
<b>Collagen</b>				<b>ECM-affiliated</b>			
<i>Coll1a1</i>	Collagen alpha-1 (I)	X	X	<i>Anxa1</i>	Annexin A1	X	X
<i>Coll1a2</i>	Collagen alpha-2 (I)	X	X	<i>Anxa2</i>	Annexin A2	X	X
<i>Col2a1</i>	Collagen alpha-1 (II)		X	<i>Anxa3</i>	Annexin A3		X
<i>Col3a1</i>	Collagen alpha-1 (III)	X	X	<i>Anxa4</i>	Annexin A4	X	X
<i>Col4a1</i>	Collagen alpha-1 (IV)	X	X	<i>Anxa5</i>	Annexin A5		X
<i>Col4a2</i>	Collagen alpha-2 (IV)	X	X	<i>Anxa6</i>	Annexin A6	X	X
<i>Col5a1</i>	Collagen alpha-1 (V)	X	X	<i>Anxa7</i>	Annexin A7	X	X
<i>Col5a2</i>	Collagen alpha-2 (V)	X	X	<i>Anxa8</i>	Annexin A8	X	X
<i>Col6a1</i>	Collagen alpha-1 (VI)	X	X	<i>Anxa11</i>	Annexin A11		X
<i>Col6a2</i>	Collagen alpha-2 (VI)	X	X	<i>Lgals1</i>	Galectin-1	X	X
<i>Col6a6</i>	Collagen alpha-6 (VI)	X	X	<i>Lgals3</i>	Galectin-3		X
<i>Col7a1</i>	Collagen alpha-1 (VII)	X	X	<i>Lgals7</i>	Galectin-7	X	X
<i>Coll12a1</i>	Collagen alpha-1 (XII)	X	X	<b>ECM-regulator</b>			
<i>Coll14a1</i>	Collagen alpha-1 (XIV)	X	X	<i>C1qtnf3</i>	Complement C1q tumor necrosis factor-related protein		X
<i>Coll15a1</i>	Collagen alpha-1 (XV)		X	<i>Cstb</i>	Cystatin B		X
<b>Glycoprotein</b>				<i>Ctsk</i>	Cathepsin K		X
<i>Aebp1</i>	Adipocyte enhancer-binding protein-1	X	X	<i>F12</i>	Coagulation factor XII	X	
<i>Coch</i>	Cochlin		X	<i>F13a</i>	Coagulation factor XIII A		X
<i>Dpt</i>	Dermatopontin	X	X	<i>Hrg</i>	Histidine-rich glycoprotein		X
<i>Ecm1</i>	Extracellular matrix protein-1		X	<i>Itih3</i>	Inter-alpha-trypsin inhibitor heavy chain H3		X
<i>Ecm2</i>	Extracellular matrix protein-2	X	X	<i>Kng1</i>	Kininogen-1		X
<i>Fbn1</i>	Fibrillin-1	X	X	<i>Serpina1a</i>	Serine protease inhibitor A1a	X	
<i>Fga</i>	Fibrinogen alpha chain	X	X	<i>Serpina1d</i>	Serine protease inhibitor A1d	X	X
<i>Fgb</i>	Fibrinogen beta chain	X	X	<i>Serpina3k</i>	Serine protease inhibitor A3k	X	X
<i>Fgg</i>	Fibrinogen gamma chain	X	X	<i>Serpina3m</i>	Serine protease inhibitor A3m		X
<i>Fn1</i>	Fibronectin	X	X	<i>Serpina1a</i>	Serine protease inhibitor B1a		X
<i>Fbln5</i>	Fibulin-5	X		<i>Serpina2</i>	Serine protease inhibitor B2	X	X
<i>Lama2</i>	Laminin subunit alpha-2	X	X	<i>Serpina5</i>	Serine protease inhibitor B5	X	X
<i>Lama4</i>	Laminin subunit alpha-4	X		<i>Serpina6a</i>	Serine protease inhibitor B6a	X	
<i>Lama5</i>	Laminin subunit alpha-5	X	X	<i>Serpina2</i>	Serine protease inhibitor F2		X
<i>Lamb1</i>	Laminin subunit beta-1	X	X	<i>Serpina1</i>	Serine protease inhibitor H1	X	X
<i>Lamb2</i>	Laminin subunit beta-2	X	X	<i>Tgm2</i>	Transglutaminase-2	X	X
<i>Lamb3</i>	Laminin subunit beta-3	X		<b>Secreted factor</b>			
<i>Lamc1</i>	Laminin subunit gamma-1	X	X	<i>Flg</i>	Filaggrin	X	
<i>Lamc2</i>	Laminin subunit gamma-2		X	<i>Flg2</i>	Filaggrin-2		X
<i>Mfap2</i>	Microfibrillar-associated protein-2	X	X	<i>S100a3</i>	Protein S100 A3		X
<i>Mfap4</i>	Microfibrillar-associated protein-4	X	X	<i>S100a4</i>	Protein S100 A4		X
<i>Mfap5</i>	Microfibrillar-associated protein-5	X	X	<i>S100a10</i>	Protein S100 A10		X
<i>Nid1</i>	Nidogen-1	X	X	<i>S100a11</i>	Protein S100 A11		X
<i>Nid2</i>	Nidogen-2	X	X	<i>S100a14</i>	Protein S100 A14		X
<i>Postn</i>	Periostin	X	X	<b>Total</b>		18	34
<i>Srpx</i>	Sushi-repeat-containing protein SRPX	X	X				
<i>Tgfbi</i>	TGF- $\beta$ -induced protein ig-h3 (TGFB1)	X	X				
<i>Vwa5a</i>	von Willebrand factor A domain-containing protein-5A		X				
<b>Proteoglycan</b>							
<i>Aspn</i>	Asporin	X	X				
<i>Bgn</i>	Biglycan	X	X				
<i>Dcn</i>	Decorin	X	X				
<i>Hspg2</i>	Perlecan	X	X				
<i>Lum</i>	Lumican	X	X				
<i>Ogn</i>	Mimcan	X	X				
<i>Prelp</i>	Prolargin	X	X				
<i>Prg2</i>	Bone marrow proteoglycan	X	X				
<b>Total</b>		43	48				

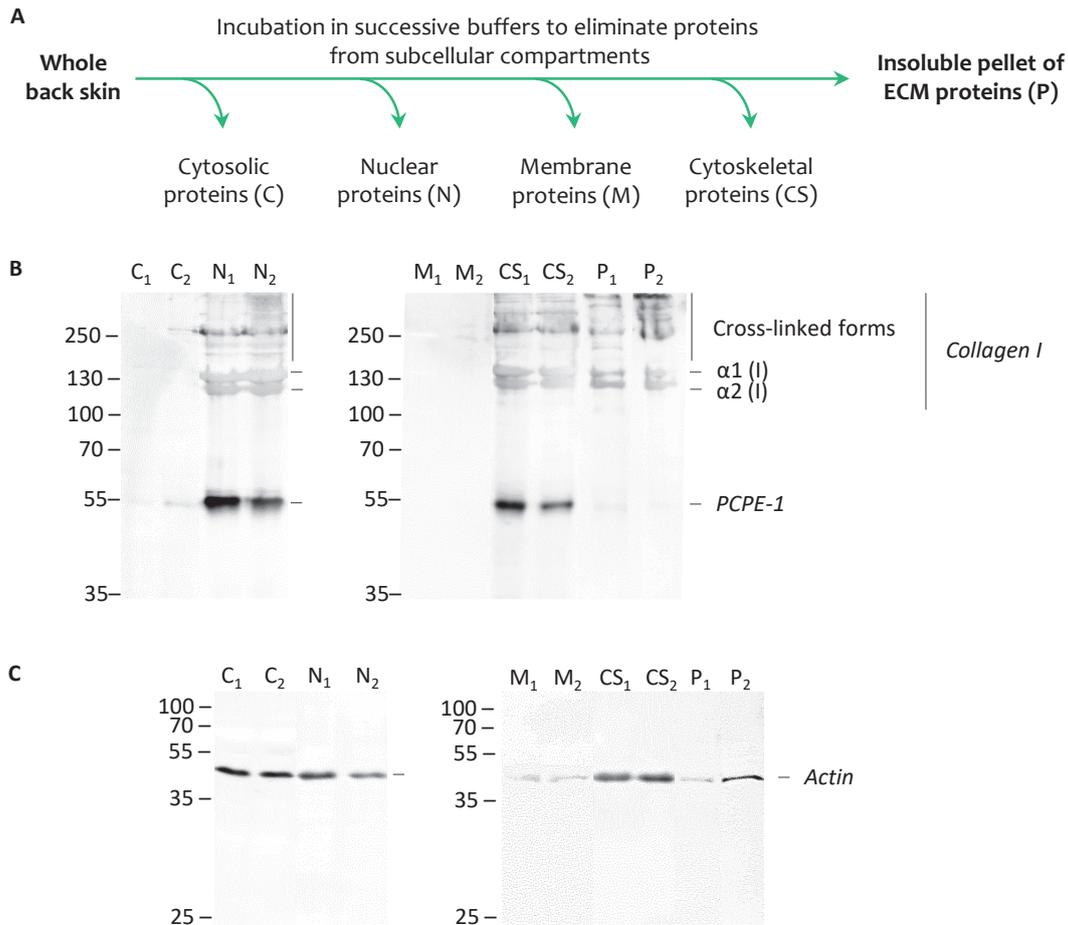
**Table 12. Matrisome proteins detected by MS/MS in the insoluble pellet after compartmental enrichment from mouse skin.**

Previous page. Two methods (urea 8 M and NaOH 0.1 M) were tested to dissolve the pellet prior to preparation for mass spectrometry. The best coverage of the matrisome, especially for associated proteins, was obtained with NaOH 0.1 M. Database used for protein identification: SwissProt.

Regarding the core matrisome, elastin, tenascin X which are found in the dermis and the  $\alpha 3$  and  $\gamma 2$  laminin chains of LM-332 in the basement membrane were surprisingly not detected. The low proportion of associated matrisome – about 30 % of the matrisome detected in this experiment – was also intriguing because it normally accounts for 75 % of the total mouse matrisome proteins according to matrisome database. If the absence of enzymes such as metzincins, transglutaminases or lysyl-oxidases was expected because their abundance is low in healthy skin, the non-detection of regulator proteins such as PCPE-1 and S100-proteins was very surprising because these components are relatively abundant. Also, because complete solubilization of the pellet was difficult to obtain with 8 M urea, we adapted the reference protocol and used 0.1 M NaOH instead. Pellet dissolution seemed improved by visual inspection of the sample and the total number of proteins detected by MS/MS was actually about 1.5 times higher than in the first experiment using urea (**Table 12**). The benefit in terms of matrisome analysis was however limited, with a slight increase in the number of identified proteins through the detection of more associated molecules, but some key matrix proteins mentioned above were still missing.

We next briefly evaluated the distribution of some representative matrix proteins of the dermis in the different buffers used during tissue processing (**Figure 23A**). The loss of PCPE-1, mainly extracted in the buffers dedicated to the removal of nucleus and cytoskeleton proteins but not found in the insoluble pellet, was observed. A distribution of collagen I in all soluble and insoluble fractions was observed (**Figure 23B**). In addition, actin was also detected in all fractions (**Figure 23C**). Regarding the components of the DEJ, a partial loss of collagen VI in the cytoskeleton buffer was also evidenced (not shown).

These first results were disappointing and clearly showed that compartmental enrichment was probably not optimized for the study of the ECM of the native skin: the final insoluble pellet obtained did not contain enough matrix proteins to allow a satisfactory study of the skin matrisome for our purpose.

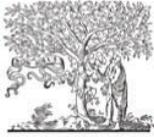


**Figure 23. Protein extraction during compartmental enrichment on mouse skin.**

Back skin samples from 2 mice (1 and 2) were processed with the CNMCS kit (2145, Merck) to enrich the ECM proteins. **A.** Simplified description of tissue processing steps and theoretical protein distribution expected with the compartmental enrichment protocol. **B.** Detection of the matrix proteins PCPE-1 and collagen I by Western blot on the enrichment fractions. PCPE-1 first and then collagen I were immunoblotted on the same membrane (reducing conditions, AF2239 antibody used to detect PCPE-1,  $\alpha$ -collagen I antibody directed against the triple helical domain). **C.** Detection of the cellular protein actin by Western blot in the enrichment fractions.

## 2. Comparison of different ECM enrichment protocols for mouse skin

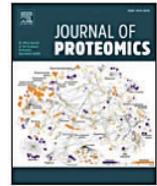
To get a better coverage of the matrisome and see if it was possible to avoid the loss of matrix proteins during tissue processing, additional ECM enrichment strategies were tested on mouse back skin. Chemical decellularization<sup>193,195</sup> and quantitative detergent solubility profiling (QDSP)<sup>99</sup>, which were previously used to study the matrisome of cardiovascular tissues and lungs respectively, were tested. These protocols were selected because they seemed to allow a higher coverage of matrix proteins, especially through the study of the intermediate fractions generated during tissue processing while compartmental enrichment was designed to focus on the insoluble pellet only. An enzyme-based decellularization strategy previously applied to human and rat skin was also adapted<sup>200,201</sup>. The comparison of all these ECM-enrichment protocols on mouse skin led to the publication of a manuscript in the *Journal of Proteomics* in 2021 (presented below) and resulted in the first description of the matrisome of mouse healthy skin.



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Proteomics

journal homepage: [www.elsevier.com/locate/jprot](http://www.elsevier.com/locate/jprot)

Full Length Article



## Comparison of extracellular matrix enrichment protocols for the improved characterization of the skin matrisome by mass spectrometry

Mélissa Dussoyer<sup>a</sup>, Adeline Page<sup>b</sup>, Frédéric Delolme<sup>b</sup>, Patricia Rousselle<sup>a</sup>, Alexander Nyström<sup>c,d</sup>, Catherine Moali<sup>a,\*</sup>

<sup>a</sup> University of Lyon, CNRS, Tissue Biology and Therapeutic Engineering Laboratory, LBTL, UMR5305, F-69367 Lyon, France

<sup>b</sup> University of Lyon, INSERM, ENS Lyon, CNRS, Protein Science Facility, SFR BioSciences, UAR3444/US8, F-69366 Lyon, France

<sup>c</sup> Department of Clinical Dermatology/Medical Center, University of Freiburg, Freiburg, Germany

<sup>d</sup> Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Freiburg, Germany

### ARTICLE INFO

#### Keywords:

Extracellular matrix  
Skin  
Matrisome  
Proteomics  
Decellularization  
Mouse

### ABSTRACT

A striking feature of skin organization is that the extracellular matrix (ECM) occupies a larger volume than the cells. Skin ECM also directly contributes to aging and most cutaneous diseases. In recent years, specific ECM enrichment protocols combined with *in silico* approaches allowed the proteomic description of the matrisome of various organs and tumor samples. Nevertheless, the skin matrisome remains under-studied and protocols allowing the efficient recovery of the diverse ECM found in skin are still to be described. Here, we compared four protocols allowing the enrichment of ECM proteins from adult mouse back skin and found that all protocols led to a significant enrichment (up to 65%) of matrisome proteins when compared to total skin lysates. The protocols based on decellularization and solubility profiling gave the best results in terms of numbers of proteins identified and confirmed that skin matrisome proteins exhibit very diverse solubility and abundance profiles. We also report the first description of the skin matrisome of healthy adult mice that includes 236 proteins comprising 95 core matrisome proteins and 141 associated matrisome proteins. These results provide a reliable basis for future characterizations of skin ECM proteins and their dysregulations in disease-specific contexts.

**Significance:** Extracellular matrix proteins are key players in skin physiopathology and have been involved in several diseases such as genetic disorders, wound healing defects, scleroderma and skin carcinoma. However, skin ECM proteins are numerous, diverse and challenging to analyze by mass spectrometry due to the multiplicity of their post-translational modifications and to the heterogeneity of their solubility profiles. Here, we performed the thorough evaluation of four ECM enrichment protocols compatible with the proteomic analysis of mouse back skin and provide the first description of the adult mouse skin matrisome in homeostasis conditions. Our work will greatly facilitate the future characterization of skin ECM alterations in preclinical mouse models and will inspire new optimizations to analyze the skin matrisome of other species and of human clinical samples.

### 1. Introduction

Skin is the one of the largest organs in mammals. It provides protection against water loss and biological, chemical and physical assaults [1]. These properties are conferred on skin by its specific tissue architecture consisting of two main layers: the epidermis, a stratified squamous epithelium which provides the barrier function, and the dermis which mainly consists of a dense extracellular matrix (ECM), synthesized by fibroblasts, and responsible for skin mechanical strength and elasticity. At the interface between the epidermis and dermis is the

dermal-epidermal junction (DEJ), a thin network of specialized ECM components involved in maintaining the dermal-epidermal cohesion [2,3].

ECM structural proteins are the main components of the dermis and DEJ. They comprise collagens, proteoglycans and glycoproteins that all share the common ability to participate in supramolecular assemblies [4]. Collagens have triple helical domains and can form a large variety of networks [5,6], from large fiber bundles (collagens I, III, V) or microfibrils (collagen VI) in the dermis to anchoring fibrils (collagen VII) or lattice networks (collagen IV) in the DEJ. Proteoglycans are decorated

\* Corresponding author at: Tissue Biology and Therapeutic Engineering Laboratory, LBTL, 7, Passage du Vercors, 69367 Lyon Cedex 07, France.

E-mail address: [catherine.moali@ibcp.fr](mailto:catherine.moali@ibcp.fr) (C. Moali).

<https://doi.org/10.1016/j.jprot.2021.104397>

Received 15 May 2021; Received in revised form 23 September 2021; Accepted 3 October 2021

Available online 20 October 2021

1874-3919/© 2021 Elsevier B.V. All rights reserved.

by long glycosaminoglycan chains with specific binding properties [7]. They regulate collagen fibrillogenesis, as is the case for small leucine-rich proteoglycans (SLRPs, e.g. decorin, biglycan, fibromodulin), play major roles in angiogenesis (perlecan) and act as signaling molecules or as reservoirs/antagonists for signaling molecules [8–10]. Glycoproteins are a more heterogeneous group of proteins which can also form supramolecular networks, such as fibrils for fibronectin in the dermis or anchoring filaments for laminin 332 in the DEJ. Besides, they are often involved in interactions with other ECM proteins to regulate matrix architecture and cell behavior (e.g. fibrillins, thrombospondins, tenascins, periostin) [8,11]. Notably, the three types of ECM structural proteins display some degree of overlap as some collagens are decorated with glycosaminoglycan chains (e.g. collagen XVIII) and O- and N-glycosylations are not only found in glycoproteins but also in collagens and proteoglycans.

In addition to these well-characterized structural proteins, many additional proteins are found in variable amounts in the ECM and are critical players in the control of skin homeostasis and in responses to injury or stress signals. These include the numerous extracellular enzymes involved in ECM biosynthesis or degradation (procollagen proteinases, matrix metalloproteinases and cross-linking enzymes) [5,12,13], inhibitors of these enzymes (serpins, cystatins and tissue inhibitors of metalloproteinases or TIMPs) [14], growth factors, cytokines or angiogenic molecules stored in the ECM (e.g. TGF- $\beta$ , chemokines, VEGF) [9]. Components released from the cell surface can also bind to ECM components and contribute to regulate ECM-dependent signaling (e.g. CD109, glypicans, annexins) [15–17].

In order to better define ECM and ECM-associated proteins, a systematic approach combining the *in silico* analysis of protein domains with an extensive literature review has been proposed a few years ago [18]. In human and mouse, more than 1000 proteins were selected to define the “matrisome” and grouped in two main subcategories. The “core matrisome” is composed of architectural proteins namely collagens, proteoglycans and glycoproteins, as described above. The “associated matrisome” comprises ECM regulators, secreted factors and ECM-affiliated proteins. These lists have proven very useful to annotate omics data and now serve as references for the analysis of the ECM composition of diverse tissues and organisms [19].

In parallel, new protocols have been developed to better extract ECM proteins, which are often insoluble and highly cross-linked. Among them, a method based on the successive extraction of proteins found in different cell compartments to yield an insoluble pellet enriched in ECM components [20] has been widely used and shown to be compatible with proteomic analysis. This approach helped to identify the matrix content of healthy and diseased tissues in human and mouse, including pancreatic islets [21], breast [22], lung [23], liver and colon [24]. The main pitfall of this approach is the loss of loosely-bound ECM components. Another method based on protein solubilization in buffers containing increasing detergent concentrations (named quantitative detergent solubility profiling or QDSP) was also developed and found to give very good matrisome coverage of the lung tissue [25]. Finally, protocols based on tissue decellularization, using detergents [26,27] or enzymes such as phospholipase A2 (PLA2) [28], which were initially developed for the analysis of the ECM from cardiovascular tissues or the production of acellular scaffolds [29], also proved efficient to enrich matrisome proteins and facilitate their identification by tandem mass spectrometry (MS/MS). These decellularization protocols were recently applied to human skin and led to the identification of between 137 and 267 ECM proteins [28,30,31]. However, when applied to rat skin, the PLA2 decellularization approach only yielded 102 matrisome proteins [28], suggesting that it may not be optimal to analyze murine skin.

Here, we compared four ECM enrichment protocols with a more classical skin lysis method in order to optimize skin matrisome recovery for MS-based proteomics. We then investigated the nature of identified matrisome proteins and their distribution among the fractions generated by the various enrichment methods. Finally, we combined all the data to

generate the first experimentally-based, detailed description of the mouse skin matrisome.

## 2. Material and methods

### 2.1. Chemicals

NaCl (1112-A) and HEPES (10–110) were purchased from Euro-medex. Acetonitrile (ACN; 34998), calcium chloride (2233506), chloroform (34854), dithiothreitol (DTT; D0632), DNAse I (04536282001), ethylenediaminetetraacetic acid (EDTA; E5134), formic acid (33015), glycerol (G9012), guanidinium hydrochloride (GuHCl; G4630), NP-40 (13021), iodoacetamide (I1149), magnesium chloride (M1028),  $\beta$ -mercaptoethanol (M7154), methanol (34860), phospholipase A2 (PLA2; P9279), cOmplete™ protease inhibitor cocktail (11836153001, prepared according to the manufacturer's instructions), cOmplete™ EDTA-free protease inhibitor cocktail (11836170001, prepared according to manufacturer's instructions), sterile phosphate buffered saline solution (PBS; D8537), sodium acetate (S2889), sodium deoxycholate (SD; D6750), sodium dodecyl sulfate (SDS; L3771), sodium hydroxide (NaOH; S5881), trifluoroacetic acid (TFA; 91707), TRIS-HCl (T3253), Triton X-100 (T8787) and urea (33247) were from Merck. PNGase-F (P0709S), Lys-C (12505061) and trypsin (V5111) were purchased from New-England Biolabs, Promega and Wako respectively. All solutions were prepared in ultra-pure water or in dedicated buffers supplied by the manufacturer and sterilized by 0.22  $\mu$ m-filtration.

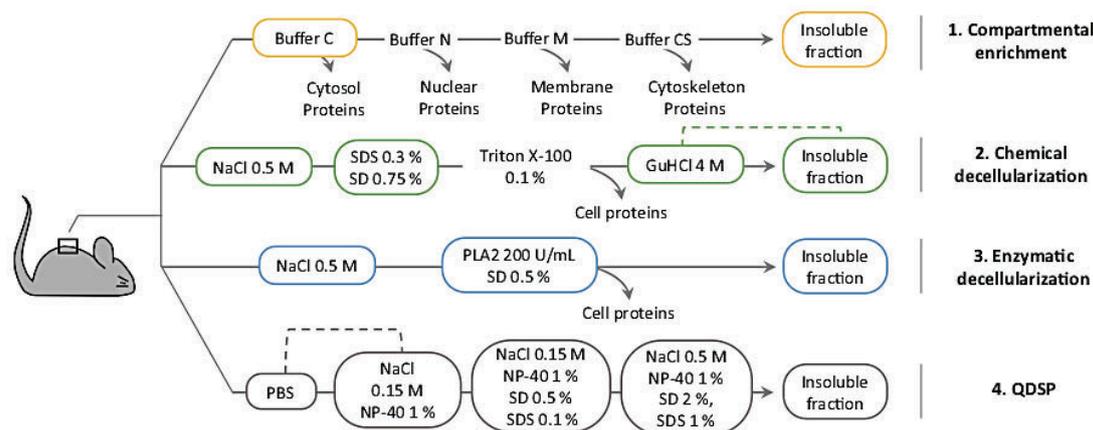
### 2.2. Skin samples

Animal experiments followed governmental and international guidelines. Mouse back skin biopsies were collected from pathogen- and disease-free C57BL/6 J animals (22 weeks-old) obtained from the Jackson Laboratory. After sacrifice, the back skin was shaved and depilated twice using depilatory cream (3059944026734, Veet). The skin was dissected using scissors and separated from the underlying tissues, taking care of removing as much subcutaneous fat as possible, and stored in Eppendorf LoBind tubes. Tissues were flash-frozen by immersion in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until use.

### 2.3. Extraction of matrisome proteins

Four protocols were adapted and compared to enrich ECM proteins from mouse skin (Fig. 1): 1. compartmental enrichment [20]; 2. chemical decellularization (adapted from [27,32]); 3. enzymatic decellularization using phospholipase A2 (adapted from [33,34]); 4. quantitative detergent solubility profiling (QDSP) [25]. All four protocols were started the same day, each using 35 mg of frozen back skin, sampled from the same initial back skin piece. The experiment was repeated three times starting from three different mice.

- Compartmental enrichment: skin samples were processed using the CNMCS commercial kit (2145, Merck Millipore), as previously described [20], to progressively remove cellular proteins from the native tissue using combinations of salts and detergents (precise composition unknown, but more information can be found in [20]). Briefly, the skin tissue was homogenized in buffer C using a Tissue-Lyser II (Qiagen) and successively incubated under agitation in buffers W, N, M and CS supplemented with the protease inhibitors from the kit. Extractions with buffer C and final insoluble pellet were kept, flash-frozen immediately after recovery and stored at  $-80^{\circ}\text{C}$  until MS analysis.
- Chemical decellularization: skin samples were cut into small pieces with a scalpel and gently rotated in 500 mM NaCl, 100 mM HEPES, pH 7.4 for 2 h 30. Samples were then centrifuged for 10 min at  $23^{\circ}\text{C}$  and 16,000 g, and the supernatant was kept aside and replaced by ultra-pure water. After 10 min, water was discarded and the skin was



**Fig. 1.** Overview of the steps and buffers used for tissue processing in selected ECM enrichment protocols. The colored circles indicate fractions that were analyzed by MS/MS.

treated with 0.75% (w/v) SD and 0.3% (w/v) SDS for 16 h, then a 1% (v/v) Triton X-100 bath was applied for 1 h to remove remaining cell fragments. The resulting tissue was washed twice in ultra-pure water for 1 h 30 and centrifuged for 10 min at 23 °C and 16,000 g. The pellet was finally incubated in 4 M GuHCl and 50 mM sodium acetate for 72 h. All incubation buffers, including ultra-pure water, were supplemented with cOmplete™ protease inhibitors and EDTA (to reach a final concentration of 25 mM EDTA). Successive extractions were performed using a Thermomixer (Thermo Fisher Scientific) at 23 °C and 600 rpm. Extractions in 500 mM NaCl, in SD/SDS and in 4 M GuHCl as well as the final insoluble tissue fragments were kept, flash-frozen immediately after recovery and stored at –80 °C until MS analysis.

- Enzymatic decellularization: the first step was the same as in the previous protocol. After the first extraction in 500 mM NaCl, 100 mM HEPES pH 7.4, cOmplete™ protease inhibitors and 25 mM EDTA, chopped skin pieces were washed twice in sterile PBS for 10 min at 4 °C and 600 rpm and centrifuged for 10 min at 4 °C and 16,000 g. Decellularization was performed using 200 U/mL PLA2 in 150 mM NaCl, 100 mM HEPES pH 7.4 for 18 h at 37 °C, 500 rpm. The enzyme incubation buffer was supplemented with 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 750 U/mL DNase I, EDTA-free cOmplete™ protease inhibitors and 0.5% SD. The sample was centrifuged for 10 min at 23 °C and 16,000 g and briefly washed with sterile PBS. A final centrifugation was applied for 15 min at 4 °C and 16,000 g and PBS was discarded. PLA2/SD extraction solution and final insoluble tissue fragments were kept, flash-frozen immediately after recovery and stored at –80 °C until MS analysis.
- Quantitative detergent solubility profiling (QDSP): skin samples were first homogenized in sterile PBS with 25 mM EDTA using a TissueLyser II and processed as previously reported [25]. Briefly, the tissue was successively incubated for 20 min under agitation (300 rpm) in PBS, Buffer 1 (150 mM NaCl, 50 mM TRIS-HCl pH 7.5, 1 mM MgCl<sub>2</sub>, 5% Glycerol, 1% NP-40), Buffer 2 (150 mM NaCl, 50 mM TRIS-HCl pH 7.5, 5% Glycerol, 1% NP-40, 0.1% SDS, 0.5% SD) at 4 °C and in Buffer 3 (500 mM NaCl, 50 mM TRIS-HCl pH 7.5, 5% Glycerol, 1% NP-40, 1% SDS, 2% SD) at 23 °C. All buffers were supplemented with cOmplete™ protease inhibitors and, between extractions, samples were centrifuged for 20 min at 16,000 g and 4 °C except after Buffer 3 for which the temperature was set at 23 °C. All extractions solutions were kept after centrifugation. PBS and Buffer 1 extractions were pooled in equal volumes while other extraction solutions were kept separately. All samples were flash-frozen in liquid nitrogen and stored at –80 °C until MS analysis.

For comparison with the four protocols described above, total skin lysates were also prepared for two of the mice. In this case, 20 mg of mouse skin were homogenized using a TissueLyser II in a lysis buffer containing 4% SDS, 5% β-mercaptoethanol, 12.5% glycerol, 100 mM TRIS-HCl pH 6.8, supplemented with cOmplete™ protease inhibitors. Samples were strongly vortexed, boiled for 15 min and incubated for 30 min at 23 °C and 600 rpm. Supernatants were then centrifuged for 15 min at 23 °C and 16,000 g to eliminate cell fragments, then flash-frozen in liquid nitrogen and stored at –80 °C until MS analysis.

#### 2.4. Sample preparation for MS/MS

200 μL of each liquid extraction sample collected above were precipitated at room temperature using the methanol/chloroform method. Briefly, 800 μL of methanol, 200 μL of chloroform and 600 μL of ultra-pure water were added to each sample and, after vortexing, the mixture was centrifuged for 10 min at 16,000 g. The upper liquid phase was carefully discarded and 800 μL of methanol were added to wash proteins. Samples were strongly vortexed, centrifuged for 10 min at 16,000 g and methanol was discarded. Protein pellets were briefly vacuum-dried and solubilized in 55 μL of 100 mM NaOH at room temperature. Protein concentrations were determined using a commercial BCA kit (23225, Thermo Fisher Scientific).

For each liquid fraction, 10 μg of proteins were further processed for LC-MS/MS analysis. Proteins were reduced with 10 mM DTT for 40 min at 60 °C and alkylated with 25 mM iodoacetamide for 40 min at 23 °C in the dark. Samples were diluted 4-fold with 100 mM ammonium bicarbonate pH 8.0 and pH was adjusted to 8.0 with 1.8 M HEPES pH 8.0. Proteins were deglycosylated using PNGaseF (10 U) for 2 h and digested into peptides by successive treatments with Lys-C (1:150 enzyme to protein mass ratio) for 2 h and trypsin (1:50) overnight. The following day, trypsin was added again (1:100) and the incubation continued for 2 h. All enzyme treatments were performed at 37 °C and 700 rpm. Samples were acidified to pH 2.0 using 50% TFA, desalted on C18 columns (89873, Life Technologies) and vacuum-dried.

Around 5 mg of each insoluble pellet were dissolved in 120 μL of 100 mM NaOH, strongly vortexed and incubated for 2 h at 23 °C and 600 rpm. Proteins were reduced, alkylated and digested as described for liquid samples. Amounts of enzymes were increased as follows to allow the complete digestion of proteins: 50 U PNGaseF, 2 μg Lys-C and 2 μg or 1 μg trypsin for the overnight incubation or the 2 h-treatment, respectively. Samples were acidified to pH 2.0 using 50% TFA and desalted on C18 columns. Peptide concentrations in the resulting fractions as well as in the liquid GuHCl extraction of the chemical decellularization protocol were determined using the Pierce™ quantitative fluorimetric peptide

assay (23290, Thermo Fisher Scientific). 10 µg of each sample were used for MS analysis. The GuHCl extraction and the insoluble fraction of the chemical decellularization protocol were pooled in equal peptide amounts. Samples were finally dried in a vacuum centrifuge.

### 2.5. MS/MS analysis

Dried samples were resuspended in 20 µL of 2% ACN/0.1% formic acid and 1 µL was injected and loaded on a C18 Acclaim PepMap100 trap-column 0.3 × 5 mm, 5 µm, 100 Å (Thermo Fisher Scientific) for 3 min at 20 µL/min with 2% ACN, 0.05% TFA in ultra-pure water and then separated on a C18 Acclaim PepMap100 nano-column, 50 cm × 75 µm i.d., 2 mm, 100 Å (Thermo Fisher Scientific) with a four-step linear gradient from 4% to 20% buffer B in 100 min (A: 0.1% formic acid in ultra-pure water, B: 80% ACN, 0.1% formic acid in ultra-pure water), from 20% to 32% of B in 20 min and then from 32% to 90% of B in 2 min, followed by 90% of B for 10 min. The flow rate was set at 300 nL/min. The column oven temperature was kept constant at 40 °C. Peptides were analyzed with an online-coupled Q-Orbitrap mass spectrometer (Q Exactive HF; Thermo Fisher Scientific) in a Top 20 HCD (Higher Collision Dissociation) data-dependent acquisition. The resolution of the survey scan was 60,000 at *m/z* 200 Th and for MS/MS scan the resolution was set to 15,000 at *m/z* 200 Th. Parameters for acquiring HCD MS/MS spectra were as follows: collision energy = 27 and isolation width = 2 *m/z*. Precursor ions with unknown charge state, charge states of 1 and 8 or greater than 8 were excluded. Peptides selected for MS/MS acquisition were then placed on an exclusion list for 20 s using the dynamic exclusion mode to limit duplicate spectra. Data were processed by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 2.4 software (Thermo Fisher Scientific) against the Uniprot mouse database (101,412 entries, December 2020). Precursor and fragment mass tolerance were set to 10 ppm and 0.02 Da respectively. Trypsin was selected as enzyme, and up to 2 missed cleavages were allowed. Acetylation (protein N-terminus, +42.011 Da), methionine loss (protein N-terminus, -131.040 Da), carbamidomethylation (Cysteine, +57.021 Da) and oxidation (Methionine, Proline and Lysine, the latter only when indicated, +15.995 Da) were set as variable modifications. Peptides were filtered with a false discovery rate of 1%. Protein abundances were obtained by summing intensities of precursor ions and normalized to total peptide counts. Mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository [35] with the dataset identifier PXD025842

### 2.6. Data processing

Matrisome proteins were annotated using the mouse matrisome lists (available on-line: <http://matrisomeproject.mit.edu/>; last update: August 2014) [19] and the tools of the on-line platform ProteoRE (<http://www.proteore.org/>) [36]. Proteins identified by MS/MS were searched against matrisome lists using Uniprot IDs. To confirm the results of this first screening and include new protein variants identified since 2014 and not referenced in the mouse matrisome database, a second search was performed using Gene IDs and then followed by a final manual validation. Heatmaps and Venn diagrams were prepared with ProteoRE.

## 3. Results and discussion

### 3.1. Skin ECM enrichment workflows

Four protocols dedicated to ECM enrichment or tissue decellularization were selected from literature, based on their user-friendliness and compatibility with MS analysis, and adapted to enrich matrisome proteins from adult mouse skin (Fig. 1). The first protocol (hereafter called the “compartmental enrichment” method) relies on the sequential enrichment of different cellular compartments to yield an insoluble

pellet mainly consisting of ECM proteins and is currently used as a reference protocol for matrisome studies [20]. Two other protocols (Fig. 1, “chemical decellularization” and “enzymatic decellularization”), adapted from [27,32] and [33,34] respectively, aim at destabilizing cell membranes through the use of detergents or PLA2, to remove cellular components and enhance the recovery of extracellular proteins. Last, the QDSP method consists in successive incubations in buffers with increasing detergent concentrations and extracts both cellular and extracellular proteins according to their solubility [25]. Notably, solubility profiling allowed by the QDSP method and to a lesser extent by the chemical decellularization method, can be an advantage to study ECM proteins that often exist in different forms (precursor proteins, mature proteins, proteolytic products) and are more or less tightly integrated in ECM networks, therefore exhibiting different solubility levels. The composition of all the buffers used in chemical and enzymatic decellularizations or in QDSP are known but the exact composition of the commercial solutions used for compartmental enrichment is proprietary (see the Material and Methods section). All insoluble pellets and the extraction solutions which were the most likely to contain matrisome proteins, based on preliminary MS and western-blot evaluation, were analyzed by LC-MS/MS (Fig. 1). Some fractions, which were found in preliminary experiments to have similar compositions were also pooled and analyzed together (the first two fractions of the QDSP protocol and the last two fractions of the chemical decellularization method).

### 3.2. The number of matrisome proteins identified in mouse skin is strongly increased by enrichment methods

Between 2091 and 2703 proteins were identified by LC-MS/MS with the four ECM enrichment protocols when mouse skin was processed in three biological replicates (Table 1, Supplementary Table 1). Among them, between 161 and 202 proteins corresponded to matrisome components, giving a constant percentage of 8% of the total number of identified proteins. Very similar numbers of matrisome proteins were respectively obtained using the chemical decellularization (202 proteins), enzymatic decellularization (201 proteins) and QDSP (201 proteins) approaches, while the compartmental enrichment method led to a lower number of identifications (161). Importantly, the choice of a database containing both reviewed and unreviewed sequences was made because some important mouse matrisome proteins such as tenascin X, α3(V) and α3(VI) collagen chains, trichohyalin, transglutaminase-6 or ADAMTS-like protein 5 were missed when only reviewed sequences were included.

The comparison of protein identities revealed that 139 matrisome proteins were shared by all enrichment protocols, while a total of 234 matrisome proteins were detected when the results of the four protocols were combined (Fig. 2A). Remarkably, the three most effective methods (chemical and enzymatic decellularizations, QDSP) display 84% overlap in matrisome components, showing that only minor improvement of matrisome coverage is brought by protocol combination. Furthermore, only 2.2% of identified matrisome proteins were specifically identified in the compartmental enrichment method. However, even the least efficient ECM enrichment protocol showed a clear benefit over the analysis of a classical skin lysate as only 122 matrisome components could be annotated (from a total of 1107 identified proteins) in the total lysate analyzed without fractionation (Table 1). Also, only 2 matrisome proteins (Col11a1 and Col25a1) were detected in the total skin lysate but not after enrichment (Supplementary Fig. 1).

The number of 161 matrisome proteins identified with the compartmental enrichment approach compares well with the 113 or 185 proteins detected in mouse lung [23] or pancreatic islets [21] with the same approach. However, the latter analyses were conducted using the insoluble fraction alone; on mouse skin, the study of this single fraction led to the identification of 102 proteins only (Supplementary Fig. 2). A recent study actually suggested that this method leads to high losses of matrisome components in the various extraction buffers when applied to

**Table 1**

Overview of protein numbers identified in enrichment protocols and total lysates. The data obtained with three mice were pooled to analyze the numbers of total and matrisome-related proteins identified for each protocol.

	Compartmental enrichment	Chemical decellularization	Enzymatic decellularization	QDSP	Total lysate
Principle	Protein fractionation of specific cell compartments	Progressive decellularization using SDS, SD, Triton X-100	Decellularization using PLA2 and SD	Protein fractionation depending on solubility in detergents	Total tissue lysis in SDS
Total proteins	2091	2703	2596	2341	1107
Matrisome proteins, including:	161	202	201	201	122
Core matrisome	38.9%	42.1%	41.3%	41.0%	50.8%
Associated matrisome	61.1%	57.9%	58.7%	59.0%	49.2%
% of matrisome proteins found in at least 2 biological replicates	74.5%	73.8%	74.7%	69.5%	66.4%

a whole mouse powder [37]. In line with this, the analysis of the extraction made with buffer C actually proved very informative and increased the number of identified matrisome components in mouse skin by 59% compared to the insoluble fraction alone (Supplementary Fig. 2), suggesting that it contains more loosely-bound ECM components.

Chemical decellularization was previously extensively used to characterize heart and cardiovascular tissues by proteomic analysis [26,38–41] and allowed the identification of between 17 and 328 ECM proteins (note that all authors do not use the same inclusion criteria to label ECM proteins). Interestingly, most of these studies relied mainly on the analysis of the NaCl and GuHCl extracts and not on the analysis of the final insoluble pellet to reveal ECM proteins. Thus, the benefit of extending MS analysis to the first extraction fractions (NaCl fraction and to a lesser extent SDS fraction) was also suggested in these reports and is nicely confirmed here. Among the 202 proteins which were detected in mouse skin in total with this protocol, NaCl and SDS/SD extracts specifically brought 71 matrix proteins (Supplementary Fig. 3). More intriguingly, the PLA2 decellularization method gave significantly better results when applied here to mouse skin than in a previous study on rat skin [28] (201 versus 102 matrisome components).

Finally, we identified 201 matrisome proteins with the QDSP approach while previous studies reported 161 and 435 matrisome components in the ECM-poor mouse brain [42] and ECM-rich mouse lung [25], respectively. Notably, more than 8000 total proteins were identified in these two previous reports but a large panel of basal and disease conditions were characterized and all extractions solutions were analyzed after an additional simplification of the fractions using a quick stage tip fractionation that we did not perform here.

Beyond total protein numbers, it is also interesting to compare peptide spectral matches (PSMs) and protein sequence coverages which give additional information about the comprehensiveness of the analysis. While QDSP led to the highest total number of PSMs for matrisome proteins (Fig. 2B), as expected for the protocol with the largest number of analyzed fractions, chemical decellularization also behaved well for matrisome proteins giving similar numbers of PSMs per fraction. The enzymatic decellularization came after, mostly due to the relative PSM paucity of the PLA2/SD and insoluble fractions, and the compartmental extraction was by far the protocol which gave the lowest number of PSMs. Interestingly however, these marked differences tended to normalize when the median sequence coverages of matrisome proteins (Fig. 2C) were considered. Median sequence coverages ranged from 13% to 20% with the lowest value obtained for the insoluble fraction of the compartmental protocol and the highest value for the insoluble fraction of the chemical decellularization method.

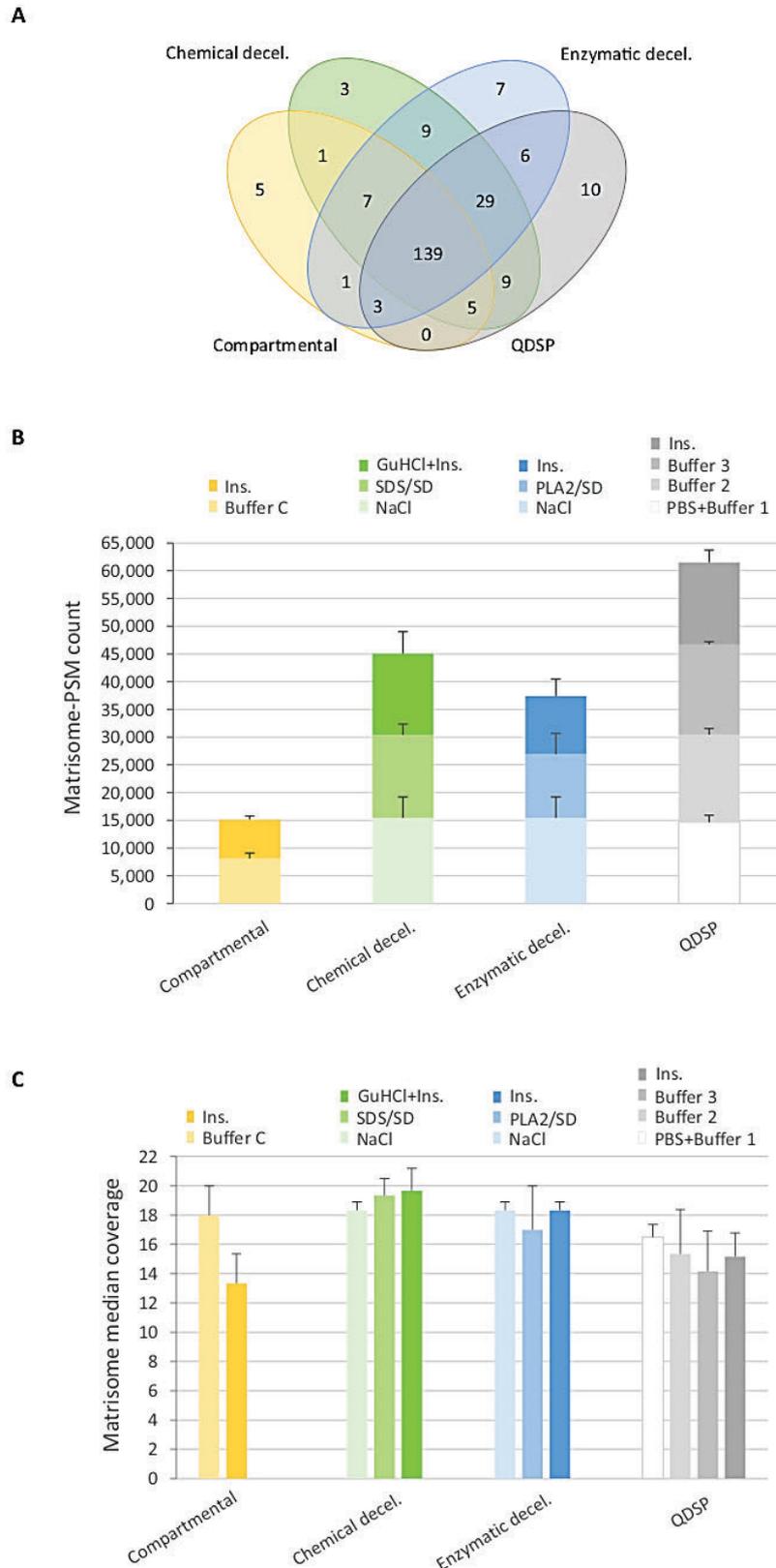
Finally, in all protocols, the total number of matrisome proteins identified in the insoluble fractions was comparatively lower than in other fractions (Supplementary Fig. 2–5), suggesting that the analysis of the more soluble fractions provide significant added-value in terms of matrisome coverage, as further discussed below.

### 3.3. The skin associated matrisome is better covered through ECM enrichment and requires extraction in less stringent buffers

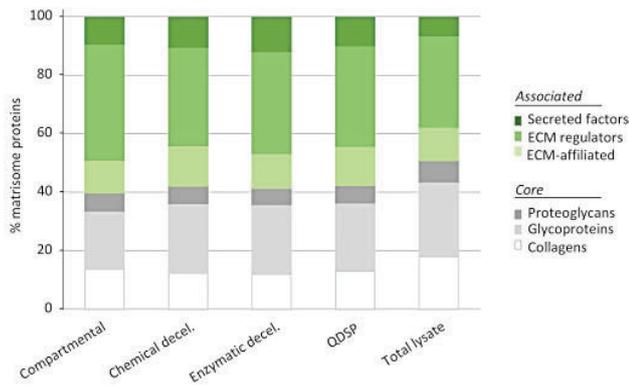
ECM enrichment protocols all led to the same distribution of proteins between the core (around 40%) and associated (around 60%) matrisomes (Fig. 3). However, the proportion of associated matrisome components decreased to 50% in the total skin lysate, suggesting that the latter approach was less efficient to reveal lower-abundance ECM components. Furthermore, a similar distribution of proteins among the 6 matrisome subcategories was observed for all enrichment methods while a substantially lower content in secreted factors was obtained with the total skin lysate. Inversely, collagens were enriched in the total skin lysate and represented 18% of the matrisome proteins, to be compared with a mean value of 13% for enrichment protocols. These results indicated that skin processing through total lysis led to samples which were too complex to allow in-depth identification of matrisome components by MS/MS.

The reproducibility of ECM enrichment protocols was also evaluated by comparing the results obtained from three biological replicates corresponding to skin samples originating from three different mice. A total of 188, 173 and 200 matrisome components was detected in the three mice and 145 proteins were common to all animals (Fig. 4A). The distribution of all identified proteins between the core and associated matrisome categories was then characterized in each fraction of the four enrichment protocols for the three mice. In total, 11 fractions were characterized for the three mice in addition to the total lysates (Fig. 1). The NaCl fractions of the chemical and enzymatic decellularization methods being identical, only one of two collected fractions was analyzed by MS/MS. The number of core matrisome proteins varied substantially in the various extraction fractions of the three mice (between 19 and 63) and was 49 or 69 for the two skin samples analyzed as total lysates (Fig. 4B). For all enrichment protocols, the highest inter-individual variability was systematically observed in the first steps of tissue processing, suggesting that the most labile ECM core components are more easily lost and difficult to recover in a reproducible manner.

Interestingly, the core matrisome proteins which could have been expected to be enriched in high stringency buffers due to their low solubility were also found in high numbers in the first extraction fractions (Fig. 4B), as confirmed by the detailed analysis of fraction content (Supplementary Table 2). This finding suggests that core matrisome proteins show a wide range of solubilities depending on their interactions with other ECM components or on their maturation levels. Indeed, several core matrisome components undergo extracellular maturation, which are known to affect their solubility. Prototypical examples are the fibrillar collagens (mainly I, III and V in skin) which are secreted as soluble precursor molecules and cannot form fibrils before the proteolytic removal of their N- and C-terminal propeptides [43]. The chains of these collagens were identified in all fractions, at least for one of the three mice. Proteolytic maturation also affect collagens VI [44] and VII [45], laminin 332 or the small leucine-rich proteoglycans



**Fig. 2.** Comparison of the matrisome content identified with the four enrichment protocols. A: Venn diagram of matrisome proteins detected with the four methods (the diagram including total lysates is available in Supplementary Fig. 1). B: Cumulative PSM counts for matrisome proteins (shown as a sum of individual fractions). C: Median sequence coverage for matrisome proteins. All bar plots represent means with SD ( $n = 3$  mice). Ins = insoluble fraction.



**Fig. 3.** Percentage of the matrisome subcategories detected in mouse skin using the enrichment protocols and the total lysate approach.

biglycan, decorin and mimecan [46] which were well distributed among extraction fractions. The exception is the  $\gamma 2$  chain of laminin 332 (Lamc2), which seemed to be preferably extracted in the less soluble skin compartments. In agreement with this observation, all the peptides identified for this laminin chain corresponded to the mature form. More generally, proteolysis is known to play a major role to regulate the interaction properties of ECM proteins. For example, thrombospondin-1 was exclusively identified in low stringency buffers with peptides belonging to its N-terminal domain and it is known that thrombospondins lose their ability to form oligomers and become soluble when they are cleaved at specific sites further in the protein sequence [47,48]. Also, small fragments released from core matrisome proteins, also known as matricryptins, can play major signaling functions [13,49] due to their increased solubility, potentially explaining the presence of several collagens, fibronectin, perlecan or elastin in the first extraction fractions of the various protocols.

Regarding the associated matrisome, between 17 and 83 proteins were identified in the 11 enrichment fractions and, in contrast to the results obtained with the core matrisome, there was a clear decrease in protein recovery with the progression of the four protocols towards more stringent buffers (Fig. 4C). Strikingly, for all protocols, the insoluble fraction had the lowest number of identified proteins while the first extraction of each protocol was the most efficient to enrich associated matrisome. For example, with the QDSP protocol, which has been specifically designed for protein solubility profiling, mean numbers of 63, 60, 49 and 27 proteins were recovered with buffer 1 combined with PBS, buffer 2, buffer 3 and in insoluble pellet respectively. This corresponds well to the solubility profile expected for the associated matrisome, with a high content in enzymes and signaling molecules, which are potentially less tightly bound to the matrix than core matrisome components. Still, some associated matrisome components were also found in the final insoluble fractions such as annexins, cystatins, galectins or plasminogen (Supplementary Table 2). Also, the cross-linking enzyme lysyl oxidase homolog-1 and the tissue inhibitor of matrix metalloproteinase 3 (TIMP-3) were only found in insoluble fractions.

#### 3.4. Core matrisome components are quantitatively more stable between individuals and protocols than associated matrisome components

For all enrichment fractions, the high similarity in the identity of core matrisome components observed between replicates suggests a highly stable composition in skin structural ECM proteins (Fig. 4B). In contrast, the composition of the associated matrisome showed more marked differences between replicates (Fig. 4C). This was observed for all enrichment protocols with the most pronounced differences in the insoluble fractions. This suggests that the low abundance of associated matrisome components makes their detection more challenging and less

reproducible or that these components are more likely to undergo individual-specific modulations due to their regulatory functions.

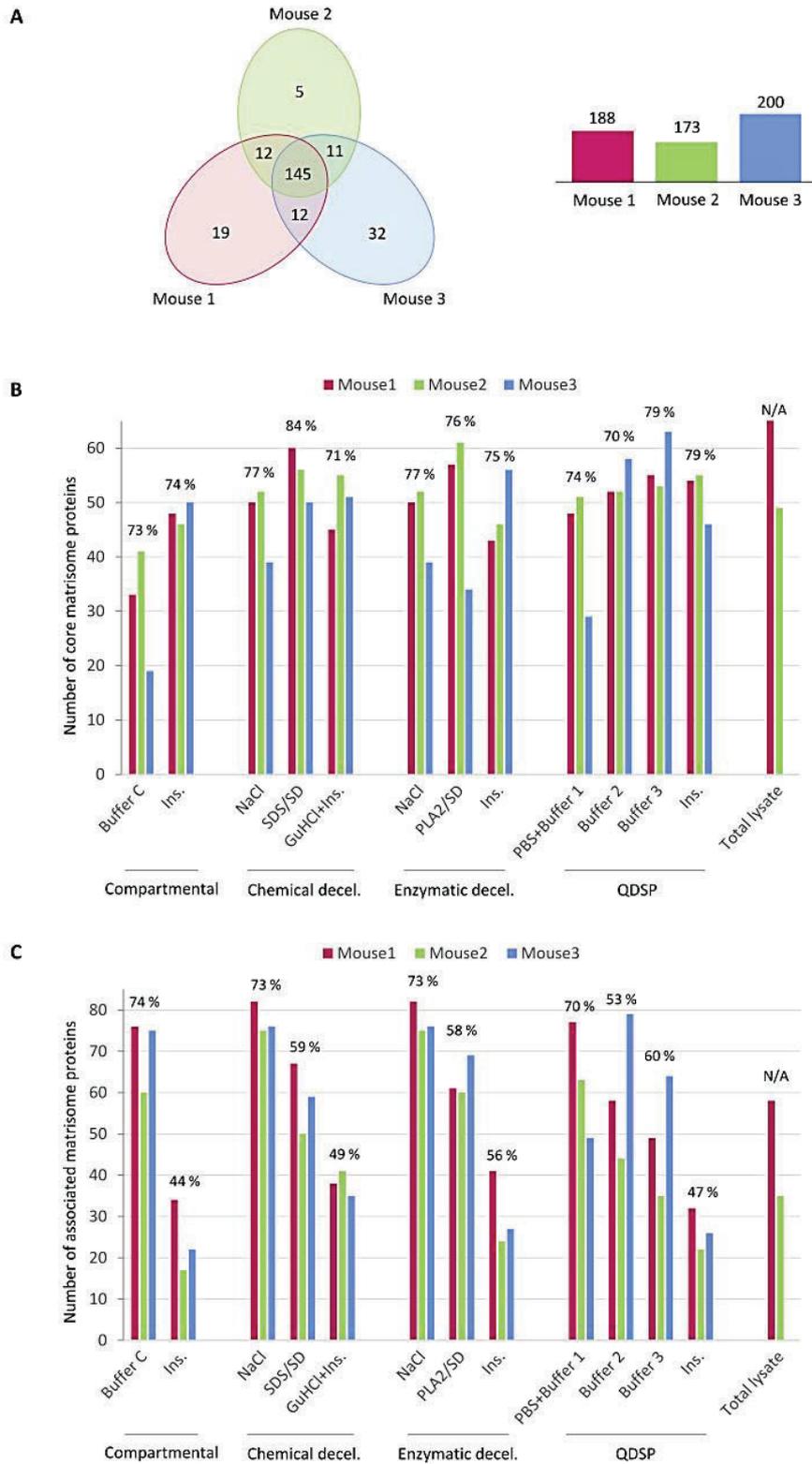
To confirm this qualitative trend and assess the compatibility of ECM enrichment protocols with quantitative analysis, we compared the abundance of matrisome proteins in the insoluble fractions of the three mice (Fig. 5, Supplementary Fig. 6 and Supplementary Table 3). All tested methods give coherent abundance values for core matrisome proteins, both when replicates are compared and when methods are compared. As expected, collagens appear as the most abundant components in the insoluble compartment of skin samples followed by proteoglycans and then by glycoproteins which are globally less abundant. It is also in the glycoprotein group that the variability between individuals is most pronounced, a trend which is further exacerbated in the associated matrisome (Supplementary Fig. 6). Therefore, abundance results confirm the observations made above. They also suggest that ECM enrichment protocols should be amenable to relative quantification if several types of skin samples are to be compared.

#### 3.5. The matrisome of healthy adult mouse skin comprises 95 core matrisome proteins and 141 associated matrisome proteins

All the data obtained with the four enrichment protocols and the total lysates were finally combined to generate a list of proteins defining the mouse skin matrisome in homeostasis conditions (Tables 2 and 3). Among the 1112 components presently described in the mouse matrisome (<http://matrisomeproject.mit.edu/>), 236 were identified in the skin of healthy adult mice (21% of all matrisome content). These comprised 308 proteoforms, with some redundancy due to protein isoforms and variants (Supplementary Table 2).

Ninety-five core matrisome proteins were detected in our experiments, among which 30 are collagen chains (68% of all collagen chains), 53 are glycoproteins (27% of core matrisome glycoproteins) and 12 are proteoglycans (33% of core matrisome proteoglycans) (Table 2). Among collagens, some of them are well-described skin components such as collagens I, III, V, VI, XII and XIV found in the dermis, collagens XV and XVI in the dermis and DEJ, collagens IV, VII, XVII and XVIII in the DEJ [3,6,50]. Collagen XIII, a normal component of the epidermis and DEJ was not detected, possibly because of its cell-membrane localization and the fact that most protocols were not designed to efficiently retrieve membrane proteins. Interestingly also, seven additional collagen chains were observed. Among them, the presence of fibrillar collagens II and XI was particularly surprising because these molecules have mainly been reported in cartilage and vitreous body [43,51]. However, they were also identified by LC-MS/MS in decellularized human skin [28] and could be low-abundance collagens of normal skin. Collagen XXVII is also found in adult cartilage and in fetal dermis [52]. Collagen XX remains poorly characterized in terms of tissue distribution and functions, but its expression is currently supposed to be limited to fetal life [53]. Collagen XXVIII [54] as well as transmembrane collagen XXV [55] are commonly associated to the nervous system and their presence could be explained by the presence of nerve endings in skin.

As lysine hydroxylation is a well-known modification of collagens, we also tested whether adding "lysine oxidation" as variable modification could impact collagen detection in the insoluble fractions of all protocols. This led to the identification of two additional collagens, collagen XXII (one peptide identified in the three mice with all protocols except QDSP) and collagen XXIII (one peptide identified in one mouse with the chemical decellularization protocol) (Supplementary Table 4). Besides, oxidized lysine residues could be attributed to around 15% of unique peptides in fibrillar collagens I, II, III and V and were also found in collagens IV ( $\alpha 1$ - $\alpha 3$ ), VI ( $\alpha 2$ ), VII and XVI with more variability. However, the inclusion of this modification had a rather modest effect on the total number of unique peptides identified for these insoluble collagens (less than 10% on average). Moreover, some proteins for which this modification has not been previously reported seem to be incorrectly identified as carrying oxidized lysines, both among



**Fig. 4.** Matrixome protein distribution in the biological replicates and extraction fractions. **A:** Global comparison of matrixome proteins detected in the biological replicates. **B:** Core matrixome proteins identified in the extraction fractions analyzed by MS/MS and in total lysates. **C:** Associated matrixome proteins analyzed by MS/MS and in the total lysates. The percentage of proteins detected in at least two replicates is indicated by numbers above the bars for each fraction. The NaCl fraction is the same for the chemical and enzymatic decellularizations but was represented twice to give a complete overview of the various protocols. Ins = insoluble fraction.

matrixome (biglycan and serpin6a; Supplementary Table 4) and non-matrixome components (e.g. keratins 6a and 13, myosins 1 and 4, nebulin). Therefore, we chose not to include this modification systematically.

Regarding the glycoproteins and proteoglycans identified in mouse skin, the key components of the dermal compartment [4,8,56] such as fibronectin, elastin and its partners in elastic fibers (fibrillin-1, fibulins 2 and 5, microfibrillar associated proteins 2, 4 and 5 and latent TGF-β

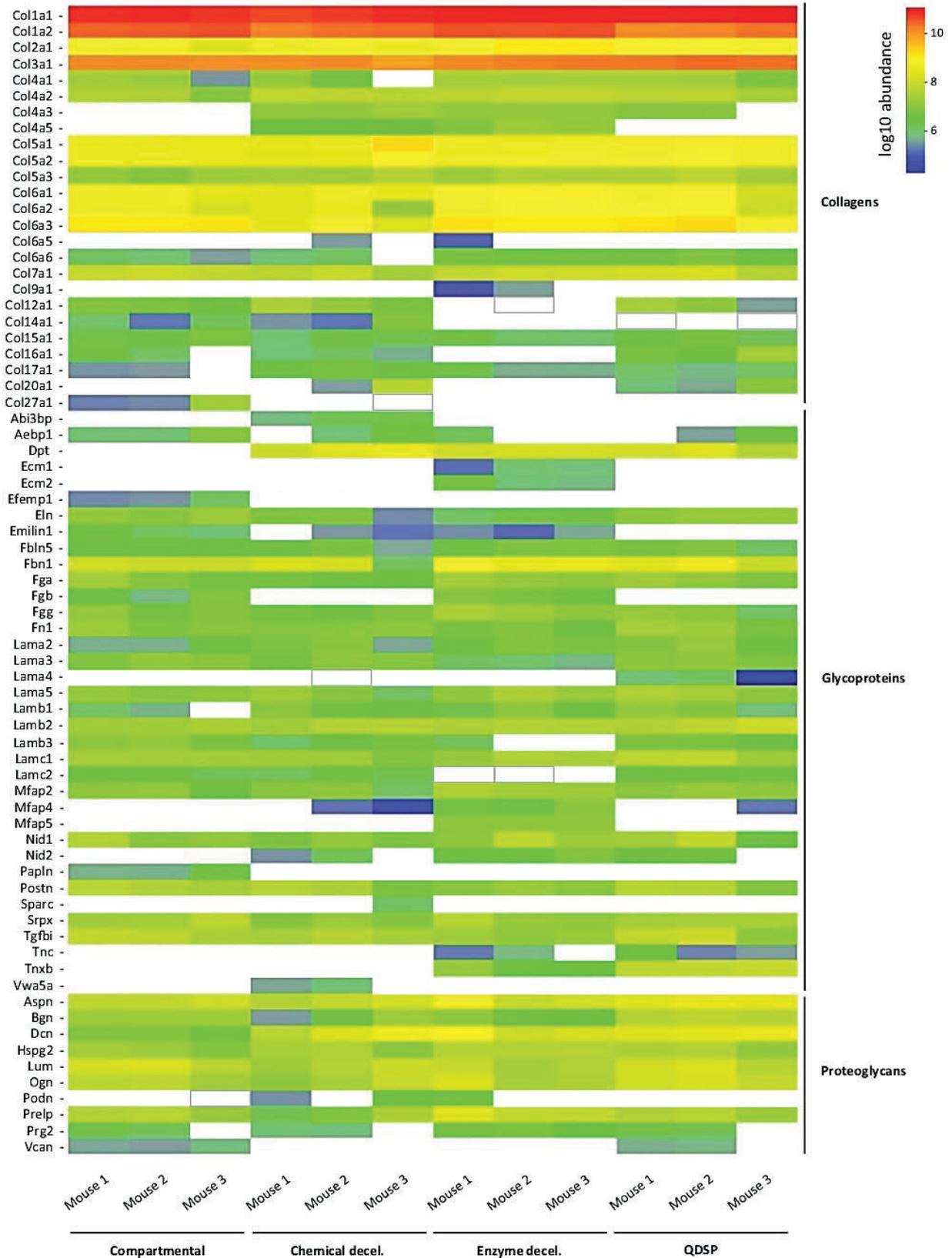


Fig. 5. Heatmap representation of the abundance of core matrisome proteins in the insoluble fractions of ECM enrichment protocols. Empty rectangles indicate proteins which were detected but could not be quantified.

**Table 2**

Core matrisome components detected in mouse skin. Numbers of proteins are specified in parentheses for each subcategory of core matrisome components. \*: two mouse gene names (*Tnxb* and *Tnxb*) are associated with tenascin-X in the Uniprot database. *Tnxb* was kept and used for all tenascin-X proteoforms. # indicates genes for which only unreviewed protein sequences are available in Uniprot database.

Core matrisome (95)			Collagens (30)		
Glycoproteins (53)			Collagens (30)		
Gene name	Protein name	Uniprot IDs	Gene name	Protein name	Uniprot IDs
<i>Abi3bp</i> #	ABI gene family, member 3 (NESH) binding protein	A0A338P6W9; Q8BKQ1	<i>Col1a1</i>	Collagen alpha-1 (XI)	Q61245
<i>Adipoq</i>	Adiponectin	E9PWU4; Q8BRW2	<i>Col1a2</i>	Collagen alpha-2(XI)	Q64739
<i>Aebp1</i>	Adipocyte enhancer-binding protein-1	Q640N1	<i>Col12a1</i>	Collagen alpha-1 (XII)	Q60847
<i>Agri</i>	Agri	A2ASQ1; Q91XS3; Z4YJS5	<i>Col14a1</i>	Collagen alpha-1 (XIV)	Q80X19
<i>AW551984</i>	Expressed sequence AW551984	Q6P8K4	<i>Col15a1</i>	Collagen alpha-1 (XV)	O35206
<i>Coch</i>	Cochlin	Q62507	<i>Col16a1</i>	Collagen alpha-1 (XVII)	Q8CIF9
<i>Crispld2</i>	Cysteine-rich secretory protein LCCL domain containing-2	Q8BZQ2	<i>Col17a1</i>	Collagen alpha-1 (XVIII)	Q07563
<i>Cthrc1</i>	Collagen triple helix repeat-containing protein-1	A2RRY6	<i>Col18a1</i>	Collagen alpha-1 (XVIII)	E9QPX1
<i>Dpt</i>	Dermatopontin	Q9QZZ6	<i>Col1a1</i>	Collagen alpha-1 (I)	P11087; Q60785
<i>Ecm1</i>	Extracellular matrix protein 1	F8WI14	<i>Col1a2</i>	Collagen alpha-2 (I)	Q01149; Q3TP88; Q7TN67; Q91VL4
<i>Ecm2</i>	Extracellular matrix protein 2	Q5FW85	<i>Col20a1</i>	Collagen alpha-1 (XX)	A0A2K6EDL8
<i>Efemp1</i>	EGF-containing fibulin-like extracellular matrix protein 1	Q8K0J4; Q8BPB5	<i>Col25a1</i>	Collagen alpha-1 (XXV)	A0A5F8MPI1
<i>Eln</i>	Elastin	Q9ESZ9	<i>Col27a1</i>	Collagen alpha-1 (XXVII)	Q5QNQ9; V9GXJ7
<i>Emilin1</i>	EMILIN-1	Q3UUU0; Q99K41	<i>Col28a1</i>	Collagen alpha-1 (XXVIII)	Q2UY11
<i>Fbln2</i>	Fibulin-2	Q3TGL4	<i>Col2a1</i>	Collagen alpha-1 (II)	P28481
<i>Fbln5</i>	Fibulin-5	A0A1Y7VJW9	<i>Col3a1</i>	Collagen alpha-1 (III)	A0A087WPJ5; Q7TT32; Q8BJU6
<i>Fbn1</i>	Fibrillin-1	O88840; Q61554	<i>Col4a1</i>	Collagen alpha-1 (IV)	P02463
<i>Fga</i>	Fibrinogen alpha chain	E9PV24	<i>Col4a2</i>	Collagen alpha-2 (IV)	B2RQK8
<i>Fgb</i>	Fibrinogen beta chain	Q8K0E8	<i>Col4a3</i>	Collagen alpha-3 (IV)	Q9QZS0
<i>Fgg</i>	Fibrinogen gamma chain	Q3UEM7	<i>Col4a5</i>	Collagen alpha-5 (IV)	Q63ZV6; Q80V57
<i>Fn1</i>	Fibronectin	A0A087WS56; Q3TCF1	<i>Col5a1</i>	Collagen alpha-1 (V)	B1AWB9
<i>Ints14</i>	Integrator complex subunit-14	Q8R3P6	<i>Col5a2</i>	Collagen alpha-2 (V)	Q3U962
<i>Lama2</i>	Laminin subunit alpha-2	Q60675	<i>Col5a3</i> #	Collagen alpha-3 (V)	Q920S4; Q9JLI2
<i>Lama3</i>	Laminin subunit alpha-3	Q61789	<i>Col6a1</i>	Collagen alpha-1 (VI)	Q04857
<i>Lama4</i>	Laminin subunit alpha-4	P97927; Q91VVO	<i>Col6a2</i>	Collagen alpha-2 (VI)	Q02788
<i>Lama5</i>	Laminin subunit alpha-5	Q61001	<i>Col6a3</i> #	Collagen alpha-3 (VI)	A0A087WS16; E9PWQ3
<i>Lamb1</i>	Laminin subunit beta-1	B9EK80	<i>Col6a5</i>	Collagen alpha-5 (VI)	A0A140T8T7
<i>Lamb2</i>	Laminin subunit beta-2	Q61292	<i>Col6a6</i>	Collagen alpha-6 (VI)	E9Q6A6; B9EK79
<i>Lamb3</i>	Laminin subunit beta-3	Q61087	<i>Col7a1</i>	Collagen alpha-1 (VII)	Q63870
<i>Lamc1</i>	Laminin subunit gamma-1	F8VQJ3	<i>Col9a1</i>	Collagen alpha-1 (IX)	Q8BSQ4
<i>Lamc2</i>	Laminin subunit gamma-2	G5E874	<b>Proteoglycans (12)</b>		
<i>Lrg1</i>	Leucine-rich alpha-2-glycoprotein	Q91XL1	<b>Gene name</b>	<b>Protein name</b>	<b>Uniprot IDs</b>
<i>Ltbp4</i>	Latent-transforming growth factor beta-binding protein-4	D3Z598	<i>Aspn</i>	Asporin	Q99MQ4
<i>Mfap2</i>	Microfibrillar-associated protein-2	Q99PM0	<i>Bgn</i>	Biglycan	P28653; Q3TAF9; Q7TMW3
<i>Mfap4</i>	Microfibrillar-associated protein-4	Q9D1H9	<i>Cspg4</i>	Chondroitin sulfate proteoglycan 4	Q8VHY0
<i>Mfap5</i>	Microfibrillar-associated protein-5	D3Z4D7	<i>Dcn</i>	Decorin	Q3TSV1; Q3UKR1
<i>Mfge8</i>	Lactadherin	P21956	<i>Fmod</i>	Fibromodulin	Q543D2
<i>Nid1</i>	Nidogen-1	P10493	<i>Hspg2</i>	Perlecan	B1B0C7; E9PZ16
<i>Nid2</i>	Nidogen-2	Q3US45; Q8C6Z2; Q8R5G0	<i>Lum</i>	Lumican	P51885
<i>Papln</i>	Papilin	B7ZN28	<i>Ogn</i>	Mimecan	Q543C5
<i>Pcolce</i>	Procollagen C-endopeptidase enhancer-1	A0A0G2JE29; Q3UN82; Q61398	<i>Podn</i>	Podocan	Q6P3D8; Q7TQ62
<i>Postn</i>	Periostin	A0A097BW21; Q62009	<i>Prelp</i>	Prolargin	Q8CAZ9; Q9JK53
<i>Sparc</i>	Basement-membrane protein-40	A0A1L1SSH9	<i>Prg2</i>	Bone marrow proteoglycan	Q61878
<i>SrpX</i>	Sushi repeat-containing protein SRPX	Q9R0M3	<i>Vcan</i>	Versican	Q8BS97
<i>SrpX2</i>	Sushi repeat-containing protein SRPX2	Q8R054			
<i>Tgfb1</i>	Transforming growth factor-beta-induced protein ig-h3	A1L353; Q3UBG4; Q3TAY9			
<i>Thbs1</i>	Thrombospondin-1	Q3TR40			
<i>Tinag1</i>	Tubulointerstitial nephritis antigen-like	Q4FJX7			
<i>Tnc</i>	Tenascin-C	Q80YX1			
<i>Tnxb</i> */#	Tenascin-X	A0A5F8MPH8; O35452; O54796			
<i>Vtn</i>	Vitronectin	P29788			
<i>Vwa1</i>	Von Willebrand factor A domain-containing protein-1	Q8R2Z5			
<i>Vwa5a</i>	Von Willebrand factor A domain-containing protein-5A	Q99KC8			



Table 3 (continued)

Associated matrisome (141)					
ECM regulators (84)					
Gene name	Protein name	Uniprot IDs	Gene name	Protein name	Uniprot IDs
<i>Cd209b</i>	CD209 antigen-like protein B	Q8CJ85	<i>Mstn</i>	Myostatin	Q3ZAS9
<i>Cd209d</i>	CD209 antigen-like protein D	Q3UW57	<i>Rptn</i>	Repetin	P97347
<i>Clec10a</i>	C-type lectin domain family-10 member A	A2CF65	<i>S100a1</i>	Protein S100-A1	Q91V77; Q9JL08
<i>Clec11a</i>	C-type lectin domain family-11 member A	O88200	<i>S100a10</i>	Protein S100-A10	Q3UF30
<i>Clec3b</i>	Tetranectin	Q8CFZ6	<i>S100a11</i>	Protein S100-A11	P50543
<i>Gpc1</i>	Glypican-1	Q3U3U; Q9QZF2	<i>S100a13</i>	Protein S100-A13	P97352
<i>Gpc4</i>	Glypican-4	P51655	<i>S100a14</i>	Protein S100-A14	Q9D2Q8
<i>Gpc6</i>	Glypican-6	Q8R3X6	<i>S100a15a</i>	Protein S100-A15	Q6S5I3
<i>Hpx</i>	Hemopexin	Q91X72	<i>S100a16</i>	Protein S100-A16	D3Z2Y6; Q9D708
<i>Lgals1</i>	Galectin-1	P16045	<i>S100a3</i>	Protein S100-A3	P62818
<i>Lgals3</i>	Galectin-3	Q3V471	<i>S100a4</i>	Protein S100-A4	A0A0G2JGD2
<i>Lgals7</i>	Galectin-7	Q9CRB1	<i>S100a6</i>	Protein S100-A6	P14069
<i>Lgalsl</i>	Galectin-related protein	Q8VED9	<i>S100b</i>	Protein S100-B	P50114
<i>Lman1</i>	Protein ERGIC-53	Q9D0F3	<i>Tchh<sup>#</sup></i>	Trichohyalin	A0A0B4J1F9; Q3V203
<i>Pldc2</i>	Plexin domain-containing protein-2	B1AY86	<i>Tchhl1</i>	Trichohyalin-like protein-1	Q9D3P1
<i>Plexn2</i>	Plexin B2	B2RXS4			
<i>Sdc4</i>	Syndecan-4	Q3UKZ1			

binding protein 4), small-leucine rich proteoglycans (asporin, biglycan, decorin, dermatopontin, fibromodulin, lumican, mimecan, PRELP and podocan) as well as numerous proteins known to regulate ECM organization and cell-matrix interactions (e.g. SPARC, tenascins C and X, thrombospondin-1, TGF- $\beta$ -induced protein/BIGH3, procollagen C-proteinase enhancer-1) were detected. The DEJ is also very well represented at the level of glycoproteins and proteoglycans [57,58] with all chains of laminins 332 and 511, nidogens 1 and 2 and perlecan being identified with most ECM enrichment protocols (laminin alpha 4 is also found but is more specific to vascular basement membranes). This suggests that our skin preparations are well adapted to study specific processes affecting the DEJ and that there is no need to separate the dermis from the epidermis to obtain a good coverage of the skin structural ECM components. It should also be noted here that a possible optimization applicable to all methods, especially if accurate quantification of proteoglycans and glycoproteins is desired, would be to combine several deglycosylation enzymes (in addition to the classical PNGaseF) to remove glycosaminoglycan (GAG) chains and other O-glycosylations, as described by others [32]. However, the systematic integration of a GAG removal step in ECM-centered analyses must be validated first as a recent study [37] suggested that it may decrease the number of PSMs for collagens and glycoproteins.

In the associated matrisome, 141 components were detected (Table 3) with the following distribution: 84 ECM regulators, 30 ECM-affiliated and 27 secreted factors. Proteases are identified as important members of the ECM regulators with several cathepsins, prothrombin and coagulation factors X and XII, plasminogen, trypsin 1 and 5, ADAM10, MMP-2 and -3. Protease inhibitors are also generally well detected with numerous serpins, cystatins, the broad-spectrum inhibitor pregnancy zone protein (or  $\alpha$ 2-macroglobulin) but only one tissue inhibitor of MMP (TIMP-3) is found. Other enzymes involved in ECM biosynthesis such as prolyl and lysyl hydroxylases and cross-linking enzymes (lysyl oxidase homolog-1 and transglutaminases) are also part of the list but some other prominent players are still missing (procollagen N- and C-proteinases, lysyl oxidase), possibly due to their low abundance in healthy skin [59].

The category of secreted factors includes some cytokines (CCL27a, IL1 family member 10), growth factors (myostatin, angiopoietin-related proteins 1 and 2), proteins from the epidermal cornified envelope (filaggrins, cornulin, hornerin and repetin) and additional markers of epithelial differentiation (S100 proteins, trichohyalin and trichohyalin-like protein-1) [60,61]. The relative paucity of this matrisome category

can be explained by the low expression level of secreted factors in homeostatic conditions, making their detection challenging. Finally, ECM-affiliated proteins are mostly annexins but several proteins normally found at the cell surface (CD209, glypican-1, -4 and -6, syndecan-4) are also detected, suggesting that their ectodomains can be shed in healthy skin.

When our results on mouse skin are compared to the recent reports describing the matrisome of the healthy human skin dermal compartment (Supplementary Table 5), comparable total numbers of matrisome proteins were detected with a similar enzymatic decellularization method using PLA2 (274 in human when the results of two studies are combined [28,31]). The chemical decellularization method was also used to analyze human dermal samples [30] and led to the identification of less matrisome proteins (99 annotated human matrisome proteins, Supplementary Table 5) but only the NaCl and GuHCl extracts were analyzed and different criteria to define ECM proteins were used, explaining the relatively low number of associated matrisome components found in this list. The most comprehensive description of human skin ECM is however given by the work of Dyring-Andersen and colleagues [62] who provided the first proteomic atlas of four skin layers, the outer epidermis, the inner epidermis, the dermis and the subcutis. Specific ECM enrichment was not the specific goal of this study but extensive peptide fractionation allowed the identification of more than 10,000 proteins. The matrisome annotation of the proteins found in the epidermal and dermal compartments yields a list of 437 proteins (Supplementary Table 5) with around 38% of core matrisome proteins (to be compared with 40% in our mouse study) and 62% of associated matrisome proteins (60% in our study). The comparison of our mouse dataset with all these previous studies made on human skin also indicates that about 84% of the skin matrisome proteins detected in mouse match with the human ones. The differences mainly lie in species-specific genetic variants, especially for the large serpin family [63]. However, there are some differences which could be more significant for skin biology. For example, if the overall number of collagen chains identified is relatively stable between species and studies, some collagen chains are only found in mouse skin (Col9a1, Col11a2, Col20a1) while some other chains are specifically detected in human skin (COL4A4, COL4A6, COL8A1, COL8A2, COL10A1, COL21A1). Whether these differences can substantially impact cutaneous functions remains to be determined.

The rat dermal matrisome was also reported in a previous study [28] but, for unclear reasons, a comparatively low number of 102 matrisome proteins was identified with a poor coverage of the associated matrisome

(42 proteins compared to 142 for mouse skin, Supplementary Table 5). Some technical differences in reagent concentrations or sample processing could explain this apparent discrepancy and we expect that the detailed protocols and analyses provided in the present study will help users to set up reproducible and efficient skin ECM enrichment protocols.

#### 4. Conclusion

Our comparison of four ECM enrichment protocols indicates that skin matrisome proteins exhibit diverse solubility profiles and that there is a real benefit in terms of matrisome coverage to analyze some of the liquid extraction fractions, not only the final insoluble pellet. If applied to the comparison of healthy and diseased skins, this could reveal subtle changes in ECM protein solubility, induced by alterations in tissue composition and stiffness/cross-linking, that would have been missed in a more global comparison of total protein amounts. Between the three methods which led to the highest numbers of identified proteins (chemical and enzymatic decellularizations, QDSP), the decellularization methods have the double advantage of requiring a lower number of MS injections compared to QDSP (3 instead of 4) and of showing a better overall reproducibility when biological replicates are compared. If PLA2 is relatively expensive, it is not the case for the reagents used in the chemical decellularization method that we recommend for routine matrisome characterizations. Importantly, this work also enabled us to report the first comprehensive description of the mouse skin matrisome. Two hundred and thirty-six matrisome proteins were detected, including the main structural components of skin ECM and several associated factors, with 84% of these components which were also found in the human skin matrisome showing a good conservation between the two species. This work provides a strong basis for the future description of ECM (dys)regulations occurring in skin upon challenge or in various disease contexts.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jprot.2021.104397>.

#### Funding

This work was supported by the French and German Research Agencies (grants ANR-18-CE92-0035 to CM and DFG NY90/5-1 to AN), the CNRS and the University of Lyon. We also thank ITMO Cancer AVIESAN for the funding of the Orbitrap mass spectrometer.

#### Declaration of Competing Interest

Authors declare no conflict of interest.

#### Acknowledgements

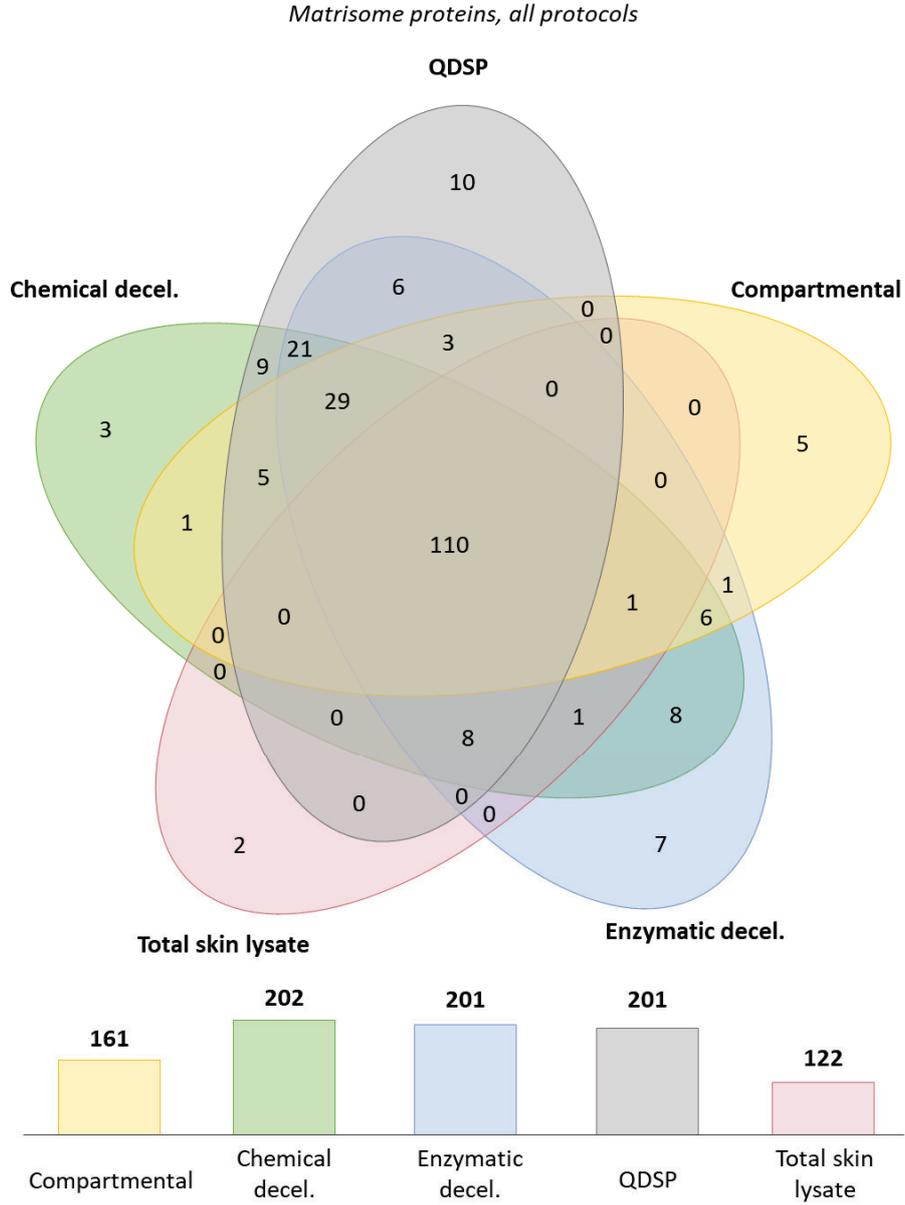
We thank Christelle Bonod, Nicolas Picard and Laura Prigent (LBTI, University of Lyon, France) for their help with the mouse work and skin sample collection.

#### References

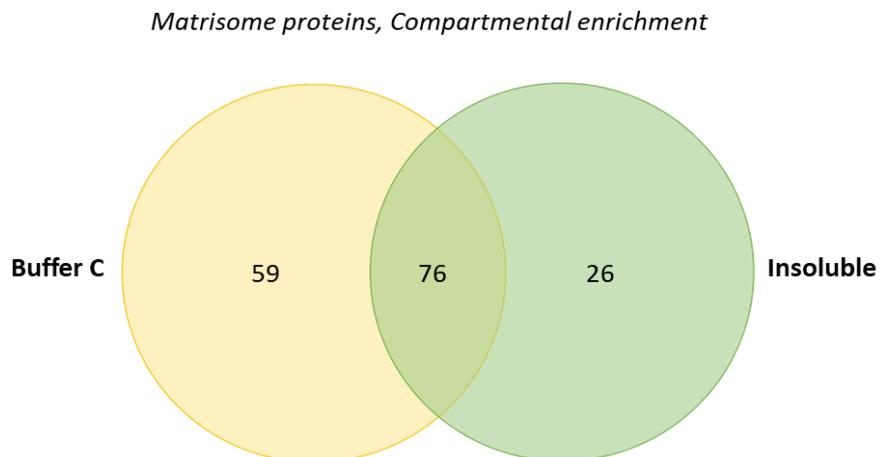
1. E. Proksch, J.M. Brandner, J.M. Jensen, The skin: an indispensable barrier, *Exp. Dermatol.* 17 (2008) 1063–1072, <https://doi.org/10.1111/j.1600-0625.2008.00786.x>.
2. D. Breitkreutz, N. Mirancea, R. Nischt, Basement membranes in skin: unique matrix structures with diverse functions? *Histochem. Cell Biol.* 132 (2009) 1–10, <https://doi.org/10.1007/s00418-009-0586-0>.
3. E. Roig-Rosello, P. Rousselle, The human epidermal basement membrane: a shaped and cell instructive platform that aging slowly alters, *Biomolecules* 10 (2020) 1607, <https://doi.org/10.3390/biom10121607>.
4. T. Krieg, M. Aumailley, The extracellular matrix of the dermis: flexible structures with dynamic functions, *Exp. Dermatol.* 20 (2011) 689–695, <https://doi.org/10.1111/j.1600-0625.2011.01313.x>.
5. J. Myllyharju, K.I. Kivirikko, Collagens, modifying enzymes and their mutations in humans, flies and worms, *Trends Genet.* 20 (2004) 33–43, <https://doi.org/10.1016/j.tig.2003.11.004>.
6. S. Ricard-Blum, The collagen family, *Cold Spring Harb. Perspect. Biol.* 3 (2011) a004978, <https://doi.org/10.1101/cshperspect.a004978>.
7. R.V. Iozzo, L. Schaefer, Proteoglycan form and function: a comprehensive nomenclature of proteoglycans, *Matrix Biol.* 42 (2015) 11–55, <https://doi.org/10.1016/j.matbio.2015.02.003>.
8. A. Nyström, L. Bruckner-Tuderman, Matrix molecules and skin biology, *Semin. Cell Dev. Biol.* 89 (2019) 136–146, <https://doi.org/10.1016/j.semcdb.2018.07.025>.
9. G.S. Schultz, A. Wysocki, Interactions between extracellular matrix and growth factors in wound healing, *Wound Repair Regen.* 17 (2009) 153–162, <https://doi.org/10.1111/j.1524-475X.2009.00466.x>.
10. J. Zeng-Brouwers, S. Pandey, J. Trebicka, M. Wygrecka, L. Schaefer, Communications via the small leucine-rich proteoglycans: molecular specificity in inflammation and autoimmune diseases, *J. Histochem. Cytochem.* 68 (2020) 887–906, <https://doi.org/10.1369/0022155420930303>.
11. J. Halper, M. Kjaer, Basic components of connective tissues and extracellular matrix: elastin, fibrillin, fibulins, fibrinogen, fibronectin, laminin and thrombospondins, *Adv. Exp. Med. Biol.* 802 (2014) 31–47, [https://doi.org/10.1007/978-94-007-7893-1\\_3](https://doi.org/10.1007/978-94-007-7893-1_3).
12. N.K. Karamanos, A.D. Theocharis, T. Neill, R.V. Iozzo, Matrix modeling and remodeling: a biological interplay regulating tissue homeostasis and diseases, *Matrix Biol.* 75–76 (2019) 1–11, <https://doi.org/10.1016/j.matbio.2018.08.007>.
13. S. Ricard-Blum, S.D. Vallet, Proteases decode the extracellular matrix cryptome, *Biochimie.* 122 (2016) 300–313, <https://doi.org/10.1016/j.biochi.2015.09.016>.
14. C. Moali, D.J. Hulmes, Extracellular and cell surface proteases in wound healing: new players are still emerging, *Eur. J. Dermatol.* 19 (2009) 552–564, <https://doi.org/10.1684/ejd.2009.0770>.
15. J. Filmus, M. Capurro, J. Rast, Glypicans, *Genome Biol.* 9 (2008) 224, <https://doi.org/10.1186/gb-2008-9-5-224>.
16. S. Mii, A. Enomoto, Y. Shiraki, T. Taki, Y. Murakumo, M. Takahashi, CD109: a multifunctional GPI-anchored protein with key roles in tumor progression and physiological homeostasis, *Pathol. Int.* 69 (2019) 249–259, <https://doi.org/10.1111/pin.12798>.
17. J. Weisz, V.N. Uversky, Zooming into the dark side of human annexin-S100 complexes: dynamic alliance of flexible partners, *Int. J. Mol. Sci.* 21 (2020), <https://doi.org/10.3390/ijms21165879>.
18. A. Naba, K.R. Clauser, S. Hoersch, H. Liu, S.A. Carr, R.O. Hynes, The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices, *Mol. Cell. Proteomics* 11 (4) (2012), <https://doi.org/10.1074/mcp.M111.014647>. M111.014647.
19. A. Naba, K.R. Clauser, H. Ding, C.A. Whittaker, S.A. Carr, R.O. Hynes, The extracellular matrix: tools and insights for the “omics” era, *Matrix Biol.* 49 (2016) 10–24, <https://doi.org/10.1016/j.matbio.2015.06.003>.
20. A. Naba, K.R. Clauser, R.O. Hynes, Enrichment of extracellular matrix proteins from tissues and digestion into peptides for mass spectrometry analysis, *J. Vis. Exp.* (2015), e53057, <https://doi.org/10.3791/53057>.
21. A. Naba, K.R. Clauser, D.R. Mani, S.A. Carr, R.O. Hynes, Quantitative proteomic profiling of the extracellular matrix of pancreatic islets during the angiogenic switch and insulinoma progression, *Sci. Rep.* 7 (2017) 40495, <https://doi.org/10.1038/srep40495>.
22. A. Naba, K.R. Clauser, J.M. Lamar, S.A. Carr, R.O. Hynes, Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters, *Elife* 3 (2014), e01308, <https://doi.org/10.7554/eLife.01308>.
23. V. Gocheva, A. Naba, A. Bhutkar, T. Guardia, K.M. Miller, C.M. Li, T.L. Dayton, F. J. Sanchez-Rivera, C. Kim-Kiselak, N. Jalkhiani, M.M. Winslow, A. Del Rosario, R. O. Hynes, T. Jacks, Quantitative proteomics identify Tenascin-C as a promoter of lung cancer progression and contributor to a signature prognostic of patient survival, *Proc. Natl. Acad. Sci. U. S. A.* 114 (2017) E5625–E5634, <https://doi.org/10.1073/pnas.1707054114>.
24. A. Naba, K.R. Clauser, C.A. Whittaker, S.A. Carr, K.K. Tanabe, R.O. Hynes, Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver, *BMC Cancer* 14 (2014) 518, <https://doi.org/10.1186/1471-2407-14-518>.
25. H.B. Schiller, I.E. Fernandez, G. Burgstaller, C. Schaab, R.A. Scheltema, T. Schwarzmayr, T.M. Strom, O. Eickelberg, M. Mann, Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair, *Mol. Syst. Biol.* 11 (2015) 819, <https://doi.org/10.15252/msb.20156123>.
26. A. Didangelos, X. Yin, K. Mandal, M. Baumert, M. Jahangiri, M. Mayr, Proteomics characterization of extracellular space components in the human aorta, *Mol. Cell. Proteomics* 9 (2010) 2048–2062, <https://doi.org/10.1074/mcp.M110.001693>.
27. A.E. Mayorca-Guiliani, C.D. Madsen, T.R. Cox, E.R. Horton, F.A. Venning, J. T. Erler, ISDoT: in situ decellularization of tissues for high-resolution imaging and proteomic analysis of native extracellular matrix, *Nat. Med.* 23 (2017) 890–898, <https://doi.org/10.1038/nm.4352>.
28. B. Liu, S. Zhang, W. Wang, Z. Yun, L. Lv, M. Chai, Z. Wu, Y. Zhu, J. Ma, L. Leng, Matrisome provides a supportive microenvironment for skin functions of diverse species, *ACS Biomater. Sci. Eng.* 6 (2020) 5720–5733, <https://doi.org/10.1021/acsbomaterials.0c00479>.
29. M. Dussoyer, A. Michopoulou, P. Rousselle, Decellularized scaffolds for skin repair and regeneration, *Appl. Sci.* 10 (2020) 3435, <https://doi.org/10.3390/app10103435>.
30. J. Barallobre-Barreiro, E. Woods, R.E. Bell, J.A. Easton, C. Hobbs, M. Eager, F. Baig, A.M. Ross, R. Mallipeddi, B. Powell, M. Soldin, M. Mayr, T.J. Shaw, Cartilage-like composition of keloid scar extracellular matrix suggests fibroblast mis-

- differentiation in disease, *Matrix Biol. Plus.* 4 (2019) 100016, <https://doi.org/10.1016/j.mbplus.2019.100016>.
- [31] S. Zhang, B. Liu, W. Wang, L. Lv, D. Gao, M. Chai, M. Li, Z. Wu, Y. Zhu, J. Ma, L. Leng, The "Matrisome" reveals the characterization of skin keloid microenvironment, *FASEB J.* 35 (2021), e21237, <https://doi.org/10.1096/fj.202001660RR>.
- [32] J. Barallobre-Barreiro, F. Baig, M. Fava, X. Yin, M. Mayr, Glycoproteomics of the extracellular matrix: a method for intact glycopeptide analysis using mass spectrometry, *J. Vis. Exp.* 55674 (2017), <https://doi.org/10.3791/55674>.
- [33] C.W. Wong, C.F. LeGrand, B.F. Kinnear, R.M. Sobota, R. Ramalingam, D.E. Dye, M. Raghunath, E.B. Lane, D.R. Coombe, In vitro expansion of keratinocytes on human dermal fibroblast-derived matrix retains their stem-like characteristics, *Sci. Rep.* 9 (2019) 18561, <https://doi.org/10.1038/s41598-019-54793-9>.
- [34] Z. Wu, Y. Zhou, N. Li, M. Huang, H. Duan, J. Ge, P. Xiang, Z. Wang, The use of phospholipase A(2) to prepare acellular porcine corneal stroma as a tissue engineering scaffold, *Biomaterials.* 30 (2009) 3513–3522, <https://doi.org/10.1016/j.biomaterials.2009.03.003>.
- [35] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Perez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarmuczak, T. Ternent, A. Brazma, J.A. Vizcaino, The PRIDE database and related tools and resources in 2019: improving support for quantification data, *Nucleic Acids Res.* 47 (2019) D442–D450, <https://doi.org/10.1093/nar/gky1106>.
- [36] L. Nguyen, V. Brun, F. Combes, V. Loux, Y. Vandenbrouck, Designing an in silico strategy to select tissue-leakage biomarkers using the galaxy framework, *Methods Mol. Biol.* 1959 (2019) 275–289, [https://doi.org/10.1007/978-1-4939-9164-8\\_18](https://doi.org/10.1007/978-1-4939-9164-8_18).
- [37] M.C. McCabe, L.R. Schmitt, R.C. Hill, M. Dzieciatkowska, M. Maslanka, W. F. Daamen, T.H. van Kuppevelt, D.J. Hof, K.C. Hansen, Evaluation and refinement of sample preparation methods for extracellular matrix proteome coverage, *Mol. Cell. Proteomics* 20 (2021) 100079, <https://doi.org/10.1016/j.mcp.2021.100079>.
- [38] J. Barallobre-Barreiro, A. Didangelos, F.A. Schoendube, I. Drodzov, X. Yin, M. Fernandez-Caggiano, P. Willeit, V.O. Puntmann, G. Aldama-Lopez, A.M. Shah, N. Domenech, M. Mayr, Proteomics analysis of cardiac extracellular matrix remodeling in a porcine model of ischemia/reperfusion injury, *Circulation.* 125 (2012) 789–802, <https://doi.org/10.1161/CIRCULATIONAHA.111.056952>.
- [39] J. Barallobre-Barreiro, R. Oklu, M. Lynch, M. Fava, F. Baig, X. Yin, T. Barwari, D. N. Potier, H. Albadawi, M. Jahangiri, K.E. Porter, M.T. Watkins, S. Misra, J. Stoughton, M. Mayr, Extracellular matrix remodeling in response to venous hypertension: proteomics of human varicose veins, *Cardiovasc. Res.* 110 (2016) 419–430, <https://doi.org/10.1093/cvr/cvw075>.
- [40] L.E. de Castro Bras, T.A. Ramirez, K.Y. DeLeon-Pennell, Y.A. Chiao, Y. Ma, Q. Dai, G.V. Halade, K. Hakala, S.T. Weintraub, M.L. Lindsey, Texas 3-step decellularization protocol: looking at the cardiac extracellular matrix, *J. Proteome Res.* 12 (2013) 43–52, <https://doi.org/10.1021/jpr.2013.005004>.
- [41] L. Mayer, R. Pechlaner, J. Barallobre-Barreiro, C. Boehme, T. Toell, M. Lynch, X. Yin, J. Willeit, E.R. Gizewski, P. Perco, G. Ratzinger, S. Kiechl, M. Mayr, M. Knoflach, Extracellular matrix protein signature of recurrent spontaneous cervical artery dissection, *Neurology* 95 (2020) e2047–e2055, <https://doi.org/10.1212/WNL.00000000000010710>.
- [42] J. Kjell, J. Fischer-Sternjak, A.J. Thompson, C. Friess, M.J. Sticco, F. Salinas, J. Cox, D.C. Martinelli, J. Ninkovic, K. Franze, H.B. Schiller, M. Gotz, Defining the adult neural stem cell niche proteome identifies key regulators of adult neurogenesis, *Cell Stem Cell* 26 (2020) 277–293, <https://doi.org/10.1016/j.stem.2020.01.002>.
- [43] J. Bella, D.J. Hulmes, Fibrillar collagens, *Subcell. Biochem.* 82 (2017) 457–490, [https://doi.org/10.1007/978-3-319-49674-0\\_14](https://doi.org/10.1007/978-3-319-49674-0_14).
- [44] S.E. Heumuller, M. Talantikite, M. Napoli, J. Armengaud, M. Morgelin, U. Hartmann, G. Sengle, M. Paulsson, C. Moali, R. Wagener, C-terminal proteolysis of the collagen VI alpha3 chain by BMP-1 and proprotein convertase(s) releases endotrophin in fragments of different sizes, *J. Biol. Chem.* 294 (2019) 13769–13780, <https://doi.org/10.1074/jbc.RA119.008641>.
- [45] A. Rattenholl, W.N. Pappano, M. Koch, D.R. Keene, K.E. Kadler, T. Sasaki, R. Timpl, R.E. Burgeson, D.S. Greenspan, L. Bruckner-Tuderman, Proteinases of the bone morphogenetic protein-1 family convert procollagen VII to mature anchoring fibril collagen, *J. Biol. Chem.* 277 (2002) 26372–26378, <https://doi.org/10.1074/jbc.M203247200>.
- [46] S. Vadon-Le Goff, D.J. Hulmes, C. Moali, BMP-1/tolloid-like proteinases synchronize matrix assembly with growth factor activation to promote morphogenesis and tissue remodeling, *Matrix Biol.* 44–46 (2015) 14–23, <https://doi.org/10.1016/j.matbio.2015.02.006>.
- [47] C. Anastasi, P. Rousselle, M. Talantikite, A. Tessier, C. Cluzel, A. Bachmann, N. Mariano, M. Dussoyer, L.B. Alcaraz, L. Fortin, A. Aubert, F. Delolme, N. El Kholti, J. Armengaud, P. Fournie, C. Auxenfans, U. Valcourt, S.V. Goff, C. Moali, BMP-1 disrupts cell adhesion and enhances TGF-beta activation through cleavage of the matricellular protein thrombospondin-1, *Sci. Signal.* 13 (2020), <https://doi.org/10.1126/scisignal.aba3880>.
- [48] D.J. Kim, E.D. Christofidou, D.R. Keene, M. Hassan Milde, J.C. Adams, Intermolecular interactions of thrombospondins drive their accumulation in extracellular matrix, *Mol. Biol. Cell* 26 (2015) 2640–2654, <https://doi.org/10.1091/mbc.E14-05-0996>.
- [49] A.G. Marneros, B.R. Olsen, The role of collagen-derived proteolytic fragments in angiogenesis, *Matrix Biol.* 20 (2001) 337–345, [https://doi.org/10.1016/S0945-053X\(01\)00151-2](https://doi.org/10.1016/S0945-053X(01)00151-2).
- [50] G. Theocharidis, J.T. Connelly, Minor collagens of the skin with not so minor functions, *J. Anat.* 235 (2019) 418–429, <https://doi.org/10.1111/joa.12584>.
- [51] D. Eyre, Collagen of articular cartilage, *Arthritis Res.* 4 (2002) 30–35, <https://doi.org/10.1186/ar380>.
- [52] D.A. Plumb, V. Dhir, A. Mironov, L. Ferrara, R. Poulsom, K.E. Kadler, D. J. Thornton, M.D. Briggs, R.P. Boot-Handford, Collagen XXVII is developmentally regulated and forms thin fibrillar structures distinct from those of classical vertebrate fibrillar collagens, *J. Biol. Chem.* 282 (2007) 12791–12795, <https://doi.org/10.1074/jbc.C700021200>.
- [53] M. Koch, J.E. Foley, R. Hahn, R.E. Burgeson, D.R. Gerecke, M.K. Gordon, alpha 1(XX) collagen, a new member of the collagen subfamily, fibril-associated collagens with interrupted triple helices, *J. Biol. Chem.* 276 (2001) 23120–23126, <https://doi.org/10.1074/jbc.M009912200>.
- [54] S. Grimal, S. Puech, R. Wagener, S. Venteo, P. Carroll, A. Fichard-Carroll, Collagen XXVIII is a distinctive component of the peripheral nervous system nodes of ranvier and surrounds nonmyelinating glial cells, *Glia.* 58 (2010) 1977–1987, <https://doi.org/10.1002/glia.21066>.
- [55] T. Wakabayashi, Transmembrane collagens in neuromuscular development and disorders, *Front. Mol. Neurosci.* 13 (2020) 635375, <https://doi.org/10.3389/fnmol.2020.635375>.
- [56] J. Dengjel, L. Bruckner-Tuderman, A. Nystrom, Skin proteomics - analysis of the extracellular matrix in health and disease, *Expert Rev. Proteome.* 17 (2020) 377–391, <https://doi.org/10.1080/14789450.2020.1773261>.
- [57] D. Breikreutz, I. Koxholt, K. Thiemann, R. Nischt, Skin basement membrane: the foundation of epidermal integrity—BM functions and diverse roles of bridging molecules nidogen and perlecan, *Biomed. Res. Int.* 2013 (2013) 179784, <https://doi.org/10.1155/2013/179784>.
- [58] C. Has, A. Nystrom, Epidermal basement membrane in health and disease, *Curr. Top. Membr.* 76 (2015) 117–170, <https://doi.org/10.1016/bs.ctm.2015.05.003>.
- [59] P.C. Trackman, Diverse biological functions of extracellular collagen processing enzymes, *J. Cell. Biochem.* 96 (2005) 927–937, <https://doi.org/10.1002/jcb.20605>.
- [60] W. Lesniak, A. Graczyk-Jarzynka, The S100 proteins in epidermis: topology and function, *Biochim. Biophys. Acta* 1850 (2015) 2563–2572, <https://doi.org/10.1016/j.bbagen.2015.09.015>.
- [61] A. Sandilands, C. Sutherland, A.D. Irvine, W.H. McLean, Filaggrin in the frontline: role in skin barrier function and disease, *J. Cell Sci.* 122 (2009) 1285–1294, <https://doi.org/10.1242/jcs.033969>.
- [62] B. Dyring-Andersen, M.B. Lovendorf, F. Coscia, A. Santos, L.B.P. Moller, A. R. Colaco, L. Niu, M. Bzorek, S. Doll, J.L. Andersen, R.A. Clark, L. Skov, M.B. M. Teunissen, M. Mann, Spatially and cell-type resolved quantitative proteomic atlas of healthy human skin, *Nat. Commun.* 11 (2020) 5587, <https://doi.org/10.1038/s41467-020-19383-8>.
- [63] C. Heit, B.C. Jackson, M. McAndrews, M.W. Wright, D.C. Thompson, G. A. Silverman, D.W. Nebert, V. Vasilioiu, Update of the human and mouse SERPIN gene superfamily, *Hum Genom.* 7 (2013) 22, <https://doi.org/10.1186/1479-7364-7-22>.

Supplementary Figure 1. Venn diagram of the matrisome proteins detected with the four enrichment protocols and the total tissue lysates.

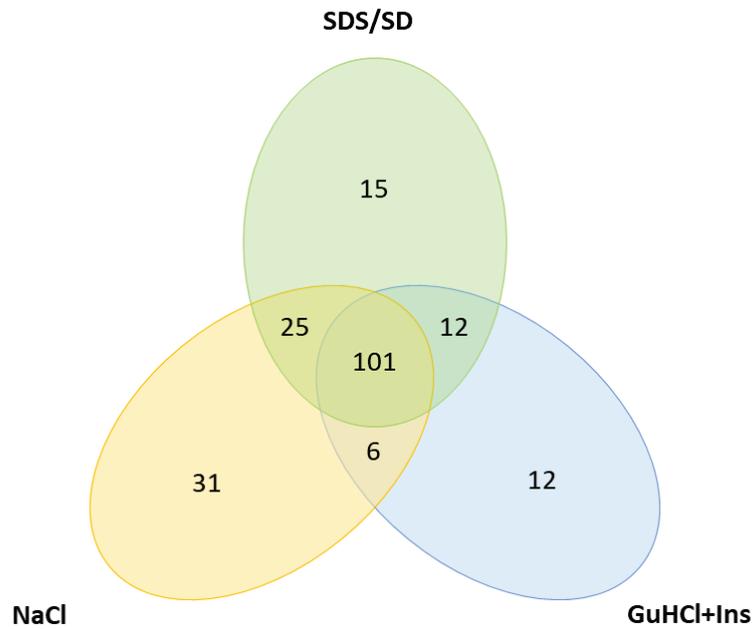


Supplementary Figure 2. Venn diagram comparing the matrisome proteins detected in the two analyzed fractions of the compartmental enrichment protocol.



**Supplementary Figure 3. Venn diagram comparing the matrisome proteins detected in the three analyzed fractions of the chemical decellularization protocol.**

*Matrisome proteins, Chemical decellularization enrichment,*



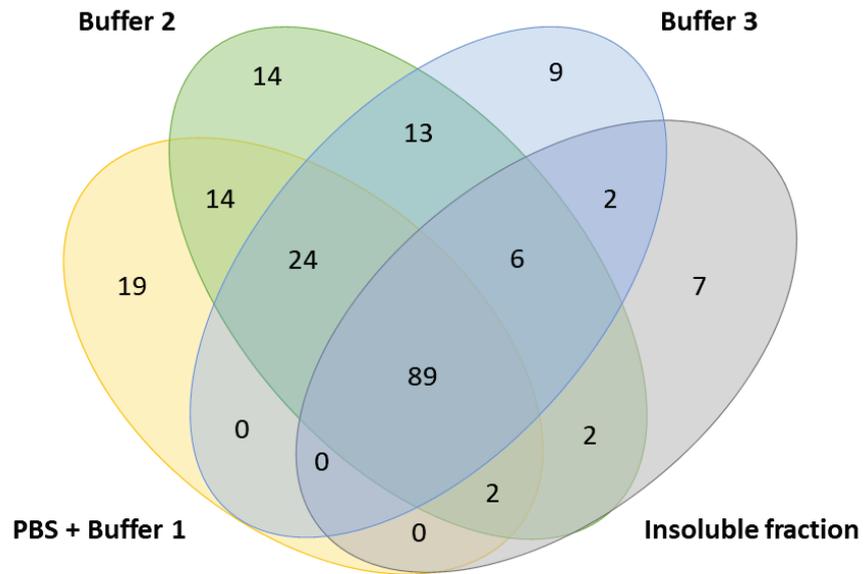
**Supplementary Figure 4. Venn diagram comparing the matrisome proteins detected in the three analyzed fractions of the enzymatic decellularization protocol.**

*Matrisome proteins, Enzymatic decellularization enrichment*

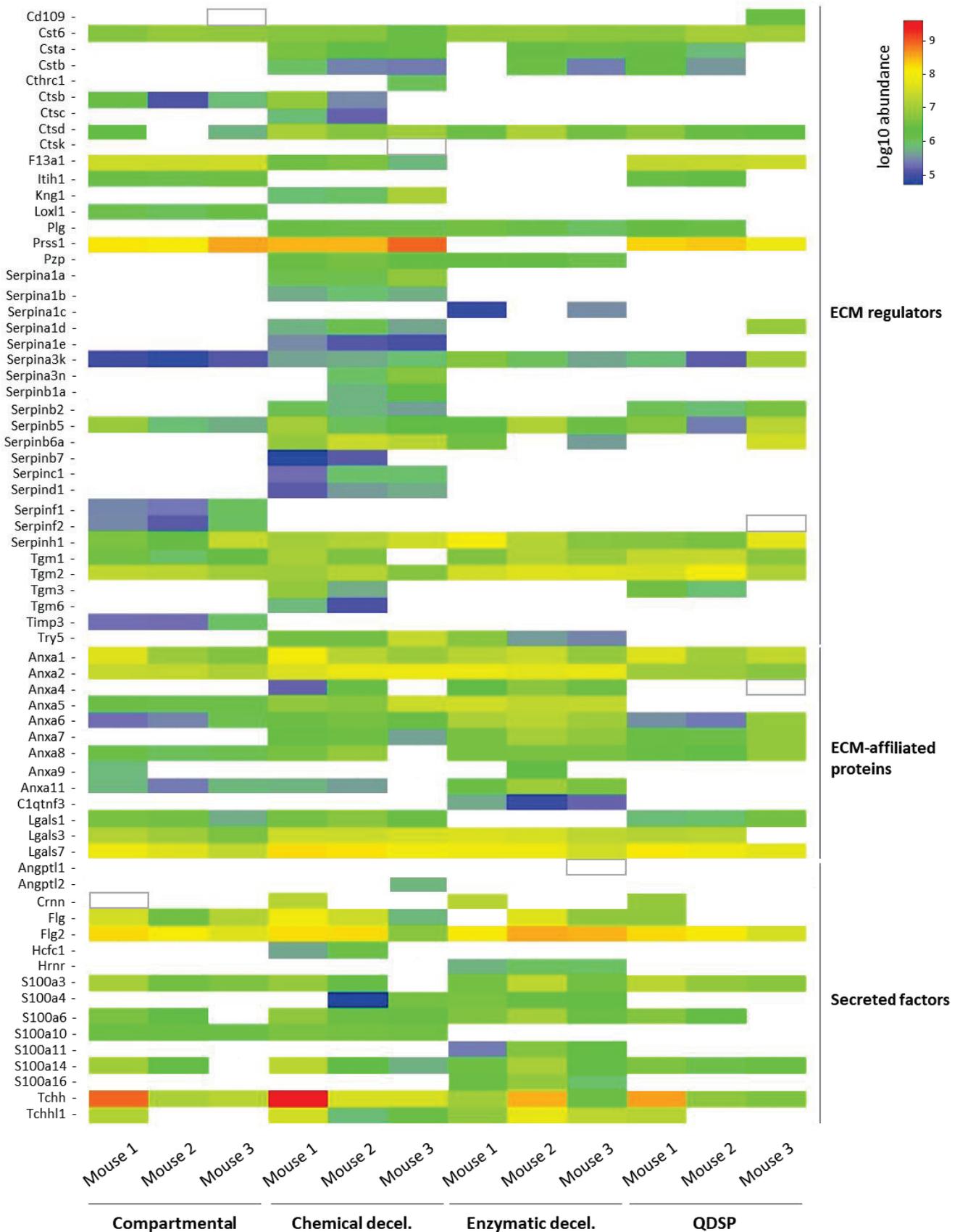


Supplementary Figure 5. Venn diagram comparing the matrisome proteins detected in the four analyzed fractions of the QDSP protocol.

*Matrisome proteins, QDSP*



**Supplementary Figure 6.** Heatmap representation of the abundance of associated matrisome proteins in the insoluble fractions of ECM enrichment protocols. Empty rectangles indicate proteins which were detected but could not be quantified.



Supplementary tables available on: <https://icloud.ibcp.fr/index.php/s/n7eE2rKLAic6ib1>

**Supplementary Table 1. Lists of proteins and peptides identified with Proteome Discoverer.**

**Supplementary Table 2. Matrisome proteins detected in the three biological replicates.**

**Supplementary Table 3. Abundance data for matrisome proteins identified in insoluble fractions.**

**Supplementary Table 4. Comparison of unique peptide numbers identified for matrisome proteins with and without “oxidized lysine” as variable modification.**

**Supplementary Table 5. Comparison of human, rat and mouse skin matrisomes.**

### 3. Final selection of ECM-enrichment protocols

The main benefits and drawbacks of each ECM-enrichment protocol tested have been already discussed in the publication presented above. If the general limitations in the study of the ECM by MS-based proteomics (mainly due to high ranges in the abundances and solubilities of matrix proteins) were also explained, some additional points can be mentioned.

Concerning tissue processing, our results showed that full decellularization is difficult to achieve. For instance, actin was detected by Western Blot in all enrichment fractions from the compartmental enrichment protocol, and additional cytoskeleton and nucleus proteins were also identified in the insoluble pellet of the same protocol by MS/MS (Supplementary Table 1 of the paper). As the intracellular compartment is involved in the synthesis and sometimes storage of both core and associated proteins, there is no certainty, in case of incomplete decellularization, on the cell or matrix origin of the proteins detected by MS/MS<sup>101,192</sup>. This limitation can be problematic when quantitative analyses are performed, because the relative contributions of the cellular and matrix compartments to the final quantification remain unknown. In addition, the range of abundances in extracellular proteins is thought to be wider than the dynamic range proposed by classical mass spectrometers. This results in an excellent detection of abundant matrix proteins such as collagens while low-abundance regulators are more difficult to detect. To overcome this limitation, it is possible to fractionate proteins from the various extracts or to fractionate peptides after protease-driven digestion<sup>99</sup>.

Regarding data exploitation for matrix analysis, the use of the whole UniProt database (reviewed + unreviewed sequences) appeared necessary. It indeed allowed the identification of important mouse matrix proteins such as tenascin X,  $\alpha 3(V)$  and  $\alpha 3(VI)$  collagen chains, trichohyalin, transglutaminase-6 or ADAMTS-like protein-5 that are absent in the SwissProt database (reviewed sequences only). Next, the matrixome project greatly improved the study of the ECM using proteomics: the matrixome datasets and tools are easy to use to identify matrix proteins from the raw protein lists obtained by MS/MS. However, human and mouse matrixome databases have not been updated since 2014: experimental results published after this date are therefore not taken into account and if new extracellular proteins have been reported these last 8 years, they are not included in the databases. Splicing variants are also missing from matrixome lists. In classical protein databases, splicing variants have distinct protein accession numbers, all associated to the same encoding gene. On the contrary, in the matrixome classification, one gene is associated to one protein only, meaning that splicing variants produced from the same gene are all matched with a unique protein ID. For instance, the *BMP1* gene encodes the two astacin metalloproteinases BMP-1 and mTLD resulting from distinct splicing. In matrixome lists, only BMP-1 is included; if mTLD is present in the initial MS/MS dataset, it is mis-referenced as BMP-1 when exploiting data with matrixome tools and the information on mTLD is

consequently lost. To overcome this limitation, we decided to exploit MS/MS data and mouse matrisome database with the tools proposed by the on-line platform ProteoRE<sup>180</sup>.

In order to perform our analyses on collagen VII-hypomorphic mice, we decided to keep two ECM enrichment protocols. First, the compartmental strategy was selected as a “reference” protocol, as it is extensively used and highly discussed in the literature, and informative regarding the core matrisome in spite of the loss of associated proteins. Only the insoluble pellet was analyzed in the following parts. Another protocol was also needed to better characterize the soluble proteins of the matrisome, especially the associated matrisome and the soluble precursors of the core matrisome components. Chemical decellularization and QDSP were good candidates as they allowed very good matrisome coverage combined with limited costs. In the decellularization protocol, the GuHCl extraction leads to the recovery of insoluble proteins mainly originating from the matrix while only very insoluble proteins remain in the pellet after this treatment. Based on the physico-chemical properties of the buffers and on the results obtained by MS/MS, we could conclude that the selective analysis of the NaCl and SDS/SD fractions as well as the combination of the GuHCl extract with the insoluble pellet allows the good characterization of matrisome proteins originating from the extracellular compartment. In QDSP, the protein recovery process is based on a progressive solubilization of proteins in buffers of increasing stringencies rather than on tissue decellularization. Most cell proteins are thus probably extracted in SDS/SD-rich buffers 2 and 3, thereby increasing sample complexity and complicating the cellular/extracellular assignment of each fraction. Therefore, the chemical decellularization protocol was selected as the second ECM enrichment protocol for this study.

In conclusion, a thorough characterization of four ECM enrichment protocols applied to mouse skin was conducted in the first part of this project and resulted in the first description of mouse skin matrisome. This comparison resulted in the selection of two protocols (compartmental strategy used as reference and chemical decellularization protocol to allow a deeper characterization of the associated matrisome) to characterize and compare the skin matrisome of WT and RDEB mice.

## II. GLOBAL CHARACTERIZATION OF THE EXTRACELLULAR MATRIX IN THE SKIN OF RDEB MICE

In this part, I will describe the modifications of the extracellular matrix of RDEB skin which were investigated in RDEB mice through two complementary approaches. Quantitative MS analyses were performed to study the matrisome dysregulations at the molecular level, while transmission electron microscopy (TEM) was used to evaluate ECM alterations at the structural level in mouse skin.

### 1. Selected mouse samples

Our initial objective was to describe ECM modifications at different stages of the disease in RDEB skin. As described in the introduction, samples from newborn, 4-week-old, and 12-week-old animals are representative in mouse RDEB paws of stages corresponding to tissue damage, chronic inflammation, and established dermal fibrosis, respectively. It would have been interesting to analyze them in comparison with normal skin and this is actually what we have done by TEM. However, based on our experience of ECM enrichment protocols, it appeared that the amount of skin material available in mouse paws would have been too low to allow a comprehensive ECM-centered proteomic study. Therefore, for the proteomic part, we decided to start to work with the back skin of 10-12-week-old animals which represents an early stage of the disease, as it is submitted to less mechanical friction, but could reveal the first alterations of the ECM leading to fibrosis.

### 2. Global description of the RDEB skin matrisome by MS-based proteomics

In order to get a comprehensive overview of ECM modifications in the skin of collagen VII-hypomorphic (RDEB) mice, we used MS-based quantitative proteomics to screen the matrisome components in healthy and RDEB tissues. Briefly, after enrichment using the two previously selected protocols (**Chapter III, I-3**), ECM proteins from the back skin of WT and RDEB mice (10-12 week-old) were digested into peptides and labeled with TMT prior to MS/MS analysis. Two experiments were successively run ( $n = 2$  mice/genotype, 4 multiplexed samples;  $n = 3$  mice/genotype, 6 multiplexed samples), each leading to four fraction types (NaCl, SDS/SD, GuHCl + insoluble pellet for the chemical decellularization protocol; insoluble pellet for the compartmental protocol, see Figure 1 of the paper presented in **Chapter III, I-2**) which were analyzed separately. Proteins in each fraction were identified by interrogating the whole UniProt mouse database to cover as many ECM molecules as possible (reviewed and unreviewed protein sequences).

#### i. Qualitative overview of the MS/MS analyses

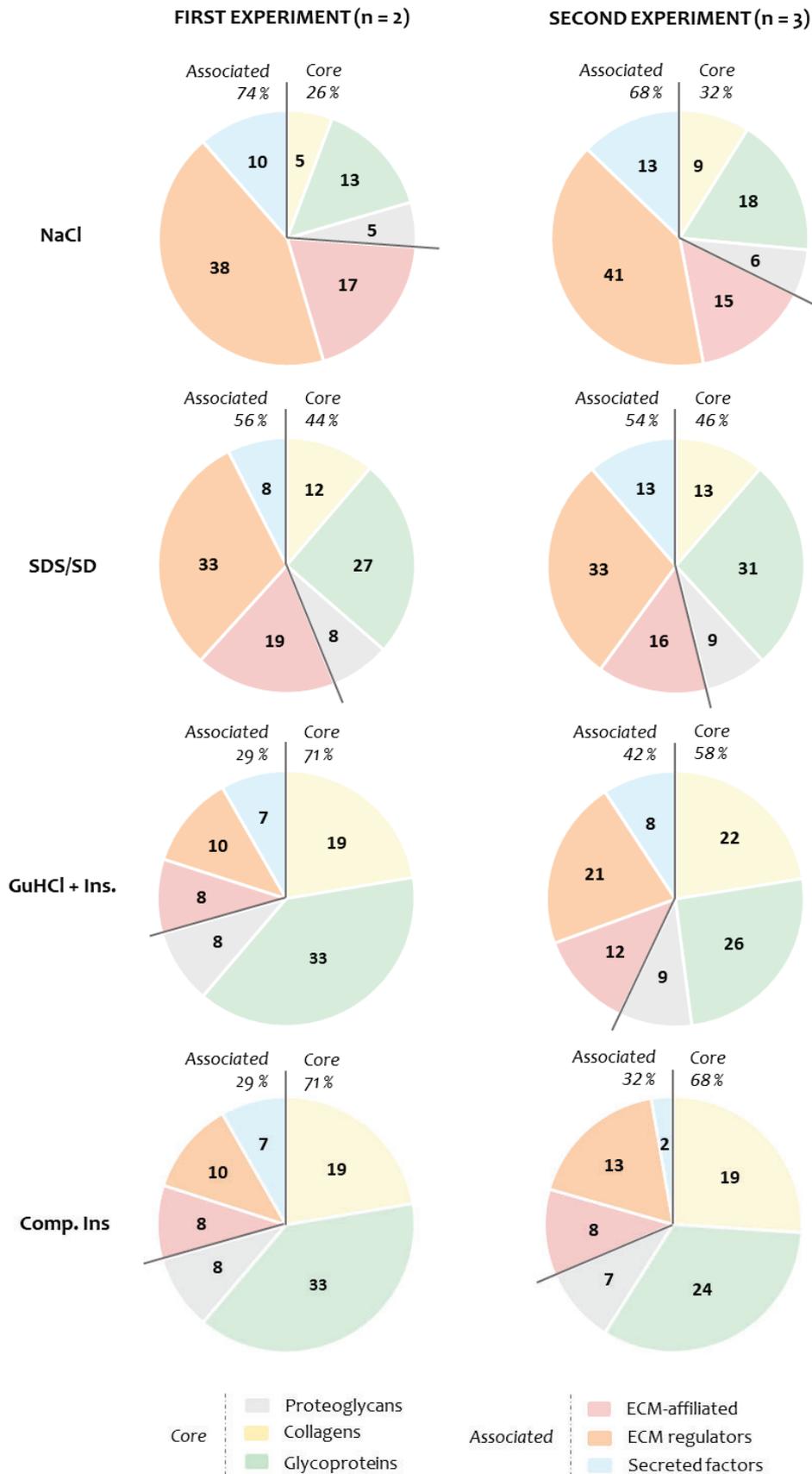
Between 506 and 1417 proteins were identified in individual enrichment fractions analyzed by MS/MS, and 2414 proteins (corresponding to 2925 proteoforms) were finally identified across all

samples (**Table 13**). A higher number of total and matrisome proteins was identified in the triplicate experiment compared to the duplicate experiment, especially for the NaCl fraction. This gain may result from the higher level of multiplexing that improves the detection of low-abundance components, especially those of the associated matrisome (**Figure 24**).

Experiment	n = 2				n = 3				Total
	NaCl	SDS/SD	GuHCl + Ins.	Comp. Ins.	NaCl	SDS/SD	GuHCl + Ins.	Comp. Ins.	
<b>Fractions</b>									All
<b>Total proteins</b>	1056	1274	600	612	1417	1318	672	506	2414
<b>Total proteoforms</b>	1077	1301	610	623	1430	1360	681	512	2925
<b>Matrisome proteins</b>	88	107	85	85	102	115	98	73	185
<b>Matrisome proteoforms</b>	91	112	91	90	107	121	101	76	232

**Table 13. Summary of total and matrisome proteins identified by MS/MS in mouse skin.**

Two experiments were successively conducted to compare the matrisome of WT and RDEB mice skin. Skin pieces from the back of 5 mice aged 10-12 weeks were analyzed for each genotype in total. Proteins were identified through the interrogation of the whole UniProt mouse database (reviewed and unreviewed sequences) and matrisome components were identified using the mouse matrisome database from the Matrisome project.



**Figure 24. Overview of the matrisome content in the enrichment fractions analyzed by MS/MS.** Matrisome proteins were identified and classified using the mouse matrisome database available online<sup>91</sup>. Associated proteins, especially ECM regulators, were enriched in “soluble” extracts (NaCl, SDS/SD) while core matrisome proteins were dominant in “insoluble” fractions.

185 ECM proteins in total were identified in MS/MS datasets using the mouse matrisome databases available online for the core and associated matrisomes<sup>91</sup>. In terms of nature and distribution of matrisome proteins in the various extracts, conclusions were similar to those drawn during the comparison of enrichment protocols. To sum up, the highest numbers of total and matrisome proteins were obtained in the “soluble” extracts (NaCl and SDS/SD). In these fractions, the associated matrisome components represented the major part while the “insoluble” extracts were enriched in proteins from the core matrisome, as expected (**Figure 24**). Very similar proportions of the various classes of matrisome proteins were observed between the duplicate and triplicate experiments, especially in “soluble” fractions. We can therefore conclude that the level of multiplexing and the switch from qualitative to TMT-based quantitative analysis did not affect the characterization of mouse skin matrisome. The results and conclusions reported during the previous analyses also apply to quantitative analyses.

Twenty-six collagens, 46 glycoproteins and 12 proteoglycans from the core matrisome category were detected (**Table 14**; the detailed matrisome content for each fraction is available in **Appendix 7**); all the key collagens, glycoproteins and proteoglycans of the skin ECM were detected (**Chapter I, parts I and II**). Twenty-five ECM-affiliated proteins, 55 ECM regulators and 21 secreted factors from the associated matrisome were also identified. Compared to the qualitative analyses conducted previously, some regulators such as transglutaminase-6, cathepsins C and L, lysyl oxidase and MMP-2 and -3 were missing from this new dataset. The loss of these components is not surprising because they are all enzymes which are known to be present in very low amounts and they have a low probability to be selected for fragmentation by the mass spectrometer despite the repeated injections (three technical replicates). The absence of some of the secreted factors detected in **Chapter III, I-2** (e.g. angiopoietin related-protein-2, myostatin) can also be explained by the fact that they were recovered through other enrichment protocols. In contrast, some new proteins were identified, especially in the associated matrisome. The ECM regulators serine protease HTRA-1 (high-temperature requirement protein-A1, encoded by the gene *Htra1*), the cysteine protease inhibitor stefin-A3 (*Stfa3*), the secreted factor hyaluronan binding protein-2 (*Habp2*) and the ECM-affiliated plexin-4 (*Plxn4*) were detected. In the core matrisome, some new glycoproteins (cartilage intermediate layer protein (*Cilp*), hemicentin-2 (*Hmcn2*) and insulin-like growth factor binding protein-7 (*Igfbp7*)) were also identified. Interestingly, the content in proteoglycans seems very stable as very similar proteins were identified (**Appendix 7**) in this study and in the optimization study described in **Chapter III, I-2**.

CORE MATRISOME		ASSOCIATED	
Gene	Protein name	Gene	Protein name
<b>Collagens (26)</b>		<b>ECM-affiliated (25)</b>	
<i>Coll1a1</i>	Collagen alpha-1 (I)	<i>Anxa1</i>	Annexin A1
<i>Coll1a2</i>	Collagen alpha-2 (I)	<i>Anxa2</i>	Annexin A2
<i>Col2a1</i>	Collagen alpha-1 (II)	<i>Anxa3</i>	Annexin A3
<i>Col3a1</i>	Collagen alpha-1 (III)	<i>Anxa4</i>	Annexin A4
<i>Col4a1</i>	Collagen alpha-1 (IV)	<i>Anxa5</i>	Annexin A5
<i>Col4a2</i>	Collagen alpha-2 (IV)	<i>Anxa6</i>	Annexin A6
<i>Col4a3</i>	Collagen alpha-3 (IV)	<i>Anxa7</i>	Annexin A7
<i>Col4a5</i>	Collagen alpha-5 (IV)	<i>Anxa8</i>	Annexin A8
<i>Col5a1</i>	Collagen alpha-1 (V)	<i>Anxa11</i>	Annexin A11
<i>Col5a2</i>	Collagen alpha-2 (V)	<i>Clqa</i>	Complement C1q subcomponent subunit A
<i>Col5a3</i>	Collagen alpha-3 (V)	<i>Clqtnf3</i>	Complement C1q tumor necrosis factor-related protein-3
<i>Col6a1</i>	Collagen alpha-1 (VI)	<i>Cd209b</i>	CD209 antigen-like protein B
<i>Col6a2</i>	Collagen alpha-2 (VI)	<i>Cd209d</i>	CD209 antigen-like protein D
<i>Col6a3</i>	Collagen alpha-3 (VI)	<i>Clec10a</i>	C-type lectin domain family-10 member A
<i>Col6a5</i>	Collagen alpha-5 (VI)	<i>Clec11a</i>	C-type lectin domain family-11 member A
<i>Col6a6</i>	Collagen alpha-6 (VI)	<i>Clec3b</i>	Tetranectin
<i>Col7a1</i>	Collagen alpha-1 (VII)	<i>Colec12</i>	Collectin-12
<i>Col8a1</i>	Collagen alpha-1 (VIII)	<i>Gpc1</i>	Glypican-1
<i>Col10a1</i>	Collagen alpha-1 (X)	<i>Hpx</i>	Hemopexin
<i>Col12a1</i>	Collagen alpha-1 (XII)	<i>Lgals1</i>	Galectin-1
<i>Col14a1</i>	Collagen alpha-1 (XIV)	<i>Lgals3</i>	Galectin-3
<i>Col15a1</i>	Collagen alpha-1 (XV)	<i>Lgals7</i>	Galectin-7
<i>Col17a1</i>	Collagen alpha-1 (XVII)	<i>Lgalsl</i>	Galectin-related protein
<i>Col18a1</i>	Collagen alpha-1 (XVIII)	<i>Lman1</i>	Protein ERGIC-53
<i>Col27a1</i>	Collagen alpha-1 (XXVII)	<i>Plxna4</i>	Plexin A4
<i>Col28a1</i>	Collagen alpha-1 (XXVIII)		
<b>Glycoproteins (46)</b>		<b>ECM regulators (55)</b>	
<i>Abi3bp</i>	ABI gene family, member 3 (NESH) binding protein	<i>Adamts15</i>	ADAMTS-like protein-5
<i>Adipoq</i>	Adiponectin	<i>Cd109</i>	CD109 antigen
<i>Aebp1</i>	Adipocyte enhancer-binding protein-1	<i>Cst3</i>	Cystatin-3
<i>Agrn</i>	Agrin	<i>Cst6</i>	Cystatin-6
<i>Cilp</i>	Cartilage intermediate layer protein	<i>Csta</i>	Cystatin A
<i>Coch</i>	Cochlin	<i>Cstb</i>	Cystatin B
<i>Cthrc1</i>	Collagen triple helix repeat-containing protein-1	<i>Ctsa</i>	Cathepsin A
<i>Dpt</i>	Dermatopontin	<i>Cstb</i>	Cathepsin B
<i>Ecm1</i>	Extracellular matrix protein-1	<i>Ctsd</i>	Cathepsin D
<i>Ecm2</i>	Extracellular matrix protein-2	<i>Ctsh</i>	Cathepsin H
<i>Efemp1</i>	EGF-containing fibulin-like extracellular matrix protein-1	<i>Ctsk</i>	Cathepsin K
<i>Eln</i>	Elastin	<i>Ctsz</i>	Cathepsin Z
<i>Emilin1</i>	EMILIN-1	<i>F2</i>	Prothrombin
<i>Fbln5</i>	Fibulin-5	<i>F13a1</i>	Coagulation factor XIII A chain
<i>Fbn1</i>	Fibrillin-1	<i>Gm5409</i>	Try10-like trypsinogen
<i>Fga</i>	Fibrinogen alpha chain	<i>Hrg</i>	Histidine-rich glycoprotein REG
<i>Fgb</i>	Fibrinogen alpha chain	<i>Htra1</i>	Serine protease HTRA-1
<i>Fgg</i>	Fibrinogen alpha chain	<i>Itih1</i>	Inter-alpha-trypsin inhibitor heavy chain H1
<i>Fn1</i>	Fibronectin	<i>Itih2</i>	Inter-alpha-trypsin inhibitor heavy chain H2
<i>Hmnc2</i>	Hemicentin-2	<i>Itih3</i>	Inter-alpha-trypsin inhibitor heavy chain H3
<i>Igfbp7</i>	Insulin-like growth factor binding protein-7	<i>Itih4</i>	Inter-alpha-trypsin inhibitor heavy chain H4
<i>Lama2</i>	Laminin subunit alpha-2	<i>Kng1</i>	Kininogen-1
<i>Lama3</i>	Laminin subunit alpha-3	<i>P4ha1</i>	Prolyl 4-hydroxylase subunit alpha-1
<i>Lama4</i>	Laminin subunit alpha-4	<i>Plg</i>	Plasminogen
<i>Lama5</i>	Laminin subunit alpha-5	<i>Plod3</i>	Lysyl hydroxylase-3
<i>Lamb1</i>	Laminin subunit beta-1	<i>Pzp</i>	Pregnancy zone protein
<i>Lamb2</i>	Laminin subunit beta-2	<i>Serpina1a</i>	Serine protease inhibitor A1a
<i>Lamb3</i>	Laminin subunit beta-3	<i>Serpina1b</i>	Serine protease inhibitor A1b
<i>Lamc1</i>	Laminin subunit gamma-1	<i>Serpina1d</i>	Serine protease inhibitor A1d
<i>Lamc2</i>	Laminin subunit gamma-2	<i>Serpina1e</i>	Serine protease inhibitor A1e
<i>Lrg1</i>	Leucine-rich alpha-2-glycoprotein	<i>Serpina3b</i>	Serine protease inhibitor A3b
<i>Mfap2</i>	Microfibrillar-associated protein-2	<i>Serpina3k</i>	Serine protease inhibitor A3k
<i>Mfap4</i>	Microfibrillar-associated protein-4	<i>Serpina3m</i>	Serine protease inhibitor A3m
<i>Mfap5</i>	Microfibrillar-associated protein-5	<i>Serpina3n</i>	Serine protease inhibitor A3n
<i>Nid1</i>	Nidogen-1	<i>Serpina6</i>	Serine protease inhibitor A6
<i>Nid2</i>	Nidogen-2	<i>Serpina1a</i>	Serine protease inhibitor B1a
		<i>Serpina2</i>	Serine protease inhibitor B2
		<i>Serpina5</i>	Serine protease inhibitor B5
		<i>Serpina6a</i>	Serine protease inhibitor B6a

<i>Postn</i>	Periostin	<i>Serpnb6c</i>	Serine protease inhibitor B6c
<i>Sparc</i>	Basement-membrane protein-40	<i>Serpnb6d</i>	Serine protease inhibitor B6d
<i>Srpx</i>	Sushi repeat-containing protein SRPX	<i>Serpnb7</i>	Serine protease inhibitor B7
<i>Tgfb1</i>	TGF- $\beta$ -induced protein ig-h3 (TGFB1)	<i>Serpnb8</i>	Serine protease inhibitor B8
<i>Thbs1</i>	Thrombospondin-1	<i>Serpnb10</i>	Serine protease inhibitor B10
<i>Tnc</i>	Tenascin C	<i>Serpnc1</i>	Serine protease inhibitor C1
<i>Tnxb</i>	Tenascin X	<i>Serpnd1</i>	Serine protease inhibitor D1
<i>Vtn</i>	Vitronectin	<i>Serpinf1</i>	Serine protease inhibitor F1
<i>Vwa1</i>	Von Willebrand factor A domain-containing protein-1	<i>Serpinf2</i>	Serine protease inhibitor F2
<i>Vwa5a</i>	Von Willebrand factor A domain-containing protein-5A	<i>Serping1</i>	Serine protease inhibitor G1
		<i>Serpinh1</i>	Serine protease inhibitor H1
		<i>Stfa3</i>	Stefin A3
		<i>Tgm1</i>	Transglutaminase-1
		<i>Tgm2</i>	Transglutaminase-2
		<i>Tgm3</i>	Transglutaminase-3
		<i>Try5</i>	Trypsin-5
<b>Proteoglycans (12)</b>		<b>Secreted factors (21)</b>	
<i>Aspn</i>	Asporin	<i>Flg</i>	Filaggrin
<i>Bgn</i>	Biglycan	<i>Flg2</i>	Filaggrin-2
<i>Cspg4</i>	Chondroitin sulfate proteoglycan 4	<i>Habp2</i>	Hyaluronan-binding protein-2
<i>Dcn</i>	Decorin	<i>Hcfc1</i>	Host cell factor-1
<i>Fmod</i>	Fibromodulin	<i>Hrnr</i>	Hornerin
<i>Hspg2</i>	Perlecan	<i>Il1rn</i>	Interleukin-1 receptor antagonist protein
<i>Lum</i>	Lumican	<i>M6pr</i>	Cation-dependent mannose-6- phosphate receptor
<i>Ogn</i>	Mimecan	<i>Rptn</i>	Repetin
<i>Podn</i>	Podocan	<i>S100a1</i>	Protein S100 A1
<i>Prelp</i>	Prolargin	<i>S100a3</i>	Protein S100 A3
<i>Prg2</i>	Bone marrow proteoglycan	<i>S100a4</i>	Protein S100 A4
<i>Vcan</i>	Versican	<i>S100a6</i>	Protein S100 A6
		<i>S100a8</i>	Protein S100 A8
		<i>S100a9</i>	Protein S100 A9
		<i>S100a10</i>	Protein S100 A10
		<i>S100a11</i>	Protein S100 A11
		<i>S100a13</i>	Protein S100 A13
		<i>S100a14</i>	Protein S100 A14
		<i>S100a16</i>	Protein S100 A16
		<i>S100b</i>	Protein S100 B
		<i>Tchh</i>	Trichohyalin
<b>Total</b>	<b>84</b>	<b>Total</b>	<b>101</b>

**Table 14. Matrisome proteins identified in the skin of WT and RDEB mice by MS-based quantitative proteomics.**

All the matrisome proteins identified in the two described experiments were reported in this table (database used for the identification of proteins: UniProt).

## ii. Analysis of quantitative data

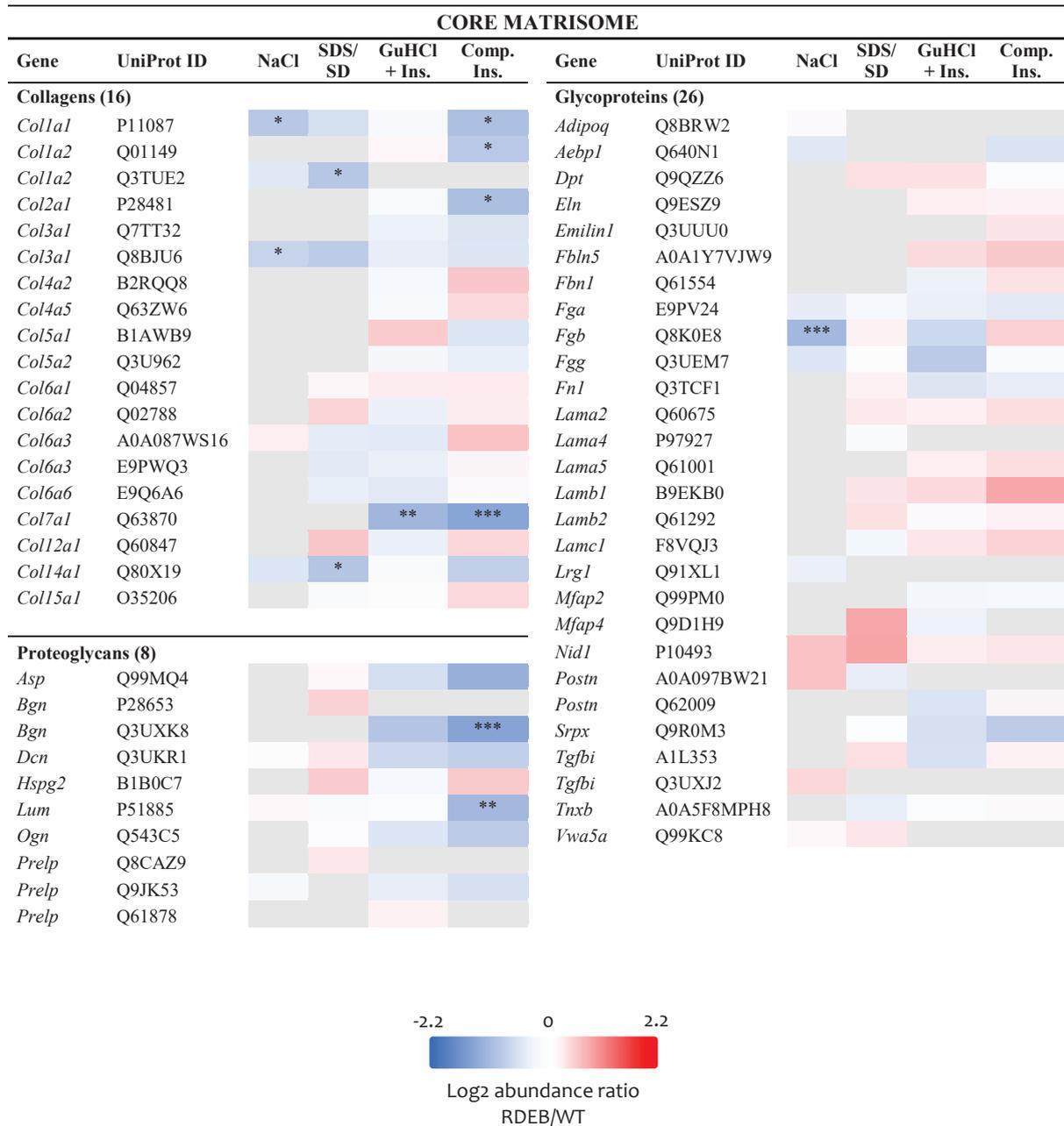
Among all the proteins detected in the two experiments ( $n = 2$  and  $n = 3$ ), several of them were not quantified, including some matrisome components. In order to maximize the detection of genotype-associated modifications and avoid false interpretations resulting from inter-individual variations, the decision was made to focus on the proteins which were quantified in the two multiplex experiments (for the matrisome, this corresponds to 60 % of identified proteins). The results of the two MS/MS experiments were therefore pooled after normalization (**Chapter II, III-3-iv**) and analyzed together. All proteoforms were included to avoid losses in quantitative information, and the abundance ratio between the WT and hypomorphic genotypes was computed for analysis.

The abundance of 1146 proteins (1227 proteoforms) was finally compared in WT and RDEB samples. Ninety-five of them were significantly dysregulated, including 22 matrisome proteins.

Regarding the core matrisome, 50 proteins were quantified (**Figure 25**), with multiple proteoform sequences identified for collagens  $\alpha 2(I)$ ,  $\alpha 1(III)$ ,  $\alpha 3(VI)$ , for biglycan and prolargin (proteoglycans) and for periostin and TGFBI (glycoproteins). Importantly, the strong diminution of collagen  $\alpha 1(VII)$  in RDEB skin was validated in all the fractions where the protein was detected (insoluble fractions of the two protocols). Furthermore, this under-expression was the strongest detected among all core matrisome proteins in insoluble extracts.

More unexpectedly, the abundance of fibrillar collagens appeared to be disturbed in RDEB skin. Soluble forms of collagen  $\alpha 1(I)$  and  $\alpha 2(I)$  were significantly reduced in the NaCl and SDS/SD fractions respectively, and the two chains were also dysregulated in the insoluble extract from the compartmental enrichment protocol. However, in the last fractions from the chemical decellularization protocol (GuHCl + Ins.), no variation was noticed, suggesting that only the very insoluble and cross-linked forms of collagen I are affected. Significant diminutions in other fibrillar collagens were also observed in RDEB skin, either in the NaCl extract for collagen III or in the compartmental insoluble pellet for collagen II. Collagen XIV is another collagen with diminished abundance. It is not a fibrillar collagen but as a FACIT and it can regulate the network of collagen fibrils and their surface properties<sup>46</sup>. Regarding proteoglycans, a global reduction in the abundance of SLRPs in insoluble fractions was observed. Thus, the collagen organizers biglycan (*Bgn*), lumican (*Lum*), decorin (*Dcn*) and osteoglycin (*Ogn*) were all decreased compared to WT in the insoluble pellet from the compartmental insoluble fraction (and to a lesser extent in the chemical decellularization GuHCl + Ins. fraction) and this reduction was significant for the first two proteins. For biglycan and decorin, this decrease was accompanied by slight increases in soluble extracts. Also, TGF- $\beta$ -induced protein ig-h3 (*Tgfbi*, abbreviated TGFBI or big-h3), a protein possibly involved in the deposition of collagen and in the regulation of its interactions with cells<sup>202</sup>, tended to be increased in the soluble fractions at the expense of the insoluble compartment, suggesting a modification in the solubility of the protein in disease. Periostin (*Pstn*), which is known to regulate cross-link formation in collagen fibrils<sup>152</sup>, showed a similar pattern. All these data suggested that the collagen network of the dermis could be significantly affected by collagen VII deficiency.

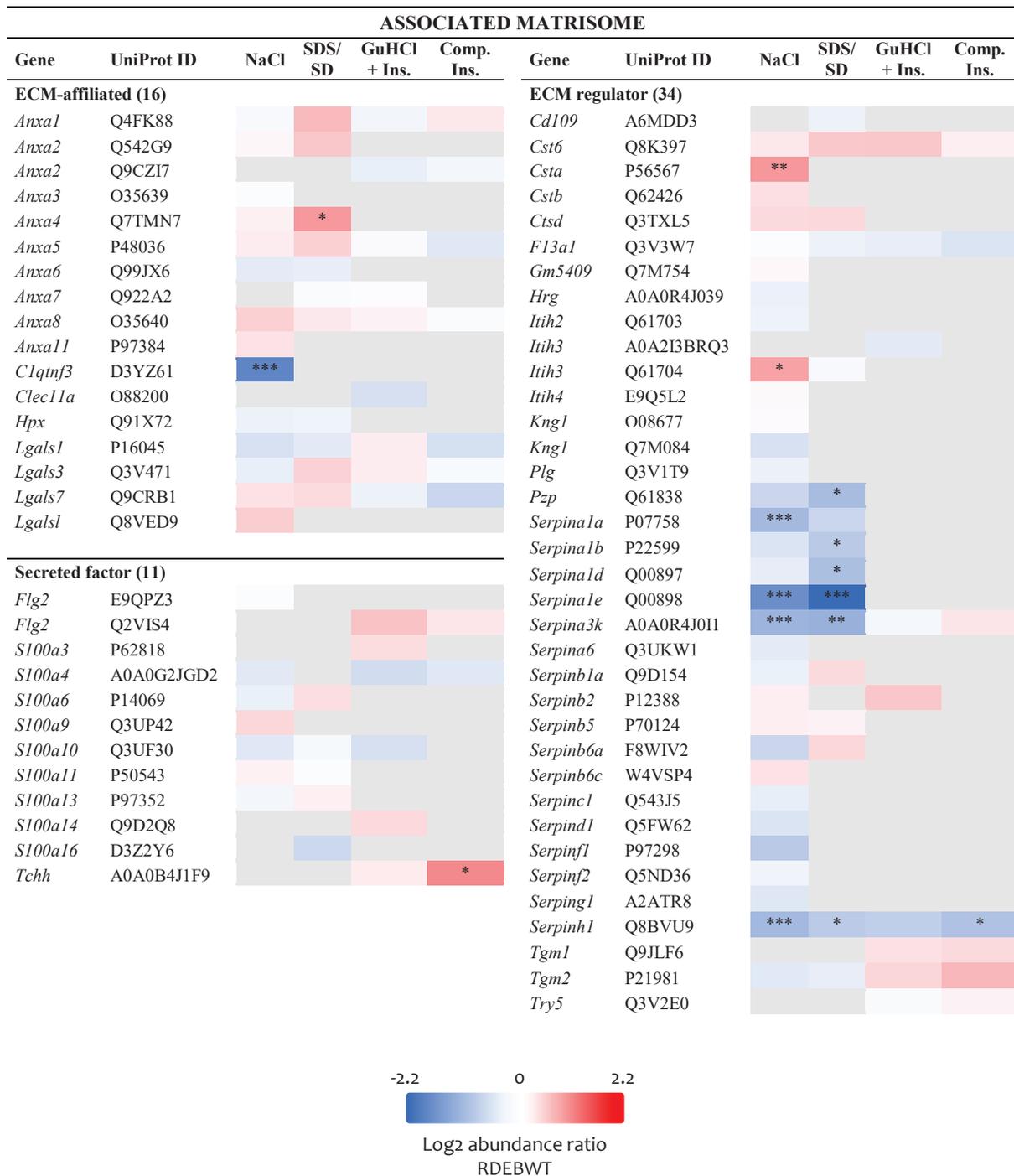
Interestingly, a trend toward augmented abundance was observed for basement membrane proteins other than collagen VII. Indeed, the abundances of collagens IV, VI and XV chains, of laminin chains, especially those belonging to laminin-511, and of perlecan (*Hspg2*) were globally increased. Nidogen-1 (*Nid1*) was also increased in all the fractions studied, and close to significance in the soluble ones. Quantitative data were unfortunately not available for some critical basement membrane components such as laminin chains  $\beta 3$  and  $\gamma 2$  or collagen XVII but the present data seem to indicate that there is some compensation of the lack of collagen VII by other basement membrane components.



**Figure 25. Heatmap of the abundance ratio of core matrisome proteins from WT and RDEB mouse skin.** For each sample, quantitative data from the two MS/MS experiments were normalized and used to compute the median abundances of core matrisome proteins for WT and RDEB genotypes (n = 5 mice/genotype). The abundance ratio represented here corresponds to RDEB/WT (Log<sub>2</sub> scale). Only proteins quantified in the two experiments were compared. Grey rectangles: protein not detected or not quantified in the fraction. Statistics: see **Chapter II, III-2-iv**. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

Quantitative data were also obtained for about 60 % of the associated matrisome proteins detected, mainly in the NaCl and SDS/SD fractions (**Figure 26**). Among the proteins which showed an overall increase in RDEB skin, we found some annexins (especially annexin-A4, a protein involved in the modulation of membrane fusion events and exocytosis<sup>203</sup> and encoded by *Anxa4*), cystatins (*Cst*, inhibitors of cysteine proteases) and Trichohyalin (*Tchh*, a substrate of transglutaminase which participates in keratin intermediate filaments<sup>204</sup>). In contrast to the possible increase in cysteine protease

inhibitors, a strong tendency toward decrease in several negative regulators of proteases was observed: pregnancy zone protein (*Pzp*, also known as “ $\alpha$ 2-macroglobulin” which is a circulating broad-spectrum protease inhibitor<sup>205</sup>) and several members belonging to the serpin family (serine protease inhibitors) exhibited a strong decrease of their abundance. The proteins produced from *Serpina1a*, *Serpina1b*, *Serpina1d*, *Serpina1e*, a group of paralogs encoding  $\alpha$ 1-anti-trypsin in mouse, and from *Serpina3k*, were the most affected. The only exceptions which did not follow this trend (and were rather increased) were inhibitors of the clade B serpin family. They are expressed in monocytes and neutrophils<sup>206</sup> and could participate in restraining inflammation at this early stage of the disease in back skin. Finally, *Serpinh1* also encode a serpin-like protein but it is not involved in the modulation of the activity of proteases<sup>207</sup>. Instead, it acts as a chaperone during the synthesis of collagens, especially the fibrillar ones, and is rather known in this context as HSP-47. Very interestingly in relation with the fibrillar collagen alterations reported above, it is clearly and significantly diminished in RDEB skin in 3 fractions out of 4 (and close to significant decrease in the fourth one).

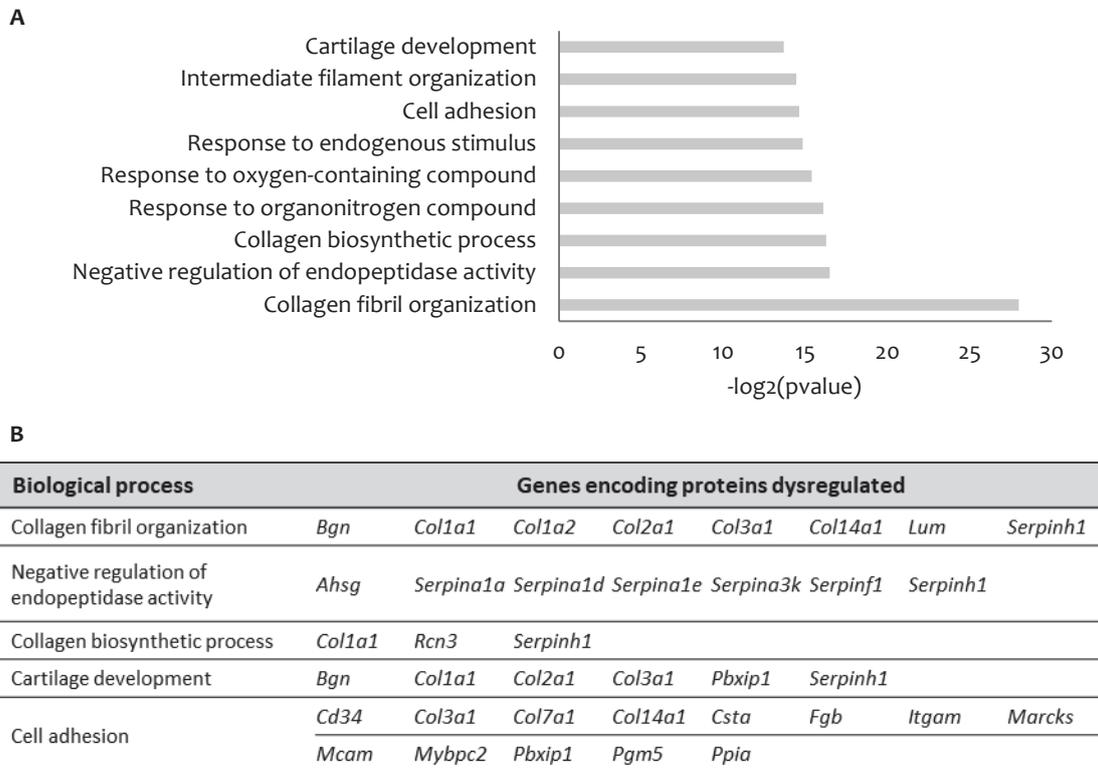


**Figure 26. Heatmap of the abundance ratio of associated matrisome proteins from WT and RDEB mouse skin.**

For each sample, quantitative data from the two MS/MS experiments were normalized and used to compute the median abundances of associated matrisome proteins for WT and RDEB genotypes ( $n = 5$  mice/genotype). The abundance ratio represented here corresponds to RDEB/WT (Log<sub>2</sub> scale). Only proteins quantified in the two experiments were compared. Grey rectangles: protein not detected or not quantified in the fraction. Statistics: see Chapter II, III-2-iv. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

### iii. Impaired biological processes

The importance of the alterations affecting the extracellular matrix relatively to all significant dysregulations in the skin was next investigated. A statistical over-representation test was performed on all the 95 proteins (including matrisome and non matrisome proteins), of which the abundance was significantly modified in disease, in order to determine the most affected biological processes.



**Figure 27. Altered biological processes in RDEB mouse skin.**

**A.** Significantly over-represented biological processes. All the genes encoding the 85 proteins with a significantly modified abundance in RDEB skin (MS/MS data) were submitted to a statistical over-representation test using Panther (“GO biological processes complete” mode)<sup>179</sup>. **B.** Detailed gene content of each matrix-related dysregulated process.

Among the top 9 significantly over-represented processes (**Figure 27A**), 5 were directly linked to ECM proteins: collagen fibril organization, collagen biosynthetic pathway, cell adhesion, cartilage development and negative regulation of endopeptidase activity (**Figure 27B**). The organization of the collagen network appeared to be the most significantly affected process in RDEB skin, as well as the production of collagens. Regarding the latter process, it is interesting to note that reticulocalbin-3 (encoded by *Rcn3*) is a cell protein with chaperone functions in the endoplasmic reticulum. It has been directly linked to collagen synthesis and deposition<sup>208,209</sup> and was significantly decreased in all the enrichment fractions where it was identified (3 of the 4 analyzed fractions, abundance reduced by half in RDEB compared to WT; **Appendix 9**), consistent to what was observed above for HSP-47 (collagen-specific chaperone). The modifications in the adhesion of cells to their ECM substrates resulted from alterations in additional matrix proteins and to some cellular proteins. Also notable is the fact that the

second over-represented pathway is negative regulation of endopeptidase activity. This was linked to the changes in abundance observed for serpins, as described in the heat map of the associated matrixome.

Cell processes related to the response to reactive molecules (oxygen, nitrogen compounds) were also detected, consistently with previous studies reporting an increase in the oxidative stress of fibroblasts originating from the unwounded skin biopsies of RDEB patients<sup>145,210</sup>. Modifications in the organization of the cytoskeleton of keratinocytes were also underlined in the intermediate filament organization biological process.

In conclusion, this analysis strengthens the idea that the deficiency of collagen VII is responsible for strong alterations in the matrix compartment including both core and associated matrixomes. Of specific interest for our study, collagen fibril organization and protease inhibition are the most dysregulated processes. However, it should be noted here that there is a strong bias in this type of over-representation studies when applied after ECM enrichment. Most cellular proteins have been excluded by the enrichment procedure and it is expected that pathways related to ECM will be over-represented. We are now working with bioinformaticians to generate a reference database focused on the extracellular compartment which will be more appropriate for our samples. In the meantime, the hierarchy between ECM-related biological processes remains valid and it helps to point the most important trends for further analysis.

### 3. Analysis of skin ultrastructure by transmission electron microscopy

In parallel to the proteomic characterizations described above, ultrastructural analyses were also conducted on mouse skin to determine if the global organization of the matrix was altered in RDEB. Different types of samples were analyzed (different ages and different skin locations) to evaluate the evolution in ECM structure with the progression of the disease. We first characterized the structure of the DEJ which is in direct interaction with collagen VII and then focused our analysis on the dermis to see if the dysregulations observed above could also be evidenced at the structural level. These two compartments contain most of the extracellular matrix of the skin.

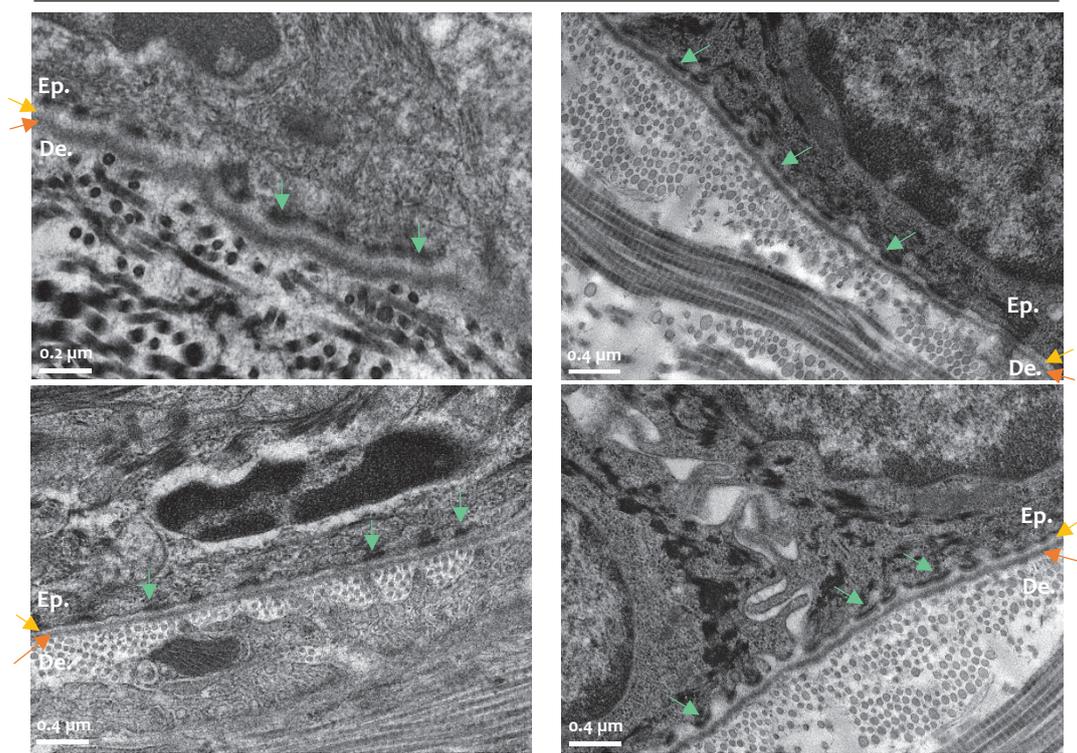
#### i. Characterization of the dermo-epidermal junction in RDEB skin

##### *DEJ in the back skin*

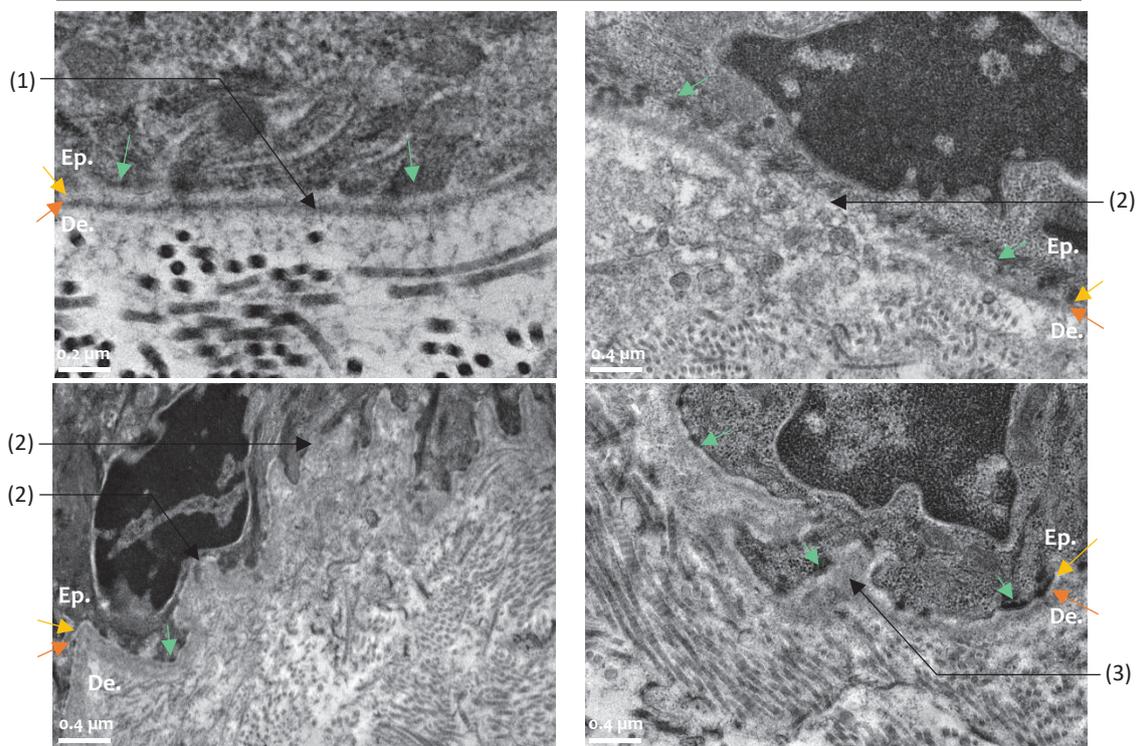
In direct relation with the studies made previously, the DEJ of the back skin from mice aged 9 to 12 weeks was first analyzed by TEM. In WT samples, the different compartments of the DEJ were clearly observed: the hemidesmosomes of basal keratinocytes were distributed all along their membrane and the *lamina lucida* and *lamina densa* of the basement membrane were clearly visible (**Figure 28**). In RDEB skin, most DEJ portions looked relatively normal but closer inspection indicated that the *lamina densa* was often thinner, as previously described by our German collaborators<sup>144</sup>. The DEJ was rarely

completely disrupted although this could be observed from time to time. Also, on some pictures, the *lamina lucida* and *lamina densa* appeared very diffuse, with the separation between the two layers being very difficult to localize. Finally, hemidesmosomes seemed smaller compared to what was observed in WT.

WT



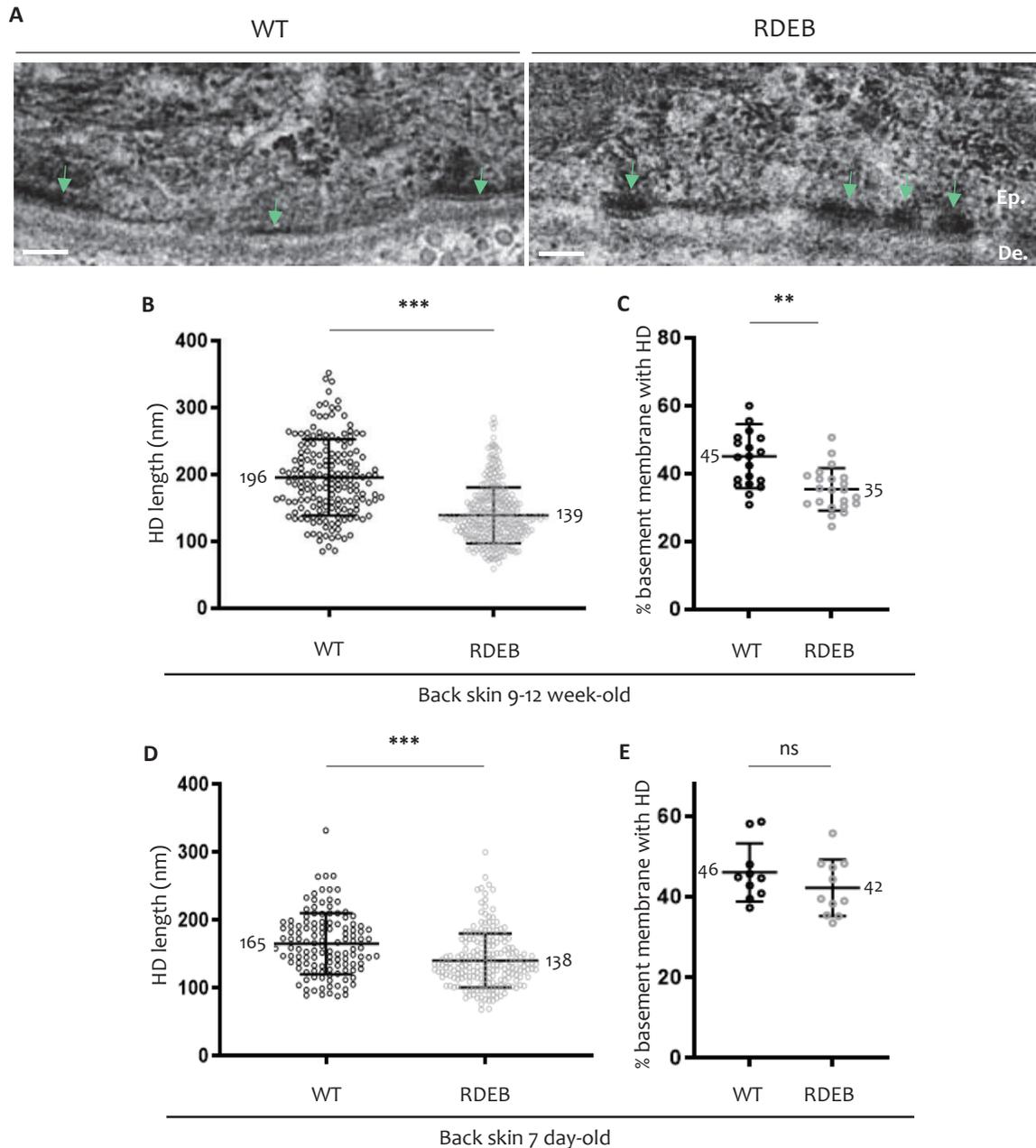
RDEB



**Figure 28. TEM observations of the DEJ in the back skin of WT and RDEB mice aged 9-12 weeks.** Representative TEM images were selected to show the DEJ. Clear abnormalities were observed in RDEB skin: (1) thin *lamina densa*, (2) diffused basement membrane, (3) ruptures in the *lamina densa*. Yellow arrows point *lamina lucida*, orange arrows point *lamina densa* and green arrows indicate hemidesmosomes. Ep: epidermis, De: dermis.

As variations in the hemidesmosome size were suspected, different parameters relative to these keratinocyte membrane structures, involved in interactions with basement membrane ECM proteins, were analyzed (measurements made by Patricia Rousselle). In the skin of WT mice aged 9 to 12 weeks (**Figure 29A**), hemidesmosomes exhibited a mean length of 196 nm but they were significantly smaller in RDEB skin with a decrease in mean length of about 30 % (mean length of 139 nm), confirming the visual observations (**Figure 29B**). In addition, the proportion of hemidesmosomes along the basement membrane was affected (- 22 % hemidesmosomes/ $\mu\text{m}$  basement membrane in RDEB, **Figure 29C**).

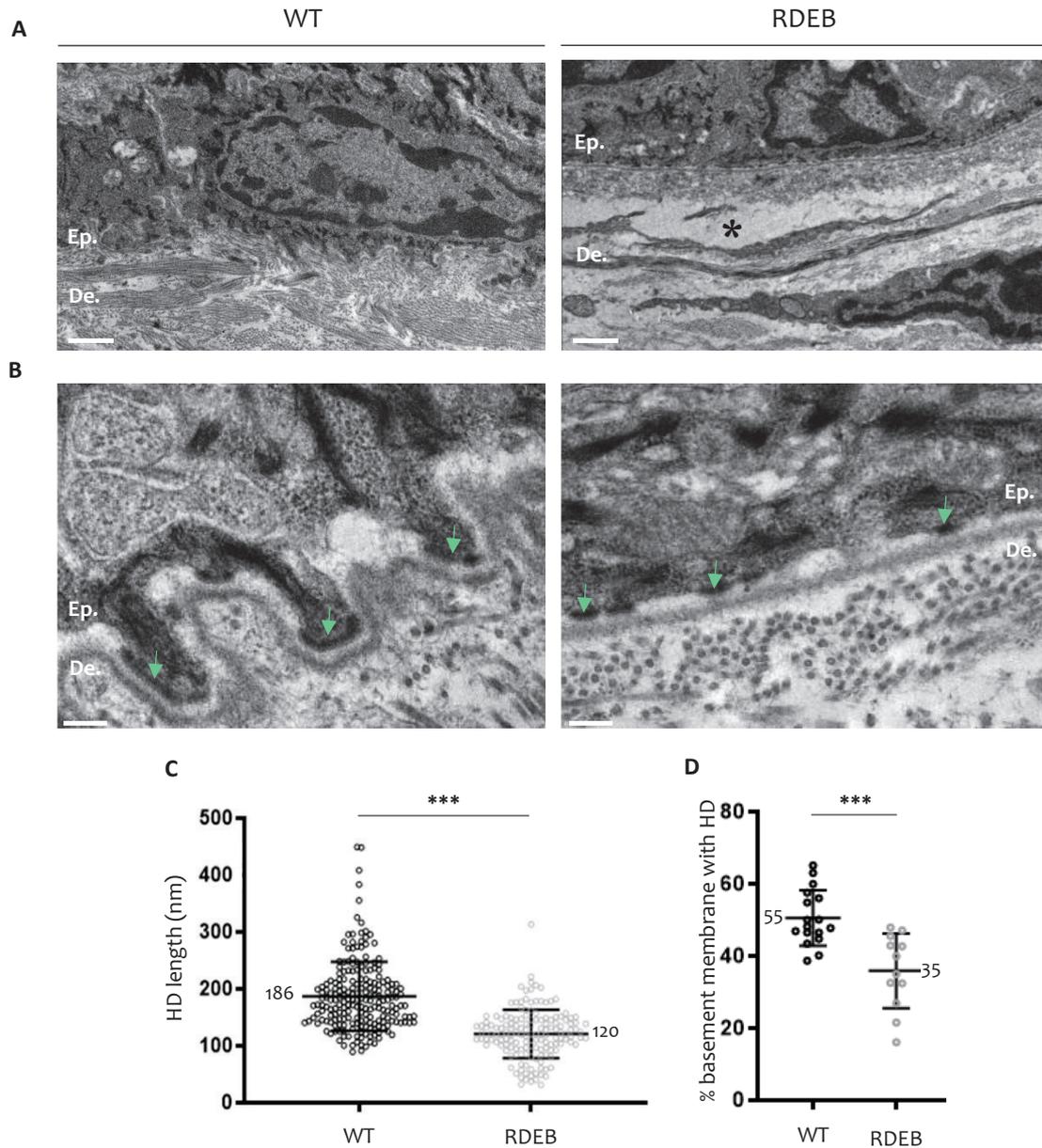
To evaluate the effect of disease progression on the DEJ ultrastructure, we next characterized back skin samples originating from new-born animals (7 day-old, less severely affected). The same results as those observed in older mice were made regarding the characteristics of the hemidesmosomes. However, the variations in the length of hemidesmosomes in RDEB were more limited compared to what was described above, with a reduction of about 16 % only compared to WT (respective mean lengths of 138 and 165 nm) (**Figure 29D**). In addition, if WT hemidesmosomes exhibited a smaller size compared to those observed in the back skin of older mice (165 nm and 196 nm respectively), it was not the case in RDEB where very similar values were surprisingly observed between the two types of samples (138 nm and 139 nm respectively). Finally, the reduction in the coverage of the basement membrane (- 8 %) was not marked enough in young animals to be significant (**Figure 29E**).



**Figure 29. Analysis of the hemidesmosomes of basal keratinocytes in the back skin of WT and RDEB mice.** **A.** Representative TEM images used to analyze hemidesmosomes (mice aged 9 to 12 week-old). Green arrows indicate the hemidesmosomes. Bar: 100 nm. **B and C.** Length of hemidesmosomes and proportion of the basement membrane covered with hemidesmosomes in the back skin of mice aged 9-12 weeks. Samples from 4 WT and 4 RDEB mice were analyzed to measure the length of 181 WT and 330 RDEB hemidesmosomes (measured from 4-6 different images/mouse). **D and E.** Length of hemidesmosomes and proportion of the basement membrane covered with hemidesmosomes in the back skin of mice aged 7 days. Samples from 2 WT and 2 RDEB mice were analyzed to measure the length of 124 WT and 202 RDEB hemidesmosomes (measured from 5-7 different images/mouse). Student t-test or Mann-Whitney test when t-test was not applicable. ns: not significant, \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ . HD: hemidesmosome.

*Analysis of the hemidesmosomes in paws*

The skin of the forepaws from mice aged 9 to 12 weeks was finally analyzed. Contrary to back skin, paws are highly submitted to friction and rapidly show fibrosis and digit fusion as early as around 10-12 weeks. They are therefore adapted to evaluate the alterations of the DEJ in more advanced stages of the disease.



**Figure 30. Analysis of the hemidesmosomes of basal keratinocytes in the forepaw skin of WT and RDEB mice aged 9-12 weeks.**

TEM images were processed to analyze different parameters. **A.** Representative images of the DEJ in paw skin. \*: epidermis blistering. Bar: 1 μm. **B.** Representative TEM images used to analyze hemidesmosomes. Green arrows indicate the hemidesmosomes. Bar: 200 nm. **C and D.** Length of hemidesmosomes and proportion of the basement membrane covered with hemidesmosomes in the forepaw skin of mice aged 9-12 weeks. Samples from 4 WT and 4 RDEB mice were analyzed to measure the diameter of 211 WT and 153 RDEB hemidesmosomes (measured from 3 to 5 different images/mouse). Statistics: Student t-test or Mann-Whitney test when t-test was not applicable. \*\*\*:  $p < 0.001$ . HD: hemidesmosome.

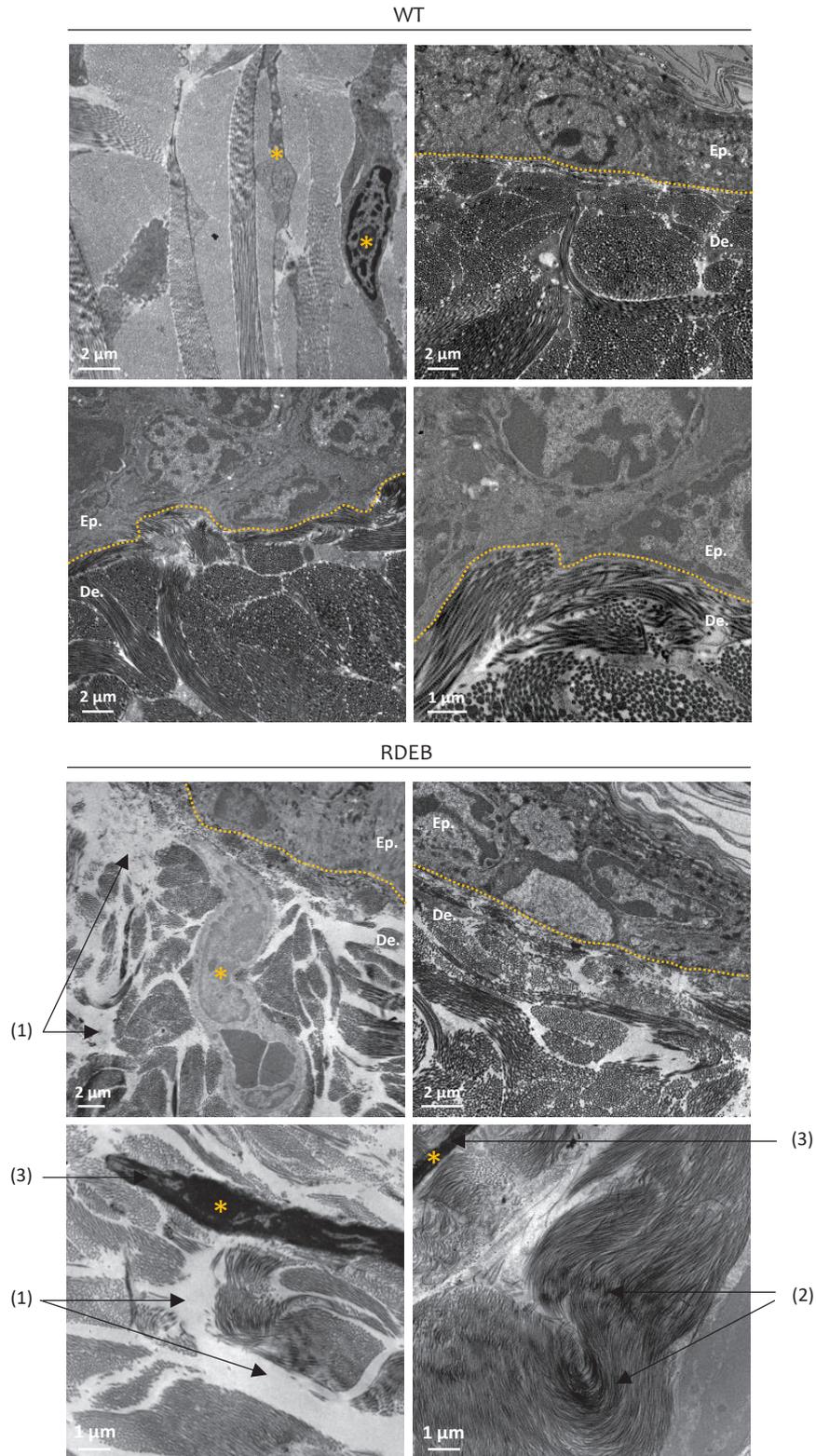
In the skin of RDEB forepaws (9-12 week-old), frequent blistering of the epidermis was observed (**Figure 30A**) and the *lamina densa* seemed thinner, as observed in back skin (**Figure 30B**). Hemidesmosomes were smaller in both WT and RDEB samples (mean lengths of 186 and 120 nm respectively, **Figure 30C**) compared to back skin (**Figure 29B**). Overall, among all the analyzed sample types (skin and paw), the length of hemidesmosomes ranged from 120 to 139 nm in RDEB, while it varied from 165 to 196 nm in WT. In RDEB paw skin compared to WT paw skin, the lowering in the length of hemidesmosomes was measured to be around 35 %. The global effect on the basement membrane was stronger compared to back skin, with a reduction in the proportion of the basement membrane covered with hemidesmosomes reaching 36 % (**Figure 30D**, compared to - 22 % for the back at the same age).

Together, it appears that the deficiency in collagen VII triggers profound modifications of the basement membrane and of the epidermal side of the DEJ. The size and distribution of hemidesmosomes along the junction are affected very early in RDEB disease and the dysregulations are further strengthened by disease progression. These results also justify the use of mouse back skin as a model of RDEB early stages as DEJ disorganization is also clearly visible in this location (although less symptomatic than in forepaws when observed from the outside).

## ii. Characterization of the collagen network in RDEB skin

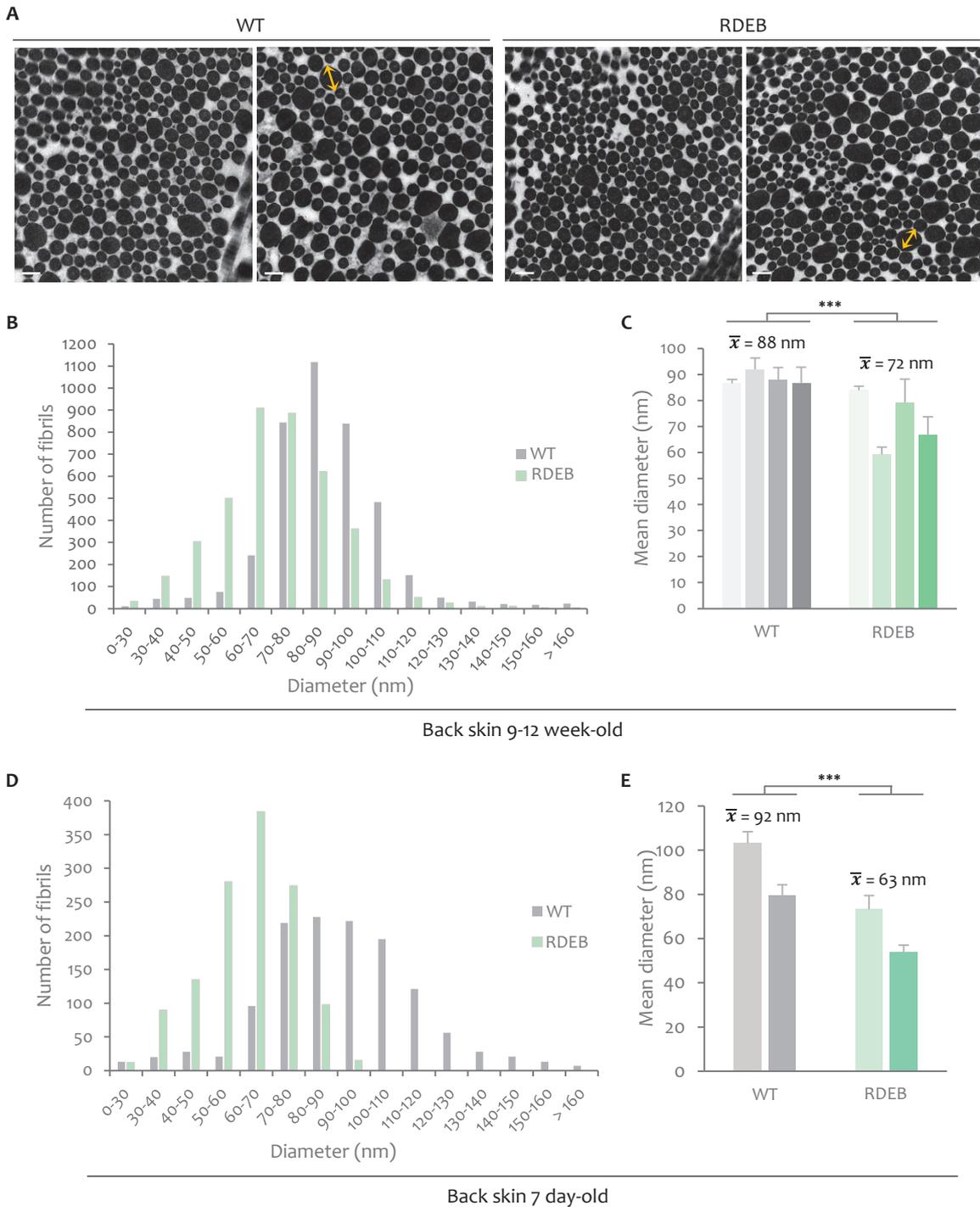
### *Characterization of the collagen network in the back skin*

To evaluate the relevance of the dysregulations detected at the molecular level for collagens, the ultrastructure of dermal ECM was next analyzed in mice of the same age (9-12 weeks, **Figure 31**). In WT dermis, the well-described and typical organization comprising sparse fibroblasts surrounded by large and dense bundles of collagen fibrils was observed. The latter have local orientations but the change in orientation from one bundle to another seems to occur randomly, allowing the tight entanglement of all bundles and the strengthening of the overall structure. In RDEB mice, the dermal collagen network appears locally dislocated: fibers are separated from each other by large “empty” (i.e. less electron-dense) spaces and abnormal bending of fibers can be observed with, for example, unusual 180° abrupt turns, as exemplified in **Figure 31**. Finally, cells with very dense nucleus content were observed, suggesting condensed chromatin.



**Figure 31. TEM observations of the collagen network in the back skin of WT and RDEB mice aged 9-12 weeks.**

Representative TEM images were selected to show the matrix of the dermis. Clear abnormalities were observed in RDEB skin: (1) less dense organization of collagen bundles, (2) atypical collagen fiber bending and (3) fibroblasts with very dense nuclear content. The yellow line shows the basement membrane and the yellow stars indicate fibroblasts. Ep: epidermis, De: dermis.



**Figure 32. Analysis of collagen fibrils by TEM in the back skin of WT and collagen VII-hypomorphic mice.** **A.** Representative pictures of transverse sections of collagen fibrils that can be found in the back skin of mice aged 9-12 weeks. Bar: 200 nm. **B.** Distribution of collagen fibril diameters in mice aged 9-12 weeks. Samples from 4 WT and 4 RDEB mice were analyzed to measure the diameter of 4,000 fibrils for each genotype (measured from 5 different images/mouse). **C.** Mean diameter of collagen fibrils ( $\pm$  SD) in individual WT and RDEB mice aged 9-12 weeks. **D.** Distribution of the diameters of collagen fibrils in mice aged 7 days. Samples from 2 WT and 2 RDEB mice were analyzed to measure the diameter of 1,288 fibrils for each genotype (measured from 5 different images/mouse). **E.** Mean diameter of collagen fibrils ( $\pm$  SD) in WT and RDEB mice aged 7 days. Statistics: Mann-Whitney non parametric test. \*\*\*:  $p < 0.001$ .

Next, we focused more specifically on collagen fibrils (**Figure 32A**). Very strikingly, the fibrils of hypomorphic skin samples were smaller than the WT ones: more small fibrils ( $< 70$  nm) were observed and the distribution was globally shifted toward smaller diameters (**Figure 32B**). Finally, a 20 % decrease in the mean diameter of collagen fibrils was observed (**Figure 32C**) and all RDEB individuals had a mean fibril diameter which was smaller than the mean diameter measured for WT individuals.

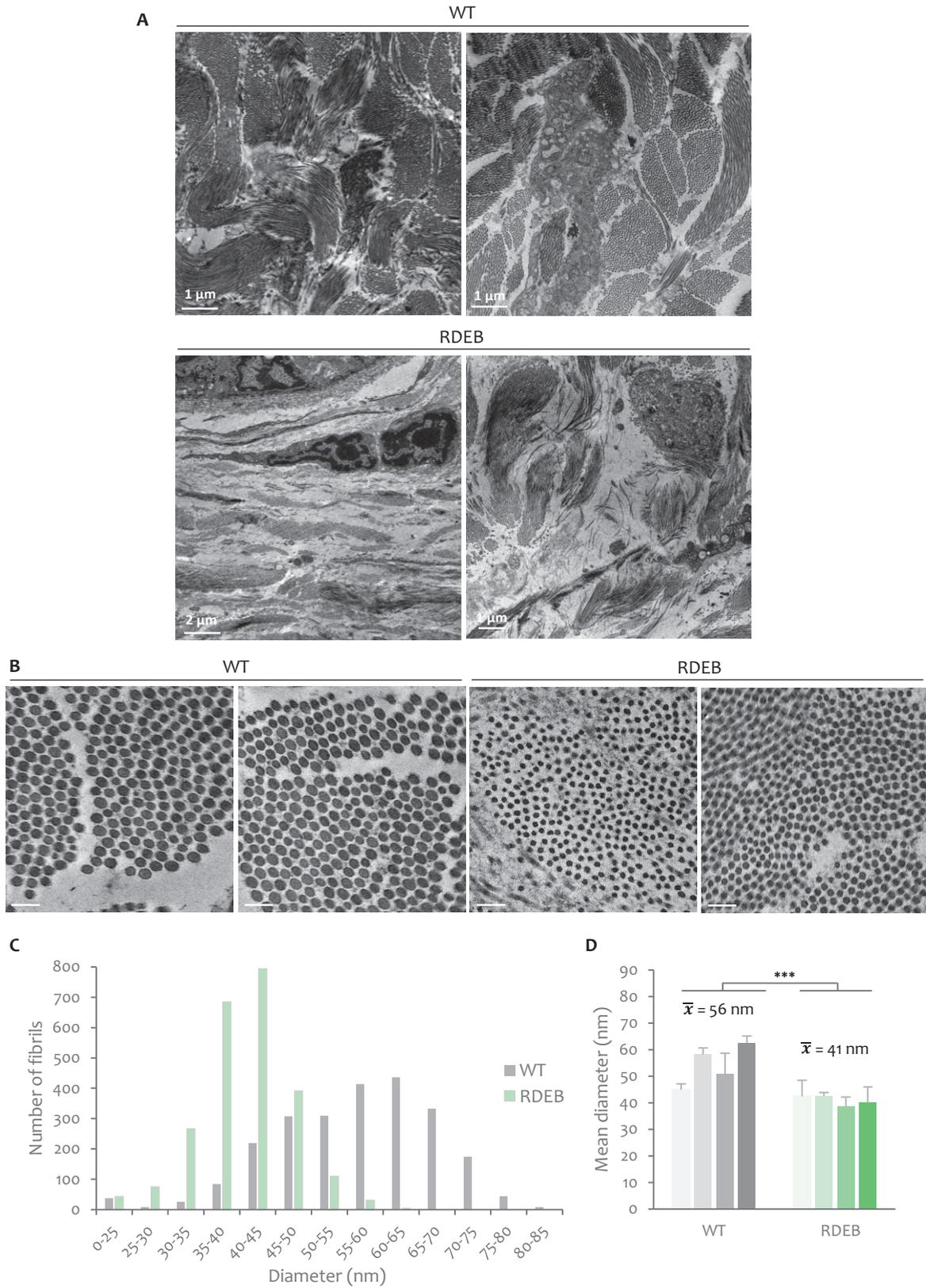
The back skin of 7 day-old mice was finally analyzed in order to get information on the very early stage of the disease. Interestingly, the decrease in fibril diameters was also detected at the early stage of the disease. The analysis of the fibrils in the skin of 7 day-old mice indeed showed that the shift of the diameter distribution detected in older mice was also observed in young RDEB animals (**Figure 32D**). The decrease in mean fibril diameter was even somewhat stronger than in older mice (around 30 %) (**Figure 32E**).

#### *Characterization of the collagen network in the paw skin*

As previously done for the characterization of hemidesmosomes, we next focused on skin samples originating from forepaws to evaluate the effect of disease progression on the network of collagens.

In the so-called “fibrotic” dermis of forepaws from mice aged 9-12 weeks, as also observed in back skin, the overall impression was significantly different in WT and RDEB individuals. Collagen bundles were usually smaller in size and separated by large “empty” spaces (**Figure 33A**). Numerous isolated collagen fibrils could also be observed, much more often in RDEB forepaws than in WT forepaws. Fibroblasts with condensed chromatin were also found in RDEB forepaws, as previously noticed in back skin. Whatever the genotype, dermal collagen fibrils were smaller in forepaw skin than in back skin (Means diameters of 56 nm and 88 nm for WT forepaws and WT back skin respectively, **Figure 33B**, **Figure 33C**). However, a further and very significant reduction of collagen fibril diameter was also evidenced in RDEB forepaws (29 %) (**Figure 33D**).

**Figure 33. Analysis of collagen fibrils in the skin of the forepaw of WT and RDEB mice aged 9-12 weeks.** Next page. **A.** Representative TEM images of forepaw dermis in 9-12 week-old WT and RDEB mice. The disorganization of the collagen network is very obvious with smaller collagen bundles, more isolated fibrils and large empty spaces. **B.** Representative pictures of transverse sections of collagen fibrils in the skin of fibrotic paws from 9-12 week-old mice, showing the clear decrease in fibril diameter. Bar: 200 nm. **C.** Distribution of collagen fibril diameters. Samples from 4 WT and 4 RDEB mice aged 9 to 12 weeks were analyzed to measure the diameter of 2,400 fibrils for each genotype (measured from 4-6 different images/mouse). **D.** Mean diameter of collagen fibrils ( $\pm$  SD) in WT and RDEB mice. Statistics: Mann Whitney non parametric test. \*\*\*:  $p < 0.001$ .



In summary, major alterations in the structure of the dermal matrix, especially in the network of fibrillar collagens, were evidenced by TEM observations. The disorganization of fibers and the diminution of diameters were detected in both early and late stages of the disease, although more spectacular at the late stage. This result suggests that the decreased abundance and abnormal organization of collagen VII directly or indirectly interferes with the elaboration of the collagen network very early during skin development, and that defects fail to normalize in adults.

#### 4. Discussion

In this part of the study, we conducted the first matrisome-focused analysis of the skin of collagen VII-hypomorphic mice, combining two complementary approaches (MS-based quantitative proteomics and TEM). The results obtained underlined major alterations in the organization of the matrix, both in the DEJ and in the dermis, at the molecular and tissue scales. In addition, our proteomic analysis highlighted marked modifications in the regulation of proteolytic activities in the skin of RDEB animals.

The disorganization of the DEJ observed by TEM was expected. Collagen VII is indeed a key component of the basement membrane and its drastic reduction in RDEB was previously reported to alter the ultrastructure of the DEJ in mouse and human forms of the disease<sup>140,211</sup>. Here, we first observed that the decreased abundance in collagen VII was accompanied by a general increase in the abundance of other basement membrane components. Although these changes were not significant at the quantitative level, it was remarkable that all the DEJ components that could be quantified (collagens IV, VI and XV, laminin-511 chains, nidogen-1, perlecan) showed an increased abundance in our proteomic analysis, especially in the insoluble fractions reflecting the stably-deposited proteins. Surprisingly, no abundance change was detected for any of these components when RDEB mouse skin lysates were analyzed by MS/MS<sup>212</sup>. These results possibly suggest alterations of the intracellular processing of basement membrane proteins in RDEB that could compensate at whole organ level those we observed in the extracellular space. In human, somewhat divergent results were previously obtained when the conditioned medium and ECM of RDEB and control dermal fibroblasts were compared, leading to a global decrease in nidogen-1, collagen IV and laminin-511<sup>146</sup>. Similarly, in cultures of human keratinocytes, the binding partner of collagen VII laminin-332 was found to be down-regulated in RDEB cells, at both the RNA and protein level<sup>213</sup>. In contrast, when whole skin biopsies from RDEB patients were analyzed by transcriptomics<sup>145</sup>, an increase in the expression of *Col4a2* and *Col6a3* was also detected, suggesting that the impairments consecutive to the loss of collagen VII may differ depending on the type of samples (human skin vs mouse skin or human skin vs human cell cultures) or on the selected analytical approach (RNA vs proteins). This highlights the difficulty of this type of omics approaches when applied to relatively small numbers of human samples with high heterogeneity (linked to different mutations, ages, disease stages, skin sites, etc.). In this context, one advantage of the mouse model is that it is more homogeneous and that adequate controls in terms of age and skin sites can be

used. In addition, we worked with 5 individuals/genotype to really identify the strongest tendencies. However, I must also point that the dysregulations reported here were not significant from a statistical point of view, it is only because all BM components show the same general trend that we were able to draw a conclusion. In the end, it appears relatively logical that compensation mechanisms may occur in response to the impairment of anchoring fibrils, especially in early disease stages, through the over-expression or over-deposition of other basement membrane proteins.

In addition to the impairments in matrix proteins in the DEJ, we also evidenced defects affecting the hemidesmosomes (length and abundance along the basement membrane). Interestingly, a brief characterization of the hemidesmosomes in the unwounded skin from 13 and 9 patients suffering from DDEB and severe RDEB, respectively, also reported a trend to decrease in the number of hemidesmosomes along the basement membrane (- 10 to - 17 %, depending on the body site)<sup>110</sup>. One surprising result was that, in RDEB mouse samples, the hemidesmosomes were rather stable in size whatever the anatomic location and the stage of the disease, and never exceeded 140 nm in length, possibly suggesting that the pathology limits their growth to a certain size. In contrast, in WT mice the length of hemidesmosomes exhibited more age- and site-dependent variations. If collagen VII does not interact directly with hemidesmosomes, its decrease may exert effects on the epidermis through laminin-332, a matrix component connecting collagen VII and integrin  $\alpha 6\beta 4$  at the cell surface of basal keratinocytes and exhibiting an impaired distribution in RDEB mouse skin<sup>144</sup>. Unfortunately, the chains of laminin-332 were not quantified in our proteomic analysis. Another protein which was detected during the optimizations of ECM enrichment protocols described in the previous part but not in this study was collagen XVII, a transmembrane collagen expressed by keratinocytes that is a central component of hemidesmosomes. This information will be important to obtain in the future even if the absence of detection in two independent experiments could reflect a decreased abundance in RDEB skin. A closer examination of changes induced by collagen VII deficiency within the basement membrane and in the epidermal compartment is currently performed by our German collaborators and by Patricia Rousselle.

Collagen VII also interacts with collagen fibrils on the dermal side of the DEJ<sup>214</sup> but the extent of the observed modifications in the amount and organization of fibrillar collagens was beyond what we had predicted. Our main results were a decreased abundance of collagens I and III (mostly in the soluble fractions for collagen III but in both soluble and insoluble fractions for collagen I) and several abnormal features in the ultrastructure of collagen fibrils, including a very strong reduction in their diameters, a less compact organization and frequent abnormal bending. The minor fibrillar collagen V was differently affected with less marked changes in abundance and even a slight increase in the final fraction of the chemical decellularization protocol. This is interesting as increased collagen V in fibrils has been largely associated to decreased fibril diameters<sup>215</sup>. Overall, the results pointing to diminished levels of major fibrillar collagens were however somewhat counter-intuitive as RDEB is presently described as a

fibrosis-promoting disease and one would rather expect to see increased fibrillar collagens. This at least what was suggested by images taken under polarized light after picosirius red staining which showed increased staining in RDEB mouse skin at the late stages<sup>141,212</sup>, indicating larger or more aligned fibers. Published expression data on fibrillar collagens however are less conclusive. If the studies using human RDEB and control fibroblasts already mentioned above indicated increased levels for the transcript of collagen  $\alpha 1(I)$ <sup>145</sup> and for the secreted protein chains of collagens  $\alpha 1(III)$  and  $\alpha 2(V)$ <sup>146</sup>, a recent paper from our German collaborators did not evidence any significant variations in the abundance of fibrillar collagens in whole back skin lysates (without ECM enrichment) obtained from the same RDEB mouse model as the one we used<sup>212</sup>. These apparent discrepancies and the striking changes evidenced by TEM were the two main reasons why we decided to focus on the possible mechanisms explaining the disruption of the normal collagen organization in the final part of our study (**Chapter III, III**).

Among the proteins which placed collagen fibril organization as the top dysregulated biological process in our proteomic study, the SLRPs (small leucine rich proteoglycans) were the second most prominent contributors after fibrillar collagens. This is explained by the fact that one of their main function is to control collagen fibril diameter and spacing through direct interactions with collagen fibrils<sup>53</sup>. Surprisingly, biglycan and decorin tended to over-abundance in SDS/SD fractions from RDEB mice, while their amount was decreased in the insoluble extracts from the two ECM enrichment protocols (very strong trends, significant decrease for biglycan in Comp. Ins. fraction), possibly underlying variations in solubility rather than variations in total protein abundance. In addition, their reduced abundance in the insoluble matrix is probably not directly associated to the reduction in fibril diameter since the opposite effect (increased diameters) was reported in the skin of *Bgn* and *Dcn* KO mice<sup>216,217</sup>. Nevertheless, SLRPs were previously suggested to act as disease modifiers in RDEB, especially through the modulation of signaling events. In particular, variations in the amount of decorin were proposed to influence disease severity in homozygous human twins affected by RDEB, with the highest expression of the proteoglycan being associated with the lowest severity of symptoms<sup>115</sup>. The ability of decorin to limit TGF- $\beta$  signaling by sequestering the growth factor in the matrix<sup>61,62</sup> could be an explanation to these observations. Similarly, in the RDEB mouse model, since biglycan and decorin can both sequester TGF- $\beta$  in the matrix<sup>62</sup>, their early decrease could contribute in the long term to promote disease progression.

Regarding proteolysis in the matrix, our data suggest a significant decrease in the inhibition of serine proteases and a more moderate increase in the inhibition of cysteine proteases in RDEB mouse skin. In the literature, elevations in the amounts of TIMP-2 and TIMP-3 involved in the negative regulation of the activity of matrix metalloproteinases were also reported in skin explants from RDEB patients<sup>218</sup>. More generally, proteolysis is thought to play a major role in RDEB progression. In human RDEB keratinocytes, the lysosomal proteases cathepsins B and Z were found to be up-regulated<sup>213</sup>. Increases

in the abundance and activity of MMP-2 at the protein level were also reported in different human models of RDEB, including cultures of fibroblasts<sup>146</sup> and skin explants from patients<sup>218</sup>. Augmented levels of MMP-9 and MMP-14 were also suggested but the correlation with an increased proteolytic activity remained unclear for these enzymes<sup>146</sup>. In mouse skin finally, no variation in the amounts of MMP-2 and MMP-9 were detected by MS-based proteomics<sup>212</sup>, again underlying the difficulty to draw clear-cut and general conclusions. However, this will be important to decipher as MMPs can contribute to collagen and proteoglycan degradation, two types of core matrix proteins which were diminished in our mouse model. In addition, MMPs can participate in TGF- $\beta$  activation, a mechanism which was recently shown to be relevant in RDEB using a co-culture model where fibroblasts and keratinocytes from either RDEB or control donors were mixed together<sup>219</sup>. In this context, I have made some attempts to perform N-terminomics on ECM-enriched mouse skin samples in order to get a better overview of the dysregulated proteolytic processes in RDEB. Some interesting preliminary results were obtained but additional optimizations are still required to get reliable results and increase the depth of the analysis.

In conclusion, our results indicate that it is necessary to better define what fibrosis means in RDEB. Fibrosis is classically described as the pathological accumulation of ECM in soft tissues resulting from an increase in the activity of pro-fibrotic growth factors such as TGF- $\beta$  (**Chapter I, III-3-ii**)<sup>150,220</sup>. If we have no reason to question the role of TGF- $\beta$ , the concept of “excessive” collagen deposition is apparently not so obvious in RDEB, at least in the mouse model, suggesting that it should be replaced by the concept of “aberrant” collagen accumulation. A specific form of fibrosis seems to be at play in RDEB and it will be important to better characterize its main characteristics to improve patient management. Interestingly, other forms of skin fibrosis also show some heterogeneity. This is the case for example in systemic sclerosis. If the increase in abundance of fibrillar collagens was not put into question in this disease<sup>221,222</sup>, an increase in fibrils with small diameters was also reported in patient skin biopsies<sup>223,224</sup>. Furthermore, in human keloid lesions, a significant decrease in both mature collagen  $\alpha 1(I)$  and collagen VII was reported by MS-based proteomics using an ECM-enrichment protocol similar to our chemical decellularization method<sup>225</sup>. Finally, the fact that altered organization of dermal collagens and DEJ is observed very early in the RDEB disease suggests that these structural alterations directly result from the absence of collagen VII rather than from inflammation or defective wound healing. Consequently, they fail to resolve with time and rather tend to worsen with disease progression.

### III. PRELIMINARY EVALUATION OF THE MECHANISMS INVOLVED IN THE COLLAGEN PHENOTYPE IN RDEB SKIN

In the previous part, we evidenced several alterations in the composition and organization of the ECM in RDEB skin, both in the DEJ and in the underlying dermis. The modifications in the organization of dermal collagen fibrils were especially intriguing. Firstly, to the best of our knowledge, they had never been reported before; secondly, collagen fibrils were affected throughout the whole dermis, not only in the upper layer which is in direct contact with collagen VII anchoring fibrils; thirdly, the extent of disorganization in the paws was more prominent than in the back skin, suggesting that collagen fibrils could also reflect disease progression. Finally, the features of this collagen disorganization did not fit with the typical features of fibrotic skin. The dermal part of the skin appeared loosened with a strong reduction in collagen fibril diameters and this result was associated with an overall decrease in the amount of the most abundant skin fibrillar collagens (I and III) in proteomics data. For all these reasons, we decided to have a closer look at fibrillar collagens and at some of the partners required for their biosynthesis and for fibrillogenesis in order to find further mechanistic explanations about the observed “collagen phenotype”.

Different hypotheses likely to explain the alterations of fibril diameters and organization can be made (**Table 15**). For example, modifications in the processing of fibrillar collagens (maturation, cross-linking) and in collagen organizers can be hypothesized.

	Hypothesis	Reference
(1)	Modifications in the proportion of collagens I, III and V in fibers	215,226
(2)	Retention of N- and/or C- propeptides of collagen I, III and V	227–230
(3)	Modifications in “collagen organizers” (SLRPs, FACITs)	216,217,231
(4)	Dysregulations in degradation processes	232,233
(5)	Alterations in post translational modifications (PTM) and cross-linking	152,234

**Table 15. Possible hypotheses to explain the alterations of the collagen network in RDEB mouse skin.**

The hypotheses (1) and (2) were investigated and the results are presented in this part.

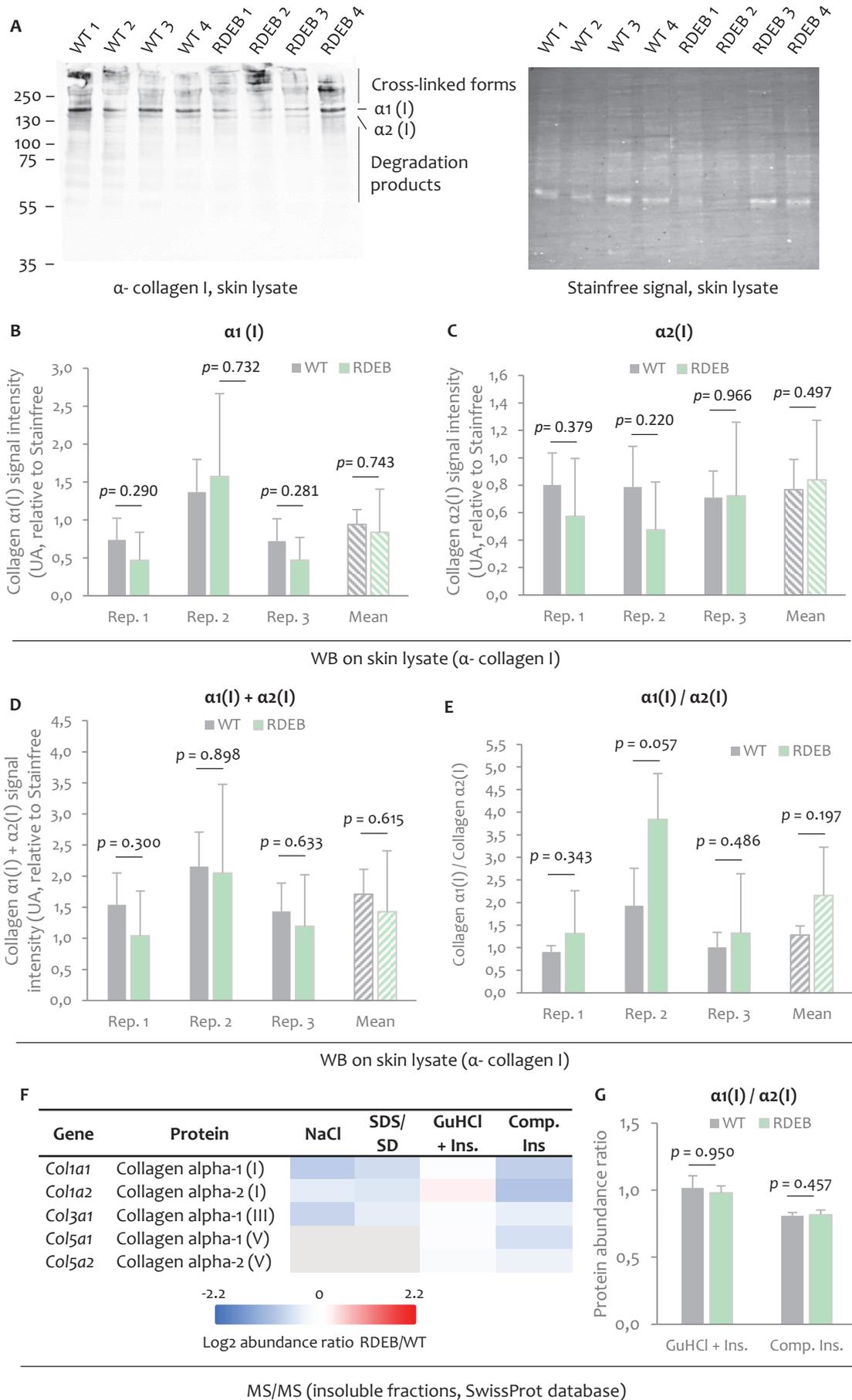
#### 1. Analyses of mature collagen forms

The results obtained by TEM indicated that collagen fibrils displayed abnormal organization and diameters in RDEB skin. The heterotypic fibrils found in skin are made of various proportions of mature collagens I, III and V. However, the proteomics results were not completely clear regarding the trends for these collagens. If the  $\alpha 1$  and  $\alpha 2$  chains of collagen I were found to be significantly decreased in the insoluble fraction of the compartmental protocol, the decrease was not confirmed in the insoluble

fraction of the chemical decellularization protocol and was much more moderate for the  $\alpha$  chains of collagen III and V (**Figure 25, Chapter III, II-2**). Consequently, we first analyzed the mature forms of collagens more in detail.

As a first approach, we investigated possible changes in collagen I abundance through Western blot (WB) on skin lysates. In the next pages, “skin lysate” will refer to a total skin sample boiled in electrophoresis sample buffer; other skin extraction samples are named according to the fractions obtained after ECM enrichment as in previous parts of this chapter. Four WT and 4 RDEB samples (originating from individuals distinct from the ones used in the previous part) were analyzed in reducing conditions using an antibody targeting the triple helix of the two chains of collagen I. As expected, this led to a complex pattern with the major bands around 150 kDa corresponding to monomeric and mature  $\alpha 1(I)$  and  $\alpha 2(I)$  collagen chains (**Figure 34A**), higher molecular weight bands attributed to cross-linked forms of collagen I and a “ladder” below 130 kDa probably corresponding to degradation products generated by collagen turnover. When the  $\alpha 1(I)$  and  $\alpha 2(I)$  free mature chains were quantified in three technical replicates and normalized to total protein (through parallel Stainfree detection on the same membrane) and to a reference WT biological replicate (randomly chosen, **Chapter II, VI-4**), no significant difference in signal intensity was detected, neither for the two chains separately (**Figure 34B, Figure 34C**), not for the total signal (sum of  $\alpha 1(I)$  and  $\alpha 2(I)$  signals, **Figure 34D**). Other bands could not be used for quantification as they are more diffuse and, in some cases, overexposed. Similarly, cross-linked forms were predominant when collagen I was blotted in the insoluble fractions of the ECM enrichment protocols and free  $\alpha 1(I)$  and  $\alpha 2(I)$  chains were not defined clearly enough to perform quantification (data not shown).

We next wanted to determine if the ratio between these two chains was affected in RDEB skin. The presence of homotrimeric  $[\alpha 1(I)]_3$  collagen I has indeed been reported in some types of cancers<sup>57</sup> and fibrotic tissues<sup>58</sup>, leading to the alteration of the  $\alpha 1/\alpha 2$  ratio. Based on the same Western blots as above, no significant variation of this ratio was observed in skin lysates (**Figure 34E**), at least for mature collagen  $\alpha 1(I)$  and  $\alpha 2(I)$  monomers, nor a possible trend of evolution from WT to RDEB.



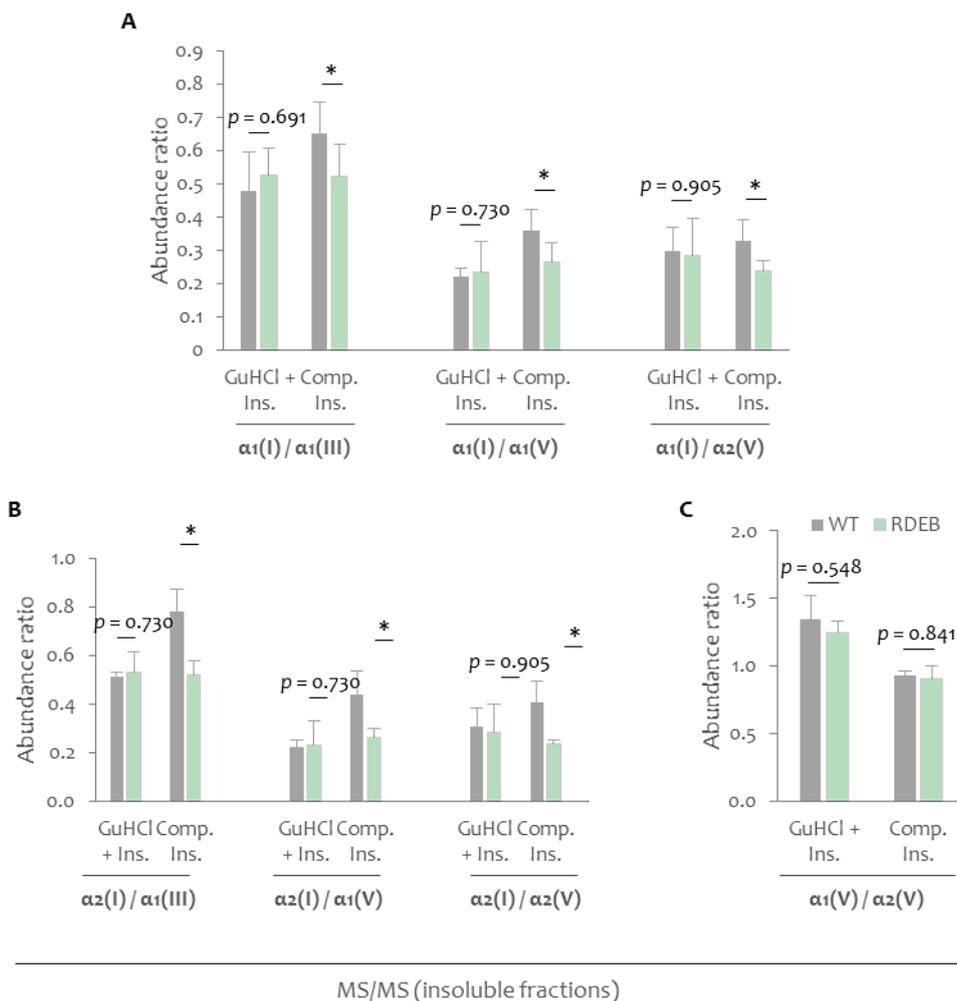
**Figure 34. Immuno- and MS/MS-detection and quantification of collagen I in the back skin of WT and RDEB mice aged 9-12 weeks.**

**A.** Immuno-detection of collagen I in skin lysates (reducing conditions, antibody directed against the triple-helical domain,  $n = 4$  mice/genotype). Three technical replicates were analyzed, only one representative membrane (or technical replicate) is shown. **B.** Quantification of total mature collagen  $\alpha 1(I)$  based on the signal of the free mature chain. The mean signal per genotype ( $\pm$  SD) was analyzed for each of the 3 technical replicates and for the pool of the technical replicates (double normalization as described in **Chapter II, V-1-iii**). **C.** Quantification of total mature collagen  $\alpha 2(I)$  based on the signal of the free mature chain. **D.** Quantification of total mature collagen I ( $\alpha 1(I) + \alpha 2(I)$  chains) based on the signal of the free mature chains (not including cross-linked forms). **E.**  $\alpha 1(I)/\alpha 2(I)$  ratio computed from WB signals. **F.** Heatmap of the RDEB/WT abundance ratio of the main dermis fibrillar collagens in mouse skin (Log<sub>2</sub> scale,  $n = 5$  mice/genotype, SwissProt database used for protein identification). The RDEB/WT ratios and the heatmap were obtained as previously described for the processing of Uniprot data (see **Figure 25, Chapter III, II-2-vi**). Grey rectangles: protein not detected in the fraction. **G.** Quantification of the  $\alpha 1(I)/\alpha 2(I)$  ratio using MS/MS signals ( $n = 5$  mice/genotype). Statistics: Student t-test or Mann-Whitney test when t-test was not applicable. \*:  $p < 0.05$ . WB: Western blot. Rep: technical replicate. Mean: mean of the mean values obtained from the 3 technical replicates for each individual mouse.

In parallel, previously obtained proteomics data were used to calculate the  $\alpha 1(I)/\alpha 2(I)$  ratio in the mature insoluble collagen network through the analysis of the insoluble fractions from the ECM enrichment protocols. In order to circumvent the problem raised by multiple proteoforms co-existing for a single protein that complicates the analysis (e.g. Q01149 and Q3TUE2 for the  $\alpha 2(I)$  chain), an additional processing of the MS/MS data using the SwissProt database (reviewed sequences only) instead of UniProt (reviewed and unreviewed sequences) for protein identification was performed in order to get only one proteoform per protein. The variations in protein abundance between WT and RDEB genotypes obtained with these data were first compared to those obtained previously using UniProt. Identical trends in the evolution of the amounts of fibrillar collagens were observed, indicating that the new dataset (**Figure 34F**) reflected the previous one (**Figure 25, Chapter III, II-2-ii**) and could be used for further analyses. When the  $\alpha 1(I)/\alpha 2(I)$  ratios obtained for WT and RDEB samples were compared (**Figure 34G**), they displayed relatively low standard deviations (for the two genotypes and the various fractions), suggesting that the computing of the  $\alpha 1(I)/\alpha 2(I)$  ratio using MS/MS data gives reproducible results between biological replicates. Again, no significant difference was seen between WT and RDEB mice for the  $\alpha 1(I)/\alpha 2(I)$  ratio calculated from the quantitative results of the insoluble fractions using the two selected enrichment protocols, confirming Western blot results. Finally, the computed ratios slightly differed from the expected value of 2, even in WT samples, and were found to be around 1. In Western blot, this may result from a differential affinity of the antibody for the two collagen chains and, in mass spectrometry, the different nature of the ions generated from the two chains is possibly responsible for the observed differences in their detection and quantification.

In order to evaluate the possible modification in the proportions of collagens I, III and V in fibrils, we next compared the abundances of collagen I individual chains with those of other fibrillar collagens in the insoluble extracts obtained after ECM enrichment. The  $\alpha 1(I)/\alpha 1(III)$ ,  $\alpha 1(I)/\alpha 1(V)$  and  $\alpha 1(I)/\alpha 2(V)$  ratios were all significantly reduced in RDEB samples in the insoluble fraction of the compartmental protocol (very insoluble part of the ECM), suggesting a decrease in the relative amount of collagen  $\alpha 1(I)$

in the fibrils (**Figure 35A**). The same observations were made when the  $\alpha 2(I)$  chain was compared to the chains of collagen III and V (**Figure 35B**). However, different results were obtained with the insoluble fraction of the chemical decellularization protocol in which no significant difference was found for the various ratios. One possible explanation could be that with the latter protocol, the so-called “insoluble” fraction is actually a mix of the final insoluble pellet (supposedly equivalent to the insoluble fraction of the compartmental protocol) and of some less tightly ECM-associated proteins extracted with the chaotropic GuHCl solution. It can thus be hypothesized that possible changes in the most insoluble collagen forms are somewhat masked by the absence of changes in the GuHCl-extracted forms. To go further, it would therefore be necessary to use another approach such as Western blot to detect and quantify collagens III and V. Unfortunately, I didn’t have the time to establish their efficient immunodetection and compare WT and RDEB samples. Finally, it can be observed that there is no change in the  $\alpha 1(V)/\alpha 2(V)$  ratio when WT and RDEB samples are compared with the two enrichment methods (**Figure 35C**), probably indicating that the normal  $[(\alpha 1)_2\alpha 2]$  collagen V heterotrimer is the dominant form in fibrils of both genotypes.



**Figure 35. Abundance ratios between the different chains of fibrillar collagens in WT and RDEB insoluble enrichment fractions.**

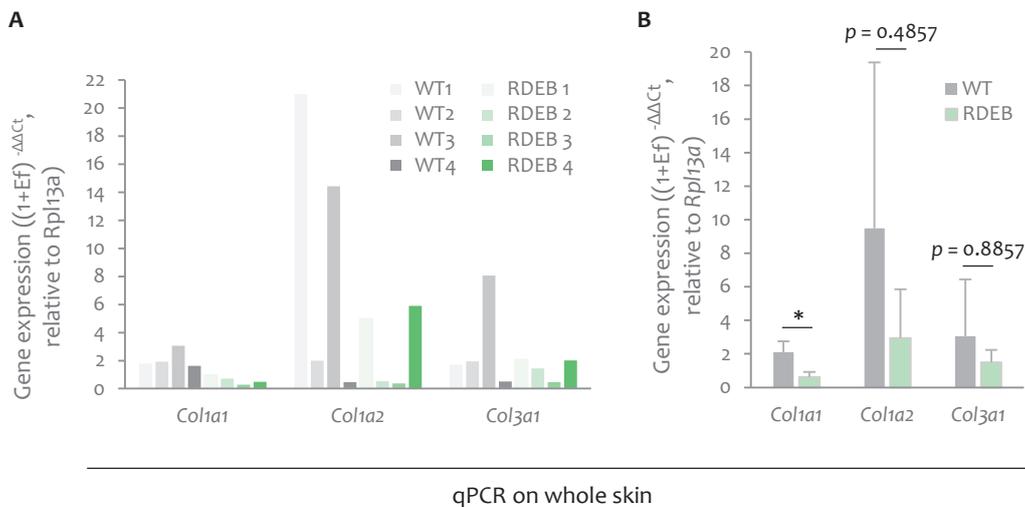
Previous page. **A.** Comparison of the abundance ratios between the  $\alpha 1(I)$  chain and chains of other fibrillar collagens found in skin. **B.** Comparison of the abundance ratios between the  $\alpha 2(I)$  chain and chains of other fibrillar collagens found in skin. **C.** Comparison of the abundances ratio between the two chains of collagen V. Abundance ratios were computed from MS/MS quantitative data ( $n = 5$  mice/genotype, 9 to 12 weeks, SwissProt database used for protein identification). SwissProt database was used to identify proteins. Statistics: Mann-Whitney test. \*:  $p < 0.05$ . Comp Ins.: insoluble pellet of the compartmental enrichment method; GuHCl + Ins.: fraction obtained by the combination of GuHCl and insoluble extracts of the chemical decellularization method.

**2. Expression of procollagens I and III in RDEB skin**

In our proteomic analyses, diminutions in the amount of fibrillar collagens I and III were also detected in the soluble extracts (NaCl or SDS/SD fractions) of the chemical decellularization protocol, suggesting that soluble fibrillar procollagens could also be affected by collagen VII deficiency. As a first step to explore this possibility, we next characterized the production of procollagens in the skin.

**i. Expression of the genes encoding collagens I and III**

The expression of the genes encoding collagen  $\alpha 1(I)$ , collagen  $\alpha 2(I)$  and collagen  $\alpha 1(III)$  was first evaluated through RT-qPCR on mRNA transcripts from whole mouse back skin.

**Figure 36. Expression of the genes encoding fibrillar  $\alpha 1(I)$ ,  $\alpha 2(I)$  and  $\alpha 1(III)$  collagen chains in the back skin of WT and RDEB mice aged 9-12 weeks.**

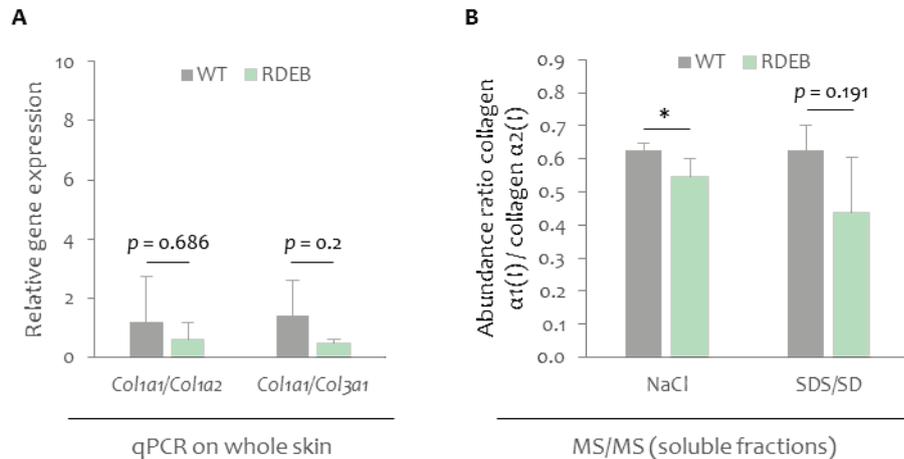
**A.** Individual gene expression. Gene expression was evaluated by RT-qPCR on mRNA transcripts ( $n = 4$  mice/genotype, 3 technical replicates). The  $\Delta\Delta C_t$  method was used to compute gene expression in WT and RDEB samples. **B.** Comparison of gene expression between the WT and RDEB genotypes. Statistics: Mann-Whitney test. \*:  $p < 0.05$ . Ef: efficiency of primer pairs.

Interestingly, the expression of *Coll1a1* was significantly decreased in RDEB samples, by more than 70 %, when compared to WT. The analysis of mRNA transcripts also revealed that while inter-individual variability was rather low for *Coll1a1* expression and moderate for *Col3a1* expression (Figure 36A), it was much higher for *Coll1a2* expression. Therefore, the similar trend observed for *Coll1a2* did not reach

significance (**Figure 36B**). However, we will see below that it is probably not modified to the same extent as *Colla1*. Finally, *Col3a1* expression did not seem to be modified.

## ii. Expression of procollagen I at the protein level

To go further, qPCR data were used to compute the *Colla1/Colla2* and *Colla1/Col3a1* ratios (**Figure 37A**). Decreasing trends were observed for all the ratios, but no significant variation was detected due to the inter-individual variations in the expression of *Col3a1* and *Colla2*. The  $\alpha 1(I)/\alpha 2(I)$  ratio was next computed at protein level in the soluble fractions, using proteomics data as already done above for the insoluble extracts. In contrast to what we observed for mature collagens, the  $\alpha 1(I)/\alpha 2(I)$  ratio was significantly decreased in the NaCl fraction of RDEB back skin, suggesting that the decrease in the amount of secreted procollagens was stronger for collagen  $\alpha 1(I)$  than for collagen  $\alpha 2(I)$  (**Figure 37B**). These results indirectly confirm the observations made on the expression of *Colla1* (significant) and *Colla2* (not significant, **Figure 36B**). In the SDS/SD fraction, only a decreasing trend was observed. As collagen  $\alpha 1(III)$  was identified from a relatively small number of peptides in the soluble extracts (protein coverage of 2 %), no comparison with this chain at the protein level was performed.



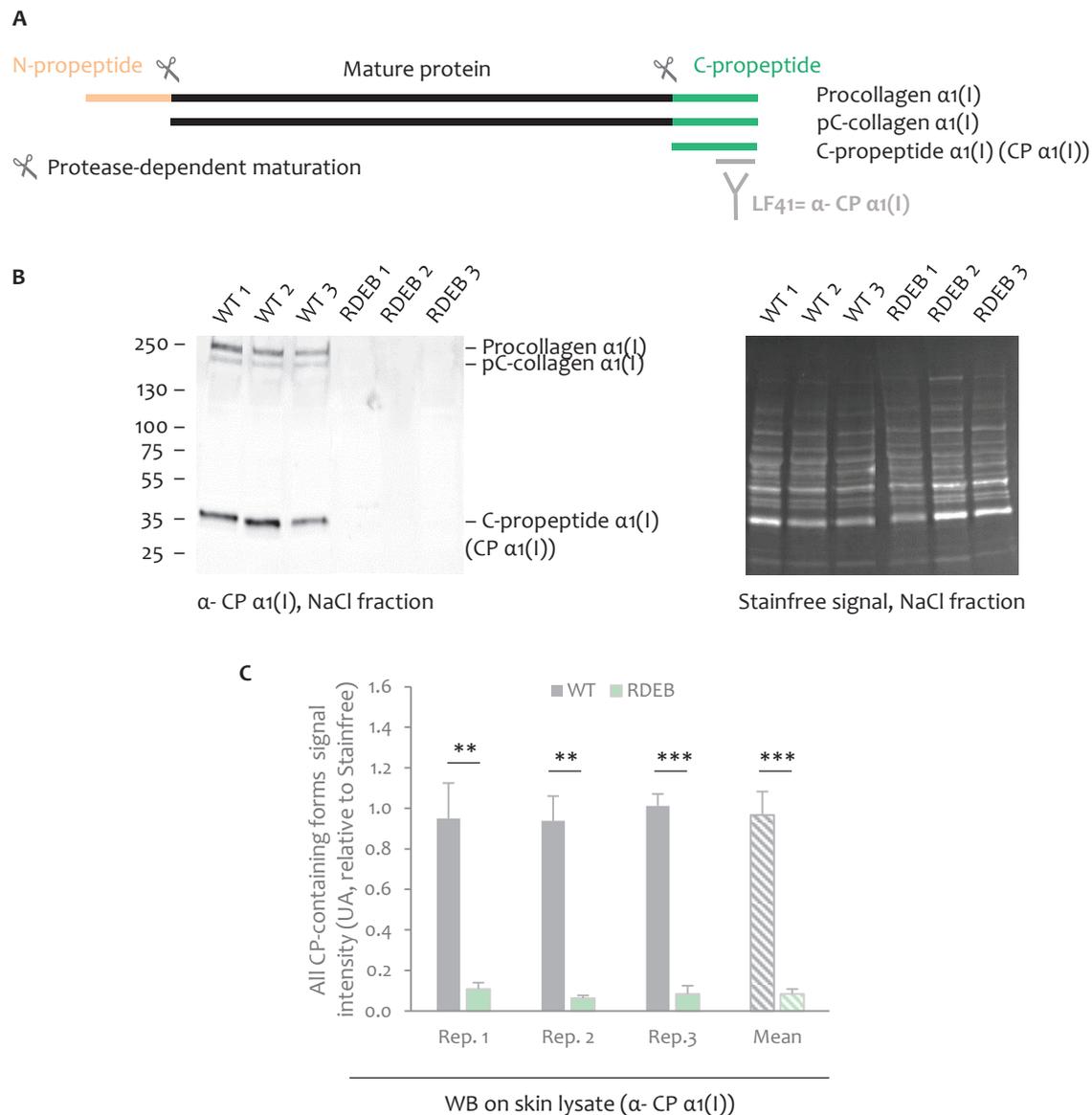
**Figure 37. Abundance ratio between collagen  $\alpha 1(I)$ ,  $\alpha 2(I)$  and  $\alpha 1(III)$  at transcript and protein levels in the back skin of WT and RDEB mice aged 9-12 weeks.**

**A.** Gene expression ratios calculated from qPCR data ( $n = 4$  mice/genotype,  $\Delta\Delta C_t$  method). **B.** Protein abundance ratio computed from MS/MS quantitative data ( $n = 4$  WT and 5 RDEB, SwissProt database used for protein identification). Statistics: Mann-Whitney test. \*:  $p < 0.05$ . Ef: efficiency of primer pairs.

We next focused on procollagen  $\alpha 1(I)$  which seemed to be the most dysregulated. To confirm the results obtained in MS-based proteomics and in RT-qPCR, we evaluated the expression of the protein by Western blot on various WT and RDEB mouse samples.

To do so, we used a well-characterized antibody (LF41)<sup>198</sup> directed against the C-terminal sequence of the procollagen  $\alpha 1(I)$  C-propeptide that allows the detection of full-length procollagen I as well as of two important products derived from the N-terminal (pC-collagen  $\alpha 1(I)$ ) and C-terminal (C-propeptide or CP  $\alpha 1(I)$ ) proteolytic maturations (**Figure 38A**). Very strikingly, if the CP  $\alpha 1(I)$ -containing forms

could be easily detected in the NaCl extracts from WT samples, almost no signal was detected in the same type of RDEB samples (**Figure 38B**). A significant decrease in the abundance of all CP  $\alpha 1(I)$ -containing proteins was consequently evidenced in RDEB (**Figure 38C**). This clearly demonstrated that procollagen  $\alpha 1(I)$  production was severely altered and confirmed (and even expanded) the MS/MS and RT-qPCR results described above.



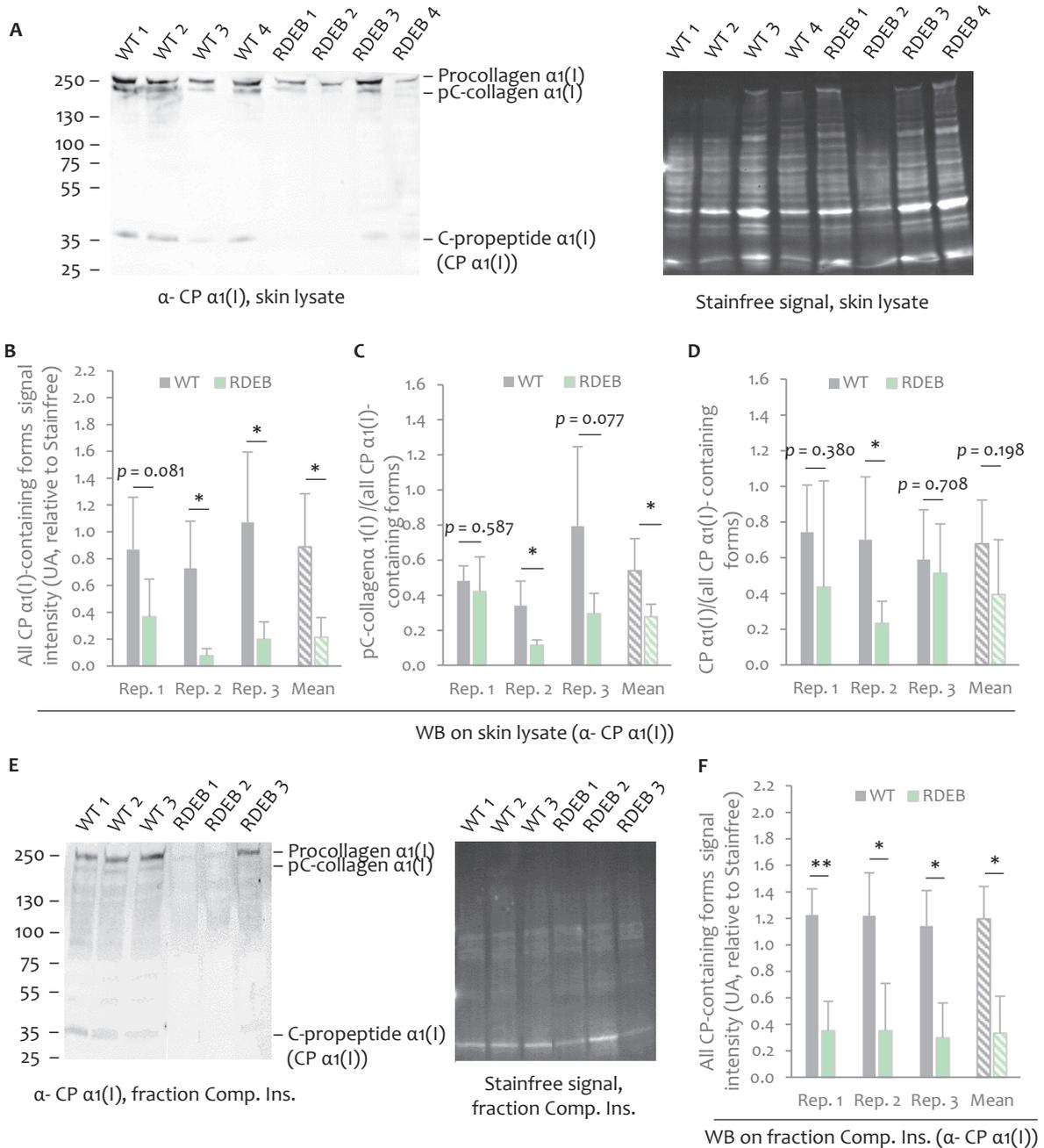
**Figure 38. Immuno-detection of procollagen  $\alpha 1(I)$  and of other C-propeptide  $\alpha 1(I)$ -containing forms in the NaCl fraction obtained from the back skin of WT and RDEB mice aged 9-12 weeks.**

**A.** Schematic representation of the protease-dependent maturation of procollagen  $\alpha 1(I)$  and of the two derived fragments (pC-collagen  $\alpha 1(I)$  and C-propeptide, CP  $\alpha 1(I)$ ) that can be detected with an anti-CP  $\alpha 1(I)$  antibody (epitope location on protein sequence is also indicated). **B.** Immuno-detection of all CP  $\alpha 1(I)$ -containing forms in the NaCl fraction obtained through the chemical decellularization protocol (reducing conditions, LF41 antibody,  $n = 3$  mice/genotype). **C.** Quantification of all CP  $\alpha 1(I)$ -containing forms in the NaCl fraction. The mean signal per genotype ( $\pm$  SD) was analyzed for each of the 3 technical replicates and for the pool of the technical replicates (double normalization). Statistics: Student t-test or Mann-Whitney test when t-test was not applicable. \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ . WB: Western blot. Rep: technical replicate. Mean: mean of the mean values obtained from the 3 technical replicates for each individual mouse.

When the same experiment was conducted on skin lysates (**Figure 39A**), CP  $\alpha 1(I)$ -containing forms were more readily detected in RDEB skin samples than in NaCl fractions. The decrease in their abundance in RDEB was significant in two out of three technical replicates and in the pool of technical replicates (**Figure 39B**). We also took advantage of the higher level of these forms in skin lysates, compared to NaCl fractions, to determine if N- and C-terminal proteolytic maturations were deficient in RDEB skin. When the ratio between pC-collagen  $\alpha 1(I)$  and all-CP  $\alpha 1(I)$  containing forms was calculated from Western blot signals, a decreasing trend was observed for RDEB skin which was significant in one technical replicate and in the pool of all technical replicates (**Figure 39C**). When the ratio of the free CP  $\alpha 1(I)$  over all CP  $\alpha 1(I)$ -containing forms was next calculated, a significant decrease was detected in only one of the 3 technical replicates (**Figure 39D**). It should be noted however that the low levels of signal for pC-collagen  $\alpha 1(I)$  and CP  $\alpha 1(I)$  in RDEB samples make quantification less reliable. We finally tried to detect CP  $\alpha 1(I)$  in the insoluble fraction obtained from compartmental enrichment (**Figure 39E**). Surprisingly, procollagen  $\alpha 1(I)$  and pC-collagen  $\alpha 1(I)$  forms were detected in all WT fractions, suggesting that incompletely-matured forms of collagen I can be integrated into fibrils. Although this goes against the dogma that pro- and pC-collagens cannot assemble into fibrils<sup>235</sup>, there are now some examples in the literature where C-propeptide retention in fibrils has also been observed<sup>236–238</sup>. Very unexpectedly, traces of free CP  $\alpha 1(I)$  were also detected, suggesting that CP  $\alpha 1(I)$  can be retained into the insoluble matrix. In RDEB samples, only traces of procollagen  $\alpha 1(I)$  were detected and the amount of CP  $\alpha 1(I)$ -containing proteins was significantly decreased in 2 of the 3 all technical replicates and in the pool of replicates (**Figure 39F**). As no other CP  $\alpha 1(I)$ -containing forms were detected, no evaluation of the fraction represented by pC-collagen  $\alpha 1(I)$  and CP  $\alpha 1(I)$  among all CP  $\alpha 1(I)$  -containing forms, as done above with skin lysates, could be performed here. Notably, we also tried to detect procollagen  $\alpha 1(I)$  with an antibody directed against the N-propeptide  $\alpha 1(I)$  (LF39<sup>198</sup>) but failed to obtain satisfactory results (procollagen I  $\alpha 1(I)$  itself was not detected, data not shown).

At this point, we can conclude that one major feature of RDEB skin is a decreased expression of *Coll1a1* resulting in a diminution of the production of procollagen  $\alpha 1(I)$ . Interestingly, the analysis of the C-propeptide  $\alpha 1(I)$  also suggested two additional alterations: (1) a reduction in the amount of all CP  $\alpha 1(I)$ -containing forms (including pC-collagen  $\alpha 1(I)$  and CP  $\alpha 1(I)$  proteolytic products) that is stronger in the extracellular environment analyzed through the NaCl and insoluble fractions than in the whole tissue including the cellular compartment and (2) a decreasing trend in the proportion of pC-collagen  $\alpha 1(I)$  (and possibly CP  $\alpha 1(I)$ ) to all CP  $\alpha 1(I)$ -containing forms that is not completely explained by the overall decrease in full-length procollagen. These observations could result from dysregulations in processes other than procollagen expression in RDEB skin, e.g. a decreased secretion or an altered N- and/or C-terminal processing of procollagen I  $\alpha 1(I)$ .

Due to the limitations of Western blots discussed above to precisely monitor collagen proteolytic maturations in the various extracellular compartments, we decided to use other approaches to determine if these important steps in collagen fibrillogenesis regulation, especially the maturation processes, were modified. This is important since mutations affecting collagen I, III or V maturation sites or associated enzymes are known to lead to severe forms of Osteogenesis Imperfecta<sup>239,240</sup> and Ehlers-Danlos<sup>241</sup> syndromes in human. They can affect both collagen organization and collagen fibril diameters<sup>239,240</sup> and suggest that deficient processing could also represent one potential mechanism to explain the collagen phenotype in RDEB skin.



**Figure 39. Immuno-detection of procollagen  $\alpha 1(I)$  and of other C-propeptide  $\alpha 1(I)$ -containing forms in the back skin of WT and RDEB mice aged 9-12 weeks.**

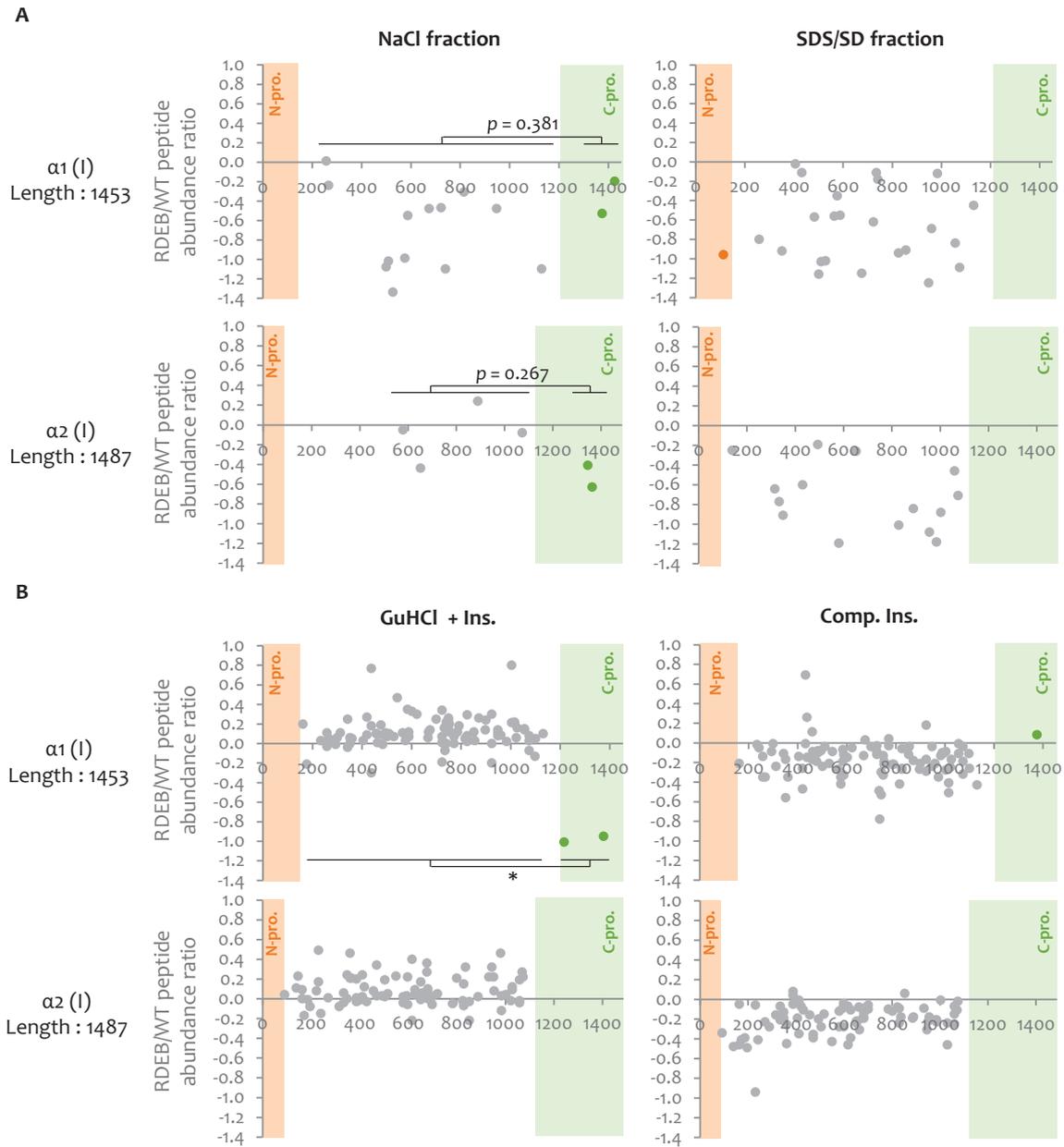
Previous page. **A.** Immuno-detection of all CP-containing forms in skin lysates (reducing conditions, LF41 antibody,  $n = 4$  mice/genotype). Three technical replicates were analyzed, only one representative membrane (or technical replicate) is shown. **B.** Quantification of all CP  $\alpha 1(I)$ -containing forms in skin lysate. The mean signal per genotype ( $\pm$  SD) was analyzed for each of the 3 technical replicates and for the pool of the technical replicates (double normalization). **C.** Ratio between pC-collagen  $\alpha 1(I)$  and all CP  $\alpha 1(I)$ -containing forms computed from WB signal in skin lysates. **D.** Ratio between CP  $\alpha 1(I)$  and all CP  $\alpha 1(I)$ -containing forms computed from WB signal on skin lysates. **E.** Immuno-detection of all CP  $\alpha 1(I)$ -containing forms in the insoluble fraction from the compartmental enrichment protocol (Comp. Ins., reducing conditions, LF41 antibody,  $n = 3$  mice/genotype). Three technical replicates were analyzed, only one representative membrane (or technical replicate) is shown. **F.** Quantification of all CP  $\alpha 1(I)$ -containing forms in Comp. Ins. fraction. The mean signal per genotype ( $\pm$  SD) was analyzed for each of the 3 technical replicates and for the pool of the technical replicates (double normalization). Statistics: Student t-test or Mann-Whitney test when t-test was not applicable. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ . WB: Western blot. Rep: technical replicate. Mean: mean of the mean values obtained from the 3 technical replicates for each individual mouse.

### 3. Evaluation of the protease-dependent maturations of fibrillar collagens

#### i. MS/MS and Western blot data for the analysis of proteolysis

First, MS/MS data were further exploited to determine if peptides belonging to N- and C-propeptides behaved similarly to peptides from the collagen central part (triple helix + telopeptides) from a quantitative point of view. First, peptide mapping was used for the  $\alpha 1(I)$  and  $\alpha 2(I)$  collagen chains in the various ECM enrichment fractions analyzed by MS/MS. In the NaCl fraction, which contains the most soluble forms of collagens (full-length procollagens and partially-matured precursors), only two peptides were identified in CP  $\alpha 1(I)$  and  $\alpha 2(I)$  (**Figure 40A**) but this number was coherent with the overall sequence coverage of the procollagen chains (CP  $\alpha 1(I)$  and  $\alpha 2(I)$  peptides represent 13 % and 33 % of identified unique peptides respectively, to be compared with 17 % and 18 % of the total chain sequences corresponding to the C-propeptide domains). CP  $\alpha 1(I)$  and  $\alpha 2(I)$  peptides displayed fold-changes which were comparable to peptides in the triple helix (leading to an overall decreased abundance in RDEB). No CP  $\alpha 1(I)$  and  $\alpha 2(I)$  peptide was found in the SDS/SD fraction but one peptide was identified in the N-propeptide  $\alpha 1(I)$ , leading to the same conclusion about the similar behavior of propeptides compared to the rest of the sequence.

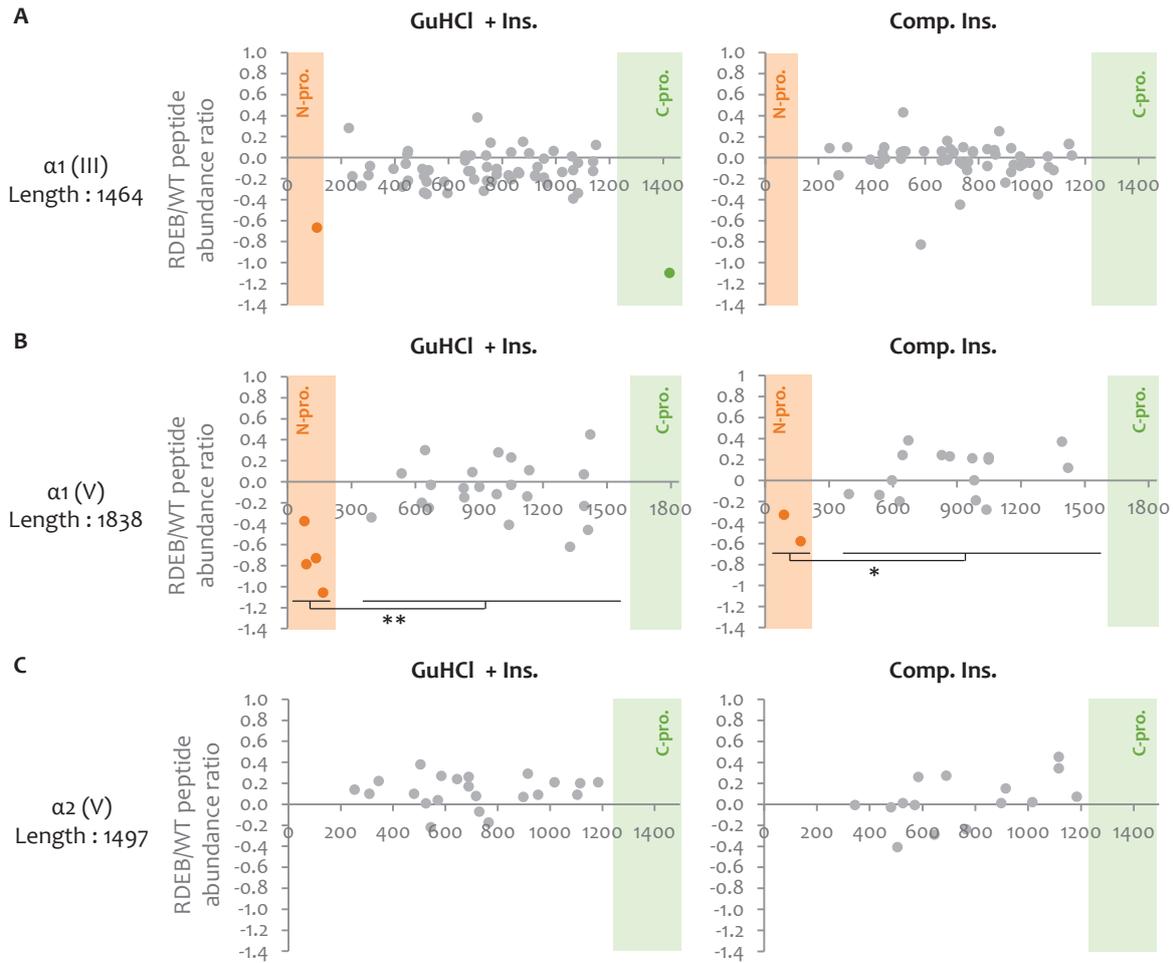
Peptides belonging to CP  $\alpha 1(I)$  were also identified in the insoluble extracts of the two enrichment protocols, consistently with the WB results obtained above showing that CP  $\alpha 1(I)$  was detected in the insoluble fraction from the compartmental enrichment protocol (**Figure 40B**). However, the two types of insoluble fractions gave divergent results with a significant decrease in fold-changes for the two CP  $\alpha 1(I)$  peptides identified in the GuHCl + Ins. fraction and no trend to change for the unique CP  $\alpha 1(I)$  peptide identified with the compartmental protocol. However, based on Western blot results of the compartmental insoluble fraction which indicate a strong decrease in all CP  $\alpha 1(I)$ -containing forms in RDEB, it seems that the GuHCl + Ins. is more in agreement with the general trends observed in this work, even if the significant result was based on only two peptides from the C-propeptide  $\alpha 1(I)$  and consequently needs to be interpreted with caution.



**Figure 40. Peptide mapping for collagens  $\alpha 1(I)$  and  $\alpha 2(I)$  in the enrichment fractions obtained from the back skin of WT and RDEB mice aged 9-12 weeks.**

**A.** Peptide mapping in soluble extracts. **B.** Peptide mapping in insoluble fractions. MS/MS data from the analysis of 12 week-old WT and RDEB mice were used to compute the RDEB/WT ratio for each identified peptide ( $n = 3$  mice/genotype, database used for the identification of peptides and proteins: SwissProt). Each dot represents a peptide quantified by MS/MS. For each peptide, the RDEB/WT ratio (Log<sub>2</sub> scale) was plotted (y axis) relative to starting amino-acid location (x axis). Orange area: N-propeptide released by BMP-1; white area: triple helix; green area: C-propeptide. Statistics: Mann-Whitney test (at least 2 peptides from propeptides detected). \*:  $p < 0.05$ .

We also analyzed other fibrillar collagens of the dermis in order to evaluate if the decrease in the abundance of C-propeptide-derived peptides observed in insoluble extracts was specific to procollagen  $\alpha 1(I)$ . As previously mentioned, I didn't have enough time to establish the immuno-detection of other fibrillar collagens but could explore the MS/MS data to extract some preliminary conclusions.



**Figure 41. Peptide mapping for collagens  $\alpha 1(\text{III})$ ,  $\alpha 1(\text{V})$  and  $\alpha 2(\text{V})$  in the insoluble fractions obtained from the back skin of WT and RDEB mice aged 9-12 weeks.**

**A.** Peptide mapping for collagen  $\alpha 1(\text{III})$ . **B.** Peptide mapping for collagen  $\alpha 1(\text{V})$ . **C.** Peptide mapping for collagen  $\alpha 2(\text{V})$ . MS/MS data from the analysis of 12 week-old WT and RDEB mice were used to compute the RDEB/WT ratio for each peptide ( $n = 3$  mice/genotype, database used for the identification of peptides and proteins: SwissProt, ). Each dot represents a peptide quantified by MS/MS. For each peptide, the RDEB/WT ratio (Log2 scale) was plotted (y axis) relative to starting amino-acid location (x axis). Orange area: N-propeptide; white area: triple helix; green area: C-propeptide. Statistics: Mann-Whitney test (at least 2 peptides from propeptides detected). \*:  $p < 0.05$ , \*\*:  $p < 0.01$ .

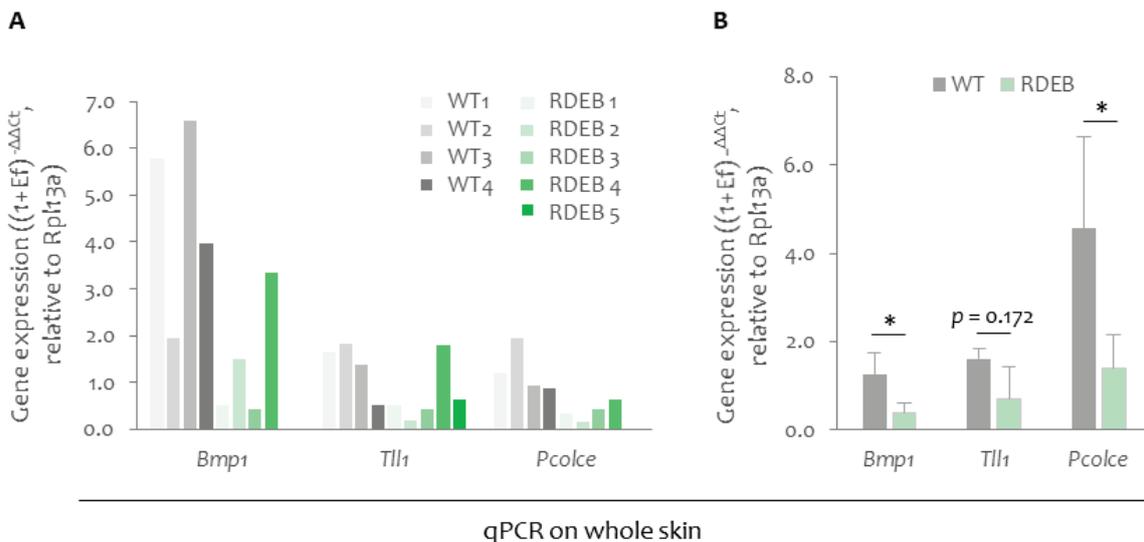
Peptide mapping was performed on insoluble extracts only, because no peptide from collagen  $\alpha 1(\text{V})$  and  $\alpha 2(\text{V})$  and only a few from collagen  $\alpha 1(\text{III})$  could be detected in soluble extracts. We first noted that some peptides belonging to the propeptides of collagens  $\alpha 1(\text{III})$  and  $\alpha 1(\text{V})$  could be identified in the insoluble skin matrix, as previously observed for collagen  $\alpha 1(\text{I})$  (Figure 41A, Figure 41B). In contrast, no peptide belonging to the  $\alpha 2(\text{V})$  C-terminal propeptide (Figure 41C) was found (there is no known N-terminal maturation for this chain). Very strikingly, the fold-changes comparing RDEB and WT samples for detected peptides were all strongly diminished compared to the fold-changes of the peptides located in the central part of these collagen chains, most probably indicating less retention in mature collagen fibrils (even though less released propeptides attached to the insoluble matrix could not be completely excluded). This was the case for the two peptides belonging to collagen  $\alpha 1(\text{III})$  propeptides

(one at each extremity, **Figure 41A**) and even more remarkably for the 2-4 peptides detected in the N-propeptide of the  $\alpha 1(V)$  chain (**Figure 41B**, no C-terminal peptides detected). This result was statistically significant when the fold-changes of the four peptides identified in the GuHCl + Ins. fraction were compared to those of the peptides located in the downstream triple helix. In the insoluble fraction from compartmental enrichment, a significant decrease was also observed for the N-terminal peptides identified. Even if this last result was based on two peptides only, the observations made on the N-terminal propeptide of collagen  $\alpha 1(V)$  seemed consistent in the two insoluble fractions. These results suggest a possible dysregulation of the BTP or ADAMTS proteases involved in propeptide removal.

## ii. Expression of procollagen C-proteinases and related partners in RDEB skin

To go further in the analysis of the potentially altered processes of procollagen maturation, we evaluated the expression of enzymes involved in the C-terminal processing of fibrillar procollagens. BMP-1/mTLD and mTLL-1 are known to be the main proteases involved in the release of procollagen I C-propeptides in skin<sup>230,242</sup>. They were never detected in our proteomic analyses, neither during the optimization phase nor in the quantitative comparison of WT and RDEB samples, suggesting a very low abundance in skin. However, Agnès Tessier, a former doctoral student in the lab, made some preliminary experiments suggesting that they could be more abundant in skin biopsies from RDEB patients than in normal skin (immuno-fluorescence, n = 4 donors/group)<sup>197</sup>. If it was also the case in mouse skin, it would supposedly lead to a more efficient conversion of pro- and pC-collagen I into mature collagen I and could explain their decreased abundance in mouse RDEB skin.

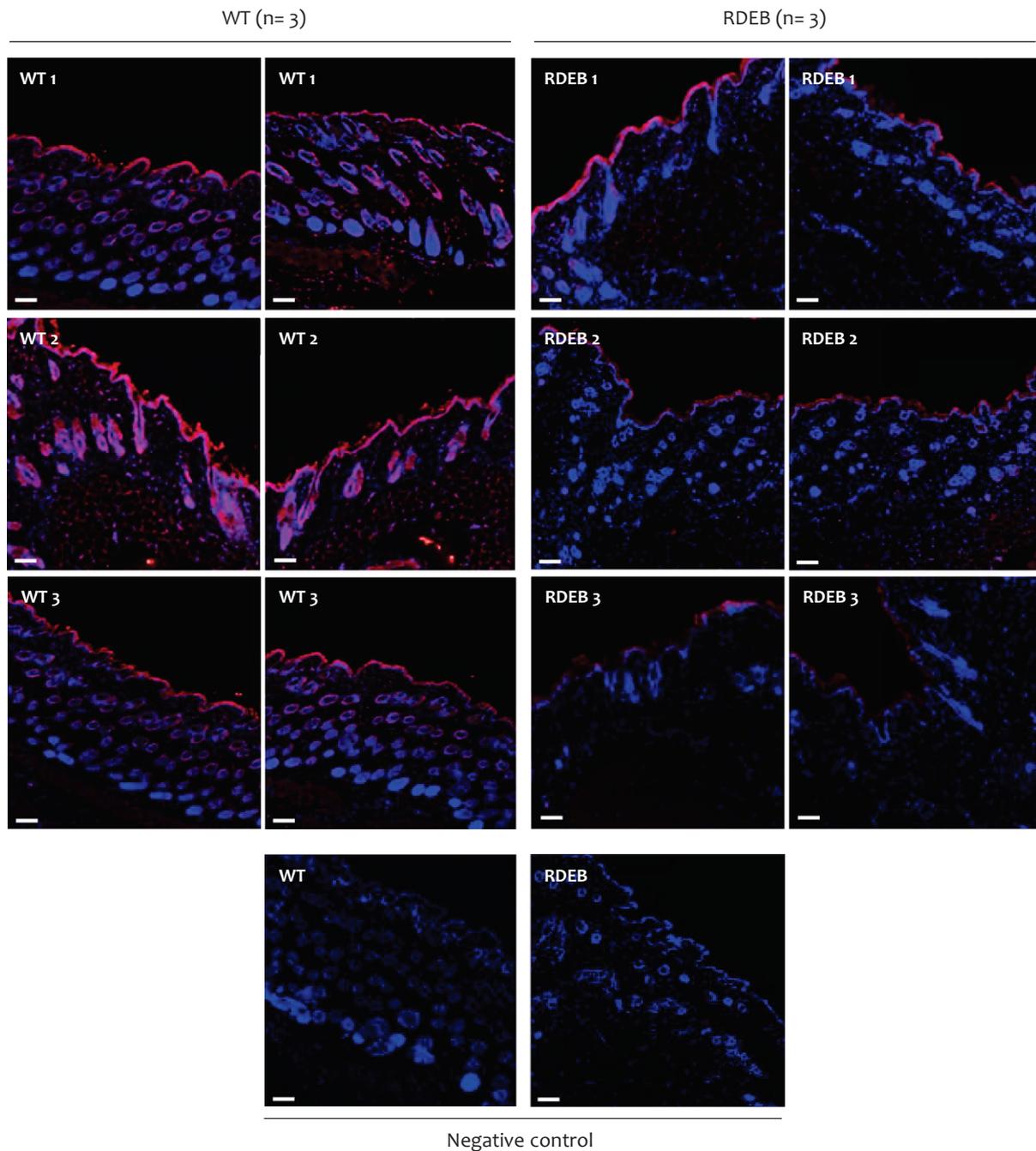
The transcript levels for the *Bmp1* gene (encoding BMP-1 and its splicing variant mTLD), the *Tll1* gene (encoding mTLL-1) and the *Pcolce* gene (encoding PCPE-1, an enhancer of the procollagen C-proteinase activity of BTPs, **Chapter I, II-3**) were first evaluated in WT and RDEB mouse skin by RT-qPCR (**Figure 42A**). Rather unexpectedly, in RDEB skin, the expression of *Bmp1* and *Pcolce* mRNA was diminished by a factor of more than 2 compared to WT skin (**Figure 42B**). Regarding *Tll1*, a clear decreasing trend was also observed but was not statistically significant.



**Figure 42. Expression of the genes encoding BMP-1/mTLD, mTLL-1 and PCPE-1 in the back skin of WT and RDEB mice aged 9-12 weeks.**

**A.** Individual gene expression. mRNA transcripts were extracted from whole skin and analyzed by RT-qPCR ( $n = 4-5$  mice/genotype for *Bmp1* and *Tll1*;  $n = 4$  mice/genotype for *Pcolce*). The  $\Delta\Delta C_t$  method was used to compute gene expression in WT and RDEB samples. **B.** Comparison of gene expression between WT and RDEB genotypes. Statistics: Mann-Whitney non parametric test. ns: not significant, \*:  $p < 0.05$ . Ef: efficiency of primer pairs.

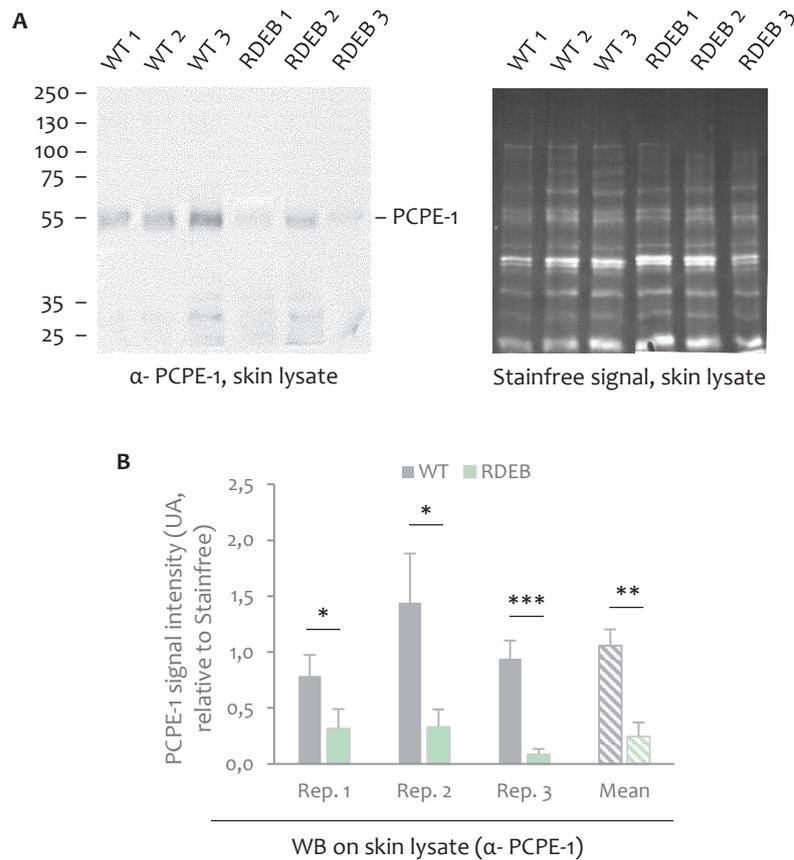
At the protein level, the detection of BTPs by Western blot on skin lysates has been shown by my group to be very challenging due to the low amounts of these proteases in normal conditions. Their detection was indeed unsuccessful in our hands but the proteases could more easily be analyzed by immunofluorescence on skin sections (**Figure 43**). Notably, the antibody which was used for these experiments is a neo-epitope antibody that only recognizes the mature and active forms of BMP-1/mTLD and mTLL-1 (all three are detected by the antibody with a slightly decreased efficiency for mTLL-1) and not the pro-enzymes<sup>243,244</sup>. In WT tissues, staining was observed in the three skin layers: epidermis, dermis and hypodermis. There was also some variation in signal intensity from one sample to another which underlined the variation in protein abundance between individuals. In RDEB dermis however, the signal was consistently weaker than in any of the WT samples. Very strikingly also, the staining of the epidermal layer in hair follicles was almost completely lost in RDEB tissues, suggesting that BTPs change of expression (or activity) could also affect the epidermal compartment. This result could have consequences for the maturation of some basement components such as LM-332 which is already very affected<sup>144</sup> by collagen VII deficiency (itself also a potential substrate of BTPs). Overall, the staining was clearly weaker in RDEB mouse skin, suggesting a decrease in the global amount of BMP-1 and mTLL-1 active proteins that is consistent with the observations made by RT-qPCR.



**Figure 43. Immunofluorescent labeling of the main BMP-1/tolloid-like proteinases (BMP-1/mTLD and mTLL-1) in the back skin of WT and RDEB mice aged 9-12 weeks.**

BTPs (red, home-made antibody<sup>244</sup>) and the nuclei of cells (blue, DAPI) were stained on 5- $\mu$ m thick skin sections from 3 WT and 3 RDEB mice (2 images shown for each mouse). Negative control: no antibody. Bar: 90  $\mu$ m.

Regarding PCPE-1, the protein was first analyzed by Western blot on skin lysates (**Figure 44A**). Significant decreases in the abundance of the protein in RDEB were detected in all the technical replicates (**Figure 44B**). The reduction was estimated to reach 75 % compared to WT samples (pool of the replicates).



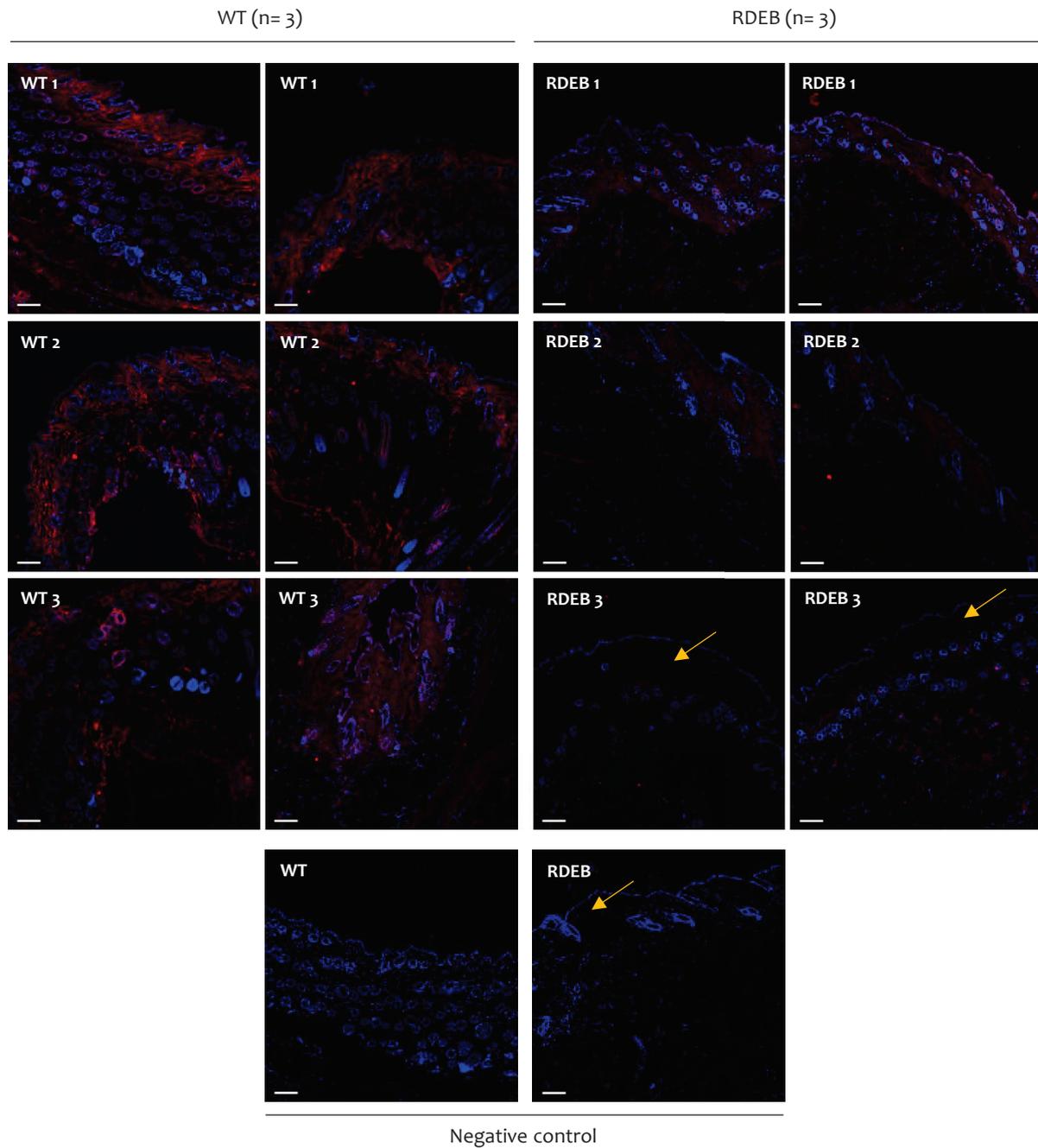
**Figure 44. Immuno-detection of PCPE-1 in skin lysates from WT and RDEB mice aged 9-12 weeks.**

**A.** Immuno-detection of PCPE-1 in skin lysate (reducing conditions, AF2239 antibody, n = 3 mice/genotype). Three technical replicates were analyzed, only one representative membrane (or technical replicate) is shown. **B.** Quantification PCPE-1 in skin lysate. The mean signal per genotype ( $\pm$  SD) was analyzed for each of the 3 technical replicates and for the pool of the technical replicates (double normalization). Statistics: Student t-test. \*:  $p < 0.05$ , \*, \*\*  $p < 0.01$ , \*\*\*:  $p < 0.001$ . WB: Western blot. Rep: technical replicate. Mean: mean of the mean values obtained from the 3 technical replicates for each individual mouse.

PCPE-1 was also stained by immunofluorescence on skin sections to confirm this result (**Figure 45**). A strong staining was observed in the dermis for all WT samples but no staining in the epidermis can be seen. In RDEB skin, the protein was also detected in the dermis but the signal was weaker than in WT. The staining was particularly low in one of the 3 RDEB samples analyzed where epidermal blisters were also observed. Taken together, the data obtained by Western blot and immunofluorescence indicated that the decrease in the expression of the *Pcolce* gene resulted in a diminution of protein abundance in the skin of RDEB mice.

In conclusion, reductions in the abundance of BTPs and PCPE-1 in RDEB mouse skin were clearly evidenced. These results could fit with the global decrease in free C-propeptide  $\alpha 1(I)$  evidenced by WB in soluble extracts after enrichment. However, they are more difficult to connect with the possible decrease in C-propeptide  $\alpha 1(I)$  retention which was observed in RDEB insoluble matrix by MS/MS and it remains difficult to determine to what extent decreased collagen expression and deficient processing

both contribute to the decreased amount of CP  $\alpha 1(I)$  in skin. At this stage, it seems that decreased collagen expression is dominant as the overall reduction in procollagen I is really drastic.



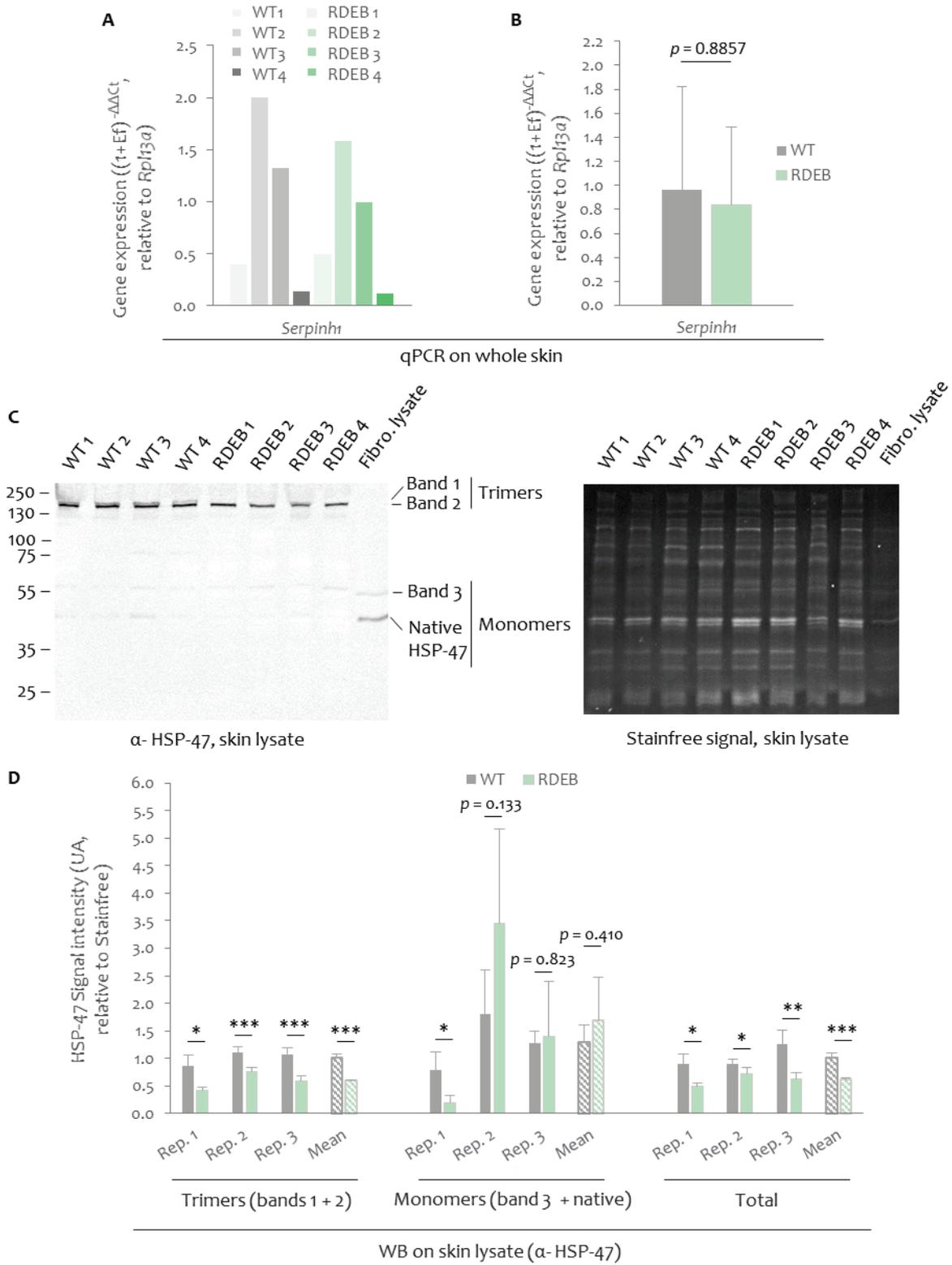
**Figure 45. Immunofluorescent labeling of PCPE-1 in the skin of WT and RDEB mice aged 9-12 weeks.** PCPE-1 (red, AF2239 antibody) and the nuclei of cells (blue, DAPI) were stained on 5- $\mu$ m thick skin sections from 3 WT and 3 RDEB mice. Yellow arrows indicate epidermis blistering in RDEB samples. Negative control: no antibody. Bar: 100  $\mu$ m.

### iii. Preliminary evaluation of the expression of HSP-47

We finally decided to evaluate the possible change in HSP-47 abundance. HSP-47 (heat-shock protein-47) is a collagen-specific chaperone involved in the control of collagen folding in the endoplasmic reticulum. It transiently binds to procollagens and prevents local unfolding and/or aggregate formation of immature collagen molecules<sup>207</sup>. It was one of the most strongly dysregulated hits in our proteomic data and was decreased in all analyzed fractions. Interestingly, HSP-47 deficiency has previously been associated with abnormal fibril formation<sup>229</sup>. Fibrils were thin, branched and displayed collagen I N-propeptide retention possibly due to delayed triple helix formation and secretion. A deficiency in HSP-47 could therefore contribute to the collagen phenotype observed in RDEB samples.

First, the expression of *Serpinh1*, the gene encoding HSP-47, was evaluated in the back skin of WT and RDEB mice by RT-qPCR. High inter-individual variability in gene expression was observed in the two genotypes (**Figure 46A**) and no significant modification could be evidenced (**Figure 46B**)

The detection of HSP-47 in skin lysates was next performed by Western blot to compare with the results obtained by MS/MS which indicated a general reduced abundance in all fractions. Various forms of HSP-47 were surprisingly detected in both WT and RDEB samples (**Figure 46C**). Monomeric HSP-47 (47 kDa) was detected (although to a rather low level), as well as a previously-described potential glycosylated form of the monomer (band 3)<sup>245</sup>. However, the strongest signals were observed around 150 kDa, in the form of two bands (1 and 2) with very similar molecular weights that may correspond to trimers of HSP-47, as previously described in the literature<sup>245,246</sup>, with and without glycosylations respectively. These forms were absent from a control human fibroblast cell lysate. The presence of trimers was unexpected because all lysates (human cell and mouse tissue) were prepared similarly prior to the SDS-PAGE electrophoresis (same sample buffer and reducing conditions). However, trimers of recombinant mouse HSP-47 were previously shown to be very stable; they remain partially folded up to 95 °C and are insensitive to 8 M urea but, from literature, are not expected to resist to reducing conditions<sup>245</sup>. Intriguingly, a significant diminution in the amount of trimeric HSP-47 was observed in RDEB tissue lysates while no variation was found for the monomeric forms of the protein (**Figure 46D**). When the summed signal of all the 4 HSP-47 forms were compared, the detected decrease in RDEB skin by MS/MS was also validated in all the technical replicates and in the pool of replicates.



**Figure 46. Expression of HSP-47 at transcript and protein levels in the back skin of WT and RDEB mice aged 9-12 weeks.**

Previous page. **A.** Individual gene expression. Gene expression was evaluated by RT-qPCR on the mRNA transcripts from whole back skin ( $n = 4$  mice/genotype, 3 technical replicates). The  $\Delta\Delta C_t$  method was used to compute gene expression in WT and RDEB samples. **B.** Comparison of gene expression between WT and RDEB genotypes. Statistics: Mann-Whitney test. **C.** Immuno-detection of HSP-47 in skin lysates (reducing conditions, 10875-1-AP antibody, theoretical weight: 47 kDa,  $n = 4$  mice/genotype). Three technical replicates were analyzed, only one representative membrane (or technical replicate) is shown. Bands 1 and 2: glycosylated and glycosylation-free HSP-47 trimers. Band 3: glycosylated HSP-47 monomer. Fibro. lysate: lysate of human dermal fibroblasts. **D.** Quantification HSP-47 in skin lysate. The mean signal per genotype ( $\pm$  SD) was analyzed for each of the 3 technical replicates and for the pool of technical replicates (Double normalization). Statistics: Student t-test or Mann-Whitney test when t-test was not applicable. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ . Ef: efficiency of primer pairs. WB: Western blot. Rep: technical replicate. Mean: mean of the mean values obtained from the 3 technical replicates for each individual mouse.

The reduction in the protein abundance of HSP-47 in RDEB evidenced in proteomic studies thus seems to be real. However, the presence of several oligomeric forms was unexpected and their definitive assignment to HSP-47 will require further validation.

#### 4. Proof-of-concept: development of a PRM assay to monitor procollagen-derived biomarkers in skin

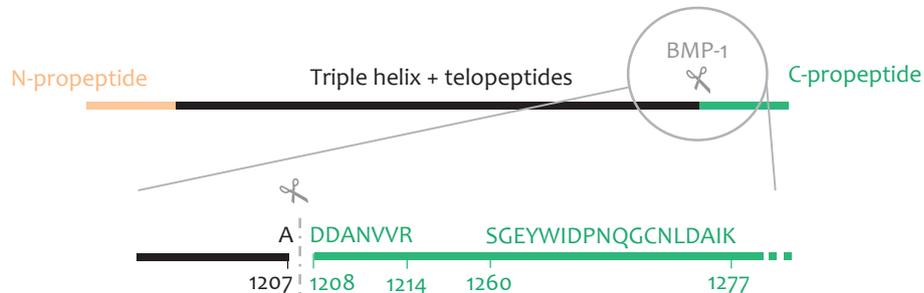
In the previous parts, we evidenced strong decreases in the abundance of BMP-1 and PCPE-1 in the early stage of RDEB disease in mice (back skin of animals aged 9-12 weeks). Conversely, in human, preliminary data suggested that the protease tended to be more abundant in the skin of RDEB patients. Whatever the variation of the protease, the release of CP  $\alpha 1(I)$  by the enzyme is likely to be affected in disease. In clinical practice, the possibility to monitor this cleavage could help to evaluate the effect of therapies on disease progression: the restoration of a normal CP  $\alpha 1(I)$  cleavage rate could indeed reflect the improvement in collagen-related processes and positive responses to therapeutic agents. In this context, we started to develop a targeted assay to monitor the abundance of full-length procollagen I and the extent of its C-terminal processing in complex biological samples. The parallel reaction monitoring technology (PRM, **Chapter I, IV-2-iii**) was selected for this assay, because it was previously reported to allow the targeted analysis of low abundance protein-derived molecules in biological fluids, especially to monitor biomarker candidates in cerebrospinal liquid (Alzheimer's disease<sup>247</sup>, multiple sclerosis<sup>248</sup>), uterine aspirates (endometrial cancers<sup>249,250</sup>) and plasma (fibrosis<sup>251,252</sup>).

In this last part of this chapter, the preliminary results obtained for the detection of procollagen  $\alpha 1(I)$  and CP  $\alpha 1(I)$  derived-peptides by PRM are presented.

##### i. PRM analysis of unlabeled samples

The release of CP  $\alpha 1(I)$  by BMP-1 involves a cleavage of the mouse procollagen  $\alpha 1(I)$  chain between <sup>1207</sup>A and <sup>1208</sup>D<sup>253</sup>. When full-length procollagen I is digested by trypsin for MS/MS analysis, the tryptic peptide carrying the BMP-1 cleavage site (also called “overlapping peptide”) is ADDANVVR. After BMP-1 cleavage and trypsin digestion, the N-terminal (semi-tryptic) peptide

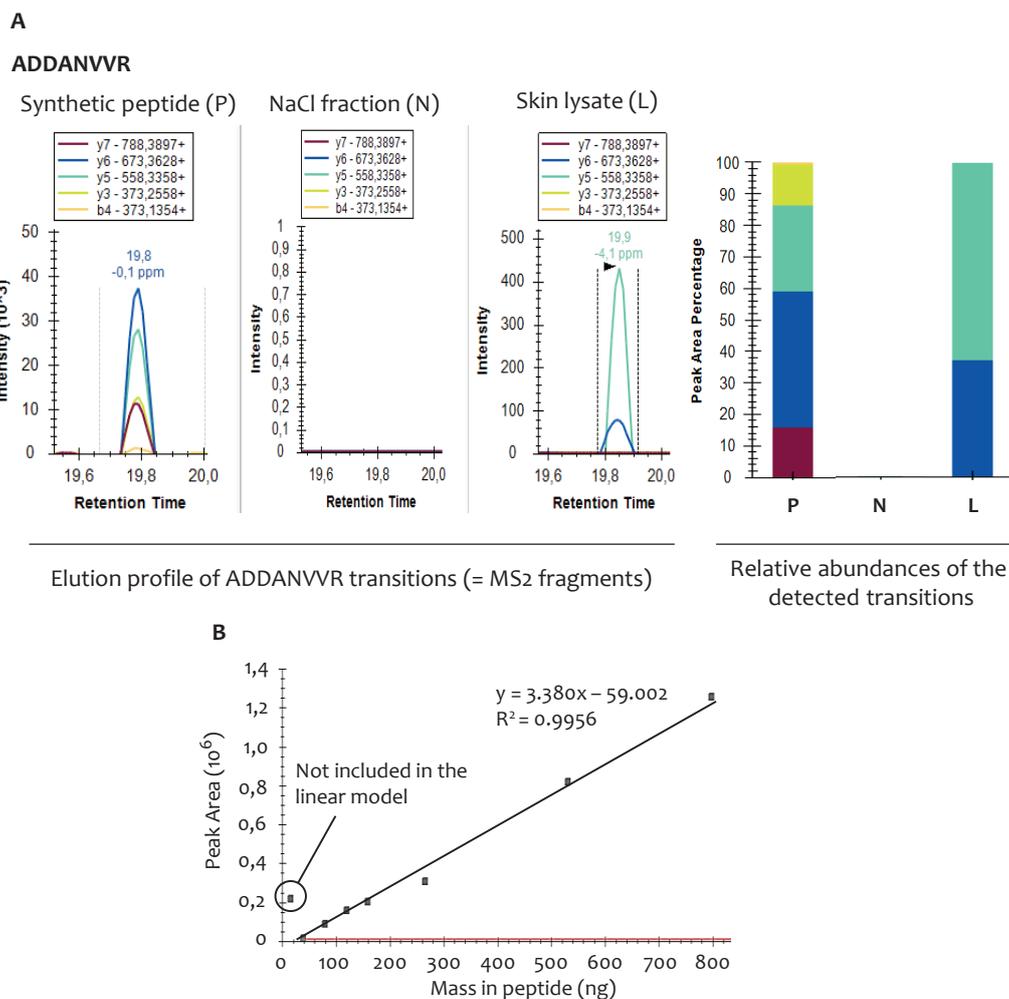
defining the beginning of CP  $\alpha 1(I)$  is DDANVVR which only differs by one residue (**Figure 47**). In this case, it is not possible to follow the C-terminal peptide of the second proteolytic product (pN-collagen  $\alpha 1(I)$ ) by MS/MS, because it is only defined by one alanine, unless another enzyme than trypsin is used for digestion. The major aim of our approach was thus to develop a PRM assay to measure the absolute abundances of DDANVVR and ADDANVVR peptides in WT and RDEB samples. This would allow us to discriminate between the intact and cleaved procollagens, to directly monitor BMP-1 activity and to evaluate the relevance of these peptides as potential biomarkers of the disease. For example, an increase in the DDANVVR/ADDANVVR ratio would reflect increased CP  $\alpha 1(I)$  release by BMP-1. Conversely, a decrease in this ratio would reveal a reduction in the C-terminal processing of procollagen  $\alpha 1(I)$ . Finally, the ADDANVVR peptide alone can also reflect C-propeptide retention in collagen fibrils or the modified abundance of full-length procollagen I.



**Figure 47. Schematic representation of the procollagen  $\alpha 1(I)$  cleavage products generated by BMP-1.**

The indicated sequences in amino-acids correspond to the peptides produced by trypsin-dependent digestion during sample preparation for MS/MS. Before BMP-1 cleavage, the tryptic peptide overlapping the cleavage site is ADDANVVR. When the C-propeptide is cleaved, ADDANVVR disappears and DDANVVR is produced concomitantly. SGEYWIDPNQGCNLDAIK is an internal tryptic peptide that is located further in the C-propeptide sequence and not affected by BMP-1.

The first step in the development of the assay was to analyze ADDANVVR and DDANVVR in the form of “light” synthetic peptides (with natural isotopes) in order to determine their elution profile and their fragmentation profile (list of transitions and relative abundances). NaCl fractions from the chemical decellularization protocol as well as skin lysates were analyzed next. More than 10 transitions were detected for the synthetic ADDANVVR peptide and its intensity exhibited a linear evolution when the injected mass was increased from 5 to 820 ng (**Figure 48A**, **Figure 48B**). However, only 2 transitions were identified in skin lysates (**Figure 48A**) and none in NaCl fractions, suggesting that the sensitivity was not sufficient for the efficient detection of this peptide. Very disappointingly also, the DDANVVR peptide was never detected, even when high amounts of the synthetic peptide were injected and in spite of the different optimizations in sample preparation which were tested (e.g. protein purification methods, sample fractionation based on molecular weight) (data not shown).



**Figure 48. Detection of the ADDANVVR peptide by PRM.**

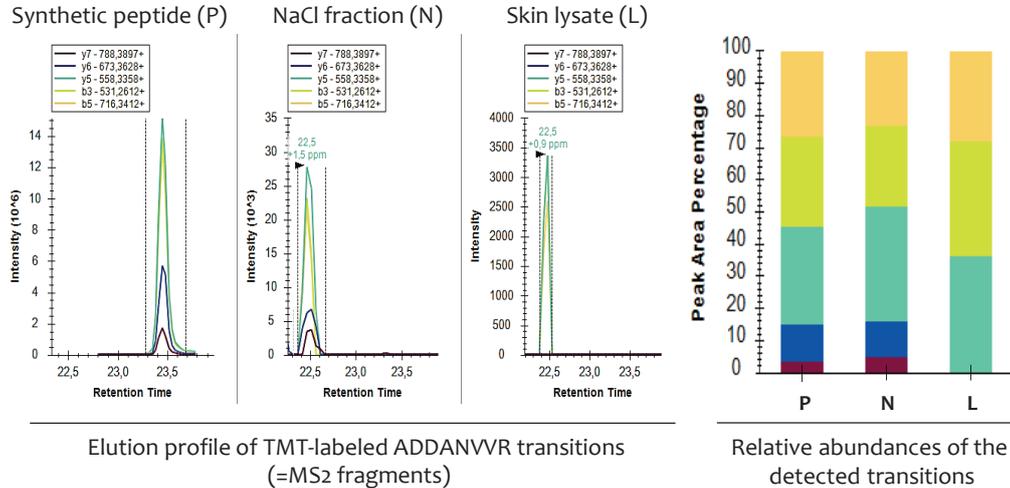
**A.** Profiles of the PRM signals obtained with the synthetic peptide, a NaCl fraction or a skin lysate. For each sample type, the elution profile and the relative abundance of the different transitions are illustrated. **B.** Dose/signal curve for 5-820 ng of synthetic peptide. The plotted peak area corresponds to the cumulative area under the curve of the different transitions detected for the peptide.

## ii. PRM-analysis of TMT-labeled samples

We next evaluated the effect of peptide TMT labeling as an artificial way to improve peptide detection. The target peptides had actually been previously detected in the lab after labeling even if the analyzed samples were less complex (fibroblast secretomes analyzed by quantitative shotgun proteomics). Labeled ADDANVVR was successfully identified in the 3 analyzed samples (**Figure 49**). 5 transitions were identified in the NaCl fraction and in 3 the skin lysate (1 more compared to unlabeled ADDANVVR). The shift in retention time (< 1 min) observed between the synthetic and natural forms observed can be explained by the fact that samples were not analyzed at the same time: the synthetic peptide was analyzed first and the native peptides at a later time point. Even if it will better in the future to inject the synthetic peptide just prior to the samples of interest, the profile of detection observed for the ADDANVVR peptide appeared compatible with a quantitative assay (enough transitions, transitions aligned to form a single peak, high signal intensity). It was still not the case for the DDANVVR peptide

which, after TMT labeling, could be identified when injected as a synthetic peptide but not in biological samples (1 to 2 transitions detected, no transition alignment).

#### ADDANVVR, TMT-labeled

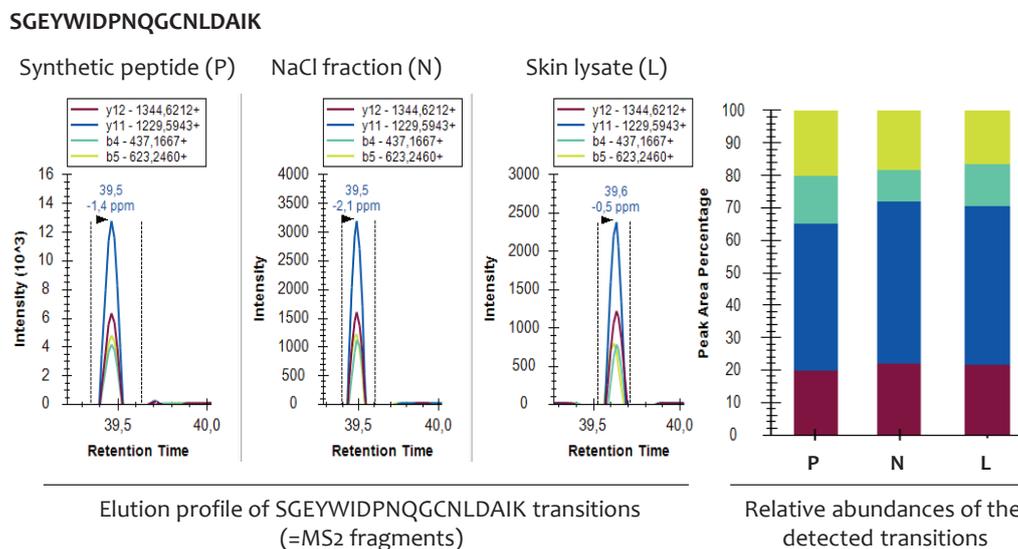


**Figure 49. Detection of TMT-labeled ADDANVVR and DDANVVR peptides by PRM.**

Profiles of the PRM signals obtained with the synthetic peptide, a NaCl fraction or a skin lysate are illustrated for each peptide.

As DDANVVR was not detected in spite of the various tested optimizations, a new approach was needed. We decided to replace the DDANVVR peptide by a peptide located within the C-propeptide  $\alpha 1(I)$  sequence, further away from the cleavage site. As this peptide is found both in collagen precursors and in released C-propeptide, it can be used to perform a kind of “normalization” of the abundance of ADDANVVR (reflecting uncleaved precursors) by the global abundance in CP  $\alpha 1(I)$ -containing proteins. Even if it is less BTP-specific, we expected that this approach would allow us to take into account the strong decrease in procollagen  $\alpha 1(I)$  observed in RDEB samples. To do this, we selected the SGEYWIDPNQGCNLDAIK peptide (Figure 47) which was efficiently detected in the synthetic form previously in the lab. Very good-quality signals (4 perfectly-aligned transitions, relative abundances

identical for the synthetic and natural peptide forms) were detected for this peptide in both the NaCl fractions and skin lysates (**Figure 50**). The nice detection profile obtained for the non-labeled form of the peptide suggests that the detection of the TMT-labeled form may also be possible.



**Figure 50. Detection of the SGEYWIDPNQGCNLDAIK peptide by PRM.**

Profiles of the PRM signals obtained with the synthetic peptide, a NaCl fraction or a skin lysate are illustrated for each peptide.

In conclusion, these preliminary results are very promising: it seems that the analysis of both uncleaved procollagen  $\alpha 1(I)$  and its C-terminal processing product may be possible after TMT labeling of the peptides. A similar approach could also be developed for other fibrillar procollagens. For now, the validation of the good detection of the TMT-labeled SGEYWIDPNQGCNLDAIK peptide and the determination of the useful concentration range for this peptide and for the ADDANVVR peptide are in progress. Regarding DDANVVR, very recently, MS-based experimentations conducted in my group suggested the existence of second cleavage site by BMP-1, a few amino acids after the one we monitored here. This new site could conduct to the loss of DDANVVR and therefore explain why it was never detected in biological samples. The PRM-monitoring of the peptide generated by the BMP-1 cleavage at the newly identified site will be tested soon.

The ability to detect all these peptides in very small skin samples could be of interest in clinical applications, as only limited amounts of biological material can be recovered from RDEB patients for analysis. Notably, PRM assays also offer high multiplexing capacity (possibility to monitor tens of transitions simultaneously) and it will be possible in the future to include additional biomarkers, either collagen-related or collagen-unrelated.

In conclusion, procollagen  $\alpha 1(I)$  and procollagen  $\alpha 1(I)$ -derived components appear likely to offer promising biomarker candidate for the monitoring of RDEB in clinic.

## 5. Discussion

In the last part of this project, we aimed at explaining the collagen phenotype previously evidenced in **Chapter IV, part II**. At the molecular level, decreases in the abundance of collagen I chains were detected in 3 of the 4 enrichment fractions (NaCl, SDS/SD, Comp. Ins.) and similar trends in the abundance of collagen III and V chains were also observed. At the tissue level, a global disorganization of the matrix was observed and a strong overall decrease in collagen fibril diameters was evidenced. This last alteration was found very early in disease, in 7-day-old mice, and appeared to worsen with disease progression as a stronger defect in fibril diameters was observed in the skin of fibrotic paws.

Different hypotheses were made at the beginning of this chapter to explain the RDEB-induced collagen phenotype, especially the decrease in fibril diameters. Among these hypotheses, we chose to investigate in priority possible defects in the composition of collagen fibrils and alterations in the collagen maturation processes which could potentially lead to propeptide retention in fibrils. A first major result was the demonstration that the proportion of collagens III and V was increased relative to collagen I in the insoluble ECM. This seemed to be mainly due to a strong diminution in the abundance of the collagen  $\alpha 1(I)$  chain, equally affecting the expression of the *Coll1a1* gene, the secreted procollagen  $\alpha 1(I)$  precursor and the mature collagen I integrated in the insoluble matrix. In addition to these impairments in protein production, our results also underlined diminutions in the global protein abundance of active BTPs, the main proteases responsible for the C-terminal processing of fibrillar procollagens, and of their PCPE-1 enhancer. Finally, we also obtained some preliminary indications that the collagen chaperone HSP-47 could be affected. The main results obtained in this part are summarized in **Table 16**.

Mechanism explaining the collagen phenotype	Conducted evaluation	Analyzed parameter	Method	Variation in RDEB mice	Possible link with phenotype
(1) Modifications in collagens I, III and/or V themselves	Composition of the insoluble matrix in collagens I, III and V	- Abundance of $\alpha 1(I)$ and $\alpha 2(I)$ chains	MS/MS (Comp. Ins., Fig. 25) WB (skin lysate, Fig. 34)	<b>Decreased</b> No defect (partial information)	<b>Yes</b>
		- Ratio $\alpha 1(I)/\alpha 2(I)$	MS/MS (insoluble fractions, Fig. 34) WB (skin lysate, Fig. 34)	No defect (partial information)	<b>Yes</b>
		- Ratios $\alpha 1(I)/\alpha 1(III)$ , $\alpha 1(V)$ and $\alpha 2(V)$	MS/MS (Comp. Ins., Fig. 35)	<b>Decreased</b>	<b>Yes</b>
		- Ratios $\alpha 2(I)/\alpha 1(III)$ , $\alpha 1(V)$ and $\alpha 2(V)$	MS/MS (Comp. Ins., Fig. 35)	<b>Decreased</b>	<b>Yes</b>
		- Ratio $\alpha 1(V)/\alpha 2(V)$	MS/MS (insoluble fractions, Fig. 35)	No defect	
	Procollagen $\alpha 1(I)$ production	- Expression of <i>Col1a1</i>	RT-qPCR (skin lysate, Fig. 36) MS/MS (NaCl, Fig. 25)	<b>Decreased</b> <b>Decreased</b>	<b>Yes</b> <b>Yes</b>
		- Abundance in procollagen $\alpha 1(I)$	WB (NaCl, Fig. 38) WB (Comp. Ins., Fig. 39) WB (skin lysate, Fig. 39)	<b>Decreased</b> <b>Decreased</b> <b>Decreased</b>	<b>Yes</b> <b>Yes</b> <b>Yes</b>
		- Ratio procollagen $\alpha 1(I)/$ procollagen $\alpha 2(I)$	MS/MS (NaCl, Fig. 37)	<b>Decreased</b>	<b>Yes</b>
		- Expression of <i>Col1a2</i>	RT-qPCR (skin lysate, Fig. 36)	No defect	
		- Expression of <i>Col3a1</i>	RT-qPCR (skin lysate, Fig. 36)	No defect	
(2) Retention of N- or/and C- propeptides	HSP-47 chaperone: - folding/secretion of procollagens 47	- Expression of <i>Serpinh1</i> (encoding HSP-47)	RT-qPCR (skin lysate, Fig. 46)	No defect	
		- Abundance in HSP-47	MS/MS (all fractions, Fig. 46) WB (skin lysate, Fig. 46)	<b>Decreased</b> <b>Decreased</b>	<b>Yes</b> <b>Yes</b>
	C-terminal processing of procollagen $\alpha 1(I)$	- Relative abundance in CP $\alpha 1(I)$	WB (skin lysate, Fig. 39)	No defect	
		- Relative abundance in pC-collagen $\alpha 1(I)$	WB (NaCl, Comp. Ins., Fig. 38, 39) WB (skin lysate, Fig. 39)	Inconclusive <b>Decreased</b>	<b>?</b>
	N- anc C-terminal processing of other fibrillar collagens	- Peptide mapping	MS/MS (all fractions, Fig. 40)	Inconclusive	
		- Peptide mapping	MS/MS (insoluble fractions, Fig. 41)	<b>Decreased N-pro <math>\alpha 1(V)</math></b>	<b>?</b>
	C-proteinase expression/activity	- Expression of <i>Bmp1</i> and <i>Pcolce</i>	RT-qPCR (skin lysate, Fig. 42)	<b>Decreased</b>	<b>?</b>
		- Expression of <i>Tll1</i>	RT-qPCR (skin lysate, Fig. 42)	No defect	
		- Abundance in active BMP-1, mTLD, mTLL1	Immunofluorescence (Fig. 43)	<b>Decreased</b>	<b>?</b>
		- Abundance in PCPE-1	WB (skin lysate, Fig. 44) Immunofluorescence (Fig. 45)	<b>Decreased</b> <b>Decreased</b>	<b>?</b> <b>?</b>

Table 16. Summary of the conducted experiments and their results in relation with the collagen phenotype in RDEB skin.

WB = Western blot, Fig. = figure.

The alteration in the proportion of the major dermal fibrillar collagens (I, III and V), especially the relative increase in collagen III and V compared to collagen I, was not really expected because all these collagens seem to be decreased to some extent on the basis of our proteomic results. However, our analyses revealed that the reduction in collagen I was clearly stronger than the reduction in other fibrillar collagens. In the literature, it is well-described that the modification in the proportion of collagen chains can disturb the insoluble collagen network. In the meniscus matrix of *Col3a1*<sup>+/-</sup> mice, the reduction in collagen III induced fibril thickening<sup>226</sup>. The same observations were made in the cornea of *Col5a1* KO mice with, in addition, a decreased inter-fibril spacing<sup>215</sup> and macroscopic defects (slimming and increased transparency of the cornea). In human also, skin aging is characterized by a decrease in collagen fiber thickness<sup>254</sup> concomitant with the augmented proportion of collagen III<sup>255</sup>. All these data suggest that the increases in the proportion of collagen III and/or V in RDEB fibers is likely to contribute to the observed decrease in fibril diameters. Additional analyses of these two collagens in the insoluble matrix of RDEB mouse skin would be interesting to conduct to validate this hypothesis. This modified I/III/V ratio could also impact skin mechanical properties and this point is currently investigated by atomic force microscopy by the group of A. Nyström.

If the reduction in protein abundance appeared similar for the two chains of collagen I in the insoluble matrix of RDEB mouse skin ( $\alpha 1(I)/\alpha 2(I)$  ratio not modified), we evidenced by mass spectrometry that collagen  $\alpha 1(I)$  was more strongly affected than collagen  $\alpha 2(I)$  at the procollagen level. The reduction in the production of procollagen  $\alpha 1(I)$  is likely to limit the formation of “normal” [ $\alpha 1(I)_2, \alpha 2(I)$ ] heterotrimers in cells, resulting in a decreased secretion of procollagen heterotrimers in the matrix. This alteration may be an explanation for the drastic reduction of procollagen I especially observed in RDEB NaCl fractions. Besides, the absence of significant decrease in collagen  $\alpha 2(I)$  in RDEB NaCl fractions (quantitative MS-based proteomics, **Chapter III, II-2-ii**) can have additional consequences. As no [ $\alpha 2(I)_3$ ] homotrimer or [ $\alpha 1(I), \alpha 2(I)_2$ ] heterotrimer can be formed (due to the “cysteine code” rule<sup>256</sup>), it is possible that free  $\alpha 2(I)$  monomers are secreted in the extracellular environment. If this phenomenon has previously been observed in our lab in cultured 293 EBNA cells, possible effects of  $\alpha 2(I)$  monomers on ECM organization are presently unknown. The specific analysis of the  $\alpha 2(I)$  procollagen chain in the NaCl fraction could thus bring additional information to better understand the dysregulation in the secretion of procollagen I chains in RDEB.

Even more unexpectedly, we also evidenced the presence of CP  $\alpha 1(I)$ -containing proteins in insoluble collagen fibers, both by mass spectrometry and Western blot. This seems to be a rare event as the number of CP  $\alpha 1(I)$  peptides detected by MS/MS is much lower than the number of peptides detected in the triple helix but the presence of the three CP  $\alpha 1(I)$ -containing forms (procollagen, pC-collagen and CP  $\alpha 1(I)$ ) in the insoluble matrix was also clearly evidenced by Western blot, at least in WT animals.

Even if this result goes against the dogma that fibrils cannot form if the C-propeptide is still present on collagen molecules, it is not the first time that C-propeptide retention is reported in collagen fibrils<sup>236,237</sup>. In the *Bmp1/Tll1* KO mice, where C-propeptide retention is expected to be maximum (and could be evidenced by immunostaining), collagen fibrils with altered morphology (“barbed-wire” appearance) and decreased diameters were observed<sup>230,257,258</sup>. Similar observations regarding fibril diameters were made for the retention of the N-propeptides of collagens I<sup>229</sup>, III<sup>227</sup> and V<sup>259</sup> which is somewhat more common. For example, previous literature reports have shown that the amount of N-propeptides from the  $\alpha 1(I)$  and  $\alpha 1(III)$  chains were specifically elevated in the smallest fibrils of both normal<sup>238</sup> and scleroderma<sup>227</sup> skins. In our case, it is difficult to make a direct link between our MS/MS observations which indicate a global decrease in propeptide retention (both N- and C-propeptides when detected) and the reduction in collagen fibril diameters. Indeed, it would have been more logical to find increased propeptide retention to explain the decrease diameters and it seems that one piece of the puzzle might still be missing.

Additional information can be brought by the analysis of the expression and activity of the enzymes and regulatory proteins involved in procollagen N- and C- terminal processing. This is indeed what we have started to do with the analysis of BTPs and PCPE-1 which control the C-terminal maturation of procollagens I, III and V and the N-terminal maturation of procollagen V. It would also be very interesting to analyze ADAMTS-2, -3 and -14 which are the main procollagen N-proteinases for collagens I and III and could also contribute to their C-terminal processing<sup>86,87</sup>. In addition, ADAMTS-2 can cleave the N-terminal domain of the  $\alpha 1(V)$  chain, at a site which is downstream (P↓<sup>435</sup>A in mouse<sup>260,261</sup>) of the site used by BTPs (S↓<sup>254</sup>Q<sup>262</sup>). Only one peptide between the cleavage sites of BTPs and ADAMTSs was detected by MS/MS and this was not enough to conclude if both sites were used or if there was any differential activity between the two enzyme families.

In the course of this study, we could however obtain interesting and conclusive results about the expression of BTPs and PCPE-1. The most active BTPs found in skin (BMP-1/mTLD) were clearly down-regulated and this was associated with a decrease in the abundance of the active proteases, as evidenced by immuno-staining with an antibody specifically directed against the active forms. Similar results were obtained for their procollagen-specific enhancer, PCPE-1, with a very nice convergence of the results obtained by PCR, Western blot and immuno-fluorescence. It was therefore a bit disappointing to see that the link with incomplete procollagen processing could not be made easily. By Western blot, the amount of CP  $\alpha 1(I)$ -containing forms was in general too low in RDEB samples to allow proper quantification and peptide mapping experiments gave results which were in apparent contradiction with a defective BTP-dependent procollagen processing (expected to lead to more propeptide retention). One way to explain peptide mapping results could be that, even if the amounts of BTPs and PCPE-1 are globally decreased, these proteins are still in large excess compared to procollagen I which seems to be the most strongly affected. However, at this point, the link between the decreased level of procollagen

proteinases, the decreased propeptide retention and the collagen phenotype remains elusive and will require further investigations.

Nevertheless, we can already predict that the decrease in active BMP-1/mTLD is likely to affect other processes, directly linked to collagen fibrillogenesis or not, in RDEB skin. For instance, the protease is involved in the maturation of prodecorin and probiglycan. In embryonic fibroblasts from *Bmp1* KO mice, a decrease in the processing of probiglycan was reported<sup>263</sup>, and in the skin of *Bmp1/Tll1* KO mice, the integration of immature biglycan into collagen fibrils was evidenced<sup>258</sup>. Furthermore, in human, a genetic deficiency in the *BMP1* gene leading to osteogenesis imperfecta was also associated to a diminution in the processing of prodecorin secreted by dermal fibroblasts<sup>240</sup>. In RDEB mice, the increase in the soluble forms of decorin and biglycan in the SDS/SD fraction (most probably corresponding to the pro-proteins) combined with the decrease in the insoluble forms seems consistent with the lowering in BMP-1/mTLD as the reduction in the processing of these two SLRPs is likely to result in the accumulation of soluble precursors. The consequences of the elevation in the amounts of probiglycan and prodecorin have not been investigated yet but effects on the binding of bioactive molecules such as TGF- $\beta$  could be possible<sup>62</sup>. Alterations in the activity of BMP-1/mTLD is also likely to negatively affect the activation of pro-lysyl oxidases<sup>76</sup> and then to inhibit the formation of cross-links between collagen trimers in fibrils. Interestingly, a reduction in the abundance of hydroxyprolines, a major PTM in collagens also involved in cross-link formation, was previously reported in the collagens produced by cultured human RDEB fibroblasts (in 3 out of 4 analyzed diseased samples)<sup>146</sup>. Whatever its origin, a reduction in cross-links could limit the stability of fibers and contribute to TEM observations suggesting a less cohesive collagen network. Furthermore, a reduction in cross-linked forms would result in an increase in free mature collagen I chains and could explain why the lowering in mature collagen I, evidenced in the compartmental insoluble fraction (Comp. Ins., containing very insoluble cross-linked collagens), was not validated by Western blot (only free mature  $\alpha 1(I)$  and  $\alpha 2(I)$  chains quantified) nor in the “GuHCl + Ins.” fraction of the chemical decellularization protocol (containing both “insoluble” collagens brought by the GuHCl extract and “very insoluble” collagens from the insoluble pellet).

Finally, our data also suggest that intracellular mechanisms involved in the folding and stabilization of fibrillar procollagen trimers can contribute to the alteration in collagen I deposition observed in the matrix. HSP-47 is a well-known chaperone for collagens (including fibrillar collagens, FACITs)<sup>207,228,264</sup> for which we evidenced a possible deficiency in RDEB skin. Interestingly, a significant decrease in its abundance was also observed recently by our German collaborators who analyzed the whole back skin (without ECM enrichment) of collagen VII-hypomorphic mice by MS/MS<sup>212</sup>. However, in the same study, the variation was not observed in paws, suggesting that its expression could depend on disease stage or anatomic location. Also intriguing in our case was the detection of HSP-47 in ECM-enriched fractions while the protein is supposedly confined in the endoplasmic reticulum<sup>265</sup>. However, to my

knowledge, none of the studies focused on this protein really evaluated its presence in the ECM. If its detection in the NaCl fraction could result from proteins released by dead cells, the presence of HSP-47 in the insoluble matrix is more difficult to explain and will have to be validated by other approaches. Nevertheless, at this point, HSP-47 remains one of the best candidates to explain the collagen phenotype in RDEB skin. The knock-out of the *Serpinh1* gene (encoding HSP-47) in NIH-3T3 cells<sup>228</sup> and mouse embryonic fibroblasts<sup>229</sup> was shown to reduce the amount of procollagen I secreted in the extracellular environment, due to collagen aggregation and retention in the endoplasmic reticulum. This could link HSP-47 deficiency to the dramatic reduction in extracellular procollagen I observed in RDEB skin. However, we were not able to evidence that ER stress could be a major pathway, as it would be expected if collagen was retained in the organelle, neither in our proteomic data (data not shown) nor in the published results obtained with skin lysates<sup>212</sup>. Defective procollagen secretion remains to be demonstrated too. The analysis of the Triton X-100 fraction generated by chemical decellularization and containing most of the cellular proteins may be helpful to go deeper in the characterization of HSP-47, and to determine if procollagen  $\alpha 1(I)$  is actually retained in cells in our model. Pulse-chase experiments on fibroblasts cultures could also be run to compare the secretion of procollagen in control and RDEB conditions. Also interesting is the fact that the release of the N-propeptide from procollagen  $\alpha 1(I)$  was reported to be delayed in the matrix deposited by *Serpinh1* KO mouse embryonic fibroblasts, thereby disturbing the formation of collagen fibrils which display abnormal morphology and smaller diameters<sup>229</sup>. This observation consequently strengthens the idea that defects in the N-terminal processing of fibrillar procollagens may play a role in RDEB skin and underlines, again, the necessity to evaluate the N-terminal maturation events and procollagen N-proteinases more thoroughly in disease. In brief, the reduction in the abundance of HSP-47 appears likely to disturb various steps in the biosynthetic pathways of fibrillar collagen, including procollagen folding in the ER, procollagen trimer secretion and their proteolytic maturation. These impairments could be exacerbated by the alterations in the conformation of HSP-47 observed in RDEB skin (decreased abundance of trimeric forms), if they are confirmed.

Finally, it is interesting to link previous results with the alteration in the abundance of reticulocalbin-3, another ER chaperone encoded by the *Rcn3* gene. It was very strongly decreased at the protein level in the three enrichment fractions where it was detected (NaCl, GuHCl + Ins. and Comp. Ins., **Appendix 9**), and in whole skin lysates analyzed by MS/MS by our German collaborators<sup>212</sup>. Very recently, *Rcn3* KO mice were reported to display over-hydroxylation of prolines and lysines in collagen  $\alpha 1(I)$  (due to delayed secretion) as well as decreased collagen fibril diameters in their tendons<sup>209</sup>. Taken together, defects in the collagen chaperone pathway appear as a very promising mechanism to explain our macroscopic observations and our preliminary results suggest that intracellular processes must not be neglected to explain alterations of the RDEB extracellular matrix.

To sum up, the preliminary investigations conducted in this part of my PhD project indicate that the collagen phenotype observed in RDEB mouse skin probably derives from multiple connected processes. We can already conclude that the main driver of the final disorganization in the fibrillar collagen network is the strong decrease in the production of procollagen I and the consequent increase in the proportion of collagens III and V in fibers, which, in the end, is likely to affect fibril diameters. However, additional alterations in other collagen-related processes, especially the decrease in HSP-47, procollagen C-proteinases and PCPE-1, could also contribute to the collagen phenotype. The defects in N- and C-terminal maturation processes possibly resulting from these alterations may also play a direct role even though the precise mechanism remains incompletely understood. Finally, if the defects in procollagen  $\alpha 1(I)$  production and maturation affect the release of collagen-derived molecules in RDEB as we suspect, the latter molecules could exhibit a biomarker potential for disease monitoring in clinics.

# GENERAL DISCUSSION AND CONCLUSION

---

Dystrophic epidermolysis bullosa is a progressive and debilitating disease consequent to mutations in the *COL7A1* gene. This gene encodes collagen VII, a protein which forms anchoring fibrils and stabilizes basement membranes. In the most severe forms of the disease, especially the recessive ones (RDEB), patients suffer from skin fragilities and progressively develop fibrosis and aggressive cancers. Currently, research efforts are mostly concentrated on the correction of the genetic deficiency (curative therapies) but new therapeutic options are also needed to limit the occurrence and progression of the most disabling symptoms (especially epidermal blistering, slow-healing wounds and dermal fibrosis) and improve the quality of life of patients. In skin, collagen VII is both key to maintain a correct matrix organization in the dermo-epidermal junction and to ensure the connection with the fibrillar collagen network in the dermis. We consequently hypothesized that major impairments in the skin ECM composition and organization would occur when collagen VII is present in abnormally low amounts in DEB. In this context, the main objective of my project was to characterize the alterations of the ECM in RDEB skin and to provide new directions for the development of symptom-relief therapies, and possibly identify new biomarkers to monitor disease evolution and/or assist clinical trials.

The early disease-associated dysregulations of the ECM were investigated by MS-based quantitative proteomics in a mouse model of RDEB, using matrisome recovery protocols, also developed in this project, and by the detailed analysis of skin ultrastructure characterized by transmission electron microscopy. Taken together, the results of these experiments first evidenced impairments in the DEJ. The abundance of basement membrane proteins was slightly but globally increased, and the length and proportion of hemidesmosomes in the DEJ were reduced. If DEJ destabilization was suspected *a priori*, based on previous literature reports and because collagen VII is involved in the structuration of basement membranes<sup>266</sup>, the massive defects in the amount and organization of fibrillar collagens (decreased fibrils diameter, less cohesive collagen network, increased proportions of mature collagen III and V), and in the abundance of collagen organizers, was not anticipated to that extent. The decrease in collagen I chains was particularly intriguing, as none of the previous works on the RDEB mouse model nor on RDEB patients ever reported such alterations<sup>156,267,268</sup>. Also surprising was the apparent strengthening of alterations affecting the collagen network at the fibrotic stage, because fibrosis is commonly associated with an increase in the abundance of fibrillar collagens<sup>55,220</sup> that could be expected to positively affect fibril diameters or at least collagen density, and we observed the inverse situation. Notably, a recent characterization of RDEB fibrotic paw skin suggested that the organization rather than the abundance of fibrillar collagens was affected<sup>212</sup>. Taken together, all these results indicate that the

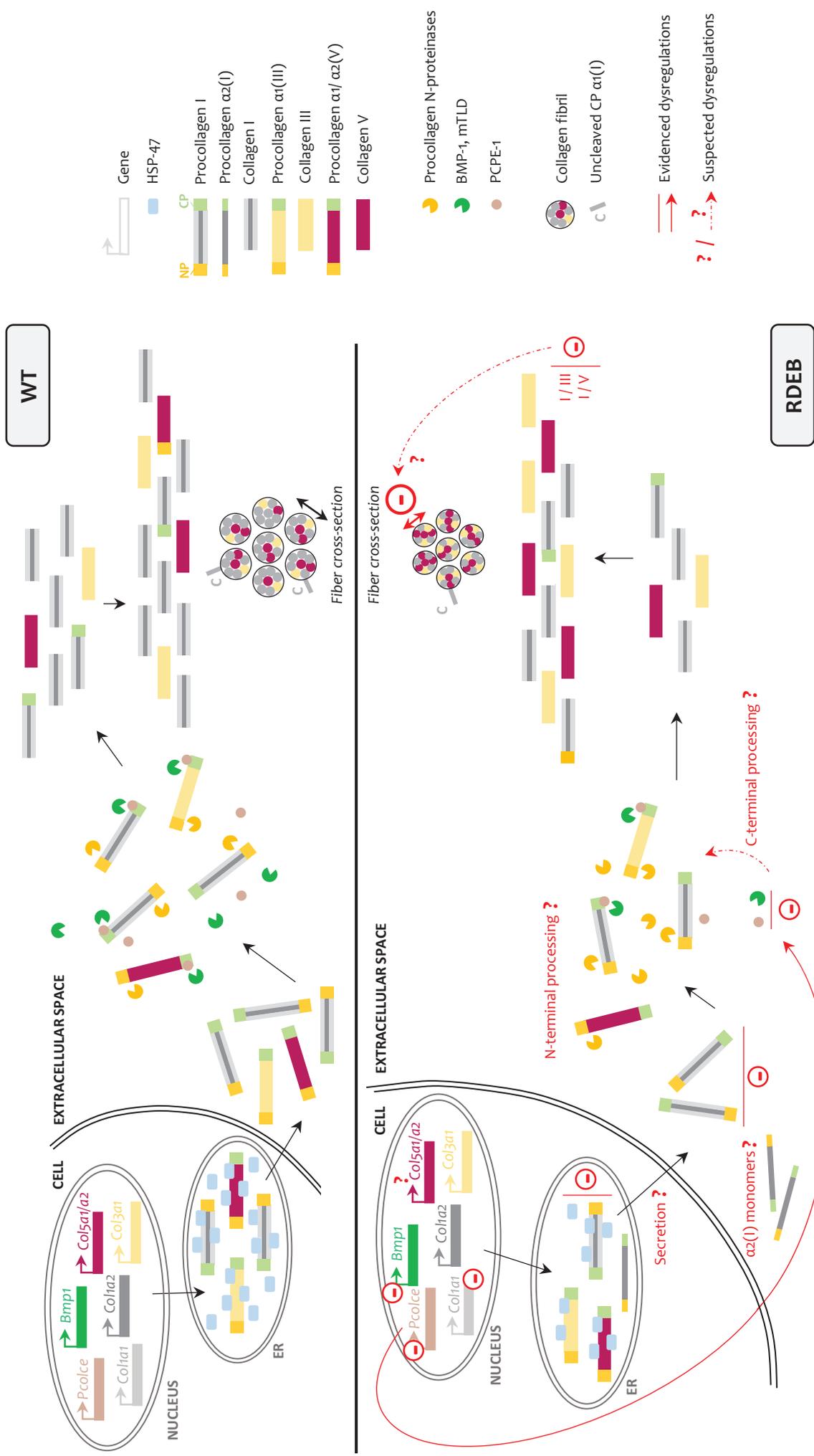
fibrosis developed in RDEB differs from the classical description of fibrosis and could be named “DEBrosis” to avoid confusion. Interestingly, in scleroderma, another fibrotic disease affecting the skin, variations in fibril diameters similar to those observed in RDEB mice were reported<sup>224</sup>, strengthening the idea that the skin-specific forms of fibrosis can differ from the commonly described ones.

When the production of fibrillar collagens was investigated in more detail, procollagen  $\alpha 1(I)$  was the most severely affected. A massive decrease in the expression of *Colla1* and in the abundance of immature collagen  $\alpha 1(I)$  forms was detected, the latter alteration being much stronger in the soluble compartment of the ECM than in whole skin samples including cells and ECM. This could be linked to defective secretion as the skin content in the collagen chaperone HSP-47 also seemed affected. In addition, lowerings in the expression and abundance of active procollagen C-proteinases and their collagen-specific regulatory protein (PCPE-1) were evidenced in RDEB skin. The evaluation of procollagen C-terminal processing was unfortunately impaired by technical limitations and we were not able to make any definitive conclusion on the potential role of defective C-terminal processing in RDEB. However, a possible dysregulation in the N-terminal processing of fibrillar procollagens was also suggested by our results including (1) the modification in the amount of pC-collagen  $\alpha 1(I)$ , a product of procollagen N-proteinases, at the whole skin level (2) the decreased abundance of HSP-47 which was previously shown to be critical for correct N-terminal processing<sup>229</sup> and (3) the observed decrease in fibril diameters possibly linked to the integration of N-propeptide-containing collagen molecules in fibers<sup>227,229,238</sup>.

All these results underlined the necessity to go deeper in the characterization of procollagen maturation processes in RDEB. To this end, the labeling of CP  $\alpha 1(I)$  by immunogold/TEM on WT and RDEB mouse skin sections was recently performed, in order to evaluate the effect of the decrease in active procollagen C-proteinases on the proportion of immature collagen I forms integrated into fibers (**Appendix 10**). In addition, we are currently optimizing a protocol to run N-terminomics analyses (N-TAILS, **Chapter I, IV-3-ii**) to analyze proteolysis (including N- and C-terminal procollagen maturation processes) in skin lysates and enrichment fractions. The evaluation of the expression of procollagen N-proteinases at transcript and protein levels in skin lysates are also planned. Experiments on RDEB mouse fibroblasts cultured *in vitro* may also be informative as secretomes can indeed be easier to handle than skin lysates for the monitoring of N- and C-terminal processing by Western blot and N-TAILS. To complete the characterization of fibrillar collagens in diseased skin, the characterization of possible alterations in PTMs, especially hydroxyylations of proline and lysine residues, and the study of cross-links for instance<sup>269,270</sup> may also be informative.

In summary, our results evidence multiple dysregulations in the ECM of RDEB mouse skin early in disease (**Figure 51**). If most of the impairments are limited to extracellular processes, preliminary results also suggest defects in the intracellular processes involved in the biosynthesis of fibrillar procollagens

(decreases in HSP-47 and reticulocalbin-3). The deficiency in collagen VII finally appears to affect all the upper skin compartments in RDEB (hemidesmosomes of keratinocytes, basement membrane, dermis) and to disturb both intracellular and extracellular pathways.



**Figure 51. Summary of collagen biosynthesis related-processes dysregulated in RDEB mouse skin**

Note that if alterations in the production of procollagen  $\alpha 1(I)$  and in the collagen I/III/V ratios are relatively well-established, it is still not the case for possible dysregulations in procollagen maturation and consequent propeptide retention which remain incompletely understood.

All the findings of this project now need to be validated in human. If our mouse model develops symptoms very similar to those observed in RDEB patients and has shown its potential to reveal relevant disease-related pathways in various studies<sup>140–142,212</sup>, it remains that it does not result from a RDEB-causing mutation in human. In addition, the symptoms and severity of the disease are mutation-dependent<sup>114,271</sup> but can also be affected by modifier genes<sup>114,115</sup>, while our model proposes a unique form of collagen VII deficiency in mice which are all identical from a genetic point of view. To be transferred in clinical practice, therapeutic target and biomarker candidates have to be applicable to the largest number of patients. The next step in the project would consequently aim at confronting our mouse results with human forms of the disease. Interestingly, preliminary results we recently generated suggested that some of the alterations in the collagen pathway detected in mouse could also be present in human. We indeed recently compared the transcripts of dermal fibroblasts from 6 RDEB and 6 control patients by Nanostring (using a collagen-focused gene panel; data not shown) and evidenced decreases in the expression of *PCOLCE* (encoding PCPE-1) in young RDEB patients, consistently with our observations on the early disease stages in mouse. The change in expression of *PCOLCE* was close to significance in older patients too. The *BMP1-3* (leading to the mTLD protease), *DCN* (decorin) and *BGN* (biglycan) transcripts were also decreased in age-matched analyses (and very close to significance). In addition, in all RDEB patients, a reduction in the expression of collagen-modifying enzymes involved in the elaboration of cross-links in the matrix, especially *LOX* (lysyl oxidase, strong trend), *PLOD3* (lysyl hydroxylase, strong trend), *TGM2* (transglutaminase-2, significant) were evidenced. Even if conducted on a limited number of patients, this preliminary analysis suggests that the collagen-related results obtained in mouse are likely to be transposable to human and to be used to define new therapeutic targets and biomarkers.

If we come back to the initial objective of this project, how could we use the dysregulations evidenced in the biosynthesis of fibrillar collagens as a basis for the development of new therapies?

We evidenced alterations at different levels (gene expression, protein folding in the ER, possibly secretion and protease-driven processes), indicating that different processes could be targeted in therapy. Additional studies are certainly needed to determine the causes of these defects but it is already possible to consider the evaluation of drugs previously reported to modulate collagen deposition, in cell cultures for example. For instance, aldosterone was reported to efficiently increase the expression of *Colla1* in rat renal fibroblasts cultured *in vitro*<sup>272</sup>. The side effect on the production of collagen III, that is already over-represented in the insoluble matrix of RDEB mice, is however to evaluate. To target collagen I only, administration of interleukin-4 was shown to activate the gene expression of *Colla1* and *Colla2* in human dermal<sup>273</sup>, liver<sup>274</sup> and lung fibroblasts<sup>275</sup>, and low doses could increase the production of collagen I chains without modulating the level of collagen III.

A first limitation to these approaches is that they are likely to affect other molecules and processes in skin, and in the whole organism in case of systemic delivery. To overcome the problem, adapting the

route of administration could limit adverse effects but remains challenging. The topical route may actually be more adapted to locally treat both wounded and unwounded skin areas. A second limitation is that these agents may also trigger pro-fibrotic pathways. Even if “DEBrosis” seems to exhibit features differing from the classical forms of fibrosis, such drugs would have to be cautiously evaluated in terms of fibrosis promotion in RDEB skin. Finally, it appears difficult to correct the abundance in BTP enzymes without triggering unwanted side effects, due to the wide diversity of substrates and biological pathways that are regulated by this protease family<sup>76</sup>. Nevertheless, if the deficiency in C-terminal processing of procollagens by BTPs is confirmed, exogenous supplies of PCPE-1 could be a promising way to specifically activate procollagen processing in the matrix and possibly restore the deposition of mature collagens. Cell culture experiments with human fibroblasts are presently conducted to further evaluate this strategy.

If the validation of new ECM-associated therapeutic targets requires additional studies, we have proposed to use procollagen-derived proteins as biomarker candidates and encouraging results were obtained for their detection by parallel reaction monitoring (PRM). Numerous collagen-derived molecules were already proposed as circulating biomarkers in various diseases, especially to monitor the production (N- and C-propeptides) and turnover (degradation products) of fibrillar collagens<sup>220</sup>. In systemic sclerosis for instance, patients were efficiently discriminated from healthy controls through their serum concentration in N-propeptide  $\alpha 1(\text{III})$  and in a collagen I degradation product generated by MMPs<sup>169</sup>. This approach also allowed to diagnose the stage of the disease, as the two biomarkers were differently abundant in early and late stages of the disease. In lung and liver fibrosis, the monitoring of serum N-propeptides  $\alpha 1(\text{I})$  and  $\alpha 1(\text{III})$  was also proposed as an indicator for the diagnosis and monitoring of disease progression<sup>276-278</sup>. Similarly, we hypothesized that the decrease in active procollagen C-proteinases in RDEB skin is likely to modulate the release of the C-propeptide  $\alpha 1(\text{I})$ , and that the fragments generated could be used as biomarkers to evaluate the progression of the disease or the response to therapy. The global amount in procollagen  $\alpha 1(\text{I})$  could also be of great interest and could represent the most reliable biomarker in our case if the extremely strong diminution of its tissue concentration is confirmed in human. However, the most useful biomarkers would be circulating biomarkers that could be easily quantified in plasma, and one main limitation of our approach is that we do not know yet if the transfer of collagen-derived molecules into blood is disturbed in disease. Furthermore, the detection of low abundance molecules in plasma by PRM remains challenging<sup>279</sup>, due to the complexity and very wide concentration range of plasma proteins, and usually requires a specific sample processing step<sup>280,281</sup>.

To overcome this limitation, we could first consider the direct analysis of very small skin biopsies as already done in clinical practice to monitor RDEB progression. Even more promising would be the PRM analysis of wound exudates, a possibility that is suggested by the efficient detection of soluble procollagen-derived peptides in the NaCl fraction of our ECM enrichment protocol. A similar targeted

proteomics-based approach was recently proposed in the literature<sup>170</sup>. Based on the analysis of pig and human wound exudates, the authors identified HTRA-1, a protease involved in matrix remodeling<sup>282,283</sup>, as a candidate biomarker to monitor the critical transition points in wound healing. Similarly, the PRM monitoring of collagen-derived molecules from wound fluids could constitute an easy and non-invasive method to evaluate the response of RDEB wounds to therapies. In addition, other dysregulated and soluble non-collagenous ECM proteins could be analyzed in wound exudates in a multiplexed experiment. PCPE-1, which is decreased at transcript and protein levels in RDEB mouse and at transcript level in RDEB patients, would be a candidate of choice. This protein was already reported to reflect the deposition of collagen especially in fibrosis<sup>284</sup>, and some of its peptides were readily detected in the secretome of human embryonic kidney cells during preliminary PRM analyses conducted in the lab. HTRA-1, which was modified at the protein level in RDEB mice (detected in the triplicate MS/MS experiment only, data not shown) and at the transcript level in fibroblasts of young RDEB patients (Nanostring analyses), could also be monitored in wound exudates, especially if it is suitable to predict healing as suggested above<sup>170</sup>.

In conclusion, we have developed and used an original ECM-centered approach to analyze RDEB skin. The results generated all along this project demonstrated, for the first time, strong alterations in the elaboration of the fibrillar collagen network in RDEB mouse skin. Intra- and extra-cellular processes involved in the biosynthesis and maturation of collagens, especially collagen I, seem to be affected. The next step will be the validation of our main results in the context of the human disease but we can already conclude that the ECM dysregulations reported here pave a very promising way for the medium-term development and validation of new biomarkers in RDEB.

## BIBLIOGRAPHY

1. Lai-Cheong, J. E. & McGrath, J. A. Structure and function of skin, hair and nails. *Medicine* **45**, 347–351 (2017).
2. Proksch, E., Brandner, J. M. & Jensen, J.-M. The skin: an indispensable barrier. *Exp Dermatol* **17**, 1063–1072 (2008).
3. Kolarsick, P. A. J., Kolarsick, M. A. & Goodwin, C. Anatomy and physiology of the skin. *J Dermatol Nurses Assoc* **3**, 204–213 (2011).
4. Vela-Romera, A. *et al.* Characterization of the human ridged and non-ridged skin: a comprehensive histological, histochemical and immunohistochemical analysis. *Histochem Cell Biol* **151**, 57–73 (2019).
5. Eckhart, L., Lippens, S., Tschachler, E. & Declercq, W. Cell death by cornification. *Biochim Biophys Acta Mol Cell Res* **1833**, 3471–3480 (2013).
6. Tobin, D. J. Biochemistry of human skin—our brain on the outside. *Chem Soc Rev* **35**, 52–67 (2006).
7. Koster, M. I. & Roop, D. R. Mechanisms regulating epithelial stratification. *Annu Rev Cell Dev Biol* **23**, 93–113 (2007).
8. Wong, R., Geyer, S., Weninger, W., Guimberteau, J.-C. & Wong, J. K. The dynamic anatomy and patterning of skin. *Exp Dermatol* **25**, 92–98 (2016).
9. Varkey, M., Ding, J. & Tredget, E. E. Superficial dermal fibroblasts enhance basement membrane and epidermal barrier formation in tissue-engineered skin: implications for treatment of skin basement membrane disorders. *Tissue Eng Part A* **20**, (2013).
10. Breitkreutz, D., Koxholt, I., Thiemann, K. & Nischt, R. Skin basement membrane: the foundation of epidermal integrity—BM functions and diverse roles of bridging molecules nidogen and perlecan. *BioMed Res Int* **2013**, 179784 (2013).
11. Breitkreutz, D., Mirancea, N. & Nischt, R. Basement membranes in skin: unique matrix structures with diverse functions? *Histochem Cell Biol* **132**, 1–10 (2009).
12. Has, C. & Nyström, A. Epidermal basement membrane in health and disease. *Curr Top Memb* **76**, 117–170 (2015).
13. Rousselle, P. & Beck, K. Laminin 332 processing impacts cellular behavior. *Cell Adh Migr* **7**, 122–134 (2013).
14. Tuusa, J., Kokkonen, N. & Tasanen, K. BP180/collagen XVII: a molecular view. *Int J Mol Sci* **22**, 12233 (2021).
15. Nyström, A. & Bruckner-Tuderman, L. Matrix molecules and skin biology. *Semin Cell Dev Biol* **89**, 136–146 (2019).
16. Behrens, D. T. *et al.* The epidermal basement membrane is a composite of separate laminin- or collagen IV-containing networks connected by aggregated perlecan, but not by nidogens. *J Biol Chem* **287**, 18700–18709 (2012).
17. Chung, H. J. & Uitto, J. Type VII collagen: the anchoring fibril protein at fault in dystrophic epidermolysis bullosa. *Dermatol Clin* **28**, 93–105 (2010).
18. Rousselle, P., Braye, F. & Dayan, G. Re-epithelialization of adult skin wounds: cellular mechanisms and therapeutic strategies. *Adv Drug Deliv Rev* **146**, 344–365 (2019).
19. Rousselle, P., Montmasson, M. & Garnier, C. Extracellular matrix contribution to skin wound re-epithelialization. *Matrix Biol* **75–76**, 12–26 (2019).

20. Iozzo, R. V. Basement membrane proteoglycans: from cellar to ceiling. *Nat Rev Mol Cell Biol* **6**, 646–656 (2005).
21. Nyström, A., Bornert, O. & Köhl, T. Cell therapy for basement membrane-linked diseases. *Matrix Biol* **57–58**, 124–139 (2017).
22. Schultz, G. S. & Wysocki, A. Interactions between extracellular matrix and growth factors in wound healing. *Wound Repair Regen* **17**, 153–162 (2009).
23. Nilforoushzadeh, M. A. *et al.* Dermal fibroblast cells: biology and function in skin regeneration. *J Skin Stem Cell* **4**, (2017).
24. Krieg, T. & Aumailley, M. The extracellular matrix of the dermis: flexible structures with dynamic functions. *Exp Dermatol* **20**, 689–695 (2011).
25. Ricard-Blum, S. The collagen family. *Cold Spring Harb Perspect Biol* **3**, a004978 (2011).
26. Wagenseil, J. E. & Mecham, R. P. New insights into elastic fiber assembly. *Birth Defect Res C* **81**, 229–240 (2007).
27. Smith, M. M. & Melrose, J. Proteoglycans in normal and healing skin. *Adv Wound Care* **4**, 152–173 (2015).
28. Limbert, G. Mathematical and computational modelling of skin biophysics: a review. *Proc R Soc A* **473**, (2017).
29. Nguyen, A. V. & Soulika, A. M. The dynamics of the skin's immune system. *Int J Mol Sci* **20**, (2019).
30. Bird, J. A., Sánchez-Borges, M., Ansotegui, I. J., Ebisawa, M. & Ortega Martell, J. A. Skin as an immune organ and clinical applications of skin-based immunotherapy. *World Allergy Organ J* **11**, (2018).
31. Swerlick, R. A. The structure and function of the cutaneous vasculature. *J Dermatol* **24**, 734–738 (1997).
32. Skobe, M. & Detmar, M. Structure, function, and molecular control of the skin lymphatic system. *J Invest Dermatol Symp Proc* **5**, 14–19 (2000).
33. Laverdet, B. *et al.* Skin innervation: important roles during normal and pathological cutaneous repair. *Histol Histopathol* **30**, 875–892 (2015).
34. Demidova-Rice, T. N., Hamblin, M. R. & Herman, I. M. Acute and impaired wound healing: pathophysiology and current methods for drug delivery, part 1: normal and chronic wounds: biology, causes, and approaches to care. *Adv Skin Wound Care* **25**, 304–314 (2012).
35. Cañedo-Dorantes, L. & Cañedo-Ayala, M. Skin acute wound healing: a comprehensive review. *Int J Inflamm* **2019**, 1–15 (2019).
36. Gurtner, G. C., Werner, S., Barrandon, Y. & Longaker, M. T. Wound repair and regeneration. *Nature* **453**, 314–321 (2008).
37. Häkkinen, L., Larjava, H. & Koivisto, L. Granulation tissue formation and remodeling. *Endodontic Topics* **24**, 94–129 (2011).
38. Bazin, S., Le Lous, M. & Delaunay, A. Collagen in granulation tissues. *Agents and Actions* **6**, 272–276 (1976).
39. Bailey, A. J., Sims, T. J., Le Lous, M. & Bazin, S. Collagen polymorphism in experimental granulation tissue. *Biochem Biophys Res Commun* **66**, 1160–1165 (1975).
40. Xue, M. & Jackson, C. J. Extracellular matrix reorganization during wound healing and its impact on abnormal scarring. *Adv Wound Care* **4**, 119–136 (2015).
41. Betarbet, U. & Blalock, T. W. Keloids: a review of etiology, prevention, and treatment. *J Clin Aesthet Dermatol* **13**, 33–43 (2020).
42. Wilkinson, H. N. & Hardman, M. J. Wound healing: cellular mechanisms and pathological outcomes. *Open Biol* **10**, 200223 (2020).
43. Mack, M. Inflammation and fibrosis. *Matrix Biol* **68–69**, 106–121 (2018).

44. Do, N. N. & Eming, S. A. Skin fibrosis: models and mechanisms. *Curr Res Transl Med* **64**, 185–193 (2016).
45. Tschumperlin, D. J., Ligresti, G., Hilscher, M. B. & Shah, V. H. Mechanosensing and fibrosis. *J Clin Invest* **128**, 74–84 (2018).
46. Bella, J. & Hulmes, D. J. S. Fibrillar Collagens. in *Fibrous proteins: structures and mechanisms* (eds. Parry, D. A. D. & Squire, J. M.) 457–490 (Springer International Publishing, 2017).
47. Gelse, K., Pöschl, E. & Aigner, T. Collagens - structure, function, and biosynthesis. *Adv Drug Deliv Rev* **55**, 1531–1546 (2003).
48. Onursal, C., Dick, E., Angelidis, I., Schiller, H. B. & Staab-Weijnitz, C. A. Collagen biosynthesis, processing, and maturation in lung ageing. *Front Med* **8**, (2021).
49. Hulmes, D. J. S. Building collagen molecules, fibrils, and suprafibrillar structures. *J Struct Biol* **137**, 2–10 (2002).
50. Goldberg, B., Epstein, E. H. & Sherr, C. J. Precursors of collagen secreted by cultured human fibroblasts. *Proc Natl Acad Sci USA* **69**, 3655–3659 (1972).
51. Goldberg, B. & Sherr, C. J. Secretion and extracellular processing of procollagen by cultured human fibroblasts. *Proc Natl Acad Sci USA* **70**, 361–365 (1973).
52. Exposito, J.-Y., Valcourt, U., Cluzel, C. & Lethias, C. The fibrillar collagen family. *Int J Mol Sci* **11**, 407–426 (2010).
53. Iozzo, R. V. & Schaefer, L. Proteoglycan form and function: a comprehensive nomenclature of proteoglycans. *Matrix Biol* **42**, 11–55 (2015).
54. Theocharidis, G. & Connelly, J. T. Minor collagens of the skin with not so minor functions. *J Anat* **235**, 418–429 (2019).
55. Mak, K. M., Png, C. Y. M. & Lee, D. J. Type V collagen in health, disease, and fibrosis. *Anat Rec* **299**, 613–629 (2016).
56. Bonod-Bidaud, C. *et al.* In vivo evidence for a bridging role of a collagen V subtype at the epidermis–dermis Interface. *J Invest Dermatol* **132**, 1841–1849 (2012).
57. Rojkind, M., Giambrone, M.-A. & Biempica, L. Collagen types in normal and cirrhotic liver. *Gastroenterology* **76**, 710–719 (1979).
58. Minafra, S., Luparello, C., Rallo, F. & Pucci-Minafra, I. Collagen biosynthesis by a breast carcinoma cell strain and biopsy fragments of the primary tumour. *Cell Biol Int Rep* **12**, 895–905 (1988).
59. Chen, M. *et al.* Interactions of the amino-terminal noncollagenous (NC1) domain of type VII collagen with extracellular matrix components. *J Biol Chem* **272**, 14516–14522 (1997).
60. Rattenholl, A. *et al.* Proteinases of the bone morphogenetic protein-1 family convert procollagen VII to mature anchoring fibril collagen. *J Biol Chem* **277**, 26372–26378 (2002).
61. Schonherr, E., Broszat, M., Brandan, E., Bruckner, P. & Kresse, H. Decorin core protein fragment Leu155-Val260 interacts with TGF-beta but does not compete for decorin binding to type I collagen. *Arch Biochem Biophys* **355**, 241–248 (1998).
62. Hildebrand, A., Border, W. A. & Ruoslahti, E. Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor. *Biochem J* **302**, 527–534 (1994).
63. Stöcker, W. & Bode, W. Structural features of a superfamily of zinc-endopeptidases: the metzincins. *Curr Opin Struct Biol* **5**, 383–390 (1995).
64. Gkouveris, I., Nikitakis, N. G., Aseervatham, J. & Ogbureke, K. Matrix metalloproteinases in head and neck cancer: current perspectives. *Metalloproteinases Med* **4**, 47–61 (2017).
65. Jackson, B. C., Nebert, D. W. & Vasiliou, V. Update of human and mouse matrix metalloproteinase families. *Human Genomics* **4**, 194 (2010).

66. Creemers, E. E. J. M., Cleutjens, J. P. M., Smits, J. F. M. & Daemen, M. J. A. P. Matrix metalloproteinase inhibition after myocardial infarction. *Circ Res* **89**, 201–210 (2001).
67. Chelladurai, P., Seeger, W. & Pullamsetti, S. S. Matrix metalloproteinases and their inhibitors in pulmonary hypertension. *Eur Respir J* **40**, 766 (2012).
68. Karamanos, N. K., Theocharis, A. D., Neill, T. & Iozzo, R. V. Matrix modeling and remodeling: a biological interplay regulating tissue homeostasis and diseases. *Matrix Biol* **75–76**, 1–11 (2019).
69. Iida, J. *et al.* Cell surface chondroitin sulfate glycosaminoglycan in melanoma: role in the activation of pro-MMP-2 (pro-gelatinase A). *Biochem J* **403**, 553–563 (2007).
70. Munesue, S. *et al.* A novel function of syndecan-2, suppression of matrix metalloproteinase-2 activation, which causes suppression of metastasis. *J Biol Chem* **282**, 28164–28174 (2007).
71. Sterchi, E. E., Stöcker, W. & Bond, J. S. Meprins, membrane-bound and secreted astacin metalloproteinases. *Mol Asp Med* **29**, 309–328 (2008).
72. Kronenberg, D. *et al.* Processing of procollagen III by meprins: new players in extracellular matrix assembly? *J Invest Dermatol* **130**, 2727–2735 (2010).
73. Broder, C. *et al.* Metalloproteases meprin  $\alpha$  and meprin  $\beta$  are C- and N-procollagen proteinases important for collagen assembly and tensile strength. *Proc Natl Acad Sci USA* **110**, 14219–14224 (2013).
74. Jefferson, T. *et al.* The substrate degradome of meprin metalloproteases reveals an unexpected proteolytic link between meprin  $\beta$  and ADAM10. *Cell Mol Life Sci* **70**, 309–333 (2013).
75. Karmilin, K. *et al.* Mammalian plasma fetuin-B is a selective inhibitor of ovastacin and meprin metalloproteinases. *Sci Rep* **9**, 546 (2019).
76. Vadon-Le Goff, S., Hulmes, D. J. S. & Moali, C. BMP-1/tolloid-like proteinases synchronize matrix assembly with growth factor activation to promote morphogenesis and tissue remodeling. *Matrix Biol* **44–46**, 14–23 (2015).
77. Ge, G. & Greenspan, D. S. BMP1 controls TGF $\beta$ 1 activation via cleavage of latent TGF $\beta$ -binding protein. *J Cell Biol* **175**, 111–120 (2006).
78. Delolme, F. *et al.* Proteolytic control of TGF- $\beta$  co-receptor activity by BMP-1/tolloid-like proteases revealed by quantitative iTRAQ proteomics. *Cell Mol Life Sci* **72**, 1009–1027 (2015).
79. Anastasi, C. *et al.* BMP-1 disrupts cell adhesion and enhances TGF-beta activation through cleavage of the matricellular protein thrombospondin-1. *Sci Signal* **13**, eaba3880 (2020).
80. Vadon-Le Goff, S. *et al.* The dark side of procollagen C-proteinase enhancer-2 (PCPE-2): inhibition of bone morphogenetic protein-1/tolloid-like proteinases. *Proc Natl Acad Sci USA* (Submitted).
81. Mulooly, M., McGowan, P. M., Crown, J. & Duffy, M. J. The ADAMs family of proteases as targets for the treatment of cancer. *Cancer Biol Ther* **17**, 870–880 (2016).
82. Kawaguchi, M. & Hearing, V. J. The Roles of ADAMs Family Proteinases in Skin Diseases. *Enzyme Res* **2011**, 482498 (2011).
83. Lambrecht, B. N., Vanderkerken, M. & Hammad, H. The emerging role of ADAM metalloproteinases in immunity. *Nat Rev Immunol* **18**, 745–758 (2018).
84. Kelwick, R., Desanlis, I., Wheeler, G. N. & Edwards, D. R. The ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) family. *Genome Biol* **16**, (2015).
85. De Veer, S. J., Furio, L., Harris, J. M. & Hovnanian, A. Proteases: common culprits in human skin disorders. *Trends Mol Med* **20**, 166–178 (2014).
86. Leduc, C. *et al.* In vivo N-terminomics highlights novel functions of ADAMTS2 and ADAMTS14 in skin collagen matrix building. *Front Mol Biosci* **8**, (2021).
87. Bekhouche, M. *et al.* Determination of the substrate repertoire of ADAMTS2, 3, and 14 significantly broadens their functions and identifies extracellular matrix organization and TGF- $\beta$  signaling as primary targets. *FASEB J* **30**, 1741–1756 (2016).

88. Oichi, T. *et al.* Adamts17 is involved in skeletogenesis through modulation of BMP-Smad1/5/8 pathway. *Cell Mol Life Sci* **76**, 4795–4809 (2019).
89. Naba, A., Clauser, K. R., Hoersch, S., Carr, S. A. & Hynes, R. O. The Matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics* **11**, M111.014647 (2012).
90. Heit, C. *et al.* Update of the human and mouse SERPIN gene superfamily. *Hum Genomics* **7**, (2013).
91. Matrisome project website. <http://matrisomeproject.mit.edu/>.
92. Nauroy, P., Hughes, S., Naba, A. & Ruggiero, F. The in-silico zebrafish matrisome: a new tool to study extracellular matrix gene and protein functions. *Matrix Biol* **65**, 5–13 (2018).
93. Davis, M. N., Horne-Badovinac, S. & Naba, A. In-silico definition of the Drosophila melanogaster matrisome. *Matrix Biol Plus* **4**, (2019).
94. Teuscher, A. C. *et al.* The in-silico characterization of the Caenorhabditis elegans matrisome and proposal of a novel collagen classification. *Matrix Biol Plus* **1**, (2019).
95. Online Pepchem matrisome database. <http://matrisomedb.pepchem.org>.
96. PRIDE platform. <https://www.ebi.ac.uk/pride/>.
97. Naba, A., Clauser, K. R., Lamar, J. M., Carr, S. A. & Hynes, R. O. Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *eLife* **3**, e01308 (2014).
98. Naba, A., Clauser, K. R., Mani, D. R., Carr, S. A. & Hynes, R. O. Quantitative proteomic profiling of the extracellular matrix of pancreatic islets during the angiogenic switch and insulinoma progression. *Sci Rep* **7**, 40495 (2017).
99. Schiller, H. B. *et al.* Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. *Mol Syst Biol* **11**, 819 (2015).
100. Gocheva, V. *et al.* Quantitative proteomics identify Tenascin-C as a promoter of lung cancer progression and contributor to a signature prognostic of patient survival. *Proc Natl Acad Sci USA* **114**, 5625–5634 (2017).
101. Naba, A. *et al.* Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver. *BMC Cancer* **14**, 518 (2014).
102. COL7A1 gene homepage. *Leiden Open Variation Database* <https://databases.lovd.nl/shared/genes/COL7A1> (2022).
103. Wynn, T. A. & Ramalingam, T. R. Mechanisms of fibrosis: therapeutic translation for fibrotic disease. *Nat Med* **18**, 1028–1040 (2012).
104. Guerra, L., Odorisio, T., Zambruno, G. & Castiglia, D. Stromal microenvironment in type VII collagen-deficient skin: the ground for squamous cell carcinoma development. *Matrix Biol* **63**, 1–10 (2017).
105. Maldonado-Colin, G., Hernández-Zepeda, C., Durán-McKinster, C. & García-Romero, M. Inherited epidermolysis bullosa: a multisystem disease of skin and mucosae fragility. *Indian J Paediatr Dermatol* **18**, 267–273 (2017).
106. Kiritsi, D. & Nyström, A. Recent advances in understanding and managing epidermolysis bullosa. *F1000 Res* **7**, (2018).
107. Tong, L. *et al.* The eye in epidermolysis bullosa. *Br J Ophthalmol* **83**, 323–326 (1999).
108. Ng, Y.-Z. *et al.* Fibroblast-derived dermal matrix drives development of aggressive cutaneous squamous cell carcinoma in patients with recessive dystrophic epidermolysis bullosa. *Cancer Res* **72**, 3522–3534 (2012).
109. Fine, J.-D., Johnson, L. B., Weiner, M., Li, K.-P. & Suchindran, C. Epidermolysis bullosa and the risk of life-threatening cancers: the national EB registry experience. *J Am Acad Dermatol* **60**, 203–211 (2009).

110. Tidman, M. J. & Eady, R. A. J. Evaluation of anchoring fibrils and other components of the dermal-epidermal junction in dystrophic epidermolysis bullosa by a quantitative ultrastructural technique. *J Invest Dermatol* **84**, 374–377 (1985).
111. Shinkuma, S. Dystrophic epidermolysis bullosa: a review. *Clin Cosmet Investig Dermatol* **8**, 275–284 (2015).
112. Pfendner, E. G. & Lucky, A. W. Dystrophic epidermolysis bullosa. in *Dystrophic Epidermolysis Bullosa* (2018).
113. Diociaiuti, A. *et al.* EGFR inhibition for metastasized cutaneous squamous cell carcinoma in dystrophic epidermolysis bullosa. *Orphanet J Rare Dis* **14**, (2019).
114. Nyström, A., Bruckner-Tuderman, L. & Kiritsi, D. Dystrophic epidermolysis bullosa: secondary disease mechanisms and disease modifiers. *Front Genet* **12**, (2021).
115. Odorisio, T. *et al.* Monozygotic twins discordant for recessive dystrophic epidermolysis bullosa phenotype highlight the role of TGF- $\beta$  signalling in modifying disease severity. *Hum Mol Genet* **23**, 3907–3922 (2014).
116. Hovnanian, A. *et al.* Characterization of 18 new mutations in COL7A1 in recessive dystrophic epidermolysis bullosa provides evidence for distinct molecular mechanisms underlying defective anchoring fibril formation. *Am J Hum Genet* **61**, 599–610 (1997).
117. Titeux, M. *et al.* A frequent functional SNP in the MMP1 promoter is associated with higher disease severity in recessive dystrophic epidermolysis bullosa. *Hum Mutat* **29**, 267–276 (2008).
118. El Hachem, M. *et al.* Multicentre consensus recommendations for skin care in inherited epidermolysis bullosa. *Orphanet J Rare Dis* **9**, (2014).
119. Denyer, J. Wound management in epidermolysis bullosa. *EWMA J* **16**, (2016).
120. Isaacs, M., Veerkamps, P. & Somani, A. K. Small intestinal submucosa matrix as a novel therapy for wounds in dystrophic epidermolysis bullosa. *Dermatol Surg* **45**, 863–864 (2019).
121. Le, N. K., Billington, A., Harrington, M. & Seminario-Vidal, L. Management of recessive dystrophic epidermolysis bullosa in a newborn with porcine-derived extracellular matrix. *Plast Reconstr Surg* **7**, (2019).
122. Beylin, A. K., Gurskaya, N. G. & Vorotelyak, E. A. Methods of gene therapy for treatment of inherited epidermolysis bullosa. *Moscow Univ Biol Sci Bull* **73**, 191–198 (2018).
123. Bruckner-Tuderman, L. Newer treatment modalities in epidermolysis bullosa. *Indian Dermatol Online J* **10**, (2019).
124. Latella, M. C. *et al.* Correction of recessive dystrophic epidermolysis bullosa by transposon-mediated integration of COL7A1 in transplantable patient-derived primary keratinocytes. *J Invest Dermatol* **137**, 836–844 (2017).
125. Siprashvili, Z. *et al.* Long-term type VII collagen restoration to human epidermolysis bullosa skin tissue. *Hum Gene Ther* **21**, 1299–1310 (2010).
126. Eichstadt, S. *et al.* Phase 1/2a clinical trial of gene-corrected autologous cell therapy for recessive dystrophic epidermolysis bullosa. *JCI Insight* **4**, (2019).
127. Lwin, S. M. *et al.* Safety and early efficacy outcomes for lentiviral fibroblast gene therapy in recessive dystrophic epidermolysis bullosa. *JCI Insight* **4**, (2019).
128. Naso, G. & Petrova, A. Cellular therapy options for genetic skin disorders with a focus on recessive dystrophic epidermolysis bullosa. *Br Med Bull* **136**, 30–45 (2020).
129. Gurevich, I. *et al.* In vivo topical gene therapy for recessive dystrophic epidermolysis bullosa: a phase 1 and 2 trial. *Nat Med* **28**, 780–788 (2022).
130. Woodley, D. T. *et al.* Gentamicin induces functional type VII collagen in recessive dystrophic epidermolysis bullosa patients. *J Clin Invest* **127**, 3028–3038 (2017).
131. Bremer, J. *et al.* Antisense oligonucleotide-mediated exon skipping as a systemic therapeutic approach for recessive dystrophic epidermolysis bullosa. *Mol Ther* **5**, (2016).

132. Bornert, O. *et al.* QR-313, an antisense oligonucleotide, shows therapeutic efficacy for treatment of dominant and recessive dystrophic epidermolysis bullosa: a preclinical study. *J Invest Dermatol* **141**, 883–893 (2021).
133. Remington, J. *et al.* Injection of recombinant human type VII collagen corrects the disease phenotype in a murine model of dystrophic epidermolysis bullosa. *Mol Ther* **17**, 26–33 (2009).
134. Hou, Y. *et al.* Intravenously administered recombinant human type VII collagen derived from chinese hamster ovary cells reverses the disease phenotype in recessive dystrophic epidermolysis bullosa mice. *J Invest Dermatol* **135**, 3060–3067 (2015).
135. Köhl, T. *et al.* Collagen VII half-life at the dermal-epidermal junction zone: implications for mechanisms and therapy of genodermatoses. *J Invest Dermatol* **136**, 1116–1123 (2016).
136. Dussoyer, M., Michopoulou, A. & Rousselle, P. Decellularized scaffolds for skin repair and regeneration. *Appl Sci* **10**, 3435 (2020).
137. Koullis, N., Moysidis, S. N., Siegel, L. M. & Song, J. C. Long-term follow-up of amniotic membrane graft for the treatment of symblepharon in a patient with recessive dystrophic epidermolysis bullosa. *Cornea* **35**, 1242–1244 (2016).
138. Goyal, R. Amniotic membrane transplantation in children with symblepharon and massive pannus. *Arch Ophthalmol* **124**, 1435–1440 (2006).
139. Nyström, A. & Bruckner-Tuderman, L. Gene therapy for epidermolysis bullosa: sticky business. *Mol Ther* **24**, 2035–2036 (2016).
140. Fritsch, A. *et al.* A hypomorphic mouse model of dystrophic epidermolysis bullosa reveals mechanisms of disease and response to fibroblast therapy. *J Clin Invest* **118**, 1669–1679 (2008).
141. Nyström, A. *et al.* Losartan ameliorates dystrophic epidermolysis bullosa and uncovers new disease mechanisms. *EMBO Mol Med* **7**, 1211–1228 (2015).
142. Nyström, A. *et al.* Impaired lymphoid extracellular matrix impedes antibacterial immunity in epidermolysis bullosa. *Proc Natl Acad Sci USA* **115**, 705–714 (2018).
143. Mittapalli, V. R. *et al.* Injury-driven stiffening of the dermis expedites skin carcinoma progression. *Cancer Res* **76**, 940–951 (2016).
144. Nyström, A. *et al.* Collagen VII plays a dual role in wound healing. *J Clin Invest* **123**, 3498–3509 (2013).
145. Breitenbach, J. S. *et al.* Transcriptome and ultrastructural changes in dystrophic epidermolysis bullosa resemble skin aging. *Aging* **7**, 389–411 (2015).
146. Küttner, V. *et al.* Global remodelling of cellular microenvironment due to loss of collagen VII. *Mol Syst Biol* **9**, (2013).
147. Braga, T. T., Agudelo, J. S. H. & Camara, N. O. S. Macrophages during the fibrotic process: M2 as friend and foe. *Front Immunol* **6**, (2015).
148. Nguyen, J. K., Austin, E., Huang, A., Mamalis, A. & Jagdeo, J. The IL-4/IL-13 axis in skin fibrosis and scarring: mechanistic concepts and therapeutic targets. *Arch Dermatol Res* **312**, 81–92 (2020).
149. Borthwick, L. A., Wynn, T. A. & Fisher, A. J. Cytokine mediated tissue fibrosis. *Biochim Biophys Acta Mol Basis Dis* **1832**, 1049–1060 (2013).
150. Meng, X., Nikolic-Paterson, D. J. & Lan, H. Y. TGF- $\beta$ : the master regulator of fibrosis. *Nat Rev Nephrol* **12**, 325–338 (2016).
151. Bhattacharyya, S. *et al.* Tenascin-C drives persistence of organ fibrosis. *Nat Commun* **7**, (2016).
152. Norris, R. A. *et al.* Periostin regulates collagen fibrillogenesis and the biomechanical properties of connective tissues. *J Cell Biochem* **101**, 695–711 (2007).
153. Ashley, S. L., Wilke, C. A., Kim, K. K. & Moore, B. B. Periostin regulates fibrocyte function to promote myofibroblast differentiation and lung fibrosis. *Mucosal Immunol* **10**, 341–351 (2017).
154. Zhao, S. *et al.* Periostin expression is upregulated and associated with myocardial fibrosis in human failing hearts. *J Cardiol* **63**, 373–378 (2014).

155. Chacón-Solano, E. *et al.* Fibroblast activation and abnormal extracellular matrix remodelling as common hallmarks in three cancer-prone genodermatoses. *Br J Dermatol* **181**, 512–522 (2019).
156. Atanasova, V. S. *et al.* Thrombospondin-1 is a major activator of TGF- $\beta$  signaling in recessive dystrophic epidermolysis bullosa fibroblasts. *J Invest Dermatol* **139**, 1497–1505 (2019).
157. van der Kooi-Pol, M. M. *et al.* High anti-staphylococcal antibody titers in patients with epidermolysis bullosa relate to long-term colonization with alternating types of staphylococcus aureus. *J Invest Dermatol* **133**, 847–850 (2013).
158. Thomson, J. J. Further experiments on positive rays. *Lond Edinb Dubl Phil Mag* **24**, 209–253 (1912).
159. Hossain, M. *Selected reaction monitoring mass spectrometry (SRM-MS) in proteomics*. (Springer International Publishing, 2020).
160. Sinha, A. & Mann, M. A beginner's guide to mass spectrometry-based proteomics. *Biochem* **42**, 64–69 (2020).
161. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212–217 (2006).
162. Sleno, L. & Volmer, D. A. Ion activation methods for tandem mass spectrometry. *J Mass Spectrom* **39**, 1091–1112 (2004).
163. Menneteau, T. Développement de méthodes de spectrométrie de masse pour l'étude de la dynamique des complexes de protéasome: application aux cellules stromales/souches mésenchymateuses humaines. (Université de Toulouse, 2018).
164. Fornelli, L., Schmid, A. W., Grasso, L., Vogel, H. & Tsybin, Y. O. Deamidation and transamidation of substance P by tissue transglutaminase revealed by electron-Capture dissociation fourier transform mass spectrometry. *Chem Eur J* **17**, 486–497 (2011).
165. Pappireddi, N., Martin, L. & Wühr, M. A review on quantitative multiplexed proteomics. *Eur J Chem Biol* **20**, 1210–1224 (2019).
166. El Anead, A., Cohen, A. & Banoub, J. Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Appl Spectrosc Rev* **44**, 210–230 (2009).
167. Scheltema, R. A. *et al.* The Q Exactive HF, a benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. *Mol Cell Proteomics* **13**, 3698–3708 (2014).
168. Eruzun, H. *et al.* Serum endotrophin levels in patients with heart failure with reduced and mid-range ejection fraction. *Eur J Intern Med* **64**, 29–32 (2019).
169. Juhl, P. *et al.* Serum biomarkers of collagen turnover as potential diagnostic tools in diffuse systemic sclerosis: a cross-sectional study. *PLoS ONE* **13**, (2018).
170. Sabino, F. *et al.* Comparative degradomics of porcine and human wound exudates unravels biomarker candidates for assessment of wound healing progression in trauma patients. *J Invest Dermatol* **138**, 413–422 (2018).
171. Lai, Z. W. *et al.* Formalin-fixed, paraffin-embedded tissues (FFPE) as a robust source for the profiling of native and protease-generated protein amino termini. *Mol Cell Proteomics* **15**, 2203–2213 (2016).
172. Wadsworth, C. & Buckley, M. Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Commun Mass Spectrom* **28**, 605–615 (2013).
173. Timp, W. & Timp, G. Beyond mass spectrometry, the next step in proteomics. *Sci Adv* **2**, (2020).
174. Dupree, E. J. *et al.* A critical review of bottom-up proteomics: the good, the bad, and the future of this field. *Proteomes* **8**, (2020).
175. Smith, L. M. & Kelleher, N. L. Proteoform: a single term describing protein complexity. *Nat Methods* **10**, 186–187 (2013).

176. Bichmann, L. *et al.* DIAproteomics: a multifunctional data analysis pipeline for data-independent acquisition proteomics and peptidomics. *J Proteome Res* **20**, 3758–3766 (2021).
177. Tian, X., Permentier, H. P. & Bischoff, R. Chemical isotope labeling for quantitative proteomics. *Mass Spectrom Rev* **2021**, 1–13 (2021).
178. <https://david.ncifcrf.gov/>.
179. <http://pantherdb.org/>.
180. Nguyen, L., Brun, V., Combes, F., Loux, V. & Vandenbrouck, Y. Designing an in silico strategy to select tissue-leakage biomarkers using the galaxy framework. *Meth Mol Biol* **1959**, 275–289 (2019).
181. Harlan, R. & Zhang, H. Targeted proteomics: a bridge between discovery and validation. *Expert Rev Proteomics* **11**, 657–661 (2014).
182. Wright, M. J., Thomas, R. L., Stanford, P. E. & Horvath, A. R. Multiple reaction monitoring with multistage fragmentation (MRM3) detection enhances selectivity for LC-MS/MS analysis of plasma free metanephrines. *Clin Chem* **61**, 505–513 (2015).
183. Anand, S., Samuel, M., Ang, C.-S., Keerthikumar, S. & Mathivanan, S. Label-based and label-free strategies for protein quantitation. *Meth Mol Biol* **1549**, 31–43 (2017).
184. Ankney, A. J., Muneer, A. & Chen, X. Relative and absolute quantitation in mass spectrometry-based proteomics. *Annu Rev Anal Chem* **11**, 48–77 (2018).
185. Dussoyer, M., Moali, C. & Lagoutte, P. Potentiel de la spectrométrie de masse pour l'identification de biomarqueurs des pathologies cutanées. in *Cours de Biologie de la Peau (COBIP)* 7–18 (2022).
186. Lindemann, C. *et al.* Strategies in relative and absolute quantitative mass spectrometry based proteomics. *Biol Chem* **398**, 687–699 (2017).
187. Lai, X., Wang, L. & Witzmann, F. A. Issues and applications in label-free quantitative mass spectrometry. *Int J Proteomics* **2013**, (2013).
188. Zanivan, S., Krueger, M. & Mann, M. In vivo quantitative proteomics: the SILAC mouse. *Meth Mol Biol* **757**, 453–450 (2012).
189. Frost, D. C., Feng, Y. & Li, L. 21-plex DiLeu isobaric tags for high-throughput quantitative proteomics. *Anal Chem* **92**, 8228–8234 (2020).
190. Stadlmeier, M., Bogena, J., Wallner, M. & Carell, T. A sulfoxide-based isobaric labelling reagent for accurate quantitative mass spectrometry. *Angew Chem Int Ed Engl* **57**, 2958–2962 (2018).
191. Braun, G. R. *et al.* Generation of multiple reporter ions from a single isobaric reagent increases multiplexing capacity for quantitative proteomics. *Anal Chem* **87**, 9855–9863 (2015).
192. Naba, A., Clauser, K. R. & Hynes, R. O. Enrichment of extracellular matrix proteins from tissues and digestion into peptides for mass spectrometry analysis. *J Vis Exp* 53057 (2015).
193. Didangelos, A. *et al.* Proteomics characterization of extracellular space components in the human aorta. *Mol Cell Proteomics* **9**, 2048–2062 (2010).
194. Barallobre-Barreiro, J. *et al.* Proteomics analysis of cardiac extracellular matrix remodeling in a porcine model of ischemia/reperfusion injury. *Circulation* **125**, 789–802 (2012).
195. Barallobre-Barreiro, J. *et al.* Extracellular matrix remodelling in response to venous hypertension: proteomics of human varicose veins. *Cardiovasc Res* **110**, 419–430 (2016).
196. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).
197. Tessier, A. Procollagen C-proteinase enhancers in skin wound healing: expression, functions and therapeutic potential. (Université de Lyon, Université de Freiburg, 2018).
198. Fisher, L. W., Lindner, W., Young, M. F. & Termine, J. D. Synthetic peptide antisera: their production and use in the cloning of matrix proteins. *Connect Tissue Res* **21**, 43–50 (1989).
199. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68–85 (2019).

200. Liu, B. *et al.* Matrisome provides a supportive microenvironment for skin functions of diverse species. *ACS Biomater Sci Eng* **6**, 5720–5733 (2020).
201. Zhang, S. *et al.* The “Matrisome” reveals the characterization of skin keloid microenvironment. *FASEB J* **35**, e21237 (2021).
202. Nacu, N. *et al.* Macrophages produce TGF- $\beta$ -induced ( $\beta$ -ig-h3) following ingestion of apoptotic cells and regulate MMP14 levels and collagen turnover in fibroblasts. *J Immunol* **180**, 5036–5044 (2008).
203. Boye, T. L. *et al.* Annexin A4 and A6 induce membrane curvature and constriction during cell membrane repair. *Nat Commun* **8**, 1623–1623 (2017).
204. Rothnagel, J. A. & Rogers, G. E. Trichohyalin, an intermediate filament-associated protein of the hair follicle. *J Cell Biol* **102**, 1419–1429 (1986).
205. Vandooren, J. & Itoh, Y. Alpha-2-macroglobulin in inflammation, immunity and infections. *Front Immunol* **12**, (2021).
206. Burgener, S. S. *et al.* Cathepsin G inhibition by serpinb1 and serpinb6 prevents programmed necrosis in neutrophils and monocytes and reduces GSDMD-driven inflammation. *Cell Rep* **27**, 3646-3656.e5 (2019).
207. Ishida, Y. & Nagata, K. Chapter nine - Hsp47 as a collagen-specific molecular chaperone. in *Methods in Enzymology* (eds. Whisstock, J. C. & Bird, P. I.) vol. 499 167–182 (Academic Press, 2011).
208. Martínez-Martínez, E. *et al.* Differential proteomics identifies reticulocalbin-3 as a novel negative mediator of collagen production in human cardiac fibroblasts. *Sci Rep* **7**, 12192 (2017).
209. Park, N. R. *et al.* Reticulocalbin 3 is involved in postnatal tendon development by regulating collagen fibrillogenesis and cellular maturation. *Sci Rep* **11**, 10868 (2021).
210. Chacón-Solano, E. *et al.* Oxidative stress imbalance as contributing factor in the establishment of fibrosis in recessive dystrophic epidermolysis bullosa. *J Invest Dermatol* **139**, (2019).
211. McGrath, J. A., Ishida-Yamamoto, A., O’Grady, O., Leigh, I. M. & Eady, R. A. Structural variations in anchoring fibrils in dystrophic epidermolysis bullosa: correlation with type VII collagen expression. *J Invest Dermatol* **100**, 366–372 (1993).
212. Bernasconi, R. *et al.* Pro-inflammatory immunity supports fibrosis advancement in epidermolysis bullosa. *EMBO Mol Med* **13**, e14392 (2021).
213. Thriene, K. *et al.* Combinatorial omics analysis reveals perturbed lysosomal homeostasis in collagen VII-deficient keratinocytes. *Mol Cell Proteomics* **17**, 565–579 (2018).
214. Villone, D. *et al.* Supramolecular interactions in the dermo-epidermal junction zone: anchoring fibril-collagen VII tightly binds to banded collagen fibrils. *J Biol Chem* **283**, 24506–24513 (2008).
215. Sun, M. *et al.* Collagen V is a dominant regulator of collagen fibrillogenesis: dysfunctional regulation of structure and function in a corneal-stroma-specific Col5a1-null mouse model. *J Cell Sci* **124**, 4096–4105 (2011).
216. Corsi, A. *et al.* Phenotypic effects of biglycan deficiency are linked to collagen fibril abnormalities, are synergized by decorin deficiency, and mimic Ehlers-Danlos-like changes in bone and other connective tissues. *J Bone Mineral Res* **17**, 1180–1189 (2002).
217. Chakravarti, S. *et al.* Corneal opacity in lumican-null mice: defects in collagen fibril structure and packing in the posterior stroma. *Invest Ophthalmol Vis Sci* **41**, 3365–3373 (2000).
218. Bodemer, C. *et al.* Skin expression of metalloproteinases and tissue inhibitor of metalloproteinases in sibling patients with recessive dystrophic epidermolysis and intrafamilial phenotypic variation. *J Invest Dermatol* **121**, 273–279 (2003).
219. Akasaka, E., Kleiser, S., Sengle, G., Bruckner-Tuderman, L. & Nyström, A. Diversity of mechanisms underlying Latent TGF- $\beta$  activation in recessive dystrophic epidermolysis bullosa. *J Invest Dermatol* **141**, 1450–1460 (2021).

220. Ricard-Blum, S., Baffet, G. & Th eret, N. Molecular and tissue alterations of collagens in fibrosis. *Matrix Biol* **68–69**, 122–149 (2018).
221. Ho, Y. Y., Lagares, D., Tager, A. M. & Kapoor, M. Fibrosis—a lethal component of systemic sclerosis. *Nat Rev Rheumatol* **10**, 390–402 (2014).
222. Jinnin, M. Mechanisms of skin fibrosis in systemic sclerosis. *J Dermatol* **37**, 11–25 (2010).
223. Braun-Falco, O. & Rupec, M. Collagen fibrils of the scleroderma in ultra-thin skin sections. *Nature* **202**, 708–709 (1964).
224. Hayes, R. L. & Rodnan, G. P. The ultrastructure of skin in progressive systemic sclerosis (scleroderma). *Am J Pathol* **63**, 433–442 (1971).
225. Barallobre-Barreiro, J. *et al.* Cartilage-like composition of keloid scar extracellular matrix suggests fibroblast mis-differentiation in disease. *Matrix Biol Plus* **4**, 100016 (2019).
226. Wang, C. *et al.* Type III collagen is a key regulator of the collagen fibrillar structure and biomechanics of articular cartilage and meniscus. *Matrix Biol* **85–86**, 47–67 (2020).
227. Perlish, J. S., Lemlich, G. & Fleischmajer, R. Identification of collagen fibrils in scleroderma skin. *J Invest Dermatol* **90**, 48–54 (1988).
228. K ohler, A. *et al.* New specific HSP47 functions in collagen subfamily chaperoning. *FASEB J* **34**, 12040–12052 (2020).
229. Ishida, Y. *et al.* Type I collagen in Hsp47-null cells is aggregated in endoplasmic reticulum and deficient in N-propeptide processing and fibrillogenesis. *Mol Biol Cell* **17**, 2346–2355 (2006).
230. Muir, A. M. *et al.* BMP1-like proteinases are essential to the structure and wound healing of skin. *Matrix Biol* **56**, 114–131 (2016).
231. Sch onborn, K. *et al.* Role of collagen XII in skin homeostasis and repair. *Matrix Biol* **94**, 57–76 (2020).
232. Jung, J.-C. *et al.* Collagen fibril growth during chicken tendon development: matrix metalloproteinase-2 and its activation. *Cell Tissue Res* **336**, 79–89 (2009).
233. Zigrino, P. *et al.* Fibroblast-derived MMP-14 regulates collagen homeostasis in adult skin. *J Invest Dermatol* **136**, 1575–1583 (2016).
234. Pischon, N. *et al.* Lysyl oxidase (Lox) gene deficiency affects osteoblastic phenotype. *Calcif Tissue Int* **85**, 119–126 (2009).
235. Kadler, K. E., Hojima, Y. & Prockop, D. J. 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* **262**, 15696–15701 (1987).
236. Steiglitz, B. M. *et al.* Procollagen C proteinase enhancer 1 genes are important determinants of the mechanical properties and geometry of bone and the ultrastructure of connective tissues. *Mol Cell Biol* **26**, 238–249 (2006).
237. Tian, C. *et al.* Suppression of pancreatic ductal adenocarcinoma growth and metastasis by fibrillar collagens produced selectively by tumor cells. *Nat Commun* **12**, 2328 (2021).
238. Fleischmajer R *et al.* Ultrastructural identification of extension aminopropeptides of type I and III collagens in human skin. *Proc Natl Acad Sci USA* **78**, 7360–7364 (1981).
239. Lindahl, K. *et al.* COL1 C-propeptide cleavage site mutations cause high bone mass osteogenesis imperfecta. *Hum Mutat* **32**, 598–609 (2011).
240. Syx, D. *et al.* Defective proteolytic processing of fibrillar procollagens and procollagen due to biallelic BMP1 mutations results in a severe, progressive form of osteogenesis imperfecta. *J Bone Mineral Res* **30**, 1445–1456 (2015).
241. Colige, A. *et al.* Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene. *Am J Hum Genet* **65**, 308–317 (1999).

242. Wong, H. *et al.* Multiscale modelling of the extracellular matrix. *Matrix Biol Plus* **13**, 100096 (2022).
243. Garrigue-Antar, L., François, V. & Kadler, K. E. Deletion of epidermal growth factor-like domains converts mammalian tolloid into a chordinase and effective procollagen C-proteinase. *J Biol Chem* **279**, 49835–49841 (2004).
244. Arnold, P. *et al.* Deficiency of the DSPP-cleaving enzymes meprin  $\alpha$  and meprin  $\beta$  does not result in dentin malformation in mice. *Cell Tissue Res* **367**, 351–358 (2017).
245. Dafforn, T. R., Della, M. & Miller, A. D. The molecular interactions of heat shock protein 47 (Hsp47) and their implications for collagen biosynthesis. *J Biol Chem* **276**, 49310–49319 (2001).
246. Thomson, C. A. & Ananthanarayanan, V. S. A method for expression and purification of soluble, active Hsp47, a collagen-specific molecular chaperone. *Protein Expr Purif* **23**, 8–13 (2001).
247. Andersson, A. *et al.* Development of parallel reaction monitoring assays for cerebrospinal fluid proteins associated with Alzheimer's disease. *Clinica Chimica Acta* **494**, 79–93 (2019).
248. Gulbrandsen, A. *et al.* Development of robust targeted proteomics assays for cerebrospinal fluid biomarkers in multiple sclerosis. *Clin Proteomics* **17**, (2020).
249. Martinez-Garcia, E. *et al.* Development of a sequential workflow based on LC-PRM for the verification of endometrial cancer protein biomarkers in uterine aspirate samples. *Oncotarget* (2016).
250. Ura, B. *et al.* A proteomic approach for the identification of biomarkers in endometrial cancer uterine aspirate. *Oncotarget* **8**, 109536–109545 (2017).
251. Sivakumar, P. *et al.* Integrated plasma proteomics and lung transcriptomics reveal novel biomarkers in idiopathic pulmonary fibrosis. *Respiratory Res* **22**, 273 (2021).
252. Panigrahi, A. *et al.* A rapid LC-MS/MS-PRM assay for serologic quantification of sialylated O-HPX glycoforms in patients with liver fibrosis. *Molecules* **27**, 2213 (2022).
253. Dickson, L. A. *et al.* The exon/intron structure of the 3'-region of the pro alpha 2(I) collagen gene. *J Biol Chem* **256**, 8407–8415 (1981).
254. Marcos-Garcés, V. *et al.* Age-related dermal collagen changes during development, maturation and ageing - a morphometric and comparative study. *J Anat* **225**, 98–108 (2014).
255. Manturova, N. E., Smirnova, G. O., Stupin, V. A. & Silina, E. V. The ratio of collagen types I/III as a marker of skin aging and prognosis of aesthetic facial surgery results. *J Pharm Sci Res* **10**, (2018).
256. DiChiara, A. S. *et al.* A cysteine-based molecular code informs collagen C-propeptide assembly. *Nat Commun* **9**, 4206 (2018).
257. Suzuki, N. *et al.* Failure of ventral body wall closure in mouse embryos lacking a procollagen C-proteinase encoded by *Bmp1*, a mammalian gene related to *Drosophila* tolloid. *Development* **122**, 3587–3595 (1996).
258. Pappano, W. N., Steiglit, B. M., Scott, I. C., Keene, D. R. & Greenspan, D. S. Use of *Bmp1*/Tll1 doubly homozygous null mice and proteomics to identify and validate in vivo substrates of bone morphogenetic protein 1/tolloid-like metalloproteinases. *Mol Cell Biol* **23**, 4428–4438 (2003).
259. Birk, D. E., Fitch, J. M., Babiarz, J. P., Doane, K. J. & Linsenmayer, T. F. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. *J Cell Sci* **95**, 649–657 (1990).
260. Bekhouche, M. & Colige, A. The procollagen N-proteinases ADAMTS2, 3 and 14 in pathophysiology. *Matrix Biol* **44–46**, 46–53 (2015).
261. Colige, A. *et al.* Domains and maturation processes that regulate the activity of ADAMTS-2, a metalloproteinase cleaving the aminopropeptide of fibrillar procollagens types I–III and V. *J Biol Chem* **280**, 34397–34408 (2005).

262. Imamura, Y., Steiglitz, B. M. & Greenspan, D. S. Bone morphogenetic protein-1 processes the NH<sub>2</sub>-terminal propeptide, and a furin-like proprotein convertase processes the COOH-terminal propeptide of pro- $\alpha$ 1(V) collagen. *J Biol Chem* **273**, 27511–27517 (1998).
263. Scott, I. C. *et al.* Bone morphogenetic protein-1 processes probiglycan. *J Biol Chem* **275**, 30504–30511 (2000).
264. Ito, S. & Nagata, K. Roles of the endoplasmic reticulum-resident, collagen-specific molecular chaperone Hsp47 in vertebrate cells and human disease. *J Biol Chem* **294**, 2133–2141 (2019).
265. Omari, S. *et al.* Mechanisms of procollagen and HSP47 sorting during ER-to-Golgi trafficking. *Matrix Biol* **93**, 79–94 (2020).
266. Mortensen, J. H. & Karsdal, M. A. Type VII Collagen. in *Biochemistry of Collagens, Laminins and Elastin* (ed. Karsdal, M. A.) 57–60 (Academic Press, 2016).
267. Küttner, V., Mack, C., Gretzmeier, C., Bruckner-Tuderman, L. & Dengjel, J. Loss of collagen VII is associated with reduced transglutaminase 2 abundance and activity. *J Invest Dermatol* **134**, 2381–2389 (2014).
268. Bernasconi, R. Identification of disease-progressing mechanisms and assessment of angiotensin-(1-7) as symptom relief therapy for dystrophic epidermolysis bullosa. (Albert-Ludwigs-University Freiburg, 2020).
269. Eyre, D. R., Weis, M. A. & Wu, J.-J. Advances in collagen cross-link analysis. *Methods* **45**, 65–74 (2008).
270. Joshi, A., Zahoor, A. & Buson, A. Measurement of collagen cross-links from tissue samples by mass spectrometry. in *Collagen: Methods and Protocols* (eds. Sagi, I. & Afratis, N. A.) 79–93 (Springer New York, 2019). doi:10.1007/978-1-4939-9095-5\_6.
271. Varki, R., Sadowski, S., Uitto, J. & Pfenner, E. Epidermolysis bullosa. II. Type VII collagen mutations and phenotype-genotype correlations in the dystrophic subtypes. *J Med Genet* **44**, 181–192 (2007).
272. Nagai, Y. *et al.* Aldosterone stimulates collagen gene expression and synthesis via activation of ERK1/2 in rat renal fibroblasts. *Hypertension* **46**, 1039–1045 (2005).
273. Gillery, P. *et al.* Interleukin-4 stimulates collagen gene expression in human fibroblast monolayer cultures Potential role in fibrosis. *FEBS Letters* **302**, 231–234 (1992).
274. Aoudjehane, L. *et al.* Interleukin-4 induces the activation and collagen production of cultured human intrahepatic fibroblasts via the STAT-6 pathway. *Lab Invest* **88**, 973–985 (2008).
275. Büttner, C., Skupin, A. & Rieber, E. P. Transcriptional activation of the type I collagen genes COL1A1 and COL1A2 in fibroblasts by interleukin-4: Analysis of the functional collagen promoter sequences. *J Cell Physiol* **198**, 248–258 (2004).
276. Boyle, M. *et al.* Performance of the PRO-C3 collagen neo-epitope biomarker in non-alcoholic fatty liver disease. *JHEP Rep* **1**, 188–198 (2019).
277. Organ, L. A. *et al.* Biomarkers of collagen synthesis predict progression in the PROFILE idiopathic pulmonary fibrosis cohort. *Respir Res* **20**, 148–148 (2019).
278. Karsdal, M. A. *et al.* Collagen biology and non-invasive biomarkers of liver fibrosis. *Liver Int* **40**, 736–750 (2020).
279. Geyer, P. E., Holdt, L. M., Teupser, D. & Mann, M. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol* **13**, 942–942 (2017).
280. Lee, P. Y., Osman, J., Low, T. Y. & Jamal, R. Plasma/serum proteomics: depletion strategies for reducing high-abundance proteins for biomarker discovery. *Bioanalysis* **11**, 1799–1812 (2019).
281. Cao, X. *et al.* Evaluation of spin columns for human plasma depletion to facilitate MS-based proteomics analysis of plasma. *J Proteome Res* **20**, 4610–4620 (2021).

## Bibliography

282. Canfield, A. E., Hadfield, K. D., Rock, C. F., Wylie, E. C. & Wilkinson, F. L. Htra1: a novel regulator of physiological and pathological matrix mineralization? *Biochem Soc Trans* **35**, 669–671 (2007).
283. An, E., Sen, S., Park, S. K., Gordish-Dressman, H. & Hathout, Y. Identification of novel substrates for the serine protease HTRA1 in the human RPE secretome. *Invest Ophthalmol Vis Sci* **51**, 3379–3386 (2010).
284. Lagoutte, P., Bettler, E., Vadon-Le Goff, S. & Moali, C. Procollagen C-proteinase enhancer-1 (PCPE-1), a potential biomarker and therapeutic target for fibrosis. *Matrix Biol Plus* **11**, (2021).

# APPENDICES

**Appendix 1. Composition and distribution of collagens in human tissues.** Collagen is composed three identic (homotrimer) or different (heterotrimer)  $\alpha$  chains, which are associated at least once into a triple helical domain. The trimers adopt different supramolecular assemblies: beaded filaments, fibrils and anchoring fibrils, fibril-associated collagens with interrupted triple helices (FACITs), multiplexins and networks can be found in the ECM of connective tissues especially. Some collagens are finally transmembrane.

Collagen	Number of $\alpha$ -chains	Trimer composition in physiology	Supramolecular assembly	Distribution in tissues
Type I	2	$[\alpha 1(I)]_3$ $[\alpha 1(I)]_2[\alpha 2(I)]$	Fibrils	Bone, cornea, dermis, tendon
Type II	1	$[\alpha 1(II)]_3$	Fibrils	Cartilage
Type III	1	$[\alpha 1(III)]_3$	Fibrils	Blood vessels, dermis
type IV	6	$[\alpha 1(IV)]_2[\alpha 2(IV)]$ $[\alpha 3(IV)][\alpha 4(IV)][\alpha 5(IV)]$ $[\alpha 5(IV)]_2[\alpha 6(IV)]$	Network	Basement membrane
Type V	4	$[\alpha 1(V)]_2[\alpha 2(V)]$ $[\alpha 1(V)][\alpha 2(V)][\alpha 3(V)]$ $[\alpha 1(V)]_3$	Fibrils	Bone, cornea, dermis, placenta
Type VI	4	$[\alpha 1(VI)][\alpha 2(VI)][\alpha 3(VI)]$	Beaded filaments	Bone, dermis, muscle
Type VII	1	$[\alpha 1(VII)]_3$	Anchoring fibrils	Basement membrane (including dermalepidermal junction)
Type VIII	2	$[\alpha 1(VIII)]_2[\alpha 2(VIII)]$ $[\alpha 1(VIII)]_3$ $[\alpha 2(VIII)]_3$	Network	Cornea
Type IX	3	$[\alpha 1(IX)][\alpha 2(IX)][\alpha 3(IX)]$	FACIT	Cartilage
Type X	2	$[\alpha 1(X)]_3$	Network	Cartilage
Type XI	3	$[\alpha 1(XI)][\alpha 2(XI)][\alpha 3(XI)]$	Fibrils	Cartilage
Type XII	2	$[\alpha 1(XII)]_3$	FACIT	Dermis, ligament, tendon
Type XIII	2	$[\alpha 1(XIII)]_3$	Transmembrane	Neuromuscular junctions
Type XIV	1	$[\alpha 1(XIV)]_3$	FACIT	Dermis, bone, cartilage
Type XV	1	$[\alpha 1(XV)]_3$	Multiplexin	Basement membrane
Type XVI	1	$[\alpha 1(XVI)]_3$	FACIT	Dermis, kidney
Type XVII	1	$[\alpha 1(XVII)]_3$	Transmembrane	Hemidesmosomes
Type XVIII	1	$[\alpha 1(XVIII)]_3$	Multiplexin	Basement membrane, liver, retina
Type XIX	1	$[\alpha 1(XIX)]_3$	FACIT	Basement membrane
Type XX	1	$[\alpha 1(XX)]_3$	FACIT	Cornea
Type XXI	1	$[\alpha 1(XXI)]_3$	FACIT	Kidney, stomach
Type XXII	1	$[\alpha 1(XXII)]_3$	FACIT	Muscle-tendon junction
Type XXIII	1	$[\alpha 1(XXIII)]_3$	Transmembrane	Epidermis
Type XXIV	1	$[\alpha 1(XXIV)]_3$	Fibrils	Bone, cornea
Type XXV	1	$[\alpha 1(XXV)]_3$	Transmembrane	Brain
Type XXVI	1	$[\alpha 1(XXVI)]_3$	-	Ovary, testicle
Type XXVII	1	$[\alpha 1(XXVII)]_3$	Fibrils	Cartilage
Type XXVIII	1	$[\alpha 1(XXVIII)]_3$	-	Basement membrane

**Appendix 2. Inclusion and exclusion protein domain used for the definition of the core matrisome.**  
More detailed are given in<sup>89</sup>.

CORE MATRISOME			
Inclusion domains		Exclusion domains	
InterPro accession number	InterPro domain name	InterPro accession number	InterPro domain name
IPR000020	Anaphylatoxin/fibulin	IPR000157	Toll-Interleukin receptor
IPR000034	Laminin B type IV	IPR000242	Protein tyrosine phosphatase, catalytic domain
IPR000082	SEA	IPR000276	7TM GPCR, rhodopsin-like
IPR000083	Fibronectin, type I	IPR000413	Integrin alpha chain
IPR000294	Gamma-carboxyglutamic acid-rich (GLA) domain	IPR000832	GPCR, family 2, secretin-like
IPR000372	Leucine-rich repeat, cysteine-rich flanking region, N-terminal	IPR001169	Integrin beta subunit, C-terminal
IPR000436	Sushi/SCR/CCP	IPR001245	Serine/threonine/tyrosine-protein kinase
IPR000483	Cysteine-rich flanking region, C-terminal	IPR001254	Peptidase S1/S6, chymotrypsin/Hap
IPR000538	Link	IPR001590	Peptidase M12B, ADAM/reprolysin
IPR000562	Fibronectin, type II	IPR002290	Serine/threonine-protein kinase domain
IPR000585	Hemopexin/matrixin	IPR002919	Protease inhibitor I8, cysteine-rich trypsin inhibitor-like
IPR000716	Thyroglobulin type-1	IPR006026	Peptidase, metallopeptidases
IPR000782	FAS1_domain	IPR006586	ADAM, cysteine-rich
IPR000859	CUB	IPR010294	ADAM-TS Spacer 1
IPR000867	Insulin-like growth factor-binding protein, IGFBP	IPR011009	Protein kinase-like domain
IPR000884	Thrombospondin, type 1 repeat	IPR011162	MHC classes I/II-like antigen recognition protein
IPR000885	Fibrillar collagen, C-terminal	IPR013273	Peptidase M12B, ADAM-TS
IPR001007	von Willebrand factor, type C Domain	IPR013642	Chloride channel calcium-activated
IPR001212	Somatomedin B	IPR013680	Voltage-dependent calcium channel, alpha-2/delta subunit, conserved region
IPR001304	C-type lectin	IPR020635	Tyrosine-protein kinase, catalytic domain
IPR001442	Type 4 procollagen, C-terminal repeat		
IPR001507	Endoglin/CD105 antigen		
IPR001791	Laminin G		
IPR001846	von Willebrand factor, type D		
IPR001881	EGF-like calcium-binding		
IPR002035	von Willebrand factor, type A		
IPR002038	Osteopontin		
IPR002049	EGF-like, laminin		
IPR002181	Fibrinogen, alpha/beta/gamma chain, C-terminal globular		
IPR002212	Matrix fibril-associated		
IPR002861	Reeler region		
IPR003129	Laminin G, thrombospondin-type, N-terminal		
IPR003591	Leucine-rich repeat, typical subtype		
IPR003645	Follistatin-like, N-terminal		
IPR003886	Nidogen, extracellular region		
IPR003961	Fibronectin, type III		
IPR003979	Tropoelastin		
IPR004043	LCCL		
IPR004116	Amelogenin		
IPR004850	Agrin NtA		
IPR006210	EGF		
IPR006605	G2 nidogen and fibulin G2F		
IPR007455	Serglycin		
IPR007798	Amelin		
IPR008160	Collagen triple helix repeat		
IPR008211	Laminin, N-terminal		
IPR008412	Bone sialoprotein II		
IPR008673	Microfibril-associated glycoprotein		
IPR008859	Thrombospondin, C-terminal		
IPR009254	Laminin I		
IPR009730	Micro-fibrillar-associated 1, C-terminal		
IPR009837	Osteoregulin		
IPR009889	Dentin matrix 1		
IPR011489	EMI		
IPR016352	Small leucine-rich proteoglycan, class I, decorin/asperin/byglycan		

**Appendix 3. Inclusion and exclusion protein domain used for the definition of the associated matrisome.** More detailed are given in<sup>89</sup>.

ASSOCIATED MATRISOME			
Inclusion domains		Exclusion domains	
InterPro accession number	InterPro domain name	InterPro accession number	InterPro domain name
<i>ECM-affiliated</i>			
IPR001050	Syndecan		
IPR001079	Galectin, carbohydrate recognition domain		
IPR001627	Semaphorin/CD100 antigen		
IPR001863	Glypican		
IPR003659	Plexin/semaphorin/integrin		
IPR018502	Annexin		
<i>ECM regulator</i>			
IPR000001	Kringle	IPR000157	Toll - interleukin 1 - resistance
IPR000010	Proteinase inhibitor I25, cystatin	IPR000242	Protein tyrosine phosphatase, catalytic domain
IPR000215	Protease inhibitor I4, serpin	IPR000436	Sushi/SCR/CCP
IPR000585	Hemopexin/matrixin	IPR000832	GPCR, family 2, secretin-like
IPR001506	Peptidase M12A, astacin	IPR001073	Complement C1q protein
IPR001590	Peptidase M12B, ADAM/reprolysin	IPR001245	Serine/threonine/tyrosine-protein kinase
IPR001695	Lysyl oxidase	IPR012317	Poly(ADP-ribose) polymerase, catalytic domain
IPR001762	Blood coagulation inhibitor, Disintegrin	IPR013642	Chloride channel calcium-activated
IPR001818	Peptidase M10, metallopeptidase	IPR017452	GPCR, rhodopsin-like superfamily
IPR002477	Peptidoglycan binding-like	IPR018933	Netrin module, non-TIMP type
IPR002931	Transglutaminase-like	IPR020068	Frizzled domain, subgroup
IPR005199	Glycoside hydrolase family 79, N-terminal	IPR020635	Tyrosine-protein kinase, catalytic domain
IPR006026	Peptidase, metallopeptidases		
IPR006586	ADAM, cysteine-rich		
IPR006587	Vault protein inter-alpha-trypsin, metazoa		
IPR008993	TIMP-like, OB-fold		
IPR010294	ADAM-TS Spacer 1		
IPR013128	Peptidase C1A, papain		
IPR013273	Peptidase M12B, ADAM-TS		
IPR014615	Extracellular Sulfatase		
IPR015446	Bone morphogenetic protein 1/tolloid-like protein		
IPR015643	Peptidase C1A, cathepsin B		
IPR015644	Peptidase C1A, cathepsin K		
IPR018155	Hyaluronidase		
IPR019742	Alpha-2-macroglobulin, conserved site		
IPR019828	Lysyl oxidase, conserved site		
IPR021158	Peptidase M10A, cysteine switch, zinc binding site		

## Appendices

Inclusion domains		Exclusion domains	
InterPro accession number	InterPro domain name	InterPro accession number	InterPro domain name
<i>Secreted factors</i>			
IPR000001	Kringle	IPR000242	Protein tyrosine phosphatase, catalytic domain
IPR000072	Platelet-derived growth factor (PDGF)	IPR000294	Gamma-carboxyglutamic acid-rich (GLA) domain
IPR000226	Interleukin-7/interleukin-9	IPR000436	Sushi/SCR/CCP
IPR000471	Interferon alpha/beta/delta	IPR001073	Complement C1q protein
IPR000762	Midkine heparin-binding growth factor	IPR001245	Serine/threonine/tyrosine-protein kinase
IPR000779	Interleukin-2	IPR001762	Blood coagulation inhibitor, Disintegrin
IPR000975	Interleukin-1	IPR001968	Glycoside hydrolase, family 56
IPR001007	von Willebrand factor, type C	IPR002007	Haem peroxidase, animal
IPR001811	Small chemokine, interleukin-8-like	IPR002035	von Willebrand factor, type A
IPR001839	Transforming growth factor beta	IPR002350	Proteinase inhibitor II, Kazal
IPR001881	EGF-like calcium-binding	IPR002369	Integrin beta subunit, N-terminal
IPR002072	Nerve growth factor	IPR006586	ADAM, cysteine-rich
IPR002154	Neuregulin 1-related, C-terminal	IPR017978	GPCR, family 3, C-terminal
IPR002181	Fibrinogen, alpha/beta/gamma chain, C-terminal globular	IPR017981	GPCR, family 2-like
IPR002348	Interleukin 1/heparin-binding growth factor	IPR018358	Disintegrin, conserved site
IPR002960	Tumour necrosis factor beta/lymphotoxin alpha	IPR018933	Netrin module, non-TIMP type
IPR002961	Tumour necrosis factor c/lymphotoxin-beta	IPR020635	Tyrosine-protein kinase, catalytic domain
IPR003263	Tumour necrosis factor ligand 5		
IPR003296	Interleukin-1, beta IL1B		
IPR003306	WIF domain		
IPR003587	Hedgehog/intein hint, N-terminal		
IPR003645	Follistatin-like, N-terminal		
IPR004825	Insulin/IGF/relaxin		
IPR005817	Wnt superfamily		
IPR006210	EGF-like		
IPR008983	Tumour necrosis factor-like		
IPR009079	Four-helical cytokine-like, core		
IPR010345	Interleukin-17		
IPR010510	FGF binding 1		
IPR010895	CHRD		
IPR013787	S100/CaBP-9k-type, calcium binding, subdomain		
IPR015528	Interleukin-12 beta chain		
IPR015529	Interleukin 18		
IPR015615	Transforming growth factor-beta-related		
IPR020412	Interleukin-11, mammalian		
IPR020415	Interleukin-34		
IPR020439	Interleukin-15, C-terminal		
IPR020443	Interleukin-10/19/20/24		
IPR020450	Interleukin-16		

#### Appendix 4. Script of the macro used to analyze the diameter of collagen fibrils in the dermis of WT and RDEB mouse skin.

```

////////////////////////////////////
/*
Date : 10/02/2015
Author : Denis Ressnikoff
version : 1.0
Macro used to measure the geodesic diameter of collagen fibrils on TEM pictures of tendons (fixed
calibration (1px = 1.0041 nm)
Necessary Plugin(s):
*ij_Geodesics (https://www.pfl-cepia.inra.fr/index.php?page=ijGeodesics\_en)
*/
////////////////////////////////////

//initialisation
run("Set Measurements...", "area limit display redirect=None decimal=2");
run("Colors...", "foreground=black background=white selection=green");

if (isOpen("ROI Manager")) {
selectWindow("ROI Manager");
run("Close");
}

if (isOpen("Log")) {
selectWindow("Log");
run("Close");
}

if (isOpen("Results")) {
selectWindow("Results");
run("Close");
}

rep = getDirectory("image");
img = getTitle();

//calibration
run("Set Scale...", "distance=1.0041 known=1 pixel=1 unit=nm global");
run("Duplicate...", "title=count");

//post-treatment filtering
run("Gaussian Blur...", "sigma=5");
//segmentation
setAutoThreshold("Huang");
setOption("BlackBackground", false);
run("Convert to Mask");
run("Watershed");
//run("Options...", "iterations=5 count=4 edm=Overwrite do=Open");
run("Fill Holes");

//analyse particules
run("Analyze Particles...", "size=750-Infinity circularity=0.4-1.00 show=[Count Masks] exclude clear
add");

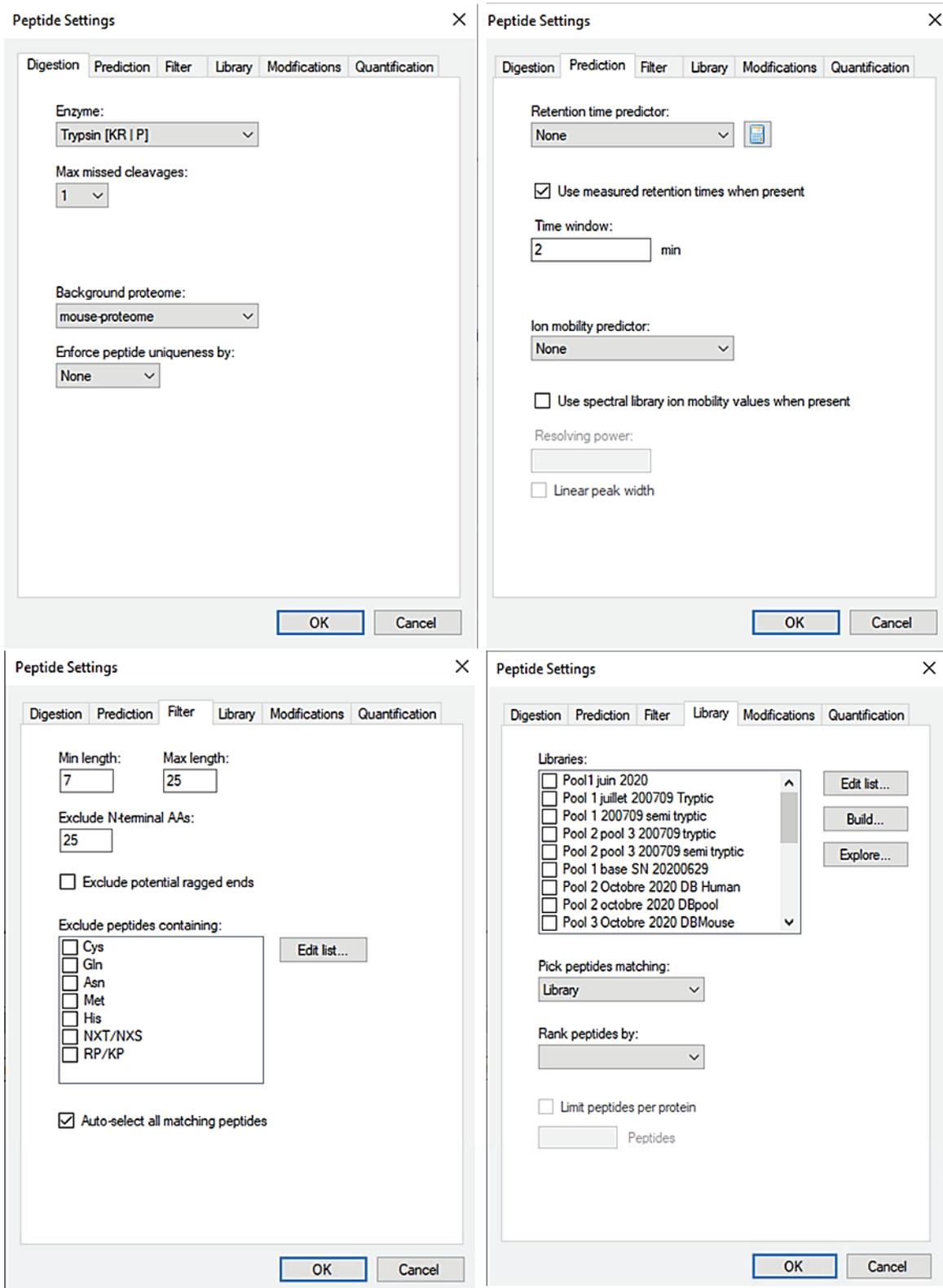
```

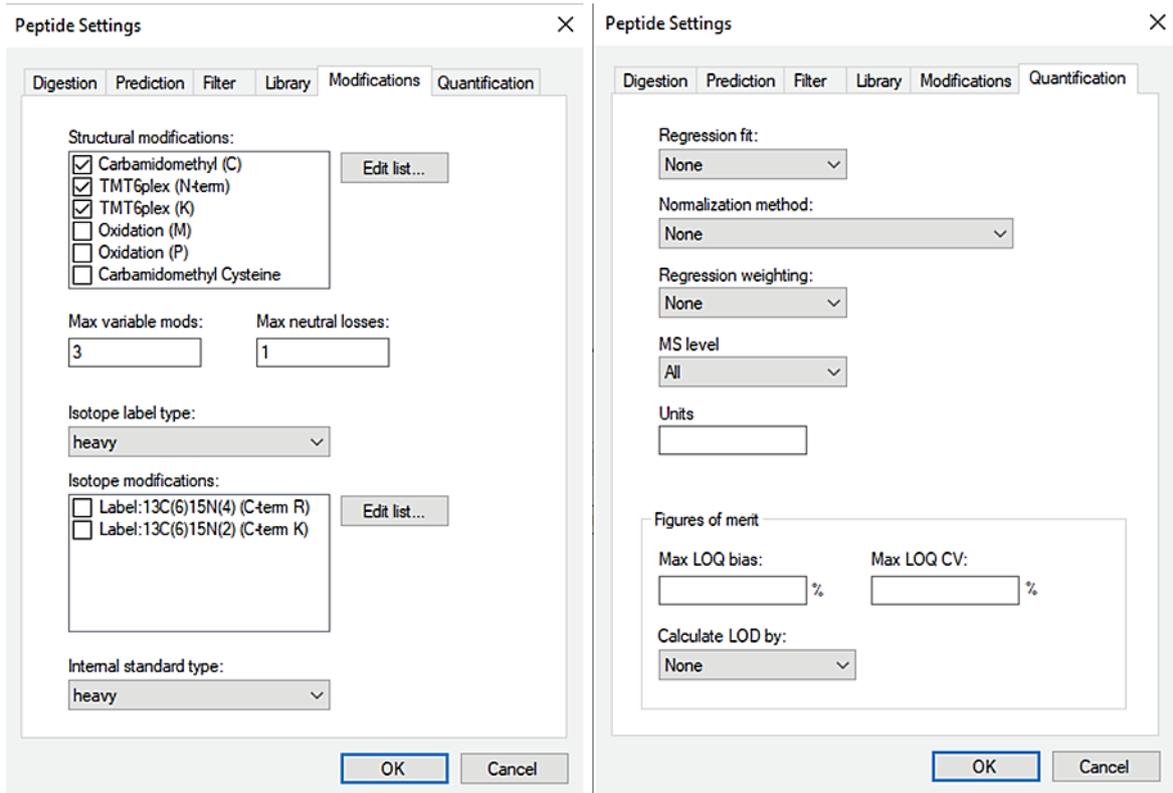
```
//diameter
run("Geodesic Diameter", "label=[Count Masks of count] distances=[Borgefors (3,4)] show
image=["+img+"]");

//saving
saveAs("Results", rep+img+".xls");
close("Count Masks of count");
close("count");

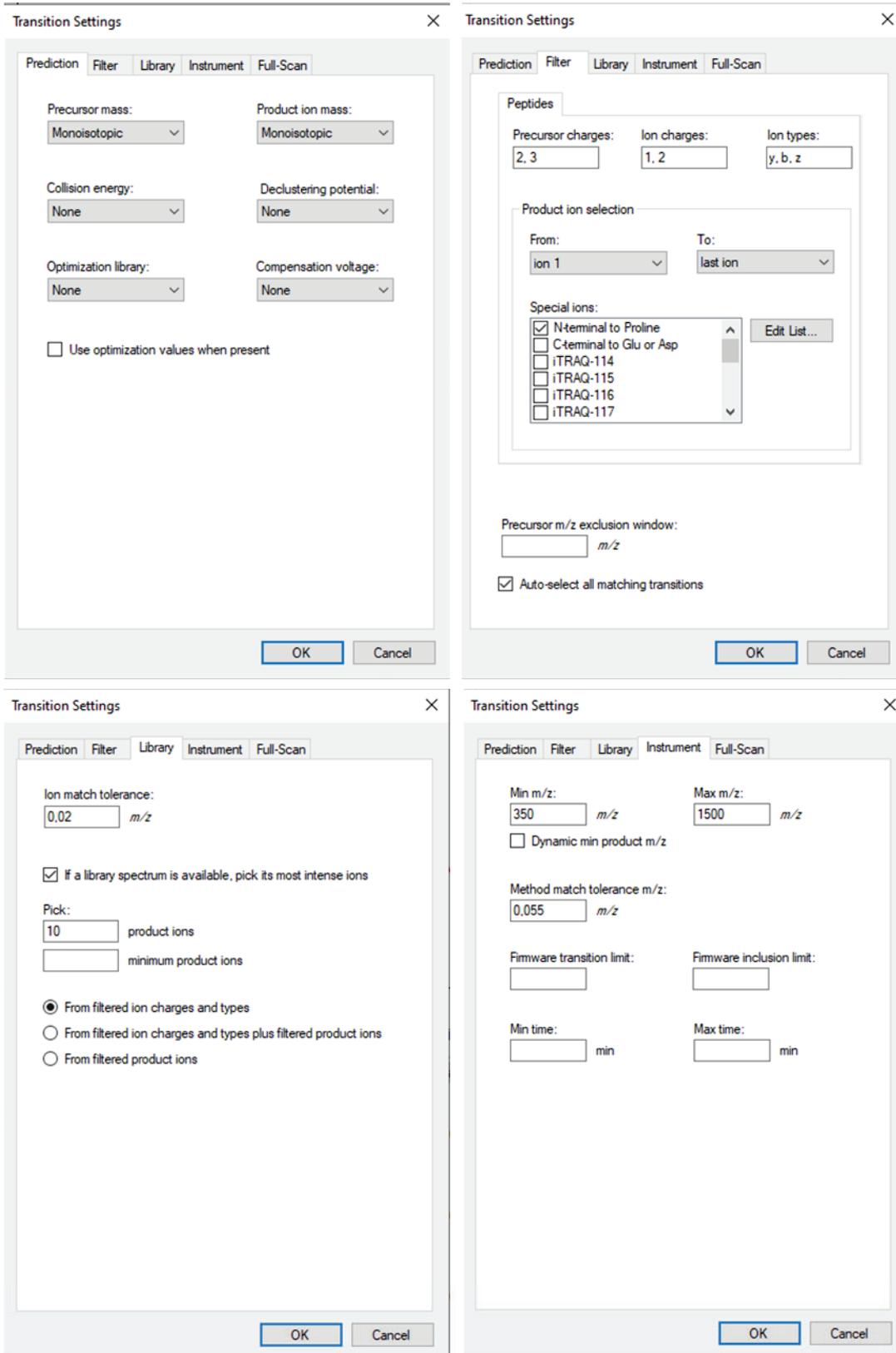
run("Flatten");
rename(img+" mesure");
saveAs("Jpeg", rep+img+" mesure.jpg");
close(img);
//close(img+" mesure.jpg");
run("Tile");
```

**Appendix 5. Peptide settings applied in Skyline software for PRM analyses.** Library settings: different libraries were generated by DDA, as described in Chapter II, VII-2, depending on the nature of the sample analyzed.





**Appendix 6. Transition settings applied in Skyline software for PRM analyses.**



Transition Settings ×

Prediction | Filter | Library | Instrument | Full-Scan

**MS1 filtering**

Isotope peaks included: Precursor mass analyzer:  
 None None

Peaks: Resolution:  
  m/z

Isotope labeling enrichment:

**MS/MS filtering**

Acquisition method: Product mass analyzer:  
 Targeted Orbitrap

Isolation scheme: Resolving power: At:  
 60 000  200 m/z

Use high-selectivity extraction

**Retention time filtering**

Use only scans within  minutes of MS/MS IDs

Use only scans within  minutes of predicted RT

**Include all matching scans**

**Appendix 7. Detailed content in core matrisome proteins in all the fractions analyzed by MS/MS.** Two ECM-enrichment experiments coupled with MS-based proteomics (n= 2 and n= 3) were conducted to compare the back skin matrisome of WT and RDEB mice. Whatever the fraction focused, most of the proteins of the matrix were detected in the two analyses.

CORE MATRISOME									
Gene	Protein	NaCl		SDS/SD		GuHCl + Ins		Comp. Ins	
		n= 2	n= 3	n= 2	n= 3	n= 2	n= 3	n= 2	n= 3
<b>Collagen</b>									
<i>Col1a1</i>	Collagen alpha-1 (I)	X	X	X	X	X	X	X	X
<i>Col1a2</i>	Collagen alpha-2 (I)	X	X	X	X	X	X	X	X
<i>Col2a1</i>	Collagen alpha-1 (II)					X	X	X	X
<i>Col3a1</i>	Collagen alpha-1 (III)	X	X	X	X	X	X	X	X
<i>Col4a1</i>	Collagen alpha-1 (IV)						X		
<i>Col4a2</i>	Collagen alpha-2 (IV)					X	X	X	X
<i>Col4a3</i>	Collagen alpha-3 (IV)					X		X	
<i>Col4a5</i>	Collagen alpha-5 (IV)					X	X	X	X
<i>Col5a1</i>	Collagen alpha-1 (V)			X		X	X	X	X
<i>Col5a2</i>	Collagen alpha-2 (V)					X	X	X	X
<i>Col5a3</i>	Collagen alpha-3 (V)						X		X
<i>Col6a1</i>	Collagen alpha-1 (VI)		X	X	X	X	X	X	X
<i>Col6a2</i>	Collagen alpha-2 (VI)			X	X	X	X	X	X
<i>Col6a3</i>	Collagen alpha-3 (VI)	X	X	X	X	X	X	X	X
<i>Col6a5</i>	Collagen alpha-5 (VI)			X	X				
<i>Col6a6</i>	Collagen alpha-6 (VI)		X	X	X	X	X	X	X
<i>Col7a1</i>	Collagen alpha-1 (VII)				X	X	X	X	X
<i>Col8a1</i>	Collagen alpha-1 (VIII)						X		X
<i>Col10a1</i>	Collagen alpha-1 (X)						X		
<i>Col12a1</i>	Collagen alpha-1 (XII)		X	X	X	X	X	X	X
<i>Col14a1</i>	Collagen alpha-1 (XIV)	X	X	X	X	X	X	X	X
<i>Col15a1</i>	Collagen alpha-1 (XV)			X	X	X	X	X	X
<i>Col17a1</i>	Collagen alpha-1 (XVII)						X		X
<i>Col18a1</i>	Collagen alpha-1 (XVIII)		X		X				
<i>Col27a1</i>	Collagen alpha-1 (XXVII)					X		X	
<i>Col28a1</i>	Collagen alpha-1 (XXVIII)					X	X	X	
<b>Glycoprotein</b>									
<i>Abi3bp</i>	ABI gene family, member 3 (NESH) binding protein				X	X		X	
<i>Adipoq</i>	Adiponectin	X	X	X	X				
<i>Aebp1</i>	Adipocyte enhancer-binding protein-1	X	X	X	X	X		X	X
<i>Agrn</i>	Aggrin			X	X				
<i>Cd109</i>	CD109 antigen								
<i>Cilp</i>	Cartilage intermediate layer protein			X					X
<i>Coch</i>	Cochlin		X		X				
<i>Cthrc1</i>	Collagen triple helix repeat-containing protein-1		X						
<i>Dpt</i>	Dermatopontin	X		X	X	X	X	X	X
<i>Ecm1</i>	Extracellular matrix protein 1				X				
<i>Ecm2</i>	Extracellular matrix protein 2			X		X		X	
<i>Efemp1</i>	EGF-containing fibulin-like extracellular matrix protein 1					X	X	X	
<i>Eln</i>	Elastin					X	X	X	X
<i>Emilin1</i>	EMILIN-1					X	X	X	X
<i>Fbln5</i>	Fibulin-5		X		X	X	X	X	X
<i>Fbn1</i>	Fibrillin-1					X	X	X	X
<i>Fga</i>	Fibrinogen alpha chain	X	X	X	X	X	X	X	X
<i>Fgb</i>	Fibrinogen alpha chain	X	X	X	X	X	X	X	X

Appendices

<i>Fgg</i>	Fibrinogen alpha chain	X	X	X	X	X	X	X	X
<i>Fn1</i>	Fibronectin	X		X	X	X	X	X	X
<i>Hmcn2</i>	Hemicentin-2				X				
<i>Igfbp7</i>	Insulin-like growth factor binding protein 7		X						
<i>Lama2</i>	Laminin subunit alpha-2			X	X	X	X	X	X
<i>Lama3</i>	Laminin subunit alpha-3				X	X		X	X
<i>Lama4</i>	Laminin subunit alpha-4		X	X	X	X	X		
<i>Lama5</i>	Laminin subunit alpha-5			X	X	X	X	X	X
<i>Lamb1</i>	Laminin subunit beta-1			X	X	X	X	X	X
<i>Lamb2</i>	Laminin subunit beta-2			X	X	X	X	X	X
<i>Lamb3</i>	Laminin subunit beta-3					X		X	
<i>Lamc1</i>	Laminin subunit gamma-1			X	X	X	X	X	X
<i>Lamc2</i>	Laminin subunit gamma-2					X	X	X	
<i>Lrg1</i>	Leucine-rich alpha-2-glycoprotein	X	X	X	X				
<i>Mfap2</i>	Microfibrillar-associated protein-2					X	X	X	X
<i>Mfap4</i>	Microfibrillar-associated protein-4			X	X	X	X	X	
<i>Mfap5</i>	Microfibrillar-associated protein-5					X		X	
<i>Nid1</i>	Nidogen-1	X	X	X	X	X	X	X	X
<i>Nid2</i>	Nidogen-2			X	X	X		X	
<i>Postn</i>	Periostin	X	X	X	X	X	X	X	X
<i>Sprac</i>	Basement-membrane protein-40		X						
<i>Srpx</i>	Sushi repeat-containing protein SRPX			X	X	X	X	X	X
<i>Tgfb1</i>	Transforming growth factor-beta-induced protein ig-h3		X	X	X	X	X	X	X
<i>Thbs1</i>	Thrombospondin-1		X		X				
<i>Tnc</i>	Tenascin-C				X				
<i>Tnxb</i>	Tenascin-X	X	X	X	X	X	X	X	X
<i>Vtn</i>	Vitronectin			X		X	X	X	
<i>Vwa1</i>	Von Willebrand factor A domain-containing protein-1					X	X	X	X
<i>Vwa5a</i>	Von Willebrand factor A domain-containing protein-5A	X	X	X	X				
<b>Proteoglycans</b>									
<i>Aspn</i>	Asporin			X	X	X	X	X	X
<i>Bgn</i>	Biglycan			X	X	X	X	X	X
<i>Cspg4</i>	Chondroitin sulfate proteoglycan 4				X				
<i>Dcn</i>	Decorin	X	X	X	X	X	X	X	X
<i>Fmod</i>	Fibromodulin		X	X			X		
<i>Hspg2</i>	Perlecan	X		X	X	X	X	X	X
<i>Lum</i>	Lumican	X	X	X	X	X	X	X	X
<i>Ogn</i>	Mimecan		X	X	X	X	X	X	X
<i>Podn</i>	Podocan		X		X				
<i>Prelp</i>	Prolargin	X	X	X	X	X	X	X	X
<i>Prg2</i>	Bone marrow proteoglycan					X	X	X	
<i>Vcan</i>	Versican			X					

**Appendix 8. Detailed content in associated matrisome proteins in all the fractions analyzed by MS/MS.** Proteins from the associated matrisome were mainly recovered in the first buffers (NaCl and SDS/SD) regardless of the experiment (n= 2 and n= 3).

ASSOCIATED									
Gene	Protein	NaCl		SDS/SD		GuHCl + Ins		Comp. Ins	
		n= 2	n= 3	n= 2	n= 3	n= 2	n= 3	n= 2	n= 3
<b>ECM-affiliated</b>									
<i>Anxa1</i>	Annexin A1	X	X	X	X	X	X	X	X
<i>Anxa2</i>	Annexin A2	X	X	X	X	X	X	X	X
<i>Anxa3</i>	Annexin A3	X	X	X	X				
<i>Anxa4</i>	Annexin A4	X	X	X	X		X		
<i>Anxa5</i>	Annexin A5	X	X	X	X	X	X	X	X
<i>Anxa6</i>	Annexin A6	X	X	X	X		X		
<i>Anxa7</i>	Annexin A7	X	X	X	X	X	X	X	
<i>Anxa8</i>	Annexin A8	X	X	X	X	X	X	X	X
<i>Anxa11</i>	Annexin A11	X	X	X	X				
<i>C1qa</i>	Complement C1q subcomponent subunit A	X							
<i>C1qtnf3</i>	Complement C1q tumor necrosis factor-related protein-3	X	X						
<i>Cd209b</i>	CD209 antigen-like protein B			X	X				
<i>Cd209d</i>	CD209 antigen-like protein D			X					
<i>Clec10a</i>	C-type lectin domain family-10 member A					X			
<i>Clec11a</i>	C-type lectin domain family-11 member A			X			X	X	
<i>Clec3b</i>	Tetranectin	X		X					
<i>Colec12</i>	Collectin-12				X				
<i>Gpc1</i>	Glypican-1				X				
<i>Hpx</i>	Hemopexin	X	X	X	X		X		
<i>Lgals1</i>	Galectin-1	X	X	X	X	X	X	X	X
<i>Lgals3</i>	Galectin-3	X	X	X	X	X	X	X	X
<i>Lgals7</i>	Galectin-7	X	X	X	X	X	X	X	X
<i>Lgalsl</i>	Galectin-related protein	X	X						
<i>Lman1</i>	Protein ERGIC-53			X					
<i>Plxna4</i>	Plexin-A4	X							
<b>ECM regulator</b>									
<i>Adamts15</i>	ADAMTS-like protein-5				X				
<i>Cd109</i>	Cd109 antigen		X	X	X		X		X
<i>Cst3</i>	Cystatin-3	X	X	X					
<i>Cst6</i>	Cystatin-6	X	X	X	X	X	X	X	X
<i>Csta</i>	Cystatin-A	X	X				X		X
<i>Cstb</i>	Cystatin-B		X						
<i>Ctsa</i>	Cathepsin A	X	X		X				
<i>Ctsb</i>	Cathepsin B	X		X	X				
<i>Ctsd</i>	Cathepsin D		X	X	X		X		X
<i>Ctsh</i>	Cathepsin H	X	X		X				
<i>Ctsk</i>	Cathepsin K								
<i>Ctsz</i>	Cathepsin Z		X						
<i>F2</i>	Prothrombin	X	X						
<i>F13a1</i>	Coagulation factor XIII A chain	X	X	X	X	X	X	X	X
<i>Gm5409</i>	Try10-like trypsinogen		X	X			X		X
<i>Hrg</i>	Histidine-rich glycoprotein	X	X	X	X				
<i>Htra1</i>	Serine protease HTRA-1				X				
<i>Itih1</i>	Inter-alpha-trypsin inhibitor heavy chain H1	X			X				
<i>Itih2</i>	Inter-alpha-trypsin inhibitor heavy chain H2	X	X	X					
<i>Itih3</i>	Inter-alpha-trypsin inhibitor heavy chain H3	X	X	X	X	X	X	X	
<i>Itih4</i>	Inter-alpha-trypsin inhibitor heavy chain H4	X	X						
<i>Kng1</i>	Kininogen-1		X	X	X				
<i>P4hal</i>	Prolyl 4-hydroxylase subunit alpha-1	X		X					

Appendices

<i>Plg</i>	Plasminogen		X		X				
<i>Plod3</i>	Lysyl hydroxylase-3	X		X					
<i>Pzp</i>	Pregnancy zone protein	X	X	X	X				
<i>Serpina1a</i>	Serine protease inhibitor A1a	X	X	X	X		X		
<i>Serpina1b</i>	Serine protease inhibitor A1b	X	X	X	X		X		
<i>Serpina1d</i>	Serine protease inhibitor A1d	X	X	X	X				
<i>Serpina1e</i>	Serine protease inhibitor A1e	X	X	X	X		X		
<i>Serpina3b</i>	Serine protease inhibitor A3b	X			X				
<i>Serpina3k</i>	Serine protease inhibitor A3k	X	X	X	X	X	X	X	X
<i>Serpina3m</i>	Serine protease inhibitor A3m			X					
<i>Serpina3n</i>	Serine protease inhibitor A3n	X	X		X		X		
<i>Serpina6</i>	Serine protease inhibitor A6	X	X						
<i>Serpinb1a</i>	Serine protease inhibitor B1a	X	X	X	X		X		X
<i>Serpinb2</i>	Serine protease inhibitor B2	X	X	X		X	X	X	
<i>Serpinb5</i>	Serine protease inhibitor B5	X	X	X	X		X		X
<i>Serpinb6a</i>	Serine protease inhibitor B6a	X	X	X	X		X		
<i>Serpinb6c</i>	Serine protease inhibitor B6c	X	X	X					
<i>Serpinb6d</i>	Serine protease inhibitor B6d								
<i>Serpinb7</i>	Serine protease inhibitor B7				X				
<i>Serpinb8</i>	Serine protease inhibitor B8	X	X	X					
<i>Serpinb10</i>	Serine protease inhibitor B10	X	X						
<i>Serpinc1</i>	Serine protease inhibitor C1	X	X	X	X		X		
<i>Serpind1</i>	Serine protease inhibitor D1	X	X		X				
<i>Serpinf1</i>	Serine protease inhibitor F1	X	X	X	X				
<i>Serpinf2</i>	Serine protease inhibitor F2	X	X	X					
<i>Serping1</i>	Serine protease inhibitor G1	X	X	X	X				
<i>Serpinh1</i>	Serine protease inhibitor H1		X	X	X	X	X	X	X
<i>Stfa3</i>	Stefin-A3		X						
<i>Tgm1</i>	Transglutaminase-1	X			X	X	X	X	X
<i>Tgm2</i>	Transglutaminase-2		X	X	X	X	X	X	X
<i>Tgm3</i>	Transglutaminase-3	X				X		X	
<i>Try5</i>	Trypsin-5	X		X		X	X	X	X

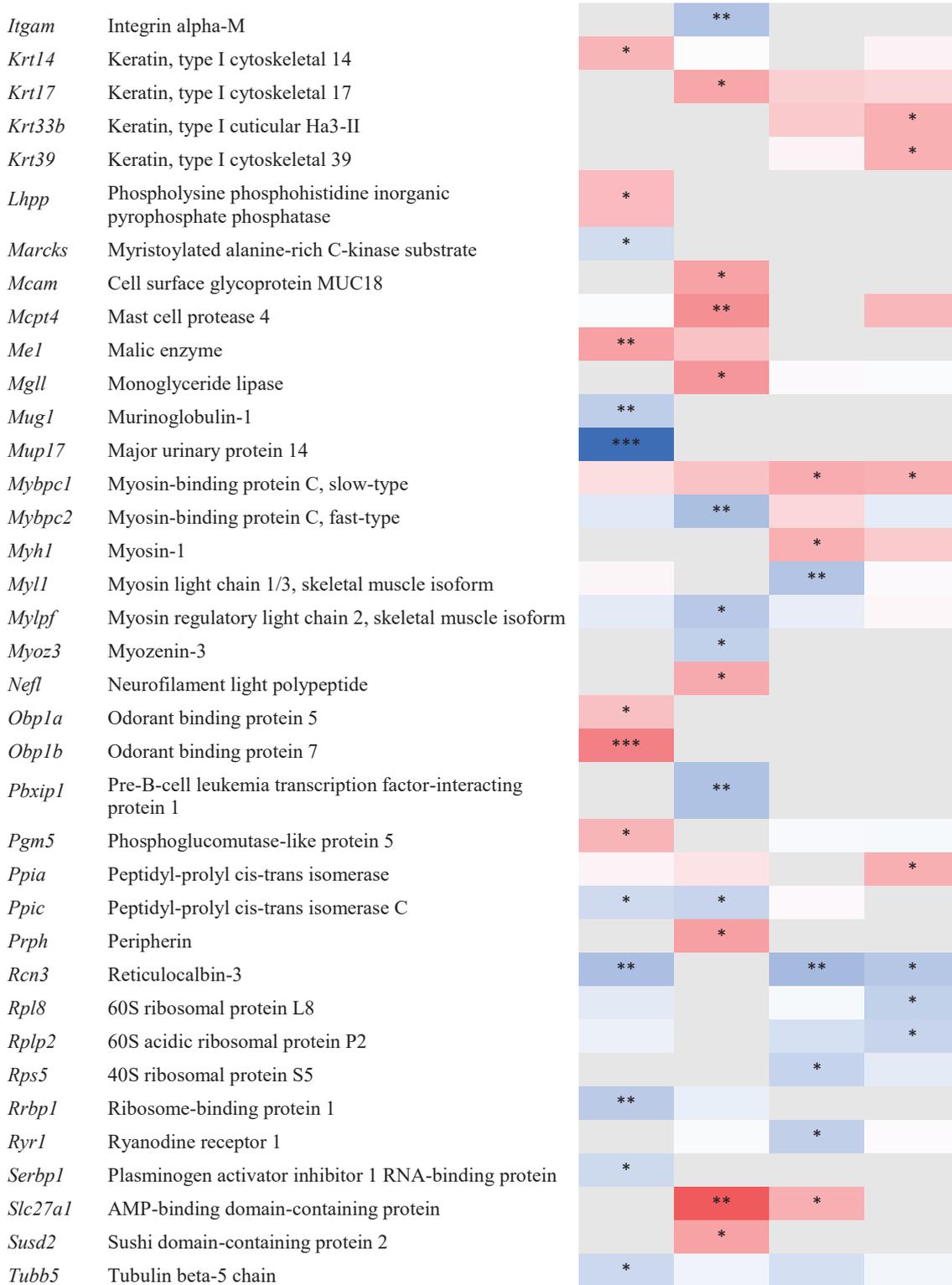
**Secreted factor**

<i>Flg</i>	Filaggrin	X	X		X				
<i>Flg2</i>	Filaggrin-2	X	X		X	X	X	X	X
<i>Habp2</i>	Hyaluronan-binding protein 2								
<i>Hcfc1</i>	Host cell factor-1				X				
<i>Hrn</i>	Hornerin		X						
<i>Il1rn</i>	Interleukin-1 receptor antagonist protein		X						
<i>M6pr</i>	Cation-dependent mannose-6- phosphate receptor			X					
<i>Rptn</i>	Repetin				X				
<i>S100a1</i>	Protein S100-A1				X				
<i>S100a3</i>	Protein S100-A3	X				X	X	X	
<i>S100a4</i>	Protein S100-A4	X	X	X	X	X	X	X	X
<i>S100a6</i>	Protein S100-A6		X	X	X		X		
<i>S100a8</i>	Protein S100-A8	X	X						
<i>S100a9</i>	Protein S100-A9	X	X	X	X				
<i>S100a10</i>	Protein S100-A10	X	X	X	X	X	X	X	
<i>S100a11</i>	Protein S100-A11	X	X	X	X	X		X	
<i>S100a13</i>	Protein S100-A13		X						
<i>S100a14</i>	Protein S100-A14			X	X	X	X	X	
<i>S100a16</i>	Protein S100-A16	X	X	X	X		X		
<i>S100b</i>	Protein S100-B								
<i>Tchh</i>	Trichohyalin		X		X	X	X	X	X

**Appendix 9. Heatmap of the abundance ratio of non-matrisome proteins dysregulated in the back skin of RDEB mice compared to WT mice.** For each sample, quantitative data from the two MS/MS experiments were normalized and used to compute the median abundances of proteins in WT and RDEB genotypes (n = 5 mice/genotype, mice aged of 9 weeks or more). Only proteins quantified in the two experiments were compared. The abundance ratio represented here corresponds to RDEB/WT (Log2 scale). Grey rectangles: protein not detected or not quantified in the fraction. Statistics: see Chapter II, III-2-v. \*: p < 0.05, \*\*: p < 0.01, \*\*\*: p < 0.001.

Gene	Protein	NaCl	SDS/SD	GuHCl + Ins.	Comp. Ins.
<i>Acaca</i>	Acetyl-CoA carboxylase 1	*			
<i>Acads</i>	Short-chain specific acyl-CoA dehydrogenase, mitochondrial		**		
<i>Acly</i>	ATP-citrate synthase	*			
<i>Ahsg</i>	Alpha-2-HS-glycoprotein		*		
<i>Aldh3a1</i>	Aldehyde dehydrogenase, dimeric NADP-preferring		*		
<i>Apoa2</i>	Apolipoprotein A-II	***			
<i>Apobec2</i>	C->U-editing enzyme APOBEC-2				**
<i>Arpc3</i>	Actin-related protein 2/3 complex subunit 3			**	
<i>Atp2b1</i>	Plasma membrane calcium-transporting ATPase 1		*		
<i>Btf3</i>	Transcription factor BTF3	*			
<i>Ca3</i>	Carbonic anhydrase 3			**	
<i>Calm4</i>	Calm4 protein			*	
<i>Casq1</i>	Calsequestrin-1			*	
<i>Cd34</i>	Hematopoietic progenitor cell antigen CD34		*		
<i>Ces1d</i>	Carboxylesterase 1D		*		
<i>Ckmt2</i>	Creatine kinase S-type, mitochondrial			*	
<i>Cma1</i>	Chymase		*		
<i>Cpa3</i>	Mast cell carboxypeptidase A		*		*
<i>Dusp14</i>	Dual specificity protein phosphatase 14				*
<i>Eef1b2</i>	Eukaryotic translation elongation factor 1 beta 2		*		
<i>Eno3</i>	2-phospho-D-glycerate hydro-lyase			*	
<i>Fabp4</i>	Fatty acid-binding protein, adipocyte			*	
<i>Fabp5</i>	Fatty acid binding protein 5, epidermal		*		
<i>Fasn</i>	Fatty acid synthase	**			
<i>Fkbp3</i>	Peptidyl-prolyl cis-trans isomerase FKBP3	*			
<i>Fkbp7</i>	Peptidylprolyl isomerase	***			
<i>Glud1</i>	Glutamate dehydrogenase 1, mitochondrial		***		
<i>Glul</i>	Glutamine synthetase	**			
<i>H2ac21</i>	Histone H2A type 2-B			*	
<i>Hbb-b2</i>	Beta-2-globin	***	**		
<i>Hist2h2bb</i>	Histone H2B type 2-B		*		
<i>Hmgb2</i>	High mobility group protein B2	***			
<i>Hsp90b1</i>	HATPase_c domain-containing protein	**			
<i>Ica</i>	Inhibitor of carbonic anhydrase	**			
<i>Ighg1</i>	IgG1			**	
<i>Ipo5</i>	Importin N-terminal domain-containing protein			*	

Appendices

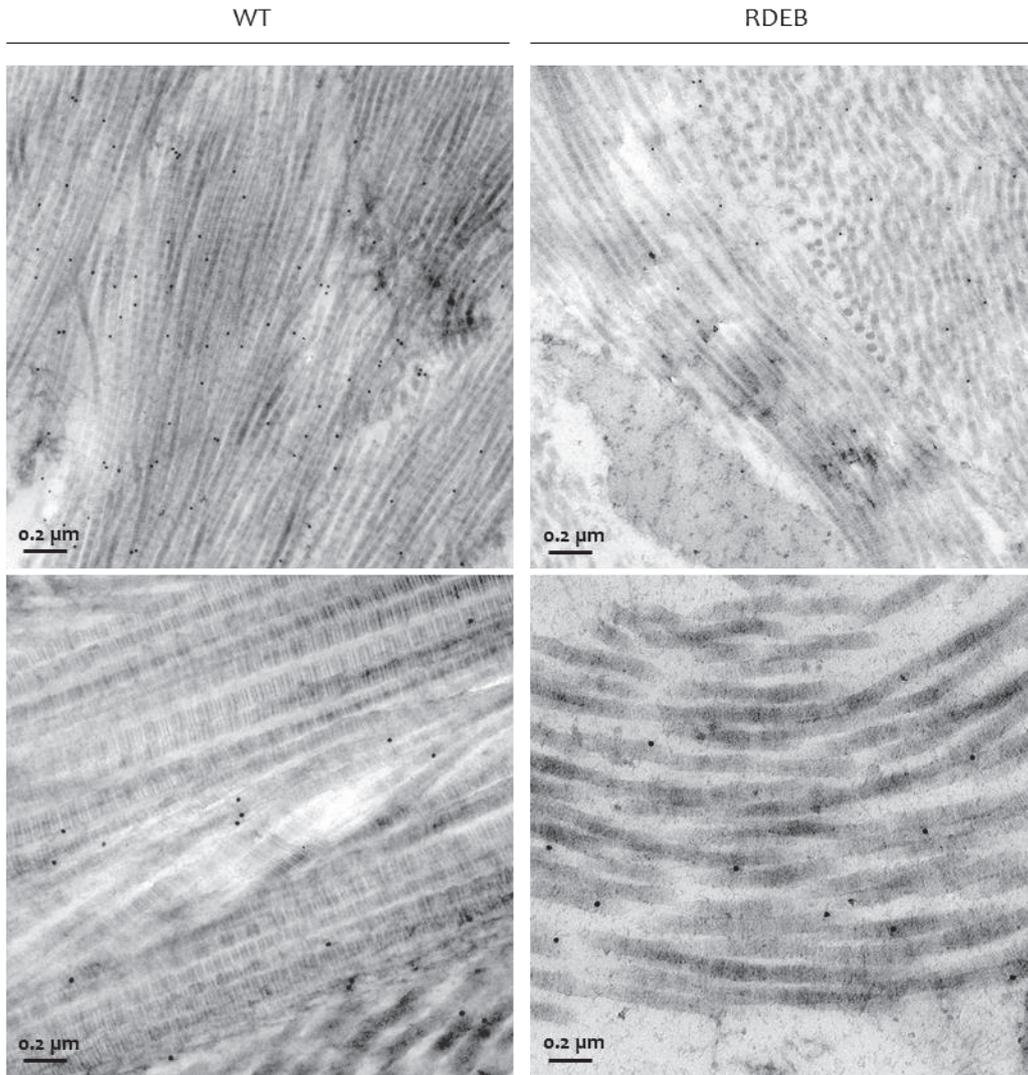


-2.7                      0                      2.7



Log<sub>2</sub> abundance ratio  
RDEB/WT

**Appendix 10. Preliminary images obtained by immunogold labeling of the C-propeptide  $\alpha 1(I)$  in the dermis of WT and RDEB mice back skin.** After processing and labeling, samples were observed by TEM. Small dense black points indicate the presence of CP  $\alpha 1(I)$  in fibers. Experimental procedure and image acquisition were conducted at Center of quantitative imaging Lyon-Est (CIQLE, SFR Santé-Lyon Est, CNRS UAR3453, INSERM US7, UCBL, Lyon, France).



## PUBLICATIONS

---

Article 1 (**Chapter III, I-2**):

Dussoyer, M., Page, A., Delolme, F., Rousselle, P., Nyström, A. & Moali, C. Comparison of extracellular matrix enrichment protocols for the improved characterization of the skin matrixome by mass spectrometry. *J Proteomics* **251**, 104397 (2022)

Article 2:

Anastasi, C., Rousselle, P., Talantikite, M., Tessier, A., Cluzel, C., Mariano, N., Dussoyer, M., Alcaraz, L.B., Fortin, L., Aubert, A., Delolme, F., El Kholti, N., Armengaud, J., Fournié, P., Auxenfants, C., Valcourt, U., Vadon-Le Goff, S. & Moali, C. BMP-1 disrupts cell adhesion and enhances TGF-beta activation through cleavage of the matricellular protein thrombospondin-1. *Sci Signal* **13**, eaba3880 (2020)

Article 3:

Vadon-Le Goff, S., Tessier, A., Dussoyer, M., Napoli, M., Lagoutte, P., Peyronnel, C., Essayan, L., Kleiser, S., Tueni, N., Bettler, E., Dieryckx, C., Mariano, N., Errazuriz-Cerda, E., Fruchart-Gaillard, C., Ruggiero, F., Becker-Pauly, C. Allain, J.-M., Bruckner-Tuderman, L., Nyström, A. & Moali, C. The dark side of procollagen C-proteinase enhancer-2 (PCPE-2): inhibition of bone morphogenetic protein-1/toiloid-like proteinases. *Proc Natl Acad Sci USA* (Submitted)

Article 4 (review):

Dussoyer, M., Michopoulou, A. & Rousselle, P. Decellularized scaffolds for skin repair and regeneration. *Appl Sci* **10**, 3435 (2020)

Book chapter:

Dussoyer, M., Moali, C. & Lagoutte, P. Potentiel de la spectrométrie de masse pour l'identification de biomarqueurs des pathologies cutanées. *Cours de Biologie de la Peau, COBIP 2022*. Ed. Marek Haftek, Edition Matrix. 7-18 (2022)

## SCIENCE SIGNALING | RESEARCH ARTICLE

## WOUND HEALING

# BMP-1 disrupts cell adhesion and enhances TGF- $\beta$ activation through cleavage of the matricellular protein thrombospondin-1

Cyril Anastasi<sup>1\*</sup>, Patricia Rousselle<sup>1\*</sup>, Maya Talantikite<sup>1</sup>, Agnès Tessier<sup>1</sup>, Caroline Cluzel<sup>1</sup>, Alice Bachmann<sup>1</sup>, Natacha Mariano<sup>1</sup>, Mélissa Dussoyer<sup>1</sup>, Lindsay B. Alcaraz<sup>2†</sup>, Laëtitia Fortin<sup>1</sup>, Alexandre Aubert<sup>1</sup>, Frédéric Delolme<sup>1,3</sup>, Naïma El Kholti<sup>1</sup>, Jean Armengaud<sup>4</sup>, Pierre Fournié<sup>5,6</sup>, Céline Auxenfans<sup>1,7</sup>, Ulrich Valcourt<sup>1,2</sup>, Sandrine Vadon-Le Goff<sup>1</sup>, Catherine Moali<sup>1‡</sup>

Copyright © 2020  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim  
to original U.S.  
Government Works

Bone morphogenetic protein 1 (BMP-1) is an important metalloproteinase that synchronizes growth factor activation with extracellular matrix assembly during morphogenesis and tissue repair. The mechanisms by which BMP-1 exerts these effects are highly context dependent. Because BMP-1 overexpression induces marked phenotypic changes in two human cell lines (HT1080 and 293-EBNA cells), we investigated how BMP-1 simultaneously affects cell-matrix interactions and growth factor activity in these cells. Increasing BMP-1 led to a loss of cell adhesion that depended on the matricellular glycoprotein thrombospondin-1 (TSP-1). BMP-1 cleaved TSP-1 between the VWFC/procollagen-like domain and the type 1 repeats that mediate several key TSP-1 functions. This cleavage induced the release of TSP-1 C-terminal domains from the extracellular matrix and abolished its previously described multisite cooperative interactions with heparan sulfate proteoglycans and CD36 on HT1080 cells. In addition, BMP-1-dependent proteolysis potentiated the TSP-1-mediated activation of latent transforming growth factor- $\beta$  (TGF- $\beta$ ), leading to increased signaling through the canonical SMAD pathway. In primary human corneal stromal cells (keratocytes), endogenous BMP-1 cleaved TSP-1, and the addition of exogenous BMP-1 enhanced cleavage, but this had no substantial effect on cell adhesion. Instead, processed TSP-1 promoted the differentiation of keratocytes into myofibroblasts and stimulated production of the myofibroblast marker  $\alpha$ -SMA, consistent with the presence of processed TSP-1 in human corneal scars. Our results indicate that BMP-1 can both trigger the disruption of cell adhesion and stimulate TGF- $\beta$  signaling in TSP-1-rich microenvironments, which has important potential consequences for wound healing and tumor progression.

## INTRODUCTION

Bone morphogenetic protein 1 (BMP-1) and related proteinases belonging to the BTP (BMP-1/tolloid-like proteinases) family are recognized as key regulators of several developmental and physiological processes. Their crucial role in development was first established in *Drosophila melanogaster* and *Xenopus laevis* embryos, in which they directly control dorsoventral patterning (1, 2). This was further supported by knockout studies in mice showing that deletion of the *Bmp1* or *Tll1* (tolloid-like 1) gene leads to early lethality and is associated with developmental defects such as gut herniation (3) and abnormal septation and positioning of the heart (4). Mutations of the *BMP1* gene in humans cause osteogenesis imperfecta (5), a bone disorder characterized by brittle bones that are prone to fracture.

A similar phenotype is observed in adult mice upon *Bmp1* conditional ablation with the additional phenotypes of periodontal disease and skin fragility (6–8). In addition, BTPs are known to be involved in the control of muscle growth and homeostasis (9) and in wound healing and tissue repair (6, 10). Last, several pathological conditions seem to involve increased BMP-1 abundance, including osteoarthritis (11), heart fibrosis (12–14), pelvic organ prolapse (15), corneal scarring (10), keloids (16), and type 1 diabetes (17). BMP-1 is also reported to belong to a group of signature extracellular matrix (ECM) proteins associated with the high metastatic potential of breast tumors (18), and its increased abundance correlates with high-grade tumors and/or poor survival prognosis in several types of cancers (19–21).

BMP-1 is a member of the stacin subgroup of metzincin metalloproteinases (22). The latent proform of the proteinase is processed in the secretory pathway to yield the mature enzyme, which is composed of a catalytic domain followed by three CUB (complement, uEGF, BMP-1) domains and one EGF (epidermal growth factor) domain. The noncatalytic domains also play a major, although indirect, role in enzyme activity by regulating substrate access to the active site and processing (23, 24). More than 30 BMP-1 substrates have been identified and characterized biochemically [reviewed in (25, 26)]. An important group of substrates consists of ECM precursor proteins that acquire the ability to form three-dimensional (3D) networks upon proteolytic maturation. Well-known examples are procollagens I to III, which can form fibers after C-terminal propeptide removal by BTPs to provide the principal scaffolding component of most connective tissues. Other prominent BMP-1 substrates

<sup>1</sup>University of Lyon, CNRS UMR 5305, Tissue Biology and Therapeutic Engineering Laboratory (LBTI), F-69367 Lyon, France. <sup>2</sup>University of Lyon, Centre Léon Bérard, INSERM U1052, CNRS UMR 5286, Cancer Research Center of Lyon (CRCL), F-69373 Lyon, France. <sup>3</sup>University of Lyon, ENS de Lyon, INSERM U58, CNRS UMS3444, SFR Biosciences, F-69366 Lyon, France. <sup>4</sup>CEA Marcoule, Innovative Technologies for Detection and Diagnostics Laboratory (DRF/Joliot/DMTS/SPI/Li2D), F-30200 Bagnols-sur-Cèze, France. <sup>5</sup>Purpan University Hospital, Ophthalmology Department, F-31059 Toulouse, France. <sup>6</sup>University of Toulouse, CNRS UMR 5165, INSERM U1056, Epithelial Differentiation and Rheumatoid Autoimmunity Unit (UDEAR), F-31059 Toulouse, France. <sup>7</sup>Hospices Civils de Lyon, Tissue and Cell Bank, F-69437 Lyon, France.

\*These authors contributed equally to this work.

†Present address: University of Montpellier, Montpellier Cancer Institute (ICM), INSERM U1194, Cancer Research Institute of Montpellier (IRCM), F-34298 Montpellier, France.

‡Corresponding author. Email: c.moali@ibcp.fr

belonging to this group are minor fibrillar procollagens (V and XI), small leucine-rich proteoglycans (decorin, osteoglycin, and biglycan), basement membrane components (laminin 332 and collagen VII), and mineralization factors (dentin matrix protein 1 and dentin sialophosphoprotein).

The second major group of substrates consists of latent forms of growth factors and cytokines or antagonists that sequester these ligands, which, upon cleavage, result in increased signaling activities. The first substrate in this group to be described was chordin, a major antagonist of BMP-2 and BMP-4 during development (2, 27). In addition, the latent forms of growth differentiation factor-11 (GDF-11) (28) and of the inhibitor of muscle growth GDF-8 (also known as myostatin) (9) as well as insulin-like growth factor-binding protein 3 (IGFBP3), which can bind and sequester IGF-I and IGF-II (29), can be cleaved by BMP-1 to release the active growth factors. A natural ligand of myostatin (WFIKKN1) was also found to behave as both a substrate of BMP-1 and a positive regulator of BMP-1-dependent myostatin activation (30), and three proteins involved in the control of transforming growth factor  $\beta$  (TGF- $\beta$ ) activity [latent TGF- $\beta$  (LTGF- $\beta$ ) binding protein-1 (LTBP-1), TGF- $\beta$  receptor III, and CD109] have been shown to be under the proteolytic control of BMP-1 (31, 32). Last, BMP-1 is also considered as a major player in lipid metabolism by mediating the maturation of apolipoprotein A1, the processing of the low-density lipoprotein receptor, and the release of the endotrophin adipokine from the C terminus of the  $\alpha 3$  chain of collagen VI (33–35).

In addition, several potential BMP-1 substrates identified in a proteomic screen (31) are still awaiting validation, suggesting that the substrate repertoire of the proteinase is still not fully characterized and that many new functions remain to be discovered. However, linking the phenotype of *Bmp1*-deficient animals and disease mechanisms to this large panel of substrates remains challenging because most BMP-1 activities are strongly contextual and interconnected. In the present study, we investigated the notable de-adhesion phenotype induced by BMP-1 overexpression in two human cell lines and revealed that this was mainly due to the cleavage of thrombospondin-1 (TSP-1), an ECM protein classified as “matricellular” for its ability to regulate cell-matrix interactions. Furthermore, we showed that the BMP-1-mediated cleavage of TSP-1 enhanced its ability to activate TGF- $\beta$  and to promote the differentiation of primary human keratocytes into myofibroblasts, both in 2D cultures and in an *ex vivo* model of corneal wound healing.

## RESULTS

### BMP-1 disrupts cell adhesion in 293-EBNA and HT1080 cells

293-EBNA cells, a variant of human embryonic kidney 293 cells expressing the Epstein-Barr nuclear antigen, have been widely used to produce recombinant BMP-1 (36–38). However, when these cells are grown to high density while producing BMP-1, they detach from culture plates en masse in the form of cell-cell adherent sheets (Fig. 1, A and B, and fig. S1A). This can be avoided by using plates coated with poly-D-lysine (37), but the mechanism underlying this observation has remained unclear. As a first step to elucidate the BMP-1-induced loss of adhesion, we added the synthetic BMP-1 inhibitor S33A [a hydroxamate inhibitor referred to as hxt in this study (31, 39)] to the culture medium and observed that the BMP-1-producing cells could be efficiently maintained in the adherent state in the presence of the inhibitor (Fig. 1A). This was the first indication that the loss

of adhesion involved the catalytic activity of BMP-1. We also observed that the secreted, active form of BMP-1 was present both in the cell supernatant and in the ECM of 293-EBNA cells (Fig. 1C).

To determine whether the loss of cell adhesion was a general property of BMP-1-transfected cells, we also transfected a construct encoding BMP-1 into the human fibrosarcoma cell line HT1080 to produce HT1080-BMP-1 cells (31). We used cells transfected with a construct encoding the catalytically inactive E94A mutant form of BMP-1 [HT1080-E94A; (31)] in parallel as a control. The adhesion of HT1080 cells also seemed to depend on the presence of active BMP-1 because HT1080-BMP-1 cells, but not HT1080-E94A cells, lost their ability to adhere when grown to high cell densities. In contrast to 293-EBNA cells, however, the loss of adhesion in HT1080 cells was preceded by notable modifications in cell morphology that became more pronounced as the cell density increased, as evidenced through phase contrast imaging, time-lapse microscopy, and actin staining (Fig. 1, D and E, and movies S1 and S2). The determination of the mean cell area after actin staining revealed a statistically significant decrease in HT1080-BMP-1 cell size before detachment (Fig. 1F). None of these modifications were observed in HT1080-E94A cells, which behaved very much like parental HT1080 cells. Moreover, addition of exogenous sizzled (Szl), a *Xenopus* protein that specifically inhibits BMP-1 activity (40, 41), to HT1080-BMP-1 cells inhibited both morphological changes and cell detachment (Fig. 1D). We therefore concluded that in two different cell lines, HT1080 and 293-EBNA cells, BMP-1 induced a loss of cell adhesion in a manner that depended on the catalytic activity of the proteinase.

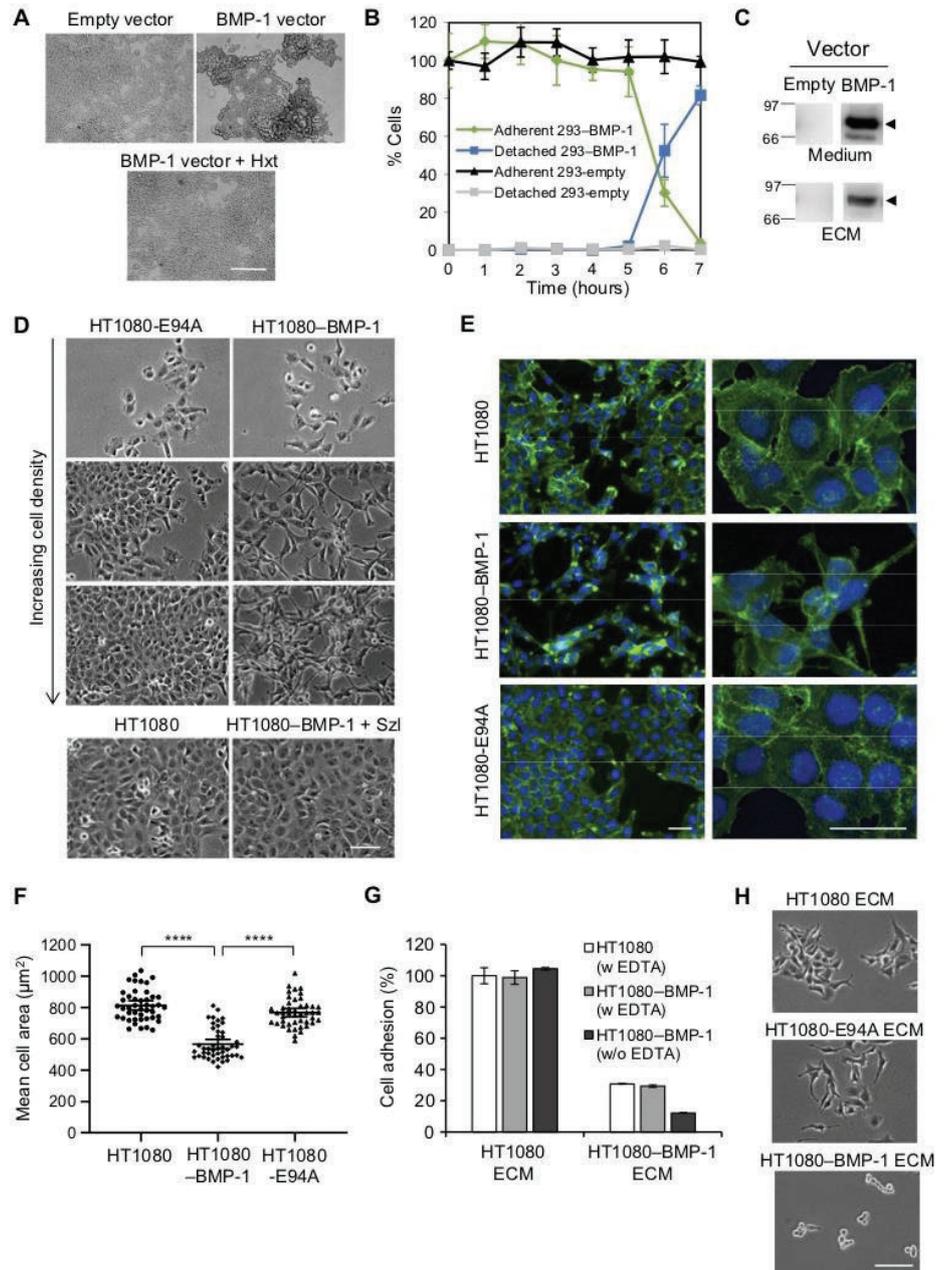
Last, we used flow cytometry to determine whether cell detachment could be a consequence of cell apoptosis by sorting cells labeled with an antibody recognizing annexin V and propidium iodide, which label apoptotic and dead cells, respectively. At the two cell densities that were analyzed (20 and 50%), we found that apoptotic cells represented less than 3% of the total cell population and that cell viability was more than 94% for all three HT1080 cell types (fig. S1B). These results demonstrate that apoptosis was not a major determinant of the observed changes in cell adhesion.

### The adhesion phenotype observed in HT1080-BMP-1 cells is due to changes in the ECM

To determine whether the loss of cell adhesion observed with HT1080-BMP-1 cells was triggered by a modification of cell adhesion receptors or a modification of the underlying ECM, we compared the adhesion of control HT1080 cells (chemically detached with EDTA) or HT1080-BMP-1 cells (either chemically or spontaneously detached) on the ECMs previously deposited by these cells. Although both cell types could adhere normally on the ECM deposited by control cells, their adhesion to the ECM of HT1080-BMP-1 cells was severely impaired (Fig. 1G). When observed under a microscope (fig. S2), few cells plated on the ECM deposited by HT1080-BMP-1 cells had adhered, and those remaining were round in shape and not spread; in contrast, cells could spread and firmly adhere to the ECM deposited by control cells. To determine whether cell adhesion on HT1080-BMP-1 ECM was completely impaired or just slightly delayed, HT1080 cells were allowed to adhere for up to 3.5 hours on the conditioned ECMs, but no progress regarding cell spreading was observed with the ECM from BMP-1-expressing cells (Fig. 1H). These results led to the conclusion that the most important changes induced by the activity of BMP-1 occurred in the ECM, whereas the cell surface was not substantially altered.

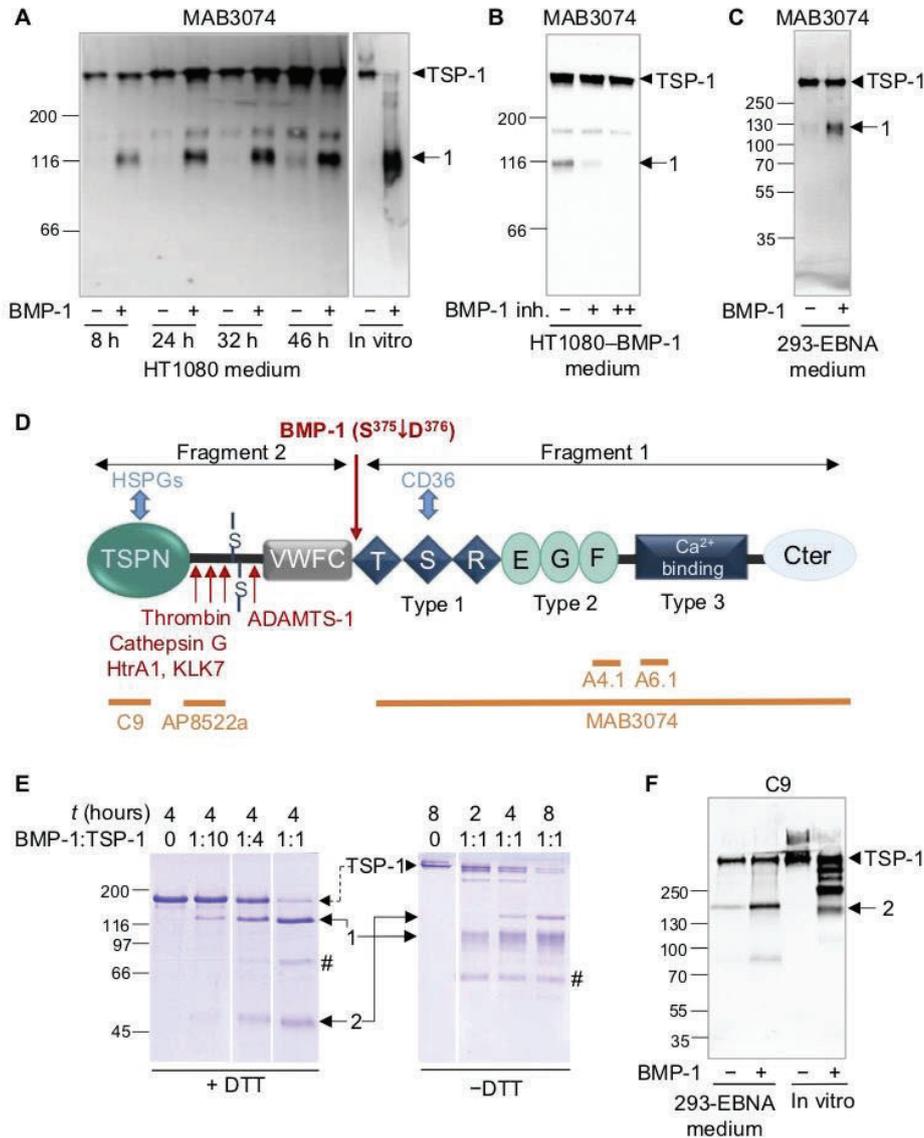
**Fig. 1. Changes in cell adhesion induced by BMP-1 cleavage of ECM proteins.**

**(A)** Phase contrast images of subconfluent 293-EBNA cells stably transfected with empty vector or a BMP-1-encoding vector, in the absence or presence of the BMP-1 inhibitor hydroxamate (Hxt). Images are representative of  $n = 3$  independent experiments. Scale bar, 150  $\mu\text{m}$ . **(B)** Percentage of adherent and detached subconfluent 293-EBNA cells expressing empty vector or BMP-1 over time after serum deprivation. Data from one of two independent experiments are presented as a percentage of total cells at the 0-hour time point, and error bars represent SDs of measurements from six technical replicates per group. Data from the other independent experiment are shown in fig. S1A. **(C)** Immunoblotting for BMP-1 (arrowhead) in the medium and ECM of 293-EBNA cells expressing empty vector or BMP-1. The four noncontiguous lanes belong to the same blot and are representative of  $n = 3$  independent experiments for medium and  $n = 2$  independent experiments for ECM. **(D)** Phase contrast images of HT1080-E94A and HT1080-BMP-1 cells at increasing cell densities, showing the morphological changes preceding the loss of cell adhesion. For comparison, parental HT1080 cells and HT1080-BMP-1 cells grown in the presence of the BMP-1 inhibitor Szl are also shown. Images are representative of  $n = 4$  independent experiments. Scale bar, 50  $\mu\text{m}$ . **(E)** Fluorescence imaging showing actin (phalloidin, green) and nuclei (blue) in HT1080, HT1080-BMP-1, and HT1080-E94A cells at intermediate cell density. Images are representative of  $n = 3$  independent experiments. Scale bars, 50  $\mu\text{m}$ . **(F)** Mean cell area of HT1080, HT1080-BMP-1, and HT1080-E94A cells measured using images taken as described in (E). The graph shows means of cell area  $\pm$  95% confidence interval from three independent experiments (15 fields analyzed per condition per experiment). Conditions were compared using the Kruskal-Wallis test with Dunn's multiple comparison test. \*\*\*\*,  $P < 0.0001$ . **(G)** Adhesion of HT1080 and HT1080-BMP-1 cells on decellularized ECM deposited by the same cells. Before the assay, the HT1080 cells to be tested were detached from culture plates with EDTA, and the HT1080-BMP-1 cells were either detached with EDTA (w EDTA) or collected after spontaneous detachment (w/o EDTA). Data show means  $\pm$  SD of  $n = 3$  independent experiments. **(H)** Phase contrast images of parental HT1080 cells seeded, in the absence of serum, on the decellularized ECM of HT1080, HT1080-E94A, or HT1080-BMP-1 cells after 3.5 hours. Scale bar, 50  $\mu\text{m}$ . Images are representative of  $n = 2$  independent experiments.

**BMP-1 cleaves TSP-1**

Our next aim was to identify the ECM substrate(s) processed by BMP-1. On the basis of our knowledge of the repertoire of proteins secreted by HT1080 cells characterized in a previous study (31), we screened the most prominent ECM components produced by these cells using Western blotting and identified TSP-1 as a promising candidate substrate. A fragment of TSP-1 migrating at around 120 kDa accumulated in the conditioned medium of HT1080-BMP-1 cells

(Fig. 2, A and B). This fragment was not detected in the medium of nontransfected cells (Fig. 2A) or when the HT1080-BMP-1 cells were treated with the BMP-1 inhibitor hxt (Fig. 2B). In vitro, the same fragment was released when TSP-1 (purified from human platelets) was incubated with recombinant BMP-1 (Fig. 2A). The fragment, called fragment 1, was also found to be present in the medium of 293-EBNA cells transfected with a BMP-1-encoding vector, whereas it was barely detectable in cells transfected with an empty vector



**Fig. 2. TSP-1 is cleaved by BMP-1.** (A) Immunoblotting for TSP-1 and one of its proteolytic products in the conditioned medium of HT1080 (–) and HT1080–BMP-1 (+) cells at different time points after serum starvation (nonreducing conditions). A similar product is observed when purified human TSP-1 is incubated with recombinant BMP-1 (in vitro). Blot is representative of  $n = 5$  independent experiments. (B) Immunoblotting for TSP-1 in the conditioned medium of HT1080–BMP-1 cells treated with 0 (–), 1 (+), or 5 (++)  $\mu\text{M}$  BMP-1 inhibitor Hxt and collected after 15 hours (nonreducing conditions). Blot is representative of  $n = 3$  independent experiments. (C) Immunoblotting for TSP-1 in the conditioned medium of 293-EBNA cells transfected with an empty vector (–) or a vector encoding BMP-1 (+) (nonreducing conditions). Medium was collected 9 hours after serum starvation. Blot is representative of  $n = 6$  independent experiments. (D) Domain structure of TSP-1, antibody epitopes, and positions of cleavage sites for BMP-1, thrombin, cathepsin G, HtrA1, KLK7, and ADAMTS-1. TSPN, thrombospondin N-terminal; VWFC, Von Willebrand Factor type C; TSR, thrombospondin type 1 repeats; EGF, epidermal growth factor; C-ter, thrombospondin C-terminal domain. (E) Analysis of TSP-1 cleavage at increasing BMP-1:TSP-1 ratios for the indicated incubation times under reducing (+DTT) and nonreducing (–DTT) conditions. “#” indicates the position of BMP-1. The image is representative of  $n = 2$  independent experiments. (F) Immunoblotting for TSP-1 in the conditioned medium of 293-EBNA cells collected as in (C) (nonreducing conditions) and comparison with TSP-1 cleaved in vitro by BMP-1. Blot is representative of  $n = 4$  independent experiments.

(Fig. 2C). The latter result suggested that this specific TSP-1 cleavage event was common to the two cell lines that shared similar adhesion phenotypes. Last, using antibodies with well-defined epitopes located

linker between the trimerization domain [Von Willebrand factor C (VWFC)] and TSP type 1 repeats (TSRs; Fig. 2D), we can infer that the N-terminal product (fragment 2) is a trimer in nonreducing

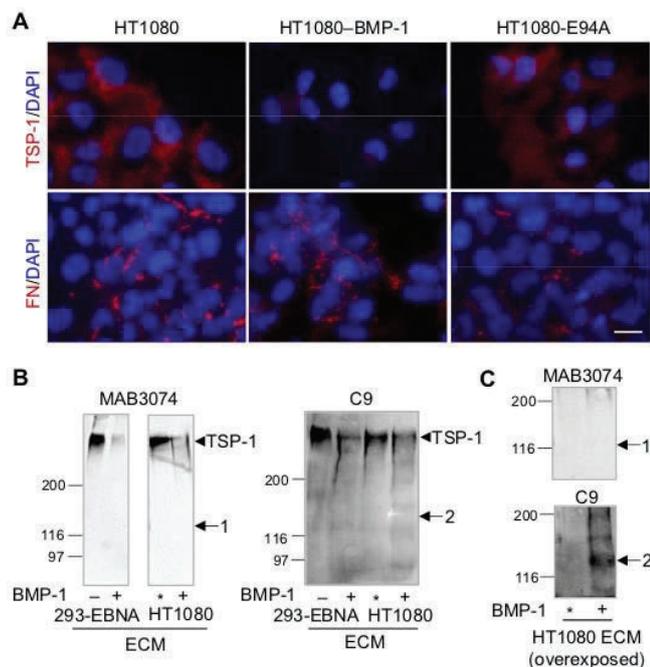
in the TSP-1 type 2 and 3 repeats (A4.1 and A6.1, respectively; Fig. 2D) (42), we could assign fragment 1 to the C-terminal part of TSP-1 (fig. S3, A and B).

To get more insight into the biochemical features of the proteolytic processing involved, we performed in vitro cleavage assays with varying proteinase concentrations and incubation times. The resulting products were analyzed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) in reducing and nonreducing conditions (Fig. 2E). As expected, the cleavage products increased when the protease:substrate ratio or the incubation time was increased. However, the cleavage pattern was very dependent on the presence of reducing agents. Whereas two main cleavage products were observed at around 120 and 50 kDa in reducing conditions (fragment 1 and fragment 2, respectively), four products—two above 200 kDa, one around 150 kDa, and one around 120 kDa—were detected in nonreducing conditions. This apparent discrepancy can be explained by the trimeric nature of TSP-1, wherein molecules are held together by interchain disulfide bonds involving two cysteines in their N-terminal half (Fig. 2D); this resulted in intermediate species when only one or two chain(s) were cleaved, as observed in nonreducing conditions. We then analyzed the upper fragment (fragment 1) observed in reducing conditions by N-terminal Edman sequencing and could unambiguously localize the BMP-1 cleavage site between Ser<sup>375</sup> and Asp<sup>376</sup> (N-terminal sequence: DSADX), in precise agreement with the known specificity of BMP-1 (25). Moreover, cleavage site identification using a mass spectrometry–based approach involving isotopic labeling of neopeptides (ATOMS) (31, 43) confirmed that the peptide starting with Asp<sup>376</sup> and ending with Arg<sup>402</sup> was much more abundant in the presence of BMP-1 than in its absence (proteinase:control ratio = 47 based on five identifications). Furthermore, two unusual posttranslational modifications on Trp<sup>385</sup> (C-mannosylation) and Ser<sup>394</sup> (disaccharide Glc-Fuc-O) were unambiguously detected on this peptide (fig. S4), confirming the findings of a previous study (44). From the identification of the cleavage site in the short

conditions (migrating around 150 kDa) and a monomer in reducing conditions (migrating around 50 kDa; Fig. 2, E and F, and fig. S3), whereas fragment 1 is observed as a monomer in both conditions (migrating around 120 kDa).

### BMP-1-mediated cleavage releases the C-terminal fragment of TSP-1 from the ECM

Our observations indicated that TSP-1 was cleaved in the conditioned medium of 293-EBNA and HT1080 cells if the cells expressed BMP-1, but the loss of cell adhesion also involved major modifications of the ECM (Fig. 1G). Using fluorescence microscopy with the MAB3074 antibody directed against the TSP-1 domains located downstream of the BMP-1 cleavage site, we observed a strong decrease of TSP-1 staining in the cell layer of HT1080–BMP-1 cells before detachment compared to HT1080 and HT1080–E94A cells (Fig. 3A). In contrast and despite the fact that fibronectin is a known ligand of TSP-1, fibronectin fibrils did not seem to be modified by the presence of active BMP-1 (Fig. 3A). ECM extracts of the different cell types analyzed by Western blotting also confirmed that full-length TSP-1 was present in the ECM deposited by control cells but almost completely absent from the ECM of BMP-1-expressing cells (Fig. 3B).



**Fig. 3. The C-terminal fragment of TSP-1 is released from the ECM.** (A) Immunostaining of HT1080, HT1080–BMP-1, and HT1080–E94A cells with antibodies specific for TSP-1 (MAB3074) and fibronectin (FN). Nuclei are labeled with DAPI. Scale bar, 20  $\mu$ m. Images are representative of  $n = 3$  independent experiments. (B) Immunoblotting for TSP-1 in the ECM of HT1080 or 293-EBNA cells transfected with an empty vector (–), a vector encoding the E94A mutant form of BMP-1 (\*), or a vector encoding BMP-1 (+) under nonreducing conditions. Antibody MAB3074 recognizes TSP-1 domains located downstream of the BMP-1 cleavage site; antibody C9 recognizes TSP-1 domains located upstream of the BMP-1 cleavage site. Blot is representative of  $n = 2$  independent experiments, except for 293-EBNA ECM detected with the MAB3074 antibody ( $n = 6$ ). (C) Immunoblotting for TSP-1 fragments in the ECM of HT1080 cells expressing BMP-1 (+) or E94A (\*) with higher antibody concentrations and longer exposition times (nonreducing conditions). Blots are representative of  $n = 2$  independent experiments.

Last, whereas TSP-1 fragment 1 was never detected in the ECM samples even using long exposure times, at least a fraction of fragment 2 remained, entrapped in the ECM compartment (Fig. 3C). These results indicated that BMP-1 cleaved the matrix-embedded TSP-1 and thereby triggered the release of the C-terminal part of the molecule from the ECM.

### Cleavage abolishes the ability of TSP-1 to promote cell adhesion

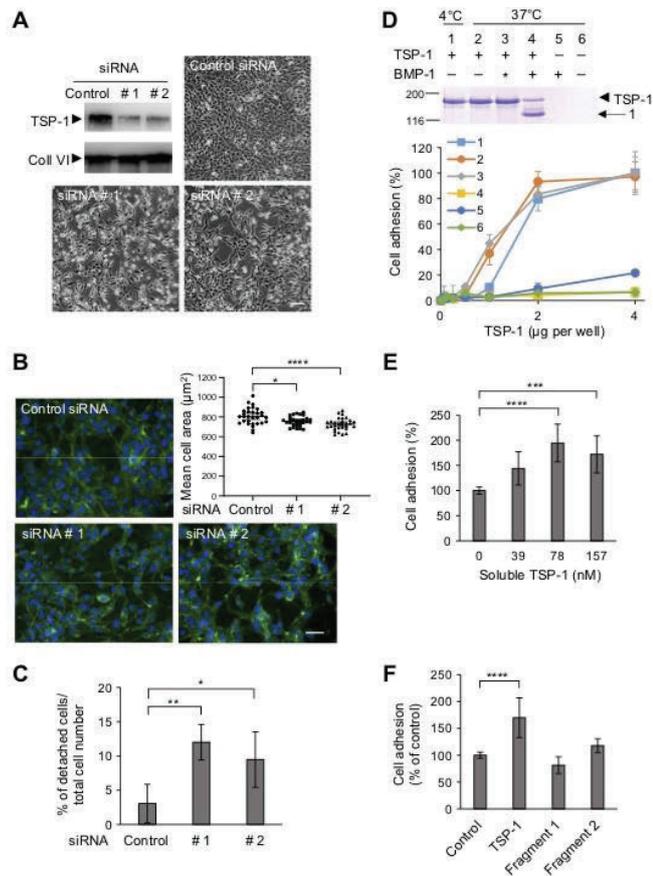
Both the N- and C-terminal fragments of TSP-1 have the potential to play a role in cell adhesion. The N-terminal portion can interact with several cellular receptors, including integrins, heparan sulfate proteoglycans, and calreticulin, and C-terminal domains found in fragment 1 can interact with integrins, CD36 or CD47 (also known as integrin-associated protein) (45, 46). Using small interfering RNAs (siRNAs) directed against human TSP-1, we first demonstrated that TSP-1 was critical for the adhesion of HT1080 cells. When TSP-1 was knocked down (by around 60% as determined by Western blotting; fig. S5), the cells changed their morphology, became rounded, and seemed to be more loosely attached to the plates (Fig. 4A). This was associated with a downward trend in cell area already visible 48 hours after transfection (Fig. 4B) and a statistically significant increase in the number of detached cells 72 hours after transfection (Fig. 4C), similarly to previous observations made in HT1080–BMP-1 cells (Fig. 1, E and F).

In a second series of experiments, we analyzed whether cleaved TSP-1 was still capable of promoting HT1080 adhesion. Purified TSP-1 was first subjected to BMP-1 proteolytic activity for 4 hours and then incubated overnight in a 96-well plate, and adhesion of HT1080 cells to the cleaved and intact proteins was measured. We observed that processing by BMP-1 completely abolished the ability of TSP-1 to promote adhesion (Fig. 4D, condition 4). The loss of TSP-1 adhesive properties was clearly due to proteolysis because incubation with the inactive mutant form of BMP-1 did not lead to the same effect (Fig. 4D, condition 3). Control experiments were also performed to exclude (i) protein unfolding after prolonged incubation at 37°C (the intact protein incubated in the same conditions was even slightly more active than the protein incubated at 4°C; conditions 1 and 2) or (ii) differential coating efficiency of 96-well plates by intact and cleaved TSP-1 (fig. S6, A and B).

Last, the contribution of soluble fragments 1 and 2 to the loss of cell adhesion was assessed through the adhesion of parental HT1080 cells to wells coated with TSP-1 in competition with intact, soluble TSP-1, purified fragment 1, or semipurified fragment 2 (fig. S7). We observed an increase in the adhesion of HT1080 cells (Fig. 4E) with increasing concentrations of intact, soluble TSP-1, possibly explained by the previously described ability of TSP-1 to recruit other TSP molecules through intermolecular contacts (47, 48). In contrast, equimolar amounts of either fragment 1 or fragment 2 did not display statistically significant effects on cell adhesion to TSP-1-coated wells (Fig. 4F). Individual fragments released into cell supernatants from matrix-embedded or soluble TSP-1 are therefore not expected to contribute substantially to the adhesion phenotype of HT1080 cells.

### Cleavage enhances the ability of TSP-1 to release active TGF- $\beta$

In addition to mediating ECM-cell interactions, another major function of TSP-1 is to release the active form of TGF- $\beta$  that, after secretion, remains noncovalently bound to its latency-associated peptide (LAP) prodomain and, in this tightly-interacting complex, is kept in



**Fig. 4. Intact TSP-1 promotes the adhesion of HT1080 cells but cleaved TSP-1 does not.** (A) Phase contrast images of parental HT1080 cells transfected with two siRNAs targeting TSP-1 or with a control siRNA 72 hours after transfection. Scale bar, 200 µm. Extinction efficiency was assessed through the immuno-detection of TSP-1 in cell medium (loading control: collagen VI). Images and blots are representative of  $n = 4$  and  $n = 3$  independent experiments, respectively. (B) Fluorescence microscopy showing actin (green) and nuclei (blue) in HT1080 cells transfected with the indicated siRNAs and quantification of cell area. Scale bar, 50 µm. Data are representative of two independent experiments and show means  $\pm$  95% confidence intervals from 15 fields per condition and replicate. \*,  $P < 0.05$  and \*\*\*\*,  $P < 0.0001$ . (C) Percentage of detached cells 72 hours after transfection with the indicated siRNAs. The data show means  $\pm$  SD of four independent experiments. \*,  $P < 0.05$  and \*\*,  $P < 0.01$ . (D) Adhesion of HT1080 cells on intact and cleaved TSP-1. TSP-1 was incubated for 4 hours with (+) or without (-) recombinant BMP-1 or the inactive E94A mutant (\*), at 4°C or 37°C before the assay (conditions 1 to 4). Control conditions with BMP-1 alone or buffer are also shown (conditions 5 and 6). The gel is representative of  $n = 3$  independent experiments, and the graph shows means  $\pm$  SD of the same experiments (performed in triplicate for each TSP-1 coating concentration). (E) Adhesion of HT1080 cells on TSP-1-coated plates in the presence of increasing concentrations of soluble TSP-1. The graph shows means  $\pm$  SD of three independent experiments each performed with three or four technical replicates. Conditions were compared to control using the Kruskal-Wallis test with Dunn's multiple comparison test. \*\*\*,  $P < 0.001$  and \*\*\*\*,  $P < 0.0001$ . (F) Adhesion of HT1080 cells on TSP-1-coated plates in the presence of 157 nM soluble full-length TSP-1, purified fragment 1, or semipurified fragment 2. The graph shows means  $\pm$  SD of three independent experiments each performed with three or four technical replicates. Conditions were compared to control using the Kruskal-Wallis test with Dunn's multiple comparison test. \*\*\*\*,  $P < 0.0001$ .

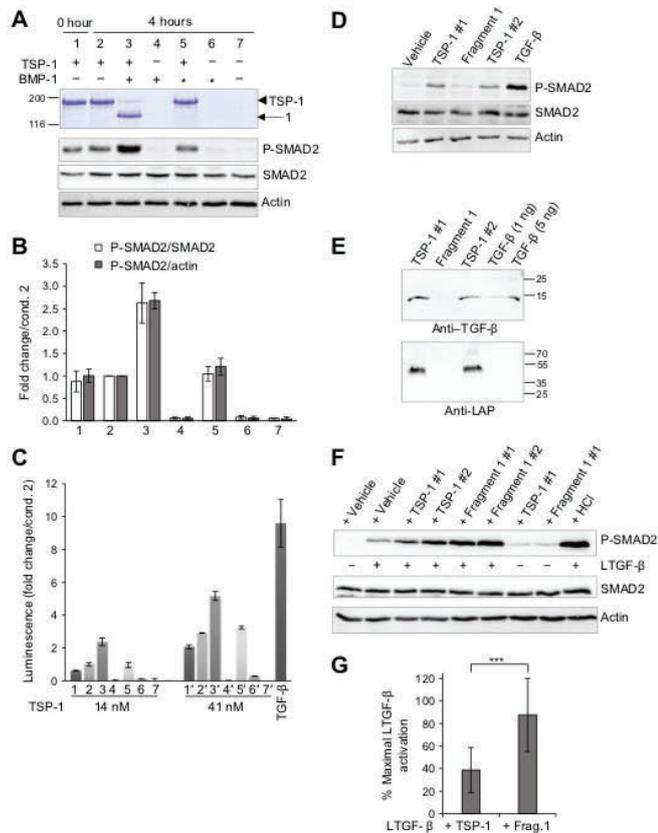
the latent form. When TSP-1 was added to HT1080 cells, the canonical TGF- $\beta$  pathway was indeed found to be activated, as evidenced by the detection of the phosphorylated form of the downstream effector SMAD2 60 min after TSP-1 addition (Fig. 5A). We also analyzed the ability of cleaved TSP-1 (provided to the cells as a mixture of TSP-1 fragments plus BMP-1 without any further purification of the in vitro reaction) to trigger TGF- $\beta$  signaling and found that it even more efficiently induced SMAD2 phosphorylation compared to the intact protein (Fig. 5A). After quantitation and normalization to the amount of actin or total SMAD2, the enhancement ratio was calculated to be around 2.7 (Fig. 5B). The increased TGF- $\beta$  activation capacity of cleaved TSP-1 was confirmed using a luciferase reporter assay (31). At the two concentrations of TSP-1 that were tested (14 and 41 nM), luciferase activity was more than twofold higher after incubation with active BMP-1 than with inactive BMP-1 or buffer alone (Fig. 5C). This suggested a higher TGF- $\beta$  activation capacity of cleaved TSP-1 compared to the full-length form.

The TSRs of TSP-1, which are located in fragment 1 just after the BMP-1 cleavage site, mediate the capacity of TSP-1 to activate TGF- $\beta$  (49). We tested whether purified fragment 1 was also capable of promoting TGF- $\beta$  activation in the same conditions in which intact TSP-1 did but we found that it did not elicit any increase in SMAD2 phosphorylation (Fig. 5D). To explain the apparent discrepancy, we analyzed the amounts of residual TGF- $\beta$  and LAP that were bound to intact TSP-1 and fragment 1. Using purified mature TGF- $\beta$  as a control, we estimated that our TSP-1 preparations reproducibly carried about 1 molecule of dimeric small LTGF- $\beta$  complex for every 100 molecules of trimeric TSP-1 (Fig. 5E) despite the high-pH stripping procedure used in the purification protocol. In contrast, purified fragment 1 was found to be totally devoid of TGF- $\beta$  and LAP, suggesting that cleavage by BMP-1 promoted the dissociation of LTGF- $\beta$  from TSP-1. However, this decreased stability did not preclude the activation of LTGF- $\beta$  by fragment 1 because SMAD signaling increased when an exogenous source of LTGF- $\beta$  was added to HT1080 cells after a short incubation at 37°C with fragment 1 (Fig. 5F). This effect was consistent between two different batches of fragment 1 with an activation efficiency similar to that achieved by treatment of LTGF- $\beta$  with 10 mM HCl and superior to that of intact TSP-1 (Fig. 5G). Note that TSP-1 final concentration in cell medium is too low (1.2 nM) in the conditions of this assay to trigger detectable SMAD2 phosphorylation when provided alone to the cells.

Together, these experiments suggest that (i) purified platelet TSP-1 promoted SMAD signaling through the release of its own TGF- $\beta$  load rather than through the activation of cell-secreted LTGF- $\beta$  (Fig. 5, A, D, and E), (ii) cleavage of TSP-1 by BMP-1 and the subsequent loss of TSR trimeric organization enhanced the activation and release of the TGF- $\beta$  bound to TSP-1 (Fig. 5, E to G), and (iii) fragment 1 did not act as a reservoir of LTGF- $\beta$  but was very efficient to activate the growth factor from exogenous sources of the latent complex (Fig. 5, E to G).

### TSP-1 is physiologically cleaved in human primary keratocytes and induces myofibroblast differentiation

To get further insight into the effect of the type of ECM proteins synthesized by cells on the BMP-1-dependent adhesion phenotype, we assessed the effect of increasing or decreasing the activity of BMP-1 in the culture medium of human primary keratocytes (corneal stromal cells), which secrete a collagen-rich ECM and have been demonstrated to be a suitable model studying BMP-1 activity (31, 35).



**Fig. 5. TSP-1 cleavage enhances TGF- $\beta$  activation in HT1080 cells.** (A) Immunoblotting for SMAD2 and phosphorylated SMAD2 (P-SMAD2) in lysates of HT1080 cells treated for 60 min with the indicated combinations of TSP-1 (visible on the Coomassie-stained gel) and BMP-1 or E94A mutant (\*) pre-incubated for the indicated time. Final TSP-1 concentration in culture medium was 16.4 nM. Actin is a loading control. The gel and blots are representative of  $n = 3$  independent experiments. (B) Quantitation of P-SMAD2 in treatment groups 1 to 7, corresponding to the same conditions as in (A). Intensity was normalized relative to total SMAD2 or actin. The data show means  $\pm$  SD of three independent experiments. (C) Quantitation of luminescence in lysates of HT1080 cells transfected with the TGF- $\beta$  reporter pGL3(CAGA)<sub>3</sub>-Luc for 24 hours and then treated with the same solutions 1 to 7 as in (A) (final TSP-1 concentrations: 14 or 41 nM) or with TGF- $\beta$  (10 ng/ml) for another 24 hours. The data show means  $\pm$  SD of three independent experiments. (D) Immunoblotting for P-SMAD2 and SMAD2 in lysates of HT1080 cells stimulated for 60 min with 41 nM full-length TSP-1 (two different batches indicated as #1 and #2) or purified fragment 1. Blots are representative of four independent experiments. (E) Immunoblotting for mature TGF- $\beta$  and the LAP prodomain in 7.1  $\mu$ g of full-length TSP-1 (two different batches), purified fragment 1, or with the indicated amounts of recombinant active TGF- $\beta$  under reducing conditions. Blots are representative of three independent experiments. (F) Immunoblotting for P-SMAD2 and SMAD2 in lysates of HT1080 cells treated for 60 min with 390 pM final concentration of LTGF- $\beta$  [equimolar to TGF- $\beta$  (10 ng/ml)]. Before cell treatment and dilution in culture medium, LTGF- $\beta$  was incubated with vehicle, or with full-length TSP-1 or purified fragment 1 (two different batches of each) or with HCl for 15 min. Blots are representative of three independent experiments. (G) Quantitation of LTGF- $\beta$  activation by intact and cleaved TSP-1 based on immunoblots obtained as in (F). The graph shows means  $\pm$  SD of percentages of maximal LTGF- $\beta$  activation compared to 100% for acid activation.  $n = 3$  independent experiments. Activations with TSP-1 and fragment 1 were compared using the Student's *t* test. \*\*\*,  $P < 0.001$ .

We found no change in the adhesion profile of keratocytes in BMP-1-supplemented culture medium, a condition in which HT1080 cell adhesion was greatly reduced (Fig. 6A). However, when we assayed TSP-1 in the conditioned medium of the keratocytes, we detected three forms, corresponding to full-length TSP-1 and to cleaved products similar in size to the fragments 1 and 2 described above (Fig. 6B). These fragments were diminished in the presence of BMP-1 inhibitors and increased upon the addition of exogenous BMP-1 (Fig. 6, B and C), demonstrating the direct involvement of the proteinase in generating the fragments. When the ECM of keratocytes was analyzed, full-length TSP-1 was still present in large amounts (Fig. 6B) despite clear evidence of BMP-1-dependent cleavage of the matrix-embedded TSP-1. Like in HT1080 cells, fragment 2 was detected in the keratocyte ECM but fragment 1 was missing. These results indicated that the endogenous amounts of BMP-1 produced by human keratocytes (Fig. 6D) were sufficient to cleave soluble and insoluble TSP-1 but that these cells do not rely on the same mechanisms as HT1080 cells to adhere to their ECM.

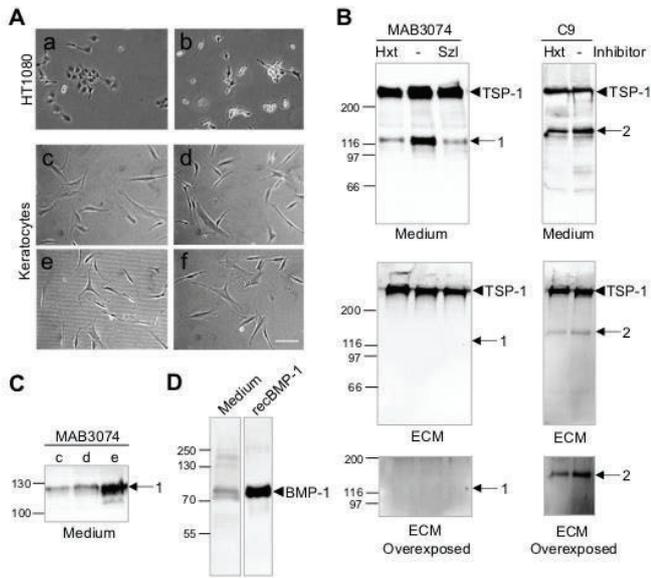
We also analyzed the ability of cleaved TSP-1 to trigger TGF- $\beta$  signaling in primary keratocytes. Similarly to what was observed for HT1080 cells, we observed an increase in the amount of SMAD phosphorylation in the presence of cleaved TSP-1 compared to full-length TSP-1 (Fig. 7, A and B). In addition, cleaved TSP-1 was statistically significantly more efficient than intact TSP-1 in inducing the production of TGF- $\beta$  target genes such as  $\alpha$ -smooth muscle actin ( $\alpha$ -SMA; Fig. 7, C and D). This further shows that the cleavage of TSP-1 by BMP-1 promoted the differentiation of keratocytes into myofibroblasts.

#### TSP-1 is cleaved in human corneal scars

Because both BMP-1 and TSP-1 have previously been proposed to play a role in the development of corneal scarring (10, 50), we analyzed four human corneas from patients presenting with permanent corneal scars and collected at the time of corneal transplantation. Scars were caused by perforating traumas or corneal ulcers that occurred between 6 and 26 months before surgery. In the lysates of all four corneas, a TSP-1 product migrating at the same size as the fragment 1 identified in keratocyte supernatant was detected (Fig. 8A). This result suggested that BMP-1-dependent processing of TSP-1 also occurred *in vivo* and could contribute to the development of corneal scars.

#### BMP-1-dependent proteolysis of TSP-1 induces myofibroblast differentiation in an *ex vivo* model of mouse corneal wound healing

To get further insights into the potential roles of TSP-1 fragments in corneal wound healing, we adapted a previously described *ex vivo* model of corneal wound healing (51). Because it was not possible to gather enough human corneas to compare all possible conditions, we used murine corneas and created calibrated wounds in the center of the corneas with 1-mm-diameter biopsy punches (Fig. 8, B and C). The wounded corneas were then placed in culture medium for 1 week and treated twice with intact TSP-1, cleaved TSP-1, cleaved TSP-1 in combination with the TGF- $\beta$  type I receptor inhibitor SB431542 (52), or appropriate controls (vehicle, BMP-1, and TGF- $\beta$ ) at days 2 and 4. We measured wound contraction once a day, calculated as the percentage of initial wound area, and observed that the wound area was statistically significantly decreased in corneas treated with cleaved TSP-1 at days 6 and 7 compared to corneas treated with vehicle



**Fig. 6. Effect of TSP-1 cleavage on the adhesion of keratocytes.** (A) Parental HT1080 cells were seeded in serum-free conditioned medium from HT1080 (a) or HT1080-BMP-1 (b) cells as controls. Images were taken 24 hours after seeding and are representative of  $n = 2$  independent experiments. Keratocytes were seeded in freshly collected, serum-free conditioned medium of keratocytes (c to e) in the absence of any additive (c) or in the presence of vehicle (d) or of 185 nM recombinant BMP-1 (e) or in serum-free conditioned medium of keratocytes cultivated with 5  $\mu$ M BMP-1 inhibitor Hxt (f). Images were taken 6 hours after seeding and are representative of  $n = 3$  independent experiments. Scale bar, 50  $\mu$ m. (B) Immunoblotting for TSP-1 and its proteolytic products in the conditioned medium and ECM of keratocytes cultured in serum-free conditions for 48 hours in the absence or presence of 30  $\mu$ M Hxt or of 300 nM Szl. Blots probed with MAB3074 are representative of  $n = 4$  (medium) or 3 (ECM) independent experiments, and blots probed with C9 are representative of  $n = 2$  independent experiments. (C) Immunoblotting for fragment 1 in the supernatants of cells grown in conditions c to e at the end of the adhesion experiment described in (A). Blots are representative of  $n = 3$  independent experiments. (D) Immunoblotting for endogenous BMP-1 in the conditioned medium of human primary keratocytes under reducing conditions. Medium was collected after 48 hours and concentrated using ultrafiltration. Recombinant purified human BMP-1 (reBMP-1; 50 ng) is a control. Blot is representative of  $n = 2$  independent experiments.

(Fig. 8D). In addition, the combination of cleaved TSP-1 and SB431542 followed a curve that nearly superimposed on the vehicle curve and led to a statistically significant reduction of wound contraction compared to treatment with cleaved TSP-1 alone.

After 7 days in culture, corneas were stained for  $\alpha$ -SMA. As observed in the first days of corneal wound healing in vivo (53), keratocytes located close to the wound died by apoptosis, whereas stromal cells from the wound margins proliferated and activated to acquire a myofibroblast phenotype (Fig. 8E). This showed that our model was relevant to analyze keratocyte activation, and we proceeded to evaluate the number of cells that differentiated into myofibroblasts by immunofluorescence. This analysis indicated a strong increase in the percentage of cells at the wound margins that were myofibroblasts when corneas were treated with cleaved TSP-1 or TGF- $\beta$  compared to treatment with vehicle, BMP-1, or intact TSP-1 (Fig. 8, E and F). In addition, the effect of cleaved TSP-1 was confirmed to be mediated by TGF- $\beta$  because SB431542 completely reversed the induction of  $\alpha$ -SMA production elicited by cleaved TSP-1 alone.

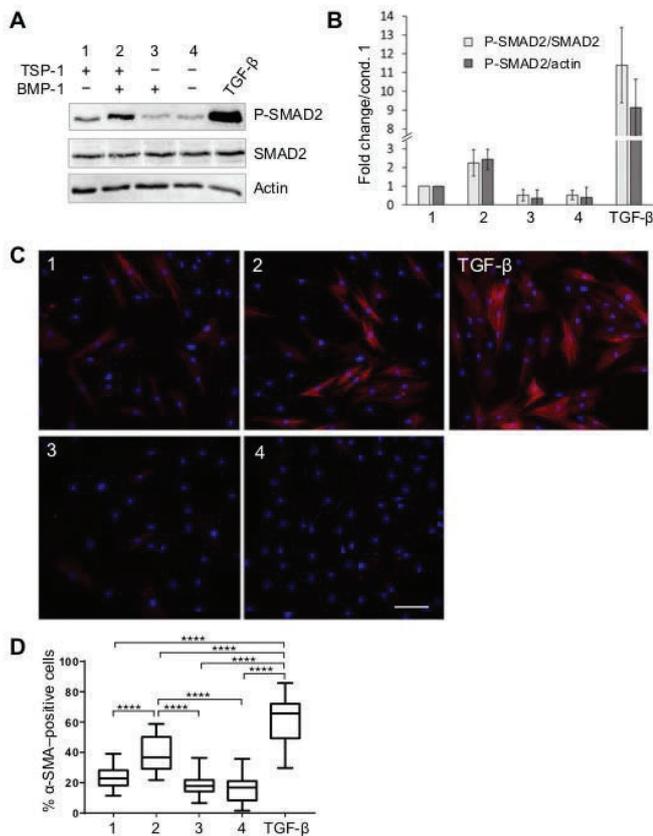
This established that cleaved TSP-1 was a potent inducer of myofibroblast differentiation through the release of bound TGF- $\beta$ , both in 2D cultures and in the 3D physiological environment of the cornea, and, based on the observation of cleaved TSP-1 in human corneal scars, that it was highly likely to play a role in wound healing in vivo.

## DISCUSSION

Cell growth, survival, and differentiation result from the integration of numerous mechanical and molecular cues present in the external environment of cells. One prominent source of these signals is the ECM that surrounds cells and provides a great diversity of components with various abilities to promote cell adhesion through specific interactions with cellular receptors. Here, we show that, by directly modifying the ECM, BMP-1 strongly alters the ability of two cell types to adhere.

Several metalloproteinases have previously been shown to induce a similar disruption of cell-matrix contacts, both through proteolytic and nonproteolytic activities (54, 55). In the present case, the drastic alterations of cell properties induced by BMP-1 in HT1080 and 293-EBNA cells clearly depended on its catalytic activity, and we linked the loss of adhesion to the cleavage of TSP-1. Several sites in TSP-1 that interact with cellular receptors have been identified (45, 46, 56), including domains that mediate interactions with calreticulin (TSPN domain; Fig. 2D),  $\beta$ 1 integrins (TSPN domain, type 1 to 3 repeats),  $\beta$ 3 integrins (type 3 repeats), heparan sulfate proteoglycans (TSPN domain, type 1 repeats), CD36 (type 1 repeats), and CD47 (C-terminal domain). As a result, the adhesive properties of TSP-1 will depend on several parameters such as (i) the cell type (which determines the specific repertoire of cell receptors), (ii) the local TSP-1 concentration and availability (which is potentially limited by the occupancy of receptor binding sites by other TSP-1 partners), and (iii) whether TSP-1 is soluble or matrix bound. In addition, previous reports have shown that proteolysis of TSP-1 can also affect cell adhesion (57–59). In one of these studies, Asch *et al.* (58) demonstrated that HT1080 cells can attach and spread on surfaces coated with TSP-1 through a multisite cooperative attachment process involving both the interaction of the N-terminal TSPN domain with heparan sulfate proteoglycans and the interaction of TSRs with CD36 (Fig. 2D). They also observed that HT1080 cells can attach but do not spread on a 140-kDa C-terminal fragment of TSP-1, generated by thrombin cleavage (Fig. 2D) and lacking the heparin-binding TSPN domain (58). In contrast, they can neither attach nor spread on the TSPN domain alone (58). Last, the same authors ruled out the involvement of integrins such as the  $\alpha$ v $\beta$ 3 integrin, an important TSP-1 cell receptor present on HT1080 cells and mediating adhesion through interactions with the RGD sequence in type 3 repeats, because the RGD peptide has no effect on cell adhesion.

BMP-1 cleaves TSP-1 between the procollagen-like/VWFC domain and the TSRs (Fig. 2D) at a site (Ser<sup>375</sup>↓Asp<sup>376</sup>) that is highly conserved in TSP-1 from a wide range of species but not present in human TSP-2 (fig. S8). The cleavage site location is different from the one targeted by thrombin and other serine proteases reported to make single cuts in TSP-1 [Val<sup>214</sup>↓Arg<sup>215</sup> for HtrA1 (60), Arg<sup>255</sup>↓Thr<sup>256</sup> for cathepsin G (61), and Tyr<sup>258</sup>↓Ile<sup>259</sup> for kallikrein 7 (62)] and seems to have an even stronger effect on the adhesion of HT1080 cells on TSP-1 than does the thrombin cleavage site. Even the C-terminal fragment 1 containing the CD36 interaction site, when provided in



**Fig. 7. TSP-1 cleavage potentiates keratocyte differentiation into myfibroblasts.** (A) Immunoblotting for P-SMAD2 and SMAD2 in cell lysates of keratocytes treated for 30 min with intact and cleaved TSP-1 (41 nM final concentration) or TGF- $\beta$  (10 ng/ml). Actin is a loading control. Blots are representative of  $n = 3$  independent experiments. (B) Quantitation of P-SMAD2 in the same conditions as in (A). Intensity was normalized relative to total SMAD2 and actin. The graph shows means  $\pm$  SD of  $n = 3$  independent experiments. (C) Staining for  $\alpha$ -SMA in primary keratocytes maintained for 72 hours in serum-free medium under the same treatment conditions 1 to 4 as in (A) or with TGF- $\beta$  (10 ng/ml). Scale bar, 100  $\mu$ m. Images are representative of  $n = 2$  independent experiments. (D) Quantitation of cells with  $\alpha$ -SMA fibers. The box plot shows data from 20 images per condition from two independent experiments. All conditions were compared using two-way ANOVA and Tukey post-ANOVA test. \*\*\*\*,  $P < 0.0001$ .

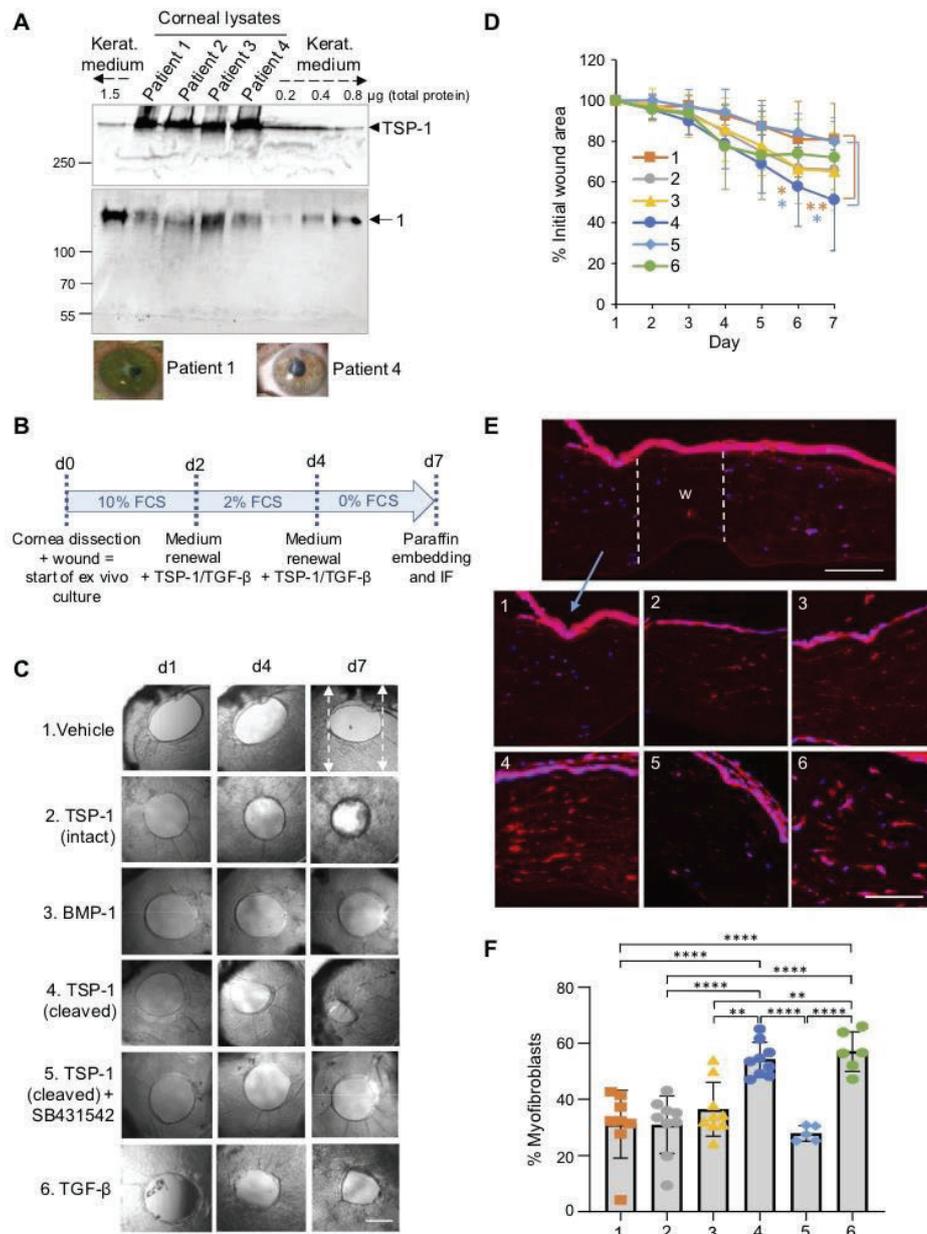
equimolar concentration with fragment 2, lost its ability to promote cell attachment (Fig. 4D). Because processing occurs just N-terminal to the known CD36 interaction site located within the second type 1 repeat (56), it is most likely that the cleavage by BMP-1 induces a major conformational rearrangement of this region, at the level of each polypeptide chain and/or of the quaternary structure (through the loss of the trimeric organization), thereby leading to disruption of the CD36 interaction site. This effect on the CD36 binding site is likely to affect other cell types, especially endothelial cells, for which CD36 is a major TSP-1 receptor (63), with possible consequences for the regulation of angiogenesis (Fig. 9). This also suggests new hypotheses linking BMP-1 expression with neovascularization (17, 64) and could explain some of the tumor-promoting activities described for BMP-1 (19–21, 65, 66).

In addition, we observed that the BMP-1-dependent cleavage of TSP-1 triggered the release of the TSR-bearing fragment (fragment

1) from the matrix, whereas the N terminus was still present in the ECM. A similar mechanism of transfer from the matrix to the soluble phase has been previously observed upon cleavage of TSP-1 by ADAMTS-1 (67) at an upstream site located just before the VWFC domain (Glu<sup>311</sup>↓Leu<sup>312</sup>; Fig. 2D), which also generates monomeric C-terminal fragments. These observations are in agreement with the previous demonstration that TSP-1 accumulation in the ECM is mediated by its L-type lectin C-terminal domain, through intermolecular interactions with itself or TSP-5 (also known as cartilage oligomeric protein), a process that requires the trimeric form of this domain and is not supported by monomers (47, 48). Last, we tested whether the competition between insoluble TSP-1 and soluble fragments could be another possible de-adhesive mechanism and found that fragments 1 and 2 had no notable inhibitory or enhancing effect on HT1080 adhesion to TSP-1-coated surfaces despite the presence of the hep I sequence, which is known to promote focal adhesion disassembly (68), in fragment 2. In summary, the observed loss of adhesion in cultures of HT1080 and 293-EBNA cells probably arises from the combined effects of (i) structural rearrangements disrupting receptor binding sites in fragment 1, (ii) the interruption of cooperativity mechanisms between both sides of the cleavage site, and (iii) the decreased amount of the C-terminal region of the protein in the ECM.

As illustrated here in the case of keratocytes, the processing of TSP-1 by BMP-1 does not always lead to the disruption of cell-matrix interactions. Actually, despite relatively high endogenous amounts of processed TSP-1, keratocyte adhesion does not seem affected by the inhibition of endogenous BMP-1 or by the addition of exogenous BMP-1. Although receptors possibly involved in the adhesion of keratocytes to TSP-1 have not been described, they are likely to differ from those on HT1080 cells and could be differently affected by TSP-1 cleavage. In addition, other molecules in keratocyte ECM could compensate for the loss of intact TSP-1. In this regard, an important difference between primary keratocytes and the other cell lines studied is that HT1080 and 293-EBNA cells do not produce any fibrillar collagens, whereas these collagens are known to represent the main anchoring molecules for keratocytes in the corneal environment. We can therefore conclude that, in the case of keratocytes, the pro-adhesive signals that are also mediated by BMP-1 (for example, through the promotion of collagen fibrillogenesis) largely outweigh the anti-adhesive signals potentially elicited by TSP-1 cleavage (Fig. 9). An additional level of complexity is brought by the fact that TSP-1 is also known to be required for proper collagen fibrillogenesis. The mechanism has been partially elucidated and involves TSP-1 interaction with fibrillar procollagens and pro-lysyl oxidase (pro-LOX), leading to the inhibition of LOX-mediated cross-linking in collagen fibers (69). Because procollagens and pro-LOX also undergo proteolytic maturation by BMP-1, we can hypothesize that TSP-1 inhibits LOX processing through competition for the BMP-1 active site. Inversely, we can expect that TSP-1 cleavage will be inhibited in the presence of pro-LOX and procollagens. However, the spatio-temporal regulation of the various proteolytic events to obtain correctly assembled and cross-linked fibers is not known and needs further investigation.

In parallel to the cell-dependent effect of TSP-1 cleavage on cell adhesion, we demonstrate that TSP-1 proteolysis by BMP-1 strongly enhances TGF- $\beta$  activity. The mechanism by which TSP-1 converts LTGF- $\beta$  into active TGF- $\beta$  has been well studied (49). This involves the disruption of specific noncovalent interactions between the mature TGF- $\beta$  growth factor and its LAP propeptide through competing

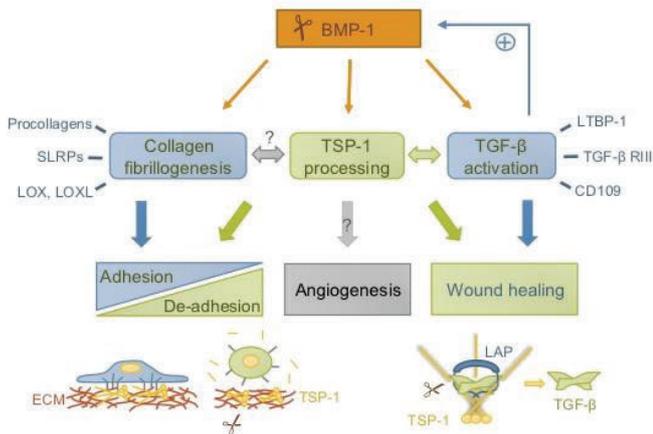


The data show the percentages of cells at the wound border that are myofibroblasts in each condition, expressed as means  $\pm$  SD [cornea numbers as in (E); three to four fields per cornea were used for quantification]. \*\*,  $P < 0.01$  and \*\*\*\*,  $P < 0.0001$ .

**Fig. 8. TSP-1 processing may play a role in corneal wound healing.** (A) Immunoblotting for TSP-1 fragment 1 under nonreducing conditions using MAB3074 in lysates of  $n = 4$  corneas from patients with permanent corneal scars. Volumes of corneal lysates loaded in each lane were adjusted to equivalent total protein amounts. Various amounts of keratocyte conditioned medium (CM) corresponding to 0.2 to 1.5  $\mu\text{g}$  of total protein were also loaded on the gel to facilitate the identification of fragment 1 in the human samples. Before detection, the membrane was split into two parts, and two different exposure times were used to reveal intact TSP-1 and fragment 1. Blot is representative of three technical replicates using the same samples. (B) Schematic representation of the ex vivo murine cornea wound healing experiments. IF, immuno-fluorescence. (C) Micrographs of the corneas taken at day 1 (d1), d4, and d7 after healing in the presence of vehicle (condition 1), intact TSP-1 (condition 2), BMP-1 (condition 3), cleaved TSP-1 (condition 4), cleaved TSP-1 plus the TGF- $\beta$  type I receptor inhibitor SB431542 (condition 5), or TGF- $\beta$  (condition 6). Scale bar, 500  $\mu\text{m}$ . Images are representative of  $n = 8$  to 9 corneas per condition except for condition 5 ( $n = 5$ ). (D) Evolution of wound contraction measured over time presented as percentage of initial wound area. The graph shows means  $\pm$  SD of  $n = 5$  corneas for condition 5 and  $n = 8$  to 9 corneas for other conditions. Condition 4 was compared to other conditions using one-way ANOVA and Dunnett's multiple comparison test. \* $P < 0.05$  and \*\* $P < 0.01$ . (E) Immunofluorescence showing  $\alpha$ -SMA (red) at the wound borders at the end of the experiment. Nuclei are labeled with DAPI (blue). Tissue sections were localized as shown with dashed arrows in (C). A full view of a cornea treated with vehicle is shown in the top panel with the wound (w) marked. Examples of regions used to count myofibroblasts are shown below for each condition. Scale bars, 100  $\mu\text{m}$ . Images are representative of  $n = 8$  to 9 corneas per condition for conditions 1 to 4,  $n = 5$  corneas for condition 5, and  $n = 6$  corneas for condition 6. (F) Quantification of  $\alpha$ -SMA-positive cells in the same wounded corneas.

interactions between the type 1 repeats in TSP-1 and both components of the LTGF- $\beta$  complex. In particular, the KRFK sequence located at the beginning of the second TSR can efficiently replace the RPKK sequence in TGF- $\beta$  to interact with the LSKL sequence in LAP, thereby leading to the exposure of TGF- $\beta$ , which otherwise remains buried within the LTGF- $\beta$  complex (49, 70). This mechanism implies a tight interaction between TSP-1 and LTGF- $\beta$  well illustrated by the fact that the TSP-1 purified from human platelets always contains some contaminating LTGF- $\beta$ . In our cell-based assays, this bound LTGF- $\beta$  seemed to account for most of the observed induction of SMAD signaling. In contrast, purified fragment 1

displayed an enhanced ability to activate LTGF- $\beta$  but was completely deprived of bound TGF- $\beta$ . Thus, we hypothesize that TSP-1 cleavage potentiates the activity of TGF- $\beta$  by speeding up its release from all remaining interactions with TSP-1, thereby increasing its mobility and facilitating its access to signaling receptors. This represents another mechanism by which BMP-1 enhances TGF- $\beta$  signaling, to be added to the previously reported cleavages of LTBP-1, TGF- $\beta$  receptor III (also known as betaglycan), and CD109 (31, 32). These mechanisms can all synergize to increase the pool and bio-availability of active TGF- $\beta$ , and we can expect major consequences for the control of tumor progression, immunity, and wound healing (Fig. 9).



**Fig. 9. TSP-1 is at the crossroads of two main pathways regulated by BMP-1, cell-matrix interactions, and growth factor signaling.** BMP-1 promotes cell-dependent effects on adhesion and TGF- $\beta$  activity. In turn, TGF- $\beta$  signaling can increase BMP-1 expression and regulate immune responses, ECM synthesis, and wound contraction during wound healing. The effect of TSP-1 processing on collagen fibrillogenesis and angiogenesis remains to be explored. SLRP, small leucine-rich proteoglycan.

The potentiation of TGF- $\beta$  signaling by cleaved TSP-1 is common to HT1080 cells and keratocytes, and our results indicate that it can play a role in the differentiation of the latter into myofibroblasts. This is of great importance in the context of corneal wound healing, where it was previously reported that TSP-1 production precedes  $\alpha$ -SMA induction and could regulate myofibroblast recruitment into the wound (50). After wound closure, stromal remodeling remains active until myofibroblasts disappear from the wound, a process that can take several years and lead to scarring. The control of myofibroblast differentiation is therefore a major issue to prevent the formation of irreversible scars. Here, we demonstrate that the capacity of TSP-1 to induce the differentiation of keratocytes into myofibroblasts is substantially enhanced after cleavage by BMP-1, both in 2D cultures and in the 3D environment of the cornea. This adds another important role for this proteinase in corneal wounds and scars where its abundance is known to be strongly increased and it is found colocalized with  $\alpha$ -SMA (10). Our present results show that a BMP-1-dependent processed form of TSP-1 is also found in human corneal scars, therefore providing a mechanistic hypothesis to explain how BMP-1 can drive pathological healing through the regulated proteolysis of TSP-1. Despite the numerous proteases known to process TSP-1, only the BMP-1-mediated cleavage product was detected by immunoblotting in human samples. However, it remains to be established whether this catabolic fragment is specific to corneal wounds or scars given that it is also produced by normal keratocytes in culture and could contribute to maintain tissue homeostasis.

The importance of regulated proteolysis, in contrast to degradation, to finely tune the properties and activities of ECM proteins has been demonstrated in several pathophysiological remodeling processes, including development, wound healing, fibrosis, and cancer. The unique location of the BMP-1 cleavage site in TSP-1, just before the type 1 repeats that mediate a large number of its activities, identifies another important regulatory mechanism to control cell adhesion and TGF- $\beta$  signaling in a concerted manner. This also confirms that proteolysis is an important aspect of TSP-1 regulation and estab-

lishes BMP-1 as a key molecular hub to coordinate the cross-talk between ECM-derived cues and growth factor signaling.

## MATERIALS AND METHODS

### Cell culture

HT1080 cells (ATCC CCL-121) were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS) and gentamycin (50  $\mu$ g/ml). HT1080 cells, stably transfected with a piRESneo2 vector encoding human BMP-1 fused with a 6His tag at the C terminus or the same construct with the E94A mutation, were obtained and grown in DMEM with 10% FCS and G418 sulfate (250  $\mu$ g/ml), as described previously (31). 293-EBNA cells overexpressing human native BMP-1 or human BMP-1 fused to a C-terminal Flag-tag were also available from previous studies (37, 71). A control 293-EBNA cell line transfected with the empty pCEP4 vector was produced using the same protocol. Primary keratocytes from human corneas were isolated as previously described (31), in accordance with ethical regulations, and grown in DMEM/Ham's F12 (1:1), 10% iron-supplemented calf serum, basic fibroblast growth factor (5 ng/ml), and 1% antibiotic/antimycotic solution. All cells were verified for the absence of mycoplasma with the MycoAlert detection kit (Ozyme).

For conditioned medium analysis by Western blotting, the cells were extensively washed with PBS (phosphate-buffered saline) and maintained between 8 and 48 hours in serum-free medium. The medium was then collected, centrifuged at 200g for 5 min, supplemented with protease inhibitors (cOmplete, Roche), and stored at  $-80^{\circ}\text{C}$ . The ECM from the same cultures was also collected after solubilization in 0.325 M tris-HCl (pH 6.9), 25% glycerol, and 10% SDS. Before this step, cells were removed by adding 20 mM  $\text{NH}_4\text{OH}$  for 5 min and the matrix was washed several times with water. Equal amounts of total proteins (medium, determined by Bradford assay) or equal volumes (ECM) were loaded onto the gels for Western blot analyses.

### Proteins and inhibitors

Recombinant human BMP-1 and *Xenopus* Szi were produced and purified as previously described (40, 72). Recombinant human TGF- $\beta$ 1 (Ala<sup>279</sup>-Ser<sup>390</sup>) was from PeproTech (ref. 100-21), and human LTGF- $\beta$  (Leu<sup>30</sup>-Ser<sup>390</sup>) was from R&D Systems (ref. 299-LT-025/CF). TSP-1 was purified from human platelets, provided by the Etablissement Français du Sang (Lyon) in agreement with French regulations, according to the protocol developed by Yee *et al.* (73) with minor modifications. Briefly, after three washes in Baezinger A buffer [0.102 M NaCl, 3.9 mM  $\text{K}_2\text{HPO}_4$ , 3.9 mM  $\text{Na}_2\text{HPO}_4$ , 22 mM  $\text{NaH}_2\text{PO}_4$ , and 5.5 mM glucose (pH 6.3)], the platelet pellet was resuspended in Baezinger B buffer [15 mM tris, 0.14 M NaCl, and 5 mM glucose (pH 7.6)] with 2 mM  $\text{CaCl}_2$  (50 ml of buffer for about 3 ml of pellet). Resuspended platelets (50 ml) were then gently mixed for 2 min at room temperature with 50-U thrombin (GE Healthcare) and kept on ice for 5 to 10 min until complete aggregation. After centrifugation at 1400g for 5 min, 1 mM Pefabloc was added to block thrombin activity. The supernatant was further centrifuged at 35,000g for 20 min at  $4^{\circ}\text{C}$ , stored at  $4^{\circ}\text{C}$  for one night, and centrifuged again at 1400g for 20 min. This was followed by heparin Sepharose chromatography (6 Fast Flow, GE Healthcare) using a 0.15 M NaCl to 2 M NaCl gradient in 15 mM tris (pH 7.6) and 2 mM  $\text{CaCl}_2$ . TSP-1-containing fractions were then pooled, concentrated in an Amicon Ultra device

(cutoff, 100 kDa; Millipore), and injected on a Superdex S200 16/600 column equilibrated in 20 mM Tris (pH 11), 0.15 M NaCl, and 2 mM CaCl<sub>2</sub>. The pH of the major peak fractions was then rapidly decreased to 7.5 to 8.0, and the protein was concentrated, dialyzed against 20 mM HEPES (pH 7.4), 0.5 M NaCl, and 2.5 mM CaCl<sub>2</sub>, and flash-frozen in liquid nitrogen before storage at -80°C in LoBind microtubes (Eppendorf).

TSP-1 fragments 1 and 2 were generated by the cleavage of 400 to 700 µg of purified TSP-1 in the presence of an equimolar amount of recombinant BMP-1 in 50 mM HEPES (pH 7.4), 0.16 M NaCl, 5 mM CaCl<sub>2</sub>, and 0.02% *n*-octyl-β-D-glucopyranoside for 4 hours at 37°C. Immediately after the cleavage, the mixture was loaded on a HiTrap Heparin HP column (1 ml; GE Healthcare) and purified with a NaCl gradient (from 0.16 to 2 M). This yielded pure fragment 1 in the flow-through and enriched fragment 2 in the eluate. Useful fractions were concentrated and injected on a Superdex S200 16/600 gel filtration column equilibrated in 20 mM Hepes (pH 7.4), 0.5 M NaCl, and 2.5 mM CaCl<sub>2</sub>. After concentration, fragments were stored in the same conditions as full-length TSP-1. By SDS-PAGE, fragment 2 still contained full-length TSP-1 and BMP-1, around 20% each (fig. S7).

The BMP-1 hydroxamate inhibitor used in cell cultures was the *N*-substituted aryl sulfonamide hydroxamate 90 described in (74), also known as S33A, which was prepared and characterized as previously reported (31, 39). It was dissolved in dimethyl sulfoxide before addition to the culture medium. The TGF-β type I receptor inhibitor SB431542 was from Sigma-Aldrich.

#### Quantification of cell numbers and detachment

Cells were counted either manually (in a Thoma chamber) or with a Cellometer Auto T4 (Nexcelom). For the monitoring of 293-EBNA cell detachment over time, 7000 cells per well were seeded in 96-well plates and grown in DMEM, 10% FCS, and hygromycin (300 µg/ml). After 7 days, cells were washed and switched to serum-free medium. At each time point, cell supernatants were transferred to plates coated with poly-D-lysine. Both adherent and detached cells were fixed with 4% glutaraldehyde and stained with 0.1% crystal violet for quantification.

#### Immunoblotting

The percentages of acrylamide used for SDS-PAGE and immunoblotting were as follows: 6 or 8% for TSP-1, 10% for BMP-1 and phosphorylated SMAD2 (P-SMAD2), 8% for SMAD2 and SMAD3, and 4 to 20% for LAP and TGF-β. After SDS-PAGE, proteins were transferred to polyvinylidene difluoride (PVDF) membranes by liquid transfer (2 hours, 4°C, 250 mA) in 10 mM *N*-cyclohexyl-3-aminopropanesulfonic acid (pH 11) with 10% ethanol and blocked for 2 hours in 10% (w/v) skim milk in PBS. Primary antibodies were diluted in 5% skim milk and 0.05% Tween 20 in PBS and applied on membranes overnight at 4°C. After three washes in 0.05% Tween 20 in PBS, horseradish peroxidase (HRP)-coupled secondary antibodies (Cell Signaling Technology or Dako; dilutions 1:5000 or 1:10,000) were added for 1 hour at room temperature and signal was detected using the enhanced chemiluminescence technique with the ECL Select (GE Healthcare) or Clarity Max (Bio-Rad) kits and the FX Fusion charge-coupled device (CCD) camera (Vilber Lourmat).

Unless otherwise stated, TSP-1 detection was with the mouse monoclonal antibody MAB3074 from R&D Systems (1:1000). Other anti-TSP-1 antibodies included the mouse monoclonal antibodies

A4.1 (1:200) and C9 (1:100) from Santa Cruz Biotechnology, the mouse monoclonal antibody A6.1 (1:500) from eBioscience, and the rabbit polyclonal antibody AP8522a from Abgent (1:500). The rat monoclonal anti-human BMP-1 (MAB1927; 1:1000) antibody was from R&D Systems, the rabbit polyclonal anti-human collagen VI antibody (20611; 1:200) was from Novotec, the mouse monoclonal anti-actin (A3854; 1:3000) was from Sigma-Aldrich, and the rabbit polyclonal anti-SMAD2/3 (3102; 1:1000) and anti-P-SMAD2 (Ser<sup>465</sup>/Ser<sup>467</sup>; 3108; 1:1000) antibodies were from Cell Signaling Technology.

#### Microscopy

Phase contrast images were taken with a Nikon TE300 microscope, a Nikon Ti-E microscope, or a Zeiss Axioplan microscope equipped with a differential interference contrast module. For videomicroscopy experiments, 40,000 cells per well of each cell type were seeded in a six-well plate and grown for 24 hours in DMEM containing 10% FCS. Images were taken every 15 min for 17 hours.

For cell area measurements, 5000 cells per well were seeded in eight-well µ-slides (Ibidi) and grown for 3 days in complete growth medium. Then, cells were fixed with 4% paraformaldehyde and stained with phalloidin-FITC (fluorescein isothiocyanate) and DAPI (4',6'-diamidino-2-phenylindole). Images were acquired using a Zeiss Axio Observer Z1 microscope equipped with a 20× 0.8 NA (numerical aperture) objective or a Zeiss LSM 710 microscope equipped with a 63× 1.4 NA objective. To extract the mean cell area, the total surface covered by cells in each field was revealed by actin staining and measured using the Fiji software (75). This surface was then divided by the number of nuclei detected by DAPI staining in the field.

For extracellular TSP-1 and fibronectin immunostaining, cells were grown on coverslips for 3 days and fixed with 4% paraformaldehyde. No permeabilization was performed. Cells were incubated with MAB3074 anti-TSP-1 (1:50) or A0245 anti-fibronectin (Dako; 1:500) primary antibodies, and immunostaining was revealed with goat anti-mouse or anti-rabbit cyanine 3 secondary antibodies (115-165-146 and 111-165-144, respectively, from Jackson ImmunoResearch; 1:1000). Nuclei were stained with DAPI. Fluorescence images were taken using a Zeiss Axio Observer Z1 microscope equipped with an Orca-R2 CCD camera (Hamamatsu Photonics). For fibronectin images, maximum intensity projections of stacks, generated using Fiji, are shown.

For the staining of α-SMA in human keratocytes, 50,000 cells per well were seeded in six-well plates in complete growth medium. After 1 day, cells were washed with PBS and treated for 72 hours with TGF-β, intact TSP-1, cleaved TSP-1, or vehicle solutions in serum-free medium. Then, cells were fixed and permeabilized with cold methanol. Immunodetection was carried out with anti-α-SMA (Ab5694 from Abcam; 1:500) and anti-rabbit Alexa Fluor 594 (A-11037 from Thermo Fisher Scientific; 1:1500) antibodies. Nuclei were stained with DAPI. Fluorescence images were taken using a Zeiss Axioplan 2 microscope, and α-SMA-positive cells (with clear organized fibers) were counted using Fiji in 20 different fields per condition.

#### Flow cytometry

HT1080 cells (150,000 or 500,000) were seeded in 100-mm culture dishes and grown for 48 hours to obtain around 20 and 50% confluency. Cells were then detached with Accutase (Sigma-Aldrich) and washed with PBS and with the binding buffer provided in the annexin V-FITC apoptosis detection kit (BioVision). Cells (500,000) were resuspended in 500 µl of binding buffer, incubated for 5 min with 5 µl

of annexin V-FITC and 5  $\mu$ l of propidium iodide in the dark, and analyzed with a FACS LSR II apparatus (Becton Dickinson) equipped with the FACSDiva software.

### RNA interference

TSP-1 expression in HT1080 cells was decreased using validated siRNAs directed against human TSP-1 from Qiagen (Hs\_THBS1\_5 with target sequence 5'-TACGAATGTAGAGATCCCTAA and Hs\_THBS1\_7 with target sequence 5'-CGGCGTGAAGTGTACTAGCTA). A control siRNA against the decoy sequence 5'-AATTCTCCGAACGTGTCACGT was also used. One hundred twenty thousand or 300,000 HT1080 cells per well were seeded in 12- or 6-well plates with DMEM containing 10% FCS. Thirty minutes after seeding, transfection mixes containing siRNA, DMEM, and HiPerFect Transfection Reagent (Qiagen), prepared as recommended by the manufacturer, were added to the cells (25 nM final siRNA concentration). Pictures were taken every 24 hours for 3 days, and the medium was collected at day 3 (after 48 hours of serum starvation), either for Western blot analysis or for the quantification of detached cells. The remaining adherent cells were also counted after trypsinization and resuspension in PBS. For cell area measurements, 31,500 cells per well were seeded in eight-well  $\mu$ -slides (Ibidi; 1 cm<sup>2</sup>) in complete growth medium and transfected in the same conditions as above. After 24 hours, medium was changed to serum-free medium and cells were maintained in this medium for another day before fixation with 4% paraformaldehyde. Staining with phalloidin-FITC and DAPI, imaging, and cell area measurement were performed as described in the "Microscopy" section.

### In vitro cleavage assays and N-terminal sequencing

All cleavage assays were performed at 37°C in 40 to 200  $\mu$ l of 50 mM Hepes (pH 7.4), 0.15 M NaCl, 5 mM CaCl<sub>2</sub>, and 0.02% *n*-octyl- $\beta$ -D-glucopyranoside (buffer A). Incubation time and protease:substrate ratios were varied as indicated. Cleavage monitoring was performed by SDS-PAGE, and detection was performed with Coomassie Blue. For Edman N-terminal sequencing, 10  $\mu$ g of TSP-1 was incubated with and without 4  $\mu$ g of BMP-1 in 200  $\mu$ l of buffer A for 4 hours. The two samples were then processed for Edman sequencing on PVDF membranes and analyzed with an Applied Biosystems 473 protein sequencer.

For ATOMS experiments, half the amounts of BMP-1 and TSP-1 (keeping concentrations constant) were used. Samples were processed as described in (31), and peptides were analyzed with a Q Exactive HF tandem mass spectrometer (Thermo Scientific) coupled to an UltiMate 3000 Nano LC system (Dionex) (76). Briefly, peptides were first desalted online on a reverse-phase Acclaim PepMap 100 C18 precolumn [5  $\mu$ m, 100  $\text{Å}$ , 300  $\mu$ m internal diameter (id)  $\times$  5 mm] for 5 min at 2% CH<sub>3</sub>CN at a flow rate of 10  $\mu$ l/min and then resolved at a flow rate of 0.300  $\mu$ l/min with a 180-min gradient of CH<sub>3</sub>CN, 0.1% formic acid applied onto the reverse-phase Acclaim PepMap 100 C18 column (3  $\mu$ m, 100  $\text{Å}$ , 75  $\mu$ m id  $\times$  500 mm) from 4 to 40% CH<sub>3</sub>CN. The Q Exactive HF instrument was operated according to a Top20 data-dependent method consisting of a scan cycle initiated with a full scan of peptide ions in the ultrahigh-field Orbitrap analyzer, followed by selection of the precursor with a fixed first mass parameter set at 80 *m/z* (mass/charge ratio), high-energy collisional dissociation with a normalized energy collision fixed at 33 eV, and tandem mass spectrometry (MS/MS) scans on the 20 most abundant precursor ions. Full-scan mass spectra were acquired from *m/z* 350 to 1800 with an

automatic gain control (AGC) target set at  $3 \times 10^6$  ions and a resolution of 60,000. MS/MS scans were initiated at a resolution of 15,000 when the ACG target reached  $10^5$  ions with a threshold intensity of  $10^5$  and potential charge states of 2<sup>+</sup> and 3<sup>+</sup> after ion selection performed with a dynamic exclusion of 10 s.

### Adhesion assays

Conditioned ECMs were prepared in 96-well plates from subconfluent cell layers. Cells were washed with PBS, detached using 20 mM NH<sub>4</sub>OH, and discarded. The remaining ECM layer was washed with PBS, incubated with 1% BSA (bovine serum albumin) in PBS, and washed again. Before adhesion assays, cells from additional culture plates were detached with 5 mM EDTA in PBS (HT1080 and HT1080-BMP-1) or collected after spontaneous detachment (HT1080-BMP-1) and rinsed in serum-free medium. After centrifugation, cells were resuspended in serum-free medium and seeded (60,000 cells per well) on conditioned ECM. After 45 min at 37°C, nonadherent cells were washed with PBS and the extent of adhesion was determined after fixation of adherent cells, staining with 0.1% crystal violet and absorbance measurements at 570 nm, as described previously (77). A blank value corresponding to BSA-coated wells was subtracted. Alternatively, cells could be seeded in larger wells containing decellularized ECMs for phase contrast imaging.

To assess adhesion of HT1080 cells on intact and cleaved TSP-1, wells were coated with protein solutions overnight at 4°C and saturated with 1% BSA in PBS. HT1080 cells were prepared and detached from the plates with EDTA, as described above. For competition experiments with TSP-1 and its fragments, cells (40,000 cells per well) were mixed with proteins for 1 hour before cell adhesion onto TSP-1-coated wells (4  $\mu$ g of TSP-1 per well).

### Enzyme-linked immunosorbent assay

Ninety-six-well Costar culture-treated microtiter plates (also used for adhesion assays) were coated overnight at 4°C with 50  $\mu$ l of each protein solution diluted in TBS [10 mM tris-HCl (pH 8) and 0.15 M NaCl] with 5 mM CaCl<sub>2</sub>. Wells were then blocked with 5% (w/v) skim milk in TBS and incubated successively with the anti-TSP-1 antibodies MAB3074 (1:1000) or C9 (1:500) and horse anti-mouse antibody coupled to HRP (Cell Signaling Technology; 1:3000). 2,2'-Azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) ammonium salt was used as HRP substrate, and color development was measured at 405 nm after 15 min in a microtiter plate reader (Tecan Infinite M1000).

### Detection of TGF- $\beta$ activity

TGF- $\beta$  activity was monitored first through the immunodetection of P-SMAD2. HT1080 cells (100,000) or keratocytes (30,000) were grown in 12-well plates with 1 ml of complete medium for 48 hours. Solutions containing TSP-1 (beforehand incubated or not with active or inactive BMP-1) and control conditions (in 100  $\mu$ l) were then added to the cells for 30 min (keratocytes) or 1 hour (HT1080 cells). For the experiments with LTGF- $\beta$ , cells were first washed with PBS and the medium was replaced by DMEM without serum before the addition of LTGF- $\beta$  (17.5 nM, preincubated for 15 min at 37°C with a 3 molar excess of TSP-1 or fragment 1 or at room temperature with 10 mM HCl) or control solutions. At the end of the cell treatment time, cell supernatants were discarded, cells were washed with cold PBS, and cell lysates were collected in radioimmunoprecipitation assay buffer [50 mM tris (pH 8), 0.15 M NaCl, 1% NP-40,

0.5% sodium deoxycholate, and 0.1% SDS] containing protease (cOmplete, Roche) and phosphatase (PhosSTOP, Roche) inhibitors. Twenty micrograms of total protein per lane (Bradford assay) were loaded on gels for immunoblotting for P-SMAD2, SMAD2, or actin.

TGF- $\beta$  activity was also assessed through a reporter gene assay, as previously described (31). In this experiment, 50,000 HT1080 cells were seeded in 24-well plates with DMEM and 10% FCS. After 24 hours, they were cotransfected with (i) the pGL3-basic reporter plasmid (Promega) encoding firefly luciferase under the control of nine repeats of a SMAD-binding element [(CAGA)<sub>9</sub>-Luc] and (ii) the pRLCMV vector (Promega) encoding *Renilla* luciferase under the control of the cytomegalovirus ubiquitous promoter to determine the transfection efficiency and to normalize luciferase values. The following day, samples containing TSP-1 (incubated or not with active or inactive BMP-1) and control conditions were added for 24 hours. Proteins were then extracted from HT1080 cells in Passive Lysis Buffer (Promega), and transcriptional responses were measured with the Dual-Luciferase Reporter Assay System (Promega).

### Human corneas

Four corneal samples from patients presenting permanent stromal scars were collected at the time of corneal transplantation, according to the tenets of the declaration of Helsinki. Patient 1 (male, age 23) had his cornea perforated by broken glass 17 months before surgery. Patient 2 (male, age 66) had a handiwork accident 6 months before surgery. Patient 3 (male, age 68) suffered from a corneal ulcer for 2 years. Patient 4 (male, age 50) was injured by a perforating metal foreign body more than 2 years before. Corneas were frozen at  $-80^{\circ}\text{C}$  after surgery before being processed for immunoblotting. Briefly, frozen corneas were ground in a mortar, homogenized in Laemmli buffer, and boiled at  $100^{\circ}\text{C}$  for 10 min. After centrifugation at  $15,000g$  for 15 min, the supernatant was collected and analyzed on a stain-free 4 to 20% TGX precast gel (Bio-Rad) to allow total protein relative quantification. Equivalent amounts of total protein for each human cornea were then analyzed by immunoblotting and compared to known amounts of keratocyte conditioned medium. Before detection, the membrane was split into two parts and two different exposure times were used to reveal full-length TSP-1 (7 s) and fragment 1 (4 min).

### Mouse ex vivo model of corneal wound healing

Mouse corneas were dissected from 13-week-old male C57BL/6J mice (Janvier Labs). After eye washing with 5% povidone-iodine alcoholic solution (MEDA Pharma) and 0.09% NaCl (Aguettant), dissection of mouse corneas was performed with an ophthalmic knife and Vannas scissors under a binocular magnifier. Corneas were then washed in four baths of culture medium without serum [DMEM/Ham's F12 (1:1), sodium ascorbate (50  $\mu\text{g}/\text{ml}$ ), and 2% antibiotic/antimycotic solution] and wounded in the center with disposable 1-mm-diameter biopsy punches (Integra Miltenyi). Wounded corneas were distributed in the wells of a 96-well plate containing 200  $\mu\text{l}$  of the same medium as above with 10% iron-supplemented calf serum and maintained in a humid atmosphere at  $37^{\circ}\text{C}$  with 5%  $\text{CO}_2$ . At days 2 and 4, cleaved TSP-1 was prepared by incubating 450 nM TSP-1 with 356 nM BMP-1 for 4 hours at  $37^{\circ}\text{C}$  in the buffer described above for all cleavage assays ("vehicle"). Controls with TSP-1 alone and BMP-1 alone were also prepared. Cornea medium was changed on days 2 and 4 with a progressive reduction in serum concentration to 2% (day 2) and 0% (day 4) and

incubations were added to represent 10% of the volume in each well (45 nM final TSP-1 concentration). Additional corneas were also treated with TGF- $\beta$  (10 ng/ml) or cleaved TSP-1 with 10  $\mu\text{M}$  SB431542. Bright-field images were taken daily with a Zeiss Primovert or Nikon Ti-E microscope and used to measure the diameter of the wound with Fiji. On day 7, the corneas were fixed overnight in buffered 4% paraformaldehyde and embedded in paraffin wax. Two sections of 5  $\mu\text{m}$  per cornea corresponding to two opposite wound borders were prepared (see Fig. 8C) and stained with the anti- $\alpha$ -SMA antibody Ab5694 (Abcam; 1:100 in 3% BSA) after de-waxing and rehydration. The secondary antibody was a goat anti-rabbit antibody coupled to Alexa Fluor 555 (A-21428 from Thermo Scientific; 1:350 in 0.2% BSA). Corneal sections were counterstained with DAPI and examined with a Nikon Ti-E fluorescence microscope.  $\alpha$ -SMA-positive cells were counted in three to four fields per cornea.

### Statistical analysis

All statistical analyses were performed using the GraphPad Prism 7 or 8 software. Unless otherwise stated, experimental values for the different conditions were compared using one-way analysis of variance (ANOVA) and Tukey post-ANOVA test. Data were tested for normality with the Shapiro-Wilk test and for equality of variances with the Brown-Forsythe test. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , and \*\*\*\* $P < 10^{-4}$ .

### SUPPLEMENTARY MATERIALS

stke.sciencemag.org/cgi/content/full/13/639/eaba3880/DC1

Fig. S1. Effect of BMP-1 on cell adhesion and apoptosis.

Fig. S2. BMP-1 alters cell adhesion through ECM modification.

Fig. S3. Characterization of TSP-1 fragments with various antibodies against the C-terminal and N-terminal domains of TSP-1.

Fig. S4. Mass spectrum and attribution of b and y ions for the peptide spanning residues 376 to 402 of human platelet TSP-1, defining BMP-1 cleavage site.

Fig. S5. Quantification of TSP-1 knockdown efficiency.

Fig. S6. Efficiency of coating 96-well plates with intact and cleaved TSP-1.

Fig. S7. SDS-PAGE analysis of purified fragment 1 and semipurified fragment 2 in nonreducing and reducing conditions.

Fig. S8. Alignment of TSP-1 from 12 species and human TSP-2 around the BMP-1 cleavage site.

Movie S1. Time-lapse recording of subconfluent HT1080 cells.

Movie S2. Time-lapse recording of the morphological changes preceding cell detachment in subconfluent HT1080-BMP-1 cells.

References (78, 79)

View/request a protocol for this paper from *Bio-protocol*.

### REFERENCES AND NOTES

1. S. A. Goodman, R. Albano, F. C. Wardle, G. Matthews, D. Tannahill, L. Dale, BMP1-related metalloproteinases promote the development of ventral mesoderm in early *Xenopus* embryos. *Dev. Biol.* **195**, 144–157 (1998).
2. G. Marques, M. Musacchio, M. J. Shimell, K. Wunnenberg-Stapleton, K. W. Cho, M. B. O'Connor, Production of a DPP activity gradient in the early *Drosophila* embryo through the opposing actions of the SOG and TLD proteins. *Cell* **91**, 417–426 (1997).
3. N. Suzuki, P. A. Labosky, Y. Furuta, L. Hargett, R. Dunn, A. B. Fogo, K. Takahara, D. M. Peters, D. S. Greenspan, B. L. Hogan, Failure of ventral body wall closure in mouse embryos lacking a procollagen C-proteinase encoded by *Bmp1*, a mammalian gene related to *Drosophila* *tolloid*. *Development* **122**, 3587–3595 (1996).
4. T. G. Clark, S. J. Conway, I. C. Scott, P. A. Labosky, G. Winnier, J. Bundy, B. L. Hogan, D. S. Greenspan, The mammalian *Tolloid*-like 1 gene, *Tll1*, is necessary for normal septation and positioning of the heart. *Development* **126**, 2631–2642 (1999).
5. D. Syx, B. Guillemin, S. Symoens, A. B. Sousa, A. Medeira, M. Whiteford, T. Hermans-Le, P. J. Coucke, A. De Paepe, F. Malfait, Defective proteolytic processing of fibrillar procollagens and prodecorin due to biallelic BMP1 mutations results in a severe, progressive form of osteogenesis imperfecta. *J. Bone Miner. Res.* **30**, 1445–1456 (2015).

6. A. M. Muir, D. Massoudi, N. Nguyen, D. R. Keene, S. J. Lee, D. E. Birk, J. M. Davidson, M. P. Marinkovich, D. S. Greenspan, BMP1-like proteinases are essential to the structure and wound healing of skin. *Matrix Biol.* **56**, 114–131 (2016).
7. A. M. Muir, Y. Ren, D. H. Butz, N. A. Davis, R. D. Blank, D. E. Birk, S. J. Lee, D. Rowe, J. Q. Feng, D. S. Greenspan, Induced ablation of Bmp1 and Tll1 produces osteogenesis imperfecta in mice. *Hum. Mol. Genet.* **23**, 3085–3101 (2014).
8. J. Wang, D. Massoudi, Y. Ren, A. M. Muir, S. E. Harris, D. S. Greenspan, J. Q. Feng, BMP1 and TLL1 are required for maintaining periodontal homeostasis. *J. Dent. Res.* **96**, 578–585 (2017).
9. N. M. Wolfman, A. C. McPherron, W. N. Pappano, M. V. Davies, K. Song, K. N. Tomkinson, J. F. Wright, L. Zhao, S. M. Sebald, D. S. Greenspan, S. J. Lee, Activation of latent myostatin by the BMP-1/tolloid family of metalloproteinases. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15842–15846 (2003).
10. F. Malecaze, D. Massoudi, P. Fournie, C. Tricoire, M. Cassagne, M. Malbouyres, D. J. Hulmes, C. Moali, S. D. Galiacy, Upregulation of bone morphogenetic protein-1/mammalian tolloid and procollagen C-proteinase enhancer-1 in corneal scarring. *Invest. Ophthalmol. Vis. Sci.* **55**, 6712–6721 (2014).
11. E. Sanchez-Sabate, L. Alvarez, E. Gil-Garay, L. Munuera, N. Vilaboa, Identification of differentially expressed genes in trabecular bone from the iliac crest of osteoarthritic patients. *Osteoarthr. Cartil.* **17**, 1106–1114 (2009).
12. J. Beaumont, B. Lopez, N. Hermida, B. Schroen, G. San Jose, S. Heymans, F. Valencia, J. J. Gomez-Doblas, E. De Teresa, J. Diez, A. Gonzalez, microRNA-122 down-regulation may play a role in severe myocardial fibrosis in human aortic stenosis through TGF- $\beta$ 1 up-regulation. *Clin. Sci.* **126**, 497–506 (2014).
13. J. Friebe, A. Weithauser, M. Witkowski, B. H. Rauch, K. Savvatis, A. Dömer, T. Tabaraie, M. Kasner, V. Moos, D. Bosel, M. Gotthardt, M. H. Radke, M. Wegner, P. Bobbert, D. Lassner, C. Tschöpe, H. P. Schutheiss, S. B. Felix, U. Landmesser, U. Rauch, Protease-activated receptor 2 deficiency mediates cardiac fibrosis and diastolic dysfunction. *Eur. Heart J.* **40**, 3318–3332 (2019).
14. B. Lopez, A. Gonzalez, J. Beaumont, R. Querejeta, M. Larman, J. Diez, Identification of a potential cardiac antifibrotic mechanism of torasemide in patients with chronic heart failure. *J. Am. Coll. Cardiol.* **50**, 859–867 (2007).
15. A. Borazjani, N. Kow, S. Harris, B. Ridgeway, M. S. Damaser, Transcriptional regulation of connective tissue metabolism genes in women with pelvic organ prolapse. *Female Pelvic Med. Reconstr. Surg.* **23**, 44–52 (2017).
16. J. Fuentes-Duculan, K. M. Bonifacio, M. Suarez-Farinas, N. Kunjraiva, S. Garcet, T. Cruz, C. Q. F. Wang, H. Xu, P. Gilleadeau, M. Sullivan-Whalen, M. H. Tirgan, J. G. Krueger, Aberrant connective tissue differentiation towards cartilage and bone underlies human keloids in African Americans. *Exp. Dermatol.* **26**, 721–727 (2017).
17. P. Perimenis, T. Bouckenoghe, J. Delplanque, E. Moitrot, E. Eury, S. Lobbens, P. Gosset, L. Devisme, B. Duvillie, A. Abderrahmani, L. Storme, P. Fontaine, P. Froguel, A. Vamberger, Placental antiangiogenic prolactin fragments are increased in human and rat maternal diabetes. *Biochim. Biophys. Acta* **1842**, 1783–1793 (2014).
18. A. Naba, K. R. Clauser, J. M. Lamar, S. A. Carr, R. O. Hynes, Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *eLife* **3**, e01308 (2014).
19. R. da Silva, M. Uno, S. K. Marie, S. M. Oba-Shinjo, LOX expression and functional analysis in astrocytomas and impact of IDH1 mutation. *PLOS ONE* **10**, e0119781 (2015).
20. Y. Y. Hsieh, S. Y. Tung, H. Y. Pan, C. W. Yen, H. W. Xu, Y. F. Deng, Y. J. Lin, W. T. Hsu, C. S. Wu, C. Li, Upregulation of bone morphogenetic protein 1 is associated with poor prognosis of late-stage gastric cancer patients. *BMC Cancer* **18**, 508 (2018).
21. W. Xiao, X. Wang, T. Wang, J. Xing, Overexpression of BMP1 reflects poor prognosis in clear cell renal cell carcinoma. *Cancer Gene Ther.* (2019).
22. J. S. Bond, R. J. Beynon, The astacin family of metalloendopeptidases. *Protein Sci.* **4**, 1247–1261 (1995).
23. R. Berry, T. A. Jowitz, J. Ferrand, M. Roessle, J. G. Grossmann, E. G. Canty-Laird, R. A. Kammerer, K. E. Kadler, C. Baldock, Role of dimerization and substrate exclusion in the regulation of bone morphogenetic protein-1 and mammalian tolloid. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8561–8566 (2009).
24. C. Wemter, M. Howel, V. Hintze, B. Bombosch, K. Aufenvenne, I. Yiallourou, W. Stocker, The protease domain of procollagen C-proteinase (BMP1) lacks substrate selectivity, which is conferred by non-proteolytic domains. *Biol. Chem.* **388**, 513–521 (2007).
25. C. Moali, D. J. Hulmes, in *Extracellular Matrix: Pathobiology and Signaling*, M. Karamanos, Ed. (De Gruyter, 2012), pp. 539–561.
26. S. Vadon-Le Goff, D. J. Hulmes, C. Moali, BMP-1/tolloid-like proteinases synchronize matrix assembly with growth factor activation to promote morphogenesis and tissue remodeling. *Matrix Biol.* **44–46C**, 14–23 (2015).
27. P. Blader, S. Rastegar, N. Fischer, U. Strähle, Cleavage of the BMP-4 antagonist chordin by zebrafish tolloid. *Science* **278**, 1937–1940 (1997).
28. G. Ge, D. R. Hopkins, W. B. Ho, D. S. Greenspan, GDF11 forms a bone morphogenetic protein 1-activated latent complex that can modulate nerve growth factor-induced differentiation of PC12 cells. *Mol. Cell. Biol.* **25**, 5846–5858 (2005).
29. B. Kim, G. Huang, W. B. Ho, D. S. Greenspan, Bone morphogenetic protein-1 processes insulin-like growth factor-binding protein 3. *J. Biol. Chem.* **286**, 29014–29025 (2011).
30. G. Szlama, V. Vasarhelyi, M. Trexler, L. Patthy, Influence of WFIKKN1 on BMP1-mediated activation of latent myostatin. *FEBS J.* **283**, 4515–4527 (2016).
31. F. Delolme, C. Anastasi, L. B. Alcaraz, V. Mendoza, S. Vadon-Le Goff, M. Talantikite, R. Capomaccio, J. Mevaere, L. Fortin, D. Mazzocut, O. Damour, I. Zanella-Cleon, D. J. Hulmes, C. M. Overall, U. Valcourt, F. Lopez-Casillas, C. Moali, Proteolytic control of TGF- $\beta$  co-receptor activity by BMP-1/tolloid-like proteases revealed by quantitative iTRAQ proteomics. *Cell. Mol. Life Sci.* **72**, 1009–1027 (2015).
32. G. Ge, D. S. Greenspan, BMP1 controls TGF $\beta$ 1 activation via cleavage of latent TGF $\beta$ -binding protein. *J. Cell Biol.* **175**, 111–120 (2006).
33. S. Banerjee, R. J. Andrew, C. J. Duff, K. Fisher, C. D. Jackson, C. B. Lawrence, N. Maeda, D. S. Greenspan, K. A. B. Kellett, N. M. Hooper, Proteolysis of the low density lipoprotein receptor by bone morphogenetic protein-1 regulates cellular cholesterol uptake. *Sci. Rep.* **9**, 11416 (2019).
34. P. Chau, P. E. Fielding, C. J. Fielding, Bone morphogenetic protein-1 (BMP-1) cleaves human proapolipoprotein A1 and regulates its activation for lipid binding. *Biochemistry* **46**, 8445–8450 (2007).
35. S. E. Heumüller, M. Talantikite, M. Napoli, J. Armengaud, M. Morgelin, U. Hartmann, G. Sengle, M. Paulsson, C. Moali, R. Wagener, C-terminal proteolysis of the collagen VI  $\alpha$ 3 chain by BMP-1 and proprotein convertase(s) releases endotrophin in fragments of different sizes. *J. Biol. Chem.* **294**, 13769–13780 (2019).
36. L. Garrigue-Antar, N. Hartigan, K. E. Kadler, Post-translational modification of bone morphogenetic protein-1 is required for secretion and stability of the protein. *J. Biol. Chem.* **277**, 43327–43334 (2002).
37. C. Moali, B. Font, F. Ruggiero, D. Eichenberger, P. Rousselle, V. Francois, A. Oldberg, L. Bruckner-Tuderman, D. J. Hulmes, Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1. *J. Biol. Chem.* **280**, 24188–24194 (2005).
38. I. C. Scott, I. L. Blitz, W. N. Pappano, Y. Imamura, T. G. Clark, B. M. Steigltz, C. L. Thomas, S. A. Maas, K. Takahara, K. W. Cho, D. S. Greenspan, Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. *Dev. Biol.* **213**, 283–300 (1999).
39. M. Talantikite, P. Lecorche, F. Beau, O. Damour, C. Becker-Pauly, W. B. Ho, V. Dive, S. Vadon-Le Goff, C. Moali, Inhibitors of BMP-1/tolloid-like proteinases: Efficacy, selectivity and cellular toxicity. *FEBS Open Bio* **8**, 2011–2021 (2018).
40. C. Bijakowski, S. Vadon-Le Goff, F. Delolme, J. M. Bourhis, P. Lecorche, F. Ruggiero, C. Becker-Pauly, I. Yiallourou, W. Stocker, V. Dive, D. J. Hulmes, C. Moali, Sizzled is unique among secreted frizzled-related proteins for its ability to specifically inhibit bone morphogenetic protein-1 (BMP-1)/tolloid-like proteinases. *J. Biol. Chem.* **287**, 33581–33593 (2012).
41. H. X. Lee, A. L. Ambrosio, B. Reversade, E. M. DeRobertis, Embryonic dorsal-ventral signaling: Secreted frizzled-related proteins as inhibitors of tolloid proteinases. *Cell* **124**, 147–159 (2006).
42. D. S. Annis, J. E. Murphy-Ullrich, D. F. Mosher, Function-blocking antithrombospondin-1 monoclonal antibodies. *J. Thromb. Haemost.* **4**, 459–468 (2006).
43. A. Doucet, C. M. Overall, Broad coverage identification of multiple proteolytic cleavage site sequences in complex high molecular weight proteins using quantitative proteomics as a complement to edman sequencing. *Mol. Cell. Proteomics* **10**, M110.003533 (2011).
44. J. Hofsteenge, K. G. Huwiler, B. Macek, D. Hess, J. Lawler, D. F. Mosher, J. Peter-Katalinic, C-mannosylation and O-fucosylation of the thrombospondin type 1 module. *J. Biol. Chem.* **276**, 6485–6498 (2001).
45. J. C. Adams, J. Lawler, The thrombospondins. *Cold Spring Harb. Perspect. Biol.* **3**, a009712 (2011).
46. C. A. Elzie, J. E. Murphy-Ullrich, The N-terminus of thrombospondin: The domain stands apart. *Int. J. Biochem. Cell Biol.* **36**, 1090–1101 (2004).
47. J. C. Adams, A. A. Bentley, M. Kvasakul, D. Hatherley, E. Hohenester, Extracellular matrix retention of thrombospondin 1 is controlled by its conserved C-terminal region. *J. Cell Sci.* **121**, 784–795 (2008).
48. D. J. Kim, E. D. Christofidou, D. R. Keene, M. Hassan Milde, J. C. Adams, Intermolecular interactions of thrombospondins drive their accumulation in extracellular matrix. *Mol. Biol. Cell* **26**, 2640–2654 (2015).
49. J. E. Murphy-Ullrich, M. J. Suto, Thrombospondin-1 regulation of latent TGF- $\beta$  activation: A therapeutic target for fibrotic disease. *Matrix Biol.* **68–69**, 28–43 (2018).
50. M. Matsuba, A. E. Hutcheon, J. D. Zieske, Localization of thrombospondin-1 and myofibroblasts during corneal wound repair. *Exp. Eye Res.* **93**, 534–540 (2011).
51. H. Janin-Manificat, M. R. Rovere, S. D. Galiacy, F. Malecaze, D. J. Hulmes, C. Moali, O. Damour, Development of ex vivo organ culture models to mimic human corneal scarring. *Mol. Vis.* **18**, 2896–2908 (2012).
52. N. J. Laping, E. Grygielko, A. Mathur, S. Butter, J. Bomberger, C. Tweed, W. Martin, J. Fornwald, R. Lehr, J. Harling, L. Gaster, J. F. Callahan, B. A. Olson, Inhibition

- of transforming growth factor (TGF)-beta1-induced extracellular matrix with a novel inhibitor of the TGF-beta type I receptor kinase activity: SB-431542. *Mol. Pharmacol.* **62**, 58–64 (2002).
53. B. Barrientz, S. E. Nicholas, A. Whelchel, R. Sharif, J. Hjortdal, D. Karamichos, Corneal injury: Clinical and molecular aspects. *Exp. Eye Res.* **186**, 107709 (2019).
  54. J. Dubail, F. Kesteloot, C. Deroanne, P. Motte, V. Lambert, J. M. Rakic, C. Lapiere, B. Nusgens, A. Colige, ADAMTS-2 functions as anti-angiogenic and anti-tumoral molecule independently of its catalytic activity. *Cell. Mol. Life Sci.* **67**, 4213–4232 (2010).
  55. N. M. Moss, M. V. Barbolina, Y. Liu, L. Sun, H. G. Munshi, M. S. Stack, Ovarian cancer cell detachment and multicellular aggregate formation are regulated by membrane type 1 matrix metalloproteinase: A potential role in l.p. metastatic dissemination. *Cancer Res.* **69**, 7121–7129 (2009).
  56. H. Chen, M. E. Herndon, J. Lawler, The cell biology of thrombospondin-1. *Matrix Biol.* **19**, 597–614 (2000).
  57. J. C. Adams, J. Lawler, Diverse mechanisms for cell attachment to platelet thrombospondin. *J. Cell Sci.* **104**, 1061–1071 (1993).
  58. A. S. Asch, J. Tepler, S. Silbiger, R. L. Nachman, Cellular attachment to thrombospondin. Cooperative interactions between receptor systems. *J. Biol. Chem.* **266**, 1740–1745 (1991).
  59. D. D. Roberts, J. A. Sherwood, V. Ginsburg, Platelet thrombospondin mediates attachment and spreading of human melanoma cells. *J. Cell Biol.* **104**, 131–139 (1987).
  60. C. Y. Chen, E. Melo, P. Jakob, A. Friedlein, B. Elsasser, P. Goettig, V. Kueppers, F. Delobel, C. Stucki, T. Dunkley, S. Fauser, O. Schilling, R. Iacone, N-Terminomics identifies HtrA1 cleavage of thrombospondin-1 with generation of a proangiogenic fragment in the polarized retinal pigment epithelial cell model of age-related macular degeneration. *Matrix Biol.* **70**, 84–101 (2018).
  61. K. Seif, L. Alidzanovic, B. Tischler, N. Ibrahim, B. Zagraban, S. Rauscher, M. Salzmann, L. Hell, L. M. Mauracher, U. Budde, J. A. Schmid, B. Jilma, I. Pabinger, A. Assinger, P. Starlinger, C. Brostjan, Neutrophil-mediated proteolysis of thrombospondin-1 promotes platelet adhesion and string formation. *Thromb. Haemost.* **118**, 2074–2085 (2018).
  62. L. M. Silva, T. Kryza, T. Stoll, C. Hoogland, Y. Dong, C. R. Stephens, M. L. Hastie, V. Magdolen, O. Kleifeld, J. J. Gorman, J. A. Clements, Integration of two in-depth quantitative proteomics approaches determines the kallikrein-related peptidase 7 (KLK7) degradome in ovarian cancer cell secretome. *Mol. Cell. Proteomics* **18**, 818–836 (2019).
  63. P. R. Lawler, J. Lawler, Molecular basis for the regulation of angiogenesis by thrombospondin-1 and -2. *Cold Spring Harb. Perspect. Med.* **2**, a006627 (2012).
  64. B. S. Croix, C. Rago, V. Velculescu, G. Traverso, K. E. Romans, E. Montgomery, A. Lal, G. J. Riggins, C. Lengauer, B. Vogelstein, K. W. Kinzler, Genes expressed in human tumor endothelium. *Science* **289**, 1197–1202 (2000).
  65. Y. Tong, M. Wang, Y. Dai, D. Bao, J. Zhang, H. Pan, LncRNA HOXA-AS3 sponges miR-29c to facilitate cell proliferation, metastasis, and EMT process and activate the MEK/ERK signaling pathway in hepatocellular carcinoma. *Hum. Gene Ther. Clin. Dev.* **30**, 129–141 (2019).
  66. X. Wu, T. Liu, O. Fang, L. J. Leach, X. Hu, Z. Luo, miR-194 suppresses metastasis of non-small cell lung cancer through regulating expression of BMP1 and p27(kip1). *Oncogene* **33**, 1506–1514 (2014).
  67. N. V. Lee, M. Sato, D. S. Annis, J. A. Loo, L. Wu, D. F. Mosher, M. L. Iruela-Arispe, ADAMTS1 mediates the release of antiangiogenic polypeptides from TSP1 and 2. *EMBO J.* **25**, 5270–5283 (2006).
  68. S. Goicoechea, A. W. Orr, M. A. Pallero, P. Eggleton, J. E. Murphy-Ullrich, Thrombospondin mediates focal adhesion disassembly through interactions with cell surface calreticulin. *J. Biol. Chem.* **275**, 36358–36368 (2000).
  69. S. Rosini, N. Pugh, A. M. Bonna, D. J. S. Hulmes, R. W. Farndale, J. C. Adams, Thrombospondin-1 promotes matrix homeostasis by interacting with collagen and lysyl oxidase precursors and collagen cross-linking sites. *Sci. Signal.* **11**, eaar2566 (2018).
  70. M. Shi, J. Zhu, R. Wang, X. Chen, L. Mi, T. Walz, T. A. Springer, Latent TGF-β structure and activation. *Nature* **474**, 343–349 (2011).
  71. G. Blanc, B. Font, D. Eichenberger, C. Moreau, S. Ricard-Blum, D. J. Hulmes, C. Moali, Insights into how CUB domains can exert specific functions while sharing a common fold: Conserved and specific features of the CUB1 domain contribute to the molecular basis of procollagen C-proteinase enhancer-1 activity. *J. Biol. Chem.* **282**, 16924–16933 (2007).
  72. C. W. Hung, T. Koudelka, C. Anastasi, A. Becker, C. Moali, A. Tholey, Characterization of post-translational modifications in full-length human BMP-1 confirms the presence of a rare vicinal disulfide linkage in the catalytic domain and highlights novel features of the EGF domain. *J. Proteomics* **138**, 136–145 (2016).
  73. K. O. Yee, M. Duquette, A. Ludlow, J. Lawler, Purification and analysis of thrombospondin-1. *Curr. Protoc. Cell Biol.* **Chapter 10**, Unit 10.10 (2003).
  74. E. Turtle, N. Chow, C. Yang, S. Sosa, U. Bauer, M. Brenner, D. Solow-Cordero, W. B. Ho, Design and synthesis of procollagen C-proteinase inhibitors. *Bioorg. Med. Chem. Lett.* **22**, 7397–7401 (2012).
  75. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
  76. G. Klein, C. Mathe, M. Biola-Clier, S. Devineau, E. Drouineau, E. Hatem, L. Marichal, B. Alonso, J. C. Gaillard, G. Lagniel, J. Armengaud, M. Carriere, S. Chedin, Y. Boulard, S. Pin, J. P. Renault, J. C. Aude, J. Labarre, RNA-binding proteins are a major target of silica nanoparticles in cell extracts. *Nanotoxicology* **10**, 1555–1564 (2016).
  77. S. Bachy, F. Letourneur, P. Rouselle, Syndecan-1 interaction with the LG4/5 domain in laminin-332 is essential for keratinocyte migration. *J. Cell. Physiol.* **214**, 238–249 (2008).
  78. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
  79. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).

**Acknowledgments:** We are grateful to J. Murphy-Ullrich (UAB, Birmingham) for very useful tips on how to handle TSP-1, to C. Lethias (LBTI, Lyon) for providing most of the antibodies used in the preliminary screening of candidate substrates, to D. Sigaud-Roussel and her group (LBTI, Lyon) for helping us during cornea dissection, and to D. Hulmes (LBTI, Lyon) for critical reading of the manuscript. We also thank the staff of the LYMIC-PLATIM, ANIRA-CYTOOMETRY, and PSF facilities from the SFR Biosciences (UMS 3444, Lyon) for help in performing microscopy, flow cytometry, and automated Edman sequencing experiments, respectively. Last, we thank C. Bardel (UMR 5558 CNRS, University of Lyon) for guiding us on how to perform statistical analyses. **Funding:** This work was supported by the CNRS, the University of Lyon, the Région Auvergne Rhône-Alpes (ARC 2013 to M.T. and C.M.), the ANR-07-PHYSIO-022-01 and ANR-18-CE92-0035-01 (to C.M.), the ANR-13-RPIB-0003-01 (to P.R.), the Ligue Nationale contre le Cancer (to L.B.A. and P.R.), and the Ligue contre le Cancer-Comité du Rhône (to U.V.). **Author contributions:** C.M. and P.R. designed the project. C. Anastasi, P.R., M.T., A.T., C.C., A.B., N.M., M.D., L.B.A., L.F., A.A., F.D., N.E.K., J.A., C. Auxenfans, S.V.-L.G., and C.M. performed experiments. C. Auxenfans and P.F. collected and provided human primary cells and corneas. C.M., P.R., C. Anastasi, C.C., M.D., S.V.-L.G., F.D., and U.V. analyzed the data. C.M. wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

Submitted 6 December 2019

Accepted 5 June 2020

Published 7 July 2020

10.1126/scisignal.aba3880

**Citation:** C. Anastasi, P. Rousselle, M. Talantikite, A. Tessier, C. Cluzel, A. Bachmann, N. Mariano, M. Dussoyer, L. B. Alcaraz, L. Fortin, A. Aubert, F. Delolme, N. El Kholi, J. Armengaud, P. Fournié, C. Auxenfans, U. Valcourt, S. V.-L. Goff, C. Moali, BMP-1 disrupts cell adhesion and enhances TGF-β activation through cleavage of the matricellular protein thrombospondin-1. *Sci. Signal.* **13**, eaba3880 (2020).

Article 3: Vadon-Le Goff, S. *et al. Proc Natl Acad Sci USA* (Submitted)

## **The dark side of procollagen C-proteinase enhancer-2 (PCPE-2): inhibition of bone morphogenetic protein-1/tolloid-like proteinases**

Sandrine Vadon-Le Goff<sup>a,1</sup>, Agnès Tessier<sup>a,b,1</sup>, Mélissa Dussoyer<sup>a</sup>, Manon Napoli<sup>a</sup>, Priscillia Lagoutte<sup>a</sup>, Célian Peyronnel<sup>a</sup>, Lucie Essayan<sup>a</sup>, Svenja Kleiser<sup>b,c</sup>, Nicole Tueni<sup>d,e</sup>, Emmanuel Bettler<sup>a</sup>, Cindy Dieryckx<sup>a</sup>, Natacha Mariano<sup>a</sup>, Elisabeth Errazuriz-Cerda<sup>f</sup>, Carole Fruchart Gaillard<sup>g</sup>, Florence Ruggiero<sup>h</sup>, Christoph Becker-Pauly<sup>i</sup>, Jean-Marc Allain<sup>d,e</sup>, Leena Bruckner-Tuderman<sup>b</sup>, Alexander Nyström<sup>b,j</sup>, Catherine Moali<sup>a,\*</sup>

<sup>a</sup>University of Lyon, CNRS UMR5305, Tissue Biology and Therapeutic Engineering Laboratory (LBTI), 69367 Lyon, France.

<sup>b</sup>Department of Dermatology, Medical Faculty, Medical Center - University of Freiburg, 79104 Freiburg, Germany.

<sup>c</sup>University of Freiburg, Faculty of Biology, 79104 Freiburg, Germany.

<sup>d</sup>Ecole Polytechnique, CNRS UMR7649, Laboratoire de Mécanique des Solides (LMS), 91120 Palaiseau, France.

<sup>e</sup>INRIA, Inria Saclay-Ile de France, 91120 Palaiseau, France.

<sup>f</sup>University of Lyon, Centre d'Imagerie Quantitative Lyon-Est (CIQLE), SFR Santé-Lyon Est, 69373 Lyon, France.

<sup>g</sup>Université Paris-Saclay, CEA, INRAE, Médicaments et Technologies pour la Santé (MTS), SIMoS, 91191, Gif-sur-Yvette, France.

<sup>h</sup>ENS Lyon, CNRS UMR 5242, Institut de Génomique Fonctionnelle de Lyon (IGFL), 69007 Lyon, France.

<sup>i</sup>University of Kiel, Biochemical Institute, Unit for Degradomics of the Protease Web, Kiel, Germany.

<sup>j</sup>Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Freiburg, Germany.

<sup>1</sup>These authors contributed equally to the work

\* Corresponding author : Catherine Moali, Email: [c.moali@ibcp.fr](mailto:c.moali@ibcp.fr)

**Author contributions:** SVLG, AT, LBT, AN and CM designed research. SVLG, AT, MD, MN, PL, CP, LE, SK, NT, EB, CD, NM, EEC, JMA, AN and CM performed research. CF, FR and CBP provided reagents. SVLG, AT, MD, MN, PL, CP, LE, NT, EB, CBP, JMA, LBT, AN and CM analyzed the data. SVLG, AT, AN and CM wrote the paper.

**Classification:** Biological sciences/Biochemistry

**Keywords:** protease inhibition, collagen fibrillogenesis, angiogenesis, BMP-1/tolloid-like proteinases, PCPE, LDLR, TGF- $\beta$ , CUB domains

**Abstract:**

Procollagen C-proteinase enhancers (PCPE-1 and PCPE-2) are known for their ability to stimulate the C-terminal maturation of fibrillar collagens by BMP-1/tolloid-like proteinases (BTPs). This proteolytic event is the ultimate trigger of collagen fibrillogenesis, a process involved in all steps of life in mammals and in many life-threatening diseases. All the mechanistic information gained on this enhancing activity has been obtained with PCPE-1 and PCPE-2 is currently thought to work in a similar manner. Here, we show that deficiency of PCPE-2 has a rather mild effect on the organization and morphology of collagen fibrils *in vivo* and does not cause significant delay in procollagen maturation in fibroblast cultures. At the molecular level, we found that PCPE-2 plays a complex dual role in controlling BTP activity. At low concentrations, PCPE-2 is capable of activating procollagen processing, albeit to a substantially lesser extent than PCPE-1. At higher concentrations however, the stimulation of procollagen maturation decreases strongly due to the co-occurrence of a potent inhibitory mechanism targeting the protease and equally affecting fibrillar procollagens and non-collagenous BTP substrates. Furthermore, we establish that both the enhancing and inhibitory activities of PCPE-2 rely on its CUB domains and that its interactions with procollagens and BMP-1 are mutually exclusive. The non-catalytic domains of BMP-1 are required for inhibition, probably explaining why closely-related meprins are not similarly inhibited by PCPE-2. Collectively, our results identify PCPE-2 as the first endogenous and specific inhibitor of human BTPs, suggesting that its primary function may not be the regulation of collagen assembly.

**Significance statement:** It is a well-accepted paradigm in the protease field that every protease should work with a specific inhibitor to regulate its activity and avoid the deleterious effects of uncontrolled proteolytic activity. BMP-1/tolloid-like proteinases (BTPs) play important roles in extracellular matrix assembly, development, angiogenesis, muscle growth and lipid metabolism but, strikingly, no endogenous specific inhibitor is presently reported for this enzyme family in mammals. In this work, we identified PCPE-2 (procollagen C-proteinase enhancer 2) as a potent and specific inhibitor of BTPs. This result was totally unexpected because, as its name says, PCPE-2 is presently thought to enhance the procollagen processing activity of BTPs. Therefore, we also deciphered the molecular mechanisms explaining this apparent discrepancy.

**INTRODUCTION**

Collagen fibrillogenesis is an important biological process that shapes tissues and organs, maintains their homeostasis and allows their regeneration after injury. Besides the numerous enzymes required to convert individual procollagen chains into properly folded triple helices and then into collagen fibrils (1), two extracellular glycoproteins called procollagen C-proteinase enhancers (PCPE-1 and PCPE-2) are thought to play a major role.

PCPE-1 was first evidenced by Adar and colleagues and found to co-purify with so-called procollagen C-proteinases (PCPs) (2), now better recognized as BMP-1/tolloid-like proteinases (BTPs) which include bone morphogenetic protein-1 (BMP-1), its splicing variant mammalian tolloid (mTLD) and the homologous proteins mammalian tolloid-like 1 and 2 (mTLL-1 and mTLL-2) (3-5). BTPs can cleave the C-terminal propeptide of procollagens I-III, V and XI in the extracellular environment to decrease the solubility of collagen monomers and trigger fibrillogenesis (6). Previous biochemical and structural studies have revealed that PCPE-1 has no catalytic activity by itself but exerts its PCP enhancing activity by directly binding the procollagen substrate in the C-propeptide region (2, 7-9). The tight PCPE/procollagen interaction induces a local conformational change in the stalk region of the C-propeptide (10, 11), in close proximity to the cleavage site, that facilitates the cleavage of the compact procollagen trimer by BMP-1 and other BTPs (12).

Noteworthy, PCPE-1 also stimulates the N-terminal maturation of procollagen V by BMP-1 (13) but has no effect on all other activities of BTPs that have been described so far (14-16). These include the processing of other ECM proteins (e.g. laminin-332, collagens VI and VII, small leucine-rich proteoglycans, thrombospondin-1, lysyl oxidases) and of numerous proteins involved in the regulation of growth factor signaling (e.g. chordin, GDF8 and GDF11, betaglycan, latent-TGF- $\beta$  binding protein 1), angiogenesis (e.g. perlecan, prolactin), lipid metabolism (apolipoprotein A1, ApoA1 and low-density lipoprotein receptor, LDLR) and peripheral nervous system development (gliomedin) (6, 17-19).

In humans, PCPE-1 is encoded by the *PCOLCE* gene (20). This is generally well expressed in tissues rich in fibrillar collagens such as skin dermis, corneal stroma, tendons, bones, and cartilage but also in more cellularized organs where the interstitial extracellular matrix (ECM) plays an important role (e.g. heart, kidney, liver, lung, digestive tract) (21, 22). The increased abundance of PCPE-1 has been strongly correlated with liver fibrosis (23-25) and is also observed in cardiac and muscle fibrosis or corneal and cutaneous scars, as reviewed in (26). *Pcolce*-null mice do not show any severe abnormality but aberrantly-shaped collagen fibrils, affecting tissue mechanical properties, were observed in their bones and tendons (27, 28).

PCPE-2 was identified later than PCPE-1 (29) and has been much less studied. It shares 43 % sequence identity with PCPE-1 and has the same domain structure consisting of two CUB (Complement, Uegf, BMP-1) domains followed by a linker (28 amino-acids in length in mouse and human) and an NTR (netrin-like) domain. Its tissue distribution seems to be more restricted than that of PCPE-1 with major sites of expression including heart, digestive tract, lung, endocrine and adipose tissues and bladder (21, 22, 29, 30). Although Xu and colleagues made the early observation that PCPE-2 appears as a very faint band in immunoblots of cell extracts of human dermal fibroblasts and is not detected in the supernatant (29), one later report showed that exogenous PCPE-2 can stimulate the cleavage of procollagen I when

added to the supernatant of human foreskin fibroblasts (31). As PCPE-2 was also found to enhance the *in vitro* cleavage of procollagen II by BMP-1 and mTLL-1 to an extent roughly similar to PCPE-1 (22), the current view is that PCPE-2 behaves like PCPE-1 in terms of enhancing procollagen maturation.

*Pcolce2*-null mice are viable and fertile and do not show any obvious phenotype (32). Although collagen organization in these mice has not been characterized, *Pcolce2* deficiency seems to protect mice from aberrant collagen deposition and myocardial stiffening when challenged in a model of pressure overload induced by transverse aortic constriction (30). In addition, when these *Pcolce2*-null mice are crossed with mice deficient in LDLR and fed with a high-fat diet, they accumulate more lipids and show higher immune cell infiltration in atherosclerotic lesions (33). The mechanism by which PCPE-2 controls lipid metabolism is still unclear but could involve the regulation of ApoA1 maturation by BMP-1 (34, 35) or the control of high-density lipoprotein (HDL)-cholesterol ester uptake by scavenger receptor class BI (SR-BI) (33, 36).

Here, we have taken a closer look at the role of PCPE-2 in collagen maturation and assembly. We first observed that collagen fibrils in *Pcolce2*-null mice only show minor defects in homeostasis conditions. The effect of PCPE-2 on procollagen processing could be more easily evidenced *in vitro* but was unexpectedly found to be concentration-dependent, ranging from moderate enhancement to partial inhibition. This led us to analyze its mechanism of action in more detail and we found that PCPE-2, in contrast to PCPE-1, has the unique ability to form a tight complex with BMP-1. This specific feature endows PCPE-2 with new possibilities to fine-tune BTP activities and extends its potential functions far beyond procollagen maturation.

## RESULTS

### ***Pcolce2*-null mice show minor defects in collagen fibrils of the skin, heart and tail tendon**

To determine if PCPE-2 is important for the homeostasis of tissues rich in collagen I, we first analyzed the organization and morphology of collagen fibrils in the skin of wild-type (WT) and *Pcolce2*-null (KO) mice by transmission electron microscopy (TEM). In both genotypes, fibroblasts were surrounded by intertwined bundles of oriented collagen fibrils (Fig. 1A, Fig. S1), which are typical of skin, and there was no sign of fibril fusion or irregular shape. We also measured the diameter of more than 3,000 fibrils in TEM pictures of WT and KO skin and this analysis revealed that they were slightly increased in KO skin (Fig. 1B). The difference between the two genotypes was found to be significant when all the fibrils were analyzed together but not when median diameters were compared for individual mice (Fig. 1B), suggesting a rather mild effect. To determine if the observed trend of increased fibril diameters had any impact on skin mechanical properties, we performed uniaxial traction assays, as previously described for mouse skin (37). The tangent modulus, heel region length, failure stretch and ultimate tensile stress

were measured for WT and KO mice (Fig. S2) and none of these parameters differed significantly between the two types of mice, indicating that PCPE-2 does not play a major role in the control of skin mechanical properties.

As the heart is reported to be the major site of *Pcolce2* expression (21, 22, 29, 30), we also looked at collagen fibrils in this organ and did not observe any change in fibril shape and organization in low or high magnification TEM pictures (Fig. 1A, Fig. S1). Fibril diameter distribution showed a slight shift toward lower diameters but it was also rather moderate and not associated with a significant difference between the median diameters calculated for individual mice (Fig. 1B).

Finally, since it was previously described that the deletion of the *Pcolce* gene has major consequences on the morphology of collagen fibrils in tendons with 100 % of the fibrils presenting irregular (scalloped) shapes (27, 28), we analyzed transverse sections of the tail tendon in WT and KO mice. Again, apart from a 13 nm-decrease in the median fibril diameter of KO mice, no major irregularity in fibril shape could be evidenced (Fig. S1 and S3), suggesting that PCPE-2 is more dispensable than PCPE-1 for the proper assembly of collagen fibrils during tendon development and homeostasis.

To check if there was a possible compensation of the absence of PCPE-2 by PCPE-1, we used quantitative RT-PCR and immunoblots to compare *Pcolce* expression in the skin of WT and KO mice (Fig. S4A-B). No significant difference was observed and this was also the case at the mRNA level for the BTP enzymes (Fig. S4C: variants *Bmp1-1* and *Bmp1-3* encoding the BMP-1 and mTLD proteins, *Tll1*) which are expressed in mouse skin (38). The lack of major effect of PCPE-2 deficiency on collagen organization is therefore not explained by a compensatory mechanism affecting procollagen maturation.

***PCPE-2 is expressed by fibroblasts at a low level and exogenous PCPE-2 negatively regulates the procollagen I processing activity of BMP-1 in fibroblast medium***

To get further insights into the effect of PCPE-2 on procollagen maturation, we extracted fibroblasts from mouse and human skins and compared the relative levels of mRNA present in these cells for PCPE-1 and PCPE-2. Interestingly, the mean PCPE-1 gene expression was more than 40-fold higher than the mean PCPE-2 gene expression in fibroblasts from both species (Fig. 1C), confirming the previous result obtained at the protein level by Xu and colleagues who found that PCPE-1 was much more abundant than PCPE-2 in human fibroblast cell extracts (29).

We then analyzed procollagen I processing in the conditioned medium of WT and *Pcolce2* KO fibroblasts 8 h after serum starvation. In contrast to what has been previously described for *Pcolce*-null mouse embryonic fibroblasts (MEFs) and corneal keratocytes (28, 39) but in agreement with the low abundance of PCPE-2 in dermal fibroblasts, there was no delay in the C-terminal maturation of

procollagen I, as revealed by similar ratios of C-propeptide I to uncleaved precursors (procollagen I and pC-collagen I) in mouse fibroblast supernatants (Fig. 1D).

To get a better understanding of its activity, we also produced recombinant PCPE-2 fused to a 6 Histidine-tag at the N-terminus, in 293-T cells, and purified the protein by successive steps of heparin sepharose and Ni-NTA agarose chromatographies. The resulting protein was more than 95 % pure (Fig. S5A) and was not contaminated by endogenous PCPE-1, as confirmed by immunoblotting (Fig. S5B). Recombinant PCPE-2 was then incubated with the supernatant of human adult fibroblasts in the presence of BMP-1 and its procollagen C-proteinase (PCP) enhancing activity on procollagen I was compared to that of recombinant PCPE-1 in the same conditions. While PCPE-1 addition led to a significant increase in C-propeptide release when compared to the condition with BMP-1 alone, PCPE-2 was unable to enhance BMP-1 PCP activity and even displayed inhibitory activity (Fig. 1E-F). Even more surprising, PCPE-2 seemed to counteract the activity of PCPE-1 as the simultaneous addition of equimolar amounts of PCPE-1 and -2 failed to reproduce the enhancement obtained by PCPE-1 alone (Fig. 1E-F).

#### ***PCPE-2 does not behave like a canonical procollagen C-proteinase enhancer***

To get further insight into this unexpected behavior of PCPE-2, we ran a series of *in vitro* experiments with three different mini-procollagens corresponding to truncated versions of fibrillar procollagens I, II and III, lacking the N-propeptides and made of short triple helices and of full-length C-telopeptides and C-propeptides (40). These mini-procollagens were cleaved by BMP-1 and their cleavages were efficiently stimulated by PCPE-1 when the latter was present in equimolar amounts with the substrate (Fig. 2A). Interestingly, when tested in the same conditions, the enhancing effect of PCPE-2 was hardly detectable with mini-procollagens I and III but was more clearly evidenced with mini-procollagen II (Fig. 2A). The enhancement factor, calculated from band intensities as described in Materials and Methods, remained however substantially lower in all cases than with PCPE-1 which allowed 100 % of processing for all mini-procollagens in the conditions of the assay.

Since the enhancing activity of PCPE-1 is known to depend on its molar ratio to procollagen (7, 41), we next tried to vary the concentrations of PCPE-1 and PCPE-2 in the assay. To achieve this, we used another model substrate called CPIII-Long and previously found to be well-suited for the precise quantification of BMP-1 C-proteinase activity (7). This substrate is made of the three final repeats of the triple helix, the C-telopeptide and the C-propeptide of procollagen III. Whereas the enhancing activity of PCPE-1 reached a plateau between 0.5 and 1 equivalent (PCPE:CPIII-Long trimer molar ratio) and remained stable at the maximum cleavage rate with 4 equivalents (enhancement factor > 3), PCPE-2 activity also reached a plateau between 0.5 and 1 equivalent, at the relatively high enhancement factor of 1.7-1.8, but then decreased to the basal level defined by BMP-1 alone at 4 equivalents (Fig. 2B). In order to better compare the effects of PCPE-1 and PCPE-2, we also quantified C-propeptide

(CP) III release using an even larger range of PCPE:CPIII-Long ratios (ranging from 0:1 to 6:1) and the more linear fluorescent Sypro Ruby dye for detection. This experiment (Fig. 2C) clearly showed that the effect of PCPE-2 on BMP-1 procollagen-processing activity followed a bell-shaped curve with activation at lower concentrations (ratios from 0.1:1 to 1:1) and inhibition at higher concentrations (ratios  $\geq 3:1$ ). In the conditions of this experiment, the maximum enhancement factor with PCPE-2 was 1.3-fold and the maximum inhibition reached 50 % at the 6:1 ratio while PCPE-1 activity stayed at a plateau with an enhancement factor of 1.7 for all ratios superior to 0.3:1. This suggests that the enhancing activity of PCPE-2 is hindered by another process that is potent enough to abrogate BMP-1 stimulation and can even lead to the inhibition of CPIII-Long cleavage.

Together, our data reveal that PCPE-2 plays a complex dual role when regulating the PCP activity of BMP-1 and does not behave like the canonical enhancer PCPE-1. To explain these observations, the mechanism of action of PCPE-2 was investigated further and compared to that of PCPE-1, as described below.

#### ***PCPE-2 and PCPE-1 bind to the same site on procollagens***

As also observed in the experiment with fibroblast supernatants (Fig. 1E-F), when PCPE-1 and PCPE-2 were added simultaneously to the mini-procollagen cleavage assays, the resulting effect was below the activity of PCPE-1 alone and close to the activity of PCPE-2 alone or even slightly lower (Fig. 2A). This result suggested that in addition to being a poor PCP enhancer, PCPE-2 was able to antagonize the enhancing activity of PCPE-1 in a very efficient manner. Since we know from previous studies that the interaction of PCPE-1 with procollagens is the main driver of its activity (8, 42, 43), one possible mechanism to explain the antagonist effect of PCPE-2 could be binding the procollagens with a higher affinity than PCPE-1 but in a less productive manner. Interestingly, the PCPE-1 residues that were previously shown to be involved in the interaction with procollagen III were perfectly conserved in PCPE-2 (Fig. S6A). In addition, when a model of the CUB domains of PCPE-2 was superimposed on the structure of the complex between the C-propeptide of collagen III and the CUB domains of PCPE-1 (12) (Fig. S6B), the two PCPEs seemed to adopt a similar conformation with a rmsd (root-mean-square deviation) between the atomic positions of alpha carbons in CUB domains of 3.25 Å.

To confirm the binding of PCPE-2 to procollagens, we analyzed the interaction of recombinant PCPE-2 with immobilized mini-procollagen III by surface plasmon resonance (SPR). We observed that PCPE-2 also interacted efficiently with mini-procollagen III (Fig. 2D). However, the association and dissociation rates were faster than for PCPE-1 (Fig. S7A), leading to kinetic constants which were outside the instrument specifications. To circumvent this problem, we fitted the curves in the steady-state mode and obtained a dissociation constant of 11 nM (inset in Fig. 2D). For comparison, the  $K_D$  obtained for PCPE-1 in the same conditions was 2.2 nM (Fig. S7A), in good agreement with previous

measurements (0.3-7 nM; (8, 43)), indicating that PCPE-1 leads to a slightly more stable complex with mini-procollagen III than PCPE-2. As for PCPE-1, the main site of interaction of PCPE-2 on mini-procollagen III was found to be the C-propeptide and interaction with this region alone led to a dissociation constant of 45 nM (Fig. S7B), also slightly higher than for PCPE-1 (8).

Furthermore, we could demonstrate that both proteins bound to overlapping sites on mini-procollagen III as there was no substantial increase in the recorded SPR signal when PCPE-1 was injected on a mini-procollagen III surface which was already saturated with PCPE-2 (Fig. 2E). This was also true when PCPE-2 was injected on a mini-procollagen III surface saturated with PCPE-1 (Fig. S7C). However, the injection of PCPE-1 after PCPE-2 led to a slower dissociation (Fig. 2E) while no change in dissociation rate was observed when PCPE-2 was injected after PCPE-1 (Fig. S7C). This demonstrated that PCPE-1 could efficiently displace PCPE-2 from mini-procollagen III whereas the opposite was not true, ruling out the hypothesis that the antagonist effect of PCPE-2 was due to its capacity to form more stable complexes with procollagens.

#### ***The CUB domains are responsible for PCPE-2 enhancing activity***

To better understand the differences between PCPE-1 and PCPE-2, we produced the CUB and NTR domains of PCPE-2 separately (Fig. S5C). Since direct overexpression in HEK 293 cell lines gave low yields of unstable protein, we designed a PCPE-2 construct containing an internal HRV 3C-protease cleavage site (named PCPE-2 3C hereafter) which allowed us to separate the CUB1CUB2 region from the NTR domain (Fig. S5C-D).

After purification, these domains were compared to full-length PCPE-2 and PCPE-2 3C in a CPIII-Long cleavage assay (Fig. 2B and S8). Interestingly, while the NTR domain had no impact on substrate processing, PCPE-2 3C and the derived CUB1CUB2 domains could recapitulate the activity of full-length PCPE-2 with a weak enhancement at 0.5 equivalent which tended to decrease when molar ratios to CPIII-Long were increased. No further change in CUB1CUB2 activity was observed when the NTR domain was also present (Fig. S8). In agreement with the observation that the CUB domains of PCPE-2 play the most important role, these bound to mini-procollagen III with a  $K_D$  around 30 nM (Fig. S9A) while no interaction between the NTR domain and mini-procollagen III was found (Fig. S9B). This suggests that the CUB domains are responsible for PCPE-2 activity, similarly to what was previously described for PCPE-1 (43, 44). Also, we can conclude that the interaction of PCPE-2 with procollagens alone is unlikely to explain the complex activity profile described in Fig. 2C as the main function-related features of PCPE-1 are conserved in PCPE-2.

***PCPE-2 is a potent inhibitor of BMP-1***

In addition to its PCP enhancing activity, PCPE-1 is also known to stimulate the cleavage of the N-terminal (TSPN) domain of the  $\alpha 1$  chain of procollagen V by BMP-1 (13), another activity related to collagen fibril assembly (45). To determine if PCPE-2 could play a similar role, we monitored the release of the TSPN domain from a truncated form of the procollagen V N-propeptide, called pN $\alpha 1$ (V) (13), by SDS-PAGE. As expected, more TSPN was detected when PCPE-1 was present (at molar ratios to pN $\alpha 1$ (V) between 0.1:1 and 2:1) than when BMP-1 alone was incubated with the substrate (Fig. 3A). In contrast, there was no visible TSPN product when PCPE-2 was added at all the tested molar ratios, suggesting that PCPE-2 strongly inhibited BMP-1 activity on pN $\alpha 1$ (V). Even more surprisingly, we observed that PCPE-2 also reduced the processing of several other known BMP-1 substrates on which PCPE-1 had no effect. This was the case for thrombospondin-1 (TSP-1), a recently described substrate involved in the control of cell adhesion and TGF- $\beta$  activation (46). Its cleavage products became hardly visible when PCPE-2 was added to the assay at a 1:1 molar ratio (Fig. 3B). Similarly, the BMP-1-dependent processing of betaglycan, another BMP-1 substrate acting both as a TGF- $\beta$  receptor (TGF- $\beta$  RIII) at the cell surface and as a TGF- $\beta$  antagonist when it is released in the extracellular environment (47), was reduced in the presence of PCPE-2 (Fig. 3C). PCPE-2 also inhibited the cleavage of chordin, a major antagonist of the BMP-2 and BMP-4 growth factors (5), and of endorepellin, a stable angiostatic fragment of perlecan (48) (Fig. 3D-E). Finally and of great interest for the described role of PCPE-2 in lipid metabolism, the BMP-1-dependent cleavage (17, 49) of LDLR that mediates the endocytosis of LDL particles was completely blocked by PCPE-2 (Fig. 3F). In summary, PCPE-2 inhibited the cleavages of all the BMP-1 non-collagenous substrates that we were able to test, suggesting that it behaves as a broad-range inhibitor of BMP-1 activities.

In order to check that this newly observed inhibitory activity of PCPE-2 was not an artifact of the protocol used to purify the recombinant protein, we looked at the processing of an endogenous substrate naturally expressed in 293-EBNA cells. Among the various possible substrates that we tested, the C-terminal fragment of LDLR generated by BMP-1 (17, 49) could be readily detected in the cell lysate of 293-EBNA cells. The dependence on endogenous BMP-1 was confirmed by the comparison of three 293-cell lines stably expressing either PCPE-2, the corresponding empty vector or the *Xenopus* protein sizzled (not present in mammals) which is known to be a potent and specific exogenous inhibitor of BTPs (50). We found that LDLR C-terminal proteolytic product was clearly diminished in PCPE-2- and sizzled-transfected cells compared to cells transfected with the corresponding empty vector (Fig. 3G).

Finally, the regulation of BMP-1 activity by PCPE-2 was assessed using the fluorogenic peptide Mca-YVADAPK(Dnp)-OH that can be cleaved by BMP-1 between the first alanine and the aspartate. Whereas the addition of PCPE-1 did not modify the cleavage rate of the peptide up to a 30:1 PCPE-1:BMP-1 molar ratio, PCPE-2 was also found to inhibit this cleavage at all tested ratios (from 1:1 to

30:1; Fig. 4A). The apparent inhibition constant ( $K_i^{\text{app}}$ ) value for PCPE-2 was calculated to be  $2.8 \pm 0.6$  nM (Fig. 4B) which indicated that PCPE-2 was a tight-binding inhibitor of BMP-1. However, it failed to decrease BMP-1 activity down to zero with around 25 % of residual activity at the highest PCPE-2 concentration (690 nM). This result suggested that PCPE-2 was not a competitive inhibitor directly binding to BMP-1 active site.

***The inhibitory activity of PCPE-2 is borne by its CUB domains and requires BMP-1 non-catalytic domains***

Our next goal was to decipher the mechanism of BMP-1 inhibition by PCPE-2. Using the separate CUB and NTR domains generated above, we first showed that the CUB domains were necessary and sufficient for the inhibition of the cleavage of the fluorogenic peptide and that the NTR domain had no effect (Fig. 4A). On the BMP-1 side, the catalytic domain of the protease could also cleave the peptide but was not affected by the presence of PCPE-2 (Fig. 4C), also excluding the possibility that PCPE-2 acts as a competitive inhibitor binding into the protease active site. In line with these results, PCPE-2 and CUB1CUB2 but not NTR bound to immobilized BMP-1 by SPR (Fig. 4D and S9C). Furthermore, full-length BMP-1 but not its catalytic domain bound to immobilized PCPE-2 (Fig. 4E). The dissociation constant for the interaction of PCPE-2 with BMP-1 was found to be 8.4 nM with the steady-state fit (Fig. 4F), in the same range as the  $K_i^{\text{app}}$  measured above and as the dissociation constant for the interaction of PCPE-2 with procollagens (Fig. 2D). In the same conditions, PCPE-1 only gave a very weak signal (Fig. 4D), as previously described (42). Therefore, we could hypothesize that the difference in activity between the two proteins comes from the increased ability of PCPE-2 to form a stable complex with BMP-1, an interaction which further leads to BMP-1 inhibition.

To better understand the consequences of this dual binding capacity, we ran a competition experiment which consisted of injecting increasing concentrations of BMP-1 on a CPIII surface to which PCPE-2 was already bound (Fig. 4G). Interestingly, BMP-1 was found to be able to disrupt the complex between PCPE-2 and CPIII in a concentration-dependent manner, leading to 77 % inhibition when both BMP-1 and PCPE-2 were injected at the same concentration (200 nM). Here again, PCPE-2 behaved differently from PCPE-1 as the simultaneous injection of PCPE-1 and BMP-1 over immobilized CPIII did not lead to a decreased signal but rather to an increased signal (Fig. S9D), revealing that BMP-1 and PCPE-1 could bind together on CPIII. These important results demonstrate that the interactions of PCPE-2 with the procollagen substrate and the BMP-1 protease are mutually exclusive and that a stable ternary complex involving BMP-1, PCPE-2 and a procollagen substrate is unlikely to occur.

***PCPE-2 specifically inhibits BTPs but also interferes with the PCP activity of meprins through its interaction with procollagens***

As PCPE-2 is only the second endogenous inhibitor of BMP-1 to be described in mammals after  $\alpha$ 2-macroglobulin (51), a plasma protein known to inhibit a wide range of proteolytic enzymes, we wanted to know if it was specific to BMP-1. We first tested its effect on mTLL-1 in a CPIII-Long cleavage assay (Fig. 5A). Interestingly, the enhancement of the PCP activity of mTLL-1 by PCPE-2 followed the same trends as for BMP-1 with a maximum enhancement factor of 1.6 at 0.5:1 and 1:1 molar ratios which then dropped off at higher PCPE-2 concentrations. In addition, PCPE-2 could efficiently inhibit mTLL-1 in the fluorogenic peptide assay (Fig. 5B), albeit to a slightly lower extent than with BMP-1, indicating that PCPE-2 can affect both the collagenous and non-collagenous substrates of the enzyme, in agreement with what was observed above for BMP-1.

Meprin  $\alpha$  and meprin  $\beta$  are two other members of the astacin-like subgroup of metalloproteinases that were tested next. They were differentially impacted in the fluorogenic peptide and CPIII-Long assays with no effect of PCPE-2 on the peptide cleavage but a slight inhibition of the maturation of the procollagen-derived protein by meprins (Fig. 5C-D). The latter result was in agreement with our previous finding that PCPE-1 could inhibit the PCP activity of meprins by blocking their access to the cleavage site through its tight interaction with procollagens (52). A similar mechanism is probably also at play with PCPE-2 but to further confirm that PCPE-2 does not inhibit other meprin activities, another physiological substrate of these proteases (CD99, (53)) was tested. As expected, CD99 cleavage was not affected by PCPE-2 (Fig. S10) and this strongly suggested that PCPE-2 inhibition specifically targeted BTPs. Consequently, PCPE-2 appears as the first specific endogenous inhibitor of this enzyme family in mammals.

## DISCUSSION

In this study, we provide evidence that collagen fibrils of three different tissues (skin, heart and tail tendon) do not show any obvious defects of their overall organization and morphology when PCPE-2 is missing. Only the distributions of fibril diameters display some alterations but this has no significant impact on tissue mechanical properties, at least in skin. The absence of pronounced collagen phenotype also applies to the heart, a site of high PCPE-2 expression (22, 54), and appears to be in marked contrast to the results obtained with *Pcolce*-null mice which display severe alterations of fibril morphology in bones and tendons (27, 28). This led us to analyze the PCP enhancing activity of PCPE-2 in more detail and we found that, even if enhancement of procollagen maturation can be observed *in vitro* in a specific range of PCPE-2 concentrations, the level of stimulation achieved by this protein is generally much lower than by PCPE-1. This is especially true for procollagens I and III but less for procollagen II, for

which the enhancement is more easily observed. Therefore, the use of the procollagen II substrate for the early characterization of PCPE-2 (22) may have led to the partially flawed conclusion that, at least *in vitro*, PCPE-2 works exactly like PCPE-1. When we extended the PCPE-2 concentration range around the classical 1:1 ratio to procollagen, we actually revealed that the regulation of substrate processing follows a bell-shaped curve with an increase in enhancing activity at low concentrations which then rapidly drops off when the PCPE-2 concentration becomes higher, to finally fall below the basal level observed with BMP-1 alone. This indicates that PCPE-2 can also lead to the inhibition of procollagen processing in specific conditions, a result which was totally unexpected.

How results obtained *in vitro* translate *in vivo* in terms of collagen fibrillogenesis regulation is difficult to anticipate in the absence of precise data regarding the relative tissue concentrations of procollagens, PCPEs and BTPs. In line with the relatively low expression of *PCOLCE2* in dermal fibroblasts (this study and (29)) and with the absence of major alterations in collagen fibrils, it seems that PCPE-2 does not play a crucial role in the organization of the collagen network and in the maintenance of its homeostasis. The absence of effect of PCPE-2 on collagen fibrillogenesis *in vivo* might thus be due to the combination of a low expression at sites of active collagen deposition and to a limited capacity to enhance or inhibit procollagen maturation at this low abundance. This preliminary conclusion applies to homeostatic conditions but it cannot be excluded that *PCOLCE2* expression might be induced upon challenge. However, even in conditions known to trigger high collagen synthesis such as tissue injury, *PCOLCE2* does not seem to follow the up-regulation trend often displayed by *BMP1* and *PCOLCE* (6, 26). Its protein or RNA level was actually found to be down-regulated in murine models of corneal scarring (55) and lung injury (56, 57) and unchanged in a model of cardiac fibrosis (58). Interestingly also, in skin and kidney, *PCOLCE2* expression seems restricted to specific fibroblast subpopulations with progenitor potential (59, 60) rather than with myofibroblast/matrix-producing potential (59, 61). In contrast, *PCOLCE2* is more highly expressed than *PCOLCE* in human macrophages (62, 63) and could regulate neutrophil functions (64). Even though the regulation of *PCOLCE2* expression during injury-driven inflammation remains to be established, these data suggest that PCPE-1 and PCPE-2 might play distinct and non-overlapping roles in tissue repair. In support of this hypothesis, we have observed here that PCPE-2 can efficiently counteract PCPE-1 activity, potentially leading to conflicting outcomes when both are present simultaneously in the same locations. More intriguingly, our present results do not fit with the findings, described in an earlier report (30), suggesting that *Pcolce2* deficiency protects mice against collagen accumulation and myocardial stiffening in a model of chronic pressure overload and that PCPE-2, like PCPE-1 (26), could represent a relevant target in fibrosis. These findings will have to be investigated again to determine if they really depend on the direct regulation of BTP-dependent collagen maturation by PCPE-2 or on unrelated mechanisms. More generally, the specific functions of the two PCPEs in fibrosis will be important to clarify if novel anti-fibrotic strategies targeting one or the other are developed.

In this context, another remarkable result of our study is that PCPE-2 completely inhibits the N-terminal maturation of procollagen V while, as expected in the same conditions (13), PCPE-1 behaves as a weak enhancer of this cleavage. Collagen V N-propeptide retention is known to limit fibril growth (45) but the absence of substantial effect of PCPE-2 deficiency on fibril diameters also argues against both proteins being synthesized simultaneously in the analyzed tissues. While the exact mechanism of stimulation of N-propeptide processing by PCPE-1 is unknown, this is also thought to be mediated by substrate binding since PCPE-1 can efficiently interact with the TSPN domain of collagen V (13). In the case of PCPE-2, the observation that procollagen V was not the only BMP-1 substrate to be affected suggested that its inhibitory activity was most probably associated to a direct effect on the protease rather than on the substrate. Actually, all the other BMP-1 substrates that we were able to test, including five physiological substrates (TSP-1, betaglycan, chordin, endorepellin and LDLR) and a short model peptide, were all strongly inhibited by PCPE-2 despite the fact that they were all unrelated in terms of structure and function. This unexpected finding also points at novel potential functions for PCPE-2. For example, the proteolytic products resulting from some of the characterized substrates are known to play a role in the control of cell adhesion (TSP-1), TGF- $\beta$  activation and bioavailability (TSP-1, betaglycan) and/or angiogenesis (TSP-1, endorepellin). Their regulation by PCPE-2 could thus give new mechanistic hints to explain that the latter recently emerged as a signature gene predicting survival in several types of cancers (65-68). Similarly, the cleavage of LDLR is strongly inhibited by PCPE-2 suggesting a connection with PCPE-2 relatively well-established roles in the regulation of diet-induced atherosclerosis (33) and adipose tissue expansion (36). However, these roles were mainly described in mouse models where the LDLR cleavage site that is targeted by BMP-1 is not conserved (17, 49). Other mechanisms are therefore also certainly involved such as a direct interaction of PCPE-2 with scavenger receptors (e.g. SR-BI, as previously proposed (33, 36)) or the cleavage of other BMP-1 substrates linked to lipid metabolism. In this context, a BMP-1 substrate that was unfortunately not available to us but will require further investigation is pro- ApoA1 (35). Its proteolytic maturation by BMP-1 was previously reported to be enhanced in the presence of PCPE-2 (31) while our present results would rather predict that it should be inhibited.

The broad inhibition of BMP-1 activities by PCPE-2 suggested the possibility of a direct and strong interaction between the regulatory protein and the protease. This hypothesis was confirmed by SPR experiments while, in the same conditions, PCPE-1 appeared as a weak ligand of the protease. Interestingly also, PCPE-2 binds to BMP-1 through the same domains as those involved in the interaction with collagen C-propeptides, through overlapping binding sites, as demonstrated by competition experiments. This results in mutually exclusive interactions of PCPE-2 with the procollagen or the protease and leads to the unusual bell-shaped curve described above, with successive phases of stimulation by PCPE-2 (through procollagen binding) and of inhibition by PCPE-2 (through protease

binding). It seems that PCPE-2 literally hijacks BMP-1 and prevents its interaction with substrates, independently of the presence of PCPE-1 which then becomes totally powerless. It even appears that PCPE-1 potentiates the inhibition of BMP-1 PCP activity by PCPE-2, probably through the saturation of binding sites on procollagens which makes more PCPE-2 molecules available to bind BMP-1. This is reflected by the higher inhibition observed when both PCPEs are present than when PCPE-2 alone is present. Since the catalytic domain of BMP-1 alone cannot interact with PCPE-2, we propose that the latter regulates protease activity through an interaction with its non-catalytic domains. These have been suggested to adopt a closed horseshoe conformation in solution, bringing CUB3 in close proximity to the catalytic domain (69). However, the precise molecular mechanism explaining why PCPE-2, in contrast to PCPE-1, can form tight complexes with BTPs remains unclear and will require further investigation.

Together, the data presented here establish that PCPE-2 is a potent endogenous inhibitor of the BMP-1/tolloid-like proteinases. It is probably also very specific as suggested by its mechanism of action and its lack of direct effect on meprins. As such, at least in mammals, PCPE-2 seems to be the first endogenous and specific inhibitor of BTPs capable of regulating all their activities. Other natural inhibitors have been described in the past but these were either not specific to BTPs ( $\alpha$ 2-macroglobulin (51)), not conserved in humans (sizzled (50, 70)) or highly controversial (sFRP-2 (50, 71, 72) and BMP-4 (73, 74)). In that sense, it is intriguing to see that PCPE-2 has a netrin-like domain in common with the natural inhibitors of the matrix metalloproteinases (TIMPs 1-4) but that its inhibitory activity is not borne by this domain which is fully dispensable for inhibition.

A summary of the finely-tuned regulation of BMP-1 activity permitted by the two PCPEs is provided in Fig. 6. The entire range of possible outcomes, from no effect to enhancement and inhibition, is potentially available but will depend on the relative concentrations and availability of the various partners (substrate, protease, PCPE) in specific situations. Future work should therefore focus on the identification of the cell types that are the most relevant in terms of PCPE-2 production and on the characterization of its possible co-expression with BTPs in time and space in order to better define the scope of the new regulatory mechanisms unveiled by the present study.

## MATERIALS AND METHODS

### *Pcolce2*-null mice

*Pcolce2*-null embryos were obtained from the group of Dr. Thomas Boehm at the Max-Planck-Institute of Immunobiology and Epigenetics in Freiburg, Germany. They were generated by replacing a large part of the exon 3 of the *Pcolce2* gene by the IRES-lacZ/neomycin resistance cassette, as described in

(32). Embryos were injected into C57BL/6 females and the offspring were mated with C57BL/6 mice to generate both wild-type (WT) and *Pcolce2*-null (KO) mice. These were housed in the pathogen-free facility of the University of Freiburg according to ethical regulations.

### ***Protein production***

The human PCPE-2 sequence was obtained from the pOTB7-PCPE-2 plasmid purchased from Source BioScience (clone ID 3951739). This sequence contained two mutations (829G>T and 874C>A) which were corrected. The PCPE-2 sequence was then cloned into a pJET1.2/blunt vector using the CloneJET PCR Cloning strategy (ThermoFisher). PCPE-2 cDNA was further amplified by PCR with an N-terminal 6His-tag and inserted into the mammalian expression vector pHLsec (75), through the AgeI/XhoI sites, using the In-Fusion cloning strategy (Ozyme). The PCPE-2 3C construct was designed to insert one HRV 3C-protease cleavage site between the 6His-tag and the beginning of PCPE-2 (giving the following N-terminal sequence after signal peptide removal: ETGHHHHHHLEVLFGQPAS) and one HRV 3C-protease cleavage site between amino-acids 270 and 271 of full-length human PCPE-2 (as described in Fig. S5C). The corresponding synthetic gene was ordered from Twist Bioscience after codon optimization. The cDNA was digested with AgeI/XhoI and subcloned into linearized pHLsec. All products were checked by Sanger sequencing (Eurofins).

For the production of PCPE-2 and PCPE-2 3C, (HEK) 293-T cells were seeded in HYPERflask or CellSTACK cell culture vessels (Corning) in DMEM AQ medium (Merck) containing 10 % of fetal bovine serum (FBS, Gibco or Eurobio) and 1 % antibiotic-antimycotic solution (AAS, Merck). At 80-90 % of confluency, cells were transfected with pHLSec-6His-PCPE-2 or pHLSec-6His-PCPE-2 3C plasmid, using polyethylenimine (PEI, Merck) as transfection reagent, in DMEM medium with only 2 % of FBS and 1 % of Non-Essential Amino Acids (Merck). Culture medium was collected 3 and 5 days after transfection, centrifuged at 1000 g and mixed with protease inhibitors (0.25 mM Pefabloc (Roth) and 2 mM N-ethylmaleimide (Merck)). The first purification step was on Heparin Sepharose 6 Fast Flow resin (Cytiva) which was prepacked into a column and equilibrated in 20 mM HEPES pH 7.4, 0.15 M NaCl. A linear gradient of NaCl was used to elute the protein (from 0.15 to 2 M NaCl) and the PCPE-2 protein was obtained between 0.5 to 0.8 M of NaCl. Imidazole was then added to the pool of PCPE-2-containing fractions to a final concentration of 5 mM before loading the protein solution on Ni-NTA resin (Qiagen). The resin was washed successively with 20 mM HEPES pH 7.4, 0.6 M NaCl containing 20 mM then 50 mM imidazole and the protein was eluted with a gradient of imidazole from 50 to 500 mM. Fractions containing PCPE-2 were pooled, diluted with 20 mM HEPES pH 7.4 to reduce salt concentration down to 0.25 M and loaded on a HiTRAP Heparin-HP column (Cytiva). The protein was eluted with 0.6 M NaCl in 20 mM HEPES pH 7.4, flash-frozen in the same buffer supplemented with 5 mM CaCl<sub>2</sub> and 0.1 % of n-octyl-β-D-glucopyranoside (βOG, Roth) and stored at -80 °C until use. The average yield for PCPE-2 production was 0.5 mg per liter of culture.

The CUB1CUB2 and NTR domains of PCPE-2 were generated by cleavage of the PCPE-2 3C protein purified as above until elution from the Ni-NTA column. Then, HRV 3C-protease was added to PCPE-2 3C in a 1/30 w/w ratio and the mixture was incubated for 1 h 30 at room temperature, followed by dialysis overnight at 4 °C in a 7 kDa Slide-a-Lyzer (ThermoFisher) to eliminate imidazole. Removal of uncleaved protein and HRV 3C-protease was achieved by purification on Ni-NTA agarose (0.5 ml) equilibrated in 20 mM HEPES pH 7.4, 0.3 M NaCl, 0.1 %  $\beta$ OG. PCPE-2 fragments were collected in the flow-through and loaded on a HiTRAP Heparin-HP column equilibrated in the same buffer. The CUB1CUB2 domains were recovered in the flow-through while the NTR domain was eluted with 0.6 M NaCl. Before concentration on Amicon 3 kDa devices (Merck), the NaCl concentration of the CUB1CUB2 solution was also adjusted to 0.6 M and calcium (2.5 mM) was added. Proteins were then flash-frozen and stored at -80 °C until use.

### ***Statistical analysis***

Statistical analyses were performed with GraphPad Prism 8. Using a 5 % threshold, we first verified the normality of the distribution using the Shapiro-Wilk test as well as the equality of variance with the Fisher test. If these tests were passed successfully, the significance of the mean difference was assessed using an unpaired t-test. In other situations, the test used was as indicated in figure legends. For multiple comparisons, one-way ANOVA with paired data and Dunnett post-test was used.

***Further information is provided in SI Materials and Methods.***

**Acknowledgements:** We are grateful to Jacques Brocard and Christophe Chamot (PLATIM facility, Lyon Multiscale Imaging Center and SFR Biosciences) and Denis Ressnikoff and Bruno Chapuis (CIQLE facility, Lyon Multiscale Imaging Center and SFR Santé-Lyon Est) for their help with the quantitative analysis of collagen fibrils. We also thank Fernando Lopez-Casillas (Universidad Nacional Autonoma de Mexico, Instituto de Fisiologia Celular, Mexico) for providing recombinant betaglycan and David Hulmes (LBTL, CNRS-University of Lyon) for critical reading of the manuscript. Finally, we acknowledge the contribution of Alexis Cellier and Paulin Arès to the optimization of PCPE-2-derived protein production.

**Funding:** This work was supported by the CNRS, the University of Lyon, ANR grants 18-CE92-0035-01 and 21-CE11-0020-01 to CM, ANR grant 17-CE14-0033 to SVLG, DFG grants SFB850, project B11 and NY90/5-1 to AN and DFG grant SFB877, project A9 to CBP.

## REFERENCES

1. J. Myllyharju, K. I. Kivirikko, Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* **20**, 33-43 (2004).
2. R. Adar, E. Kessler, B. Goldberg, Evidence for a protein that enhances the activity of type I procollagen C-proteinase. *Coll Relat Res* **6**, 267-277 (1986).
3. E. Kessler *et al.*, Bone morphogenetic protein-1: the type I procollagen C-proteinase. *Science* **271**, 360-362 (1996).
4. S. W. Li *et al.*, The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenetic protein-1. *Proc Natl Acad Sci USA* **93**, 5127-5130 (1996).
5. I. C. Scott *et al.*, Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. *Dev Biol* **213**, 283-300 (1999).
6. S. Vadon-Le Goff, D. J. Hulmes, C. Moali, BMP-1/tolloid-like proteinases synchronize matrix assembly with growth factor activation to promote morphogenesis and tissue remodeling. *Matrix Biol* **44-46C**, 14-23 (2015).
7. J. M. Bourhis *et al.*, Procollagen C-proteinase enhancer grasps the stalk of the C-propeptide trimer to boost collagen precursor maturation. *Proc Natl Acad Sci USA* **110**, 6394-6399 (2013).
8. S. Vadon-Le Goff *et al.*, Procollagen C-proteinase enhancer stimulates procollagen processing by binding to the C-propeptide region only. *J Biol Chem* **286**, 38932-38938 (2011).
9. E. Kessler, R. Adar, Type I procollagen C-proteinase from mouse fibroblasts. Purification and demonstration of a 55-kDa enhancer glycoprotein. *Eur J Biochem* **186**, 115-121 (1989).
10. J. M. Bourhis *et al.*, Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat Struct Mol Biol* **19**, 1031-1036 (2012).
11. U. Sharma *et al.*, Structural basis of homo- and heterotrimerization of collagen I. *Nat Commun* **8**, 14671 (2017).
12. D. Pulido *et al.*, Structural Basis for the Acceleration of Procollagen Processing by Procollagen C-Proteinase Enhancer-1. *Structure* **26**, 1384-1392 e1383 (2018).
13. S. Symoens *et al.*, Identification of binding partners interacting with the alpha1-N-propeptide of type V collagen. *Biochem J* **433**, 371-381 (2010).
14. C. Moali *et al.*, Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1. *J Biol Chem* **280**, 24188-24194 (2005).
15. V. Petropoulou, L. Garrigue-Antar, K. E. Kadler, Identification of the minimal domain structure of bone morphogenetic protein-1 (BMP-1) for chordinase activity: chordinase activity is not enhanced by procollagen C-proteinase enhancer-1 (PCPE-1). *J Biol Chem* **280**, 22616-22623 (2005).
16. Z. von Marschall, L. W. Fisher, Dentin sialophosphoprotein (DSPP) is cleaved into its two natural dentin matrix products by three isoforms of bone morphogenetic protein-1 (BMP1). *Matrix Biol* **29**, 295-303 (2010).
17. S. Banerjee *et al.*, Proteolysis of the low density lipoprotein receptor by bone morphogenetic protein-1 regulates cellular cholesterol uptake. *Sci Rep* **9**, 11416 (2019).
18. S. E. Heumuller *et al.*, C-terminal proteolysis of the collagen VI alpha3 chain by BMP-1 and proprotein convertase(s) releases endotrophin in fragments of different sizes. *J Biol Chem* **294**, 13769-13780 (2019).
19. Y. Eshed-Eisenbach *et al.*, Precise Spatiotemporal Control of Nodal Na(+) Channel Clustering by Bone Morphogenetic Protein-1/Tolloid-like Proteinases. *Neuron* **106**, 806-815 e806 (2020).
20. K. Takahara *et al.*, Type I procollagen COOH-terminal proteinase enhancer protein: identification, primary structure, and chromosomal localization of the cognate human gene (PCOLCE). *J Biol Chem* **269**, 26280-26285 (1994).
21. HPA (2020) The Human Protein Atlas. (Sweden).
22. B. M. Steiglitz, D. R. Keene, D. S. Greenspan, PCOLCE2 Encodes a Functional Procollagen C-Proteinase Enhancer (PCPE2) That Is a Collagen-binding Protein Differing in Distribution of Expression and Post-translational Modification from the Previously Described PCPE1. *J Biol Chem* **277**, 49820-49830 (2002).
23. E. Hassoun *et al.*, Procollagen C-Proteinase Enhancer 1 (PCPE-1) as a Plasma Marker of Muscle and Liver Fibrosis in Mice. *PLoS One* **11**, e0159606 (2016).
24. I. Ogata *et al.*, Up-regulation of type I procollagen C-proteinase enhancer protein messenger RNA in rats with CCl4-induced liver fibrosis. *Hepatology* **26**, 611-617 (1997).
25. P. Sansilvestri Morel *et al.*, Procollagen C-Proteinase Enhancer-1 (PCPE-1) deficiency in mice reduces liver fibrosis but not NASH progression. *PLoS One* **17**, e0263828 (2022).
26. P. Lagoutte, E. Bettler, S. Vadon-Le Goff, C. Moali, Procollagen C-proteinase enhancer-1 (PCPE-1), a potential biomarker and therapeutic target for fibrosis. *Matrix Biol Plus* **11**, 100062 (2021).
27. P. Sansilvestri-Morel *et al.*, Deficiency of Procollagen C-Proteinase Enhancer 1 in Mice has No Major Impact on Cardiac Collagen and Function Under Basal Conditions. *J Cardiovasc Pharmacol* **78**, e703-e713 (2021).

28. B. M. Steiglitz *et al.*, Procollagen C proteinase enhancer 1 genes are important determinants of the mechanical properties and geometry of bone and the ultrastructure of connective tissues. *Mol Cell Biol* **26**, 238-249 (2006).
29. H. Xu, T. S. Acott, M. K. Wirtz, Identification and expression of a novel type I procollagen C- proteinase enhancer protein gene from the glaucoma candidate region on 3q21-q24. *Genomics* **66**, 264-273 (2000).
30. C. F. Baicu *et al.*, Effects of the absence of procollagen C-endopeptidase enhancer-2 on myocardial collagen accumulation in chronic pressure overload. *Am J Physiol. Heart Circ Physiol* **303**, H234-H240 (2012).
31. J. Zhu *et al.*, Regulation of apoAI processing by procollagen C-proteinase enhancer-2 and bone morphogenetic protein-1. *J Lipid Res* **50**, 1330-1339 (2009).
32. K. Heinzel, C. C. Bleul, The Foxn1-dependent transcripts PCOLCE2 and mPPP1R16B are not required for normal thymopoiesis. *Eur J Immunol* **37**, 2562-2571 (2007).
33. R. D. Pollard *et al.*, Procollagen C-endopeptidase Enhancer Protein 2 (PCPE2) Reduces Atherosclerosis in Mice by Enhancing Scavenger Receptor Class B1 (SR-BI)-mediated High-density Lipoprotein (HDL)-Cholesteryl Ester Uptake. *J Biol Chem* **290**, 15496-15511 (2015).
34. O. L. Francone *et al.*, Disruption of the murine procollagen C-proteinase enhancer 2 gene causes accumulation of pro-apoA-I and increased HDL levels. *J Lipid Res* **52**, 1974-1983 (2011).
35. P. Chau, P. E. Fielding, C. J. Fielding, Bone morphogenetic protein-1 (BMP-1) cleaves human proapolipoprotein A1 and regulates its activation for lipid binding. *Biochemistry* **46**, 8445-8450 (2007).
36. H. Xu *et al.*, Pcp2, a Novel Extracellular Matrix Protein, Regulates Adipocyte SR-BI-Mediated High-Density Lipoprotein Uptake. *Arterioscler Thromb Vasc Biol* **41**, 2708-2725 (2021).
37. B. Lynch *et al.*, A novel microstructural interpretation for the biomechanics of mouse skin derived from multiscale characterization. *Acta Biomater* **50**, 302-311 (2017).
38. D. Kruppa *et al.*, Distinct contributions of meprins to skin regeneration after injury - Meprin alpha a physiological processor of pro-collagen VII. *Matrix Biol Plus* **11**, 100065 (2021).
39. D. Massoudi, C. J. Germer, J. M. Glisch, D. S. Greenspan, Procollagen C-proteinase enhancer 1 (PCPE-1) functions as an anti-angiogenic factor and enhances epithelial recovery in injured cornea. *Cell Tissue Res* **370**, 461-476 (2017).
40. N. Mariano *et al.*, Production and characterization of mini-procollagens I, II and III for the analysis of procollagen C-proteinase activity and regulation. (in preparation).
41. L. Moschovich *et al.*, Folding and activity of recombinant human procollagen C-proteinase enhancer. *Eur J Biochem* **268**, 2991-2996 (2001).
42. M. Bekhouche *et al.*, Role of the netrin-like domain of procollagen C-proteinase enhancer-1 in the control of metalloproteinase activity. *J Biol Chem* **285**, 15950-15959 (2010).
43. D. Kronenberg *et al.*, Strong cooperativity and loose geometry between CUB domains are the basis for procollagen c-proteinase enhancer activity 1. *J Biol Chem* **284**, 33437-33446 (2009).
44. D. J. Hulmes, A. P. Mould, E. Kessler, The CUB domains of procollagen C-proteinase enhancer control collagen assembly solely by their effect on procollagen C-proteinase/bone morphogenetic protein-1. *Matrix Biol.* **16**, 41-45 (1997).
45. D. E. Birk, Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron* **32**, 223-237 (2001).
46. C. Anastasi *et al.*, BMP-1 disrupts cell adhesion and enhances TGF-beta activation through cleavage of the matricellular protein thrombospondin-1. *Sci Signal* **13**, eaba3880 (2020).
47. F. Delolme *et al.*, Proteolytic control of TGF-beta co-receptor activity by BMP-1/tolloid-like proteases revealed by quantitative iTRAQ proteomics. *Cell Mol Life Sci* **72**, 1009-1027 (2015).
48. E. M. Gonzalez *et al.*, BMP-1/tolloid-like metalloproteases process endorepellin, the angiostatic C-terminal fragment of perlecan. *J. Biol. Chem.* **280**, 7080-7087 (2004).
49. T. B. Strom, K. Bjune, T. P. Leren, Bone morphogenetic protein 1 cleaves the linker region between ligand-binding repeats 4 and 5 of the LDL receptor and makes the LDL receptor non-functional. *Hum Mol Genet* **29**, 1229-1238 (2020).
50. C. Bijakowski *et al.*, Sizzled is unique among secreted frizzled-related proteins for its ability to specifically inhibit bone morphogenetic protein-1 (BMP-1)/tolloid-like proteinases. *J Biol Chem* **287**, 33581-33593 (2012).
51. Y. Zhang, G. Ge, D. S. Greenspan, Inhibition of bone morphogenetic protein 1 by native and altered forms of alpha2-macroglobulin. *J Biol Chem* **281**, 39096-39104 (2006).
52. D. Kronenberg *et al.*, Processing of procollagen III by meprins: new players in extracellular matrix assembly? *J Invest Dermatol* **130**, 2727-2735 (2010).
53. T. Bedau *et al.*, Ectodomain shedding of CD99 within highly conserved regions is mediated by the metalloprotease meprin beta and promotes transendothelial cell migration. *Faseb J* **31**, 1226-1237 (2017).
54. S. Doll *et al.*, Region and cell-type resolved quantitative proteomic map of the human heart. *Nat Commun* **8**, 1469 (2017).

55. F. Malecaze *et al.*, Upregulation of bone morphogenetic protein-1/mammalian tolloid and procollagen C-proteinase enhancer-1 in corneal scarring. *Invest Ophthalmol Vis Sci* **55**, 6712-6721 (2014).
56. H. B. Schiller *et al.*, Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. *Mol Syst Biol* **11**, 819 (2015).
57. T. Tsukui *et al.*, Qualitative rather than quantitative changes are hallmarks of fibroblasts in bleomycin-induced pulmonary fibrosis. *Am J Pathol* **183**, 758-773 (2013).
58. P. Lagoutte *et al.*, Procollagen C-Proteinase Enhancer 1 (PCPE-1) is a marker of myocardial fibrosis and impaired cardiac function in a murine model of pressure overload. *BioRxiv* 10.1101/2021.03.05.434071 (2021).
59. C. Kuppe *et al.*, Decoding myofibroblast origins in human kidney fibrosis. *Nature* **589**, 281-286 (2021).
60. T. Tabib *et al.*, SFRP2/DPP4 and FMO1/LSP1 Define Major Fibroblast Populations in Human Skin. *J Invest Dermatol* **138**, 802-810 (2018).
61. V. Bergmeier *et al.*, Identification of a myofibroblast-specific expression signature in skin wounds. *Matrix Biol* **65**, 59-74 (2018).
62. J. Etich *et al.*, Gene Expression Profiling of the Extracellular Matrix Signature in Macrophages of Different Activation Status: Relevance for Skin Wound Healing. *Int J Mol Sci* **20** (2019).
63. F. Finkernagel *et al.*, The transcriptional signature of human ovarian carcinoma macrophages is associated with extracellular matrix reorganization. *Oncotarget* **7**, 75339-75352 (2016).
64. H. S. Yoon *et al.*, Procollagen C-Endopeptidase Enhancer 2 Secreted by Tonsil-Derived Mesenchymal Stem Cells Increases the Oxidative Burst of Promyelocytic HL-60 Cells. *Biology (Basel)* **11** (2022).
65. L. Chen *et al.*, Identification of biomarkers associated with diagnosis and prognosis of colorectal cancer patients based on integrated bioinformatics analysis. *Gene* **692**, 119-125 (2019).
66. C. Shi *et al.*, Identification of Ferroptosis-Related Genes Signature Predicting the Efficiency of Invasion and Metastasis Ability in Colon Adenocarcinoma. *Front Cell Dev Biol* **9**, 815104 (2021).
67. S. Tian, G. Meng, W. Zhang, A six-mRNA prognostic model to predict survival in head and neck squamous cell carcinoma. *Cancer Manag Res* **11**, 131-142 (2019).
68. X. Zhang, Y. Wang, Identification of hub genes and key pathways associated with the progression of gynecological cancer. *Oncol Lett* **18**, 6516-6524 (2019).
69. R. Berry *et al.*, Role of dimerization and substrate exclusion in the regulation of bone morphogenetic protein-1 and mammalian tolloid. *Proc Natl Acad Sci USA* **106**, 8561-8566 (2009).
70. H. X. Lee, A. L. Ambrosio, B. Reversade, E. M. De Robertis, Embryonic dorsal-ventral signaling: secreted frizzled-related proteins as inhibitors of tolloid proteinases. *Cell* **124**, 147-159 (2006).
71. W. He *et al.*, Exogenously administered secreted frizzled related protein 2 (Sfrp2) reduces fibrosis and improves cardiac function in a rat model of myocardial infarction. *Proc Natl Acad Sci USA* **107**, 21110-21115 (2010).
72. K. Kobayashi *et al.*, Secreted Frizzled-related protein 2 is a procollagen C proteinase enhancer with a role in fibrosis associated with myocardial infarction. *Nat Cell Biol* **11**, 46-55 (2009).
73. R. Jasuja *et al.*, Bone morphogenetic protein 1 prodomain specifically binds and regulates signaling by bone morphogenetic proteins 2 and 4. *J Biol Chem* **282**, 9053-9062 (2007).
74. H. X. Lee, F. A. Mendes, J. L. Plouhinec, E. M. De Robertis, Enzymatic regulation of pattern: BMP4 binds CUB domains of Tolloids and inhibits proteinase activity. *Genes Dev* **23**, 2551-2562 (2009).
75. A. R. Aricescu, W. Lu, E. Y. Jones, A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr* **62**, 1243-1250 (2006).

## FIGURE LEGENDS

**Figure 1: The absence of endogenous PCPE-2 does not have a major impact on collagen fibrils *in vivo* but exogenous PCPE-2 regulates procollagen I maturation in fibroblast supernatants. (A)** TEM pictures of collagen fibrils in the skin and heart of WT and *Pcolce2*-null (KO) mice (6 week-old). Representative images of n = 8 (skin) or 4 (heart) mice/genotype. Bar = 100 nm (heart) or 200 nm (skin). **(B)** Distribution of collagen fibril diameters for WT and KO mice and median values of all measurements; 3092 fibrils from n = 8 mice were analyzed for each genotype in skin and 807 fibrils from n = 4 mice in heart (\*\*\*\*, p < 0.0001; Mann-Whitney test). The graph of the median diameters calculated for individual mice is also shown (means ± SD; n = 8 mice/genotype for skin and 4 mice/genotype for heart). **(C)** Comparison of the mRNA abundance for the genes encoding PCPE-1 and PCPE-2 by qRT-PCR in normal mouse and human dermal fibroblasts (NMF and NHF respectively). Graphs show means ± SD (n = 4 mice or 6 donors for human; unpaired t-test with Welch's correction: \*, p-value < 0.05). **(D)** Immuno-detection of procollagen I, pC-collagen I and C-propeptide of collagen I (see schematic for a definition; PNP = procollagen N-proteinase and PCP = procollagen C-proteinase) with the anti-C-propeptide I (LF-41) antibody in the supernatant of WT and KO fibroblasts after 8 h of culture without serum. SDS-PAGE was run in reducing conditions and the same membrane was exposed for 1 min to detect the pro- and pC-collagens and for 3 min to detect the C-propeptide. Quantification of band intensities is shown for 3 mice/condition (means ± SD). **(E)** Detection of the same processing products in the supernatant of human adult fibroblasts after incubation for 2 h at 37 °C with buffer or 10 nM BMP-1. When indicated, 750 nM PCPE-1 and/or PCPE-2 were also added to the cleavage assays. The uncropped membrane is available in Fig. S11. **(F)** Quantification was with n = 7-8 independent experiments (conditions shown in the same order as on the western-blot). Means ± SD are shown and conditions were compared using one-way ANOVA with paired data and Dunnett test for multiple comparisons (\*, p-value < 0.05; \*\*, p-value < 0.01).

**Figure 2: The effect of PCPE-2 on procollagen C-terminal maturation is concentration-dependent and involves a direct interaction with procollagens. (A)** Mini-procollagens I, II and III (430 nM) were incubated with 21 nM BMP-1 in the absence or presence of 430 nM PCPE-1 or / and PCPE-2 in a total volume of 30 µL for 1 h. Reaction products were analyzed by SDS-PAGE (4-20 % polyacrylamide gel; non-reducing conditions) and Coomassie Blue staining. Enhancement factors estimated by densitometry are indicated below the gel. Gels are representative of n = 3 independent experiments. (\*) indicates the position of the trimeric N-terminal cleavage product of mini-procollagen III (when visible). **(B)** Cleavage of CPIII-Long (360 nM) by BMP-1 (20 nM) for 1 h in the absence or presence of PCPE-1, PCPE-2, the CUB domains of PCPE-2, the NTR domain of PCPE-2 or combinations of PCPE-1 and PCPE-2 (or its domains). Molar ratios of PCPEs and domains to CPIII-Long are indicated in the table above the gel and enhancement factors are reported below the gel. The gel is representative of n = 3 independent experiments. **(C)** Quantification of CPIII-Long (360 nM) processing by BMP-1 (16.5 nM) in the presence of increasing amounts of PCPE-1 and PCPE-2 (0-2.2 µM). Analysis was by SDS-PAGE (reducing conditions) on a 15 % acrylamide gel with Sypro Ruby staining and quantification of individual substrate bands. **(D)** PCPE-2 binding to immobilized mini-procollagen III (435 RU). PCPE-2 was injected at the following concentrations: 0.78, 1.56, 3.12, 6.25, 12.5, 25 and 50 nM. The best fit of the binding response at equilibrium (steady-state conditions) is also shown (inset). **(E)** Successive injections of PCPE-2 + PCPE-1 on immobilized mini-procollagen III (435 RU) compared to successive injections of PCPE-2 alone; 200 nM PCPE-2 were injected first for 60 s followed by the injection of PCPE-2 (200 nM or 400 nM) alone or in combination with 200 nM PCPE-1 for another 60 s.

**Figure 3: PCPE-2 inhibits the cleavage of BMP-1 substrates. (A-E)** Effect of PCPE-1 and -2 on various BMP-1 substrates at 1:1 PCPE:substrate molar ratios (unless otherwise indicated) at 37 °C. **(A)** The pNα1(V) protein (800 nM) is the monomeric N-terminal region of the α1 chain of collagen V and it was incubated with 20 nM BMP-1 for 2 h to monitor the release of the N-terminal TSPN domain (PCPE to pNα1(V) molar ratios as indicated above the gel). **(B)** TSP-1 (200 nM) was incubated with 70 nM BMP-1 for 4 h to generate one N-terminal fragment (45 kDa) and one C-terminal fragment (120 kDa). **(C)** The ectodomain of betaglycan (780 nM) is cleaved by BMP-1 (50 nM) in two main products

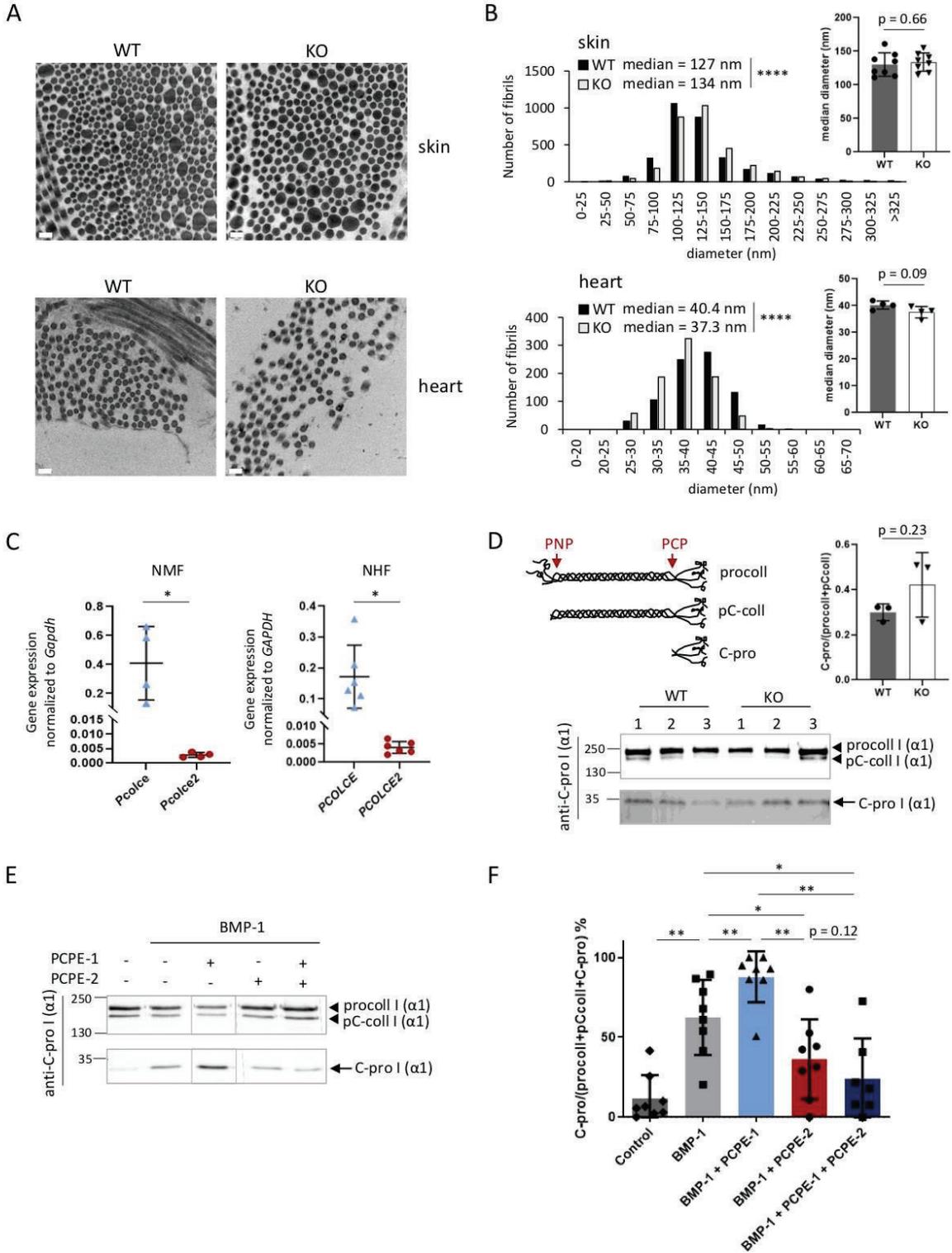
(N-ter, C-ter) in the conditions of the assay (4 h). **(D)** Chordin (present in two forms in the starting material, resulting from alternative splicing according to the manufacturer; 370 nM) yielded 3 products (referred to as N-ter, internal and C-ter) when incubated with 20 nM BMP-1 for 4 h. **(E)** Endorepellin (265 nM) corresponds to the C-terminal part of perlecan and was cleaved by BMP-1 (18 nM) in two fragments (LG1-2 and LG3) after 1 h. **(F)** The ectodomain of LDLR (390 nM) was converted to a shorter fragment (C-ter) in the presence of 1.5 nM BMP-1 for 1 h. In (A-F), detection was by SDS-PAGE (reducing conditions) with Coomassie Blue staining except for endorepellin and LDLR which were detected with Sypro Ruby staining. (\*) indicates BMP-1 position when visible. Uncropped gels for (B-D, F) can be seen in Fig. S11. **(G)** Immuno-detection of endogenous LDLR in 293-EBNA cells stably transfected with an empty vector or the same vector containing sizzled sequence or PCPE-2 sequence. Expression of sizzled and PCPE-2 was also evidenced by immunoblotting with anti-His antibody in the supernatant of transfected cells.

**Figure 4: The CUB domains of PCPE-2 are responsible for its inhibitory activity and the non-catalytic domains of BMP-1 are required for inhibition.** **(A)** Quantification of BMP-1 activity on the fluorogenic peptide Mca-YVADAPK(Dnp)-OH in the absence or presence of increasing concentrations of PCPE-2, PCPE-2 3C, the CUB1CUB2 and NTR domains of PCPE-2 or PCPE-1, at the indicated molar ratios to BMP-1 (7 nM). Means  $\pm$  SD of  $n = 3$  independent experiments run in duplicate. **(B)** The apparent inhibition constant ( $K_i^{app}$ ) of PCPE-2 was measured in the same conditions with PCPE-2 concentrations between 0.07 and 690 nM and calculated using the Morisson's equation. Means  $\pm$  SD of  $n = 3$  independent experiments run in duplicate. **(C)** Effect of PCPE-2 on the cleavage of the fluorogenic peptide by the catalytic domain of BMP-1 (BMP-1cat, 8 nM). Means  $\pm$  SD of  $n = 3$  independent experiments run in duplicate. **(D)** Comparison of the binding of 50 nM of PCPE-1, PCPE-2 and its domains (CUB1CUB2 and NTR) on immobilized BMP-1 (974 RU). **(E)** Comparison of the binding of 50 nM BMP-1 and BMP-1cat on immobilized PCPE-2 (646 RU). **(F)** Fits of sensorgrams obtained with the kinetic (black dotted lines; heterogeneous ligand) or steady-state (inset) model when increasing concentrations of PCPE-2 (1.56-100 nM, prepared as serial two-fold dilutions) were injected over immobilized BMP-1 (974 RU). **(G)** Competition experiments consisting of an injection of 200 nM PCPE-2 for 60 s followed by an injection of PCPE-2 (200 nM) alone or in combination with 50, 100 or 200 nM BMP-1 for another 60s. The immobilized protein was CPIII (351 RU).

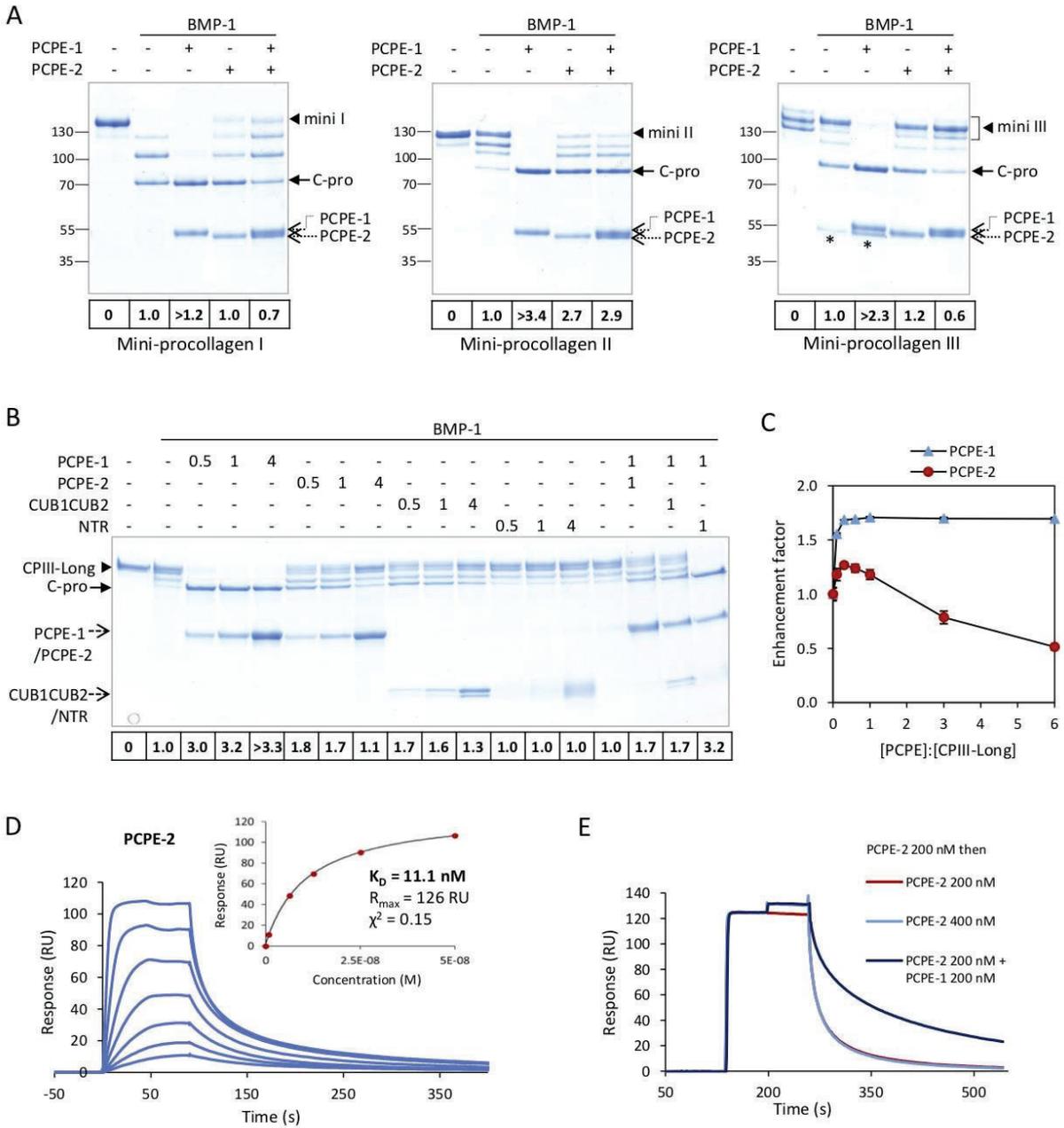
**Figure 5: Effect of PCPE-2 on mTLL-1 and meprins.** **(A)** SDS-PAGE analysis in non-reducing conditions of the cleavage of CPIII-Long (320 nM) by mTLL-1 (17 nM) in the absence or presence of PCPE-1 or PCPE-2 for 4 h. Molar ratios of PCPEs to CPIII-Long are indicated above the gel and the enhancement factors are indicated below the gel. The gel is representative of  $n = 3$  independent experiments. **(B)** Comparison of mTLL-1 (50 nM) and BMP-1 (10 nM) activities on the fluorogenic peptide Mca-YVADAPK(Dnp)-OH, in the absence or presence of PCPE-1 or PCPE-2 (concentrations corresponding to indicated molar ratios to protease). Means  $\pm$  SD ( $n = 3$  independent experiments run in duplicate). **(C)** SDS-PAGE analysis in non-reducing conditions of the cleavage of CPIII-Long (470 nM) by meprin  $\alpha$  (3 nM) or meprin  $\beta$  (1.5 nM), in the absence or presence of equimolar PCPE-1 or PCPE-2 concentrations (470 nM), for 15 min. The uncropped gel is available in Fig. S11 and is representative of  $n = 3$  independent experiments. (\*) indicates the N-terminal cleavage product of PCPE-1 by meprins (52) when visible. **(D)** Comparison of BMP-1 (10 nM), meprin  $\alpha$  (1 nM) or meprin  $\beta$  (0.5 nM) activities on the fluorogenic peptide Mca-YVADAPK(Dnp)-OH, in the absence or presence of PCPE-1 or PCPE-2. Means  $\pm$  SD ( $n = 3-5$  independent experiments run in duplicate).

**Figure 6: Summary of the differences in activity for PCPE-1 and PCPE-2 in the context of BMP-1 regulation.** **(A)** PCPE-1 enhances the C-terminal maturation of procollagens I-III through the binding of CUB1CUB2 to the stalk region of collagen C-propeptides. PCPE-2 binds procollagens and BMP-1 with similar affinities but cannot interact with both partners simultaneously, thereby precluding the efficient processing of procollagens. **(B)** With regard to non-collagenous substrates, PCPE-1 has no effect on their BMP-1-dependent cleavage while PCPE-2 behaves as a direct inhibitor of these cleavages. Note that the incompletely-characterized case of fibrillar collagens V and XI is not represented.

Figure 1



**Figure 2**



**Figure 3**

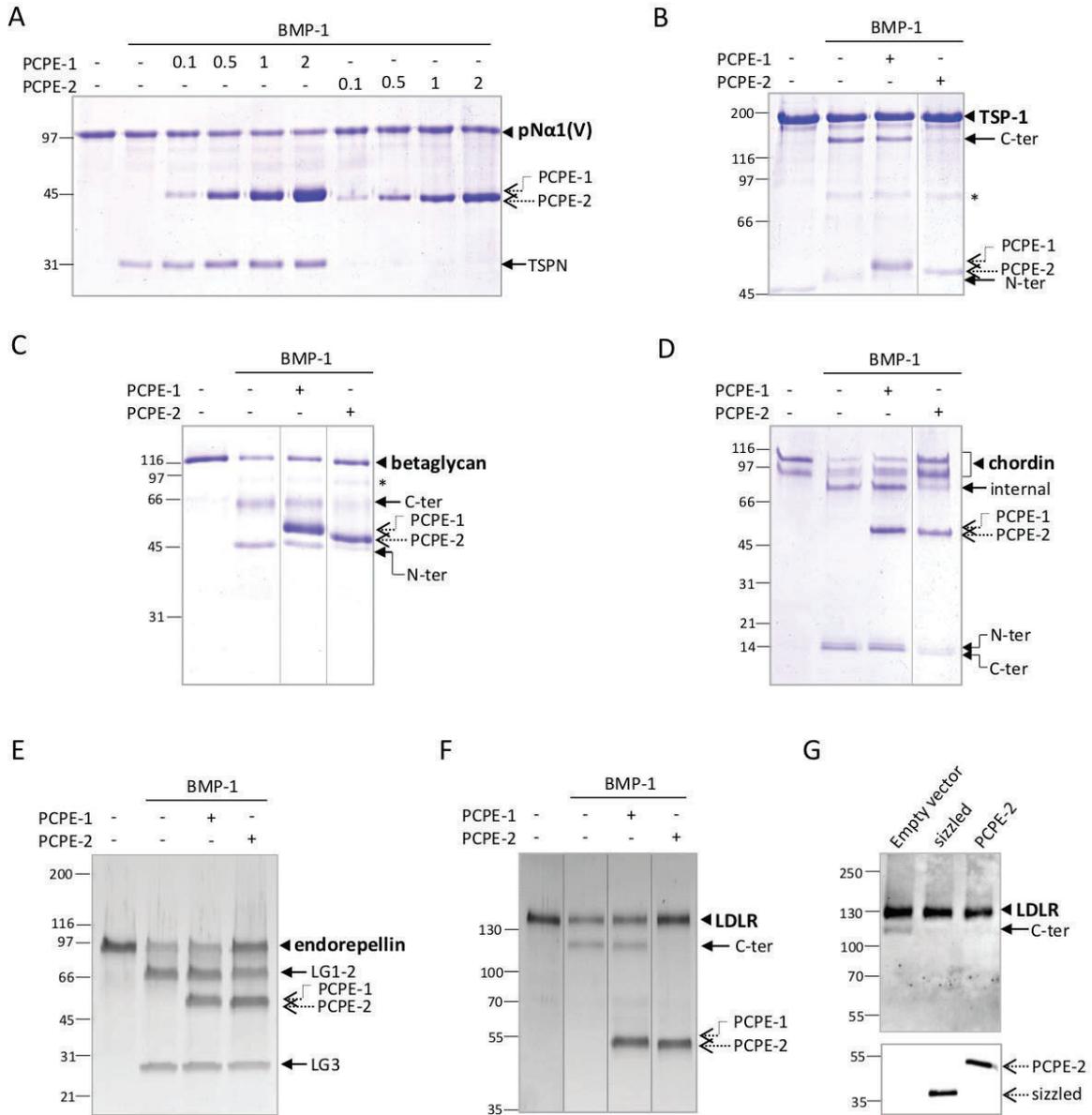
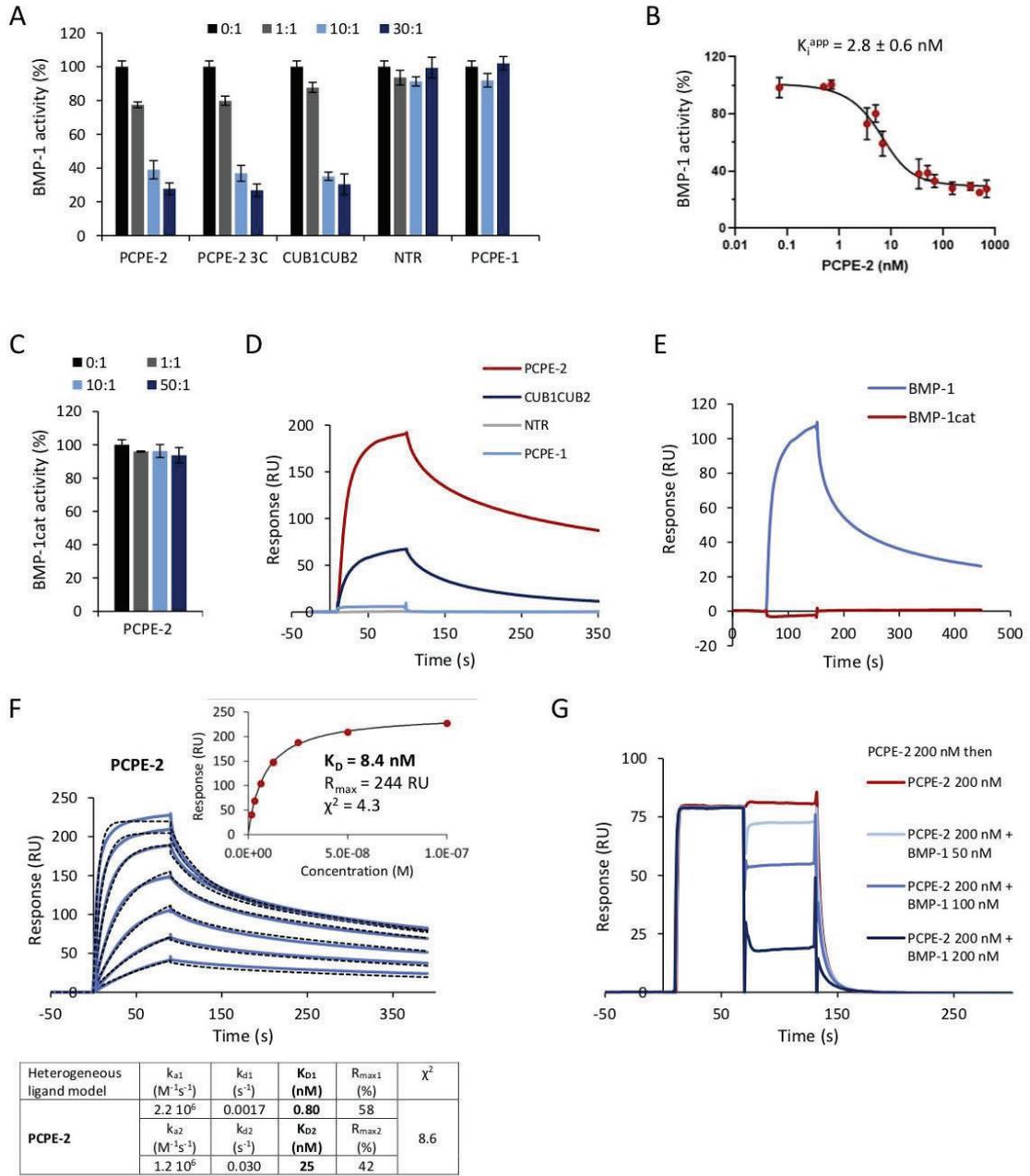
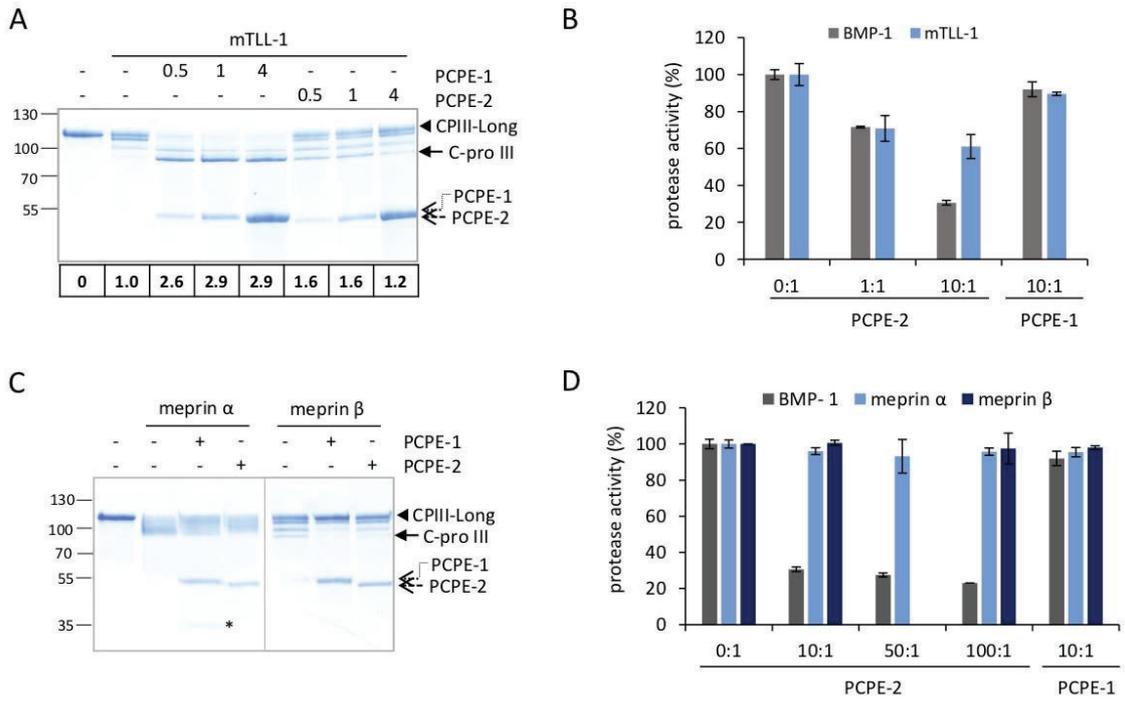


Figure 4

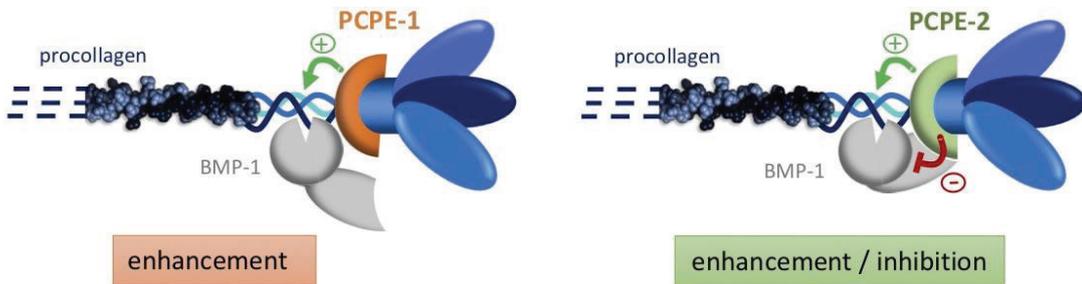


**Figure 5**

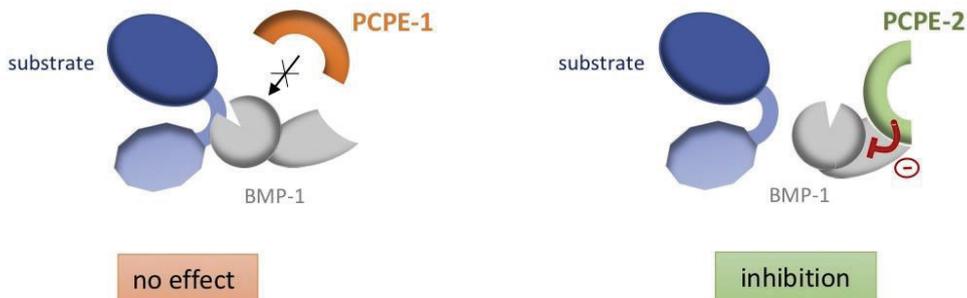


**Figure 6**

**A- Fibrillar collagens I-III**



**B- Non-collagenous BTP substrates**



Article 4: Dussoyer, M., Michopoulou, A. & Rousselle, P. *Appl Sci* (2020)

Review

# Decellularized Scaffolds for Skin Repair and Regeneration

 Mélissa Dussoyer <sup>1</sup>, Anna Michopoulou <sup>2</sup> and Patricia Rousselle <sup>1,\*</sup>

<sup>1</sup> Laboratoire de Biologie Tissulaire et Ingénierie Thérapeutique, UMR 5305, CNRS—Université Lyon 1, Institut de Biologie et Chimie des Protéines, SFR BioSciences Gerland-Lyon Sud, 7 Passage du Vercors, 69367 Lyon, France; melissa.dussoyer@ibcp.fr

<sup>2</sup> BIOHELLENIKA SA, Company of Biotechnology, 65 Georgikis Scholis Av., 57001 Thessaloniki, Greece; amichopoulou@biohellenika.gr

\* Correspondence: patricia.rousselle@ibcp.fr

Received: 23 April 2020; Accepted: 12 May 2020; Published: 15 May 2020



**Abstract:** The skin is the largest organ in the body, fulfilling a variety of functions and acting as a barrier for internal organs against external insults. As for extensive or irreversible damage, skin autografts are often considered the gold standard, however inherent limitations highlight the need for alternative strategies. Engineering of human-compatible tissues is an interdisciplinary and active field of research, leading to the production of scaffolds and skin substitutes to guide repair and regeneration. However, faithful reproduction of extracellular matrix (ECM) architecture and bioactive content capable of cell-instructive and cell-responsive properties remains challenging. ECM is a heterogeneous, connective network composed of collagens, glycoproteins, proteoglycans, and small molecules. It is highly coordinated to provide the physical scaffolding, mechanical stability, and biochemical cues necessary for tissue morphogenesis and homeostasis. Decellularization processes have made it possible to isolate the ECM in its native and three-dimensional form from a cell-populated tissue for use in skin regeneration. In this review, we present recent knowledge about these decellularized biomaterials with the potential to be used as dermal or skin substitutes in clinical applications. We detail tissue sources and clinical indications with success rates and report the most effective decellularization methods compatible with clinical use.

**Keywords:** skin repair; wound healing; decellularization; scaffolds; skin tissue engineering; skin alternate; biomaterial; skin regeneration

## 1. Introduction

Skin is the largest organ in vertebrates and fulfills a variety of functions. Acting primarily as a barrier protecting the internal organs from external insults, it also fulfills other functions such as controlling fluid homeostasis, sensory detection, vitamin D synthesis, immune surveillance, and self-healing. Vertebrate skin consists of the epidermis anchored firmly to an underlying dermis, conferring elasticity and mechanical resistance [1]. The epidermis, which consists primarily of keratinocytes, is continuously renewed by the proliferation of stem cells, which progressively undergo terminal differentiation as they move upward from the basal layer toward the surface, where they die and slough off [2,3]. The dermis consists of the predominating extracellular matrix (ECM), composed of collagens, elastin, and glycosaminoglycans (GAGs) with embedded fibroblasts, the major cellular constituents. The dermis mainly provides the physical strength and flexibility of skin and supports its extensive vasculature, lymphatic system, and nerve bundles. The upper sublayer of the dermis, called the papillary dermis, is a loose ECM consisting of a fine network of collagen fibers. The papillary dermis includes a large amount of nerve fibers, capillaries, water, and fibroblasts [4,5]. The reticular

layer, constituting the lower part of the dermis, has a denser and thicker ECM network with fewer nerve fibers and capillaries. In this sublayer, collagen fibers are aggregated into thick bundles that are largely aligned parallel to the skin surface [6,7]. Beneath the reticular dermis is the adipose tissue, which provides insulation from heat and cold and performs endocrine functions involved in food intake, glucose homeostasis, lipid metabolism, inflammation, and angiogenesis. Recent findings henceforth discriminate the dermal white adipose tissue, which consists of intradermal pre-adipocyte and adipocyte populations, from the subcutaneous white adipose tissue [8].

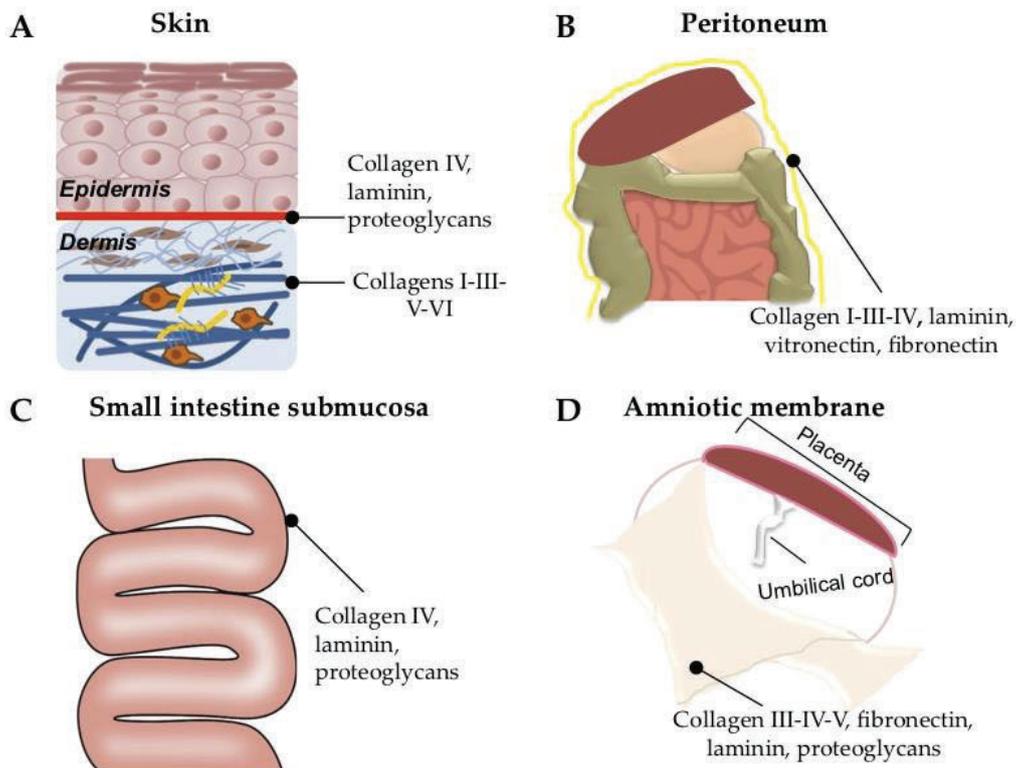
Damage or loss of integrity of the skin caused by wounds may impair the functions of the skin to varying extents. Wounds can be caused by mechanical trauma, surgical procedures, reduced blood circulation, burns, or aging. Most skin wounds can heal naturally through dynamic and interactive processes, which involve soluble factors, blood elements, ECM components, and cells [9]. However, in cases of extensive or irreversible damage, surgical strategies are often needed to provide immediate skin coverage using skin substitutes or scaffolds to guide repair and regeneration. This scenario is typical for chronic wounds that do not proceed through the normal phases of wound healing for multifactorial reasons [10]. A chronic wound may enter a persistent inflammatory state and perpetual non-healing state characterized by chronicity and frequent relapses. A number of pathophysiological factors can cause a failure of normal wound healing, including inflammation, infection, malnutrition, age, diabetes, tissue maceration, pressure necrosis, and renal impairment [11–13]. The restoration of dermis requires three-dimensional (3D) scaffolds to provide elasticity and strength to the epidermal graft and to feed the keratinocytes in the epidermal layer. Thus, one crucial factor in skin tissue engineering is the construction of a tissue scaffold that serves as a template to guide restructuring of cells and subsequent host infiltration of the graft.

In the case of important tissue injury or disease, tissue autografts are often considered the gold standard, but inherent limitations including donor site morbidity, low availability, and troubling failure rates highlight the need for alternative strategies [14,15]. To date, engineering human tissues is an interdisciplinary and active field of research [16–18], but fully reproducing the properties of the ECM is a great challenge [19–21]. The involvement of ECM proteins in tissue repair is well documented [22–24], and the development of ECM-biomimetic and -bioinspired materials is undergoing considerable growth. Moreover, despite technological progress and advances in polymer science, the vast majority of artificial scaffolds do not perfectly mimic ECM scaffolding and bioactive cell-instructive and cell-responsive properties. These aspects are particularly challenging for skin repair because regeneration of the different skin layers requires different cell types and matrix elements.

Among the essential noncellular components, the ECM is a heterogeneous, connective network composed of fibrous glycoproteins, proteoglycans, and small molecules that coordinate *in vivo* to provide the physical scaffolding, mechanical stability, and biochemical cues necessary for tissue morphogenesis and homeostasis. Decellularization allows acquisition of cell-free, natural ECMs. The pronounced effect of decellularized ECM scaffolds in supporting tissue regeneration is based on two major characteristics: first, the maintenance of the 3D structure, providing support, tensile strength, and attachment sites for cell surface receptors; and second, the availability of bioactive components that modulate angiogenesis, cell migration, and cell proliferation and orientation in wound healing [25].

## 2. Tissue Sources for Preparation of ECM-Based Biomaterials for Skin Repair

In the last decades, decellularized scaffolds for use as skin substitutes have been intensively developed from various tissue sources such as dermis, skin flaps, peritoneum, intestinal tissues, and amnion/chorion tissues (Figure 1).



**Figure 1.** Schematic representation of tissue sources for the preparation of decellularized matrix to support skin repair and regeneration. Basic structure, composition and location of each tissue source are indicated for (A) Skin, (B) Peritoneum, (C) Small intestine submucosa, (D) Amniotic membrane.

Allogeneic sources, such as tissues from cadaveric donors, and xenogeneic materials, such as porcine, bovine, and goat tissues, have been used, with different long-term results [26]. The decellularization protocols typically combine physical and chemical treatments to remove cellular antigens, involving freeze-thawing or detergents and/or enzymes that can alter the 3D structure of the ECM [27]. Additionally, the US Food and Drug Administration (FDA) requirements impose that acellular mammalian-derived products undergo “viral inactivation,” which implies the use of detergents resulting in the removal of many important biologic components from the tissue, such as lipids, glycans, and elastins [28]. As an alternative, skin grafts from fish such as Atlantic cod have been proposed because viral and prion transmission risk is nonexistent. These tissues can be subjected to gentle processing that does not disrupt the structure or bioactive composition of the original ECM [29].

Thus, decellularized tissues can be obtained from a variety of sources and implanted either orthotopically, i.e., the reconstructed tissue exhibits similar properties to those of the source, or ectopically, i.e., to a site that is completely different from origin site of the decellularized tissue [30]. In this review, we focus on decellularized biomaterials with the potential to be used as dermal or skin substitutes for clinical applications.

### 2.1. Acellular Dermal Matrix (ADM)

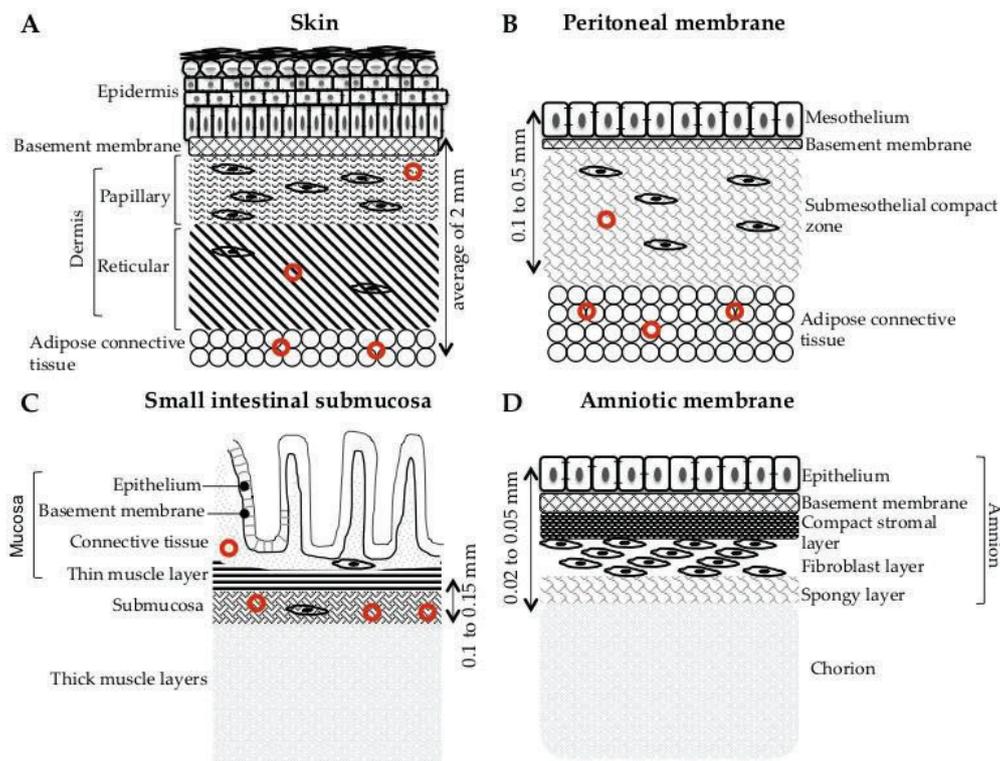
A decellularized dermal matrix is an acellular tissue made by taking a full-thickness section from a donor source, which in most cases is a human cadaver or of porcine or bovine origin. In the case of human donors, the tissue is screened for infectious agents, such as HIV, hepatitis, and syphilis [31]. In the last two decades, ADM has become increasingly popular for the coverage of soft-tissue open wounds. Artificial dermal substitutes are routinely used in the treatment of chronic and acute injuries, particularly when autologous intact cutaneous tissue is not available [32]. Although skin grafting is significantly less expensive, the use of ADM can be a successful alternative to a painful and aesthetically

undesirable donor site [31]. In addition, as has long been recognized, the presence of dermis enhances the success of subsequent skin transplantation, inhibits granulation tissue formation, and prevents scar formation, thereby reducing vascular contraction [26]. Thus, ADM is often used in conjunction with split-thickness grafts for the treatment of full-thickness wounds [31]. Aside from burn wounds and diabetic foot ulcers (DFUs), more recently, ADMs have been suggested as a successful alternative to vascularized flaps for extremity wounds with exposed tendons, bones, and joints [32]. ADMs are currently produced and developed as various commercial products, and here we focus on commercially available skin substitutes with substantial relevant evidence of efficacy.

The gold standard for temporary coverage is a skin allograft, i.e., the skin is transplanted from one person to another. As with autografts involving graft transfer from the same individual, allografts undergo vascularization within 2–3 days and can provide a variable barrier after a severe burn. Unlike autografts, however, the skin is highly immunogenic, and when a viable skin allograft is transplanted onto a healthy recipient, rejection occurs within 6–14 days [33]. Allograft ADM products derived from donated human skin are supplied by tissue banks compliant with standards of the American Association of Tissue Banks and FDA guidelines. A long list of products derived from minimally processed human donor tissue is available, with different durations of covering depending on their design and composition [34] including the following: AlloDerm™ regenerative tissue matrix (RTM; LifeCell Corp., Branchburg, New Jersey, USA), AlloPatch HD® Pliable (Musculoskeletal Transplant Foundation), Cymetra™ Micronized AlloDerm™ (injectable form of AlloDerm™, LifeCell/KCI), Dermacell® (LifeNet Health), Flex HD® (Ethicon and Musculoskeletal Transplant Foundation), GammaGraft® (Promethan Lifesciences, Inc.), GraftJacket® RTM (LifeCell, licensed to Wright Medical Technology and KCI), Glyaderm® (Euro Skin Bank, Beverwijk, The Netherlands), Matrix HD™ (RTI Biologics), Memoderm™ (Memometal, Inc.), Puros Dermis® (Zimmer Dental), and Repliform® (LifeCell and Boston Scientific) [35]. At least five different manufacturing processes and products (AlloDerm™, DermaMatrix, Glyaderm®, GraftJacket®, and SureDerm®) are currently registered for wound care (Table 1) [33]. We discuss here in detail the most cited versions considering their efficacy in treating burns and non-healing/difficult-to-heal wounds.

#### 2.1.1. AlloDerm™ RTM (LifeCell Corp., Branchburg, NJ, USA)

AlloDerm™ (LifeCell Corp., Branchburg, New Jersey, USA) is one of the first FDA-approved acellular matrix materials and among the most extensively investigated [25]. It has been used since 1992 to treat burns with successful engraftment of both the AlloDerm™ and the subsequent overlying skin graft in many individuals [36]. This product was FDA approved for the replacement of inadequate integument tissue, but the most common cited application has been to cover the skin flaps of donor sites as a sheet graft without an overlying split-thickness graft (Table 1) [37]. AlloDerm™ is processed directly from fresh cadaver skin treated with high salt to remove the cellular components. The treated skin is then freeze-dried, leaving an ADM with an intact basement membrane complex (Figure 2A) [25]. AlloDerm™ is hydrated before use, typically in an antibiotic saline solution according to the manufacturer's recommendations. The meshing of the ADM at a 1:1 ratio optimizes revascularization and nutrition of the overlying skin graft [25]. Finally, *in vitro* recellularization of AlloDerm™ with allogeneic keratinocytes, adipose-derived stem cells, fibroblasts, and umbilical vein endothelial cells allows for production of biological skin substitutes [25].



**Figure 2.** Detailed schematic histological representation of tissues used for decellularization for skin repair and regeneration. The various tissue compartments of skin (A), peritoneal membrane (B), small intestinal mucosa (C) and amniotic membrane (D) are represented by separate graphics and annotated on the side. When a tissue compartment is cellularized, cells are drawn in black and when it is vascularized, vessels are represented by red circles. The tissue layers used for decellularization and clinical applications are shown by arrows on the side indicating their average thickness.

### 2.1.2. AlloPatch<sup>®</sup> Pliable (Musculoskeletal Transplant Foundation, Edison, NJ, USA)

AlloPatch<sup>®</sup> Pliable is an aseptically processed human reticular (HR)-ADM, intended for use in chronic or acute wound coverage [34]. It requires no rehydration or refrigeration prior to use and can be stored at ambient temperature according to the manufacturer's recommendations. This dermis differs from many other ADMs in that it does not contain a papillary portion, resulting in a homogeneous matrix that retains its original architecture and its key ECM components (Figure 2). The absence of asymmetry or orientation is considered beneficial in the clinical setting, having a positive impact on cell infiltration and native tissue remodeling [38]. In a prospective, randomized, controlled, multicenter study, this HR-ADM was shown to be superior in promoting DFU healing as compared to standard of care involving daily dressing changes with a collagen alginate matrix [39].

### 2.1.3. DermACELL<sup>™</sup> (LifeNet Health)

DermACELL<sup>™</sup> is a decellularized regenerative human tissue matrix allograft that contains both the reticular and papillary compartments with a basement membrane (Figure 2). Upon application, the reticular site is placed against the surgical wound. The patented preparation process for DermACELL<sup>™</sup> includes the use of anionic detergents and endonucleases to eliminate more than 97% of the nucleic acids. It is preserved and stored at room temperature and thus can be delivered hydrated [40]. DermACELL<sup>™</sup> is indicated for chronic non-healing wounds [34] (Table 1). In a prospective, multicenter, single-arm clinical trial for the treatment of large complex DFUs with exposed tendon or bone, DermACELL<sup>™</sup> rapidly reduced the size of the wound. However, that study favored cases that required a single application of the ADM [41]. Additionally, in a case report [40], DermACELL<sup>™</sup> was successfully used to treat a patient with scarring from a severe burn. More recently, a case report publication

provided evidence that DermACELL™ can eventually be used to treat fingertip injuries, e.g., after amputation [42]. An ongoing open-label trial will assess the effectiveness of this ADM in patients with chronic venous ulcers.

#### 2.1.4. GraftJacket® RTM (Wright Medical Technology, Inc., Licensed to KCI)

GraftJacket® RTM is an intact human ADM containing a basement membrane (Figure 2), processed with a patented method to minimize damage to the scaffold and freeze-dried to remove moisture. GraftJacket® RTM is intended for replacement of damaged or inadequate integumental tissue, such as DFUs, venous leg ulcers, pressure ulcers, or other homologous uses on human integument [43]. A quantitative analysis of comparative literature including randomized, prospective controlled clinical studies with full-thickness DFU evaluated the efficacy of this ADM compared to standard moist wound care. The results demonstrated improved healing of DFUs in a significantly reduced time with this ADM. The authors highlighted the fact that a single application achieved this outcome, whereas multiple applications are necessary with other commercialized ADMs [44].

Despite their effectiveness, particularly toward wound bed vascularization, skin allografts present some drawbacks. These include limited availability, potential for pathogen transmission, and cost [33].

#### 2.1.5. Animal-Derived ADMs

Grafts from xenogenic dermal substitutes are often applied to extensive skin defects caused by severe burns. Although human ADMs and engineered human skin matrices have shown satisfactory effects in the treatment of extended second- and/or third-degree burns when xenogenic ADMs lack histocompatibility, human ADMs are so expensive that their use is often impossible in these indications. The necessity of novel cost-effective dermal substitutes with equivalent efficiency has led to the search for alternative solutions [45]. The critical barrier to xenogenic scaffolds in translational application is the recipient immune response to the antigenic components of the xenogenic tissue [46]. Despite initial speculation, the decellularization process is insufficient to eliminate antigens from xenogenic tissues because both cells and the ECM components themselves can elicit the host immune response. In addition, the remaining cell debris can be immunogenic [47]. Some strategies to achieve removal of specific antigens from xenogenic scaffolds are being developed and are presented later in this review. These include enzymatic removal of known xenoantigens, such as  $\alpha$ -gal epitopes, stepwise solubilization-based antigen removal, or a combination of these approaches [47].

Although the animal skin can be extremely adherent and occasionally incorporated into the wound, xenografts are not yet considered true grafts or transplants. Xenografts do not vascularize, but they can still be beneficial to wound management by reducing the pain in partial-thickness injuries and improving healing rates. Such ADMs are being used for exfoliative skin disorders [48].

An extensive list of animal-derived ADM products is currently available, most of which achieve success in clinical use. Since the 1960s, pigs have served as the primary donors for xenografts in the United States, by virtue of their relative affordability and histological structural similarities to the human skin [48]. Several reports have attested to the benefits of using porcine skin in the treatment of extensive wounds. These advantages include reduced healing rates for partial-thickness wounds and granulating wounds and reduced pain with placement over burns [48]. Currently available acellular porcine dermal matrices include Permacol™ (Tissue Sciences Laboratories), Strattice™ (LifeCell Corp.), Collamend™ (Bard), Xenoderm (mbp), and XenMatrix™ (Davol, Inc.) [31]. According to the UnitedHealthCare® Medicare Advantage Policy Guideline for 2020, porcine skin dressings have been suggested as possibly necessary as an occlusive dressing for burns, donor sites of a homograft, and decubiti and other ulcers [49]. Permacol™ is composed of porcine dermal collagen that is subjected to crosslinking, which is suggested to improve tensile strength. Strattice™ is a non-crosslinked porcine-derived ADM in pliable and firm versions, which can be stored at room temperature and is available fully hydrated.

Materials from different sources and processing exhibit different biochemical and mechanical properties and host responses upon implantation [50]. Therefore, suitable dressing material must be used based on the wound type.

PriMatrix™ (TEI Biosciences) is an ADM processed from fetal bovine skin. It was cleared for marketing by the FDA for partial- and full-thickness wounds, for diabetic, pressure, and venous stasis ulcers, and for surgical wounds [51]. This fetal bovine collagen scaffold is processed to remove all lipids, carbohydrates, and non-collagenous proteins, leaving primarily intact fetal dermal collagen fibers with a great proportion of fibrillar collagen III [52]. This PriMatrix™ collagen III rich substrate is associated with healing in developing tissues, whereas the unique structural matrix of fetal tissues has been identified as a contributing factor in scarless wound healing [50]. SurgiMend™ PRS (TEI Biosciences) is another ADM processed from fetal bovine dermis.

An acellular fish dermis has also been FDA approved. Kerecis™ Omega3 Wound (Kerecis) is an ADM rich in omega-3 fatty acids, intended for use in burn wounds, chronic wounds, and other applications. Human skin differs from fish skin mainly because of the adaptation of the latter to the aqueous environment through the development of scales, secretion of mucus from epithelial cells, lack of superficial keratinized layers, and presence of two basement membranes [53]. Additionally, fish skin heals faster than human skin without leaving scar tissue [54]. The histological analysis of human skin after a Kerecis™ graft demonstrated that the ADM is incorporated into the damaged area and infiltrated by autologous cells, is resistant to bacteria, and exhibits anti-inflammatory properties because of the omega-3 fatty acids [29,54].

## 2.2. Decellularized Mesothelium

The mesothelium is a simple squamous epithelium lining the walls of large cavities (peritoneum, pericardium, and pleura) (Figure 1B) [55]. The mesothelial tissues have a thin and strong layer of ECM that supports epithelial cells, which are capable of rapid wound healing (Figure 2B) [56]. Decellularized mesothelium is obtained using detergent agents in a process designed to maximize the preservation of the ECM architecture and composition [57]. A list of decellularized mesothelial matrices of porcine, bovine, equine, and human origins is currently available on the market for use in regeneration of different tissues and organs, mainly soft tissues including skin, vascular tissues, and valve replacements [26,58].

### 2.2.1. Decellularized Intestine and Urinary Bladder

One of the most studied decellularized matrices for a variety of biomedical and tissue-engineering applications derives from the intestine. The ileum, jejunum, and small intestinal submucosa (SIS) are the sources of decellularized intestine scaffolds (Figure 1C) [26]. Porcine SIS is derived from the thin, translucent tunica submucosa layer of the small intestine, which remains after removing the mucosal and muscular layers (Figure 2C). Once processed, SIS biomaterials constitute cross-linked 3D matrices consisting of constitutive proteins that are mostly collagen, with smaller amounts of carbohydrates and lipids [59]. This material supports the formation of a skin barrier 14 days after seeding with human keratinocytes, fibroblasts, and endothelial cells.

The role of SIS in promoting wound closure has been extensively investigated [25]. Incubation with SIS inhibits matrix metalloproteinases, which are important contributors to wound chronicity and are abundantly expressed in chronic wounds [60,61]. Moreover, SIS scaffolds contain transforming growth factor (GF)- $\beta$ , basic fibroblastic GF, vascular endothelial GF, and epidermal GF, thus promoting angiogenesis, cell growth, and differentiation, and consequently tissue regeneration [26]. Advantageous properties of SIS scaffolds in healing wounds include their effective barrier function of the wound bed dehydration because of their low porosity, and their strength and flexibility even after a long-term storage in a lyophilized form [59].

Oasis® Wound Matrix (Cook Biotech, Inc., West Lafayette, IN, USA) is a collagen scaffold derived from porcine intestine submucosa cleared for marketing by the FDA for the management of partial and full-thickness wounds, including pressure, venous, diabetic, and chronic vascular ulcers, surgical wounds, and draining wounds. A clinical study revealed that in a cohort of patients carrying with chronic leg ulcer treated with SIS, complete wound closure was achieved in 55% cases. The use of SIS seemed to accelerate healing rates and to reduce the incidence of ulcer recurrence [62]. Decellularized

porcine urinary bladder matrix (UBM) also has been reported to be an alternative to complex surgical procedures involving flap-associated donor site morbidity to treat complex extremity wounds. Use of UBM seemed to be rather beneficial for the management of patients suffering from comorbidities such as obesity, diabetes mellitus, and cardiopulmonary compromise [27]. After UBM decellularization, the remaining components include collagen, fibronectin, laminin, GAGs, and GFs. The UBM is likely to act through induction of bioactive GFs that support angiogenesis and facilitate proliferation of connective tissues, and the intact basement membrane is conducive to epithelial and endothelial cell proliferation [63].

Commercially available porcine UBM products include MicroMatrix® (Acell, Inc., Columbia, MD, USA) and Cytal® Wound Matrix (Acell, Inc., Columbia, MD, USA). Regarding the mechanism of action, a preclinical study showed rapid cell infiltration into the UBM after application, followed by a robust angiogenic response and new host tissue deposition [64]. Another study demonstrated induction of a different immune response as compared to classical wound healing, one that involves M2-phenotype macrophages related to tissue remodeling rather than the M1 macrophages that are associated with scar tissue formation [65].

### 2.2.2. Decellularized Amniotic Membrane (AM)

The AM surrounds the fetus in the uterus and consists mainly from top to bottom of an epithelium layer, basement membrane, compact ECM layer, fibroblast layer, and spongy layer (Figures 1D and 2D). Fresh human AM contains collagen, fibronectin, and hyaluronic acid along with a combination of GFs, cytokines, and anti-inflammatory proteins [66]. Several studies have shown that specific AM products may be useful for treating several conditions in adults, including non-healing lower-extremity DFUs. AMs are harvested immediately after birth, cleaned, sterilized, and either cryopreserved or dehydrated, maintaining their properties [67]. Fully decellularized AM can be prepared by incubation of fresh tissue with hypotonic Tris buffer containing protease inhibitors, sodium dodecyl sulfate (SDS), DNAase, RNAase, and 1% trypsin-EDTA (ethylene-diamine-tetraacetic acid) for complete elimination of epithelial and mesenchymal cells [68]. Human AM is characterized by anti-bacterial, anti-inflammatory, and non-immunogenic properties, promotes reduced pain and dehydration, and favors the re-epithelialization process [69].

Some commercially available acellular AM products suggested for use in DFUs include AmnioBand® Membrane (MTF Biologics), Biovance® (Celularity, Inc.), and EpiFix® (MiMedx). Despite these many advantages, disadvantages of AM include poor mechanical properties and a high biodegradability rate, which complicate their extensive use in clinic [69]. Finally, the *in vitro* use of acellular human AM allows the construction of living skin equivalents through seeding of human keratinocytes on the epithelial side and fibroblasts on the chorionic side (Figure 2) [70,71].

### 2.2.3. Decellularized Skin Flaps

A skin flap consists of healthy skin and adipose tissue (Figure 2A) or skin, fat, and muscle. Clinically, skin flaps are used to cover full-thickness, serious wounds [26] but are not always an option and can result in donor site morbidity and poor aesthetics. Additionally, split-thickness, autologous skin grafts often used to cover large wounds transplant only a small portion of the dermis and are thus subject to scarring and contracture. Full-thickness skin grafts, including epidermis and full dermis, are difficult to obtain but are subjected to less contracture and result in a more functional graft and a better cosmetic final outcome. However, they depend on vascular ingrowth. Therefore, generating a decellularized full-thickness skin flap with a perfusable vascular pedicle could eventually yield substantial regenerative potential. In two recent studies, decellularized skin flaps from animal sources were developed and recellularized and/or integrated into host tissue. Zhang et al. [72], using a perfusion protocol combined with agitation, produced a skin/fat decellularized matrix that maintained the native ECM architecture and composition of mainly collagen, GAGs, and vascular endothelial GF and absence of the major class I-histocompatibility complex. Immunohistochemical analysis of the matrix showed peripheral nerves and vessel structure maintenance, and angiography confirmed the existence of the

vascular pedicle and microcirculatory networks. Those scaffolds could subsequently be re-populated with human adipose-derived stem cells and endothelial cells and promote neovascularization after implantation in a rat model [72]. Jank et al. [73], using a porcine model, developed a full-thickness decellularized skin flap by applying a perfusion technique. The resulting scaffold comprised a vascular pedicle, maintained the ECM architecture and composition, and could be successfully perfused and anastomosed to the recipient vascular system. The scaffold demonstrated good biocompatibility and ability to support human keratinocyte engraftment and formation of an epidermis with intact barrier function [73]. These results are rather promising and may represent an adequate solution for the treatment of wounds, especially those characterized by loss of tissue volume.

**Table 1.** Commercially available decellularized scaffolds with potential use in skin repair.

Source Tissue	Product	Application Focus
<b>Acellular Dermal Matrix</b>		
<b>Human skin</b>	AlloDerm™ Regenerative Tissue Matrix (RTM, LifeCell Corp.)	Burns, Soft tissue ridge augmentation
	AlloPatch® Pliable (Musculoskeletal Transplant Foundation, Edison, NJ, USA)	Chronic or acute wound coverage
	Cymetra™ Micronized AlloDerm (injectable form of AlloDerm, LifeCellKCI)	Damaged tissue, facial reconstructive procedures
	DermaCell™ (LifeNetHealth)	Reconstructive procedures, chronic wounds
	Flex HD®, Structural, Acellular Hydrated Dermis (Ethicon and Musculoskeletal Transplant Foundation)	Abdominal wall repair
	GammaGraft® (Promethean LifeSciences, Inc.)	Temporary grafts on burns and chronic wounds
	GraftJacket® Regenerative Tissue Matrix (LifeCell, licensed to Wright Medical Technology and KCI)	Soft tissue repair
	Glyaderm® (EuroSkin Bank, Beverwijk, The Netherlands)	Replacement in deep burns, soft tissue defects, unstable scars
	Matrix HD™ (RTI Biologics)	Reconstructive surgery, chronic skin wounds
	Memoderm™ (Memometal Inc.)	Diabetic foot ulcers, soft tissue repairs
Puros Dermis® (Zimmer Dental)	Soft tissue augmentation	
Repliform® (LifeCell and Boston Scientific)	Repair/replacement of soft tissue	
Permacol™ (Tissue Sciences Laboratories)	Burns, ulcers, abdominal wall repair	
Strattice™ (LifeCell Corp.)	Soft tissue repair and body wall defects	
Bard CollaMend™ Implant (Davol, Inc.)	Soft tissue defects	
<b>Porcine skin</b>	Xenoderm (mbp)	Burns and chronic wounds
	XenMatrix™ (Davol, Inc.)	Surgical repair of damaged soft tissues, plastic and reconstructive surgery
<b>Fetal bovine skin</b>	PriMatrix™ (TEI Biosciences)	Partial-/full-thickness wounds, burns, diabetic foot ulcers
	SurgiMend™ PRS (TEI Biosciences)	Hernia repair, plastic and reconstructive surgery
<b>Fish skin</b>	Kerecis™ Omega <sup>3</sup> Wound (Kerecis)	Partial-/full-thickness wounds, trauma, chronic wounds and diabetic ulcers
<b>Decellularized Small Intestine Submucosa (SIS) and Urinary Bladder Matrix (UBM)</b>		
<b>Porcine SIS</b>	Oasis® Wound Matrix (Cook Biotech, Inc., West Lafayette, IN, USA)	Partial-/full-thickness wounds chronic wounds and diabetic ulcers
<b>Porcine UBM</b>	MicroMatrix® (Acell, Inc., Columbia, MD, USA)	Acute and chronic wounds
	Cyta™ Wound Matrix (Acell, Inc., Columbia, MD, USA)	Acute and chronic wounds
<b>Decellularized Amniotic Membrane Matrix</b>		
<b>Human allograft</b>	AmnioBand® (MTF Biologics)	Acute and chronic wounds
	Bioavance® (Celularity, Inc.)	Surgical covering, part-/full-thickness acute and chronic wounds
	EpiFix® (MiMedx)	Acute and chronic wounds
<b>Decellularized Mesothelium Matrix</b>		
<b>Porcine mesothelium</b>	Meso BioMatrix® (MTF Biologics)	Soft tissue
<b>Bovine pericardium</b>	Veritas®, Dura-Guard®, Peri-Guard®, Vasu-Guard® (Baxter Healthcare)	Soft tissue
	CopiOs® (Zimmer)	Dentistry
	Lyoplast® (B. Braun Melsungen)	Dura mater
	Perimount (Edwards Lifesciences)	Valve replacement
<b>Equine pericardium</b>	Unite™ Biomatrix (Synovis Orthopedic and Woundcare)	Soft tissues and chronic wounds
<b>Human pericardium</b>	DurADAPT™ (Pegasus Biologics)	Dura mater
	IOPatch™ (IOP)	Ophthalmology

### 3. Decellularization Methods

Both academic and industrial researchers have developed protocols to decellularize various native tissues, including those mentioned above. They are classically sorted into three categories of chemical, enzymatic, and physical methods. Almost all protocols, however, consist of combinations of these three approaches. To investigate these candidates in as many trials as possible, decellularization studies furthermore apply and compare very different protocols combining various agents instead of varying only one substance to assess its effects.

We here review decellularization methods specifically used for the production of acellular matrices for skin repair and evaluate the properties and effects of each substance commonly used in literature. The results of this analysis led us to propose another classification of these substances into main decellularizing agents that can disrupt cells and eliminate all DNA content when used alone (detergents, acids and bases, physical-based methods) and complementary agents that are added to improve these effects. Particularly, enzymes such as proteases are reviewed as common complementary agents because their exclusive use, which has been tested several times, never demonstrated total decellularization or satisfying *in vitro* cell repopulation of the matrices obtained.

#### 3.1. Complementary Agents

##### 3.1.1. Enzymes

Nucleases are the most common complementary agents included in decellularization protocols independently of the source of native tissue. Nucleic acids, especially DNA, are well known for their high immunogenic properties. To limit nucleic acid levels as much as possible, treatments with DNase I, DNase II, or benzonase are performed after exposure of the tissue to the main decellularizing agent(s) [68,74,75]. Applied concentrations, ranging from 50 U to 10 kU/mL, depend on the properties of the tissue (type, thickness, cell density) and the exposure to main decellularizing agent(s) (nature, duration, concentration, temperature). Regarding the choice between benzonase or DNase, no clear rationale is available in the literature. Finally, for highly cellularized tissues, processing can also include RNase treatments.

As mentioned above, strict protease-based approaches are not sufficient to remove all cells from the tissues used for the production of skin-repair matrices. Accutase, a commercial mix of proteolytic and collagenolytic enzymes classically used for soft cell-detachment from two-dimensional (2D) culture supports, failed to decellularize human AM: the epithelial layer was totally removed after 1 h of incubation, but many mesenchymal cells were still detected [76]. Trypsin has also been tested on human AM combined with EDTA, allowing for complete elimination of cells in both epidermal and mesenchymal compartments and retention of basement membrane components. However, the ability of the matrix to drive sufficient primary fibroblast infiltration and proliferation after a 7-day *in vitro* culture has not been demonstrated. The same group investigated the effects of dispase, a bacterial protease targeting collagen IV, laminin, and fibronectin [77]. They achieved complete decellularization but observed ECM alterations and important losses in basement membrane structure, as had been already reported [78,79]. Cell seeding was not assayed because of the low probability of epithelialization success using a basement membrane with such damage.

Although still undocumented as an efficient decellularized agent, trypsin can be combined with physical, acid-, base-, or detergent-based protocols [80–82]. The action of trypsin is achieved by immersion of the tissue in an isotonic solution containing from 0.25% to 0.1% w/v of enzyme, separately from the other treatments of the protocol. This exposure is often combined with EDTA (0.1%) and preferentially applied between delipidation and principal decellularization steps.

Tissue exposure to trypsin needs to be controlled; its ability to cleave a broad range of substrates means that deleterious effects on ECM structure can be exerted under overly long or too concentrated trypsin treatments. In porcine skin, different kinds of “fracturations” and density loss were observed in the ECM of the dermis compartment after an 18-h 0.25% trypsin exposure [80], whereas a denser

and less porous ECM network was obtained for the same decellularization protocol excluding this preliminary step. Similar observations have been reported for human amnion with shorter exposure at the same concentration [76]. Surprisingly, none of the studied patents report strict trypsin incubation during the processing of human skin, human AM, and porcine SIS for acellular matrix production for skin repair [83–85]. Only one recent patent presenting a decellularization method for fish skin includes a trypsin step, but at a concentration lower than 0.001% [86].

Dispase, mentioned above as deleterious to human AM in isotonic solution, has surprisingly been described in two publications as a good complement for skin decellularization. On porcine skin, application of a soft dispase treatment (0.56 U/mL, 4 °C in isotonic solution), followed by SDS and soft trypsin exposures, resulted in good cell removal without any adverse effect on ECM post- 4- and 12-h treatments. These results suggest a softening effect of both temperature and concentration decreases on dispase treatment [80]. On mouse skin, solubilization of 2 U/mL of dispase in Dulbecco's Modified Eagle Medium (DMEM, 37 °C, 2 h) prior to decellularization led to the production of an acellular matrix compatible with skin re-epithelialization in a mouse model [87]. The use of DMEM cell culture medium containing various components instead of simple isotonic solution could affect enzyme activity and explain the absence of deleterious effects. Furthermore, skin is an ECM-rich tissue containing more layers of epithelial cells than the amnion, so different accessibility to basement membrane between the two tissues also could explain the differential results with skin and human AM.

### 3.1.2. Chemicals

EDTA is a well-known chelator of di- and trivalent metal ions, commonly used in biochemistry to inhibit proteases. Its addition during exposures to trypsin or main decellularizing agents is a common tactic to prevent ECM degradation by cation-dependent proteases, which are abundant in both intra- and extracellular compartments. It also allows disruption of the protein–protein interactions based on metal cation bounding. Concentrations from 0.05% to 0.1% w/v are commonly applied, but levels up to 0.75% w/v have also been reported [80], without any specific association with tissue or decellularizing agent types. EDTA has been tested as a main decellularizing agent in human AM under various conditions (0.02–0.1%, 0.5–16 h, 4 or 37 °C), but does not completely decellularize the tissue, with a maximum DNA reduction limited to 50% of the initial level [88]. Other commercial protease inhibitors consisting of mixes of chemicals and small peptides can also be added to prevent ECM damage.

Hypo- and hypertonic shocks are applied from 1 to 12 h before the main decellularization steps or after to wash tissues and ensure the total lysis of remaining cells. To create hypertonic conditions, NaCl is commonly dissolved at 1 M to weaken cell membranes, as reported in the literature and in patented protocols to process fibroblast sheets, fish skin, and human skin [83,86,89]. Hypertonic NaCl also has been used to dissociate epidermis and dermis [90,91]. A hypertonic solution composed of 1 M KI has been used to process mouse skin [87]. Regarding hypotonic buffers, Tris (tris(hydroxymethyl)aminomethane) is often used at 10 mM, and immersion baths before decellularization and for solubilization of SDS have been reported for porcine bladder wall or fibroblast sheets processing [89,92].

### 3.1.3. Organic Solvents

Organic solvents are used to remove lipid content, an essential step for the preparation of acellular matrices using SIS and mesothelium, two fat-rich tissues.

A protocol combining methanol and chloroform (1/1, v/v, 12-h immersion) has been reported with porcine SIS, following manual elimination of the tunica serosa and tunica muscularis and prior intensive washings and SDS decellularization [81,82]. A degreasing step did not elicit notable changes in ECM architecture, and no remaining fat cells were detected in the final acellular matrix. The decrease in GFs reported in the final scaffold has been attributed to the manual removal of some compartments of the native tissue, not to organic solvent treatment.

Isopropanol (IPL) exposure is another method to remove lipids [76,93]. Two protocols have been compared with porcine skin: a single final exposure of the tissue to 100% IPL, and two 70%

IPL exposures distributed during the whole process, followed by a final immersion in 100% IPL [93]. The single exposure was associated with the removal of half of the total lipid content, whereas a decrease of 90% was estimated for repeated treatments. Regarding nucleic acid removal, no significant variation in remaining DNA levels was detected between the two methods, confirming that IPL is not a main decellularizing agent. A combination of IPL and acetone is also mentioned in the patent of the commercial Oasis<sup>®</sup> matrix derived from porcine SIS [84].

Use of acetone/hexane (40/60, v/v) and pure acetone with porcine mesothelium was reported [94]. Lipid amounts were not assessed, and only a visual evaluation suggested better elimination of fat using the acetone/hexane combination.

### 3.2. Main Decellularizing Agents

#### 3.2.1. Detergents

Detergents are the most common decellularizing agents used to process skin, SIS, mesothelium, bladder, and AM, regardless of the tissue origin. SDS is the strongest detergent used for tissue decellularization. It is ionic and disrupts all interactions between biomacromolecules. This agent is the most reported in decellularization protocols, and numerous teams have used it to process porcine SIS, porcine bladder wall, human amnion, human skin, rat flap, or fibroblast cell sheets [68,80,81,89,92,95]. Patents often mention its use for the production of commercial acellular matrices used in clinic [83,84,86]. SDS treatments are commonly performed by immersion in a 0.5% w/v solution (range of concentrations reported in literature: 0.1–1%) and last for hours. Thick and complex tissues containing a dense conjunctive layer or a multilayered epithelium, such as mesothelium, skin, and bladder wall are processed from 12 to 24 h. In contrast, thinner or less complex structure treatments last from 0.5 h for fibroblast sheets to 4–12 h for SIS. All decellularization studies that include SDS in concentrations >0.1% in their protocols, have reported a total removal of cells and cellular debris from the native tissue and low levels of remaining DNA. Some combinations with nucleases have been reported on amnion and bladder wall for SDS concentrations lower than 0.1% [68,92].

Although SDS is an excellent agent for decellularization, it has been frequently reported to damage ECM. Loss of global ECM density as well as GAG content and damage to the collagen network have particularly been observed in the matrix produced by fibroblast sheets [89] or in complex tissues such as human skin [96] and porcine SIS [81]. The increase in SDS concentration from 0.05% to 0.5% on fibroblast sheets or in exposure time from 6 to 12 h on SIS has been directly correlated with the amplification of ECM alterations. Scanning electron microscopy studies also suggest rearrangements of ECM fibers and variations in scaffold porosity after SDS treatment, even if no common and systematic effects can be generalized. The exposure of porcine SIS to the combination of 1% SDS for 12 h and 1% Triton X-100 for 0.5 h leads to complete cell removal but exerts similar adverse effects on ECM quality, coupled with a low cell development and metabolic activity after hemangioma stem cell (HEMSC) seeding [75]. Different SDS/Triton X-100 combinations have also been tested on rat fasciocutaneous flaps, leading to high DNA content removal after perfusion of 1% SDS for 24 or 72 h, followed by a 0.5-h 1% Triton X-100 treatment [95]. Modifications of the 2D structure of the ECM have been noted for the 72-h SDS perfusion using histological staining, but no other investigation has characterized the effects on ECM in detail. The switch of the 72-h SDS treatment from perfusion to immersion was associated with a decrease in DNA removal.

Basement membrane preservation after exposure to SDS is essential for epithelialization of the scaffold. The preservation of collagen IV, collagen VI, and laminin has been confirmed by immunostaining in porcine bladder wall (Tris HCl 10 mM overnight, 0.1% SDS for 24 h) [92] and softly decellularized human skin (1 M NaCl 24 h, 0.5% SDS 1 h) [97], but not in porcine SIS (0.05% trypsin 12 h, 0.5% SDS 4 h). Trypsin was shown to be compatible with basement membrane preservation at a higher concentration [81]. In human AM softly decellularized by a 24-h immersion in 0.03% SDS coupled with DNase I and RNase I treatments, maintenance of the basement membrane components but damage

to the mesenchymal matrix architecture have been observed [68]. However, the mechanical properties were not significantly altered, and the final scaffold exhibited characteristics, e.g., failure strain, failure stress, an elastin phase slope, and a collagen phase slope similar to native tissue. This study is the only one to report decellularization of human AM using SDS, and the soft treatment applied could be an interesting approach to preserving ECM mechanics and integrity. Finally, no other comparison of tissue biomechanics before and after decellularization has been reported.

Even if ECM damage is often underlined after SDS treatments, biocompatibility of the resulting scaffolds has been demonstrated for several native tissues. *In vitro*, matrices derived from porcine skin [80], porcine SIS [82], human AM [68], and fish skin [98] do not exert cytotoxic or proliferation-limiting effects. Cytotoxic tests performed by direct or conditioned-medium contact, metabolic arrays, or visual observations by light microscopy several days after cell seeding have confirmed the compatibility of the acellular matrices with cell development. These studies used NIH 3T3 fibroblasts, primary fibroblasts, bone marrow mesenchymal stem cells (BMSCs), or human lung carcinoma cells. Implantation tests have also been reported for some acellular matrices obtained using SDS. Porcine SIS-derived matrix implanted subcutaneously in rat has shown a very satisfying colonization and establishment of host fibroblasts 4 weeks post-operatively. A classic immune response involving giant cells and infiltrated inflammatory cells has been described but started to decrease after 14 days of implantation [81].

Triton X-100 is a non-ionic detergent that can disrupt interactions between biomacromolecules more gently than SDS. It is very commonly used alone or with other agents to generate matrices for skin repair derived from porcine SIS [82,99], porcine skin [68], porcine mesothelium [94], mouse skin [87], and rat flap [72]. Almost all protocols apply this detergent at the same concentration of 0.5% v/v, with a few using up to 1%. The ability of Triton X-100 to fully decellularize a native tissue has been evaluated with porcine SIS exposure to increasing concentrations of detergent from 0.5% to 2% [99]. Acellular scaffolds were obtained after 48 h of immersion in all conditions, but ECM damage, especially loss of density, was observed for concentrations higher than 1%. On mouse skin, successful decellularization using 0.5% Triton X-100 for 24 h also has been reported, associated with matrix alterations, especially loss of elastin and increased stiffness compared to native skin [87]. Limited colonization by primary fibroblasts and poor metabolic activity was also described. Surprisingly, during wound healing experiments in mouse, new blood vessel formation was observed in the grafted matrix, whereas a low dermis fibroblast infiltration occurred in accordance with *in vitro* experiments. Deleterious effects of 0.5% Triton X-100 on mouse skin may be the result of its limited thickness and cell abundance compared to SIS.

To modulate the effects on ECM and favor cell development, combinations of Triton X-100 with other agents have been tested at reduced exposure times. As mentioned above, the combination of 0.1% SDS for 12 h with 1% Triton X-100 for 0.5 h failed to produce a satisfying porcine SIS acellular matrix [75] because the scaffold obtained was not compatible with cell growth *in vitro*. In contrast, co-exposure of porcine skin to 0.5% Triton X-100 and hypertonic 1 M NaCl, followed by 2% sodium deoxycholate and DNase I treatments, appeared more promising [74]. Density and integrity of dermis ECM were preserved, especially collagen I and fibronectin networks, and the preservation of basement membrane was confirmed by the detection of collagen VII and laminin. Other observations revealed that, after a 6-month abdominal implantation in monkey, new blood vessels and dense colonization by host cells could be observed (global immune response not estimated). In rat, a dermis/adipose flap was successfully decellularized by perfusion using a combination of 1 M NaCl (epidermis removal and tonic shock), trypsin, and 1% Triton X-100 for 48 h [72]. DNA in dermis and fat compartments was quantified under 50 ng/mg of dry flap (10-fold reduction compared to native tissue). The general structure (2D and 3D) of the ECM and GAG content were preserved, but no evaluation of the basement membrane was described. Vascular endothelial GF was easily detectable by immunohistochemistry in fat tissue, but very low staining appeared in dermis. The growth of human umbilical vein endothelial cells and adipose stromal cells was shown between days 1 and 7 post-seeding.

Sodium deoxycholate (SD) is another ionic detergent applied as a decellularizing agent. It is commonly used at concentrations ranging from 2% to 4% and disrupts interactions between biomacromolecules. Its action, however, is frequently associated with DNA precipitation in spite of the loss of DNA–histone interactions, as has been confirmed [100]. The incubation of tissues with nucleases is thus automatically applied after exposure to SD, lasting for a few hours at concentrations from 5 to 10 kU/mL.

The capacity of SD to decellularize tissue alone has been reported in a study dedicated to the whole decellularization of rat intestine [101]. All cells and cell debris were removed with perfusion of a 4% SD solution for 4 h (connection to the superior mesenteric artery), followed by a DNase I treatment for 3 h, and ECM density was preserved. In contrast, an incomplete decellularization of porcine SIS was observed after a 12 h immersion in 4% SD and a DNase I treatment. Some cell nuclei were still visible in the conjunctive layer, associated with a level of remaining DNA that was 40% higher than that detected in tissue treated with SDS (40–45 versus 30 ng/mg dry tissue, respectively) [82]. However, other tissues that are compatible with perfusion and rich in ECM, such as aorta vein or skeletal muscle [102,103], can be totally decellularized using similar SD treatments.

A possible explanation for the incomplete cell removal in porcine SIS could be the limited SD penetration into the tissue during immersion compared to perfusion. As perfusion cannot be applied to all tissues, especially to skin that lacks an accessible vascular tree, immersion strategies have been developed by combining SD with other agents to aid in cell removal. In porcine skin where the multilayered epithelium is strongly anchored to the basement membrane, an exposure of the tissue to 1 M NaCl and 0.5% Triton X-100 for 24 h prior to a 4%, 24-h SD and DNase I treatments, successfully removed cells from both the epidermis and dermis compartments [74]. In rat, a fasciocutaneous flap decellularization protocol using 4% SD and DNase I was also reported, but further characterization of the acellular matrix is needed to characterize its properties [90].

SD-based decellularization approaches have been reported to be excellent for ECM preservation when used on tissues such as aorta, muscle, and lungs. This probably explains the effort invested in the development of protocols dedicated to skin-repair acellular matrix production despite problems with the perfusion-to-immersion switch. In porcine SIS, comparison of SD treatment to a double SDS/Triton X-100 exposure indicated a better quality of the final ECM matrix, even if some cells were still detected in the tissue. The 3D structure of the matrix was very close to that of the native tissue architecture, whereas other detergents increased porosity and modified fiber organization. Mechanical properties (failure strain, elastic modulus) were also similar to control tissue, and metabolic activity of HEMSCs at 96 h post-seeding was 10 times that of the activity of cells grown on the SDS/Triton X-100 matrix [75]. As previously described, a combination with NaCl and Triton X-100 on porcine skin is associated with basement membrane and ECM density preservation and a very satisfying biocompatibility *in vivo* [74].

Sodium lauroylsarcosinate (NLS, ionic) is a detergent that is less commonly reported for skin-repair matrix production. In mouse skin, a 1% NLS treatment for 24 h leads to complete removal of cells [87]. In a mouse wound healing model, despite important loss of GAG content, alterations in elastic fibers, and important tissue stiffening compared to native skin, this acellular matrix supported the establishment of host fibroblasts and formation of new blood vessels after a 6-month graft. A patented protocol, reported to be one of the MatrACELL<sup>®</sup> decellularization methods applied to produce the commercial human acellular matrix DermACELL<sup>™</sup>, also mentions use of NLS combined with nucleases to decellularize tissues [104]. These authors have published a comparison of this substitute with other commercial matrices, finding the lowest DNA content in the DermACELL<sup>™</sup> scaffold; however, other independent studies are needed to confirm and complete the characterization of this matrix. More studies also are needed to describe in greater detail the effects of NLS on ECM during decellularization processing and to evaluate its compatibility with acellular matrix production using other tissues.

No use of CHAPS, a zwitterionic detergent reported for the decellularization of thin tissues (lungs and trachea), has been tested for producing scaffolds for skin repair. The thickness of the tissues used for skin repair may not be compatible with adequate decellularization using this detergent [28].

### 3.2.2. Acids and Bases

Despite promising results, tissue exposures to acids or bases are not as commonly used as detergents to decellularize tissues. Sodium hydroxide (NaOH) treatments have been reported to process porcine SIS and human skin. High pH can modify the solubility of proteins and disrupt their interactions with other proteins and biomacromolecules in general, leading to destabilization of the cell membrane. Exposure of human skin to 0.06 M NaOH for several weeks was assayed to remove cells, followed by a 0.1 M HCl neutralization for 10 min. After 6 weeks, a cell-free tissue was produced with a very dense ECM, conserving its elastin and collagen networks [105]. In a porcine wound healing model, this type of acellular matrix has driven a complete epithelialization of the wound at 14 days post-grafting, with a limited immune response. This protocol was patented by a non-profit organization and is now available as a commercial product under the name Glyaderm®. A study describing patient care and follow-up a few years after Glyaderm® application was published in 2015, indicating very satisfying results with wounds such as burns or forearm flap loss [106]. Complete wound closure, pigmentation of the graft, and elasticity recovery were observed. NaOH treatment is also mentioned in the patent used for the production of the commercial Oasis® matrix derived from porcine SIS [84].

NaOH combined with hypertonic treatment and low-pH exposure can also be used on porcine SIS [82]. Exposure to 0.01 M NaOH, followed by 1 M HCl and 1 M NaCl treatments, has been associated with a complete decellularization (remaining DNA amount <30 ng/mg of dried ECM) and a satisfying preservation of the ECM architecture. The development of BMSCs has also been validated for 7 days post-seeding. Additional tests are needed to evaluate the effect of NaOH on a broader diversity of tissues.

Peracetic acid exposure, commonly used for tissue sterilization, has been used rarely in processing porcine SIS, with contradictory observations. A first study tested two 0.1% v/v treatments in a perfusion chamber (12 h) or by immersion (2 h) [75]. In both cases, levels of remaining DNA similar to native tissue attested to poor matrix decellularization. The longest exposure modified the biomechanics of the SIS, with increased stiffness and reduced resistance to strain. However, growth of HESMCs was possible in vitro. In contrast, a 2-h immersion using the same peracetic acid concentration combined with 4% ethanol was associated with a low remaining DNA level below 40 ng/mg of dried ECM [82]. A clear alignment of matrix fibers and increased scaffold porosity compared to native tissue were also found. Proliferation of BMSCs was also demonstrated in vitro.

Biocompatibility of these kinds of scaffolds for in vivo uses needs to be assessed. More studies are needed before any conclusions can be drawn about the relevance of peracetic acid in the production of acellular matrices for skin repair.

### 3.2.3. Physical Approaches

‘Supercritical conditions’ refers to combinations of particularly high temperature and high pressure so that a fluid reaches a physical state combining the properties of both liquid and gas. This physical state allows a good penetration of the fluid inside tissues, just like a gas, with a macromolecule solubilization capacity similar to that of a liquid. Supercritical CO<sub>2</sub> has been used initially to gently sterilize natural materials and tissues [107]. In recent years, supercritical CO<sub>2</sub> decellularization was developed on vascular tissues [108], and most recently, processing of human skin and porcine skin has been reported.

On human skin, processing with supercritical CO<sub>2</sub> after a hypertonic 1 M NaCl exposure has produced results similar to those with SDS in terms of cellular content removal (DNA level post-treatment <30 ng/mg of dried ECM) [91]. Small fractures reducing global ECM density were detected, but the fibronectin network seems to be unaffected. A surprising 3D structure was observed on scanning electron microscopy, however, characterized by punctate globular structures in the ECM network. Proliferation and viability of primary bovine chondrocytes were documented, but fibroblasts growth assessment was not presented. That study additionally underlined the importance

of tissue saturation with the CO<sub>2</sub>-philic detergent LS-54 prior to CO<sub>2</sub> exposure to enhance removal of cellular components.

A recent paper presented an approach combining supercritical CO<sub>2</sub> with NaOH (from 0.1 to 1 M, exact concentration unknown) applied on porcine skin [109]. Observations in the ECM network were similar to those previously reported with human skin, and promising results were obtained in a porcine wound healing model with the matrices seeded with autologous cells before grafting. This protocol is patented and used to produce ABCcolla<sup>®</sup> Collagen Matrix (ACRO Biomedical, Kaohsiung, Taiwan).

Freeze-thawing is a special case that is difficult to categorize as a main or complementary agent because the complete elimination of cells using only extreme variations of temperature has not been fully demonstrated. With cell sheets, no complete decellularization has been achieved applying three cycles of freeze-thawing [89]. However, with human mesothelium, a combination of three freeze-thawing cycles with trypsin, DNase I, and lipase VI exposures and isopropanol treatment to remove lipids has been reported as effective in removing all cells from the tissue. As trypsin and IPL cannot fully decellularize a tissue, the effect of freeze-thawing may be determinant in eliminating cells content [76]. Supplementary studies of the tissue after each step of the protocol are necessary to characterize the effects of each agent. Regardless, this study suggests that the combination of several complementary agents could be an interesting approach to investigate [76].

Regarding hydrostatic pressure, no protocol has been reported that has used it to decellularize tissues to produce acellular matrices for skin repair.

### 3.3. Promising New Approaches

Investigations are under way to find new main and complementary agents for tissue decellularization. A main issue when using acellular matrices for grafts in clinic is the immune response elicited by the exogenous tissue, sometimes leading to graft rejection. Galactose- $\alpha$ -(1, 3)-galactose ( $\alpha$ -gal) is an oligosaccharide sequence in some sugar-linked proteins and lipids of non-primate cell membranes. This oligosaccharide is a strong immunogenic epitope in humans and is especially responsible for the rejection of pig-derived organs. A method has been specifically developed to eliminate this antigen from porcine skin [74]. A classic decellularization protocol combining hypotonic NaCl, Triton X-100, SD, and DNase (already described above) has been completed by the addition of an  $\alpha$ -galactosidase incubation step, with no observed modification in the ECM structure compared to galactosidase-free treatment. After abdominal implantation in a monkey model, anti- $\alpha$ -gal antibody levels were similar for matrices derived from galactosidase-treated and KO gal +/- pig tissues, whereas amounts for non-galactosidase-treated matrix were 16 and 25 times higher at 20 and 70 days post-surgery, respectively. Despite such promising results, no other similar studies have been published.

In terms of main decellularizing agents, latrunculin B, a sea sponge toxin that inhibits actin polymerization [110], has recently been used to decellularize mouse skin [87]. Combined with 0.6 M NaCl, 1 M KI, dispase II, and benzonase exposures, latrunculin B treatment (50 nM in DMEM, 2 h) led to a complete cell removal, with a satisfying conservation of ECM density and quality (GAG and hydroxyprolin levels, elastin distribution similar to native tissue). No changes in the biomechanics of the tissue were detected, in contrast to results with NLS or Triton X-100 treatments tested in parallel. In vivo, in a mouse wound healing model, infiltration of numerous host fibroblasts, formation of new blood vessels, and total re-epithelialization of the wound were observed.

Many protocols have been published that propose techniques adapted for tissue decellularization (Table 2). Some of these methods are already patented to produce skin-repair matrices used in clinic. Although classic decellularizing agents are efficient at cell removal, adverse effects on the ECM are frequently reported. The main difficulty in the development of decellularization methods is thus to find the fine balance between sufficient removal of cell content and sufficiently limited ECM modification to maintain its properties driving the infiltration and the development of host cells.

**Table 2.** Summary of the different main decellularizing agents and their characteristics.

Method	Physical					Acids and Bases				Detergents			
	Freeze Thawing	Supercritical CO <sub>2</sub>	Peracetic Acid	NaOH	NLS	SD	Triton X-100	SDS	SDS				
<b>Common application</b>	(no consensus)	(no consensus)	0.1%	<0.1%	-	2-4%	0.5%	0.5%	0.5%				
<b>Cell content removal</b>	To be determined	To be determined	To be determined	Hours to weeks	(no consensus)	4-12 h	0.5-24 h	1-24 h	1-24 h				
<b>Main effects on ECM</b>	To be determined	Preservation of fibronectin, network and 3D structure	Modifications of 3D-structure, increase in stiffness	Preservation of elastin, collagens and 3D-structure	Increase in stiffness, alterations in 3D-structure	Preservation of basement membrane and mechanical properties	Preservation of basement membrane, increase in stiffness, loss of ECM density if >1%	Loss of ECM density, alterations in 3D-structure (porosity, fibers alignments)	Loss of ECM density, alterations in 3D-structure (porosity, fibers alignments)				
<b>Common combinations</b>	Trypsin, EDTA	Tonic shocks, CO <sub>2</sub> -philic agents	-	HCl (neutralization)	-	Nucleases mandatory, Triton X-100	SDS, SD	Tonic shocks, Triton, Trypsin, EDTA	Tonic shocks, Triton, Trypsin, EDTA				
<b>In vivo evaluation</b>	-	Wound healing model (pig)	-	Wound healing (pig), follow-up of patients	Wound healing (mouse), follow-up of patients	Wound healing (mouse), deep implantation (monkey)	Wound healing (mouse), deep implantation (monkey)	Wound healing (mouse), subcutaneous implantation (rat)	Wound healing (mouse), subcutaneous implantation (rat)				
<b>Mentioned in patents (commercial product)</b>	-	(ABCcolla®)	-	(Glyderm®)	X	X	X	X	X				
<b>Reported in literature</b>	+	+	+	+	+	++	+++	+++	+++				
				(DermACELL®)				(AlloDermI®)					

#### 4. Sterilization and Storage of Decellularized Scaffolds

Quite a variety of methods have been published for decellularizing tissues, but sterilization procedures are standard. They consist of short immersions in peracetic acid or ethanol solutions [76,111], gamma [112] or ultraviolet irradiation [81,101], and ethylene-oxide exposures [82]. Disinfectants such as antibiotic/antimycotic substances [80,88] or bacteriostatic agents such as sodium azide [87,97] are also common.

Regarding storage of acellular matrices, three main methods are commonly applied: immersion in phosphate-buffered saline containing antibiotics or bacteriostatics, lyophilization, and freezing [74,82,94]. The two latter methods require specific protocols including exposure of the tissue to preserving agents to protect macromolecules from the extreme pressures and temperatures applied.

Reviews presenting the different processes of sterilization and storage methods used for natural materials and tissues have been published elsewhere [113,114].

#### 5. Conclusions

Great advances have been made in skin regenerative medicine with regard to the preparation and use of decellularized tissues, leading to a growing number of products used in clinics with great success. Improvements are now expected to come at the level of protocol optimization. Ideal methods will reproducibly generate ECM that is free of cellular and nucleic acid material while retaining the structural, biochemical, and biomechanical properties crucial to its inherent cell-colonization function. Techniques may be optimized to achieve this balance, specifically in eliminating components of the tissue that could trigger an immunogenic response while maintaining components that will support and regulate reconstruction of new tissue. In particular, methods that combine strategies with chemicals, enzymes, and mechanical techniques may improve decellularization efficiency and limit the negative effects of protocols based on single agents. The goal is ultimately to administer the minimum amount of harsh chemicals or mechanical treatments to avoid unnecessary damage to the ECM microstructure and ultrastructure. Recellularization of skin devices before grafting is of great interest and should allow for the production of ready-to-use tissues. As with other organs, however, methods must be optimized to distribute the proper cell types uniformly and quickly through the tissue and enable sufficient nutrient and oxygen delivery for best cell viability. To that end, biofunctionalization of the decellularized tissue might offer promising perspectives. Interface biofunctionalization strategies should aim to modulate and control the implant–host interaction to enhance biocompatibility of implant materials and potentially even regulate implant function. The surface functionalization with ECM- or GF-derived molecules is a common approach in synthetic materials science to improve host cell–scaffold interactions at their interface and enhance biointegration of implant materials. Some groups have started to functionalize decellularized tissues by combining them with synthetic peptides to promote functions compromised by the decellularization process. This area will be intensified in the coming years with work that should lead to improved efficiency of decellularized tissues used for skin regeneration.

**Author Contributions:** A.M. and P.R. undertook chapters on skin defects and tissue sources for decellularization and clinical uses, Table 1 and both figures. M.D. managed the part on decellularization agents and Table 2. P.R. supervised the review construction and process. All authors have read and agreed to the published version of the manuscript.

**Funding:** Work in PR laboratory was supported by the Agence Nationale de la Recherche (grant number ANR-13-RPIB-0003-01). MD (PhD student) and AM (post-doctoral researcher) are funded by Agence Nationale de la Recherche (grant number ANR-18-CE92-0035-01) and EU/Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation (project T1EDK-03798), respectively.

**Conflicts of Interest:** The authors declare that there are no conflict of interest.

## References

1. Blanpain, C.; Fuchs, E. Epidermal homeostasis: A balancing act of stem cells in the skin. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 207–217. [[CrossRef](#)]
2. Koster, M.I.; Roop, D.R. Mechanisms Regulating Epithelial Stratification. *Annu. Rev. Cell Dev. Biol.* **2007**, *23*, 93–113. [[CrossRef](#)]
3. Rousselle, P.; Braye, F.; Dayan, G. Re-epithelialization of adult skin wounds: Cellular mechanisms and therapeutic strategies. *Adv. Drug Deliv. Rev.* **2019**, *146*, 344–365. [[CrossRef](#)] [[PubMed](#)]
4. Driskell, R.; Lichtenberger, B.M.; Hoste, E.; Kretzschmar, K.; Simons, B.; Charalambous, M.; Ferrón, S.R.; Hérault, Y.; Pavlovic, G.; Ferguson-Smith, A.C.; et al. Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature* **2013**, *504*, 277–281. [[CrossRef](#)] [[PubMed](#)]
5. Ghetti, M.; Topouzi, H.; Theocharidis, G.; Papa, V.; Williams, G.; Bondioli, E.; Cenacchi, G.; Connelly, J.; Higgins, C.A. Subpopulations of dermal skin fibroblasts secrete distinct extracellular matrix: Implications for using skin substitutes in the clinic. *Br. J. Dermatol.* **2018**, *179*, 381–393. [[CrossRef](#)] [[PubMed](#)]
6. Jiang, D.; Correa-Gallegos, D.; Christ, S.; Stefanska, A.; Liu, J.; Ramesh, P.; Rajendran, V.; De Santis, M.M.; Wagner, D.E.; Rinkevich, Y. Two succeeding fibroblastic lineages drive dermal development and the transition from regeneration to scarring. *Nature* **2018**, *20*, 422–431. [[CrossRef](#)]
7. Korosec, A.; Frech, S.; Gesslbauer, B.; Vierhapper, M.; Radtke, C.; Petzelbauer, P.; Lichtenberger, B.M. Lineage Identity and Location within the Dermis Determine the Function of Papillary and Reticular Fibroblasts in Human Skin. *J. Investig. Dermatol.* **2019**, *139*, 342–351. [[CrossRef](#)]
8. Driskell, R.; Jahoda, C.A.B.; Chuong, C.-M.; Watt, F.M.; Horsley, V. Defining dermal adipose tissue. *Exp. Dermatol.* **2014**, *23*, 629–631. [[CrossRef](#)]
9. Clark, R.A.F. Wound repair. Lessons for tissue engineering. In *Principles of Tissue Engineering*; Lanza, R.P., Langer, R., Chick, W., Eds.; Academic Press: Cambridge, MA, USA, 1997; pp. 737–768.
10. Järbrink, K.; Ni, G.; Sönnergren, H.; Schmidtchen, A.; Pang, C.; Bajpai, R.; Car, J. Prevalence and incidence of chronic wounds and related complications: A protocol for a systematic review. *Syst. Rev.* **2016**, *5*, 152. [[CrossRef](#)]
11. Eaglstein, W.H.; Falanga, V. CHRONIC WOUNDS. *Surg. Clin. N. Am.* **1997**, *77*, 689–700. [[CrossRef](#)]
12. Harding, K.; Morris, H.L.; Patel, G.K. Science, medicine, and the future: Healing chronic wounds. *BMJ* **2002**, *324*, 160–163. [[CrossRef](#)] [[PubMed](#)]
13. Mekkes, J.; Loots, M.; Van Der Wal, A.; Bos, J. Causes, investigation and treatment of leg ulceration. *Br. J. Dermatol.* **2003**, *148*, 388–401. [[CrossRef](#)] [[PubMed](#)]
14. Kartus, J.; Movin, T.; Karlsson, J. Donor-site morbidity and anterior knee problems after anterior cruciate ligament reconstruction using autografts. *Arthrosc. J. Arthrosc. Relat. Surg.* **2001**, *17*, 971–980. [[CrossRef](#)] [[PubMed](#)]
15. Konofaos, P.; Halen, J.V. Nerve Repair by Means of Tubulization: Past, Present, Future. *J. Reconstr. Microsurg.* **2013**, *29*, 149–164. [[CrossRef](#)]
16. Berthiaume, F.; Maguire, T.; Yarmush, M.L. Tissue Engineering and Regenerative Medicine: History, Progress, and Challenges. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 403–430. [[CrossRef](#)]
17. Khan, F.; Tanaka, M. Designing Smart Biomaterials for Tissue Engineering. *Int. J. Mol. Sci.* **2017**, *19*, 17. [[CrossRef](#)]
18. Zhu, M.; Li, W.; Dong, X.; Yuan, X.; Midgley, A.C.; Chang, H.; Wang, Y.; Wang, H.; Wang, K.; Ma, P.X.; et al. In vivo engineered extracellular matrix scaffolds with instructive niches for oriented tissue regeneration. *Nat. Commun.* **2019**, *10*, 4620. [[CrossRef](#)]
19. Zhong, S.P.; Zhang, Y.; Lim, C.T. Tissue scaffolds for skin wound healing and dermal reconstruction. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnology* **2010**, *2*, 510–525. [[CrossRef](#)]
20. Bryksin, A.V.; Brown, A.C.; Baksh, M.M.; Finn, M.; Barker, T.H. Learning from nature—novel synthetic biology approaches for biomaterial design. *Acta Biomater.* **2014**, *10*, 1761–1769. [[CrossRef](#)]
21. Niklason, L.E. Understanding the Extracellular Matrix to Enhance Stem Cell-Based Tissue Regeneration. *Cell Stem. Cell* **2018**, *22*, 302–305. [[CrossRef](#)]
22. Daley, W.P.; Peters, S.B.; Larsen, M. Extracellular matrix dynamics in development and regenerative medicine. *J. Cell Sci.* **2008**, *121*, 255–264. [[CrossRef](#)] [[PubMed](#)]

23. Kim, Y.; Ko, H.; Kwon, I.K.; Shin, K. Extracellular Matrix Revisited: Roles in Tissue Engineering. *Int. Neurobiol. J.* **2016**, *20*, S23–S29. [CrossRef] [PubMed]
24. Rousselle, P.; Montmasson, M.; Garnier, C. Extracellular matrix contribution to skin wound re-epithelialization. *Matrix Biology.* **2019**, *75–76*, 12–26. [CrossRef] [PubMed]
25. Eweida, A.; Marei, M.K. Naturally Occurring Extracellular Matrix Scaffolds for Dermal Regeneration: Do They Really Need Cells? *BioMed Res. Int.* **2015**, *2015*, 1–9. [CrossRef] [PubMed]
26. Cui, H.; Chai, Y.; Yu, Y. Progress in developing decellularized bioscaffolds for enhancing skin construction. *J. Biomed. Mater. Res. Part A* **2019**, *107*, 1849–1859. [CrossRef] [PubMed]
27. Elmashhady, H.H.; Kraemer, B.A.; Patel, K.H.; Sell, S.A.; Garg, K. Decellularized extracellular matrices for tissue engineering applications. *Electrospinning* **2017**, *1*, 87–99. [CrossRef]
28. Crapo, P.M.; Gilbert, T.; Badylak, S.F. An overview of tissue and whole organ decellularization processes. *Biomaterials* **2011**, *32*, 3233–3243. [CrossRef]
29. Patel, M.; Li, J.C.L. Fish skin acellular dermal matrix: Potential in the treatment of chronic wounds. *Chronic Wound Care Manag. Res.* **2019**, *6*, 59–70. [CrossRef]
30. Nakamura, N.; Kimura, T.; Kishida, A. Overview of the Development, Applications, and Future Perspectives of Decellularized Tissues and Organs. *ACS Biomater. Sci. Eng.* **2016**, *3*, 1236–1244. [CrossRef]
31. Fosnot, J.; Kovach, S.J.; Serletti, J.M. Acellular Dermal Matrix: General Principles for the Plastic Surgeon. *Aesthetic Surg. J.* **2011**, *31*, 5S–12S. [CrossRef]
32. Scalise, A.; Torresetti, M.; Verdini, F.; Capecci, M.; Andrenelli, E.; Mengarelli, A.; Ceravolo, M.G.; Fioretti, S.; Di Benedetto, G. Acellular dermal matrix and heel reconstruction: A new prospective. *J. Appl. Biomater. Funct. Mater.* **2017**, *15*, 15. [CrossRef] [PubMed]
33. Yamamoto, T.; Iwase, H.; King, T.; Hara, H.; Cooper, D.K.C. Skin xenotransplantation: Historical review and clinical potential. *Burns* **2018**, *44*, 1738–1749. [CrossRef] [PubMed]
34. Mohebichamkhorami, F.; Alizadeh, A. Skin substitutes; an updated review of products from year 1980 to 2017. *J. Appl. Biotechnol. Rep.* **2017**, *4*, 615–623.
35. Snyder, D.L.; Sullivan, N.; Schoelles, K.M. *Skin substitutes for treating chronic wounds*; Technology Assessment Report; Agency for Healthcare Research and Quality (AHRQ): Rockville, MD, USA, 18 December 2012. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK248353/> (accessed on 9 April 2020).
36. Wainwright, D.J.; Bury, S.B. Acellular Dermal Matrix in the Management of the Burn Patient. *Aesthetic Surg. J.* **2011**, *31*, 13–23. [CrossRef]
37. Ellis, C.V.; Kulber, D.A. Acellular Dermal Matrices in Hand Reconstruction. *Plast. Reconstr. Surg.* **2012**, *130*, 256S–269S. [CrossRef]
38. Dasgupta, A.; Orgill, D.; Galiano, R.D.; Zelen, C.M.; Huang, Y.-C.; Chnari, E.; Li, W.W. A Novel Reticular Dermal Graft Leverages Architectural and Biological Properties to Support Wound Repair. *Plast. Reconstr. Surg. Glob. Open* **2016**, *4*, e1065. [CrossRef]
39. Zelen, C.M.; Orgill, D.P.; Serena, T.; Galiano, R.; Carter, M.J.; DiDomenico, L.; Keller, J.; Kaufman, J.; Li, W.W. A prospective, randomised, controlled, multicentre clinical trial examining healing rates, safety and cost to closure of an acellular reticular allogenic human dermis versus standard of care in the treatment of chronic diabetic foot ulcers. *Int. Wound J.* **2016**, *14*, 307–315. [CrossRef]
40. Chen, S.-G.; Tzeng, Y.-S.; Wang, C.-H. Treatment of severe burn with DermACELL<sup>®</sup>, an acellular dermal matrix. *Int. J. BurnsTrauma* **2012**, *2*, 105–109.
41. Cazzell, S.; Moyer, P.M.; Samsell, B.; Dorsch, K.; McLean, J.; Moore, M.A. A Prospective, Multicenter, Single-Arm Clinical Trial for Treatment of Complex Diabetic Foot Ulcers with Deep Exposure Using Acellular Dermal Matrix. *Adv. Ski. Wound Care* **2019**, *32*, 409–415. [CrossRef]
42. Wu, T.-H.; Bertasi, G. The use of Dermacell<sup>®</sup> in Fingertip Injury. *J. Clin. Case Rep. Images* **2019**, *1*, 14–22. [CrossRef]
43. Reyzelman, A.; Crews, R.T.; Moore, J.C.; Moore, L.; Mukker, J.S.; Offutt, S.; Tallis, A.; Turner, W.B.; Vayser, D.; Winters, C.; et al. Clinical effectiveness of an acellular dermal regenerative tissue matrix compared to standard wound management in healing diabetic foot ulcers: A prospective, randomised, multicentre study. *Int. Wound J.* **2009**, *6*, 196–208. [CrossRef] [PubMed]
44. Reyzelman, A.M. Human acellular dermal wound matrix for treatment of DFU: Literature review and analysis. *J. Wound Care* **2015**, *24*, 128–134. [CrossRef] [PubMed]

45. Yu, G.; Ye, L.; Tan, W.; Zhu, X.; Li, Y.; Jiang, D. A novel dermal matrix generated from burned skin as a promising substitute for deep-degree burns therapy. *Mol. Med. Rep.* **2016**, *13*, 2570–2582. [[CrossRef](#)] [[PubMed](#)]
46. Buxton, D.; Lee, S.C.; Wickline, S.A.; Ferrari, M. Recommendations of the National Heart, Lung, and Blood Institute Nanotechnology Working Group. *Circulation* **2003**, *108*, 2737–2742. [[CrossRef](#)] [[PubMed](#)]
47. Wong, M.L.; Griffiths, L.G. Immunogenicity in xenogeneic scaffold generation: Antigen removal vs. decellularization. *Acta Biomater.* **2014**, *10*, 1806–1816. [[CrossRef](#)]
48. Imahara, S.D.; Klein, M.B. Skin grafts. In *Biomaterials for Treating Skin Loss*; Orgill, D.P., Blanco, C., Eds.; Woodhead Publishing: Cambridge, UK, 2009; pp. 58–79. [[CrossRef](#)]
49. United Health Care®. *Porcine Skin and Gradient Pressure Dressings (NCD 270.5)*; Medicare Advantage Policy Guideline; United Health Care: Minnetonka, SM, USA, 2020.
50. Lullove, E. Acellular fetal bovine dermal matrix in the treatment of nonhealing wounds in patients with complex comorbidities. *J. Am. Podiatr. Med. Assoc.* **2012**, *102*, 233–239. [[CrossRef](#)]
51. Parcells, A.L.; Karcich, J.; Granick, M.S.; Marano, M.A. The Use of Fetal Bovine Dermal Scaffold (PriMatrix) in the Management of Full-Thickness Hand Burns. *Eplasty* **2014**, *14*, 14.
52. Cornwell, K.; Landsman, A.; James, K.S. Extracellular Matrix Biomaterials for Soft Tissue Repair. *Clin. Podiatr. Med. Surg.* **2009**, *26*, 507–523. [[CrossRef](#)]
53. Le Guellec, M.; Morvan-Dubois, G.; Sire, J.-Y. Skin development in bony fish with particular emphasis on collagen deposition in the dermis of the zebrafish (*Danio rerio*). *Int. J. Dev. Biol.* **2004**, *48*, 217–231. [[CrossRef](#)]
54. Sitje, T.S.; Grøndahl, E.C.; Sørensen, J.A. Clinical innovation: Fish-derived wound product for cutaneous wound. *Wounds Int.* **2018**, *9*, 44–50.
55. Lachaud, C.C.; Rodriguez-Campins, B.; Hmadcha, A.; Soria, B. Use of Mesothelial Cells and Biological Matrices for Tissue Engineering of Simple Epithelium Surrogates. *Front. Bioeng. Biotechnol.* **2015**, *3*, 3. [[CrossRef](#)] [[PubMed](#)]
56. Hoganson, D.; Owens, G.E.; O'Doherty, E.M.; Bowley, C.M.; Goldman, S.M.; Harilal, D.O.; Neville, C.; Kronengold, R.T.; Vacanti, J. Preserved extracellular matrix components and retained biological activity in decellularized porcine mesothelium. *Biomaterials* **2010**, *31*, 6934–6940. [[CrossRef](#)] [[PubMed](#)]
57. Urciuolo, F.; Casale, C.; Imperato, G.; Netti, P.A. Bioengineered Skin Substitutes: The Role of Extracellular Matrix and Vascularization in the Healing of Deep Wounds. *J. Clin. Med.* **2019**, *8*, 2083. [[CrossRef](#)] [[PubMed](#)]
58. Costa, A.; Naranjo, J.D.; Londono, R.; Badylak, S.F. Biologic Scaffolds. *Cold Spring Harb. Perspect. Med.* **2017**, *7*, a025676. [[CrossRef](#)] [[PubMed](#)]
59. Chang, J.; DeLillo, N.; Khan, M.; Nacinovich, M.R. Review of small intestine submucosa extracellular matrix technology in multiple difficult-to-treat wound types. *Wounds* **2013**, *25*, 113–120.
60. Shi, L.; Ramsay, S.; Ermis, R.; Carson, D. In vitro and in vivo studies on matrix metalloproteinases interacting with small intestine submucosa wound matrix. *Int. Wound J.* **2011**, *9*, 44–53. [[CrossRef](#)]
61. Michopoulou, A.; Rousselle, P. How do epidermal matrix metalloproteinases support re-epithelialization during skin healing? *Eur. J. Dermatol. EJD* **2015**, *25*, 33–42. [[CrossRef](#)]
62. Cheng, C.W.; Solorio, L.D.; Alsberg, E. Decellularized tissue and cell-derived extracellular matrices as scaffolds for orthopaedic tissue engineering. *Biotechnol. Adv.* **2014**, *32*, 462–484. [[CrossRef](#)]
63. Parcells, A.L.; Abernathie, B.; Datiashvili, R. The use of urinary bladder matrix in the treatment of complicated open wounds. *Wounds* **2014**, *26*, 189–196.
64. Alvarez, O.M.; Smith, T.; Gilbert, T.W.; Onumah, N.J.; Wendelken, M.E.; Parker, R.; Markowitz, L. Diabetic foot ulcers treated with porcine urinary bladder extracellular matrix and total contact cast: Interim analysis of a randomized, controlled trial. *Wounds* **2017**, *29*, 140–146. [[PubMed](#)]
65. Cronce, M.; Faulknor, R.A.; Pomerantseva, I.; Liu, X.; Goldman, S.M.; Ekwueme, E.C.; Mwizerwa, O.; Neville, C.M.; Sundback, C.A. In vivo response to decellularized mesothelium scaffolds. *J. Biomed. Mater. Res. Part B Appl. Biomater.* **2017**, *106*, 716–725. [[CrossRef](#)] [[PubMed](#)]
66. Parolini, O.; Soncini, M.; Evangelista, M.; Schmidt, D. Amniotic membrane and amniotic fluid-derived cells: Potential tools for regenerative medicine? *Regen. Med.* **2009**, *4*, 275–291. [[CrossRef](#)] [[PubMed](#)]
67. Koob, T.J.; Rennert, R.; Zabek, N.; Masee, M.; Lim, J.J.; Temenoff, J.S.; Li, W.W.; Gurtner, G. Biological properties of dehydrated human amnion/chorion composite graft: Implications for chronic wound healing. *Int. Wound J.* **2013**, *10*, 493–500. [[CrossRef](#)] [[PubMed](#)]

68. Wilshaw, S.-P.; Kearney, J.N.; Fisher, J.; Ingham, E. Production of an Acellular Amniotic Membrane Matrix for Use in Tissue Engineering. *Tissue Eng.* **2006**, *12*, 2117–2129. [CrossRef]
69. Farhadhosseiniabadi, B.; Farahani, M.; Tayebi, T.; Jafari, A.; Biniyazan, F.; Modaresifar, K.; Moravvej, H.; Bahrami, S.; Redl, H.; Tayebi, L.; et al. Amniotic membrane and its epithelial and mesenchymal stem cells as an appropriate source for skin tissue engineering and regenerative medicine. *Artif. Cells Nanomed. Biotechnol.* **2018**, *46*, 431–440. [CrossRef]
70. Kshersagar, J.; Kshirsagar, R.; Desai, S.; Bohara, R.; Joshi, M. Decellularized amnion scaffold with activated PRP: A new paradigm dressing material for burn wound healing. *Cell Tissue Bank.* **2018**, *19*, 423–436. [CrossRef]
71. Yang, L.; Shirakata, Y.; Tokumaru, S.; Xiuju, D.; Tohyama, M.; Hanakawa, Y.; Hirakawa, S.; Sayama, K.; Hashimoto, K. Living skin equivalents constructed using human amnions as a matrix. *J. Dermatol. Sci.* **2009**, *56*, 188–195. [CrossRef]
72. Zhang, Q.; Johnson, J.A.; Dunne, L.W.; Chen, Y.; Iyyanki, T.; Wu, Y.; Chang, E.I.; Branch-Brooks, C.D.; Robb, G.L.; Butler, C.E. Decellularized skin/adipose tissue flap matrix for engineering vascularized composite soft tissue flaps. *Acta Biomater.* **2016**, *35*, 166–184. [CrossRef]
73. Jank, B.J.; Goverman, J.; Guyette, J.P.; Charest, J.M.; Randolph, M.; Gaudette, G.R.; Gershlak, J.R.; Purschke, M.; Javorsky, E.; Nazarian, R.M.; et al. Creation of a Bioengineered Skin Flap Scaffold with a Perfusable Vascular Pedicle. *Tissue Eng. Part A* **2017**, *23*, 696–707. [CrossRef]
74. Xu, H.; Wan, H.; Zuo, W.; Sun, W.; Owens, R.T.; Harper, J.R.; Ayares, D.L.; McQuillan, D.J. A Porcine-Derived Acellular Dermal Scaffold That Supports Soft Tissue Regeneration: Removal of Terminal Galactose- $\alpha$ (1,3)-Galactose and Retention of Matrix Structure. *Tissue Eng. Part A* **2009**, *15*, 1807–1819. [CrossRef]
75. Syed, O.; Walters, N.J.; Day, R.M.; Kim, H.-W.; Knowles, J.C. Evaluation of decellularization protocols for production of tubular small intestine submucosa scaffolds for use in oesophageal tissue engineering. *Acta Biomater.* **2014**, *10*, 5043–5054. [CrossRef] [PubMed]
76. Porzionato, A.; Sfriso, M.M.; Macchi, V.; Rambaldo, A.; Lago, G.; Lancerotto, L.; Vindigni, V.; De Caro, R. Decellularized omentum as novel biologic scaffold for reconstructive surgery and regenerative medicine. *Eur. J. Histochem.* **2013**, *57*, 4. [CrossRef] [PubMed]
77. Stenn, K.S.; Link, R.; Moellmann, G.; Madri, J.; Kuklinska, E.; Stenn, R.L.K.S. Dispase, a Neutral Protease From *Bacillus Polymyxa*, Is a Powerful Fibronectinase and Type IV Collagenase. *J. Investig. Dermatol.* **1989**, *93*, 287–290. [CrossRef] [PubMed]
78. Lim, L.S.; Riau, A.; Poh, R.; Tan, D.; Beuerman, R.; Mehta, J. Effect of dispase denudation on amniotic membrane. *Mol. Vis.* **2009**, *15*, 1962–1970. [PubMed]
79. Hopkinson, A.; Shanmuganathan, V.A.; Gray, T.; Yeung, A.M.; Lowe, J.; James, D.K.; Dua, H.S. Optimization of Amniotic Membrane (AM) Denuding for Tissue Engineering. *Tissue Eng. Part C Methods* **2008**, *14*, 371–381. [CrossRef] [PubMed]
80. Chen, R.-N.; Ho, H.-O.; Tsai, Y.-T.; Sheu, M.-T. Process development of an acellular dermal matrix (ADM) for biomedical applications. *Biomaterials* **2004**, *25*, 2679–2686. [CrossRef] [PubMed]
81. Luo, J.-C.; Chen, W.; Chen, X.-H.; Qin, T.-W.; Huang, Y.-C.; Yang, Z.; Li, X.-Q.; Qian, Z.; Yang, Z.-M. A multi-step method for preparation of porcine small intestinal submucosa (SIS). *Biomaterials* **2011**, *32*, 706–713. [CrossRef]
82. Ji, Y.-H.; Zhou, J.; Sun, T.-F.; Tang, K.; Xiong, Z.; Ren, Z.; Yao, S.; Chen, K.; Yang, F.; Zhu, F.; et al. Diverse preparation methods for small intestinal submucosa (SIS): Decellularization, components, and structure. *J. Biomed. Mater. Res. Part A* **2018**, *107*, 689–697. [CrossRef]
83. Livesey, S.A.; Del Campo, A.A.; Nag, A.; Nichols, K.B.; Griffey, E.S.; Coleman, C. Acellular Dermal Matrix and Method of Use Thereof for Grafting. US Patent 8,067,149, 17 June 2009. Available online: <http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetacgi%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=8,067,149.PN.&OS=PN/8,067,149&RS=PN/8,067,149> (accessed on 29 March 2020).
84. Shah, B.; Harrigan, K.L. Isolated Extracellular Matrix Material Including Subserous Fascia. US Patent 9,044,455, 26 July 2013. Available online: <http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetacgi%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=0000084.PN.+9,044,455.PN.&OS=PN/0000084+OR+PN/9,044,455&RS=PN/0000084+OR+PN/9,044,455> (accessed on 29 March 2020).

85. Qin, X.; Chen, S.; Aschenbach, L.; Chen, J. LifeNet Health. Decellularized Placental Membrane and Methods of Preparing and Use Thereof. WO Patent 2017/112934, 23 December 2016. Available online: <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2017112934&tab=PCTBIBLIO> (accessed on 29 March 2020).
86. Sigurjonsson, G.F. Kerecis EHF. A Scaffold Material for Wound Care and/or Other Tissue Healing Applications. WO Patent 2013/144727, 29 March 2013. Available online: [https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2013144727&tab=PCTBIBLIO&\\_cid=P21-KA6TQY-23091-1](https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2013144727&tab=PCTBIBLIO&_cid=P21-KA6TQY-23091-1) (accessed on 29 March 2020).
87. Farrokhi, A.; Pakyari, M.; Nabai, L.; Pourghadiri, A.; Hartwell, R.; Jalili, R.B.; Ghahary, A. Evaluation of Detergent-Free and Detergent-Based Methods for Decellularization of Murine Skin. *Tissue Eng. Part A* **2018**, *24*, 955–967. [CrossRef]
88. Sanluis-Verdes, A.; Vilar, M.T.Y.-P.; García-Barreiro, J.J.; García-Camba, M.; Ibáñez, J.S.; Doménech, N.; Rendal-Vazquez, E. Production of an acellular matrix from amniotic membrane for the synthesis of a human skin equivalent. *Cell Tissue Bank.* **2015**, *16*, 411–423. [CrossRef]
89. Xing, Q.; Yates, K.; Tahtinen, M.; Shearier, E.; Qian, Z.; Zhao, F. Decellularization of Fibroblast Cell Sheets for Natural Extracellular Matrix Scaffold Preparation. *Tissue Eng. Part C Methods* **2014**, *21*, 77–87. [CrossRef] [PubMed]
90. Henderson, P.; Nagineni, V.V.; Harper, A.; Bavinck, N.; Sohn, A.M.; Krijgh, D.D.; Jimenez, N.; Weinstein, A.L.; Spector, J.A. Development of an Acellular Bioengineered Matrix with a Dominant Vascular Pedicle. *J. Surg. Res.* **2010**, *164*, 1–5. [CrossRef] [PubMed]
91. Antons, J.; Marascio, M.G.; Aeberhard, P.; Weissenberger, G.; Hirt-Burri, N.; Applegate, L.; Bourban, P.; Pioletti, D. Decellularised tissues obtained by a CO<sub>2</sub>-philic detergent and supercritical CO<sub>2</sub>. *Eur. Cells Mater.* **2018**, *36*, 81–95. [CrossRef] [PubMed]
92. Rosario, D.J.; Reilly, G.C.; Salah, E.A.; Glover, M.; Bullock, A.; Mac Neil, S. Decellularization and sterilization of porcine urinary bladder matrix for tissue engineering in the lower urinary tract. *Regen. Med.* **2008**, *3*, 145–156. [CrossRef]
93. Ventura, R.D.; Padalhin, A.R.; Park, C.M.; Lee, B.T. Enhanced decellularization technique of porcine dermal ECM for tissue engineering applications. *Mater. Sci. Eng. C* **2019**, *104*, 109841. [CrossRef]
94. Soffer-Tsur, N.; Shevach, M.; Shapira, A.; Peer, D.; Dvir, T. Optimizing the biofabrication process of omentum-based scaffolds for engineering autologous tissues. *Biofabrication* **2014**, *6*, 035023. [CrossRef]
95. Qu, J.; Van Hogezaand, R.M.; Zhao, C.; Kuo, B.J.; Carlsen, B.T. Decellularization of a Fasciocutaneous Flap for Use as a Perfusable Scaffold. *Ann. Plast. Surg.* **2015**, *75*, 112–116. [CrossRef]
96. Carruthers, C.A.; Dearth, C.L.; Reing, J.; Kramer, C.R.; Gagne, D.H.; Crapo, P.M.; Garcia, O.; Badhwar, A.; Scott, J.R.; Badylak, S.F. Histologic Characterization of Acellular Dermal Matrices in a Porcine Model of Tissue Expander Breast Reconstruction. *Tissue Eng. Part A* **2014**, *21*, 35–44. [CrossRef]
97. Walter, R.J.; Matsuda, T.; Reyes, H.M.; Walter, J.M.; Hanumadass, M. Characterization of acellular dermal matrices (ADMs) prepared by two different methods. *Burns* **1998**, *24*, 104–113. [CrossRef]
98. Magnússon, S.; Baldursson, B.T.; Kjartansson, H.; Rolfsson, O.; Sigurjonsson, G.F. Regenerative and Antibacterial Properties of Acellular Fish Skin Grafts and Human Amnion/Chorion Membrane: Implications for Tissue Preservation in Combat Casualty Care. *Mil. Med.* **2017**, *182*, 383–388. [CrossRef]
99. Villoldo, G.; Loresi, M.; Ielpi, M.; Cambiasso, P.; Barbich, M.; De Badiola, F.; Ruiz, E.; Giudice, C.; Damia, O.; Argibay, P. UP-2.184: Different Concentrations of Triton X-100 To Decellularize Small Intestinal Submucosa. *Urology* **2009**, *74*, S289–S290. [CrossRef]
100. Smart, J.E.; Bonner, J. Selective dissociation of histones from chromatin by sodium deoxycholate. *J. Mol. Biol.* **1971**, *58*, 651–659. [CrossRef]
101. Maghsoudlou, P.; Totonelli, G.; Loukogeorgakis, S.P.; Eaton, S.; De Coppi, P. A decellularization methodology for the production of a natural acellular intestinal matrix. *J. Vis. Exp.* **2013**, *80*, e50658. [CrossRef] [PubMed]
102. Bertanha, M.; Moroz, A.; Jaldin, R.G.; Silva, R.A.; Rinaldi, J.C.; Golim, M.A.; Felisbino, S.L.; Domingues, M.A.C.; Sobreira, M.L.; Reis, P.P.D.; et al. Morphofunctional characterization of decellularized vena cava as tissue engineering scaffolds. *Exp. Cell Res.* **2014**, *326*, 103–111. [CrossRef]
103. Illien-Junger, S.; Sedaghatpour, D.D.; Laudier, D.M.; Hecht, A.C.; Qureshi, S.A.; Iatridis, J.C. Development of a bovine decellularized extracellular matrix-biomaterial for nucleus pulposus regeneration. *J. Orthop. Res.* **2015**, *34*, 876–888. [CrossRef]



---

# 1

## **Potentiel de la spectrométrie de masse pour l'identification de biomarqueurs des pathologies cutanées**

Mélissa Dussoyer, Catherine Moali\* et Priscillia Lagoutte

Laboratoire de Biologie Tissulaire et Ingénierie Thérapeutique – CNRS UMR5305, Université de Lyon

\*Auteur pour la correspondance : catherine.moali@ibcp.fr

---

### **I – Introduction**

Le recours aux biomarqueurs pour révéler l'état physiopathologique de la peau ne s'impose pas d'emblée car les signes de dysfonctionnement de cet organe sont souvent visibles à l'œil nu et plus faciles à détecter que dans les organes internes. Cependant, l'état du derme et de l'hypoderme peut être difficile à caractériser par une inspection superficielle et, dans certains cas, il serait bénéfique de disposer d'indicateurs prédictifs sensibles de l'évolution d'une pathologie cutanée. Les biomarqueurs tissulaires ou circulants ont donc toute leur place en biologie cutanée pour mettre en évidence des processus annonciateurs d'une dérive plus grave ou rendre compte de l'état des couches inférieures de la peau. Ils peuvent aussi servir à sélectionner les patients qui pourraient répondre préférentiellement à une thérapie et à suivre l'efficacité d'un traitement.

En biologie de la peau comme dans d'autres domaines, les méthodes omiques ont déjà fait la preuve de leur potentiel pour révéler les acteurs moléculaires impliqués dans une pathologie. Après l'ère de la génomique, une transition rapide s'est opérée vers la transcriptomique puis vers la protéomique pour identifier de nouveaux gènes ou protéines particulièrement dérégulés dans un contexte spécifique. Les nouvelles générations de spectromètres de masse utilisés en protéomique permettent actuellement de couvrir les protéomes avec une très bonne profondeur qui donne une information enrichie par rapport à l'étude des gènes et des ARNs. Un seul gène peut en effet donner

lieu à de multiples protéoformes (variants d'épissage, modifications post-traductionnelles), possédant souvent des activités différentes et qu'il est important de mieux connaître pour en appréhender la pertinence physiopathologique. Par ailleurs, les études transcriptomiques ne rendent compte que de la néosynthèse des protéines mais non de leur stabilité ou *turnover* global alors qu'il est connu que la demi-vie de certaines protéines importantes de la peau comme l'élastine peut atteindre plusieurs dizaines d'années.

Dans le contexte de la recherche de biomarqueurs, l'étude des protéines en spectrométrie de masse (MS) requiert leur transformation en peptides de plus petite taille (idéalement entre 5 et 25 acide aminés) dont la séquence est plus facile à identifier à partir des données obtenues que celle des protéines entières. Pour cela, des étapes de réduction et alkylation des protéines, pour éliminer les ponts disulfures, et de digestion enzymatique par des protéases de spécificité connue (par exemple la trypsine ou l'endoprotéinase GluC) sont habituellement intégrées au traitement des échantillons. Par ailleurs, plus que la description qualitative du protéome d'un échantillon biologique, c'est l'estimation des quantités de protéines qui se révèle la plus utile en recherche comme en clinique. En spectrométrie de masse, deux types d'approches quantitatives peuvent être utilisées (Emadali and Gallagher-Gambarelli, 2009; Lindemann et al., 2017) :

- la quantification relative, utilisée pour des analyses globales en protéomique de découverte (*discovery proteomics* en anglais) car permettant de comparer simultanément l'abondance de nombreuses protéines dans différents échantillons sans *a priori* sur leur nature.
- la quantification absolue, qui vise à déterminer la quantité exacte d'un ou plusieurs biomarqueurs déjà identifié(s) ou en cours de validation, dans une unité exploitable (masse, moles etc). Cette approche est également souvent assimilée à la protéomique ciblée (*targeted proteomics* en anglais) dans la mesure où les cibles intéressantes ont été préalablement définies.

Les approches de quantification relative sont utilisées en recherche fondamentale comme appliquée, pour étudier des échantillons dans leur globalité. En effet, dans de nombreux cas, c'est davantage la variation d'abondance d'une protéine entre plusieurs conditions qui est informative (sain vs malade, non traité vs traitement 1 vs traitement 2, etc.) que la quantité exacte. De plus, l'approche relative est compatible avec la comparaison simultanée de l'abondance d'une multitude de protéines, permettant ainsi le

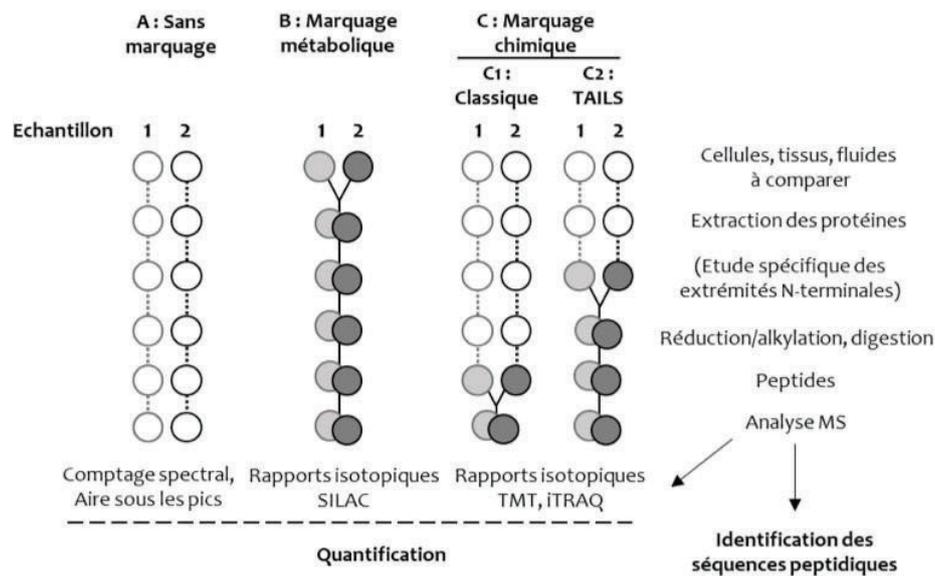
suivi de plusieurs centaines ou milliers de molécules, provenant d'un ou plusieurs échantillon(s), en une seule injection dans le spectromètre de masse. A partir de ces données, des études bioinformatiques peuvent également être menées afin de détecter des dérégulations non pas à l'échelle de la simple molécule mais de la voie métabolique, du processus biologique global ou encore des interactions entre protéines (Chen et al., 2020).

La quantification absolue requiert quant à elle l'utilisation de composés « étalons » (ou « standards ») qui permettent de dresser la correspondance entre l'intensité du signal détecté en MS et une quantité de protéine donnée (Lindemann et al., 2017). En effet, chaque protéine se comporte différemment lors de la préparation de l'échantillon et répond de façon spécifique aux phénomènes physico-chimiques se déroulant lors de l'injection dans l'appareil. Il est ainsi nécessaire d'utiliser un standard, interne ou externe, de quantité connue pour au moins un peptide issu de chaque protéine étudiée : celui-ci reflètera le comportement du peptide naturel lors de la préparation de l'échantillon et de l'injection dans le spectromètre de masse. La quantification absolue apparaît donc parfaitement indiquée pour une utilisation en clinique où seulement quelques molécules d'intérêt clairement identifiées doivent être suivies (détection de biomarqueurs, détection de toxines dans le sang etc). Elle ne constitue cependant pas une méthode adaptée à l'étude simultanée de centaines voire de milliers de protéines : la diversité des peptides générés serait beaucoup trop élevée pour assurer la présence d'un étalon pour tous.

## **II – La protéomique de découverte et la mise en évidence de nouveaux biomarqueurs**

La capacité à différencier les peptides (et donc les protéines) issus de différents échantillons est le pré-requis incontournable pour toute investigation relative en MS. Pour cela, plusieurs méthodes ont été développées. Les méthodes dites sans marquage (*label-free*) ne nécessitent pas de traitement particulier des échantillons : la capacité à différencier les échantillons est assurée par l'injection en MS de chacune des conditions de façon individuelle (Fig. 1A). Le nombre de spectres MS (méthode *spectral counting*) ou l'aire sous le pic de l'ion précurseur (méthode *area under the curve*), détecté(s) pour l'ensemble des peptides d'une protéine est comparé(e) d'une condition à l'autre, permettant de conclure à l'égalité ou à la variation de l'abondance d'un composé (Anand et al., 2017). Ces approches sont peu

coûteuses, faciles à mettre en œuvre et ne nécessitent pas de traitement complexe des échantillons avant injection MS. Cependant, elles sont très demandeuses en temps-machine, ce qui est à prendre en compte dans le coût final, et le traitement parallèle ou successif de chacun des échantillons d'intérêt peut induire des biais de quantification puisque tous les signaux n'ont pas été acquis de manière simultanée et donc strictement identique. Il est alors impératif de démontrer la reproductibilité des processus de préparation d'échantillons et d'acquisition des données d'un échantillon à l'autre, mais aussi de mettre en place une méthode de normalisation des données inter-échantillons appropriée.



**Figure 1 : Principales approches quantitatives en protéomique de découverte.** Les analyses de spectrométrie de masse donnent deux informations importantes : la séquence en acides aminés et la quantité relative des peptides qui permettent de remonter, respectivement, à l'identité et à la quantité relative des protéines dans les échantillons. Pour les méthodes avec marquage, celui-ci peut être effectué dès la préparation de l'échantillon (B), au niveau des protéines extraites de l'échantillon (C2) ou au niveau des peptides issus de la digestion enzymatique (C1). Le schéma représente le cas simple de deux échantillons à comparer mais ce nombre peut être étendu sans limites pour la méthode sans marquage (A) et jusqu'au nombre maximum de réactifs isotopiques disponibles pour les autres méthodes (B-C). Les méthodes de quantification sont différentes pour chacune des approches et requièrent des logiciels d'analyse dédiés. Adapté de (Emadali and Gallagher-Gambarelli, 2009).

Pour pallier à ce biais de quantification et augmenter la reproductibilité de l'analyse MS en réalisant l'injection simultanée de plusieurs échantillons, des techniques de protéomique avec marquage, ou *label-based*, ont été développées (Rauniyar and Yates, 2014). Elles consistent en l'intégration au contenu de chaque échantillon d'un marqueur spécifique qui permettra de le différencier du contenu des autres conditions lors de l'injection commune. Les protéines provenant des différentes conditions peuvent alors être mélangées et analysées en un seul pool. Plusieurs techniques de marquage peuvent être employées. La première consiste à marquer les protéines au sein même des cellules par intégration d'un ou plusieurs atomes lourds ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) aux acides aminés essentiels : on parle de « marquage métabolique » ou SILAC (*Stable Isotope Labelling by Amino acids in cell Culture* ; Fig. 1B). Pour chaque condition, un isotope ou une combinaison d'isotopes différent(e) est utilisé(e). Les signaux MS des peptides correspondants sont ainsi caractérisés par un gain de masse particulier propre à cette combinaison d'isotopes, donc à la condition dont ils proviennent. Il est alors possible, après mélange des peptides des différentes conditions, de quantifier l'abondance (en unités arbitraires) des peptides dans chacune d'entre elles et donc de comparer leurs teneurs entre les différents échantillons analysés. Cette approche initialement appliquée à des cellules en culture (apport des acides aminés marqués *via* le milieu) est aussi pratiquée *in vivo* aujourd'hui (apport par l'alimentation), le plus souvent pour des études courtes d'identification des protéines induites sous l'action d'un stimulus (Zanivan et al., 2012). Si le mélange des échantillons à une étape très précoce de leur préparation limite les sources de biais expérimentaux, de telles méthodes sont limitées par leur coût, un multiplexage de conditions limité et une durée de mise en œuvre souvent conséquente (temps nécessaire à l'intégration des acides aminés marqués à la totalité des protéines). Par ailleurs, elles ne sont pas applicables pour les études sur des sujets humains.

Afin d'augmenter la capacité de multiplexage ainsi que la diversité des échantillons analysables et de diminuer le temps d'expérimentation, les « marquages chimiques » (principalement à l'aide des réactifs commerciaux TMT™ et iTRAQ®), réalisés après digestion enzymatique (Fig. 1C1), ont été développés. Pour chaque échantillon, une étiquette (*tag*) unique de masse connue (marquée par une combinaison d'isotopes stables) est choisie et ajoutée de façon covalente à tous les peptides de l'échantillon. Dans le mélange analysé en MS, un peptide retrouvé dans plusieurs conditions est donc détecté sous plusieurs formes, chacune correspondant à la présence d'un

tag différent. La quantification relative est alors possible, de manière similaire au marquage métabolique. De telles approches peuvent cependant demander une optimisation de la réaction de marquage et nécessitent un traitement individuel des échantillons jusqu'à l'étape de marquage, pouvant conduire à une quantification légèrement moins fiable qu'avec le marquage métabolique.

Les méthodes de spectrométrie de masse quantitative relative ont à ce jour permis de comparer des protéomes très divers et de caractériser de multiples dérégulations pathologiques, en seulement quelques analyses. Récemment, le protéome de la peau humaine et ses variations en termes de qualité et de quantité (évaluée par la méthode sans marquage), d'un compartiment à l'autre (épiderme supérieur, épiderme inférieur, derme et hypoderme) mais aussi d'un type cellulaire à l'autre (kératinocytes, fibroblastes, cellules endothéliales, mélanocytes, mastocytes, macrophages, cellules dendritiques), ont été décrits (Dyring-Andersen et al., 2020). Cette étude décrit donc le premier « atlas » de la composition de la peau normale au niveau des protéines et met en évidence des biomarqueurs spécifiques de chaque compartiment ou type cellulaire. De nombreuses autres études protéomiques ont également contribué à mettre en évidence une modification de l'abondance de certains composants dans les pathologies cutanées et permis de révéler des voies de biosynthèse et de signalisation ainsi que différents processus cellulaires particulièrement altérés dans ces maladies. Des revues récentes font par exemple le point sur les analyses protéomiques réalisées dans le cadre de l'étude du psoriasis (Chularojanamontri et al., 2019), du mélanome (Donnelly et al., 2019) ou de l'épidermolyse bulleuse dystrophique récessive (Dengjel et al., 2020). Notons également que des optimisations ont été réalisées pour permettre l'étude protéomique quantitative des couches superficielles de l'épiderme (*stratum corneum*) par des méthodes peu invasives telles que les prélèvements sur rubans adhésifs (Kaleja et al., 2020).

Par ailleurs, de nouvelles approches ont été développées pour étudier spécifiquement certains sous-protéomes cutanés. En particulier, la peau étant un tissu riche en matrice extracellulaire, de nombreuses pathologies cutanées ont des répercussions sur la composition de ce compartiment tissulaire (cicatrices hypertrophiques, chéloïdes, sclérodémie, épidermolyse bulleuse etc). Cependant, les protéines matricielles sont porteuses de nombreuses modifications post-traductionnelles et largement insolubles, ce qui rend leur détection en MS particulièrement difficile. Pour remédier à ce problème, il peut être nécessaire d'utiliser des protocoles de préparation d'échantillons et

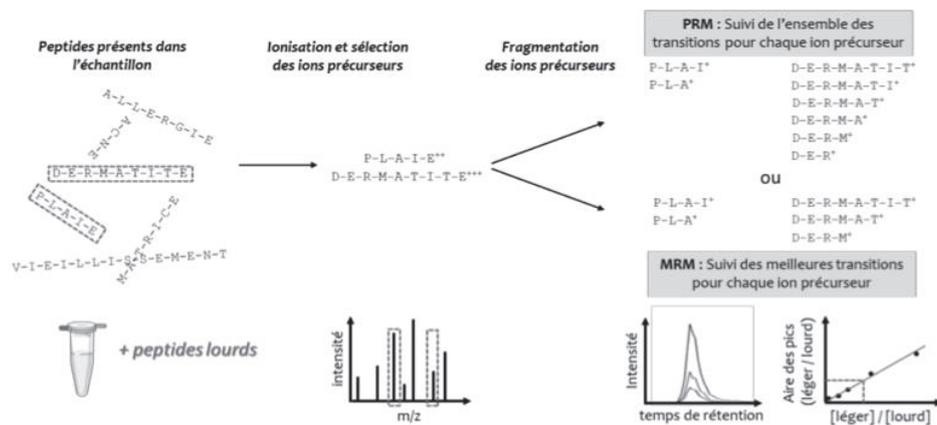
de post-traitement des données MS favorisant leur enrichissement et leur détection. Une étude récente a montré que les approches de décellularisation et de solubilisation progressive par des détergents permettent une meilleure couverture du « matrisome », particulièrement complexe de la peau, que les méthodes de lyse classiques (Dussoyer et al., 2022). Ces méthodes d'enrichissement ont par ailleurs fait leurs preuves pour l'étude de la matrice extracellulaire aberrante qui est déposée dans les cicatrices chéloïdes (Barallobre-Barreiro et al., 2019; Zhang et al., 2021).

Une autre approche intéressante, possible seulement en protéomique et pas avec les autres méthodes omiques, consiste à caractériser les altérations quantitatives des modifications post-traductionnelles observées sur les protéines suite à une perturbation de l'homéostasie. La disparition de certains peptides peut en effet révéler l'ajout anormal d'une modification sur ces peptides. Cette approche appelée *Peptide Location Fingerprint* a ainsi été appliquée avec succès à la comparaison de peaux âgées par le vieillissement naturel ou l'exposition aux UV et a permis l'identification de 120 biomarqueurs candidats dans le derme et 70 dans l'épiderme (Ozols et al., 2021). Par ailleurs, certaines modifications post-traductionnelles spécifiques se prêtent particulièrement bien à l'analyse protéomique. C'est le cas de la protéolyse qui est un reflet très pertinent des processus dynamiques en cours dans un tissu. Ainsi, les fragments générés par l'activité des protéases au sein d'un échantillon représentent des sources potentielles de biomarqueurs du remodelage tissulaire. Ces approches sont regroupées sous le nom de N-terminomique car elles sont centrées sur la détection des peptides N-terminaux des protéines et de leurs fragments protéolytiques. Une des techniques de N-terminomique appelée TAILS (*Terminal Amine Isotopic Labeling of Substrates* ; Fig. 1C2) a ainsi permis de caractériser de nouveaux processus de maturation des précurseurs des collagènes fibrillaires dans le derme (Leduc et al., 2021), de révéler une augmentation globale de la protéolyse en cas d'inflammation de la peau (auf dem Keller et al., 2013) ou encore d'identifier des processus de dégradation anormaux dans le contexte des plaies chroniques (Sabino et al., 2018).

Les résultats de ces criblages du protéome à grande échelle, permis par la quantification relative, peuvent ensuite être utilisés comme base pour la validation ou le suivi de biomarqueurs spécifiques en donnant des informations préliminaires sur les peptides susceptibles d'être détectés en protéomique ciblée.

### III – La protéomique ciblée et le suivi des biomarqueurs

Les approches quantitatives globales, comme nous venons de le voir, servent de phase exploratoire pour identifier des biomarqueurs. Pour valider ces biomarqueurs protéiques mais également les utiliser en clinique pour diagnostiquer ou suivre une pathologie, il est courant d'utiliser l'ELISA. Cependant, cette méthode est limitée car elle repose sur le suivi d'un biomarqueur unique et nécessite le développement d'anticorps sensibles, spécifiques et reconnaissant la forme native de la protéine. Grâce aux progrès technologiques instrumentaux réalisés ces dernières années, plusieurs approches, également basées sur l'utilisation de la spectrométrie de masse, ont été développées. Elles permettent de suivre et de quantifier plusieurs protéines simultanément avec une sensibilité équivalente à l'ELISA et une très bonne spécificité.



**Figure 2 : Principe des méthodes de protéomique ciblée PRM et MRM pour la quantification absolue.** Après sélection et fragmentation des ions précurseurs choisis pour la détection de la protéine d'intérêt, la quantification peut être réalisée par comparaison avec le signal des peptides synthétiques correspondants, marqués avec des isotopes et ajoutés en quantité connue dans l'échantillon de départ.

La protéomique ciblée couvre les approches de MRM (*Multiple Reaction Monitoring*), de PRM (*Parallel Reaction Monitoring*) mais également l'utilisation de données acquises en DIA (*Data Independent Acquisition*) pour quantifier des bibliothèques de peptide. Cette dernière approche ne sera pas détaillée ici car encore peu utilisée mais son intérêt pour l'identification de

biomarqueurs est décrit dans plusieurs revues (par exemple (Sajic et al., 2015)). Les approches ciblées reposent sur le suivi exclusif de peptides spécifiques appartenant à une ou plusieurs protéines (par opposition à la spectrométrie de masse décrite ci-dessus où une sélection est également réalisée mais sur la base de l'intensité des peptides et non sur la base de la protéine à laquelle ils appartiennent). Chaque protéine est détectée à partir d'un nombre limité de peptides uniques, issus du traitement de l'échantillon par des protéases telles que la trypsine (peptides protéotypiques). Ces peptides protéotypiques ou « ions précurseurs » sont ensuite fragmentés dans le spectromètre de masse et les signaux générés par ces fragments sont appelés « transitions ». Pour chaque peptide, plusieurs transitions sont suivies, et de même, pour chaque protéine, plusieurs peptides protéotypiques ou précurseurs peuvent être suivis (Fig. 2). Cette double sélection (précurseur et transitions) assure la spécificité de la méthode. La PRM et la MRM sont basées sur le même principe, toutefois la PRM utilise un spectromètre Quadripôle-Orbitrap haute résolution sans sélection préalable des transitions tandis que la MRM repose sur l'utilisation d'un spectromètre du type triple quadripôle avec sélection préalable des transitions. La MRM et la PRM, en plus d'être extrêmement sensibles et spécifiques, sont les seules méthodes de spectrométrie de masse qui permettent la mesure de la quantité absolue d'une protéine cible dans un échantillon, grâce à l'ajout d'un étalon interne (peptides synthétiques marqués par des isotopes stables) (Fig. 2). C'est pourquoi ces approches sont de plus en plus envisagées dans le contexte clinique pour valider des candidats biomarqueurs, diagnostiquer une maladie ou suivre l'effet d'un traitement thérapeutique.

Dans le cadre des maladies de la peau, la PRM a été utilisée récemment pour valider la différence d'expression de protéines chez des patients souffrant de *Psoriasis vulgaris*. Après l'identification par la méthode de marquage au TMT™ de plusieurs cibles potentielles, des biopsies de patients sains et atteints de *Psoriasis vulgaris* ont été analysées par PRM. Au total, cinq protéines dont la quantité est significativement plus élevée, et quatre dont la quantité est significativement plus faible, ont été retrouvées chez les patients atteints de *Psoriasis vulgaris* (Li et al., 2020). Bien que cette première étude soit encourageante, il reste encore à valider ces biomarqueurs potentiels sur une cohorte de patients plus importante et à réaliser des quantifications absolues. D'autre part, plusieurs autres travaux ont caractérisé le profil protéomique du sérum de patients atteints de dermatite atopique, de dermatite de contact ou de psoriasis. Si ces pathologies de la peau partagent des

dérégulations communes, il est toutefois possible d'identifier des signatures protéomiques propres à chacune d'elles (Brunner et al., 2017; Pavel et al., 2020). Grâce à ces données, il serait envisageable de développer un essai utilisant la protéomique ciblée pour suivre et quantifier plusieurs biomarqueurs spécifiques de chacune de ces pathologies afin de les diagnostiquer au mieux mais également de prédire le traitement préférentiel pour chaque patient. Un autre exemple de l'utilisation de la protéomique ciblée dans l'étude des biomarqueurs cutanés concerne les plaies chroniques. Les études effectuées par Sabino *et al.* sur un modèle porcin de cicatrisation ont permis d'identifier l'augmentation de l'abondance relative des protéines zyxine et HtrA1 comme biomarqueurs de la fermeture favorable des plaies (Sabino et al., 2018). Par la suite, une étude en PRM d'exsudats de plaies issus d'une petite cohorte de patients a permis de valider qu'un retard dans l'augmentation de l'abondance de ces protéines était caractéristique des plaies à cicatrisation lente (Sabino et al., 2018).

En résumé, la protéomique ciblée offre la possibilité de suivre plusieurs protéines dans une même analyse (multiplexage) et de les quantifier dans divers types de milieux complexes (biopsies de peau, exsudats de plaie, plasma/sérum).

#### **IV – Conclusion**

Nous avons vu que la protéomique offre de multiples possibilités pour identifier et quantifier des biomarqueurs de pathologies cutanées. Si la mise en évidence de nouveaux biomarqueurs est maintenant relativement bien couverte par les laboratoires académiques, l'enjeu actuel consiste à sélectionner les biomarqueurs les plus pertinents et à les valider sur de plus grandes cohortes de patients. Ceci nécessitera l'intégration des méthodes de protéomique ciblée dans le milieu hospitalier. Même si cette transition est lente du fait du coût des équipements et du haut degré de technicité requise, elle a déjà commencé, comme en témoigne le projet idBIORIV, coordonné par les hospices civils de Lyon (2019-2024 ; <https://idbioriv.com>), visant à développer une méthode MRM de diagnostic du sepsis en moins d'une heure. On peut donc envisager son déploiement à plus grande échelle dans un futur proche, y compris pour des applications de dermatologie. Comme nous l'avons vu au travers des exemples donnés ci-dessus, la protéomique ciblée

pourrait aider à distinguer rapidement plusieurs formes proches d'affections cutanées et donc à mieux adapter le traitement, à prédire quels patients sont plus susceptibles de présenter un retard de cicatrisation après une plaie ou encore à mieux comprendre les mécanismes impliqués dans l'effet thérapeutique d'un candidat médicament lors d'un essai clinique.

## Bibliographie

- Anand, S., M. Samuel, C.S. Ang, S. Keerthikumar, and S. Mathivanan. 2017. Label-based and label-free strategies for protein quantitation. *Methods Mol Biol.* 1549:31-43.
- auf dem Keller, U., A. Prudova, U. Eckhard, B. Fingleton, and C.M. Overall. 2013. Systems-level analysis of proteolytic events in increased vascular permeability and complement activation in skin inflammation. *Sci Signal.* 6:rs2.
- Barallobre-Barreiro, J., E. Woods, R.E. Bell, J.A. Easton, C. Hobbs, M. Eager, F. Baig, A.M. Ross, R. Mallipeddi, B. Powell, M. Soldin, M. Mayr, and T.J. Shaw. 2019. Cartilage-like composition of keloid scar extracellular matrix suggests fibroblast mis-differentiation in disease. *Matrix Biol Plus.* 4:100016.
- Brunner, P.M., M. Suarez-Farinas, H. He, K. Malik, H.C. Wen, J. Gonzalez, T.C. Chan, Y. Estrada, X. Zheng, S. Khattri, A. Dattola, J.G. Krueger, and E. Guttman-Yassky. 2017. The atopic dermatitis blood signature is characterized by increases in inflammatory and cardiovascular risk proteins. *Sci Rep.* 7:8707.
- Chen, C., J. Hou, J.J. Tanner, and J. Cheng. 2020. Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. *Int J Mol Sci.* 21:2873.
- Chularojanamontri, L., N. Charoenpipatsin, N. Silpa-Archa, C. Wongpraparut, and V. Thongboonkerd. 2019. Proteomics in Psoriasis. *Int J Mol Sci.* 20:1141.
- Dengjel, J., L. Bruckner-Tuderman, and A. Nystrom. 2020. Skin proteomics - analysis of the extracellular matrix in health and disease. *Expert Rev Proteomics.* 17:377-391.
- Donnelly, D., 3rd, P.P. Aung, and G. Jour. 2019. The "-OMICS" facet of melanoma: Heterogeneity of genomic, proteomic and metabolomic biomarkers. *Semin Cancer Biol.* 59:165-174.

- Dussoyer, M., A. Page, F. Delolme, P. Rousselle, A. Nystrom, and C. Moali. 2022. Comparison of extracellular matrix enrichment protocols for the improved characterization of the skin matrisome by mass spectrometry. *J Proteomics*. 251:104397.
- Dyring-Andersen, B., M.B. Lovendorf, F. Coscia, A. Santos, L.B.P. Moller, A.R. Colaco, L. Niu, M. Bzorek, S. Doll, J.L. Andersen, R.A. Clark, L. Skov, M.B.M. Teunissen, and M. Mann. 2020. Spatially and cell-type resolved quantitative proteomic atlas of healthy human skin. *Nat Commun*. 11:5587.
- Emadali, A., and M. Gallagher-Gambarelli. 2009. [Quantitative proteomics by SILAC: practicalities and perspectives for an evolving approach]. *Med Sci (Paris)*. 25:835-842.
- Kaleja, P., H. Emmert, U. Gerstel, S. Weidinger, and A. Tholey. 2020. Evaluation and improvement of protein extraction methods for analysis of skin proteome by noninvasive tape stripping. *J Proteomics*. 217:103678.
- Leduc, C., L. Dupont, L. Joannes, C. Monseur, D. Baiwir, G. Mazzucchelli, C. Deroanne, A. Colige, and M. Bekhouche. 2021. In vivo N-terminomics highlights novel functions of ADAMTS2 and ADAMTS14 in skin collagen matrix building. *Front Mol Biosci*. 8:643178.
- Li, Y., P. Lin, S. Wang, S. Li, R. Wang, L. Yang, and H. Wang. 2020. Quantitative analysis of differentially expressed proteins in psoriasis vulgaris using tandem mass tags and parallel reaction monitoring. *Clin Proteomics*. 17:30.
- Lindemann, C., N. Thomanek, F. Hundt, T. Lerari, H.E. Meyer, D. Wolters, and K. Marcus. 2017. Strategies in relative and absolute quantitative mass spectrometry based proteomics. *Biol Chem*. 398:687-699.
- Ozols, M., A. Eckersley, K.T. Mellody, V. Mallikarjun, S. Warwood, R. O'Cualain, D. Knight, R.E.B. Watson, C.E.M. Griffiths, J. Swift, and M.J. Sherratt. 2021. Peptide location fingerprinting reveals modification-associated biomarker candidates of ageing in human tissue proteomes. *Ageing Cell*. 20:e13355.
- Pavel, A.B., L. Zhou, A. Diaz, B. Ungar, J. Dan, H. He, Y.D. Estrada, H. Xu, M. Fernandes, Y. Renert-Yuval, J.G. Krueger, and E. Guttman-Yassky. 2020. The proteomic skin profile of moderate-to-severe atopic dermatitis patients shows an inflammatory signature. *J Am Acad Dermatol*. 82:690-699.

## ORAL AND POSTER COMMUNICATIONS, PRIZES

---

Oral communication 1:

**Proteomic characterization of skin extracellular matrix alterations in dystrophic epidermolysis bullosa.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*LBTI internal seminar. 2020, December 15<sup>th</sup>. Lyon, France.*

Oral communication 2:

**Proteomic characterization of skin extracellular matrix alterations in dystrophic epidermolysis bullosa.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*Early-career pre-conference of the oint Meeting of the German (DGMB) and the French (SFBMec) Matrix Biology Societies. 2021, May 26<sup>th</sup>-28<sup>th</sup>. Virtual format.*

Oral communication 3 (Best oral communication prize):

**Proteomic characterization of skin extracellular matrix alterations in dystrophic epidermolysis bullosa.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*26<sup>th</sup> scientific day of the Sciences-health interdisciplinary doctoral school (EDISS). 2021, November 18<sup>th</sup>. Virtual format.*

Oral communication 4:

**Strong alterations of collagen biosynthesis and fibrillogenesis in the dermis of recessive dystrophic epidermolysis bullosa (RDEB) mice.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*ECM pharmacology congress. 2022. June 23<sup>th</sup>-25<sup>th</sup>. Copenhagen, Denmark.*

Poster 1:

**Identification of extracellular matrix-derived biomarkers and targets in Dystrophic Epidermolysis Bullosa.** Dussoyer, M., Kleiser, S, Kiritsi, D., Lagoutte, P., Rousselle, P., Vadon-Le Goff, S., Lemoine J., Nyström, A. & Moali, C.

*EMBO Practical course. "Targeted proteomics, experimental design and data analysis". 2019, November 10<sup>th</sup>-15<sup>th</sup>. Barcelona, Spain.*

Poster 2 (Best poster prize):

**Proteomic characterization of skin extracellular matrix alterations in dystrophic epidermolysis bullosa.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*Joint Meeting of the German (DGMB) and the French (SFBMec) Matrix Biology Societies. 2021, May 26<sup>th</sup>-28<sup>th</sup>. Virtual format.*

Poster 3:

**Caractérisation des altérations de la matrice extracellulaire de la peau dans l'Epidermolyse bulleuse dystrophique.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*Annual congress of dermatology research (CARD). 2021, July 1<sup>st</sup>. Virtual format.*

Poster 4:

**Proteomic characterization of skin extracellular matrix alterations in dystrophic epidermolysis bullosa.** Dussoyer, M., Kleiser, S, Page, A., Delolme, F., Kiritsi, D., Vadon-Le Goff, S., Rousselle, P., Nyström, A. & Moali, C.

*DEB research association (DEBRA) international congress. 2021, September 16<sup>th</sup>-19<sup>th</sup>. Virtual format.*

# TRAINING AND COMPLEMENTARY ACTIVITIES

---

**EMBO Practical course “Targeted proteomics, experimental design and data analysis” (5 days)**

2019, November 10<sup>th</sup>-15<sup>th</sup>. Barcelona, Spain.

**Training course “Draw high-quality graphics with Inkscape” (7 h)**

2019, November 26<sup>th</sup>. University Claude Bernard – Lyon 1, Lyon, France.

**MOOC « Ethique de la recherche » (15 h)**

2020, Autumn session. University Claude Bernard – Lyon 1, Lyon, France.

**Training course “La voix comme instrument de travail” (16 h)**

2020, March 13<sup>th</sup> and 2021, October 12<sup>th</sup>. National Center for Scientific Research (CNRS).

**Training course « S'affirmer face à des situations et à des personnalités difficiles » (21 h)**

2020, October 15<sup>th</sup>-16<sup>th</sup>. University Claude Bernard – Lyon 1, Lyon, France.

**Course “Cours de Biologie de la peau” (COBIP, 3 days)**

2021, March, 15<sup>th</sup>-17<sup>th</sup>. University Claude Bernard – Lyon 1. Virtual format.

**Thematic workshop « Analyses protéomiques quantitatives : choix de méthodes et analyses des données » (4 days)**

2021, January 25<sup>th</sup>-28<sup>th</sup>. National Center for Scientific Research (CNRS). Virtual format.

**Teaching complementary activity (« activité complémentaire d’enseignement », 64 h/year)**

Biochemistry (50 %), Cell biology (50 %). Practical works & tutorials, levels Licence 1 to Licence 3.

2019-2022. University Claude Bernard – Lyon 1, Lyon, France.

## FUNDING

---

This project was part of the Franco-German DEBTARGET project, co-funded by the French national agency for research (ANR, ANR-18-CE92-0035) and the German research foundation (DFG, NY90/5-1).