Exploring CEMIP axis to identify new therapeutic strategies in glomerular diseases

Spataro, Sofia Carmen Suzanne

UNIVERSITÉ DE GENÈVE
Section de Sciences Pharmaceutiques
Laboratoire de Biochimie Pharmaceutique

FACULTÉ DES SCIENCES
Professeur Leonardo Scapozza
PD Docteur Marco Prunotto

EXPLORING CEMIP AXIS TO IDENTIFY NEW THERAPEUTIC STRATEGIES IN GLOMERULAR DISEASES

THÈSE

présentée aux Facultés de médecine et des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences en sciences de la vie,
mention Sciences pharmaceutiques

par
Sofia Carmen Suzanne SPATARO
de
Giubiasco (Tessin)

Thèse N° 261

GENÈVE
Atelier de reproduction Repromail
2023

DOCTORAT ÈS SCIENCES EN SCIENCES DE LA VIE DES
FACULTÉS DE MÉDECINE ET DES SCIENCES
MENTION SCIENCES PHARMACEUTIQUES

Thèse de Mme Sofia Carmen Suzanne Spataro

intitulée :

« Exploring CEMIP Axis to Identify New Therapeutic Strategies in Glomerular Diseases »

Les Facultés de médecine et des sciences, sur le préavis de Monsieur Leonardo SCAPOZZA, Professeur ordinaire et directeur de thèse (Section des sciences pharmaceutiques), Monsieur Marco PRUNOTTO, Docteur et co-directeur de thèse (Section des sciences pharmaceutiques), Monsieur Yogeshvar KALIA, Professeur ordinaire (Section des sciences pharmaceutiques), Monsieur Jeffrey H. MINER, Professeur (Division of Nephrology, Washington University in St. Louis, Missouri, United States), Monsieur Jonathan SLEEMAN, Professeur (Vascular Biology & Medicine, Medical Faculty of Mannheim, Heidelberg University, Mannheim, Germany) et Monsieur Rafael FRIDMAN, Professeur (Department of Pathology, School of medicine, Wayne State University, Michigan, United States) autorisent l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 8 janvier 2024

Thèse - 261 -

Le Doyen
Faculté de médecine

La Doyenne
Faculté des sciences

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

FACULTÉ DE MÉDECINE
FACULTÉ DES SCIENCES

UNIVERSITÉ DE GENÈVE

To my family,

and to Francis,
the love of my life
♥

TABLE OF CONTENTS

1 LIST OF ABBREVIATIONS

AA	amino acid
ACE	angiotensin-converting-enzyme
ANXA1	annexin A1
AS	Alport syndrome
ASO	antisense oligonucleotide
AUC	area under the curve
BUN	blood urea nitrogen
CEMIP	cell migration-inducing protein
CGN	crescentic glomerulonephritis
CKD	chronic kidney disease
CPC	cetylpyridinium chloride
Cryo-EM	cryogenic electron microscopy
Cryo-ET	cryogenic electron tomography
CryoSPARC	cryo-EM single particle ab-initio reconstruction and classification
DCI	Dubochet center for imaging
DDR1/2	discoidin domain receptor tyrosine kinase 1/2
DEL	DNA-encoded library
DMEM	Dulbecco's modified eagle medium
DSF	differential scanning fluorimetry
ECM	extracellular matrix
EGFR	epidermal growth factor receptor
EMT	epithelial–mesenchymal transition
ESRD	end-stage renal disease
FCS	fetal calf serum
FLS	fibroblast-like synoviocytes
GAGs	glycosaminoglycans
GBM	glomerular basement membrane
GFR	glomerular filtration rate
GN	glomerulonephritis

GS	GenScript
HA	hyaluronic acid
HA-tag	hemagglutinin tag
HABP	HA-binding protein
HAS	hyaluronan synthases
HB-EGF	heparin-binding epidermal growth factor-like growth factor
HYAL	hyaluronidases
HMW	high molecular weight
KO	knock out
LMW	low molecular weight
MALS	multiple angle light scattering
MoA	mode of action
MR	mineralocorticoid receptor
MW	molecular weight
NMR	nuclear magnetic resonance
NTS	nephrotoxic serum
PCR	polymerase chain reaction
PEC	parietal epithelial cell
PEG	polyethylene glycol
PEI	polyethylenimine
PFA	paraformaldehyde
PGs	proteoglycans
POMGNT1	protein O-linked-mannose β-1,2-N-acetylglucosaminyltransferase 1
RA	rheumatoid arthritis
RAAS	renin-angiotensin-aldosterone system
RI	refractive index
RIPA	radioimmunoprecipitation assay buffer
ROS	reactive oxygen species
RPGN	rapidly progressive glomerulonephritis
RTK	receptor tyrosine kinase
SAR	structure-activity relationships

SEC	size exclusion chromatography
SGLT2	sodium-glucose co-transporter-2
sHA	sulfated hyaluronic acid
SPA	single particle analysis
SPR	surface plasmon resonance
TMEM2	transmembrane protein 2
WB	Western blot

2 ABSTRACT

Glomerular diseases, also called glomerulonephritides (GNs), refer to all those conditions that affect glomeruli, the filtering units of the kidney. If left untreated or not managed effectively, progressive GNs may result in a range of complications, including chronic kidney disease and end-stage renal disease. Some pathophysiological aspects of GNs are treatable, with several novel mode of action (MoA) drugs reaching the market in recent years, but a strong unmet medical need persists and identification of novel potential GN therapeutic targets is still an active field of research in nephrology. In the present thesis work, I present evidence of the relevance of a newly identified hyaluronic acid degrading protein named Cell migration-inducing and hyaluronan-binding protein (CEMIP). CEMIP is specifically upregulated upon activation of the collagen receptor Discoidin Domain Receptor 1 (DDR1), an accepted molecular target shown to have a role in GN initiation and progression. Analysis of patient biopsies and CEMIP deletion experiment in a mouse model of GN suggested that the DDR1/CEMIP axis could play a role in renal pathologies characterized by a crescentic phase. This role may be mediated by CEMIP hyaluronidase activity, which generates low molecular weight hyaluronic acid (LMW-HA) fragments, known to be signaling molecules in inflammation. To date, very little scientific knowledge has been reported on the activity, mechanism and mode of action of these CEMIP-generated fragments. Our current grasp of CEMIP biology is dominated by the vast body of literature that has investigated CEMIP expression and/or deletion in cancer models such as in the context of tumor proliferation and invasion. After preliminary results assessed the role of CEMIP in GN, in the present work I sought to study its molecular MoA. Mutagenesis experiments confirmed the position of the key amino acid residues used by CEMIP to bind selectively to hyaluronic acid and identified, for the first time, the ones belonging to CEMIP postulated catalytic site. Additional experiments shed light on the role and importance of other domains of CEMIP as well as on the catalytic function of intracellular and secreted CEMIP. This molecular knowledge was used to design, synthesize and evaluate, using newly developed cellular assays, small molecule inhibitors on CEMIP hyaluronidase activity. A parallel effort was conducted to solve CEMIP structure by cryo-EM and one of its domains (the GG domain) by X-ray crystallography. All those efforts will hopefully result in the creation of CEMIP inhibitors active *in vivo.* The availability of such tool compounds will be then tested in preclinical model of GNs and cancer.

3 RÉSUMÉ DE LA THÈSE

Les maladies glomérulaires, également appelées glomérulonéphritides (GNs), désignent toutes les affections qui touchent les glomérules, les unités de filtration du rein. Si elles ne sont pas traitées ou si elles ne sont pas prises en charge efficacement, les GNs évolutives peuvent entraîner une série de complications, notamment une maladie rénale chronique et une insuffisance rénale terminale. Certains aspects physiopathologiques des GNs sont traitables, plusieurs nouveaux médicaments à mode d'action (MoA) ayant été mis sur le marché ces dernières années, mais un important besoin médical non satisfait persiste et l'identification de nouvelles cibles thérapeutiques potentielles pour les GN reste un domaine de recherche actif en néphrologie. Dans ce travail de thèse, je présents des preuves de la pertinence d'une protéine de dégradation de l'acide hyaluronique récemment identifiée, appelée Cell migration-inducing and hyaluronan-binding protein (CEMIP). CEMIP est spécifiquement régulée à la hausse lors de l'activation du récepteur du collagène Discoidin Domain Receptor 1 (DDR1), une cible moléculaire reconnue pour son rôle dans l'initiation et la progression de la GN. L'analyse de biopsies de patients et des expériences avec la délétion de CEMIP dans un modèle murin de GN suggèrent que l'axe DDR1/CEMIP pourrait jouer un rôle dans les pathologies rénales caractérisées par une phase *crescentic*. Ce rôle pourrait être médié par l'activité hyaluronidase de CEMIP, qui génère des fragments d'acide hyaluronique de faible poids moléculaire (LMW-HA), connus pour être des molécules de signalisation dans l'inflammation. À ce jour, très peu de connaissances scientifiques ont été rapportées sur l'activité, le mécanisme et le mode d'action de ces fragments générés par CEMIP. Notre compréhension actuelle de la biologie de CEMIP est dominée par la vaste littérature qui a étudié l'expression et/ou la délétion de CEMIP dans des modèles de cancer, par exemple dans le contexte de la prolifération et de l'invasion tumorales. Après que des résultats préliminaires ont permis d'évaluer le rôle de la CEMIP dans la GN, j'ai cherché dans le présent travail à étudier son MoA moléculaire. Des expériences de mutagenèse ont confirmé la position des résidus d'acides aminés clés utilisés par CEMIP pour se lier sélectivement à l'acide hyaluronique et ont identifié, pour la première fois, ceux qui appartiennent au site catalytique postulé de CEMIP. Des expériences supplémentaires ont mis en lumière le rôle et l'importance d'autres domaines de CEMIP ainsi que la fonction catalytique de CEMIP intracellulaire et sécrétée. Ces connaissances moléculaires ont été utilisées pour concevoir, synthétiser et évaluer, à l'aide de tests cellulaires

nouvellement développés, des petites molécules inhibitrices de l'activité hyaluronidase de CEMIP. Un effort parallèle a été mené pour résoudre la structure de CEMIP par cryo-EM et l'un de ses domaines (le domaine GG) par cristallographie aux rayons X. Tous ces efforts devraient aboutir à la création d'inhibiteurs de CEMIP actifs in vivo. La disponibilité de ces composés outils sera ensuite testée dans des modèles précliniques de GN et de cancer.

4 INTRODUCTION

The introduction of this thesis manuscript includes parts of my published review article[1].

4.1 Glomerular diseases

Glomerular diseases, also called glomerulonephritides (GN), refer to all those conditions that affect the kidney's glomeruli. They can be autoimmune diseases (lupus nephritis, IgA nephropathy), infection-related, diabetes-induced, or even of hereditary origin (Alport syndrome). When GN presents a rapid loss of renal function, is referred to as crescentic glomerulonephritis (CGN), or rapidly progressive glomerulonephritis (RPGN). Crescentic glomerulonephritis is characterized by crescents, histopathological lesions that develop from activated parietal epithelial cells (PECs) proliferating in the Bowman's capsule. The development of crescents leads to a reduction in patients' glomerular filtration rate (GFR). While cellular crescents can sometimes be reversible, if PECs undergo multilevel growth and epithelial-mesenchymal transition, they can transform into fibrous crescents, which are challenging to reverse in terms of GFR recovery.
Immunosuppressive medications (such as glucocorticoids), diuretics, angiotensin-converting enzyme (ACE) inhibitors and angiotensin II receptor blockers (ARBs) have been a longstanding treatment for glomerulonephritis. Moreover, various molecular pathways trigger PECs activation and serve as important targets for therapeutic approaches in crescentic glomerulonephritis. Studies report that silencing of CD44 and CD9 reduces PECs presence, glomerular cell proliferation, and albuminuria[2,3]. Other researchers report that targeting heparin-binding epidermal growth factor-like growth factor (HB-EGF) in parietal epithelial cells and podocytes, along with delayed epidermal growth factor receptor (EGFR) inhibition, offers promise for attenuating the condition and has translational potential in human rapidly progressive glomerulonephritis[4].

4.1.1 Alport syndrome

Alport syndrome (AS) is a genetic condition described for the first time in 1927 by Arthur Cecil Alport, who observed several case histories of patients with hereditary familial congenital hemorrhagic nephritis[5]. It is considered a rare disease, affecting 1 in 5'000-10'000 newborns[6,7]. AS is caused by mutations in collagen type IV, the major structural component of the basement

membrane. In fact, the network-forming nature of type IV collagen makes it essential for providing the backbone structure supporting growing epithelial and endothelial cells. In mammals, type IV collagen is assembled from six distinct polypeptide chains α1 to α6, encoded by six different genes COL4A1-COL4A6. Mutations in these six genes give rise to multiple diseases, such as congenital anomalies of the kidney and urinary tract and Alport syndrome (**Figure 1**).

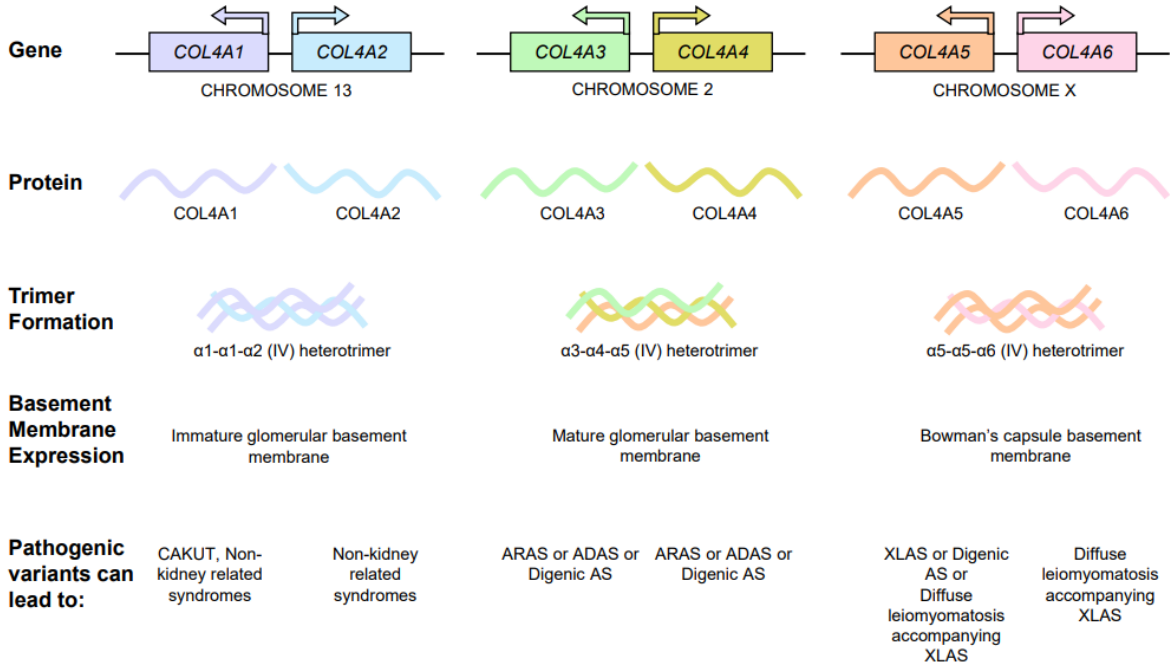


***Figure 1**: **Type IV collagen from gene to protein trimerization to disease**. The 6 collagens are arranged in a head-to-head configuration, with COL4A1 and COL4A2 on chromosome 13, COL4A3 and COL4A4 on chromosome 2, and COL4A5 and COL4A6 on chromosome X. The arrow above each gene indicates the direction in which transcription takes place. The collagen monomers come together to form heterotrimers, with the α1-α1-α2 (IV) expressed in the glomerular basement membrane during development, which later switches to the α3-α4-α5 (IV) in adulthood. Uniquely, the α5-α5-α6 (IV) heterotrimer is expressed in the basement membrane of Bowman's capsule. Pathogenic variants in these genes can lead to congenital abnormalities of the kidney and urinary tract (CAKUT) (COL4A1), autosomal recessive Alport syndrome (COL4A3 or COL4A4), autosomal dominant Alport syndrome (COL4A3 or COL4A4), digenic Alport syndrome (COL4A3 and COL4A4 or COL4A3/COL4A4 and COL4A5), or diffuse leiomyomatosis accompanying X-linked Alport syndrome (COL4A5 and COL4A6). ADAS, autosomal dominant Alport syndrome; ARAS, autosomal recessive Alport syndrome; AS, Alport syndrome; XLAS, X-linked Alport syndrome. From De Gregorio et al.*[8]

In 80% of the cases, AS is X-linked (XLAS), with a mutation in COL4A5[9]. In 15% of the cases, it is recessive, with a mutation in COL4A3 or 4 (ARAS), and in 5% of the cases it is dominant

(ADAS)[9]. Males with the X-linked inheritance of the disease show a severe phenotype with end-stage renal disease (ESRD) in the adolescence or even childhood[10]. Mutations in COL4A345 prevent the correct assembly of α3α4α5 chains, found in the glomerulus of the kidney, in the organ of Corti in the inner ear and in the retina of the eye. This explains why AS patients suffer from severe kidney problems, progressive hearing loss and eye abnormalities. Symptoms typically begin during childhood, and the first signs of the condition are usually hematuria and proteinuria, which lead to renal fibrosis, progressive loss of kidney function and almost inevitably to ESRD. In Alport syndrome, the manifestation of hematuria is attributed to the abnormal structure of the glomerular basement membrane, leading to the leakage of blood cells into the urine. Proteinuria results from the compromised filtration barrier in the kidneys, allowing proteins to escape into the urine. The progression to renal failure in Alport syndrome is primarily due to the cumulative damage to the renal tissue over time, resulting in impaired kidney function and eventual organ failure.

When mutations in COL4A345 occur, the resulting collagen protein doesn't assemble correctly, leading to impaired production and deposition of collagen networks[11]. As a result, the glomerular basement membrane (GBM) of an AS patient cannot be maintained and will progressively split (**Figure 2**).

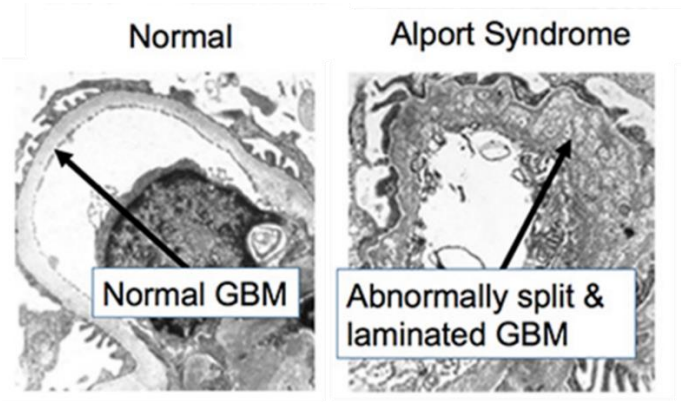


***Figure 2**: **Abnormal split of the GBM in an Alport Syndrome patient.** The normal GBM from a healthy individual (left) compared to the GBM of an AS patient (right). The GBM of the patient presents abnormal split and lamination. Modified from UNC Kidney Center.*

The etiology of Alport syndrome is yet to be fully understood. Previous studies have shown that the deletion of the collagen receptor DDR1 (discoidin domain receptor 1) in COL4A3 -/- mice, recognized model for CKD (chronic kidney disease)[12], increased their survival from 64.3 to 94.2 days[13], underlying the importance of podocytes in the pathogenesis of AS. In this scenario, the

atypical GBM composition would lead to aberrant chronic activation of collagen receptors, contributing to the onset of the disease (**Figure 3**).

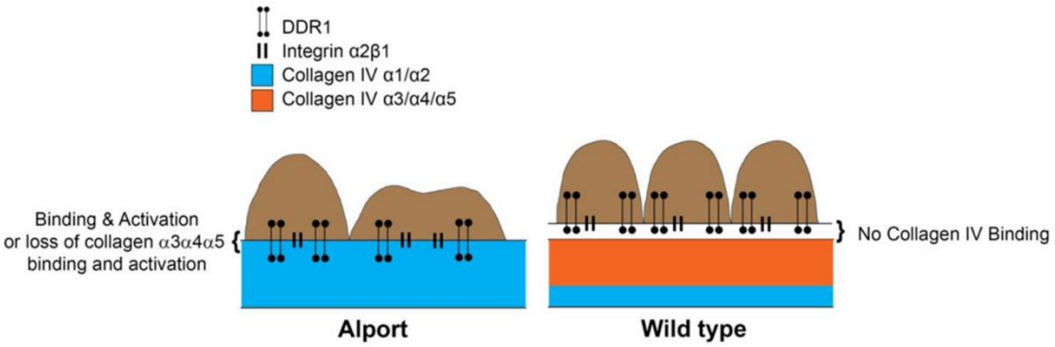


Figure 3:. ***Model for Alport glomerular disease initiation***. *Left: Alport GBM composition activates DDR1 since no collagen IV α3/ α4/ α5 is present (in orange). Right: Wild type GBM. From Cosgrove and Liu*[14].

Moreover, mechanical stress caused by proteinuria and hematuria injuries the GBM even more, generating fibroblast formation and renal fibrosis[15], thus creating a vicious cycle of GBM injury, fibrosis, thickening of the membrane, and mechanical stress.

4.1.1.1 Current treatments for Alport syndrome

Until now, no specific treatment is available for AS patients. Current treatments aim to control the progression of the disease and treat the symptoms, by reducing the blood pressure and filtration of large molecules (renin-angiotensin-aldosterone system - RAAS - inhibitors). In 2002, Gross and colleagues evaluated the effect of ramipril (angiotensin-converting enzyme inhibitor) on an established Alport syndrome mouse model, COL4A3-knockout mice, and found that ramipril is able to delay renal failure and to reduce renal fibrosis[16]. Ten years later, Oliver Gross conducted a study in AS patients revealing that early angiotensin-converting enzyme inhibition delays ESRD and improves life expectancy also in humans[17]. The sooner the treatment begins, the more the benefit for the patient, and the less the severity of the disease (**Figure 4**).

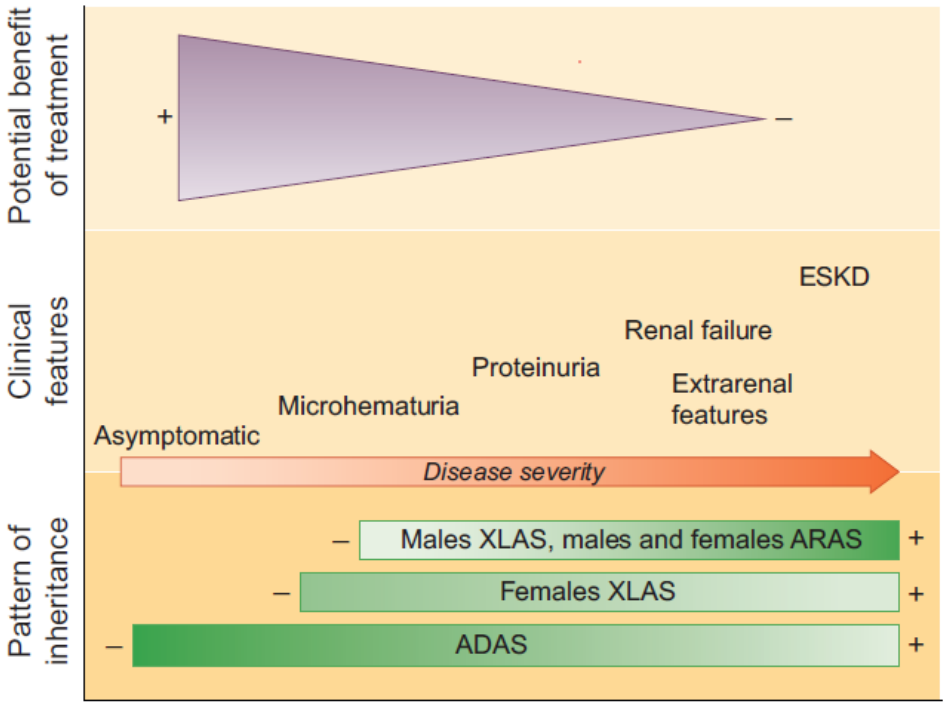


***Figure 4**: **Clinical presentation in AS and potential effect of treatment**. Upper: potential effect of treatment according to stage of disease. The potential impact decreases to the right as severity increases. Middle: clinical features that may be present in AS. Clinical features imply greater severity of disease towards the right. Lower: the stronger shadow represents the more frequent presentation of the disease, according to pattern of inheritance.. From Torra and Furlano[18].*

Other novel therapeutics include bardoxolone, anti-miRNA-21, paricalcitol, sparsentan, epidermal growth factor receptor (EGFR) inhibitors, lipid-lowering agents, chaperones treatment and stem-cell based therapies[18,19] (**Figure 5**).

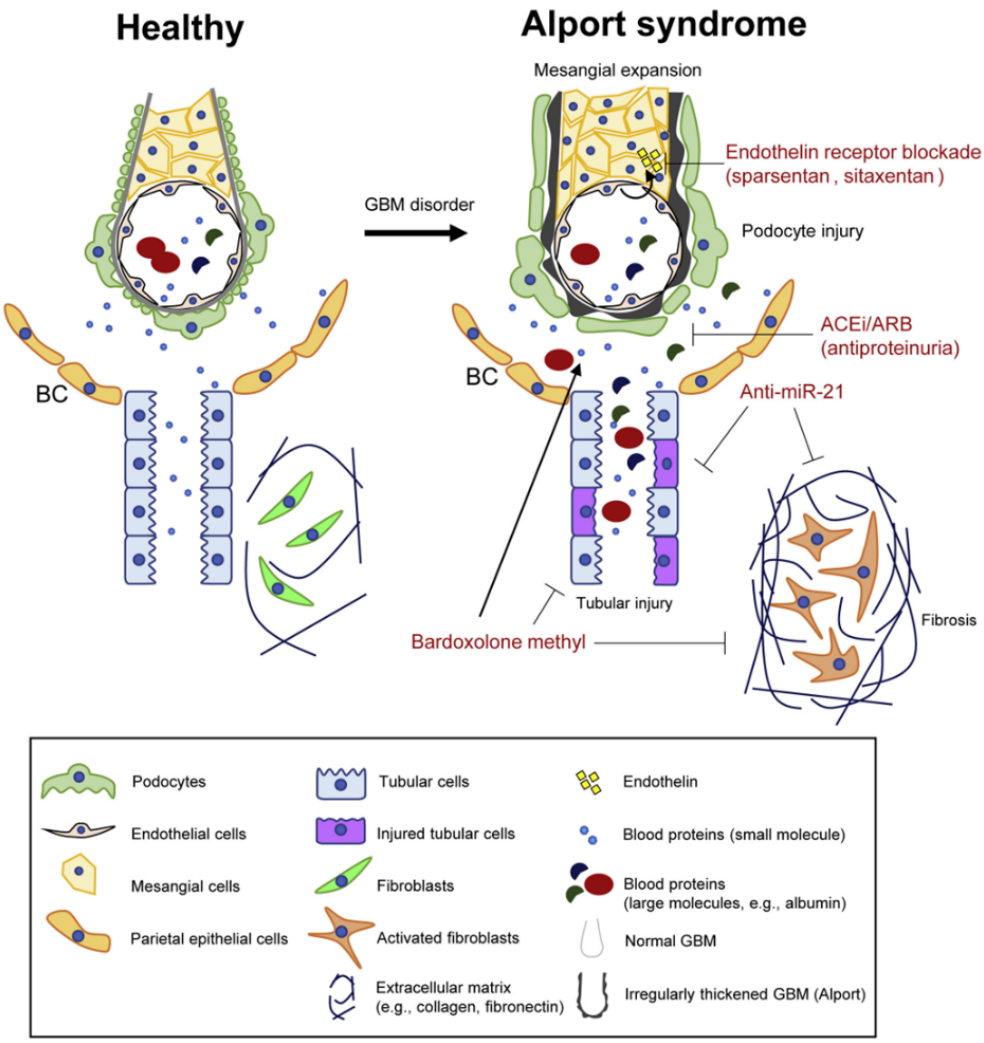


***Figure 5**: **AS at a glomerulus level**. Schematic of a Partial Glomerular Capillary, Bowman's Capsule (BC), and Proximal Tubule in Healthy and Alport Syndrome (AS) Situations. In healthy states, the glomerular basement membrane (GBM) is lined by endothelial cells and covered by podocytes with narrow foot processes; small molecules are filtered while larger ones are retained in the capillaries; and interstitial fibroblasts are quiescent. In AS, the GBM is irregularly thickened; podocyte foot processes become effaced; endothelial cells produce endothelin-1 (yellow squares), which activates endothelin receptor type A (ETRA) on mesangial cells. The mesangium expands; large molecules are filtered and enter the tubular lumen; tubular epithelial cells are injured; and interstitial fibroblasts become activated and produce fibrotic matrix proteins that contribute to nephron dysfunction. Moreover, indicated in red font are the activities of the drugs discussed in the article. Angiotensin converting enzyme inhibitor (ACEi) and angiotensin II receptor blocker (ARB) reduce blood pressure and filtration of large molecules, resulting in reduced proteinuria. Anti-miR-21 protects tubular cells by stimulating metabolic pathways and reduces interstitial fibrosis. Bardoxolone methyl increases estimated glomerular filtration rate (eGFR) and inhibits tubular cell injury and fibrosis by preserving mitochondrial function. Sparsentan and sitaxentan antagonize ETRA activation in mesangial cells and reduce blood pressure. From Omachi and Miner[19].*

Recently, an emerging concept for AS treatment has been the one of combination therapy. Zhu *et al*. describe an improvement in survival of COL4A3-knockout mice by adding finerenone (nonsteroidal mineralocorticoid receptor antagonist) to ramipril and empagliflozin (sodium-glucose co-transporter-2, SGLT-2, inhibitor), suggesting that the triple RAS/SGLT2/MR blockade might represent the next step in AS treatment[20]. The inclusion of finerenone, ramipril, and empagliflozin, collectively targeting the renin-angiotensin-aldosterone system (RAAS), sodium-glucose co-transporter-2 (SGLT-2), and mineralocorticoid receptor (MR) pathways, suggests a comprehensive approach to delaying the progression of AS. Importantly, the potential of this triple blockade extends beyond AS, serving as a promising avenue for the preventive treatment of diverse glomerulopathies characterized by proteinuria. However, a specific therapy for AS remains an unmet medical need, and inevitably chronic kidney failure progresses to end-stage kidney disease, requiring dialysis or transplantation[18].

4.2 The DDR1/CEMIP/HA axis

4.2.1 The role of DDR1 in glomerulonephritis

Renal epithelial cells are able to sense the GBM via the discoidin domain receptor 1 (DDR1), which binds collagen and plays a major role in GN. DDR1, together with its close analog DDR2, belongs to a specific family of receptors: the receptor tyrosine kinases (RTKs). RTKs are transmembrane proteins, with their tyrosine kinase domain at the C-terminal located inside the cytoplasm and their discoidin domain (DS) and discoidin-like domain located extracellularly at the N-terminal[21] (**Figure 6**). DDRs are the only collagen-activated RTKs, with a collagen binding site in the discoidin domain that selectively recognizes collagen with three conserved surface-exposed loops[22].

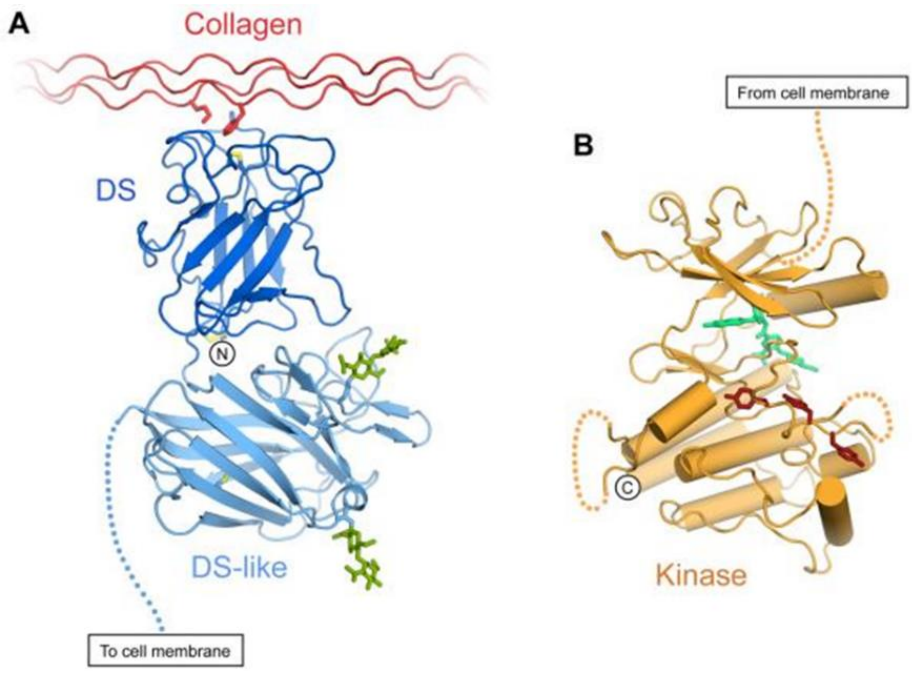


***Figure 6**: **Crystal structure of DDR1 globular domains**. A) Extracellular domains. Cartoon drawing of most of the DDR1 ectodomain bound to collagen, represented by a composite of the DDR2 DS domain bound to a collagen-like peptide and the structure of the DS and DS-like domain of DDR1. The side chains of key collagen residues (Met21 of the leading chain, Phe23 of the middle chain) are shown. Disulphide bonds are depicted in yellow. N-linked glycans in the DS-like domain are shown in green. The DDR1 N-terminus is indicated. The dashed line represents the extracellular JM region B) Intracellular domain. Structure of the DDR1 kinase domain in complex with type II inhibitor DDR1-IN-1. The inhibitor is depicted in green, tyrosine residues of the activation loop (Tyr792, Tyr796 and Tyr797; DDR1b numbering) are shown in brown. Dashed lines represent loop regions with poor electron density. The dashed line towards the top of the Figure represents the cytoplasmic JM region. The DDR1 C-terminus (Val913) is indicated. From Leitinger[21].*

In the kidney, DDR1 is expressed in various cell types, including renal tubular epithelial cells, mesangial cells, and podocytes. DDR1 is implicated in cell functions such as cell adhesion[23], tissue repair [24], inflammation[25] and fibrosis[24]. Studies have shown that DDR1 is involved in maintaining the structural integrity of the glomerular basement membrane (GBM) by facilitating cell-matrix communication[26], and its deficiency can lead to disrupted anchorage of foot processes and slit diaphragm in podocytes[27]. DDR1 also plays a role in glomerulosclerosis and renal fibrosis, as evidenced by its involvement in inflammatory responses and collagen expression[26,28,29]. Additionally, experiments with DDR1-deficient mice and therapeutic interventions using DDR1-specific antisense oligonucleotides (ASOs)[30] and inhibitors[31,32] have demonstrated its significance

in protecting against kidney diseases, including glomerulonephritis and tubule-interstitial injury. These findings highlight DDR1 as a crucial mediator in the pathogenesis of renal fibrosis and glomerulosclerosis.

As mentioned above, previous studies showed a reduction of renal damage in DDR1 genetic deletion mouse models[13]. Moreover, Dr. Marco Prunotto (co-director of this thesis) and colleagues have demonstrated that the use of a potent DDR1 inhibitor improves both histological parameters and renal function in crescentic GN[31] and Alport syndrome[32] mouse models. For this reason, a series of DDR1 inhibitors was developed and was found very specific and highly selective[31-34]. Unfortunately, the clinical value of such approaches remains limited due to the high level of target inhibition required for these ATP-binding site competitor kinase inhibitors (more than 95%). With the currently available inhibitors, the high concentrations required for clinical efficacy would also inhibit the DDR2 receptor, likely causing susceptibility to liver fibrosis[35] and perturbation of normal choroidal and retinal vascularization[36], thus leading to undesirable and unacceptable side-effects.

4.2.2 Identification of CEMIP as a downstream effector of DDR1 activation

In order to overcome the limitation of the DDR1 inhibitors, Dr. Prunotto and Dr. Cavalli identified candidate genes involved in the DDR1 downstream mode of action. A series of mRNA sequencing experiments were conducted using transgenic overexpressing cells cultivated in the presence or absence of collagen type II. Interestingly, they found that in the presence of collagen, DDR1 induces the expression, by a 16-fold change, of a protein called CEMIP (**Figure 7**).

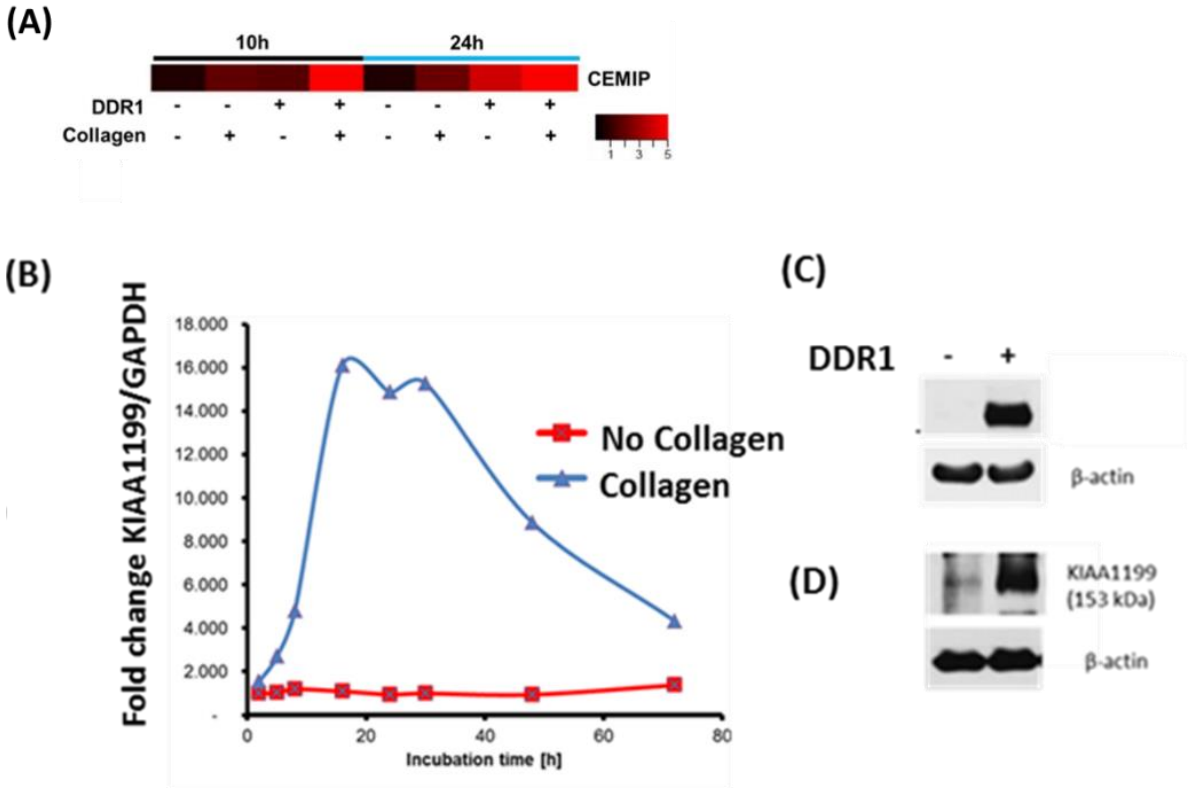


***Figure 7*: CEMIP is induced by collagen type II-activated DDR1.** *A) Original sequencing experiment showing expression of KIAA1199 induced by collagen in the presence of DDR1. B) qPCR time-course in the absence/presence of DDR1. C and D) Modulation of KIAA1199 at the protein level. Courtesy of Jing Ke, Jonathan Sleeman, Marco Prunotto and Andrea Cavalli.*

CEMIP, also known as KIAA1199 or HYBID, is known for mediating hyaluronic acid (HA) depolymerization independently of CD44 and HYAL1/2[37] (**Figure 8**). In support to the idea that CEMIP could be involved in the disease progression, alteration of HA metabolism has been demonstrated to play a crucial role in experimentally induced crescentic GN[38].

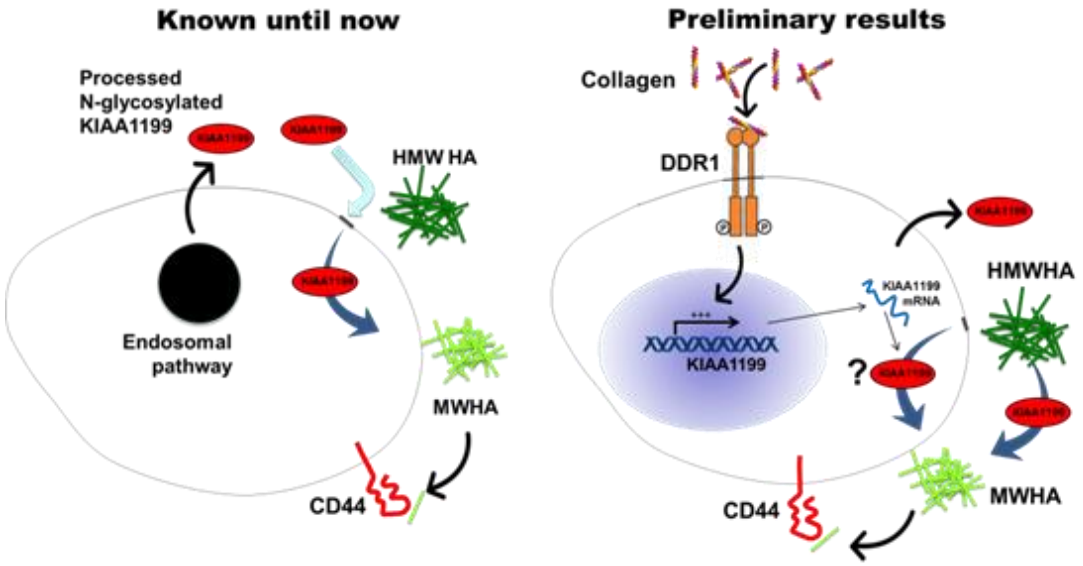


***Figure 8**: **Preliminary results of this project**. Preliminary results of this project elucidate the expression of CEMIP (= KIAA1199) induced by the collagen activated DDR1 receptor. CEMIP then depolymerizes HA to low molecular weight hyaluronic acid fragments which may activate CD44. Courtesy of Marco Prunotto and Andrea Cavalli.*

This work aims to investigate the DDR1/CEMIP/HA axis (**Figure 9**). In particular, I am interested in investigating CEMIP's role in and outside the cell, its activity and structure, and to study the possible inhibition of CEMIP via small molecules and/or antibodies. I aim to define if this protein would be a possible therapeutic target to develop a treatment for glomerulonephritis, as well as renal diseases in general, by providing renal protection for patients.

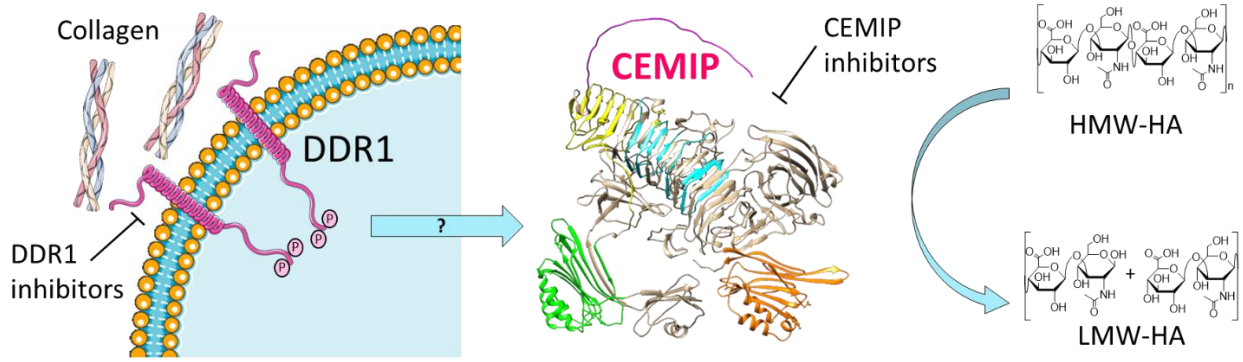


***Figure 9**: **The DDR1/CEMIP/HA axis**. After collagen stimulation, DDR1 induces the expression of CEMIP by an unknown mechanism (indicated by the question mark). CEMIP is responsible for the depolymerization of high molecular weight hyaluronic acid into smaller fragments. Note that CEMIP structure is the prediction model from AlphaFold[39].*

4.3 CEMIP: structure, function and expression

In light of the relevance of CEMIP for this research, this section aims to describe CEMIP structure, function as a hyaluronidase, putative role in HA-mediated biological functions, its role in glomerulonephritis and other diseases, regulation by positive and negative pathways, and its expression in tissues.

4.3.1 CEMIP structure: a large protein with yet unsolved atomistic structure

CEMIP is composed of 1361 amino acids. Its atomistic structure remains to date unsolved, essentially due to difficulties in expression and purification of the protein. However, as pointed out by Yamaguchi[40] and Yoshida and co-workers[41], sequence analyses provide a first overview of its domain architecture. CEMIP is characterized by the presence of an N-terminal signal sequence of 30 amino acids, which directs the protein to the secretion pathway. Additionally, CEMIP consists of one G8 domain, two GG domains, four PbH1 (parallel β-helix repeats) domains, and several N-linked glycosylation sites identified throughout the entire sequence (**Figure 10A**)[42].

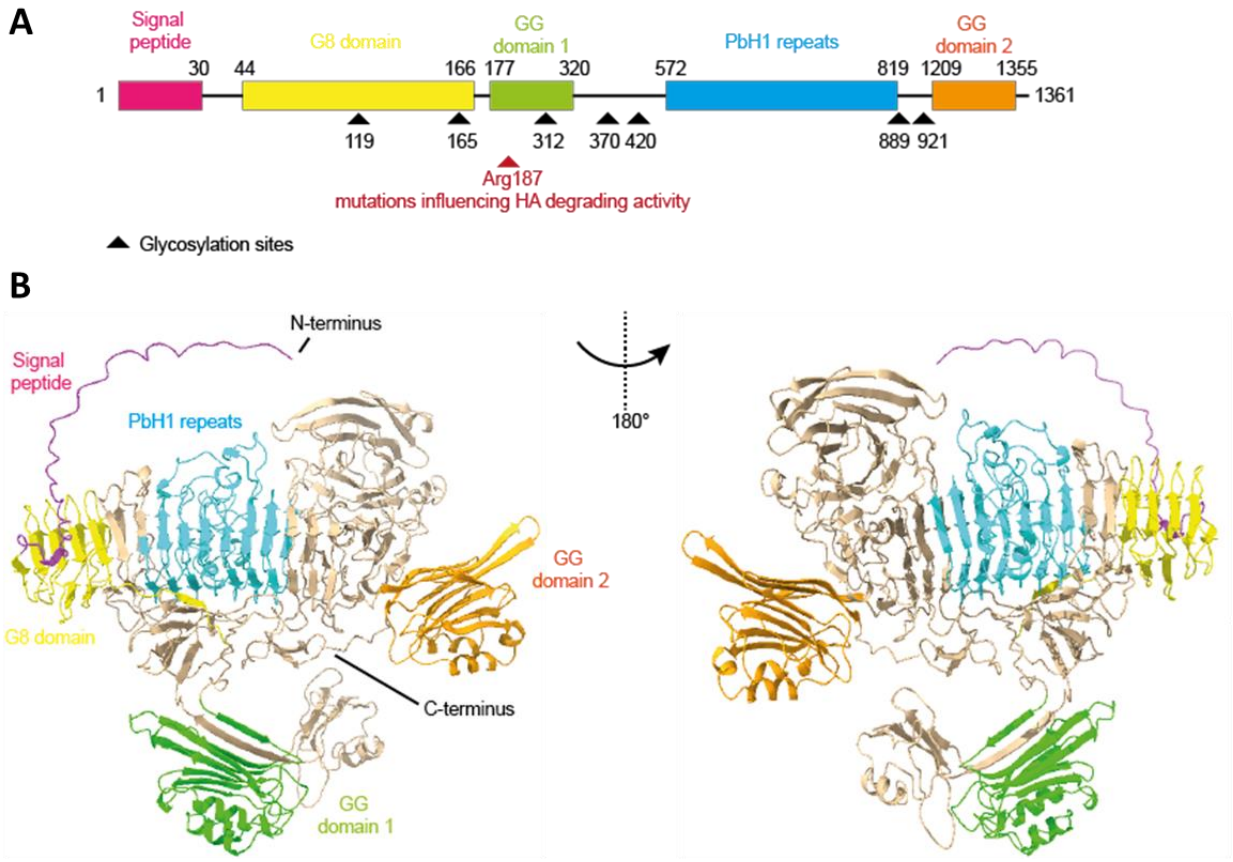


***Figure 10**: **CEMIP structure**. A) Schematic representation of the CEMIP secondary structure, annotated with mutation and glycosylation sites. Two examples, taken from the PDB of GG domain and PbH1 domains are shown as a cartoon for explicative purposes. B) Atomistic model. From AlphaFold, modified with USCF chimera.*

The G8 domain**,** identified by Lang and coworkers in 2006[43], is defined by the presence of eight conserved glycine residues. The domain is thought to form five consecutive β-strand pairs, although atomic coordinates of this type of domain have not yet been deposited in the protein data bank (PDB)[44]. G8 domains exist in several proteins from lower eukaryotes to animals, but are absent in plants, viruses and archaea. Notably, G8 domains have been identified in disease-related human proteins such as the Polycystic Kidney And Hepatic Disease 1 (PKHD1) protein, also known as Fibrocystin and Polyductin, and TMEM2, also known as hyaluronoglucosaminidase[43].

The GG domain has been identified in members of the eukaryotic FAM3 superfamily (FAM3A, FAM3B, FAM3C and FAM3D), POMGnT1 (protein O-linked mannose β-1,2-N-acetylglucosaminyltransferase), TMEM2 and the phage gp35 protein[45]. The structure of

POMGnT1 has been solved and deposited in the PDB[46]. The role of the GG domain is not yet fully understood. It might be involved in HA binding and/or degradation since a single mutation (R187H or R187C) in the first GG domain hinders the HA degrading activity[37,47] of CEMIP.

The PbH1 domain can be found in several polysaccharide degrading enzymes[48,49] and also in fibrocystin, a protein involved in the development of the autosomal recessive polycystic kidney disease[50]. However, the functional role of the PbH1 domain in CEMIP remains elusive.

As mentioned above, the atomic coordinates of CEMIP remain unsolved to date. Given the relevance of the three-dimensional structure of CEMIP for the design of therapeutics, Andrea Cavalli and Jacopo Sgrignani (IRB, Switzerland) used homology modeling techniques[51] to generate a CEMIP model. To this end, they leveraged proteins already deposited in the PDB and ran a template search using the Swissmodel web-server[46]. The search indicated that a template is only available for the first GG domain (*i.e.,* Protein O-linked-mannose beta-1,2-N-acetylglucosaminyl transferase 1, PDB code 5GGP[52], percentage of identity 30%)[51]. Other regions of the protein display only a low homology with proteins of known structure (less than 15-20%), which is insufficient to allow the generation of reliable models.

In 2021, Google Deepmind and the group of David Baker released two software packages, Alphafold2[39] and RoseTTAFold[53]. Both programs employ deep learning algorithms that enable prediction of protein structures at an unprecedented accuracy, even in the absence of suitable templates. Collaborators from IRB therefore analyzed the recently released AI-based postulated CEMIP structure available in the Alphafold2 database maintained by the EMBL (https://alphafold.ebi.ac.uk). Although deposited models are still unreliable concerning high-resolution information such as hydrogen bonds, side chain orientations or conformation of putative binding sites, these models are extremely valuable for interpreting and designing new experiments. Interestingly, the predicted CEMIP structure by Alphafold2 (**Figure 10B**) confirms the presence of a G8 domain (formed by 5 β-strand pairs) located immediately after the end of the signal peptide (Ala33 to Ser159) and followed by a GG domain (Glu175 to Gln321). The model is then characterized by the presence of two small subdomains not previously identified that encompass residues Asp322 to Pro406 and Asn418 to Leu494. In the region from Ile499 to Ser916 the

expected PbH1 domain is found, followed by a previously unreported domain formed by residues from Cys917 to Asp1208. The structural model further confirms the presence of the second GG domain in the C-terminal region of the protein (His1209 to Leu1361).

4.3.1.1 CEMIP and TMEM2

In addition to CEMIP and the HYAL family molecules, TMEM2, was recently reported to be a cell-surface transmembrane hyaluronidase in mouse organs that shares a high degree of sequence similarity with CEMIP[40]. From a structural point of view, TMEM2 has a transmembrane domain, one G8 domain, one GG domain and 3 PbH1 repeats, whereas CEMIP contains, as described below, a signal sequence, one G8 domain, two GG domains and 4 PbH1 repeats. While previous studies have indicated that the second GG domain is only present in CEMIP[40], a structural alignment of the C-terminal portion of TMEM2 (AA 1209-1355) and CEMIP (AA 1209-1353) underlines the possible presence of a second GG domain in TMEM2 (**Figure 11**). Indeed, recent findings reported the resolution of TMEM2 ectodomain structure by X-ray crystallography, confirming the presence of a second GG domain[54]. Despite their structural similarities, the current (but limited) knowledge about these two proteins suggests several differences in both function and regulation. TMEM2 is a type II transmembrane protein, which traffics to the cell membrane and catabolizes high-molecular weight HA into small fragments of ~ 5 kDa[40]. Its activity is calcium dependent and requires a pH optimum between 6 and 7[40]. In contrast to CEMIP, TMEM2-mediated HA degradation does not rely on the participation of live cells[40,55]. Studies in mice also showed that TMEM2 is more highly expressed at the transcriptional level compared to CEMIP, both during development and in adult tissues[40].

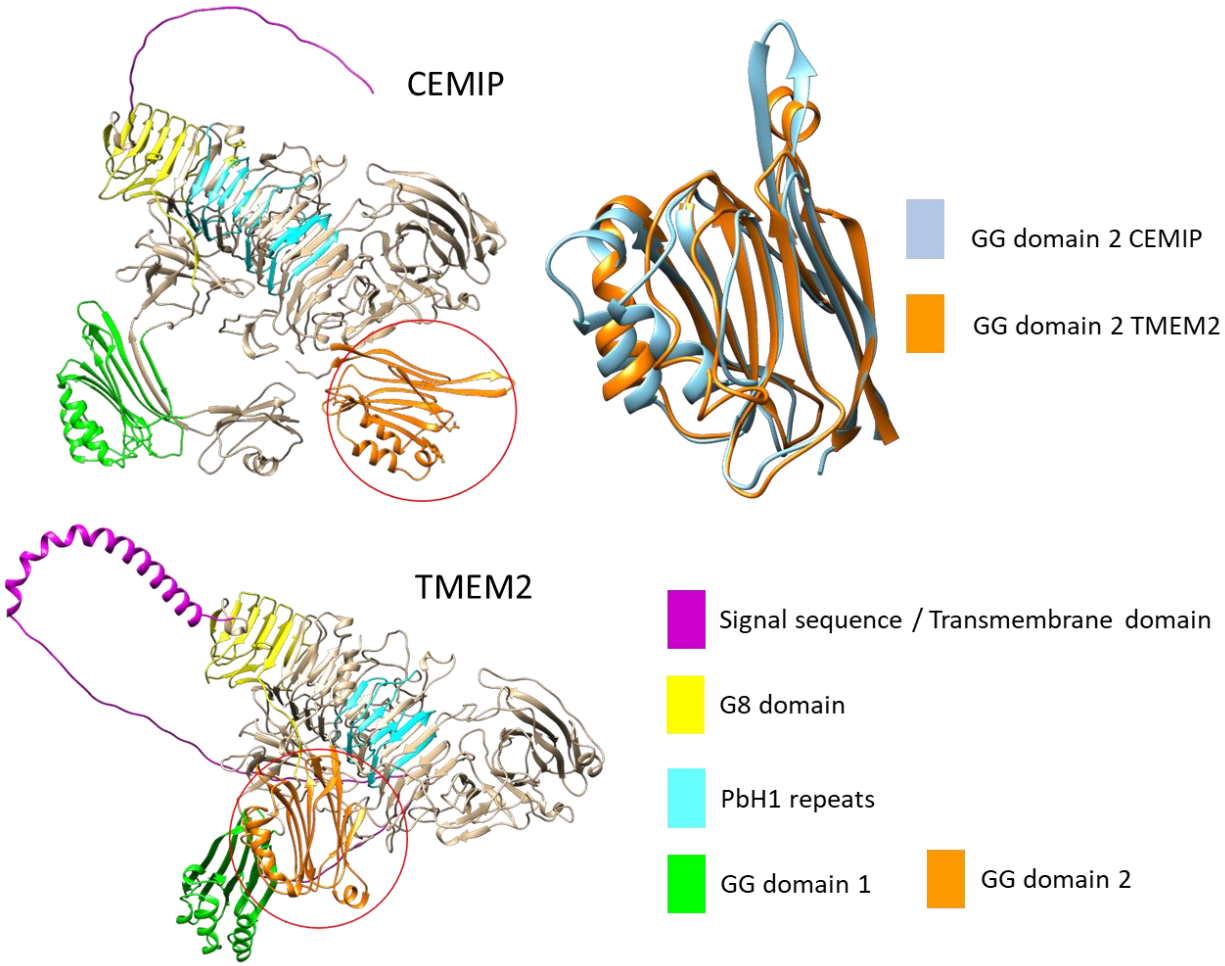


***Figure 11**: **Structure of CEMIP and TMEM2 with their domains** (left, top and bottom). Structural alignment of AA1209-1353 (CEMIP) and AA1209-1355 (TMEM2) reveals the presence of a second GG domain in TMEM2 (right, top). From AlphaFold, modified with USCF chimera.*

4.3.1.2 HA as a major component of the ECM

As CEMIP is considered a putative hyaluronidase, this section is intended to provide a foundational understanding of its substrate, hyaluronic acid (HA). HA is a major glycosaminoglycan component of the interstitial ECM. The interstitial ECM is a complex molecular structure that surrounds cells in tissues that provides structural support and regulates cell behavior in both physiological and pathological conditions[56]. The interstitial ECM is composed of two main classes of macromolecules: viscous (energy absorber) proteoglycans (PGs) such as HA, and elastic (energy storage) fibrous proteins, such as collagens, elastins, fibronectins and laminins[57]. These two classes of macromolecules confer upon the ECM its viscoelasticity and

architecture. Even though ECM is found in all types of tissues in adults, it is important to note that the composition differs depending on the tissue type. Viscous PGs fill the majority of the interstitial space within tissues in the form of a hydrated gel, for example in the brain and pancreas, whereas elastic fibrous components provide tensile strength, for example in muscle and bone[57]. HA belongs to a group of heteropolysaccharides named glycosaminoglycans (GAGs), which also includes chondroitin sulfate, dermatan sulfate, keratin sulfate, heparin sulfate and heparin[58]. HA is the most common GAG found in the tissues of humans and other vertebrates[59] and is comprised of tandem repeats of glucuronic acid and N-acetylglucosamine. It differs from other GAGs in that it is not sulfated and not synthesized by Golgi enzymes in association with proteins, but is rather produced at the inner face of the plasma membrane without any covalent bond to a protein core[58,60]. Nevertheless, HA can be covalently tethered to the serum protein inter-alpha-inhibitor by TSG6[61]. In addition, HA can reach a very high molecular weight (HMW, 10^8 Da), in contrast to other GAGs that are smaller in size ($<5 \times 10^4$ Da, usually $1.5–2 \times 10^4$ Da).

In mammals, HA is synthesized as a free linear polymer by three hyaluronan synthases named HAS1, HAS2 and HAS3. HAS are transmembrane glycosyltransferase isoenzymes whose catalytic sites are located on the inner face of the plasma membrane. Thus, growing HA chains are extruded onto the cell surface or into the ECM through the plasma membrane and HAS protein complexes[58]. HAS isoforms expression and activity are controlled differently by growth factors, cytokines and other proteins such as kinases, in a manner which appears to be cell and tissue specific. Similarly, the biochemical and synthetic properties of HAS isoforms are distinct:

- HAS1 is the least active isoenzyme and produces HMW hyaluronan from 2 × 105 to 2 × 106 Da.
- HAS2, which represents the main hyaluronan synthetic enzyme in normal adult cells, is more active than HAS1 and synthesizes HA chains greater than 2 × 106 Da. It regulates the developmental and reparation processes of tissue growth, and may be involved in inflammation, cancer, pulmonary fibrosis and keloid scarring62-66.
- HAS3 is the most active isoenzyme and produces HA molecules with a MW lower than 3 × 105 Da.

HA is degraded within hours of its synthesis[67]. Extra-large (1,000–10,000 kDa) native HA molecules within tissues are initially depolymerized into fragments of 10–100 kDa[37]. Most fragments are then released from the ECM, drained into lymphatic vessels, and catabolized within the lymph nodes. The remaining HA fragments reach the circulation and are fully degraded within the liver. HA degradation can be mediated by HYALs - enzymes that are classified as endoglycosidases[68,69] -, and through oxidative damage mediated by reactive oxygen species (ROS)[58]. With regards to the former, six HYAL gene sequences have been identified in the human genome: *HYAL-1, HYAL-2, HYAL-3, HYAL-4*, *PH20/SPAM1* and *HYAL-P1*[70]. However, because *HYAL-P1* is a pseudogene and both *HYAL-3* and *HYAL-4* have restricted expression patterns, these three genes are unlikely to have major roles in constitutive HA degradation in vivo[37]. HYAL-1, HYAL-2 and PH20/SPAM1 are the most characterized human HYAL genes[58]. HYAL-1 and HYAL-2, which are highly expressed in human somatic tissues, are acid-active with a pH optimum of 3.5 and below 4 respectively[69], and produce HA fragments of different sizes. HYAL-1 can degrade HA of any size into small hexa- and tetrasaccharides, while HYAL-2 processes HMW HA into 20 kDa oligomers in conjunction with the cell surface HA receptor CD44[69]. HYAL-1, which was found to regulate cell cycle progression and apoptosis, is the main HYAL expressed in cancers and therefore, it may regulate tumor growth and angiogenesis[67]. In contrast to HYAL-1 and HYAL-2, PH20/SPAM1 shows endoglycosidase activity at both an acidic and neutral pH and behaves as a multifunctional enzyme[58].

4.3.2 CEMIP: putative role in HA-mediated biological functions

In addition to its structural role, HA triggers different cell signaling cascades and modulates many physiological and pathological processes[56]. HA fragment size is one of the major determinants of its activity[71]. HA is rapidly depolymerized within tissues, from extra-large native molecules of 1,000–10,000 kDa to HA fragments which consist of low molecular weight HA (LMW HA, 10-100kDa) and much smaller HA oligosaccharides (8–30-dimer lengths). HMW-HA and LMW-HA can have antagonistic or inverse effects[72]. For example, HMW-HA can act as an anti-inflammatory molecules that promotes epithelial cell homeostasis and survival and accelerates wound healing[71], while LMW-HA has been shown to trigger inflammatory response[71], to promote activation of dendritic cells and monocyte maturation into macrophages[71] and to induce angiogenesis[73] and lymphangiogenesis[74].

In most cases, short oligosaccharides are produced by enzymatic cleavage mediated by hyaluronidases. However, it is interesting to note that LMW HA can also be directly synthesized by dysregulation of HAS enzymes, as well as by an impaired cellular metabolism that alters precursor availability[56]. Thus, it has been shown that phosphorylation of HAS2, the main hyaluronan synthetic enzyme which synthesizes HMW HA, leads to the inhibition of HA secretion[56]: since HA production by HAS is an energy consuming process, a low ATP:AMP ratio, as observed in an altered cellular metabolism, activates AMPK (AMP activated protein kinase) which mediates HAS2 phosphorylation and inhibits HA secretion[56].

CEMIP has been reported to degrade HMW HA into both intermediate-sized fragments of between 35 and 50 kDa (int-HA fragments) as well as LMW-HA[37,60,75]. CEMIP-produced int-HA fragments might conceivably have specific function(s) compared to other sizes of HA. Interestingly, while small HA oligosaccharides have been intensively investigated and have been shown to exert a number of biological effects not observed with HMW-HA[73,74], the biological activity of int-HA assembly largely remains to be investigated. Int-HA could potentially act as signaling molecules, for example in inflammation. It has been demonstrated that LMW and Int-HA fragments activate dendritic cells and macrophages via CD44 or TLR[76]. Int-HA fragments have also been reported to promote monocyte polarization, although with discordant results, with one study reporting polarization into M2-like macrophages[76], while the other into the M1 phenotype[77,78]. A possible explanation for these different *in vitro* results might be due to different experimental conditions. Int-HA fragments have also been demonstrated to induce wound closure, contrary to HMW-HA and LMW-HA[79], suggesting that int-HA fragments might induce a microenvironment that promotes tumorigenesis, metastasis and modulating fibrosis[80].

4.3.3 HA degradation by the CEMIP hyaluronidase activity

CEMIP mediates HA depolymerization[41,81] in a manner that involves clathrin-coated vesicles, internalization into early endosomes, and excretion of degraded HA molecules into the extracellular space[37]. Consistently, knockdown of clathrin heavy chain (CHC) and α-adaptin subunit of AP-2 impairs CEMIP hyaluronidase activity[37]. Additionally, deletion of the N-terminal signal peptide (amino acids 1-30) affects CEMIP trafficking and impairs its HA degradation activity[81]. Recent studies show that CEMIP-mediated HA depolymerization requires the

interaction with the Annexin-1 via its G8 domain, allowing CEMIP and Annexin-1 to colocalize at the cell membrane[82].

In an attempt to correlate structure and function, and based on the CEMIP model structure of **Figure 11**, it might be postulated that:

- the G8 domain is essential for HA catalytic activity and potentially interacts with multiple protein partners[82,83], some of them probably still to be discovered.
- the two GG domains might provide a key "hinge" function, comparable to the XYZ domain in protein gp35 in bacteriophages. This domain is essential for low molecular weight-HA binding (or protein catalytic function) as suggested by mutation in position 187, and potentially also relevant in CEMIP inactivation.
- the size of CEMIP-produced HA fragments might be dependent on the integrity of what seems to be a long tunnel formed by the four PbH1 domains.

4.3.4 The role of CEMIP in crescentic glomerulonephritis

The discovery of CEMIP as a functional effector of DDR1 activation upon collagen stimulation prompted Dr. Prunotto to explore the relevance of CEMIP in renal diseases. In order to achieve that, he proceeded establish various collaborations with the aim to a) assess CEMIP translational relevance using biopsies of patients affected by different kidney diseases, particularly focusing on GN, and b) assess the potential causal role of CEMIP in mouse model of GN and Alport syndrome. Firstly, collaborators at the Service de Pathologie Clinique, section néphro-pathologie (HUG, Switzerland) attested the presence of CEMIP in immunostainings of parietal epithelial cells (PECs) and in principal cells of the collecting duct in healthy kidneys. Interestingly, CEMIP was strongly upregulated in PECs in crescents in experimentally induced anti-GBM glomerulonephritis (not shown). Since PECs have been shown to be crucially involved in glomerulosclerosis and crescentic nephritis[2], these data are consistent with a potential role for the DDR1/CEMIP signaling pathway. Next, the expression of CEMIP in human GNs was investigated by Prof. Solange Moll (HUG, Switzerland) by immunohistochemistry using the generated anti-human CEMIP monoclonal antibody (described in section 7.2) to stain the human biopsy collection of the University Hospital of Geneva (Switzerland). Interestingly, in immunohistochemical analysis of patients' biopsies

CEMIP seems to be present during the initial inflammation of the glomerular crescent, but not during the fibrotic state, as shown in **Figure 12** in a case of Lupus nepgritis.

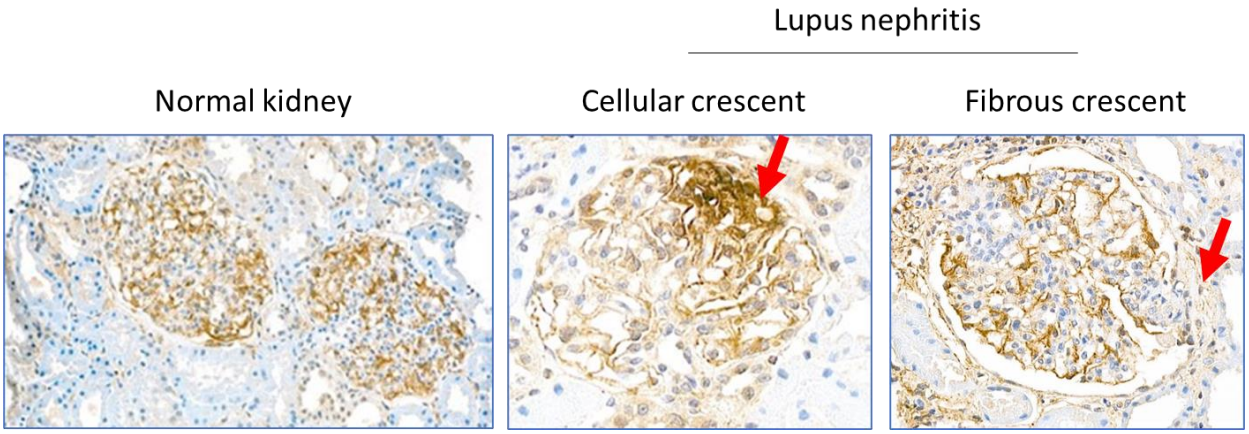


***Figure 12**: **CEMIP is involved in the initial inflammation step**. In the normal kidney, CEMIP is expressed in podocytes of the glomeruli (left). In lupus nephritis, CEMIP is present during the initial inflammation of the glomerular crescent, but not during the established fibrotic phase (red arrows; center and right). Courtesy of Thomas Cagarelli and Solange Moll.*

Once established the significance of CEMIP in translational experiments, collaborators proceeded to ascertain whether this holds true in an animal model of GN and AS. To validate the role of CEMIP in crescentic glomerulonephritis, floxed CEMIP mice obtained from Alain Chariot (University of Liege, Belgium) were used by Prof. Marcus Möller (Uniklinik RWTH Aachen, Germany) to conduct a pilot experiment. Mice were crossed with Pax8-Cre mice to induce CEMIP deletion specifically in the epithelial cells of the kidney including PECs (**Figure 13A**). As expected from reports of the relatively mild phenotype of spontaneous mutations of CEMIP in human patients of an isolated hearing loss[43], the offspring exhibited no overt spontaneous phenotype. In the anti-GBM model of crescentic glomerulonephritis, genetic deletion of CEMIP resulted in a reduced number of crescentic lesions (**Figure 13B**). These data are consistent with the observed role of DDR1 in PECs and more generally with the expression profile of CEMIP in glomeruli.

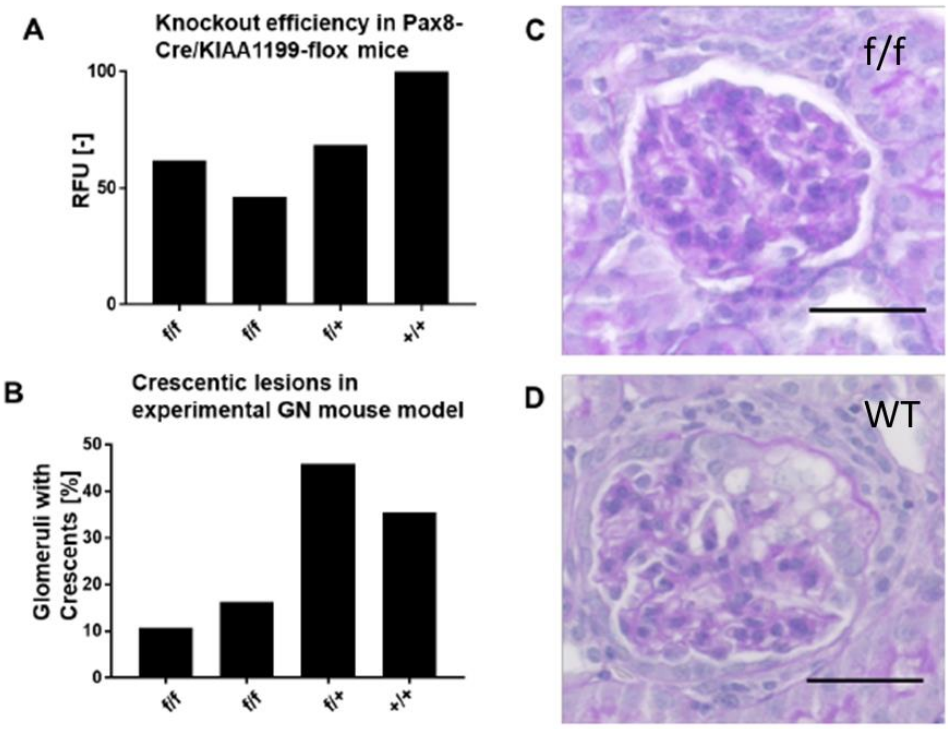


***Figure 13**: **CEMIP genetic ablation reduces crescent formation**. Preliminary evaluation of Pax8-Cre/CEMIP-floxed mice. A) Expression of CEMIP transcripts in Pax8-Cre/KIAA1199-floxed mice compared to controls. Note that CEMIP expression is not reduced to zero because the gene is inactivated only in kidney epithelial cells. B) The percentage of crescentic glomeruli is reduced in Pax8-Cre/CEMIP-floxed mice 8 days after induction of anti-GBM nephritis. C) and D) Representative images from Pax8-Cre/CEMIP-floxed and wild-type mice. f/f – homozygously floxed; f/+ heterozygously floxed; +/+ wild-type; RFU relative fluorescence units. Scale bar: 50 µm. Courtesy of Marcus Möller.*

To validate the role of CEMIP in Alport syndrome, collaborators at the Washington University in St. Louis (USA) generated CEMIP KO mice by CRISPR/Cas9-mediated non-homologous end-joining and verified CEMIP KO by RT-PCR. To examine whether knocking out CEMIP could ameliorate disease progression in *Col4a3*-/- Alport mice, they monitored kidney function and life span of CEMIP-deficient Alport mice. CEMIP KO mice at 6 and 10 weeks of age showed normal blood urea nitrogen (BUN) levels as compared to controls, an assay for glomerular filtration rate (**Figure 14A**). In contrast, Alport mice at 10 weeks of age showed an increased level of BUN as expected. Deficiency in CEMIP did not suppress the elevated level of BUN in Alport mice. Similarly, deficiency in CEMIP did not change the life span of Alport mice (**Figure 14B**; 76±3.8 days vs 78 ±3.4 days).

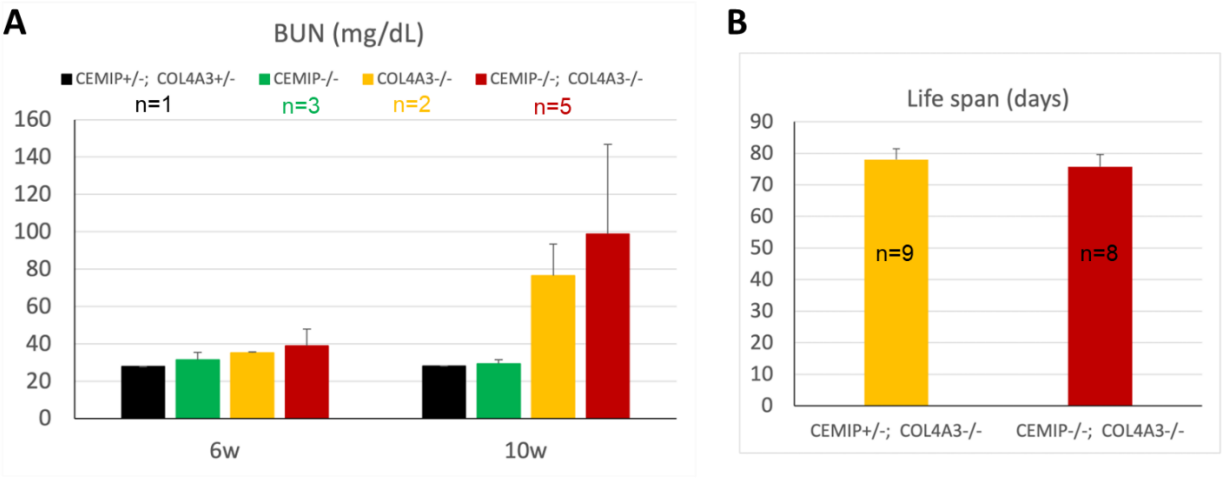


***Figure 14**: **BUN measurements and life span of CEMIP KO Col4a3-/- Alport mice**. A) GFR defect in CEMIP KO Alport mice (129 strain) at 10 weeks of age measured by BUN. B) Life span of CEMIP KO Alport mice (129 strain). Courtesy of Meei-Hua Lin and Jeff Miner.*

To examine whether knocking out CEMIP altered HA distribution in the kidney of *Col4a3*-/- Alport mice, they stained for HA on kidney sections at 10 weeks of age with biotinylated HA-binding protein (HABP) followed by avidin-Alexa594 incubation. The control kidney showed weak HA deposition in the papilla (**Figure 15A**). Knocking out CEMIP did not alter HA deposition (**Figure 15B**). In contrast, Alport kidney showed additional HA deposition in sclerotic glomeruli and interstitium (**Figure 15C**), compared to the control. However, knocking out CEMIP did not ameliorate the HA deposition in Alport kidney (**Figure 15D**).

In conclusion, these preliminary translational and animal studies suggest the following preliminary findings. Since the deletion of CEMIP in Alport mice did not improve the disease progression and did not modify the HA distribution, it appears that CEMIP is not significantly involved in the progression of this syndrome. Nonetheless, the above-mentioned findings suggest that the DDR1/CEMIP axis could have a role in renal pathologies characterized by a crescentic phase and that CEMIP could play a fundamental role in the initial inflammatory step of the formation of the crescent, but not in an advanced fibrotic stage.

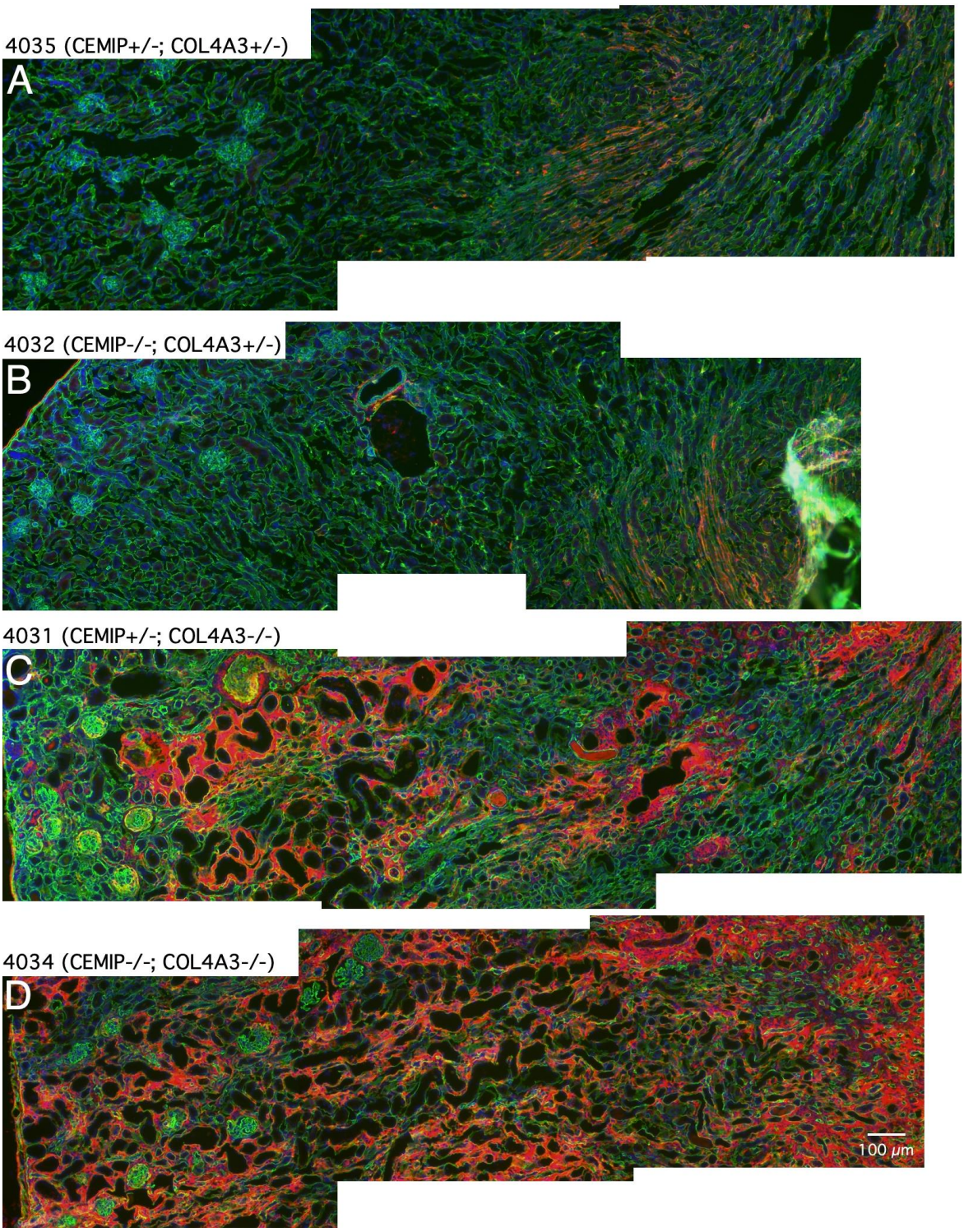


***Figure 15**: **Detection of HA and nidogen on kidney sections**. In red: hyaluronic acid; in green: nidogen. A) Control kidney. B) CEMIP KO kidney. C) Alport kidney. D) CEMIP KO Alport kidney. Courtesy of Meei-Hua Lin and Jeffrey Miner.*

4.3.5 The role of CEMIP in other diseases

CEMIP is also described in the literature in a broad range of different diseases, from nervous system diseases to inflammatory diseases and aging. In addition, CEMIP emerges as an important protein loosening the ECM and thus creating a permissive environment for EMT, cellular migration and cancer.

4.3.5.1 CEMIP in diseases and aging

CEMIP is involved in pathophysiological processes that regulate diseases of the nervous system, inflammatory diseases, as well as the aging/senescence process.

- Diseases of the nervous system

Hearing loss

CEMIP was originally described in families that are affected by non-syndromic hearing loss, which is associated with CEMIP mutations (R187C, R187H, H783Y)[47]. These authors showed that in mice, CEMIP was expressed specifically in Deiters' cells in the organ of Corti at postnatal day P0, before the onset of hearing, and that expression in those cells disappeared by day P7. In addition, CEMIP expression was observed in fibrocytes of the spiral ligament and the spiral limbus through to P21, when the murine cochlea matures. The *CEMIP* gene was therefore proposed to play a role in auditory development. These results were confirmed in subsequent studies[84-86].

Multiple sclerosis

In experimental autoimmune encephalomyelitis (EAE), an animal model for multiple sclerosis (MS) that is characterized by focal demyelinating lesions, CEMIP immunoreactivity was demonstrated to be exclusively associated with focal loci in damaged white columns of the spinal cord, and was mainly expressed by activated astrocytes that invaded damaged tissue[87]. Similar findings were observed in tissue from a patient with MS, suggesting that CEMIP expression by activated astrocytes could explain the focal HA degradation observed during MS progression and might represent a possible new therapeutic target[87].

- Inflammatory diseases

Rheumatoid arthritis and osteoarthritis

Pioneering work on the role of CEMIP was conducted in 2013 by the team of Yoshida *et al*. using osteoarthritis (OA) and rheumatoid arthritis (RA) synovial tissues. They found that transcription of CEMIP was higher in OA or RA synovium than in non-inflamed synovium[37]. These data suggested that CEMIP is a unique hyaladherin, with a key role in HA catabolism in the arthritic synovium. Relative to healthy controls, CEMIP expression was also found to be more strongly increased in synovial tissues with active RA compared to tissue with inactive disease, and is associated with CEMIP-dependent proliferation of fibroblast-like synoviocytes[88]. Angiogenesis is thought to be a key event in the formation and maintenance of the pannus in RA, and CEMIP was also implicated in promoting angiogenesis in vitro and in vivo in this study[88]. Subsequently, a positive correlation between CEMIP levels and the inflammatory markers TNF-α, IL-1β and IL-6 in both serum and synovial fluids of RA patients was reported, underscoring the clinical relevance of these findings, and treatment with an anti-CEMIP antibody partially alleviated arthritis severity and reduced serum LMW-HA levels and cytokine secretion in a collagen-induced arthritis (CIA) mouse model[82], providing proof of principle of the possible therapeutic relevance of these observations.

Mechanistically, CEMIP-KO mice exhibited resistance to CIA, which could be partially rescued by intra-articular injection of vectors encoding full-length CEMIP, whereas a CEMIP mutant with an inactive G8 domain had no effect[82]. Moreover, CEMIP expression was found to be suppressed by TGFβ-1 in normal synovial fibroblasts, but was only slightly decreased in synovial fibroblasts from OA or RA patients[89]. Coupled to this, TGFβ-1 upregulated the expression HA synthases (HAS1/2), which resulted in the increased accumulation of intermediate-sized HA fragments in cultures of synovial fibroblasts from OA or RA patients compared to normal synovial fibroblasts[89]. In addition, IL-6 significantly up-regulated CEMIP expression and HA-degrading activity in OA synovial fibroblasts[90]. Collectively, these data suggest a key role for CEMIP in the pathophysiology of RA.

In addition to synovial fibroblasts, CEMIP has been shown to be expressed by hypertrophic chondrocytes at the chondro-osseous junction. Thus, in CEMIP deficient mice, an accumulation of HMW-HA with reduced angiogenesis can be observed in this context, and is thought to be a key mechanism that modulates endochondral ossification during postnatal development[91]. In a mouse model of temporo-mandibular joint lesion, CEMIP expression was shown to be induced in injured mandibular condyles in association with increased HYAL2 expression[92]. In a study analyzing human cartilage, CEMIP transcription was significantly higher in OA cartilage compared to control cartilage, and CEMIP was highly expressed by chondrocytes in the HA-depleted area of OA cartilage[93]. Interestingly, CEMIP immunoreactivity correlates with the Mankin score, the histopathologic severity score used to evaluate OA lesions of the cartilage. OA chondrocytes exhibit HA-degrading activity, which is abolished by knock-down of CEMIP but not by downregulation of hyaluronidases HYAL1, HYAL2 or CD44[93]. Unlike RA synovial fibroblasts, only TNF-α but not histamine stimulated OA chondrocytes to overexpress CEMIP. Similar results were obtained by others[94]. In addition, these authors studied CEMIP expression *in vitro* using a chondrocyte dedifferentiation model. High-throughput RNA sequencing was performed on chondrocytes after CEMIP silencing. Most of the deregulated genes were involved in cartilage turnover, mesenchymal transition and fibrosis. Moreover, CEMIP was demonstrated to be essential for chondrocytes proliferation, promoted αSMA expression, and was co-expressed *in situ* with αSMA in all OA cartilage layers. These results strongly suggest a role for CEMIP in the trans-differentiation of chondrocytes into "chondro-myofibroblasts" which are found in OA cartilage but not in healthy cartilage.

Inflammatory bowel disease

Crohn's Disease (CD) is a chronic inflammatory disease of the gastrointestinal tract. Cultured CD fibroblasts produce increased levels of CEMIP protein through an IL-6-driven autocrine mechanism compared with control colon fibroblasts[75]. This leads to CEMIP deposition in the ECM and excessive degradation of HA, generating HA fragments that contribute to gut inflammation and fibrosis. Accordingly, antibody blockade of IL-6 receptors in CD fibroblasts decreased CEMIP protein levels in the ECM, and CEMIP silencing abrogated the ability of colon fibroblasts to degrade HA.

- The aging/senescence process

Skin

In the skin, the metabolism of HA is highly regulated. Aging leads to chronic low-grade inflammation, which is characterized by elevated levels of pro-inflammatory cytokines[95]. CEMIP is expressed in dermal fibroblasts[37] and in photoexposed skin[96,97], and is directly correlated with skin roughness in the papillary dermis[96]. Interestingly, mast cells were significantly enriched in the photoaged skin and were frequently associated with CEMIP-positive fibroblasts[97]. Furthermore, treatment of skin fibroblasts with a pro-inflammatory cytokine mixture (TNF-α, IL-1β and IL-6) or with IL1β alone suppressed HA depolymerization through downregulation of CEMIP expression, which was associated with permanently increased HA levels in the culture medium due to the upregulation of HAS2[95]. These results might suggest that cytokine-stimulated fibroblasts in unexposed skin increase the amounts of dermal HMW-HA in order to protect surrounding tissues against short-term inflammation. Notably IL-6 and IL-8 individually had no effect on HA metabolism in human skin fibroblasts, and their effect was only observed as part of a cocktail[95]. By contrast, and as mentioned above, IL-6 has been reported to promote HA depolymerization in synovial fibroblasts by inducing CEMIP expression[90]. This suggests that CEMIP might be differently regulated according to the cell- or tissue type.

Aortic valve disease

Aortic valve disease (AVD) is one of the leading causes of cardiovascular mortality. Aging is a significant clinical risk factor, and calcific AVD is the most common type of AVD occurring in 2% of the aged population[98]. During latent AVD, abnormal expression of HA and of its synthesizing/degrading enzymes has been observed[98]. Physiologically, the middle spongiosa layer of aortic valve consists primarily of proteoglycans and glycosaminoglycans, and provides lubrication and dampening functions when the valve leaflet flexes opens and closes. In normal human aortic valve tissue, CEMIP expression is present at all postnatal stages and increases with age, being mostly expressed inside the cells at the young stage and in ECM at advanced age. Increased CEMIP expression was demonstrated in calcified aortic valves, especially around calcified nodules in the fibrosa layer [99]. Finally, *in vitro* studies reported that CEMIP expression,

which was detected in porcine aortic valve interstitial cells, was modulated by stiffness, as less stiff substrates were observed to downregulate CEMIP expression[100].

4.3.5.2 CEMIP, cancer and EMT

The role of CEMIP in cancer development has been recently reviewed[101], so here I briefly summarize the insights that studies on cancer cells afford regarding CEMIP function and its involvement in patho-physiological mechanisms.

CEMIP has been extensively documented in colon cancer. Pioneering studies in 2007 by Giancarlo Marra's group at the University of Zurich searched for a distinguishing gene signature of precancerous adenomatous colorectal polyps compared to the normal colonic epithelium[102]. The authors showed that modulation of the Wnt pathway was a key feature of this transformation and that, amongst all the genes identified, transcription of CEMIP was most affected. Thus, the CEMIP expression pattern, normally confined to the lower portion of normal colonic epithelial crypts, was disrupted in dysplastic glands with widespread higher expression levels at both the mRNA and protein levels. Subsequent work demonstrated that this overexpression was also present in colorectal cancer, with a nuclear expression in the majority of UICC stage I-IV adenocarcinomas[42]. Other studies confirmed these findings, showing that increased CEMIP expression occurred in the phenotypic switch from normal to adenomatous or neoplastic lesions[103]. In addition, increased CEMIP mRNA expression was detectable in the plasma of patients with adenomatous or neoplastic lesions[103], which was associated with poor patient prognosis[104]. Recently CEMIP was found bridging MIB1 and GRAF1, promoting the degradation of GRAF1 thus promoting EMT and metastasis in colorectal cancer[105].

CEMIP expression in other types of cancer is consistent with the observations made in colon cancer. For example, CEMIP was detected with intermediate or high expression levels in human papillomavirus cervical (pre)neoplastic lesions, whereas only marginal levels were detected in normal exocervical epithelium[83]. This pattern was also observed in pancreatic cancer (PanCa) with an overexpression in benign pancreatic intraepithelial neoplasia, as well as in PanCa in both human and genetically engineered mouse models[106]. It is noteworthy that in all these cancer studies, CEMIP expression was always confined to the epithelial cells.

CEMIP has been shown to play a central role in EMT, a developmental morphogenic program in which cells lose their epithelial characteristics and gain a mesenchymal phenotype[107]. EMT is characterized by downregulation of epithelial and upregulation of mesenchymal markers, has been implicated in tumor initiation through endowing transformed cells with stemness characteristics, and has been more extensively investigated in the process in invasion and metastasis[108]. Several *in vitro* studies have reported CEMIP-induced EMT in different tumor cell lines. Thus, CEMIP silencing in non-small-cell lung cancer cell lines induced expression of E-cadherin (epithelial marker) and decreased expression of Vimentin, Snail, and Twist (mesenchymal markers)[109]. Conversely, CEMIP overexpression resulted in decreased E-cadherin levels and increased expression of Vimentin, Snail and Twist[109]. In gastric cancer cell lines CEMIP knockdown resulted in decreased mRNA expression of the EMT-related markers Slug, Snail, Vimentin and Twist[110]. Another study demonstrated an increase in E-cadherin protein levels and a decrease in Slug and Vimentin levels in BCPAP cells line after CEMIP silencing[111].

Several recent studies analyzed the relationship between CEMIP and EMT markers in human tumor tissues. For example, in 2018, Jiang and colleagues studied HCC samples and found a positive correlation between CEMIP expression and N-cadherin and Vimentin expression, and a negative one with E-cadherin expression[112]. In colon cancer samples, an inverse relationship between CEMIP localization and E-cadherin expression[113] exists, while in cholangiocarcinoma tissue samples a negative relationship of CEMIP with E-cadherin and a positive one with N-cadherin and Vimentin[114] has been found. Other evidence of CEMIP involvement in EMT includes the fact that the Epithelial Splicing Regulatory Protein 1 (ESRP1), an epithelial cell-specific regulator, has been identified as an upstream regulator of CEMIP[115].

Three main pathways have been described in the CEMIP-dependent EMT process: PI3K-Akt-mediated, EGFR-mediated and Wnt/β-catenin-mediated EMT.

- **PI3K-Akt-mediated EMT**: TGF-β is a potent inducer of EMT[116], and promotes cell survival through the PI3K-Akt pathway in cancer[117]. Tang *et al.* showed that CEMIP silencing in lung cancer cell lines decreased the levels of TGF-β, PI3K and Akt[109]. On the other hand, CEMIP overexpression increased TGF-β, PI3K and Akt levels[109]. Similar findings have been reported

in cholangiocarcinoma cell lines, where a significant downregulation of TGF-β as well as PI3K, AKT and mTOR was observed after CEMIP silencing, whereas an increase of these proteins was observed in CEMIP overexpressing cells[114].

- **EGFR-mediated EMT:** Epidermal growth factor (EGF) promotes cell growth and differentiation by binding to the Epidermal growth factor receptor (EGFR)[118]. It has been reported that CEMIP activates EGFR signaling and regulates the downstream kinases in the EGF-mediated EMT pathway[119]. Tang *et al*. observed a correlation between the expression of CEMIP and the expression of EFGR in lung cancer cell lines[109]. In HCC cells resistant to the kinase inhibitor drug Sorafenib, high levels of CEMIP triggered EGF-induced EMT, increasing the migratory and invasive ability of these cells[120]. CEMIP upregulation in HCC parental cells induces phosphorylation of EGFR and its downstream kinases, leading to EMT, while CEMIP silencing reduces EGFR expression[120]. All in all, these data suggest that CEMIP acts as an upstream regulator of EGFR signaling.

- **Wnt/β-catenin-mediated EMT**: Several studies suggest an important role of Wnt/β-catenin signaling during EMT[121,122]. In 2019, Deroyer and colleagues demonstrated that CEMIP modulates the Wnt/β-catenin signaling in human chondrocytes[94]. In gastric carcinoma cells, CEMIP knockdown reduces the expression of β-catenin, but also of c-myc and cyclinD1, two downstream players of the Wnt/β-catenin pathway[110]. These findings suggest a role for CEMIP in promoting invasion and metastasis through the Wnt/β-catenin signaling pathway[110]. Indeed, the Wnt/β-catenin signaling pathway coordinates pivotal cellular processes such as cell differentiation, proliferation, migration, and many others. A study published in 2007 by Sabates-Bellver *et al*. examined the transcriptomes of colorectal polyp samples and found a positive correlation between CEMIP upregulation and the expression of Wnt targets[102]. Moreover, northern blotting analysis confirmed an important decrease in CEMIP mRNA levels linked to Wnt pathway inhibition. In 2011, Birkenkamp-Demtroder and colleagues demonstrated that lentiviral-mediated knockdown of CEMIP in colon cell lines alters the cell cycle and the Wnt-signaling pathway[42]. Ingenuity Pathways Analysis (IPA) together with immunofluorescence analyses and western blotting showed a change in the expression (mostly downregulation) of 67 genes involved in Wnt/β-catenin signaling upon CEMIP knockdown.

In view of these observations, it is interesting to speculate about a potential role of CEMIP and the HA fragments it produces in the activation of the Wnt/β-catenin signaling pathway, which could conceivably coherently connect CEMIP expression to cell proliferation and migration initiated by the binding of Wnt followed by β-catenin dephosphorylation.

Taken together, these observations suggest an involvement of CEMIP in EMT regulation. The publications that support this notion are summarized in **Table 1**.

Table 1*: Summary of EMT markers involved in CEMIP silencing and upregulation.*

	Epithelial phenotype	**Mesenchymal phenotype**	**Cell line/Tissue**	**Reference**
CEMIP silencing	↑ E-cadherin	↓ Vimentin ↓ Snail ↓ Twist	NSCLC cell line	Tang *et al*., 2018
		↓ Vimentin ↓ Snail ↓ Slug ↓ Twist	GC cell line	Jia *et al*., 2017
	↑ E-cadherin	↓ Vimentin ↓ Slug	PTC cell lines	Jiao *et al*., 2019
CEMIP upregulation	↓ E-cadherin	↑ Vimentin ↑ Snail ↑ Twist	NSCLC cell line	Tang *et al*., 2018
	↓ E-cadherin	↑ N-cadherin ↑ Vimentin	HCC tissues	Jiang *et al*., 2018
	↓ E-cadherin	↑ N-cadherin ↑ Vimentin	CCA tissues	Zhai *et al*., 2020
EMT Markers	↑ E-cadherin	↑ N-cadherin		

	↑ Cytokeratin ↑ Laminin-1 ↑ ZO-1 ↑ Occludin	↑ Vimentin ↑ Snail ↑ Slug ↑ Twist	

In addition to cancer-relevant functions for CEMIP in tumor cells, several observations suggest that CEMIP might also regulate the tumor microenvironment. For example, as a major component of ECM, HA can act as a modulator of the tumor microenvironment, and thereby regulate tumor growth, angiogenesis, invasion, and metastasis[123]. Given that CEMIP functions as a hyaluronidase, its involvement in cancer could conceivably be directly linked to its hyaluronidase activity in this context. Furthermore, CEMIP has been identified as a candidate gene for skin tropism of cancers[124]. Although these data are anecdotal because the study only employed a single patient derived cell line, it is interesting to note that they suggest that CEMIP might create a permissive ECM environment that foster tumor engraftment in the skin. Moreover, CEMIP-containing exosomes have been reported to educate the brain microenvironment and thereby foster cranial metastases[125]. Possibly consistent with these findings, the growth of glioblastoma cells implanted into the mouse brain was suppressed in CEMIP KO mice compared to the situation in wild-type mice, which was associated with decreased macrophage infiltration into the brain tumors[126], again implicating CEMIP in the creation of a pro-tumor microenvironment.

4.3.6 Regulation of CEMIP protein expression

CEMIP protein levels are regulated by cytokines, transcription factors, the Wnt/β-catenin signaling pathway, microRNAs (miRNAs), histone methylation and hypoxia (**Figure 16**). However, it should be noted that the results published by different studies are often conflicting.

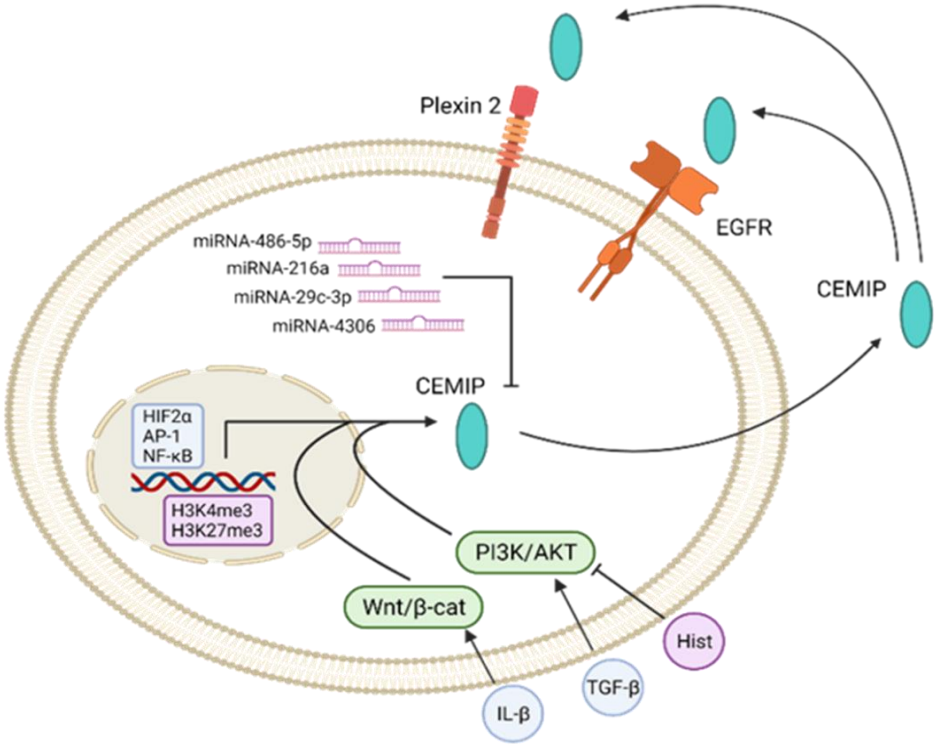


***Figure 16**: **Schematic representation of regulators of CEMIP protein expression**. Created with BioRender.*

Several cytokines have been found to play a major role in the regulation of CEMIP. In 2017, Khoi *et al.* suggested that the pro-inflammatory interleukin 1β (IL-1β) increased CEMIP transcription and migration of pancreatic ductal adenocarcinoma cells[127]. De la Motte's group investigated the effect of interleukin 6 (IL-6) in Crohn's disease fibroblasts. The authors observed low levels of CEMIP deposition in the extracellular matrix after treatment with IL-6[75]. However, contrasting evidence has been reported in a recent study by Sato *et al*. who demonstrated that treatment with a mixture of pro-inflammatory cytokines (TNF-α, IL-1β, and IL-6) in human skin fibroblasts decreased CEMIP mRNA levels and protein expression, an effect primarily mediated by IL- 1β[95].

TGF-β, a pro-fibrotic protein and inducer of cell proliferation, can regulate CEMIP expression. Deroyer *et al*. identified TGF-β as a CEMIP upregulator through the Alk5/PAI-1 pathway in dedifferentiated chondrocytes[94] and therefore as a pro-fibrotic mediator. However, another TGF-β-initiated pathway was found by Shintaro Inoue's group, in this case leading to an opposite effect in CEMIP regulation. This group investigated the TGF-β1-mediated pathway in Detroit 551 skin

fibroblasts and demonstrated that TGF-β1 downstream of PI3K-Akt signaling pathway was implicated in CEMIP downregulation after TGF-β1 stimulation[89]. Consistently, collaborators have also recently reported that TGF-β1 downregulates CEMIP expression in fibroblasts[128]. The role of TGF-β in the epithelial-mesenchymal transition (EMT) will be discussed in the next section.

The coordinate roles of TMEM2 and CEMIP expression has been analysed in human skin fibroblasts. Thus, CEMIP and TMEM2 expression together with HA depolymerization appeared to be tightly regulated by the action of TGF-β1 and histamine[129]: while TGF-β1 enhanced TMEM2 levels and decreased both CEMIP and HA processing, histamine was demonstrated to act in the opposite direction[129]. In the same cells, knockdown of CEMIP, but not of TMEM2, heavily impaired HA depolymerization[129]. In addition, pro-inflammatory cytokines were shown to significantly regulate both protein levels[95]: indeed, treatment of human skin fibroblasts with IL-β1 increased TMEM2 expression, but suppressed both CEMIP and HA depolymerization[95]. Even though CEMIP and TMEM2 are both involved in HA processing, these data clearly indicate that they act separately in different contexts and that their activity is tightly regulated at the transcriptional level through the intervention of several macromolecules.

A number of transcription factors have been shown to control CEMIP transcription. In 2014, Shostak *et al.* identified CEMIP as an oncogene whose expression is controlled by nuclear factor-kappa B (NF-kB), which protects tumor cells from apoptosis and promotes invasiveness and the epithelial-mesenchymal transition through the EGFR pathway[83]. Kuscu *et al.* reported that distal NF-kB sites and the transcription factor AP-1 are required for CEMIP gene expression in fibroblast-like monkey cell lines[130].

Several microRNAs have been shown to negatively regulate CEMIP expression. In 2019, Jiao *et al.*[111] showed that miR-486-5p directly targets the 3'-UTR region of CEMIP, and exerts a significant inhibition of CEMIP mRNA and protein expression in papillary thyroid cancer cells using a luciferase reporter assay. Wang *et al.*[119] reported that the same miR-486-5p targets CEMIP and that its overexpression reduces the proliferation and migration of non-small cell lung cancer cells via the EGFR pathway. Overexpression of miR-29c-3p, another miRNA that targets CEMIP, has been shown to decrease gastric cancer cells migration *in vitro* and *in vivo* via the Wnt/β-catenin

and EGFR signaling pathways[131]. Similarly, overexpression of miR-216a can reduce migration and invasion of colorectal cancer cells *in vitro* and inhibit metastasis *in vivo*[132]. miR-4306 was demonstrated to regulate CEMIP expression in osteosarcoma, and LINC00958, a long non coding RNA (lncRNA), was shown to promote tumour progression and metastasis by inhibiting miR-4306 expression, leading to increased CEMIP expression[133].

Histone methylation can positively or negatively modulate gene transcription. Increased tri-methylation of lysine 27 on histone H3 (H3K27me3) has been associated with inactivation of CEMIP, decreased tumour cell growth and reduced migration in triple-negative breast cancer[134]. In another study, a correlation between H3K4me3 levels and CEMIP expression was demonstrated in human colon cancer cells[113]. Hypoxic conditions (see below) inhibit the histone demethylase Jarid1A, causing an increase of H3K4me3 within the CEMIP promoter, which leads to CEMIP upregulation[113].

Hypoxic conditions present in the tumour microenvironment modulate CEMIP expression. In 2015, Evensen *et al.*, observed that hypoxia inducible factor 2α (HIF-2α) induces CEMIP expression by binding to the hypoxia response element (HRE) present in the CEMIP promoter region[113]. Moreover, Wang *et al.*[114] reported that in HCC patient samples, co-expression of CEMIP and HIF-1α correlated with a poor prognosis and overall survival rate[114].

4.3.7 CEMIP expression: tissue and cell localization

CEMIP is expressed in a wide range of tissues, with the highest levels in brain, skin, placenta, lung, testis and ovary, but its expression is notably absent in the liver, kidney, and spleen[135]. CEMIP is widely expressed in the brain, in contrast to the HYAL1 and HYAL2 enzymes, which are not expressed in this organ[136]. Therefore, it is likely that CEMIP may have an important role in HA catabolism in tissues that do not express HYAL1 or HYAL2, such as the brain. Indeed, it is noteworthy that ECM molecules fill the extracellular space, which occupies up to 20% of the adult brain volume[137]. HA is important components of the brain ECM in both condensed ECM (such as perineuronal nets) and diffuse ECM (which fills perisynaptic spaces), and plays crucial roles in neuronal development, plasticity, and pathophysiology[137]. Thus, HA plays a major structural role in ECM throughout all stages of brain maturation, specifically supporting HA-binding proteins

and proteoglycans as a central filament of aggregates[138]. Moreover, an interesting study by Wilson *et al.* highlighted the importance of HA in regulating the formation and function of synapses in a model of human cortical brain development[139]. Numerous studies showed that, from the beginning of brain development, HA critically regulates neural circuit formation. At the onset of neurulation, HA is necessary for neural crest cell migration. Later in cortical neurogenesis, HA regulates neural progenitor cell proliferation and promotes neuronal differentiation, migration and the formation of cortical layers. In addition, HA surrounds developing excitatory synapses, where it critically regulates synapse formation and the resulting balance between excitatory to inhibitory signaling[139]. Accordingly, HA removal was demonstrated to be sufficient to drive a hyperexcitable state which is characteristic of neurodevelopmental disorders, including epilepsy, intellectual disability, and autism spectrum disorders. Conversely, the observation that HA decreases excitatory synapse formation could have implications for aging and Alzheimer's Disease, both of which exhibit HA accumulation and synapse loss[139].

CEMIP was first shown to be widely expressed in the brain by the pioneering work of Izumi Horokawa's team at NIH on induced cell mortality[135]. Yoshino *et al*., subsequently showed that CEMIP is expressed in the hippocampus and cerebellum in wild-type mice, and demonstrated decreased mnemonic ability in novel object recognition in CEMIP KO mice[140]. In particular, high levels of CEMIP transcripts have been detected in the granular cell layer of the murine hippocampus and cerebellum[140], and also during the immature period of the organ of Corti in mice, in fibrocytes of the spiral ligaments and spiral limbus, and in Deiters' cells[47]. Finally, in another study, Yoshino *et al*., demonstrated that the dendritic spine density was significantly decreased in the dentate gyrus granule cells in CEMIP KO mice, suggesting that CEMIP-mediated HA degradation may be critical for the synaptic formation process by contributing to cognitive functions, such as learning and memory in the mouse brain[141]. In future work it will be interesting to investigate further the role of CEMIP in the homeostatic regulation of the HA content of the brain, and the impact this has on brain development, function and disease.

In addition to the brain, CEMIP has been frequently detected in fibroblasts. Yoshida *et al.* reported CEMIP expression in dermal fibroblasts[37], while other groups have detected CEMIP in synovial[90] and colon fibroblasts[75]. CEMIP is also expressed in hypertrophic chondrocytes at the chondro-

osseous junction[91]. Collaborators have recently reported the expression of CEMIP in mesenchymal stromal cells, and have shown that CEMIP regulates the differentiation of these cells into the osteogenic and adipogenic lineages[128].

Within the cell, CEMIP has been found to localize in different cellular sub-compartments, a fact that could explain its involvement in regulating different cellular pathways and diseases. Cytoplasmic CEMIP has been identified in different cells of the human cochlea[84] and in several tumor cells such as gastric cancer[142] and hepatocellular carcinoma (HCC)[143], where its knockdown triggers ER stress-mediated apoptosis through upregulation of CHOP and ATF4[143]. CEMIP heterogeneous localization was observed in colorectal cancer cells, with nuclear[42], cytoplasmic[42,102] and cell membrane [102,144] localizations. Despite this heterogenous localization, several lines of evidence suggest a common role in colorectal cancers for CEMIP as modulator of Wnt-β catenin signaling[42,102,144] (see below, CEMIP and EMT). CEMIP has also been found in the ER[81,145] and Golgi[81] compartments, which is prototypical for a secretory protein[81] and consistent with the existence of an N-terminal signal peptide. Interestingly, this localization pattern has been classically associated with the HA depolymerizing function of CEMIP[81]. Recent evidence also links cell membrane localization of CEMIP to its HA degradation activity[82], as CEMIP can localize to the cell membrane by binding to ANXA1 in rheumatoid arthritis fibroblast-like synoviocytes, which facilitates exogenous HA depolymerization[82].

4.4 CEMIP biology remains elusive

In this chapter I described the discovery of CEMIP as a downstream effector of DDR1 receptor. I also showed the key role of the DDR1/CEMIP axis in crescentic GN. CEMIP relevance in crescentic GN was assessed using patient biopsies and its causal role in GN explored using preclinical mouse model of GNs. I illustrated CEMIP involvement in other diseases and aging summarizing the main scientific evidence reported in literature related to its structure, role and mechanism.

In addition to a structural function, CEMIP exerts a unique enzymatic activity, producing, alone or in coordination with other proteins, a range of int-HA and LMW-HA fragments. Despite all the scientific observations reported here, CEMIP biology remains enigmatic. The current

comprehension of CEMIP's biology is primarily informed by the extensive volume of research focused on its involvement in cancer-related processes, such as tumor growth and invasion. Additional loss and gain of function studies are required to investigate the functional relationship between the HA fragments produced by CEMIP, and the physiological and pathophysiological processes that are regulated by CEMIP. Moreover, the role of CEMIP, and the HA fragments it produces might exert quite different effects in different cellular settings[93] and disease contexts[82,106,124,126]. The literature provides compelling evidence that CEMIP plays a role in modulating fibrosis (for example through the modulation of CEMIP activity by TGF-β1 and histamine), an observation that can guide future mechanistic studies in simple cellular systems[146]. The precise study of CEMIP in kidney disease is beyond the scope of this thesis. Here, I focused my study on understanding CEMIP *in vitro*, in biochemical assay of HA degradation and HA binding, revealing its activity, its 3D structure and its inhibition. Once additional structural and functional knowledge has been generated, this can be leveraged to identify compounds capable of modulating the CEMIP protein function *in vivo*. Indeed, collaborators have recently reported that highly sulfated HA is a potent inhibitor of the CEMIP hyaluronidase activity[128]. These and further next-generation compounds are expected to have a strong potential for therapeutic application, for example through modulating tissue inflammation, an unmet clinical need for a range of inflammatory disease conditions.

5 AIM OF THE PROJECT

The present work aimed to investigate the DDR1/CEMIP/HA axis to identify new therapeutic strategies in the context of glomerular diseases. In the specific, I was interested in investigating CEMIP's role in and outside the cell, its activity and structure, and studying the possible inhibition of CEMIP via small molecules and antibodies.

To elucidate CEMIP domains and amino acids' function, different CEMIP mutants were created by site-directed mutagenesis (following evidence provided by computational analysis), and tested, in order to characterize the portion of CEMIP with catalytic activity and the one with HA-binding ability. Since CEMIP mode of action is yet to be understood, a series of experiments elucidating the activity of cytosolic and secreted CEMIP was conducted. Another part of the project was dedicated to finding potential hit inhibitors among several small molecules and antibodies, by testing them in CEMIP-expressing cell lines. Finally, CEMIP's tridimensional structure was assessed by cryo-EM, and the GG domain's structure by crystallography, with the hope of improving the precision of the small molecules' design.

The end goal of this project was to define if CEMIP could be a possible therapeutic target to develop a treatment for GN by providing renal protection for patients.

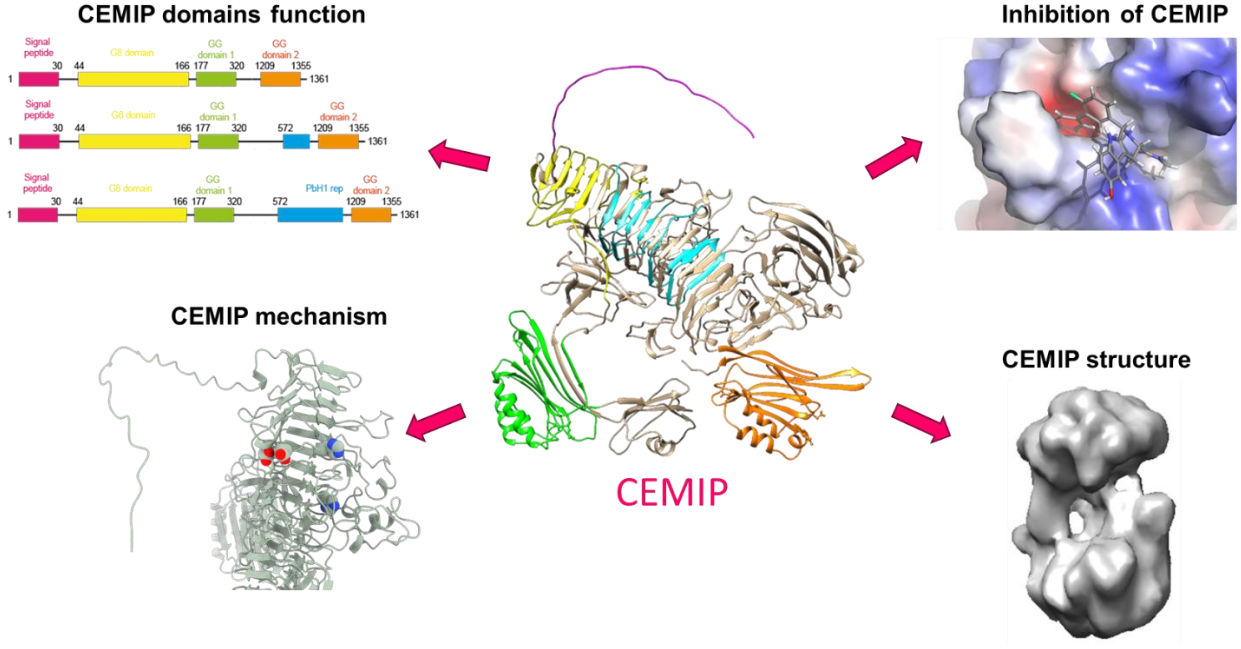


***Figure 17**: **Overview of the project**. Note that CEMIP structure is the prediction model from AlphaFold[39].*

6 CHAPTER I – ACTIVITY AND MECHANISM OF CEMIP

In this chapter, I established two main methods, HA assay and HA pulldown, as a readout of CEMIP enzymatic activity and CEMIP HA binding, respectively. Hence, I evaluated if CEMIP binds to HA of different molecular weights, and if or in which condition its secreted form is active. Furthermore, I tested a wide range of CEMIP mutants, from single amino acid mutations to entire domains' deletion, with the view to understanding CEMIP domains' role and involvement in the binding and processing of hyaluronic acid.

6.1 HA assay gel as a readout of CEMIP hyaluronidase activity

Frist, it was important to establish a readout for CEMIP activity in terms of HA depolymerization ability. For this, a technique called HA assay was learnt from Anja Schmaus (KIT, Germany) at the beginning of the project. This assay consists in culturing the cells together with HMW-HA of 1.5 MDa and assessing the hyaluronidase activity of CEMIP by measuring the degradation of the HMW-HA using agarose gel electrophoresis assay. Cells are incubated with high molecular weight HA (1.5 MDa) for 48 to 72 hours. The supernatant is then harvested, the HA is precipitated with ethanol and purified, loaded on an agarose gel and revealed with a stains-all staining (**Figure 18**).

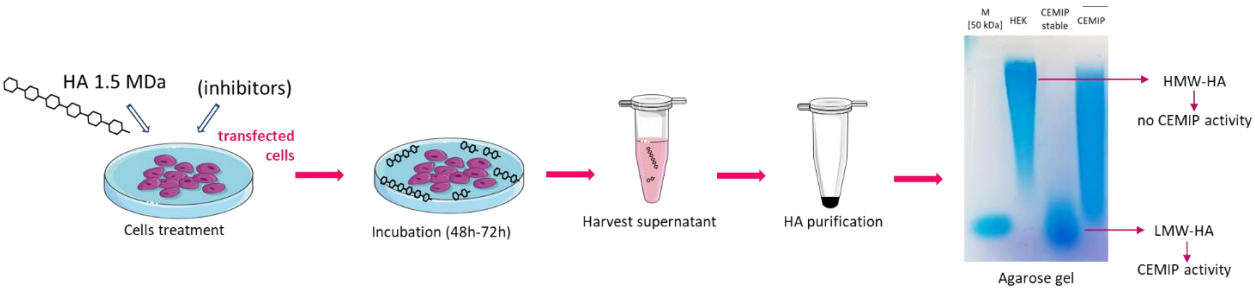


***Figure 18**: **HA assay workflow**. The HA assay is performed by culturing the cells together with HMW-HA of 1.5 MDa and assessing the hyaluronidase activity of CEMIP by measuring the degradation of the HMW-HA using agarose gel electrophoresis assay. Cells are incubated with high molecular weight HA (1.5 MDa), with or without inhibitors, for 48 to 72 hours. The supernatant is then harvested, the HA is precipitated with ethanol and purified, loaded on an agarose gel and revealed with a stains-all staining. The gel on the right shows a marker of 50 kDa HA in the first lane, the HA profile of HEK293 in the second lane, the HA profile of stably expressing CEMIP cells in the third lane, and the HA profile of CEMIP transient transfection.*

After assessing the HA on the gel without cells involved (**Figure 19**, left), some tests were performed to assess the reproducibility, the cell seeding variable, the effect of DMSO and the HA-degrading activity of the hCEMIP plasmid (**Figure 19**, right). The effect of the cell seeding at the beginning of the assay plays a crucial role because without enough cells present in the well, the readout is compromised (**Figure 19**, lane 3) and could be mistaken for a false negative. The assay relies therefore on correct cell seeding and proliferation. The cell seeding at time zero was set at 80'000 cells/well. Moreover, HA processing was unchanged when cells were cultured with 0.1% and 0.5% of DMSO (**Figure 19**, right).

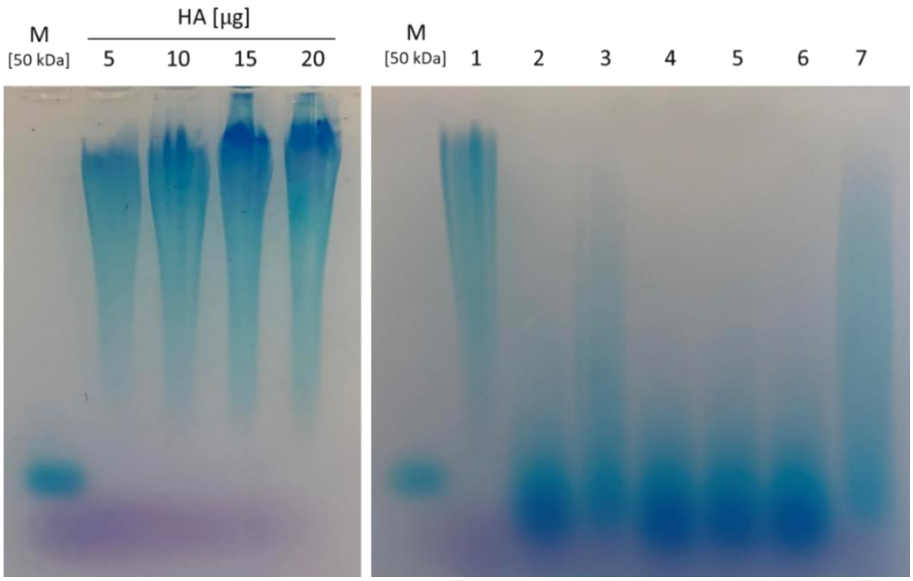


***Figure 19**: **Preliminary tests of HA assay**. Left: HA directly on the gel (no cells involved). Amounts of HA from 5 to 20 µg were loaded on the gel. Right: 1 = control vector, 80'000 cells/well; 2 = mCEMIP vector, 80'000 cells/well; 3 = mCEMIP vector, 40'000 cells/well; 4 = mCEMIP vector, 160'000 cells/well; 5 = mCEMIP vector, 80'000 cells/well, 0.1% DMSO; 6 = mCEMIP vector, 80'000 cells/well, 0.5% DMSO; 7 = hCEMIP transient transfection, 80'000 cells/well.*

The limit of detection of the HA on the gel was assessed between 50 ng and 5 ng, as shown in **Figure 20**.

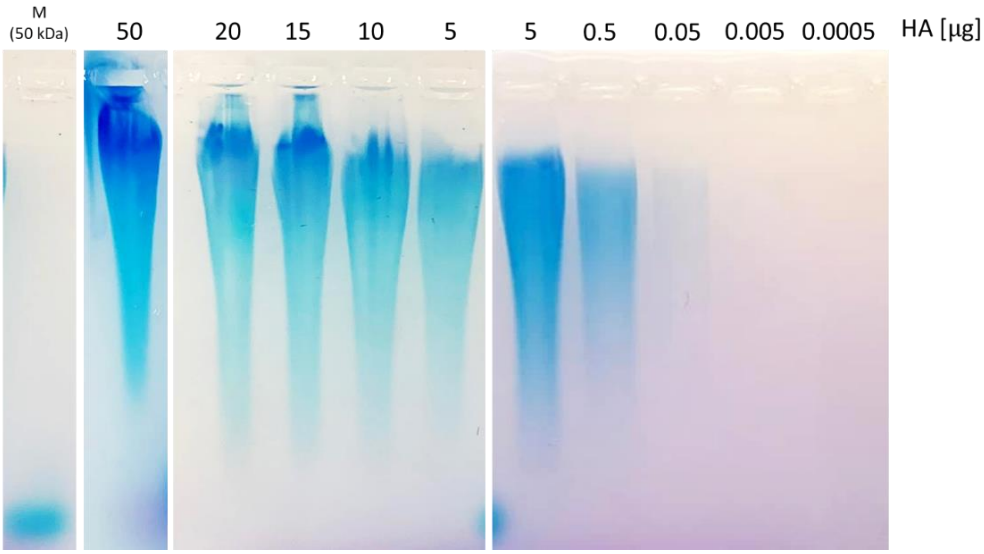


Figure 20: Limit of detection of the HA assay gel. *Amounts of HA from 50 to 0.0005 µg were loaded on the gel. The minimum amount of HA detectable is between 50 ng and 5 ng.*

6.2 HA pulldown as a readout of the HA-CEMIP binding

Second, in order to perform experiments in which CEMIP binding is tested, it was important to develop a pulldown protocol. After unsuccessful attempts using magnetic streptavidin-coated beads (Dynabeads, Thermofisher), a protocol involving the precipitation of HA by cetylpyridinium chloride (CPC) was developed from previous methods[147,148].

6.2.1 Dynabeads pulldown

The workflow of the pulldown using Dynabeads provides an incubation of biotinylated HA with proteins, followed by a second incubation with the streptavidin-coated magnetic beads. With the help of a magnet, the mixture is washed several times with SDS and/or PBS before adding Laemmli DTT and running a western blot (**Figure 21**).

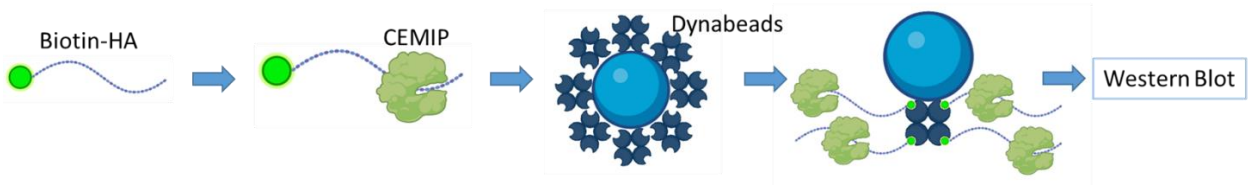


***Figure 21**: **Dynabeads pulldown workflow**. Incubation of biotinylated HA with proteins, followed by a second incubation with the streptavidin-coated magnetic beads. With the help of a magnet, the mixture is washed several times with SDS and/or PBS before adding Blue Laemmli with DTT and running a western blot. Courtesy of Filippo Donati.*

Despite changing multiple variables such as increasing biotin-HA, modifying washing steps, time of incubation and using different samples, the assay was found non-reproducible. Most of the time, CEMIP was lost during the washing steps. Other times, CEMIP was revealed also in the last step. **Figure 22** represents just a small part of the number of inconsistent results obtained.

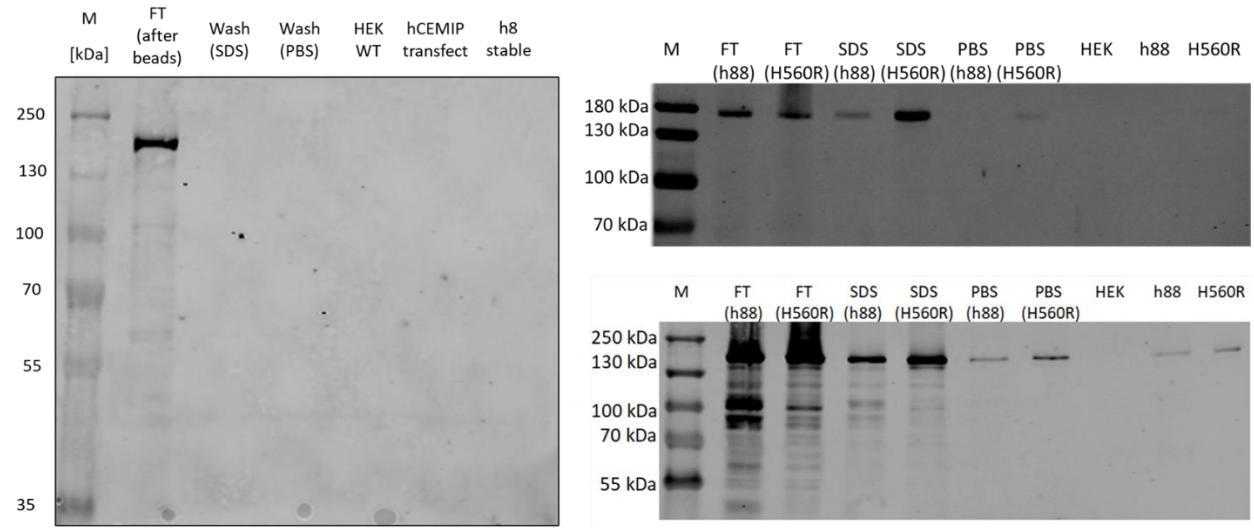


***Figure 22**: **Pulldown with the Dynabeads**. Left: CEMIP is completely lost in the flow-through (FT); top right: CEMIP is lost during the washing steps; bottom right: CEMIP is found in the last step. h8 = hCEMIP-HEK293T stable clone, first selection; h88 = hCEMIP-HEK293T stable clone, second selection; H560R = CEMIP mutant without hyaluronidase function.*

6.2.2 CPC pulldown

After assessing the experimental limitation reached with the Dynabeads magnetic separation technology, an adaptation from the CPC pulldown protocol from Lee[147] and Sleeman[148] was developed. This protocol consists of incubating the extracted proteins with HA and then precipitating the mixture with cetylpyridinium chloride (CPC). After centrifugation and washing steps, the precipitated HA together with its binding partner (CEMIP in this case) is dissolved in Laemmli blue with DTT and run on a Western blot with an anti-CEMIP antibody (**Figure 23**).

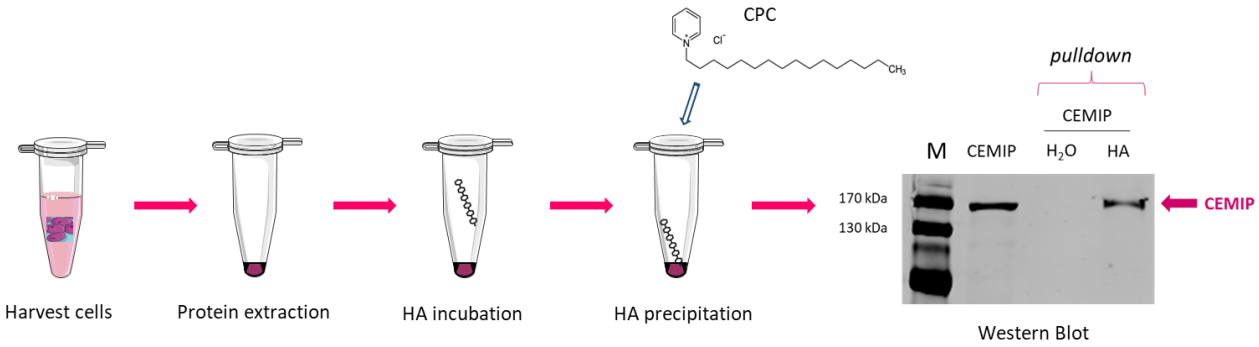


***Figure 23**: **CPC-HA pulldown workflow**. Incubation of the extracted proteins with HA and then precipitation the mixture with cetylpyridinium chloride (CPC). After centrifugation and washing steps, the precipitated HA together with its binding partner (CEMIP in this case) is dissolved in Laemmli blue with DTT and run on a Western blot with an anti-CEMIP antibody.*

At first, an issue with the negative control was observed. When incubating the protein mixture with water, CEMIP band shouldn't appear on the membrane since nothing is supposed to precipitate when adding the CPC (due to the absence of HA). Unfortunately, this was not the case, as illustrated in **Figure 24**.

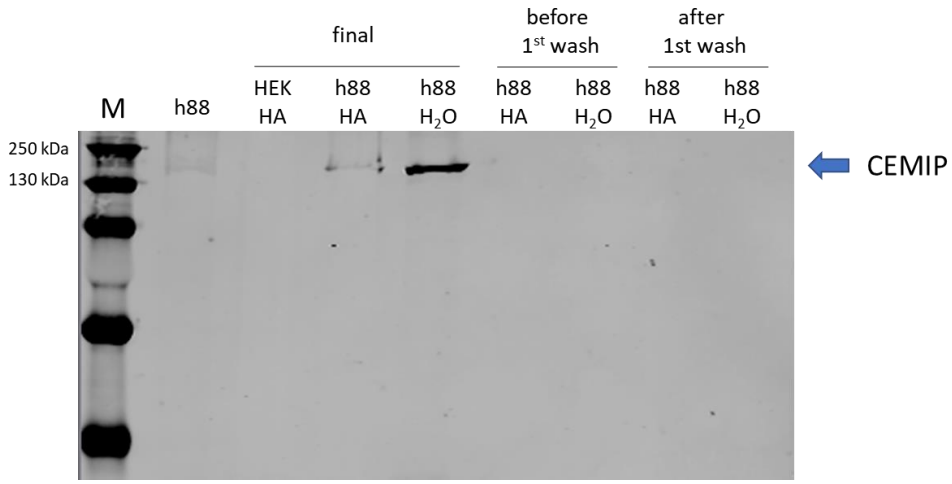


***Figure 24**: **Issue with CPC pulldown negative control**. CPC pulldown samples after and before washing steps. Negative control with water (H_2O) presents a CEMIP band, while the positive control is not revealing CEMIP. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

To solve this problem, a set of protocol variables was tested: dilution of the lysate (are proteins precipitating?), centrifugation step right after adding CPC (no second incubation), longer or shorter HA incubation at 37°C and incubation at room temperature. As exposed in **Figure 25**, the dilution of the sample plays a crucial role in the success of the negative control.

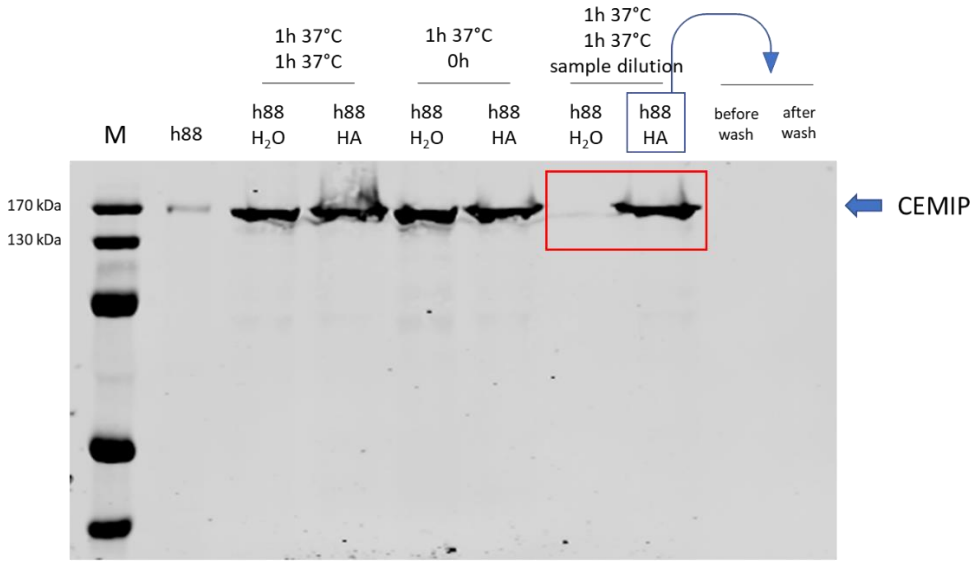


***Figure 25**: **Implementations of the CPC pulldown**. CPC pulldown with different parameters of incubation, temperature, and dilution. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

At this point, it became of prime interest to investigate the effect of the sample dilution. With a titration, I observed a clear threshold of the total protein precipitation at around a concentration of 0.8-1.5 mg/mL (**Figure 26**).

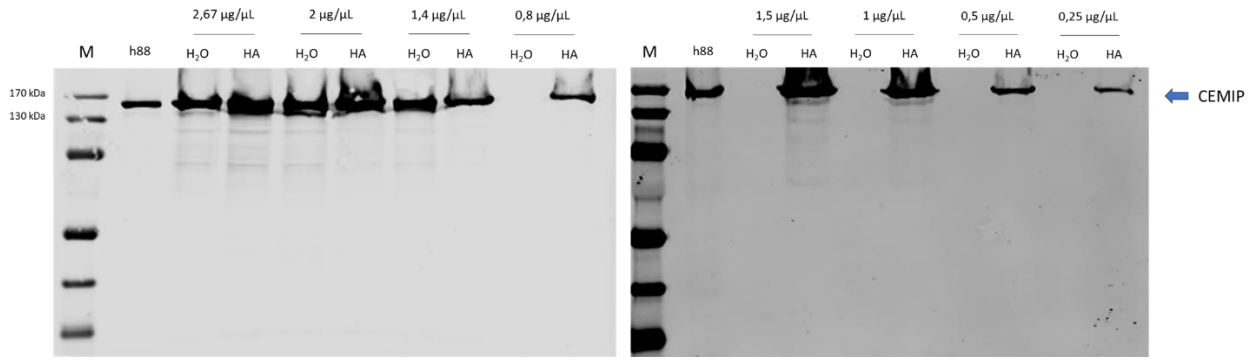


***Figure 26**: **Dilution effect on the proteins' precipitation**. CPC pulldown with protein lysate of concentrations from 2,67 to 0,25 mg/mL. Each sample was analyzed with water (negative control) or HA. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

The limit of the binding between CEMIP and HA in this experimental set-up was estimated at a protein concentration of around 0.1 mg/mL, concentration at which CEMIP band disappears (**Figure 27**).

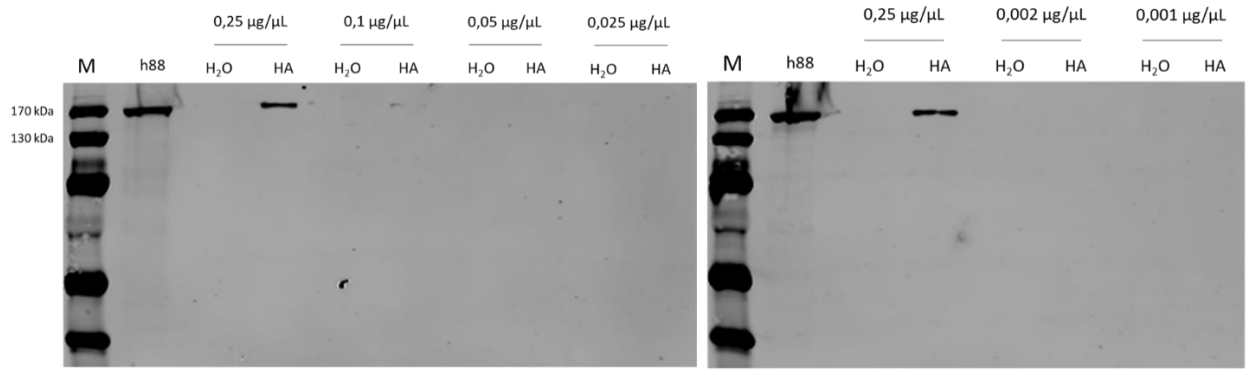


***Figure 27**: **Limit of the binding**. CPC pulldown with protein lysate of concentrations from 0,25 to 0,001 mg/mL. Each sample was analyzed with water (negative control) or HA. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

Finally, at good working concentrations (0.25 mg/mL < x < 1.5 mg/mL), incubation time and temperatures, the CPC pulldown protocol was validated with different CEMIP lysates stocks

(**Figure 28**). The sample with HA gives a clear CEMIP band, and the negative control with water shows no precipitation and absence of the CEMIP band.

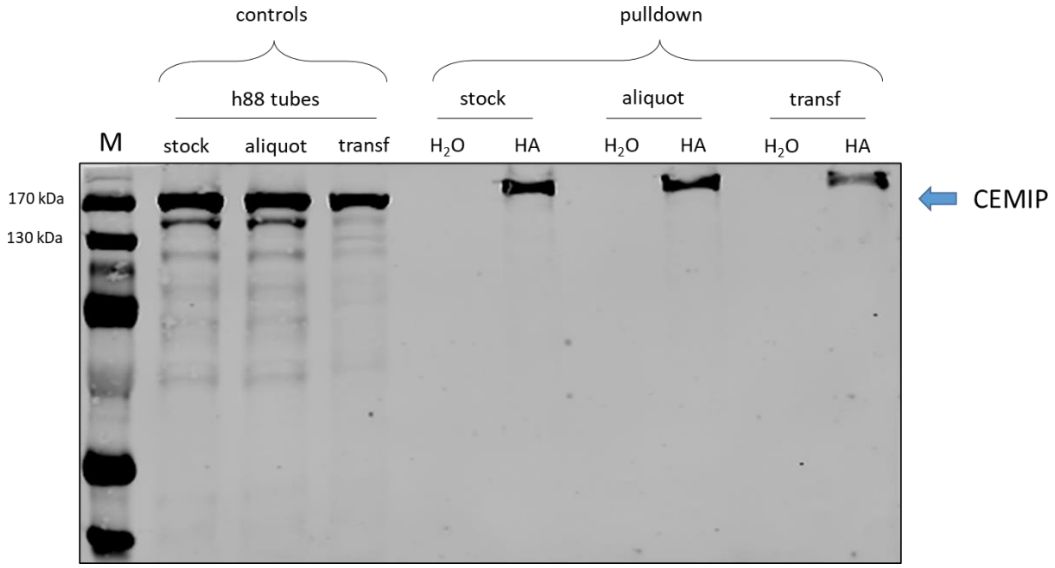


***Figure 28**: **Final validation of the CPC pulldown method**. Controls and pulldowns form different tubes: stock, aliquot and transfection (transf). Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

After observing a difference in molecular weight between the control CEMIP present in the lysate and the pulled down CEMIP, higher in MW (**Figure 28**), further analysis of the band was carried out, to verify its identity. Mass spectrometry analysis confirmed that the band appearing in the Western blot membrane is in fact CEMIP. More precisely, CEMIP was identified with 44 exclusive unique peptides allowing a protein coverage of 34% (**Figure 29**). Presumably, the HA still present in the sample is somehow dragging CEMIP through the SDS gel, increasing its molecular weight and delaying its run.

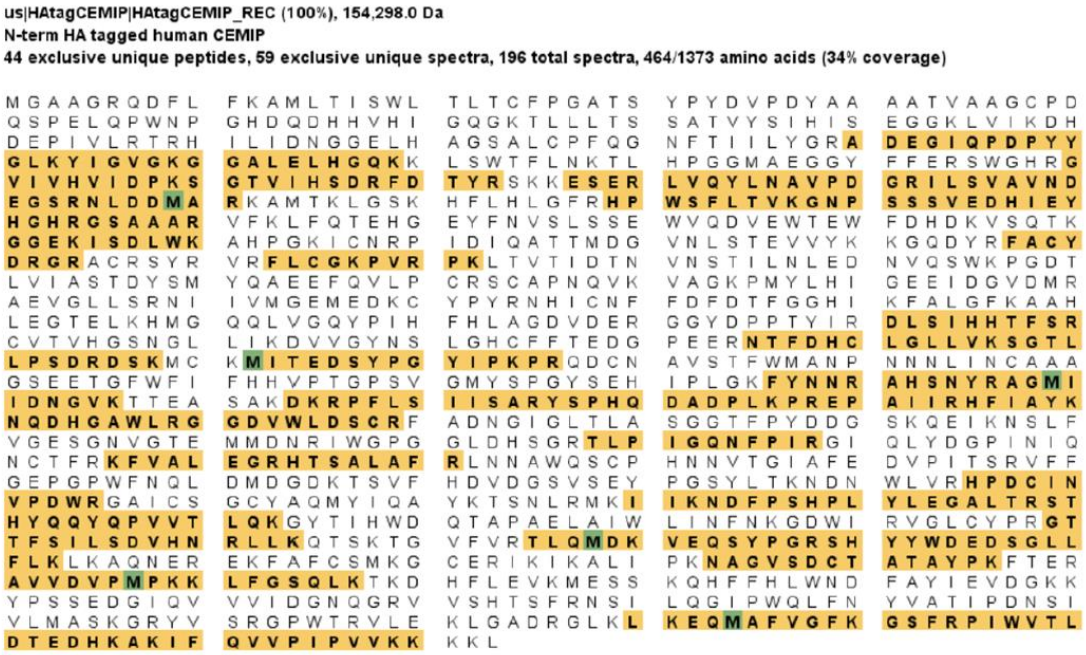


***Figure 29**: **Mass spectrometry analysis**. MS analysis of pulled-down CEMIP band confirms the presence of CEMIP with a 34% coverage (in yellow).*

With the HA assay gel and the CPC pulldown in place, one has now the means to proceed with CEMIP investigation.

6.3 Establishing a hCEMIP cell line

Since this work wants to focus on a human version of CEMIP, and the cell line available at the beginning of the project was a HEK293 line stably transfected with mouse CEMIP, I decided to establish a hCEMIP cell line. After transfection of HEK293T with the hCEMIP plasmid (VectorBuilder, vector ID VB191004-1084jkc) and a limited dilution selection, I selected clones based on CEMIP expression levels and HA depolymerizing activity. Among 8 clones, number 8 was the best in terms of HA degradation (**Figure 30**, left). This clone, h8, was tested in comparison to the murine CEMIP cell line with a known inhibitor, sulfated HA (sHA). Sulfated hyaluronic acid, a synthetic sulfated polysaccharide that inhibits the hyaluronidase HYAL1 with an IC_{50} of 0.4 – 0.8 μM[149,150], has been shown to have preventive and therapeutic effects on experimental mesangial proliferative GN[38], indicating that hyaluronidase activity plays a role in GN. Mesangial proliferative glomerulonephritis is a type of kidney disorder characterized by inflammation and

abnormal proliferation of mesangial cells within the glomeruli, leading to structural changes in the glomerulus and subsequent impairment of kidney function. This suggests a potential avenue for intervention in treating or preventing this type of glomerulonephritis, emphasizing the importance of maintaining the balance in the extracellular matrix components within the mesangium. Moreover, sulfated HA inhibits CEMIP hyaluronidase activity with an IC_{50} of 5-20 nM. To establish the inhibition by sHA on the hCEMIP clone generated (h8), an HA assay was performed revealing the correct IC_{50} (**Figure 30**, right). More details about sHA, its inhibition and toxicity evaluation can be found in section 7.1.1.

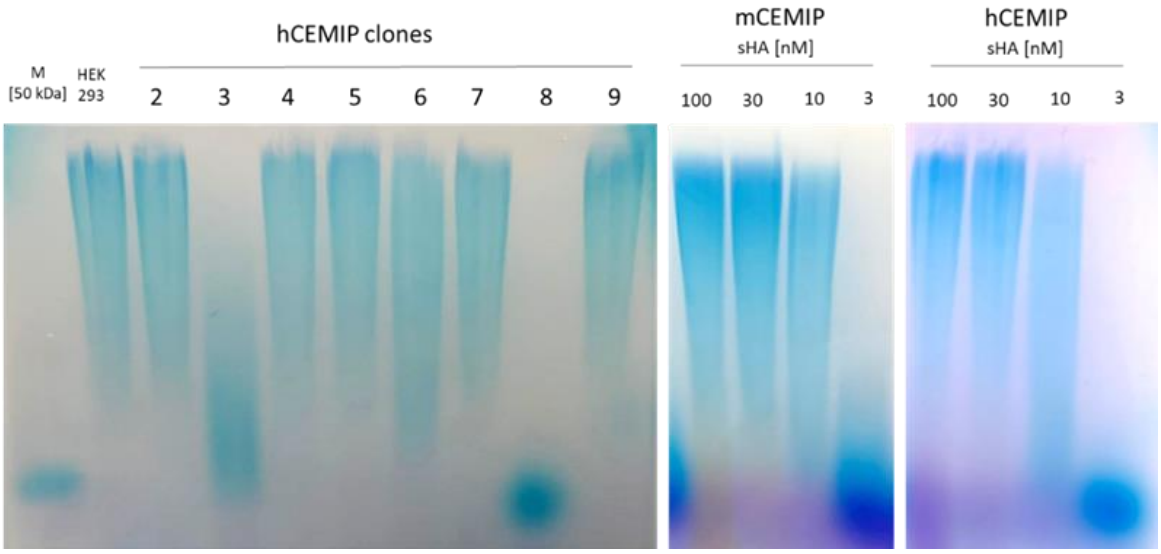


***Figure 30**: **Clone selection for a hCEMIP cell line**. Left: HA assay of HEK293 hCEMIP clones (named 2 to 9). Right: HA assay of the inhibition by sHA from 100 nM to 3 nM on murine (mCEMIP) and human clone 8 (hCEMIP).*

To gather a deeper understanding of the difference between the murine and the human clones I performed a time course experiment during 72h using the exact same conditions for mCEMIP and hCEMIP. I observed a difference in the HA depolymerization timing, where hCEMIP-expressing cells were more efficient (**Figure 31**). This raised the question of whether the observed difference was due to a difference in protein expression where the hCEMIP expression is higher than the mCEMIP.

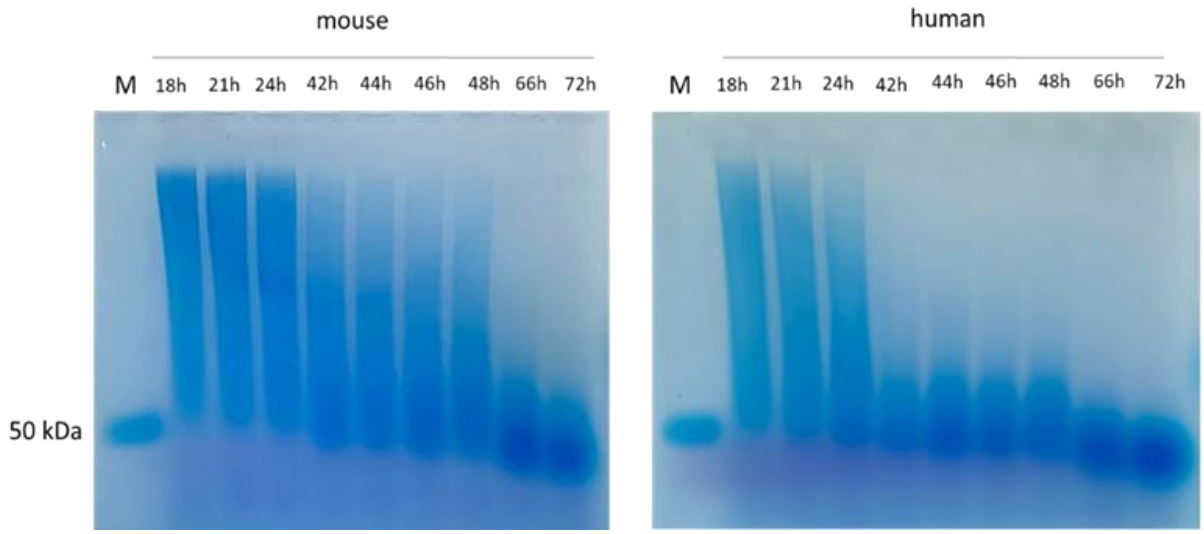


***Figure 31**: **Murine versus human clone**. HA assay at different time points revealed enhanced HA depolymerization of the human clone over 72h. Left: HA of murine clone cells at different time points. Complete depolymerization is reached after 72h. Right: HA of human clone cells at different time points. Complete depolymerization is reached after 48h.*

In order to answer this question, I performed a Western blot analysis of the lysate as well as the supernatant of these two clones (**Figure 32**). I observed a large increase of CEMIP expression in the human clone explaining the observed difference in efficiency in HA cleavage.

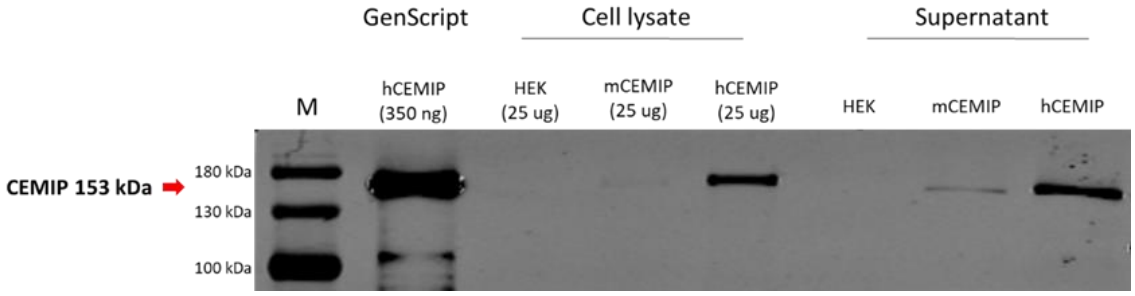


***Figure 32**: **Western blot of murine and human clones**. Control (hCEMIP protein produced by GenScript), cell lysate and supernatant from HEK293T, mCEMIP clone and hCEMIP clone tested by WB with an anti-CEMIP antibody. WB analysis revealed CEMIP expression in both cell lysate and supernatant, as well as higher levels of CEMIP in the human clone.*

Since in Western blot of **Figure 32** an anti-mouse antibody was used (our in-house anti-CEMIP of section 7.2), a follow-up Western blot to check for cross-reactivity with a commercial anti-rabbit anti-CEMIP antibody was performed. This gave a double confirmation that the h8 clone is

indeed expressing CEMIP in both lysate and supernatant, with increased levels compared to the murine clone (**Figure 33**).

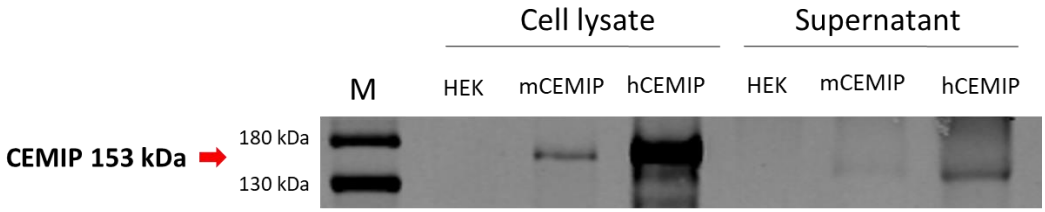


***Figure 33**: **Western blot of murine and human clone**. Cell lysate and supernatant from HEK293T, mCEMIP clone and hCEMIP clone tested by WB with a commercial anti-rabbit anti-CEMIP antibody. WB analysis revealed CEMIP expression in both cell lysate and supernatant, as well as higher levels of CEMIP in the human clone.*

As proof of hCEMIP integration in the genomic DNA, the latter was extracted form clone 8 and a PCR with primers inside the CEMIP sequence (giving a fragment of 1'200 bp) revealed the integration in the genomic DNA (**Figure 34**).

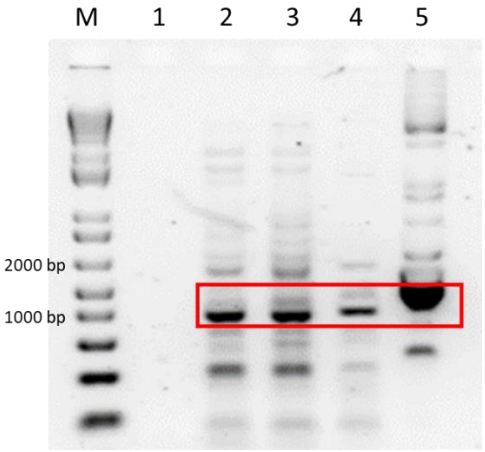


***Figure 34**: **hCEMIP is stably integrated in the genomic DNA**. M = marker; 1 = negative control; 2 = gDNA, 58°C annealing temperature; 3 =gDNA, 60°C annealing temperature; 4 =gDNA, 62°C annealing temperature; 5 =plasmid hCEMIP DNA, 58°C annealing temperature.*

One year later, a loss of CEMIP activity was observed and a second selection was performed. Again clone 8 was selected based on HA depolymerization ability and CEMIP expression (**Figure 35**). The new cell line was named "h88" for hCEMIP clone 8 subclone 8.

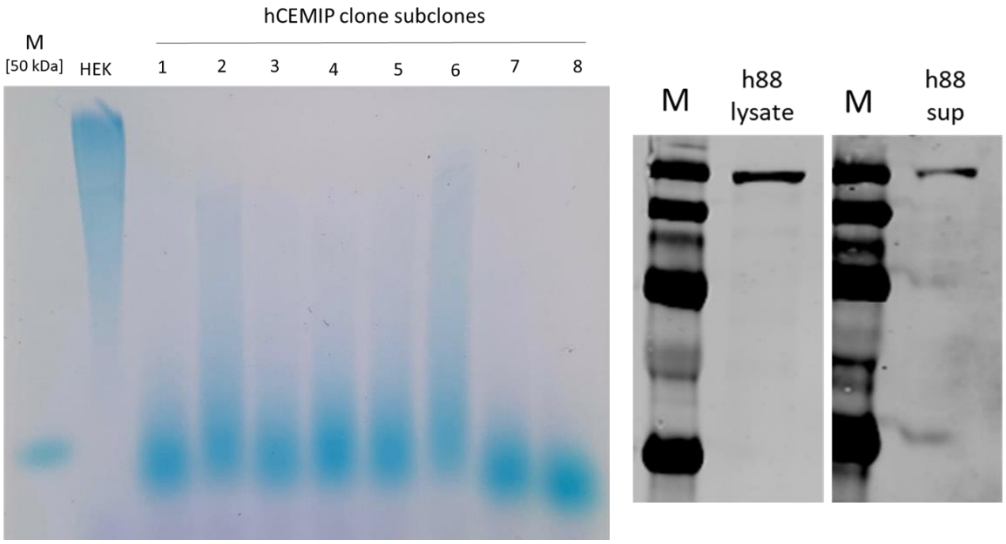


***Figure 35**: **Clone selection of HEK293-hCEMIP clone 8**. Left: HA assay of different hCEMIP clones subclones (from 1 to 8) and control HEK293T. Right: CEMIP expression of clone hCEMIP 8 subclone 8 (h88) in the lysate and supernatant (sup) by WB with an anti-CMEIP antibody. Subclone 8 displays the best HA depolymerization and CEMIP expression.*

6.4 Analysis of HA fragments with size exclusion chromatography

Despite giving a clear result about CEMIP activity, the HA gel assay does not provide a precise measure of the size of the fragments. Inspired by Yoshida's protocols about separating radio-labelled HA fragments with a Sepharose column by size exclusion chromatography[37], I decided to invest some time in developing my own protocol to separate HA fragments after cellular incubation. After some unsatisfactory results with an Ultrahydrogel column available in the department, and some different mobile phase testing, I finally developed a protocol using a Superose 6 Increase 10/300 GL column (Cytiva) in PBS (**Figure 36**).

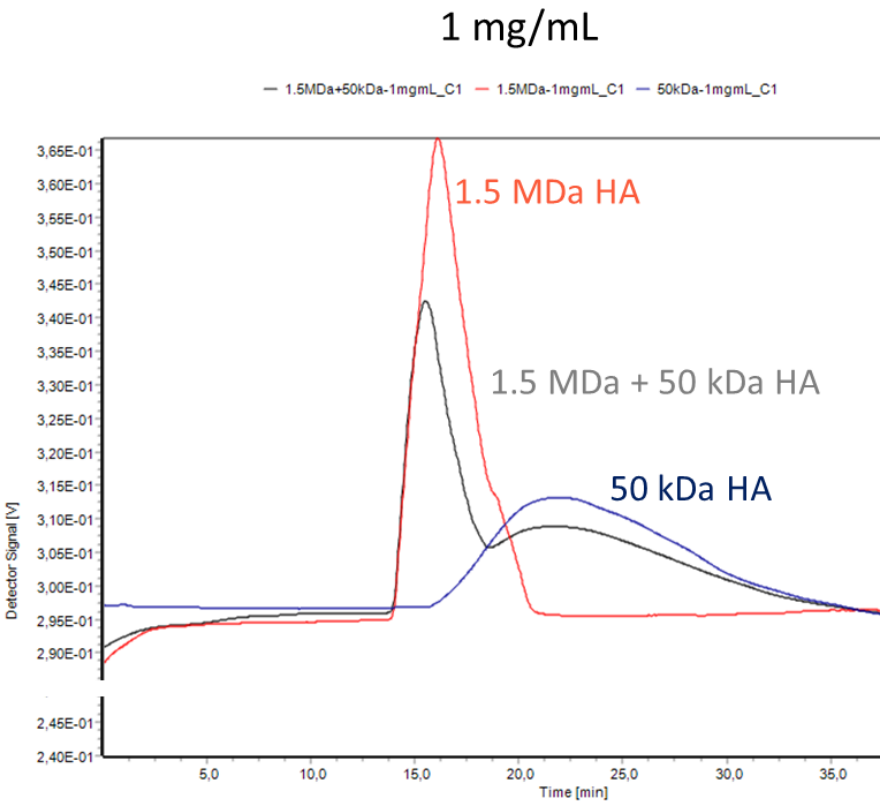


***Figure 36**: **SEC profile of commercial HA**. Separation of hyaluronic acid of 1500 kDa (red), 50 kDa (blue) and a mixture of both (grey) by SEC (PBS, 1 mg/mL). X axis: time in minutes; Y axis: detector signal.*

It should be mentioned that several trials with PBS enriched with NaCl were conducted in order to reduce the osmotic pressure, reduce the repulsive interactions of HA and reach a better separation. Anyway, simple PBS turned out to be the best choice due to its stability, while PBS with NaCl was causing shifting of the mobile phase and samples profiles.

Next, the relationship between the concentration (or the HA injected in the column) and the area under the curve was established for each HA size. For HA of 50 kDa and 100 kDa, the relationship was linear for concentrations from 0.1 mg/mL to 5 mg/mL (**Figure 37**). Interestingly, for HA of 1500 kDa the linearity was lost for a concentration equal to or higher than 2 mg/mL (200 µg).

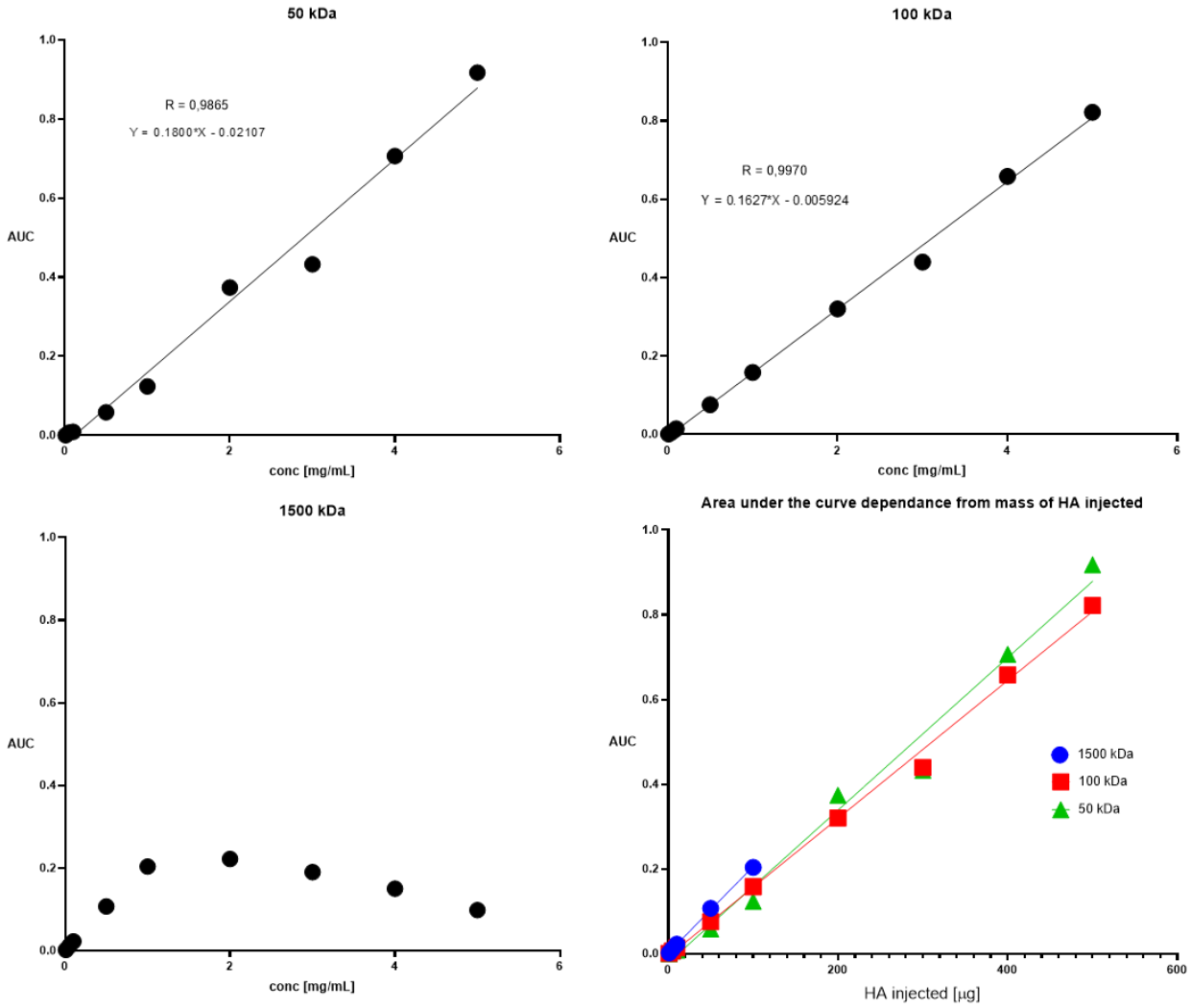


***Figure 37**: **AUC versus HA concentration**. The function of AUC (area under the curve) and HA concentration or mass of HA injected. X axis: concentration in mg/mL; Y axis: area under the curve (AUC) or HA injected in µg.*

Finally, injection of samples from cells supernatants after purification was performed. When zooming into the samples profile, we can observe the HA from HEK293T cells exiting the column at around minute 11, while HA processed by clone h88 at around minute 17 (**Figure 38**, bottom left). This reflects exactly the profile displayed by commercial HA of 1500 kDa and 50 kDa at 0.1 mg/mL (**Figure 38**, top right).

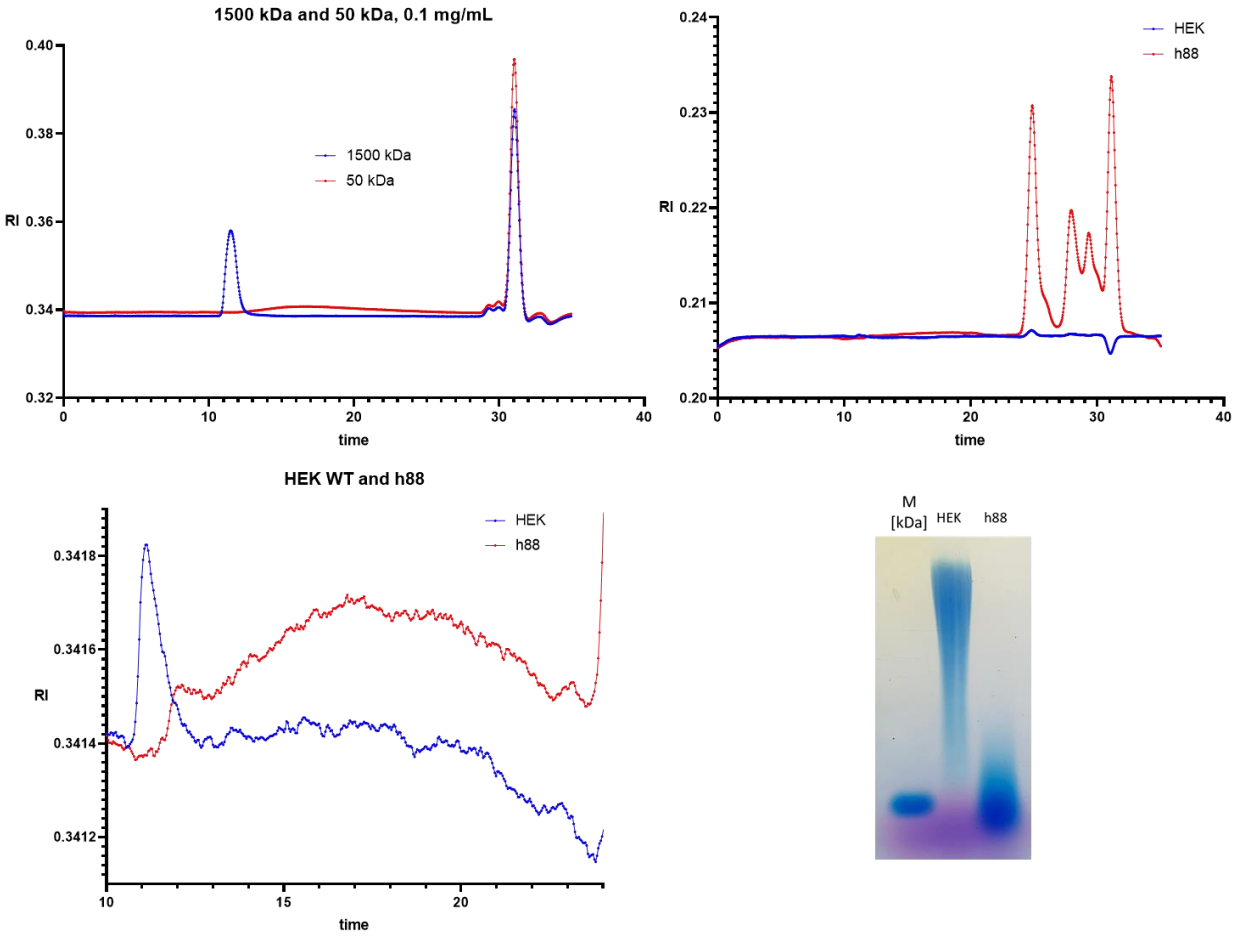


***Figure 38**: **SEC of HA from cells' supernatant**. Top left: commercial HA 1500 kDa (blue)and 50 kDa (red) at 0.1 mg/mL profile; top right: HA from HEK293T (blue) and clone h88 (red); bottom left: zoom of the previous panel; bottom right: HA gel of the same samples injected in the column (5 µg of HA).*

As the samples' concentrations were very low and required a possible scale-up of the HA assay, this part of the project was put on standby and was eventually carried on in Bellinzona using fluorescent HA.

6.5 Investigating CEMIP mode of action

Until now, the exact mode of action of CEMIP has not been elucidated yet. Yoshida *et al.*[37] demonstrated that the HA is catabolized by CEMIP via the clathrin-coated pit pathway, and previous evidence has shown that the HA depolymerization occurs in the extracellular space and via receptor-mediated endocytosis[151-153]. Other groups identified an N-terminal signal sequence

needed for the cellular trafficking of CEMIP, and that cleavage of these N-terminal 30 amino acids occurs in the mature CEMIP[81], but it still remains uncertain how CEMIP is processing the HA. Previous studies have shown that CEMIP is essential in HA depolymerization and is able to bind to HA[37,60], but it remains uncertain if it has intrinsic enzymatic activity. Moreover, recent findings reported that the HA-degrading activity of secreted CEMIP depends on its G8 domain's ability to adhere to the membrane in an annexin A1 (ANXA1) dependent manner[82].

In order to elucidate the points mentioned above, I investigated the activity of CEMIP outside the cell in the absence and presence of ANXA1 expression, the role of the binding of CEMIP (and secreted CEMIP) to HA of different sizes and the importance of the C-terminal portion of CEMIP.

6.5.1 Intracellular CEMIP is binding to HA of different sizes

In confirmation with previous studies, CEMIP is binding to hyaluronic acid of different sizes, 1'500 kDa, 100 kDa and 50 kDa, in a CPC pulldown experiment (**Figure 39**).

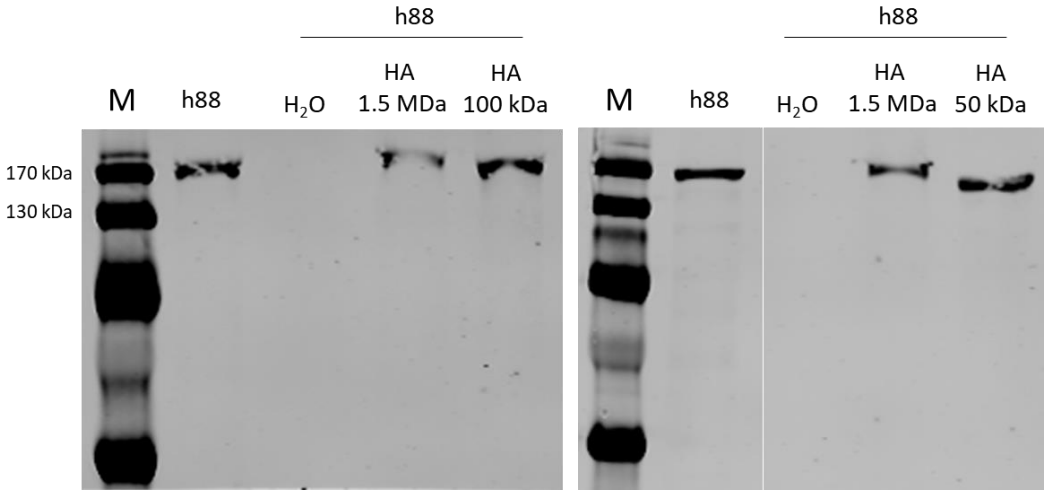


***Figure 39**: **CPC pulldown of CEMIP in the cell lysate with different HA sizes**. First lane after the marker represents the control (no pulldown). Other lanes represent the pulldown with different HA sizes and water as control. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

6.5.2 CEMIP is secreted in an inactive form

To investigate CEMIP activity and understand if the secreted protein is active, I tested different human CEMIP (produced from GenScript or from the supernatant of the stable clone generated)

in different environments: added in the media of cultured HEK293 cells, in simple serum-containing culture media, and in HEK293 cells supernatant (without cells), as illustrated in **Figure 40**. Intriguingly, CEMIP seems to be active only in cells that are producing it. This could be one more evidence supporting the hypothesis that CEMIP is membrane-bounded, internalizes the HA via the clathrin-coated pit pathway, where the depolymerization takes place, and is released in an inactive form in the extracellular space.

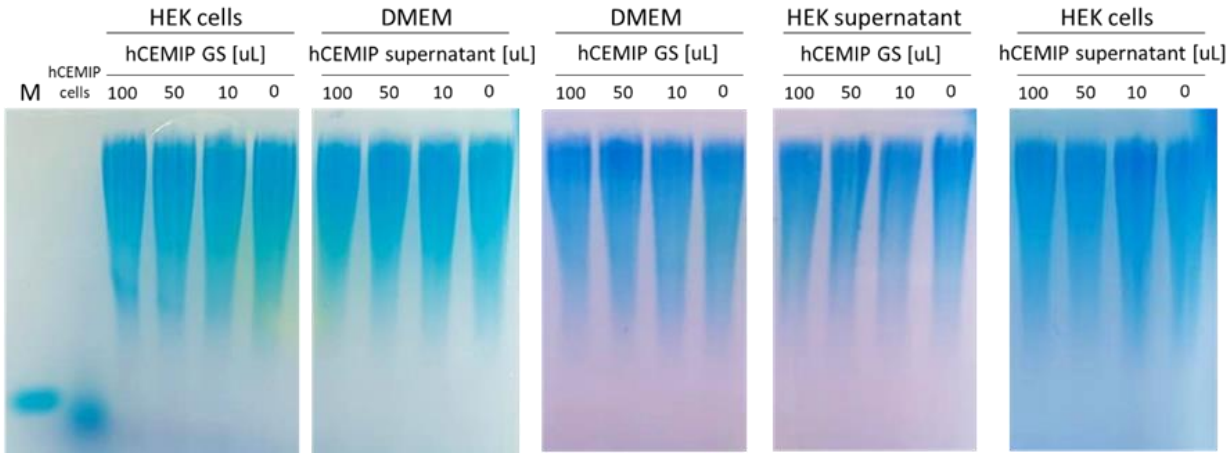


***Figure 40**: **CEMIP is active only in cells that are producing it**. Different conditions of CEMIP and cells or medium was tested in order to select the variables by which CEMIP is active. HA assay of control (hCEMIp cells) and different combinations of media and cells. M = marker 50 kDa; GS = GenScript; DMEM = Dulbecco's Modified Eagle Medium.*

6.5.3 Secreted CEMIP is binding to HA of different sizes

Despite not being able to process HA when secreted, it is fascinating to observe that in a CPC pulldown experiment, secreted CEMIP (sCEMIP) is binding to hyaluronic acid of different sizes: 1'500 kDa, 100 kDa and 50 kDa (**Figure 41** and annex **Figure 116**).

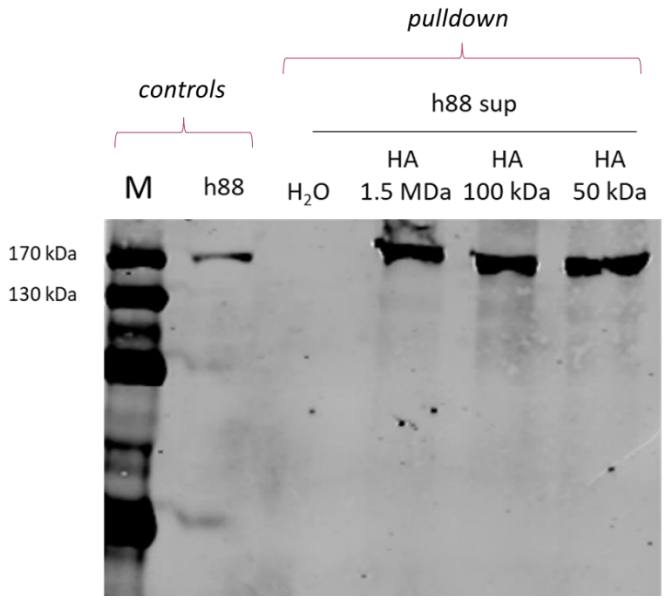


***Figure 41**: **CPC pulldown of secreted CEMIP with different HA sizes**. sCEMIP is binding to HA fragments of 1500, 100 and 50 kDa. First lane after the marker represents the control (no pulldown). Other lanes represent the pulldown with different HA sizes and water as control. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP.*

6.5.4 Secreted CEMIP's activity is not rescued by ANXA1

Recent findings reported that the HA-degrading activity of sCEMIP depends on its G8 domain's ability to adhere to the membrane in an annexin A1 (ANXA1) dependent manner[82]. When sCEMIP-rich medium was supplemented with rheumatoid arthritis fibroblast-like synoviocytes (RA FLS) membrane fractions, HA was degraded. On the other hand, this didn't happen in the HEK293T-CEMIP medium supplemented with HEK293T-CEMIP cell membranes. Then, in an immunoprecipitation experiment, Zhang and colleagues identified ANXA1 as a CEMIP-interacting membrane protein. Interestingly, sCEMIP showed HA degradation ability in media mixed with the membrane fractions of HEK293T-ANXA1 cells. They also reported co-localization of CEMIP and ANXA1 on RA FLS cell membrane by immunofluorescence staining.

I decided to test whether the presence of annexin 1 in HEK293T cells could rescue the activity of secreted CEMIP function by supplementing h88 supernatant containing sCEMIP to HEK293T cells transfected with ANXA1. Under this conditions, sCEMIP activity doesn't seem to be rescued by ANXA1 (**Figure 42**).

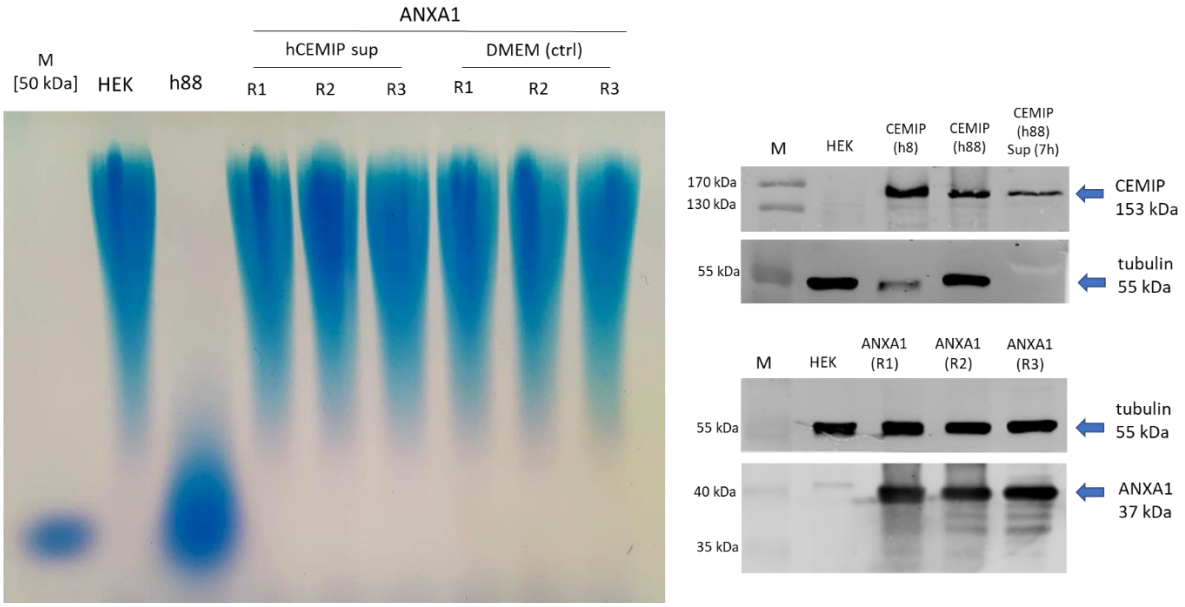


***Figure 42**: **HA assay and Western blot of sCEMIP on HEK293T-ANXA1**. Left: HA assay of controls HEK293T and h88 and conditions in triplicate (R1, R2, R3) of sCEMIP (hCEMIP sup) and DMEM on ANXA1 expressing cells. Right: Western blot of CEMIP and ANXA1 expression for each condition, and anti-tubulin as loading control. h88 = HEK293T-hCEMIP.*

I then wondered if the activity of the membranous CEMIP could be rescued by annexin 1. I proceeded to extract h88 membranes and add them to HEK293T-ANXA1 cells, but again, no HA depolymerization was observed (**Figure 43**).

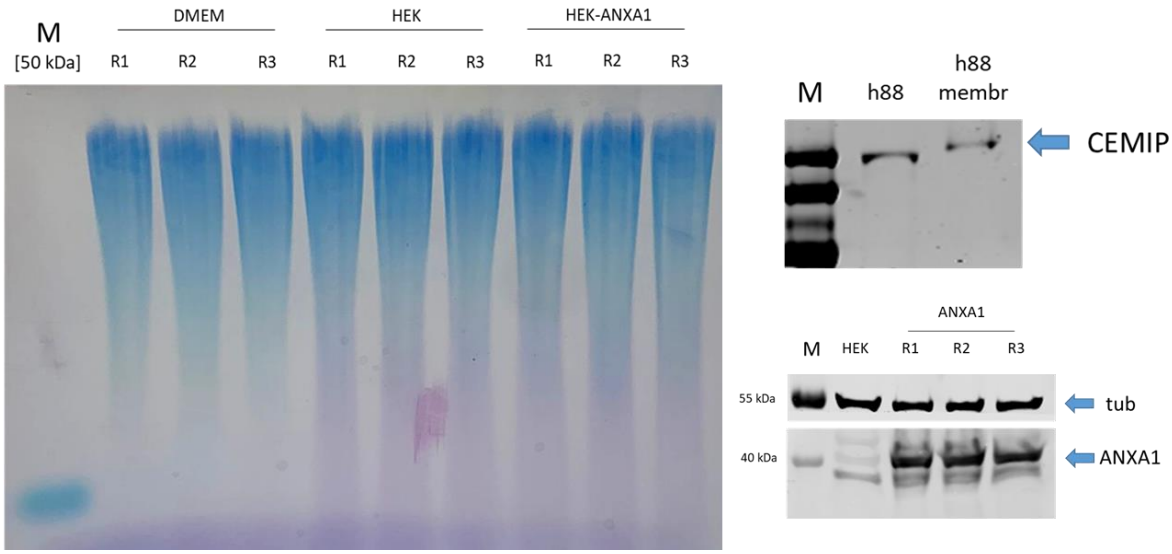


***Figure 43**: **HA assay and Western blot of membranes from h88 cells on HEK293T-ANXA1**. Left: HA assay of conditions in triplicate (R1, R2, R3) of h88 membranes fraction on DMEM, HEK293T and ANXA1 expressing cells. Right: Western blot of CEMIP and ANXA1 expression for each condition, and anti-tubulin as loading control. h88 = HEK293T-hCEMIP.*

To have final confirmation of the unrelated connection between CEMIP and ANXA1 in transfected HEK239T cells, collaborators in IRB (Bellinzona, Switzerland) tested the localization of these two proteins by immunofluorescence. As shown in **Figure 44**, CEMIP and ANXA1 do not co-localize.

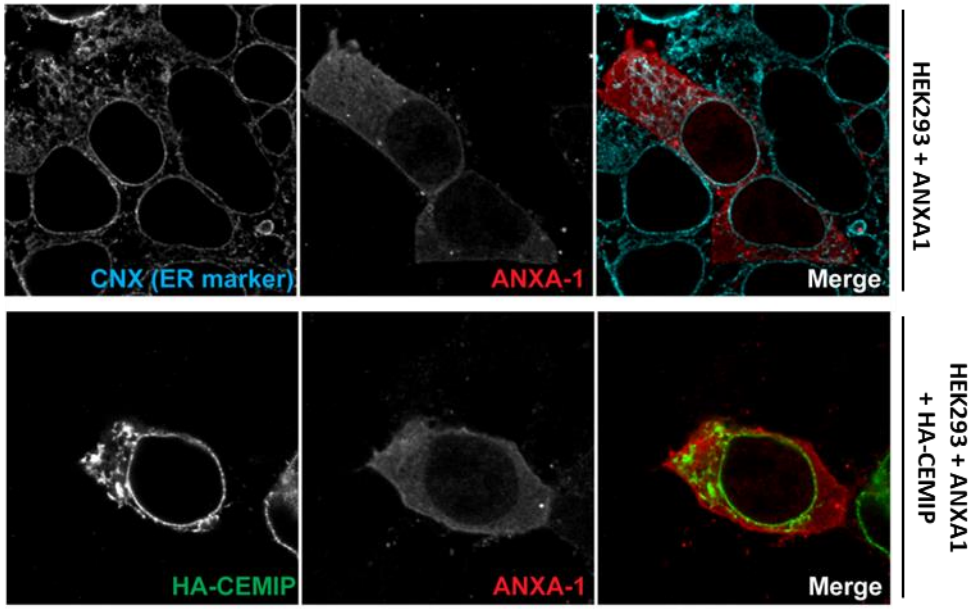


***Figure 44**: **CEMIP and ANXA1 are not co-localized**. Immunofluorescence of HEK293T-ANXA1 cells without or with CEMIP expression. ANXA1 in red, CEMIP in green and ER in blue. HA-CEMIP = hCEMIP HA-tag N-term. Courtesy of Concetta Guerra.*

6.5.5 CEMIP C-terminus preservation is important for HA degradation

When cloning a new tag into hCEMIP construct, I encounter an unexpected observation about the C-terminus of the protein. Since the original plasmid lacked a tag, later during the project it became of interest to introduce one. For this reason, an HA-tag was inserted at the N-terminus or at the C-terminus. At the N-term, the tag was inserted between the signal sequence (amino acids 1-30) and the protein sequence with an AAA linker, since the signal sequence is reported to be required for cellular trafficking and hyaluronan depolymerization[81]. At the C-term, the tag was inserted after the KKKL signal with a GSS linker at the beginning, placing the stop codon after the HA-tag. Despite immunofluorescence experiments showing both constructs expressed and localized in the ER when transfected in HEK293T cells (**Figure 45**), a crucial difference was observed in the HA degradation activity.

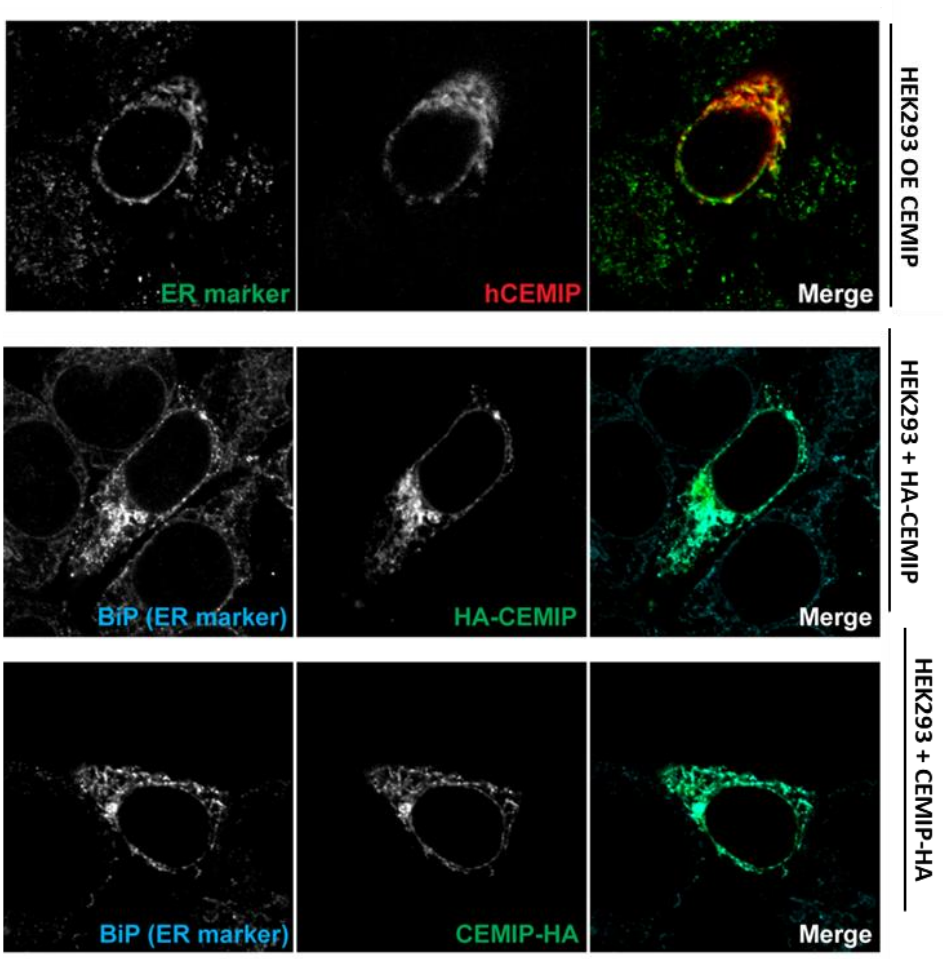


***Figure 45**: **CEMIP is found in the ER**. Immunofluorescence of hCEMIP clone (HEK293T OE CEMIP), and HEK293T transiently transfected with hCEMIP-HA tag at the N-term (HA-CEMIP and C-term (CEMIP-HA). hCEMIP satble in red, CEMIP in green and ER in green (first row) and blue (lower rows). Courtesy of Concetta Guerra.*

When tested in an HA assay and compared to the WT CEMIP in transient tranfection, HA-tag-CEMIP conserved the ability to depolymerize hyaluronic acid and was well expressed (**Figure 46**). In fact, this became the construct for future cloning of the project, giving the possibility to reveal CEMIP with an anti-HA-tag antibody in case needed (annex **Figure 127**).

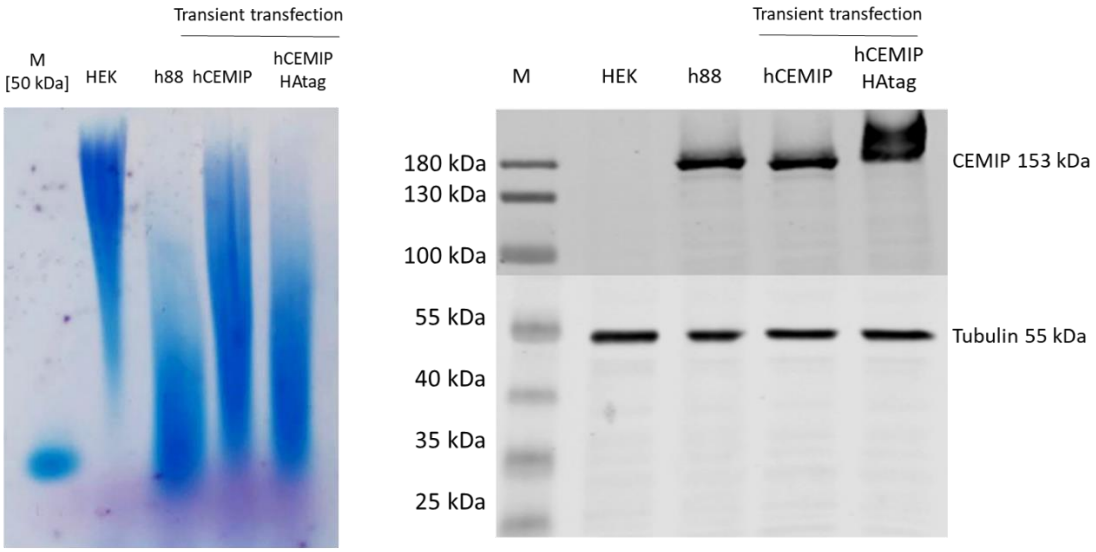


***Figure 46**: **HA assay and Western blot of CEMIP HA-tag at the N-term**. Left: HA assay of HEK293T, hCEMIP stable clone (h88), hCEMIP transient and hCEMIP-HA tag N-term transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

On the other hand, interestingly, when adding the tag at the C-terminus, this version of CEMIP was not able to process hyaluronic acid, despite being well expressed (**Figure 47**). This raised the question of whether the C-term part of CEMIP is somehow compromised in this construct, probably highlighting the importance of the KKKL retention signal, a classical ER retention signal (KKXX) for transmembrane localization.

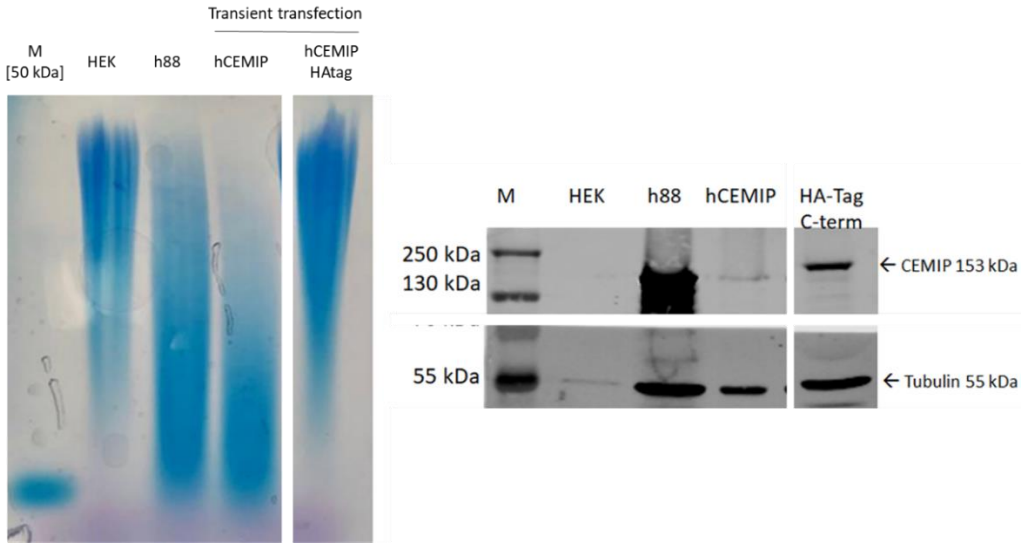


***Figure 47**: **HA assay and Western blot of CEMIP HA-tag at the C-term**. Left: HA assay of HEK293T, hCEMIP stable clone (h88), hCEMIP transient and hCEMIP-HA tag C-term transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

Further confirmation about the importance of the C-terminus and KKKL retention signal preservation was obtained when working on an hCEMIP-EGFP construct to track CEMIP localization inside cells. An HA assay revealed that this form of CEMIP is not efficient in breaking down hyaluronic acid (**Figure 48**).

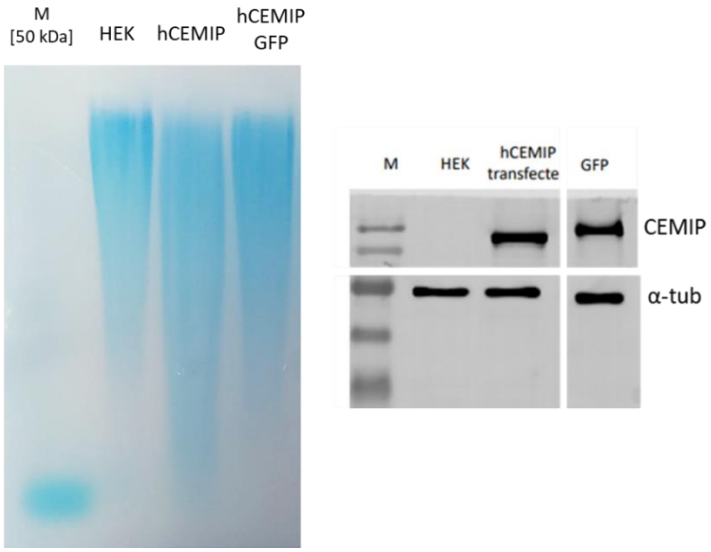


***Figure 48**: **HA assay and Western blot of hCEMIP-EGFP**. Left: HA assay of HEK293T, hCEMIP transient and hCEMIP-GFP transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

6.6 Understanding CEMIP through mutagenesis

To elucidate CEMIP different domains and amino acids' function, I relied on site-directed mutagenesis. To impair the function to the maximum, the amino acids were changed in a disruptive manner: arginine (R) into cysteine (C), aspartic acid (D) into asparagine (N), and histidine (H) into arginine (R) (**Figure 49**).

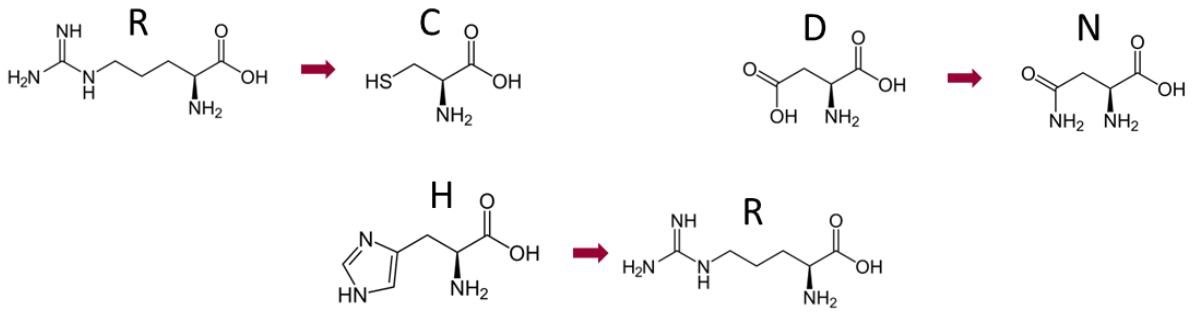


***Figure 49**: **Disruptive amino acid mutations**. Arginine (R) into cysteine (C), aspartic acid (D) into asparagine (N), and histidine (H) into arginine (R).*

Mutants were tested for their HA cleaving ability as well as their HA binding capability. In this scenario, if one mutated construct conserves both HA depolymerization and HA biding, the said amino acid or region is not considered particularly important for the protein's function. On the other hand, if a mutant binds to HA but loses its ability to process it, the mutated positions will belong to the catalytic site of CEMIP. It cannot be excluded that this type of mutation could also impact the catalytic site in another way, for example an adjacent mutation leading to a structural change of the catalytic site, or may be affecting trafficking, stability and indirect activity. A third possibility is when a mutant loses both functionalities, binding and cleaving HA. In this case, the changed amino acids or region are involved in the HA biding, therefore compromising the degradation of HA. **Table 2** summarizes the mentioned possibilities.

***Table 2**: Different outcomes from mutants' functionality assays.*

IF A MUTANT:	CLEAVES	DOESN'T CLEAVE
BINDS	NOT IMPORTANT AA	CATALYTIC AA
DOESN'T BIND	IMPOSSIBLE	BINDING AA

Quantification and normalization of CEMIP before pulldown

Since mutants of CEMIP are expressed at different levels in the cells' lysates, it was important to quantify and normalize CEMIP before testing its binding to HA, in order to be able to compare the resulting pulled-down bands. With this purpose, preliminary Western blot bands were quantified for CEMIP and α-tubulin. CEMIP values were divided by the α-tubulin ones and normalized to the CEMIP WT (100%). Dilution of the lysate for the pulldown was then calculated, making sure that the range of the working concentrations of proteins in the lysate was acceptable (< 1.5 mg/mL; annex **Figure 124** and **Figure 125**).

6.6.1 Arg[187] and Asp[208] are essential for HA processing, and Arg[187] is involved in HA binding

In order to explore the function of CEMIP first GG domain, I identified in the literature mutations that reduce the HA-depolymerizing activity of CEMIP and of the hyaluronidase TMEM2. Abe and collaborators identified in different families two mutations in CEMIP causing nonsyndromic hearing loss, R187C and R187H[47]. Ten years later, in 2013, Yoshida and colleagues demonstrated that R187C and R187H eliminate HA-degrading activity[37]. Moreover, a few years later Yamamoto reported that R265C, D273N and D286N's activity reduced in TMEM2[40]. When aligning these two proteins, the identified amino acids conserved in CEMIP and TMEM2 in different species correspond to CEMIP's amino acids R187, D195 and D208 (**Figure 50**).

All these mutations are found in the first GG domain, which is hypothesized to be essential for CEMIP function[47]. As shown in **Figure 50**, R187 and D208 are in proximity to each other, rather than D195 which is on the opposite side of the domain. Computational analysis of my IRB colleagues suggested this domain to be involved in the HA binding.

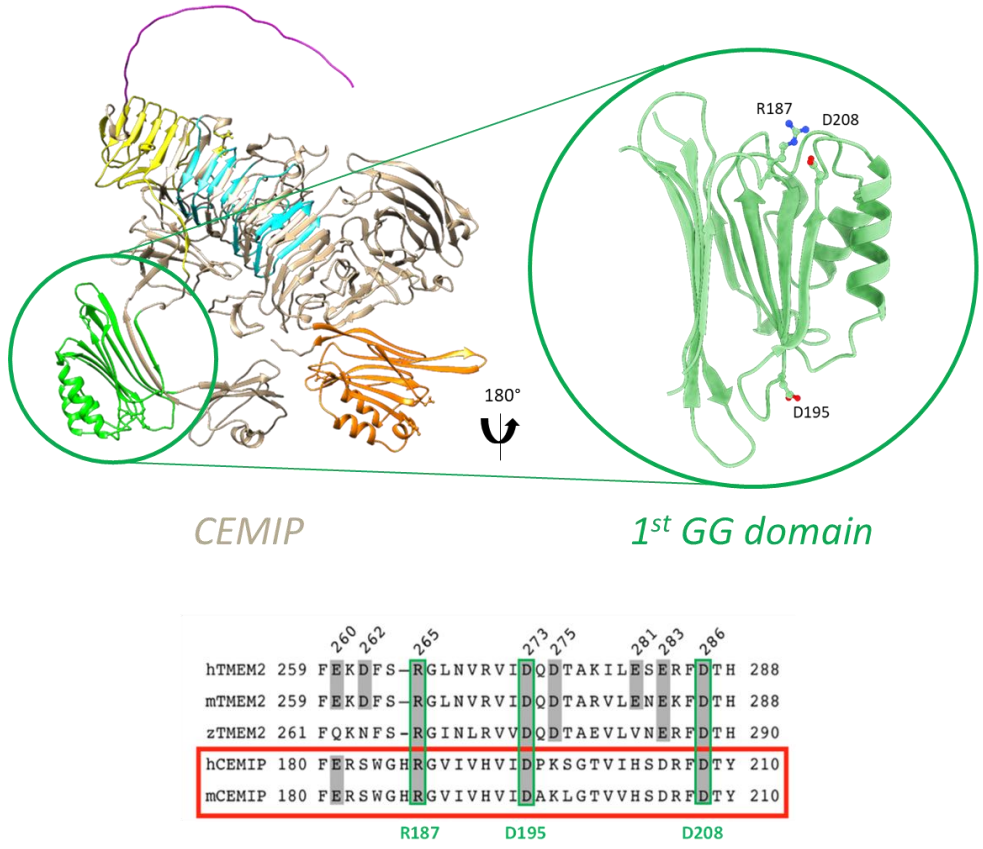


***Figure 50**: **First GG domain presumed important amino acids**. Top: AlphaFold GG domain structure. Courtesy of Margaux Heritier. Bottom: Conserved AA in CEMIP and TMEM2 among species (human, mouse, zebrafish). Modified from Yamamoto[40].*

This hypothesis was tested by transient transfection of the mutated hCEMIP plasmids in HEK293T cells followed by analysis of the HA present in the supernatant by an HA assay. Mutants R187C and D208N failed to depolymerize HA, while D195N maintained a HA depolymerization profile similar to the WT CEMIP (**Figure 51**).

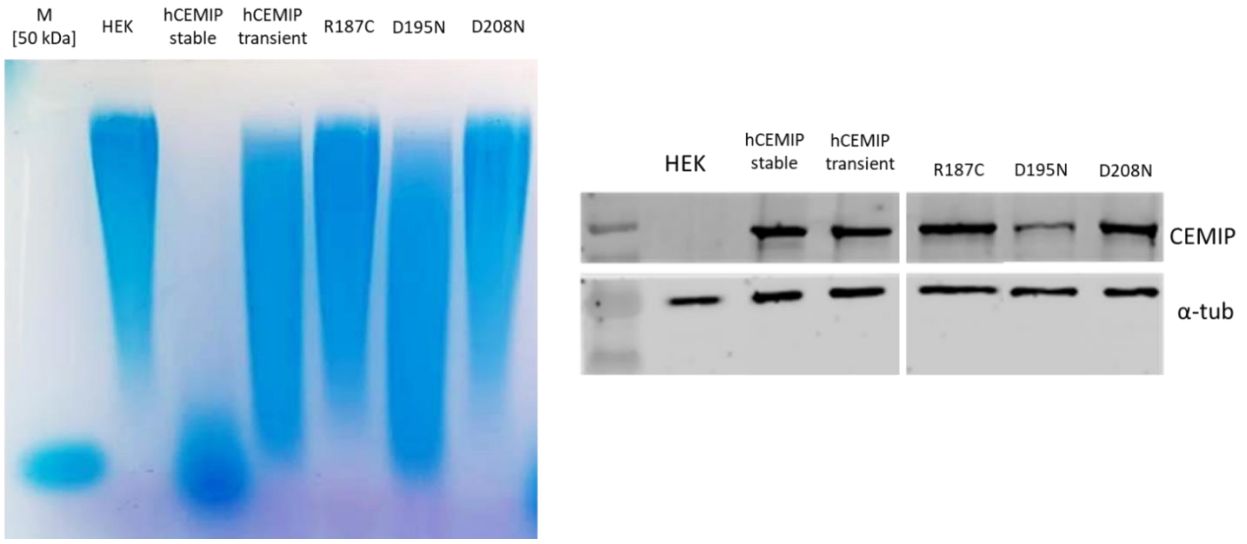


***Figure 51**: **HA assay and Western blot of the first GG domain mutants**. Left: HA assay of HEK293T, hCEMIP stable clone (h88), hCEMIP transient and mutants transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

To discriminate whether the impaired HA depolymerization of mutants R187C and D298N is due to the inability to bind HA, or to a damage of the catalytic site, I performed a pulldown experiment. Curiously, D208N conserved the HA binding ability (194%, **Figure 52**), but R187C lost almost all its HA binding ability (-81%, **Figure 52**), suggesting that the arginine in position 187 is an important AA binding hyaluronic acid.

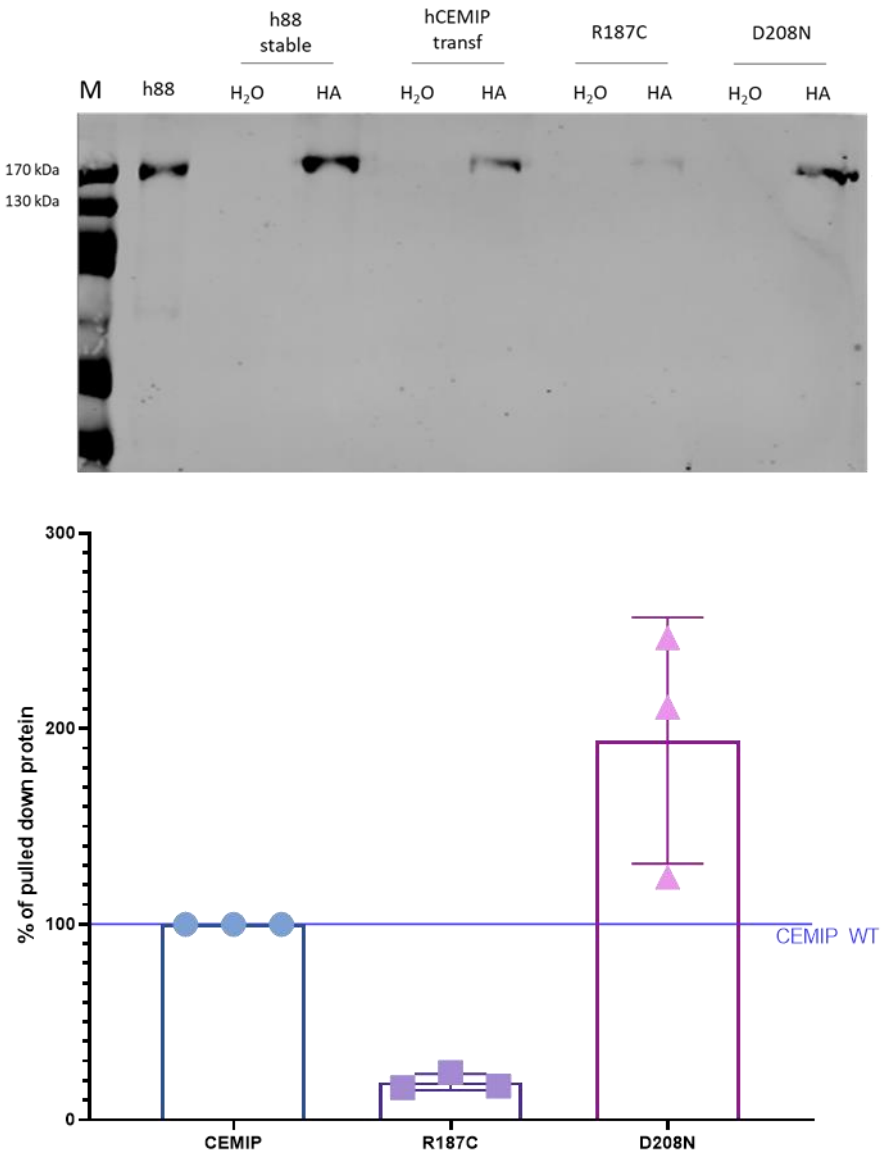


***Figure 52**: **Pulldown of R187C and D208N**. Top: Western blot membrane of the pulldown (with 1.5 MDa HA) with an anti-CEMIP antibody. Bottom: Band quantification standardized by the level of CEMIP (100%).*

6.6.1.1 Double mutant R187C and D208N

Next, I was interested in observing the activity of the double mutant R187C+D208N. Unfortunately, despite trying multiple times, the double mutant was never well expressed, invalidating the HA assay's results (**Figure 53** and annex **Figure 126**).

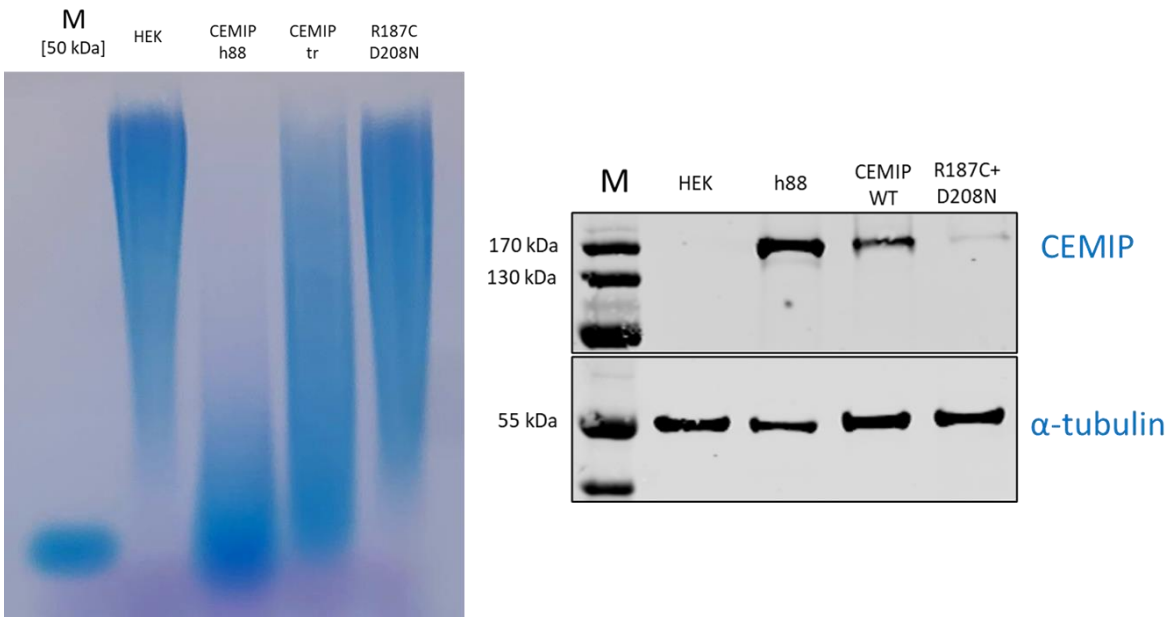


***Figure 53**: **HA assay and Western blot of mutant R187C+D208N**. Left: HA assay of HEK293T, hCEMIP stable clone (h88), hCEMIP transient and mutant transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

6.6.2 CEMIP second GG domain has HA anchoring function

Since CEMIP contains two GG domains, I thought it would be of interest to test the three corresponding amino acids of the first GG domain in the second one (D1244N, D1252N and R1265C), to elucidate whether the second GG domain is a hyaluronic acid binding domain too.

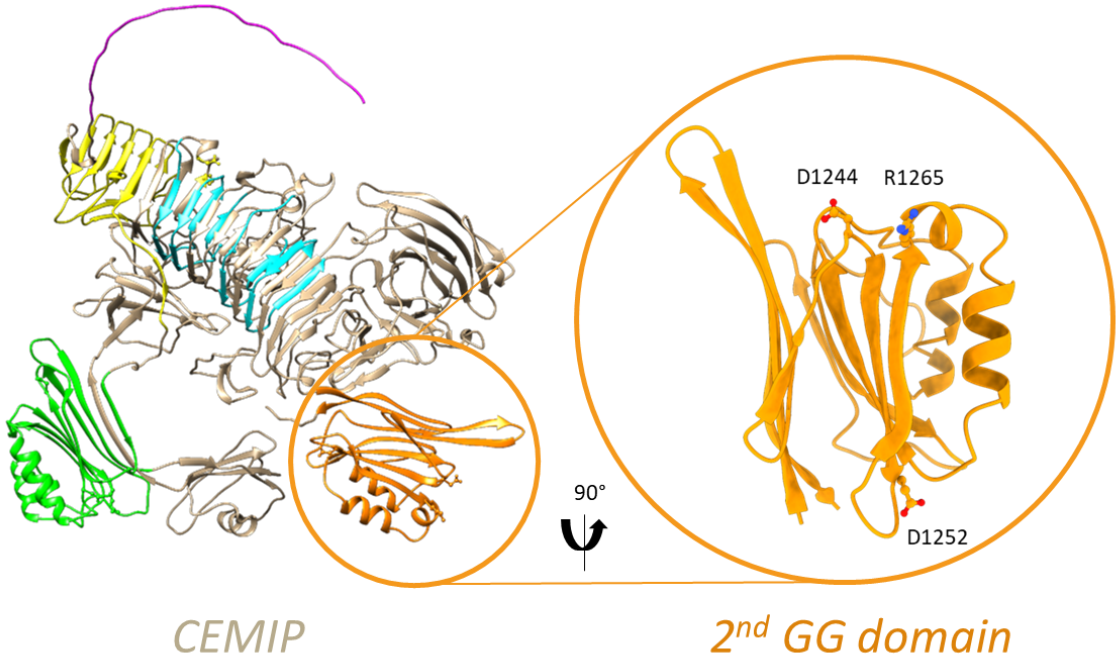


***Figure 54**: **Second GG domain amino acids of interest**. AlphaFold prediction of CEMIP with the second GG domain in orange. Courtesy of Margaux Heritier.*

The cloning of mutant D1252N was unsuccessful, but mutants D1244N and R1265C were successfully cloned and sequenced. When tested in an HA assay, mutants D1244N and R1265C present a positive HA-degrading profile (**Figure 55**), while their HA binding ability was reduced by 63% for D1244N and 66% for R1265C (**Figure 56**). This could be explained by the fact that these two amino acids are important for forming interactions that keep the HA in place and in a good orientation for the hydrolysis. The recognition site for HA binding is the Arg^{187}, and only if this binding occurs the HA can be fixed completely. Nevertheless, other interactions must be formed to keep the HA in the right position for the catalysis, and Asp^{1244} and Arg^{1265} seem to be involved in this role. Loss of the entire second GG domain compromises CEMIP catalytic activity as well as reducing the binding to hyaluronic acid by 95% (**Figure 55** and **Figure 56**). This is expected since the loss of an entire domain is a substantial change in the protein conformation. In this case, the entire loss of the second GG domain could mean that the HA is loose and not well sealed, and cannot be processed nor bound properly.

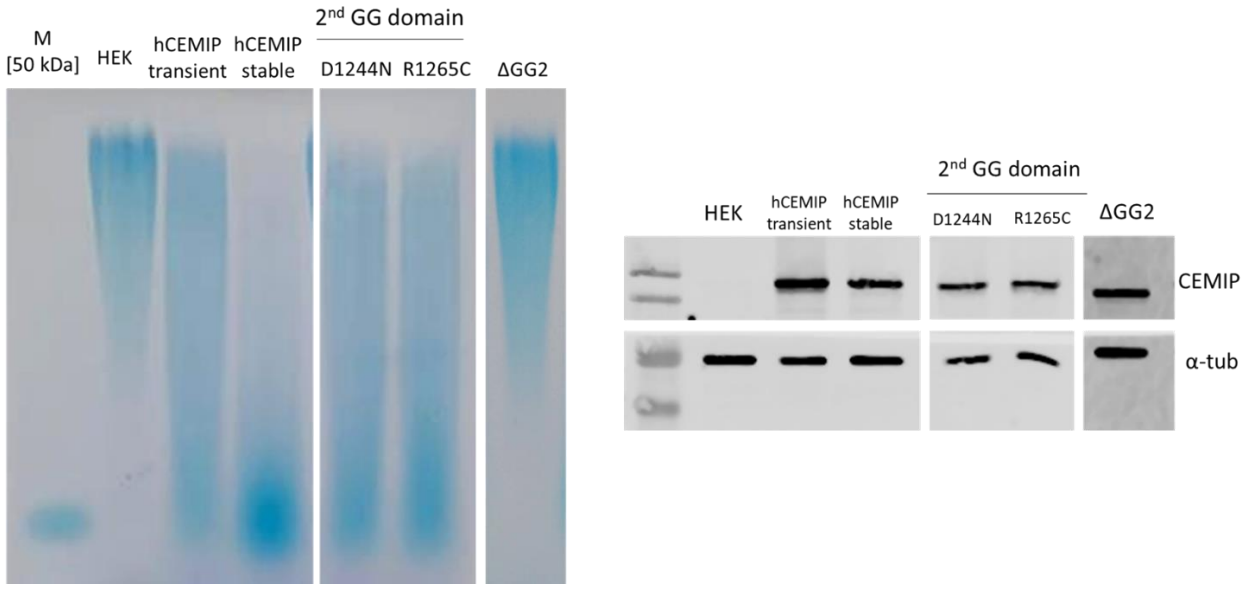


***Figure 55**: **HA assay and Western blot of mutants D1244N, R1265C and ΔGG2**. Left: HA assay of HEK293T, hCEMIP transient, hCEMIP stable clone (h88), and mutants transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

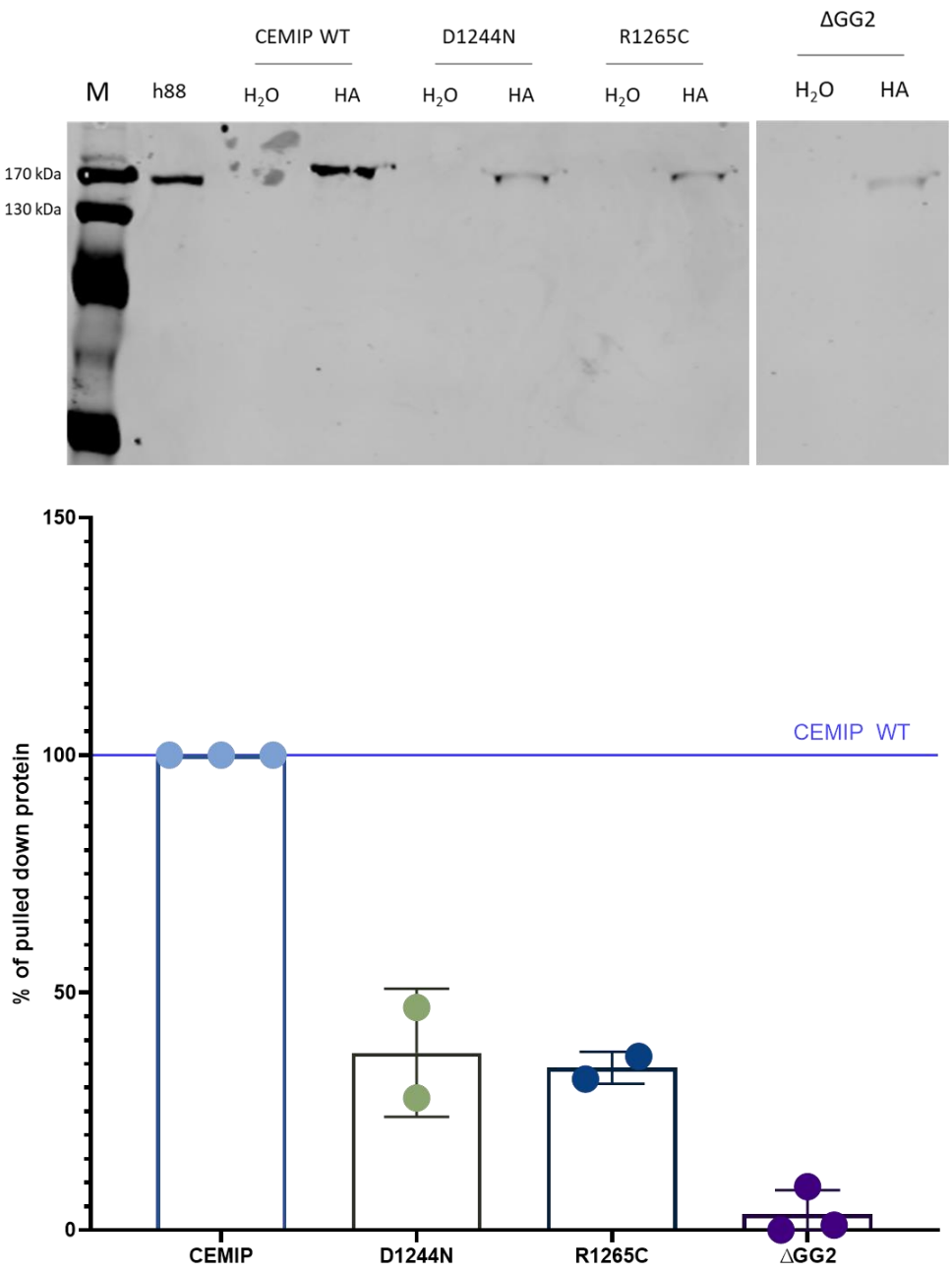


***Figure 56**: **Pulldown of mutants D1244N, R1265C and ΔGG2**. Top: Western blot membrane of the pulldown (with 1.5 MDa HA) with an anti-CEMIP antibody. Bottom: Band quantification standardized by the level of CEMIP (100%).*

6.6.3 His^{560}, Asp^{566} and His^{611} are part of CEMIP putative catalytic site

Computational analysis suggested the presence of two putative zinc-binding residues, H560 and H611[154]. Zinc is a known metal present in enzymes with glycosidic activity. It can coordinate sugar

and water as well, making it more acidic to accelerate the reaction. Moreover, a comparison with HYAL1 highlighted the interest in aspartic acids D564 and D566 in the ^{564}DVDER^{568} motif. These four amino acids are illustrated in **Figure 57**.

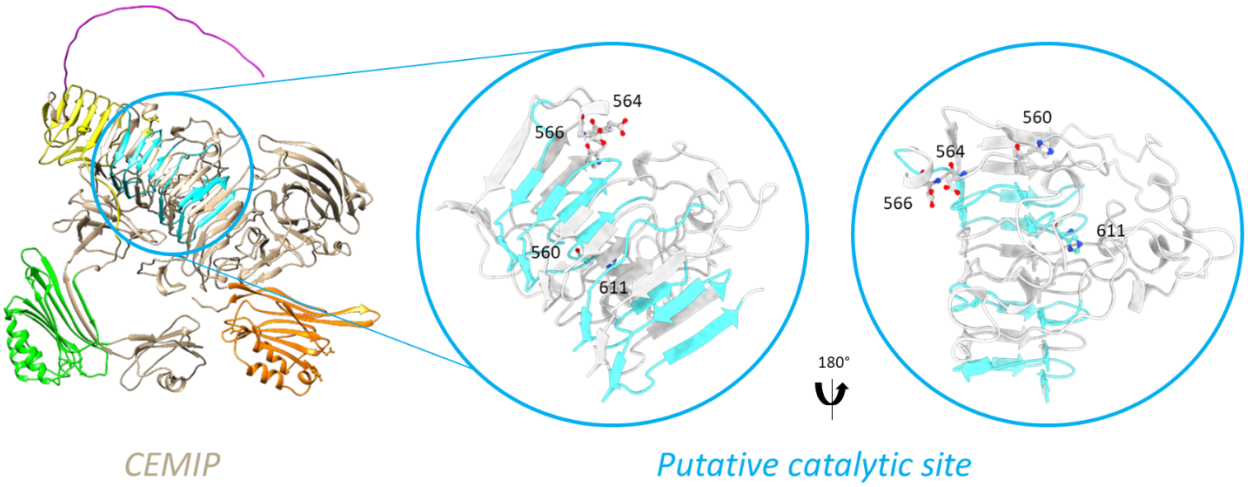


***Figure 57**: **PbH1 domain amino acids of interest**. AlphaFold prediction of CEMIP with the PbH1 repeats in blue. Amino acids H560, D564, D566 are found before the first PbH1 repeat; H611 is found in the second PbH1 repeat. Courtesy of Margaux Heritier.*

After cloning, mutants H560R, D564N, D566N and H611R were tested with an HA assay and a CPC pulldown. H560R, D566N and H611R, but not D564N, exhibited a defective catalytic activity (**Figure 58**).

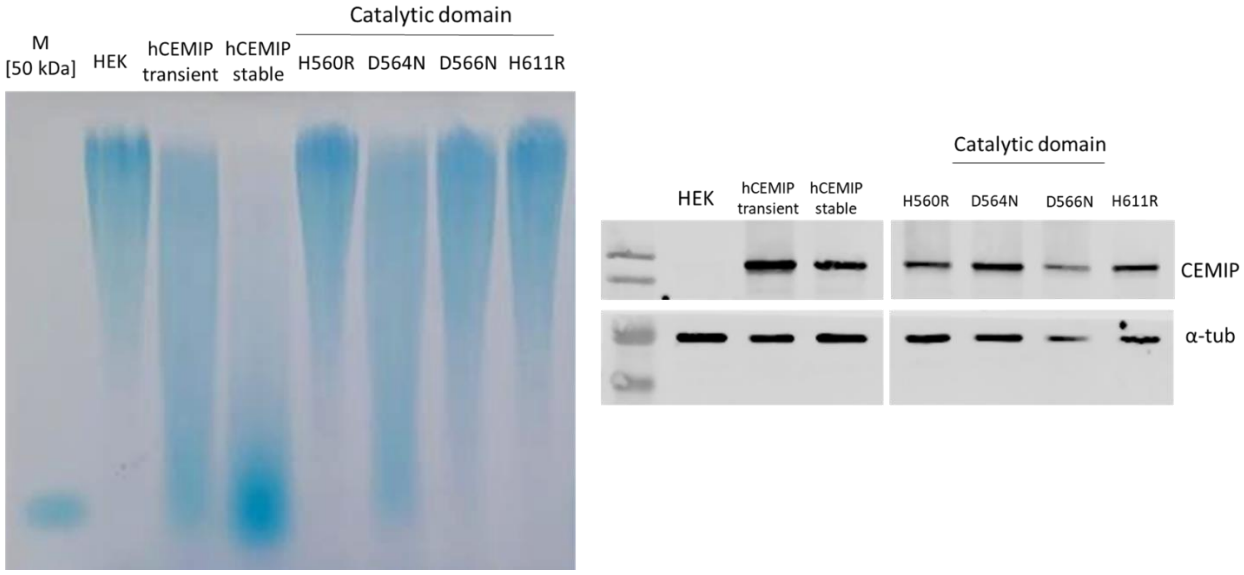


***Figure 58**: **HA assay and Western blot of mutants H560R, D564N, D566N and H611R**. Left: HA assay of HEK293T, hCEMIP transient, hCEMIP stable clone (h88), and mutants transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

CPC pulldown experiment helped clarify that mutants H560R, D566N and H611R conserve their HA binding ability (170%, 282% and 72%, respectively), indicating for the first time that amino acids H560, D566 and H611 belong to CEMIP's catalytic site, or that their integrity is crucial for the proper function of the catalytic site (**Figure 59**).

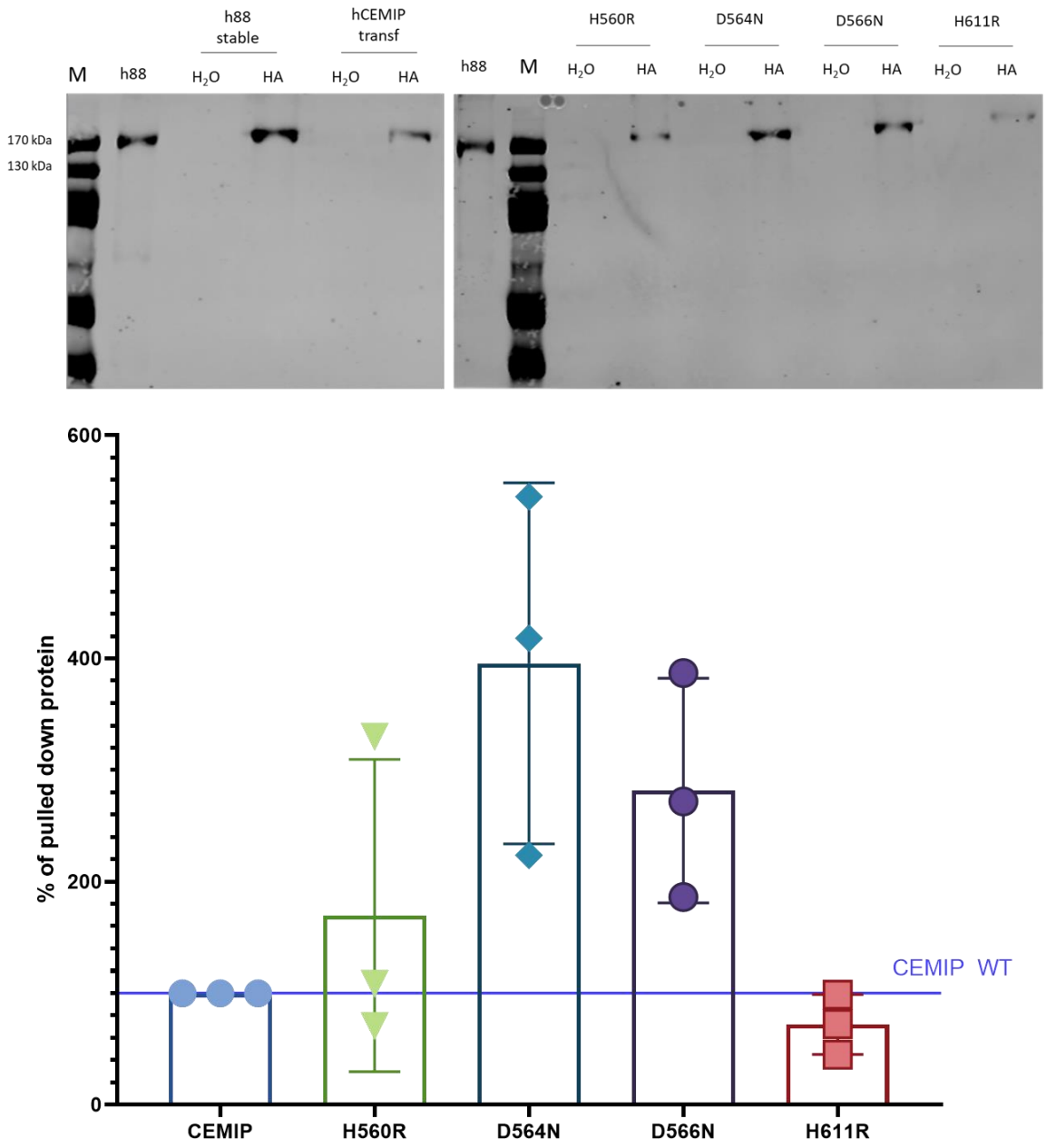


***Figure 59**: **Pulldown of mutants H560R, D564N, D566N and H611R**. Top: Western blot membrane of the pulldown (with 1.5 MDa HA) with an anti-CEMIP antibody. Bottom: Band quantification standardized by the level of CEMIP (100%).*

6.6.4 The G8 domain and Pbh1 repeats are essential for CEMIP function

In most cases, G8-containing proteins are predicted to be membrane-integral or secreted, and the G8 domain has predicted roles in extracellular ligand binding[43]. Several reports illustrate that CEMIP's binding partners interact with the G8 domain. According to Shostak *et al*., CEMIP binds

to EGFR and plexin A2 through the N-terminus and part of the G8 domain[83]. Furthermore, there is evidence that sCEMIP depends on membrane adhesion reached by the G8 domain binding to ANXA1[82]. Based on the SMART database (Simple Modular Architecture Research Tool) the predicted cellular role of the PbH1 (Parallel beta-helix repeats) domain is polysaccharide hydrolysis[155]. Already seeing the impact of deleting the second GG domain entirely (loss of HA processing and HA binding), I expect complete deletions of other CEMIP's domains to have a big impact on the conformational structure and thus activity of CEMIP. To confirm the hypothesis that G8 and PbH1 repeats are essential for CEMIP function, I cloned and tested mutants lacking these respective domains, ΔG8 and ΔPbH1, and different PbH1 repeats variations, ΔPbH1-4, ΔPbH1-3-4 and ΔPbH1-2-3-4.

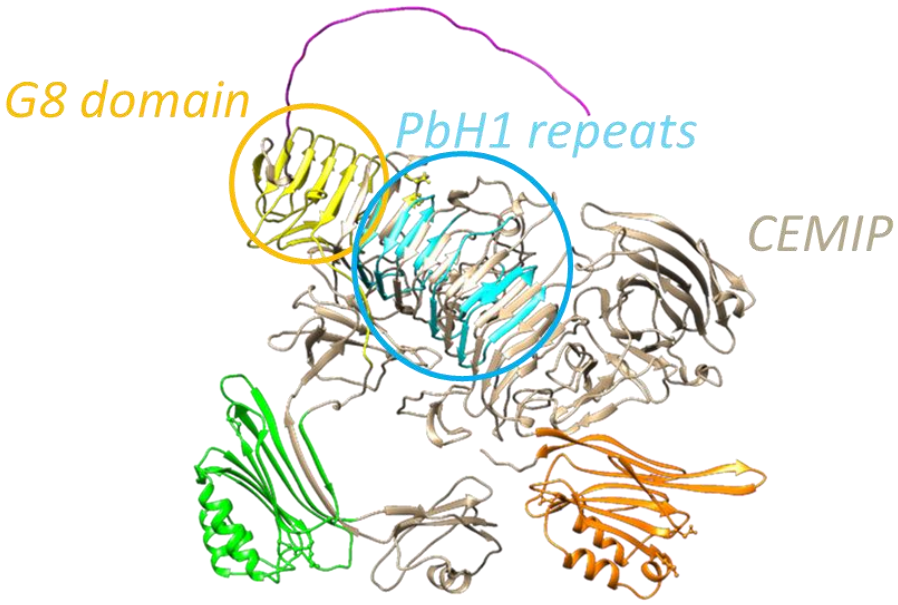


***Figure 60**: **The G8 domain and PbH1 repeats in CEMIP**. AlphaFold prediction of CEMIP with the G8 domain in yellow and the PbH1 repeats in blue.*

6.6.4.1 ΔG8 and ΔPbH1

As expected, mutants ΔG8 and ΔPbH1 lose the HA degrading ability (**Figure 61**).

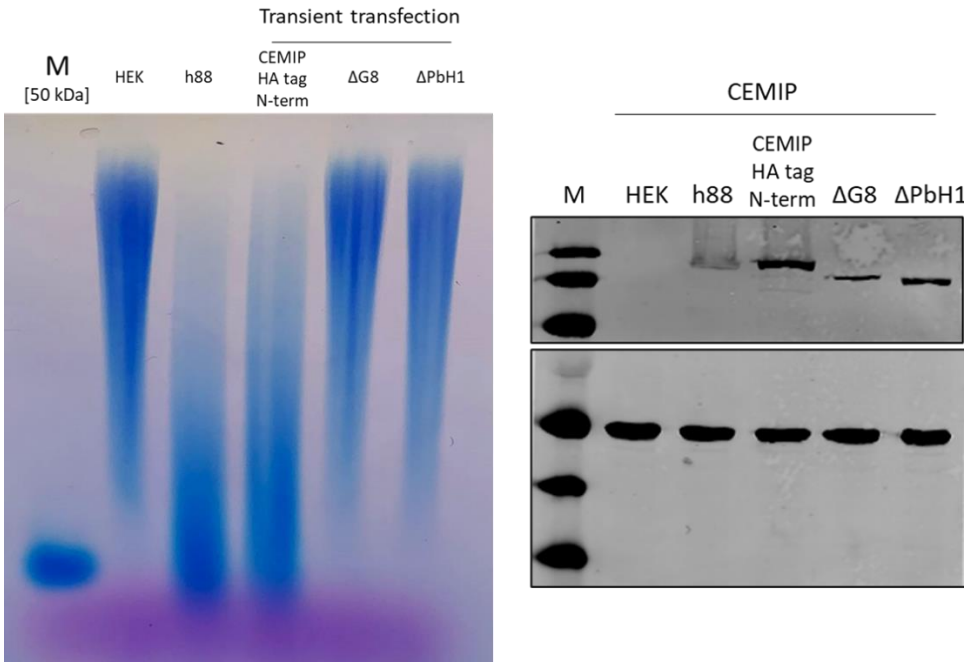


***Figure 61**: **HA assay and Western blot of mutants ΔG8 and ΔPbH1**. Left: HA assay of HEK293T, hCEMIP stable clone (h88), hCEMIP transient, and mutants transient. Right: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

6.6.4.2 PbH1 repeats variations

Since CEMIP contains four PbH1 repeats, I decided to clone three more mutants, each lacking one, two or three PbH1 repeats: ΔPbH1-4 lacking the last PbH1 repeat, ΔPbH1-3-4 lacking the last two repeats, and ΔPbH1-2-3-4 lacking the last three repeats. It is to be noted that amino acids H560 and D566 are found before the first PbH1 repeat, and AA H611 in the second one (**Figure 57**), making ΔPbH1-2-3-4 the only mutant with a compromised putative catalytic site.

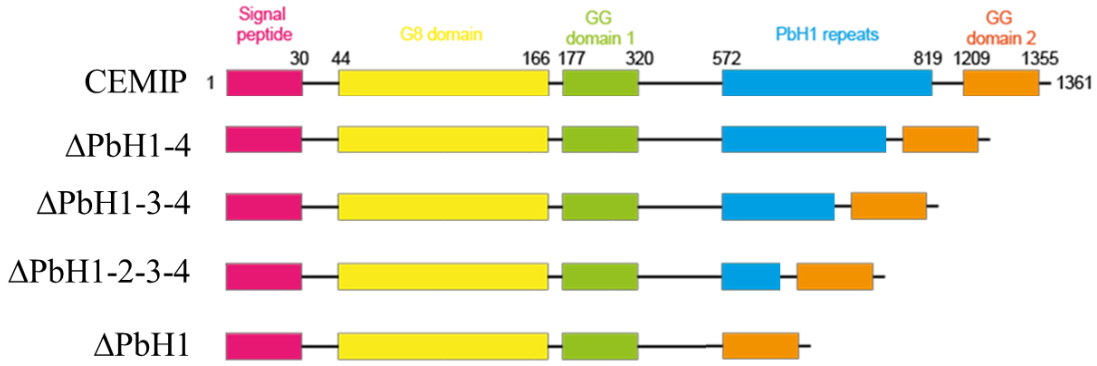


***Figure 62**: **PbH1 repeats mutants**. Wild-type CEMIP, ΔPbH1-4, ΔPbH1-3-4, ΔPbH1-2-3-4 and ΔPbH1.*

None of the PbH1 repeats mutants was capable to process hyaluronic acid (**Figure 63**).

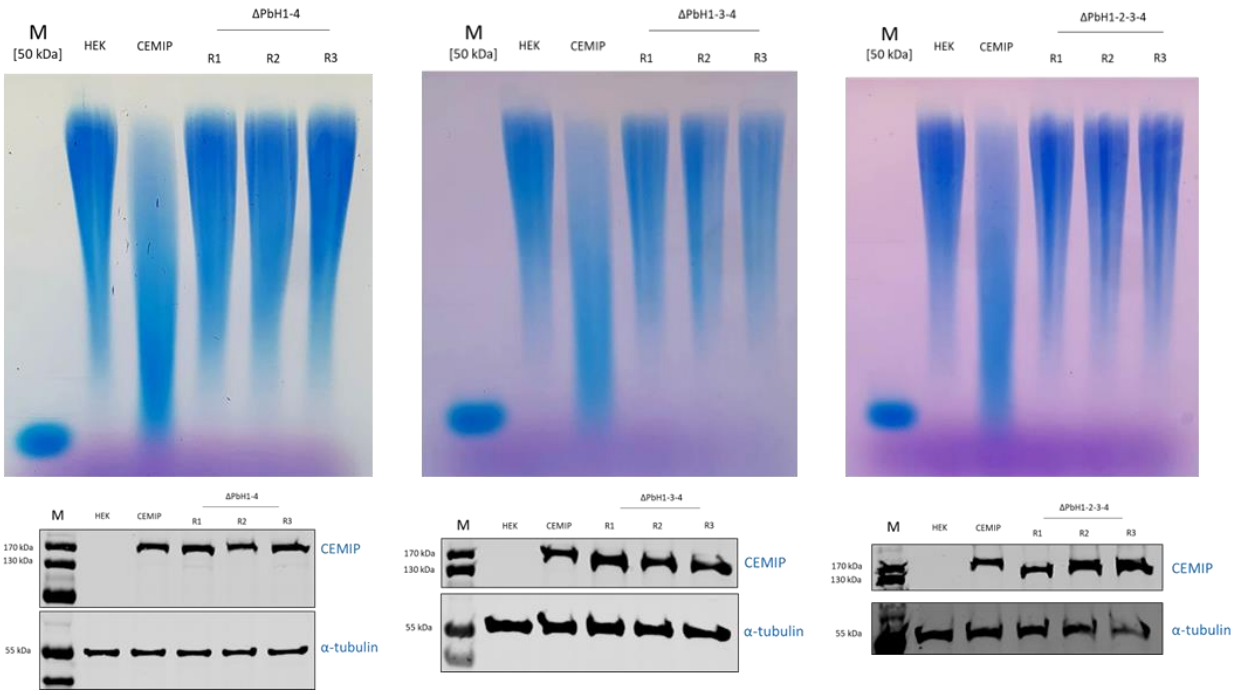


***Figure 63**: **HA assay and Western blot of mutants ΔPbH1-4, ΔPbH1-3-4 and ΔPbH1-2-3-4**. Top: HA assay of HEK293T, hCEMIP transient, and mutants transient (in replicates R1, R2 and R3). Bottom: Western blot of each sample with an anti-CEMIP antibody and anti-tubulin as loading control.*

Concerning the CPC pulldown of all these mutants, ΔG8 partially lost its ability of HA binding (-62%) and ΔPbH1 reduced its ability by 26% on average, despite showing a lot of variation. Interestingly, when looking at the PbH1 variations mutants, we can observe an increase in the capacity to bind HA, from ΔPbH1-2-3-4 (-90%) to ΔPbH1-3-4 (-69%) and ΔPbH1-4 (-35%) (**Figure 64**).

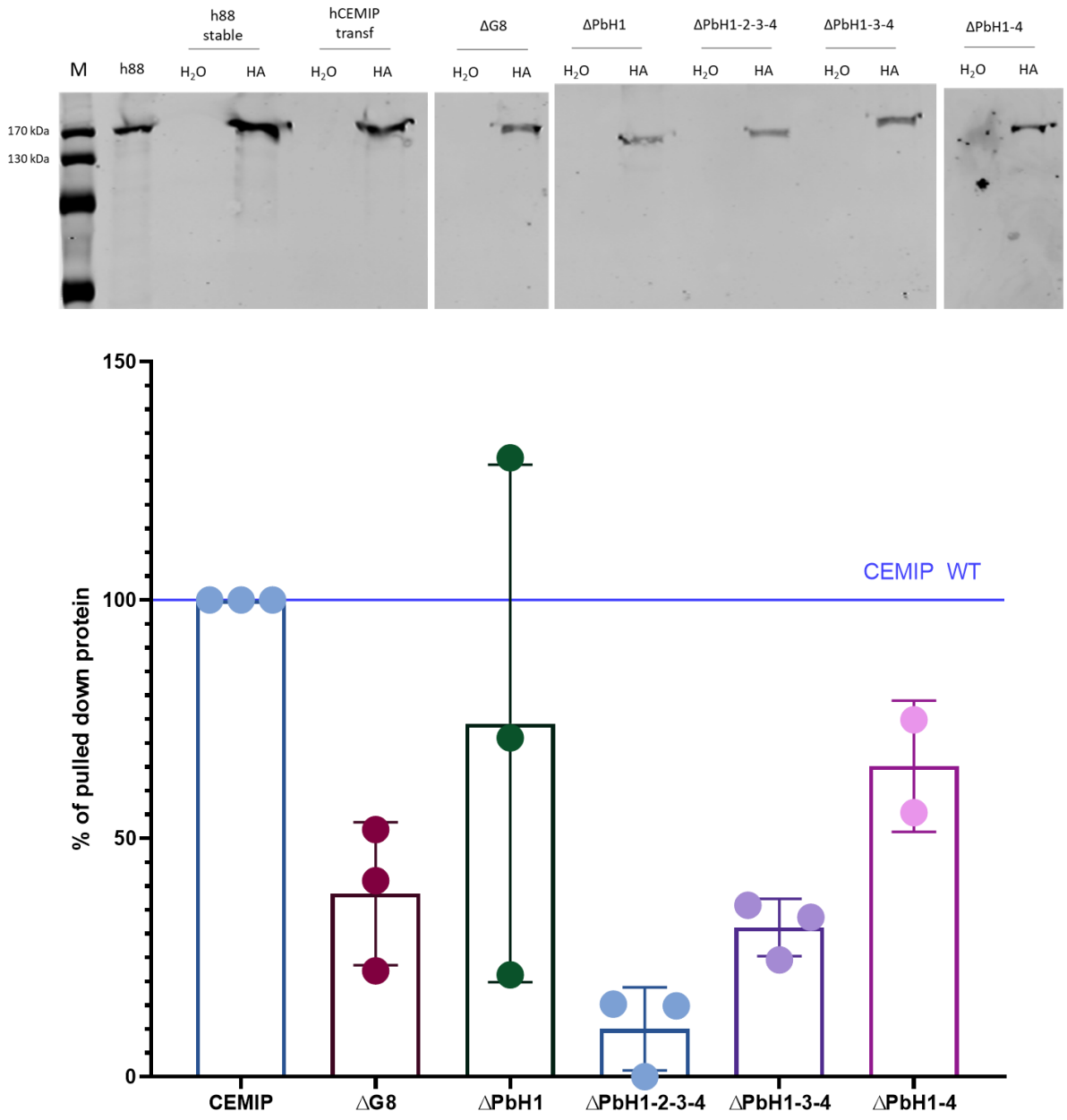


***Figure 64**: **Pulldown of mutants ΔG8, ΔPbH1, ΔPbH1-2-3-4, ΔPbH1-3-4 and ΔPbH1-4**. Top: Western blot membrane of the pulldown (with 1.5 MDa HA) with an anti-CEMIP antibody. Bottom: Band quantification standardized by the level of CEMIP (100%).*

6.6.5 Mutants' summary

The following table summarizes the mutants' ability to cleave or bind hyaluronic acid.

***Table 3**: Summary of mutants HA-cleaving and HA-binding ability.*

Mutant name	Is it cleaving HA?	Percentage of conserved HA-binding (%)
R187C	NO	19
D195N	YES	N/A
D208N	NO	194
H560R	NO	170
D564N	YES	396
D566N	NO	282
H611R	NO	72
D1244N	YES	37
R1265C	YES	34
ΔGG2	NO	3
ΔG8	NO	39
ΔPbH1	NO	74
ΔPbH1-2-3-4	NO	10
ΔPbH1-3-4	NO	31
ΔPbH1-4	NO	65
R187C + D208N	N/A	N/A

6.7 Discussion

The preliminary work of this project included the implementation of an HA assay and CPC pulldown as a readout for CEMIP hyaluronidase activity and CEMIP binding to HA, respectively. Although the HA assay was rapidly learned from collaborators, the CPC pulldown was more difficult to develop and took more time to be set up. Initially, the binding assay was attempted with the Dynabeads, but despite trying a comprehensive number of different conditions, this method never worked properly. In response to this technical issue the CPC pulldown method, suggested by collaborators (J.S. and A.S.) was developed. After some issues with protein precipitation, the correct protocol was established. This was a key step in the project allowing the discrimination of catalytically inactive mutants, distinguishing the catalytic mutants from the HA-binding mutants.
The cell line originating from HEK293T stably transfected with the human CEMIP gene was successfully created and kept under selection by puromycin-supplemented medium. After one year, the first clone lost the ability to depolymerize HA, requiring a second selection. This important information must be considered throughout the project. CEMIP activity must be tested at intervals in time to be sure that the cell line is functional for the experiments. This is why a positive control should also be included in each of the HA assays.
On another note, the analysis of HA fragments was carried out with size exclusion chromatography on a superose column. Due to time limitations, a deep investigation of this method was not conducted. Nevertheless, the quantification of HA samples from cells' supernatant was possible, and with some additional implementation, it could be exploited with the aim of a better quantification of HA fragments.

In this chapter, I demonstrated that CEMIP and sCEMIP are binding to hyaluronic acid fragments of 1'500 kDa, 100 kDa and 50 kDa. This is consistent with previous findings. Yoshida *et al.*, already observed that cellular CEMIP has specific binding to hyaluronic acid (1'452 kDa, 1'039 kDa, 219 kDa, 52 kDa and 28 kDa), and not to other GAGs such as chondroitin sulfate, dermatan sulfate, heparin and heparan sulfate[37].
My experiments show that sCEMIP is released from the cell in an inactive form, supporting the hypothesis that HA must somehow be internalized inside the cell. In a series of knock-out experiments, Yoshida's group proposed that this happens via the clathrin-coated pit pathway,

suggesting that HA is processed in acidic compartments before endosome-lysosome fusion[37]. The fact that sCEMIP binds to HA but doesn't cleave it supports this idea, excluding the notion that the catalysis is happening outside the cells. In order to test whether the reaction could happen at the membrane, I supplemented h88 membranes to HEK293T cells but observed no rescue in CEMIP activity. I then tested both h88 supernatant containing sCEMIP and membranes extracted from h88 on HEK293T cells transfected with ANXA1, since researchers reported that the HA-degrading activity of sCEMIP depends on its G8 domain's ability to adhere to the membrane in an annexin A1 (ANXA1) dependent way[82]. In both cases, CEMIP activity was not rescued by the presence of ANXA1. In their study, Zhang and colleagues also show that sCEMIP-containing culture medium supplemented with membrane fractions from HEK293-CEMIP presents no HA degrading activity, but instead, sCEMIP-containing culture medium supplemented with membrane fractions from rheumatoid arthritis fibroblast-like synoviocytes (RA FLS) presents CEMIP HA depolymerization. Given that the two cellular models, RA FLS and HEK293T-hCEMIP cells, are fundamentally different, they concluded that the presence of a certain molecule on RA FLS membranes must be required for sCEMIP-mediated HA degradation. After observing HA depolymerization by sCEMIP in media mixed with the membrane fractions of HEK293T-ANXA1 cells, they confirmed that this third factor is ANXA1 (combined with co-IP and IF experiments). Contrasting evidence emerged from my experiments, where CEMIP activity was not rescued by the presence of ANXA1. In order to shed light on this point, additional experiments are needed. An example could be replicating their study and supplement the h88 supernatant containing sCEMIP with membrane fractions from HEK293T cells transfected with ANXA1.

As CEMIP is inactive outside the cell nor at the membrane, the idea that CEMIP is internalized in endosomes in a clathrin-dependent manner seems to be the most appropriate. As Yoshida suggested, the HA could be processed into acidic compartments before endosome-lysosome fusion[37], emphasizing the fact that CEMIP activity could be pH-dependent. This could explain why both sCEMIP and membranous CEMIP are not active. Indeed, TMEM2 is ubiquitously expressed, is not active at a pH lower than 5 and could function as a "regular" hyaluronidase. On the other hand, CEMIP could function by internalizing and breaking down HA in a low pH environment, only when and if stimulated by external factors.

We know that the N-terminal portion of CEMIP acts as a cleavable signal sequence needed for the correct CEMIP translocation and CEMIP-mediated HA depolymerization[81]. Unexpectedly, during

some cloning manipulation, the importance of the C-term part of CEMIP was underlined, revealing that the C-term preservation is key for the hyaluronidase activity of CEMIP. When adding a tag (HA tag or EGFP) at the end of CEMIP, its HA degrading activity is lost. In 2018, Zhao and colleagues reported that CEMIP interacts with PP2A through the C-term domain (amino acids 881-1361), creating microtubule destabilization and CEMIP-mediated cell motility[156]. My results confirm that the C-term is extremely important for CEMIP. A possible hypothesis is that at the C-term of CEMIP there is the KKKL sequence, which can function as an ER retention signal, ensuring the correct folding of CEMIP before being secreted in the extracellular environment.

Loss of function experiments revealed Arg^{187} as an important HA binding spot. This amino acid is located at the tip of the first GG domain, in a solvent-exposed manner, and it is essential for the correct binding and consequent catalysis. When mutated into a cysteine, it loses its positive charge, and both HA cleaving and binding are lost, confirming that the arginine is essential for HA binding. Furthermore, I have evidence that amino acids Asp^{1244} and Arg^{1265} are key for hooking HA in the correct position. When mutated, CEMIP binding capability is reduced, but the HA processing is still present. This means that the correct binding is still happening through Arg^{187}, and that the catalytic site is not touched. Loss of the entire second GG domain compromises CEMIP catalytic activity as well as reducing the binding to hyaluronic acid by 95%. Entire loss of the second GG domain could mean that the HA is loose and not well sealed, and cannot be processed nor bound properly. Thus, I propose the notion that the second GG domain has an HA anchoring function. Surprisingly, Asp^{208} displays features of a catalytic amino acid profile: when Asp^{208} was mutated, the binding was conserved, while the HA cleavage was lost. It cannot be excluded that Asp^{208} is involved in the enzymatic activity, maybe as a water activator for the hydrolysis reaction (**Figure 65**). In fact, when the aspartic acid in this position is mutated into an arginine, the negative charge is completely lost, and the water's hydrogen is no attracted and activated anymore.

The putative catalytic site includes amino acids, belonging (or in proximity) to the PbH1 repeats region, predicted to be involved in the hydrolysis. This could also support the idea that the reaction is coordinated by a zinc ion: His^{560}, Asp^{566} and His^{611} could coordinate the Zn^{2+}, and put it in the right position for it to attract electrons away from the water nucleophilic attack's spot (**Figure 65**).

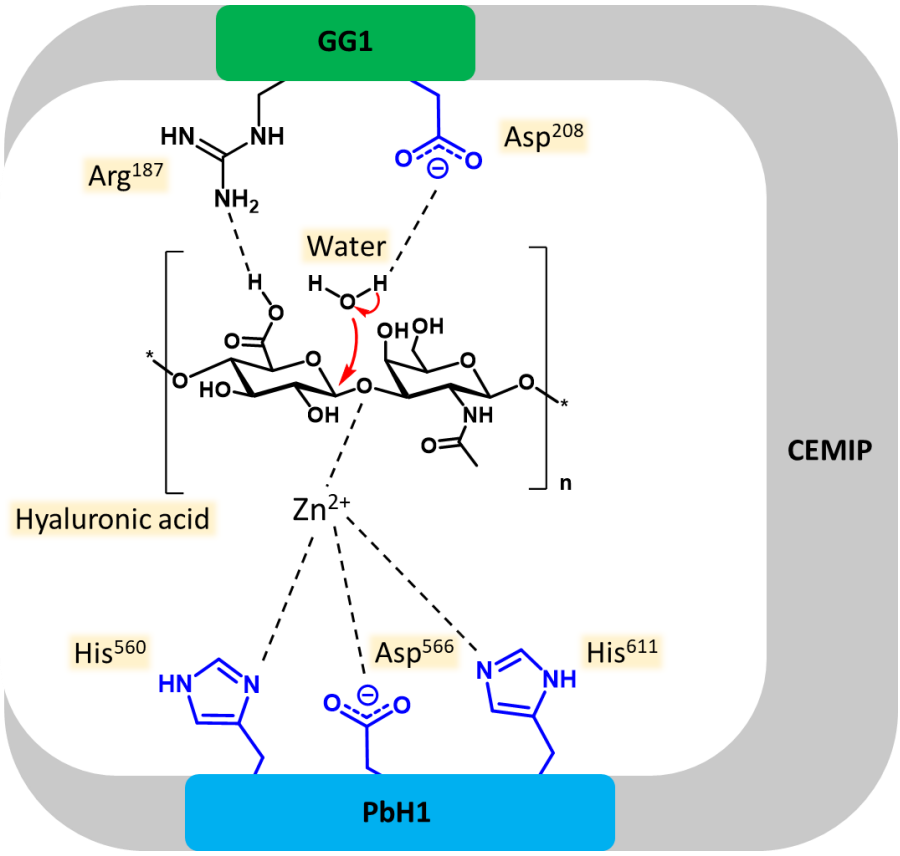


***Figure 65**: **Hypothetical mode of action of CEMIP**. In blue, catalytic amino acids; in black, binding amino acids. Courtesy of Sébastien Tardy.*

On another note, the importance of the G8 domain has been previously described. Most G8-containing proteins are predicted to be membrane-integral or secreted, and the G8 domain may be involved in extracellular ligand binding and catalysis[43]. Recent findings reported that the HA-degrading activity of sCEMIP depends on its G8 domain's ability to adhere to the membrane in an annexin A1 (ANXA1) dependent manner[82]. Moreover, CEMIP has been found to have interactions with EGFR and plexin A2 through the G8 domain[83]. When I deleted the G8 from CEMIP, the catalytic activity was completely lost, while the binding capacity was reduced by 62%, confirming the importance of this domain. Probably, the deletion is impairing the 3D structure of CEMIP and its function *in toto*, or maybe it is enabling the binding with a possible partner. Even if this thesis supports the hypothesis that CEMIP has intrinsic enzymatic activity, the idea that it is just recruiting another protein cannot be eliminated. Supplementary experiments can be conducted to eliminate any doubt. For example, a pulldown and identification by mass spectrometry of possible CEMIP binding partners could give a final answer.

According to the SMART database (Simple Modular Architecture Research Tool), the predicted cellular role of the PbH1 (parallel beta-helix repeats) domain is polysaccharide hydrolysis[155]. My results support this prediction since His^{560}, Asp^{566} and His^{611} belong to this beta-helix structure (**Figure 57**). When the repeats were deleted completely, the ΔPbH1 mutant lost the ability to process hyaluronic acid, while the pulldown showed the conservation of 74% of the binding capacity. Note that the pulldown results presented a high variation between replicates to have a solid interpretation. About the other three PbH1 mutants (ΔPbH1-4 lacking the last PbH1 repeat, ΔPbH1-3-4 lacking the last two repeats, and ΔPbH1-2-3-4 lacking the last three repeats), none of them were able to process hyaluronic acid correctly. Interestingly, the shorter the construct, the worse the HA binding, from ΔPbH1-2-3-4 (-90%) to ΔPbH1-3-4 (-69%) and ΔPbH1-4 (-35%). I could hypothesize that also the length of the beta-helix structure is important for the HA binding, but this is only a speculation since deletion of all PbH1 repeats does not present an important loss of HA binding capacity (-26%). We also have to keep in mind that big portion deletion affects the entire protein, both structurally and functionally, and this could be the reason why the catalytic activity of these mutants is always lost.

In terms of spatial orientation, AlphaFold predicts the G8 domain to constitute the beta-helix structure with the PbH1 repeats (**Figure 60**). The size of CEMIP-produced HA fragments might be dependent on the integrity of what seems to be a long tunnel formed by the G8 domain and the four PbH1 domains. I hypothesize that the "tunnel" structure formed by the beta-helix could serve as a "molecular ruler", explaining why CEMIP products are precise ~ 5-10 kDa HA fragments and not longer irregular products, produced for example by HYAL1 and 2. The recent report of the soluble TMEM2 structure is in support of this theory, endorsing the fact that the beta-helix region of CEMIP/TMEM2 is predicted correctly, also including the G8 domain[54].

My data confirm that the binding site is likely not near the PbH1 region, but rather falls under the competence of the two GG domains. The first one with its Arg^{187} binding the hyaluronic acid, and the second one as anchor of the long HA filament. In fact, Niu *et al.* also suggest that the Arg^{265} of the first GG domain, corresponding to the Arg^{187} of CEMIP, is in an equivalent critical position of a sugar-binding residue in a related domain of POMGNT1 (protein O-linked-mannose β-1,2-N-acetylglucosaminyltransferase 1)[54], consisting with my findings.

We should also consider that we are in the presence of a cellular system that overexpresses CEMIP constantly, and not an endogenous system. Various experiments should evaluate the activity of

CEMIP in an endogenous cell line. Additional loss and gain of function studies are required to investigate the functional relationship between the HA fragments produced by CEMIP, and the physiological and pathophysiological processes that are regulated by CEMIP. Moreover, the role of CEMIP and the HA fragments it produces might exert quite different effects in different cellular settings[93] and disease contexts[82,106,124,126].

6.8 Materials and methods

6.8.1 Cloning strategy and protocol

6.8.1.1 Site-directed mutagenesis

In order to insert point mutations in the hCEMIP cDNA plasmid, a mutagenic PCR was performed using overlapping primers containing different nucleotides necessary to encode a different amino acid at the desired position. Reactions were prepared following NEB protocols and enzymes (Q5® High-Fidelity DNA Polymerase). An example of the thermocycler program is shown in **Table 4**. Depending on the nature of the primers, the annealing temperature (step 3) was changed ranging between 55°C and 65°C. Extension time (step 4) depended on plasmid length and polymerase speed (30 seconds per kb). The number of cycles from step 4 to 2 varied from 16 to 25.

***Table 4**: PCR steps*

Step	Cycles	Time	Temperature
1	1	30 s	95°C
2	16-25	30 s	95°C
3		1 min	55°C
4		5 min	72°C
5	1	5 min	72°C

The reaction was digested with DpnI (NEB) to eliminate parental plasmid. Competent bacteria were transformed with heat shock using part of the reaction mix. Colonies were selected and amplified, and DNA was extracted using a NucleoSpin Plasmid kit (Macherey-Nagel). DNA was sequenced with Microsynth services by Sanger sequencing (annex **Figure 117**, **Figure 118**, **Figure 119**, **Figure 120**, **Figure 121**, **Figure 122** and **Figure 123**). Each plasmid and mutants

were sequenced again ensuring coverage of all CEMIP gene without other unwanted modifications.

***Table 5**: Primers for mutagenesis*

Name	**Sequence 5'→ 3'**
hCEMIP_R187C_F	GAAAGGAGCTGGGGCCACTGTGGAGTTATTG
hCEMIP_R187C_R	GACATGAACAATAACTCCACAGTGGCCCCAG
R187C_F_2	GCCACTGTGGAGTTATTGTTCATGTCATCGACCCCAAATC
R187C_R_2	TCCACAGTGGCCCCAGCTCCTTTCAAAAAAATAGC
hCEMIP_D195N_F	GAGTTATTGTTCATGTCATCAACCCCAAATCAGGC
hCEMIP_D195N_R	GACTGTGCCTGATTTGGGGTTGATGACATGAAC
hCEMIP_D208N_F	GTCATCCATTCTGACCGGTTTAACACCTATAGATC
hCEMIP_D208N_R	CTTTCTTGGATCTATAGGTGTTAAACCGGTCAGAATGG
D1244N_F	TACCCCAGTTCGGAGAATGGCATCCAGGTG
D1244N_R	CACCACCTGGATGCCATTCTCCGAACTGGG
D1252N_F	CAGGTGGTGGTGATTAACGGGAACCAAGGG
D1252N_R	GCGCCCTTGGTTCCCGTTAATCACCACCAC
R1265C_F	GTGAGCCACACGAGCTTCTGTAACTCCATTCTGCAA
R1265C_R	GCCTTGCAGAATGGAGTTACAGAAGCTCGTGTGG
hCEMIP-2GG-del-F	AAGAAGAAGAAGTTGTGAAC
hCEMIP-2GG-del-R	TGTTTTCAGCTGAGAACC
H560R_F	CAGTACCCGATTCACTTCCGCCTGGCCG
H560R_R	CGGCCAGGCGGAAGTGAATCGGGTACTG
D564N_F	CACTTCCACCTGGCCGGTAATGTAGACGAAAGGGGA
D564N_R	ACCTCCCCTTTCGTCTACATTACCGGCCAGG
D566N_F	CTGGCCGGTGATGTAAACGAAAGGGGAGGT
D566N_R	CCTCCCCTTTCGTTTACATCACCGGCC
H611R_F	TATAACTCTTTGGGCCGCTGCTTCTTCACG
H611R_R	TTCCGTGAAGAAGCAGCGGCCCAAAGAGTT

PbH1-deletionF	ATAACCTCCCCTTTCGTCTA
PbH1-deletionR	GGTGGAACCTTCCCGTATGA
PbH1-4-deletionF	GCGCAGCCAGGCCCCGTGGT
PbH1-3-deletionF	GTGCTCTGAATAACCTGGGG
PbH1-2-deletionF	GCCATGGACTGTGACGCAGC
G8deletionF	TTGCAACTCAGGGCTCTGGT
G8deletionR	ACCCTTCACCCAGGTGGCAT

***Table 6**: Primers for sequencing*

Name	**Sequence 5'→ 3'**
hCEMIP-for-PF	GCAGAGCTGGTTTAGTGAACC
hCEMIP-for-VB-PF1	AGACCCTTCACCCAGGTGG
hCEMIP-for-VB-PF2	AAGGCCAGGATTATAGGTTTGC
hCEMIP-for-VB-PF3	CTGCGTCACAGTCCATGGC
hCEMIP-for-VB-PF4	CCTACAAGAACCAGGACCAC
hCEMIP-for-VB-PF5	GTGCTATGCACAGATGTACATTC
hCEMIP-for-VB-PF6	CAAAGAACGCAGGCGTCAG
VB-PF3-rev	CGGTACCTGACACTGCGTC

6.8.2 HA-tag insertion

At the C-term, the HA-tag was inserted in the hCEMIP plasmid (VectorBuilder, vector ID VB191004-1084jkc) after the KKKL signal with a GSS linker at the beginning, placing the stop codon after the HA-tag. At the N-term, the tag was cloned between the signal sequence (amino acids 1-30) and the protein sequence with an AAA linker into a pcDNA3.1 after introducing the synthesized secretion signal with the necessary restriction sites (from GenScript).

***Table 7**: Primers used to insert HA-tag*

Name	**Sequence 5'→ 3'**
HA-tag-for-Cterm	GGGAAGCTTGGGGCAGACAGGGGTCTC
HA-tag-rev-Cterm	GGGGCCGGCCTCAAGCGTAATCTGGAACATCGTATGGGTA

	GCTAGAGCCCAACTTCTTCT
HA-tag-Nterm-for	GATCACACCGCGGCCGCGACAGTGGCTGCTGGGT
HA-tag-Nterm-rev	CTAGTGTGGCTCGAGTCACAACTTCTTCTTCTTCACCACAG

6.8.3 Cell culture and transfection

HEK293T cells were cultured at 37°C, 5% CO2 in DMEM (Gibco, 31966-021), 10% FCS, streptomycin, and penicillin. Clones h8 and h88 (HEK293T cells expressing hCEMIP) were cultured in the same way, with the addition of puromycin (1 µg/mL) to induce constant selection. Cells were transfected with Lipofectamine 3000 transfection reagent (ThermoFisher) or with PEI 25K at 1 mg/mL in PBS (Polysciences, 23966-1), following the manufacturer's instructions.

6.8.3.1 Clone selection

After transfection of HEK293T with hCEMIP plasmid (VectorBuilder, vector ID VB191004-1084jkc), cells were seeded with limited dilutions in four 96-well plates at 3 cells/mL and two 96-well plates at 10 cells/mL. After 3 weeks, new clones were amplified and tested with an HA assay and a WB. The clone with more CEMIP expression and HA depolymerization activity was selected.

6.8.4 Immunofluorescence

Cells were cultured on a cover slip treated with poly-ornithine in a 12-well plate at 250'000 cells/well. One day after transfection and/or HA treatment, cells were washed 3 times with PBS and fixed with PFA 4% for 20 minutes at room temperature. Primary incubation with anti-CEMIP antibody 12B5-2 1/50 in buffer (PBS, Triton, BSA) for 1h at 4°C was followed by washes (3 times with PBS) and secondary incubation with Alexa Fluor 488 dye 1/1'000 and Hoechst 1/5'000 for 1h at room temperature in the dark. After washes (3 times with PBS) cover slips were mounted on the slide with Mowiol solution and left dry for 48 h in the dark. Samples were analyzed with a confocal laser scanning microscopy Zeiss LSM700 at the Bioimaging Core Facility.

6.8.5 HA assay

Cells are seeded at 80'000 cells/well in a 24-well plate in the morning. After around 7h of incubation to allow cells to adhere to the plate, 50 µg HA were added (Lifecore, HA15M, 1.5 MDa, 5 mg/mL aqueous solution) for a total of 500 µL per well. After 48h incubation, supernatants were collected and centrifuged at 2'000 g for 5 minutes to pellet cell debris, then collected again. HA was purified by digesting proteins for 4 hours at 60°C with proteinase K (Sigma, 1.24568) in 100 mM ammonium acetate, 0.1% SDS. 400 uL EtOH (99%) were added and samples were incubated overnight at –20°C. The day after samples were centrifuged at 10'000g for 10 min, washed with 500 µL 70% EtOH, and dried for 1h at room temperature. HA was resuspended in 15 µL water and 4 µL dye (2M sucrose in 1x TAE buffer + 0.25% w/v bromophenol blue sodium salt).

Samples were loaded in an agarose gel 1% in TAE buffer and run for 2h15 at 50 V. A marker of 50 kDa HA was used in order to compare different gels (Echelon Biosciences, HYA-0050). The gel was washed for 1h in 30% EtOH, and then left in a colored solution (Stains All, Sigma, E9379) overnight at room temperature in the dark. The gel was finally washed in the dark in water for 4h and exposed to light for 15-30 minutes to remove unspecific stains. As a positive CEMIP inhibition control sHA was used (KS 1552, modified hyaluronan, MW = 108'000 g/mol).

6.8.6 Cell lysis

Frozen cells were lysed using RIPA buffer (NaCl 150 mM, 0.1% SDS, 0.5% deoxycholic acid, 1% NP-40, 50 mM Tris, pH 8). Extracts were lysed on ice for 30 minutes, then sonicated for 10 seconds and centrifuged at 4°C at 10'000g for 10 minutes to remove cell debris. Supernatants were collected and stored at -20°C.

6.8.7 Protein quantification

Protein quantification was performed with a BCA, bicinchonic acid assay, which allows the determination of total protein concentration in a solution. The BCA standard curve was performed with a BSA stock (4 ug/uL). Thermo Scientific Pierce BCA Protein Assay Kit was used and the plate was read with a CLARIOstar machine (BMG Labtech) at 562 nm.

6.8.8 Pulldown

Pulldown experiments were performed to assess the binding ability of CEMIP (or its mutants) to HA. After encountering too many issues using ThermoFisher's Dynabeads magnetic separation technology, an adaptation from the CPC pulldown protocol from Lee[147] and Sleeman[148] was developed.

6.8.8.1 Dynabeads

50 µL of protein solution were incubated with 100 µL of biotinylated HA (Echelon Biosciences, HYA-B250-200, 250 kDa, 0.5 mg/mL in water) at 37°C for 1h30. 15 µL of Dynabeads (Dynabeads M-280 Streptavidin, ThermoFisher, 11205D) were added to the sample, followed by a 4h incubation at 4°C, rotating. After centrifugation for 30 s at 2'000 g, tubes were inserted into the DynaMag slots (DynaMag-2 Magnet, ThermoFisher, 12321D) and incubated for 2 minutes. The supernatant was discarded and pellets were resuspended in 100 µL of PBS. The washing step was repeated a second time. Dynabeads were resuspended in 15 µL water and 5 µL SDS 4x blue Laemmli with DTT and heated for 8 minutes at 98°C. Western blot technique was performed with an anti-CEMIP antibody (see **Table 8** for details).

6.8.8.2 CPC

135 µL of protein solution were incubated with 50 µL of HA (Lifecore, HA15M or HA100K, 1.5 MDa or 100 kDa, 1 mg/mL in water) or water for 1h at 37°C. 460 µL of 1.4% CPC aqueous solution (Sigma, C-9002) were added and samples were incubated for 1h at 37°C. After 10 minutes of centrifugation at 13'000g, supernatants were discarded. Samples were washed 3 times with 1 mL 30 mM NaCl, 1% CPC solution (13'000 g, 2 min). Pellets were dissolved in 50 µL Laemmli blue with DTT and heated at 95°C for 10 minutes. Western blot technique was performed with an anti-CEMIP antibody (see **Table 8** for details) and bands were quantified using ImageJ and GraphPad Prism.

6.8.9 Western blot

Samples were prepared with a volume of protein calculated from the BCA results and completed with lysis buffer and Laemmli blue 1x with DTT. Before loading, samples were heated at 95°C

for 5 minutes to denature proteins. Samples were loaded in 10% polyacrylamide gels previously prepared. Electrophoresis was performed at 80 V for 20 minutes and 130 V for 1h30 with BIO-RAD materials and protocols. The running buffer contained 25 mM Tris, 190 mM glycine and 0.1% SDS. Transfer from SDS-PAGE gel to nitrocellulose membrane was performed with a run of 65 V for 2h. In order to visualize the proteins and the membranes, red Ponceau S was used. Membranes were left for 2h in a blocking solution (5% milk or 3% BSA in TBS-T) at room temperature. Membranes were incubated overnight at 4°C with the specific antibody solution (see **Table 8** for more information). The day after, membranes were washed 3 times with TBS-T for 20 minutes. Incubation with secondary antibodies (LI-COR IRDye® Secondary Antibodies) was performed for 1h at room temperature, followed by 3 washes in TBS-T for 20 minutes each. Membranes were analyzed with an Odyssey CLx Imager (LI-COR).

***Table 8**: Primary antibodies used for Western blot.*

Antibody	MW (target)	Supplier	Reference	Blocking	Species	WB dilution
Anti-TMEM2	154 kDa	Sigma SABI	SAB2105088	3% BSA	Rabbit	1/1000
Anti-CEMIP	153 kDa	ProteinTech	21129-1-AP	5% Milk	Rabbit	1/1000
Anti-CEMIP 12B5-2	153 kDa	In-House	/	5% Milk	Mouse	1/750
Anti-ANXA1	39 kDa			3% BSA	Rabbit	
Anti-α-Tubulin	55 kDa	Sigma	/	3% BSA	Mouse	1/2000
Anti-flag (DYKDDDDK)	/	GenScript	A00187	3% BSA	Mouse	1/1000
Anti-CD44	80 kDa	?	ASML1.1	5% Milk	Mouse	1/1000
Anti HA-tag (YPYDVPDYA)	/	Cell Signaling	#2367	5% Milk	Mouse	1/1000
Anti-HisTag	/	Qiagen	34660	5% Milk	Mouse	1/1000

6.8.10 SEC

Cells supernatants were treated as described in the HA assay part (see 6.8.5), with the difference of being resuspended in the mobile phase (PBS) at a concentration of 0.134 mg/mL (40 µg of HA in 300 uL of PBS). Samples were analyzed at 0.65 mL/min with a Superose 6 Increase 10/300 GL column from Cytiva in a Postnova AF2000 AF4 system coupled with a MALS (multiple angle

light scattering) and RI (refractive index) detectors. Data were collected with Software NovaFFF AF200 and analyzed with GraphPad Prism.

7 CHAPTER II – INHIBITION OF CEMIP

As mentioned in section 4.2, after observing a reduction in renal damage in DDR1 deletion mouse models, a series of DDR1 inhibitors was developed but was found clinically irrelevant due to the unacceptable side effects caused by the inhibition of DDR2. In order to overcome the limitation of the DDR1 inhibitors, collaborators identified candidate genes involved in the DDR1 downstream mode of action, and found that in the presence of collagen, DDR1 selectively induces the production of CEMIP. For this reason, in this chapter, I aim to evaluate the potential of CEMIP inhibitions of a set of small molecules and an anti-CEMIP antibody. It should be noted that only one paper reports CEMIP inhibition, indicating that extract from the plant *Geranium thunbergia* down-regulates CEMIP gene expression, inhibits CEMIP HA degradation in HEK293-CEMIP *in vitro* and has anti-wrinkle activity *in vivo*[157].

7.1 Screening of small molecules against CEMIP

Design and synthesis of small molecules: The identification of a lead compound for the development of CEMIP inhibitors was conducted in the Computational Structural Biology group (Andrea Cavalli) of the Institute of Research in Biomedicine of Bellinzona, using a combined structure-based and ligand-based approach. First, analysis of the CEMIP sequence homology and information available in crystal structures of topologically related proteins facilitated the creation of a structural homology model for the GG domain of CEMIP, then updated when AlphaFold was published[39]. In fact, as previously mentioned, this region contains the R187 residue which mutation was reported to be responsible for non-syndromic hearing loss, and is hypothesized to be essential for CEMIP activity[47]. Considering this experimental observation, they hypothesized that molecules able to bind close to this residue could impair the enzymatic activity.

Preliminary modelling of individual domains using distant homologous search[158,159] and co-evolutionary restraints[160] allowed the identification of potential druggable spots on the CEMIP protein. Blind computational docking with AUTODOCK[161] of an HA fragment and of the approximately 50’000 molecules of the EPFL NCI compound library resulted in a selection of compounds capable of interacting with CEMIP (**Figure 66**). Interestingly, all molecules bind to

the same spot in the GG domain and strongly interact with both R187 and D208, amino acids shown to be crucial for CEMIP activity[37].

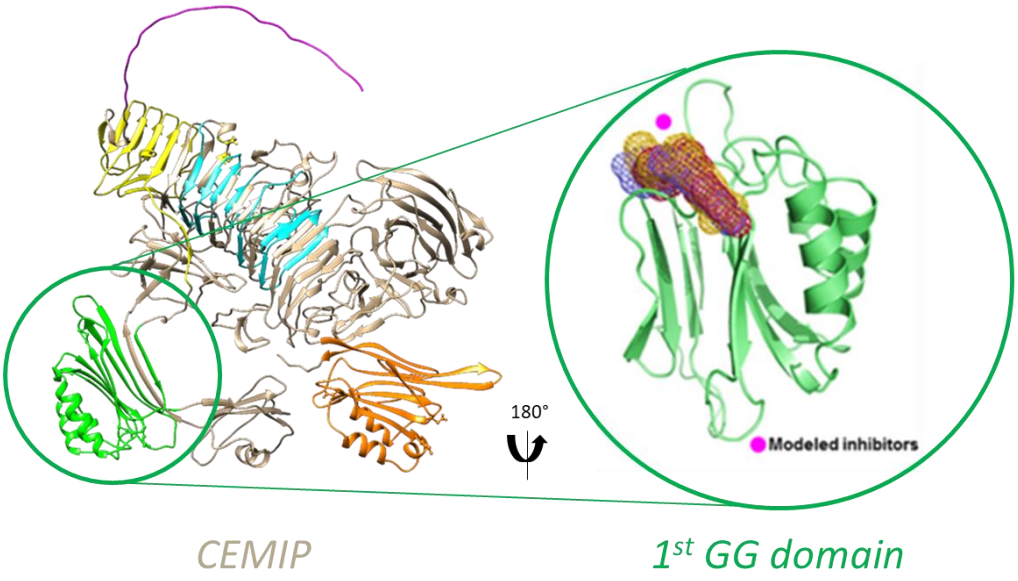


***Figure 66**: **Docking of CEMIP inhibitors on the GG domain**. Alphafold prediction of CEMIP with the first GG domain in green. Inhibitors (in pink) were modeled on the tip of the first GG domain (region containing R187 and D208).*

Then, to explore a wider chemical space, they utilized the selected compounds for screening on a library consisting of 1.5 million commercially available compounds. The screening involved two steps: (1) identification of molecules similar to the previously identified hits based on fingerprint similarity and (2) refinement of the screening through docking calculations. Following the screening process, 21 molecules were obtained from vendors and subjected to testing for their ability to inhibit CEMIP activity. The experiments revealed the inhibitory activity of one compound, prompting them to conduct both SPR (surface plasmon resonance) and NMR (nuclear magnetic resonance) experiments to validate the interaction of the molecules with the CEMIP GG domain. Both experiments confirmed the interaction, with the SPR experiment indicating a K_d (dissociation constant) greater than 200 µM.

Subsequently, they initiated an optimization process guided by structure-based calculations, that involved the IRB group and some organic chemists from the University of Genova, which led to the design of more than 20 molecules. This effort allowed them to establish some structure-activity

relationships (SAR) with improved affinity for the target (K_d < 50 µM) and demonstrated good inhibitory activity.

In parallel, a DEL screening was performed in collaboration with WuXi AppTec, allowing the identification of three hit molecules called WX-1, WX-2 and WX-3.

7.1.1 sHA inhibits CEMIP and prevents HA binding

As mentioned above, sulfated hyaluronic acid, a synthetic sulfated polysaccharide that inhibits the hyaluronidase HYAL1 with an IC_{50} of 0.4 – 0.8 µM[149], has been shown to have preventive and therapeutic effects on experimental mesangial proliferative GN[38], indicating that hyaluronidase activity plays a role in GN. Moreover, recent studies confirmed that sulfated HA inhibits CEMIP hyaluronidase activity with an IC_{50} of 5-20 nM[128]. In the frame of this thesis, sHA can be used as a positive control to inhibit CEMIP. In an experimental set-up equal to the HA assay, sHA presented no effect on cell proliferation and toxicity when tested on h88 cells (**Figure 67**).

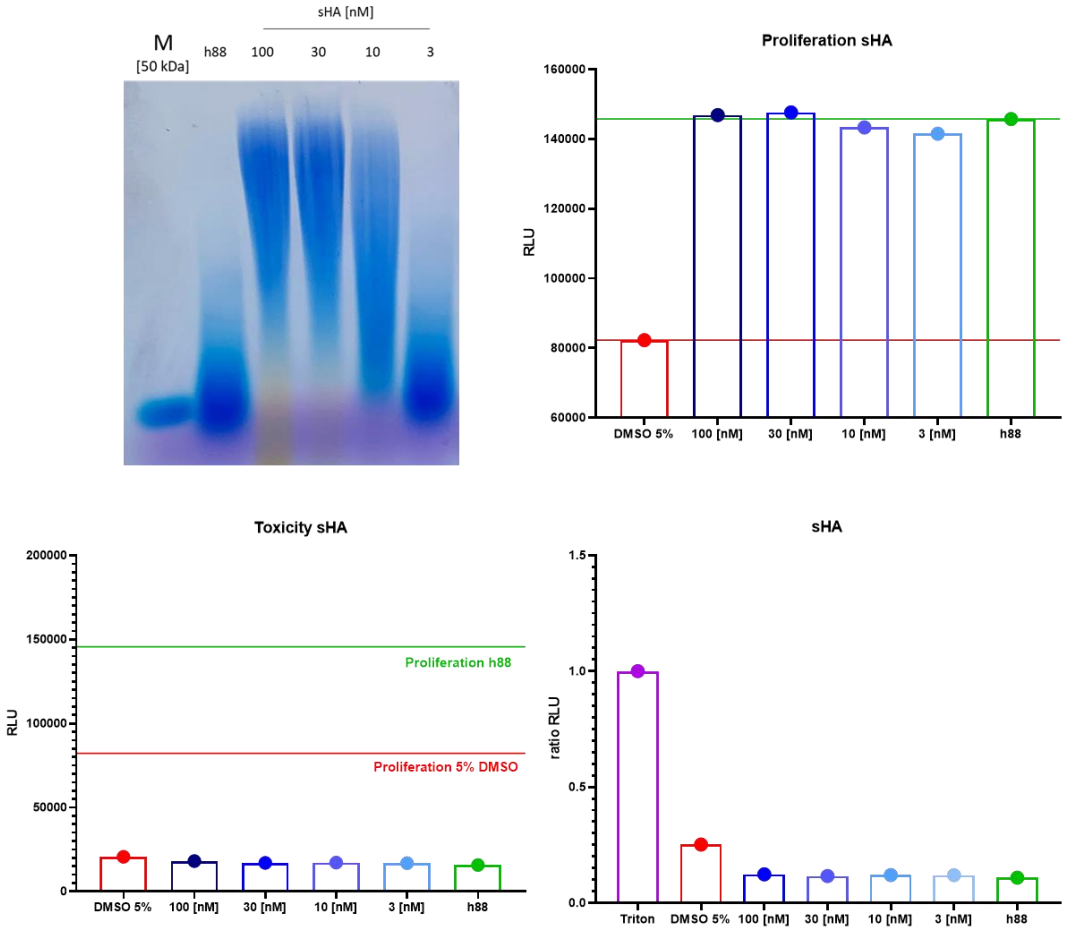


***Figure 67**: **sHA inhibits CEMIP**. HA assay of the sHA inhibition on h88 cells (top left), proliferation assay (top right), toxicity assay (bottom left) and relative toxicity (bottom right; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. h88 = HEK293T-hCEMIP stable clone.*

Until now, no clear evidence of sHA inhibition mode of action has been reported. A CPC pulldown of h88 lysate in the presence of sHA (10 times the K_d, 1 µM), added 30 minutes before HA, showed a loss of binding, indicating that sHA binds to CEMIP's HA binding site (**Figure 68**).

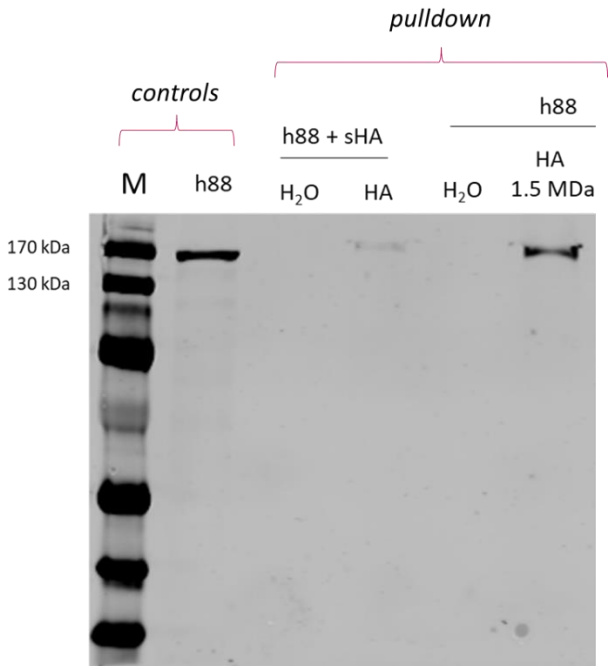


***Figure 68**: **sHA prevents CEMIP binding to HA**. First lane after the marker represents the control (no pulldown). Other lanes represent the pulldown with and without sHA, and water as control. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP stable clone.*

7.1.2 Hit molecules inhibit CEMIP HA degrading ability at 50-125 µM without impacting cell proliferation and without cell toxicity

At first, when working on the creation of the hCEMIP stable cell line, the first set of molecules was tested on the mCEMIP cell line. At an earlier stage, several conditions were tested in order to frame the most suited experiment set-up. Molecules were tested with or without an incubation time of 2h before HA, in the context of a transient transfection, with or without a transfection agent (Lipofectamine, ThermoFisher), and with different HA incubation times. After these assessments, it was agreed upon a 2h molecule incubation, without transfection agent, on the stable clone with a 72h total HA incubation time for the murine cone and 48h for the human one. Regarding the effect of DMSO, an MTT test confirmed that molecules should be tested in 1% DMSO or less since cell proliferation at higher DMSO concentrations was strongly reduced by 75% or more (**Figure 69**). All information regarding the name, the molecular weight, and the K_d of the tested molecules can be found in **Table 9** at the end of this section.

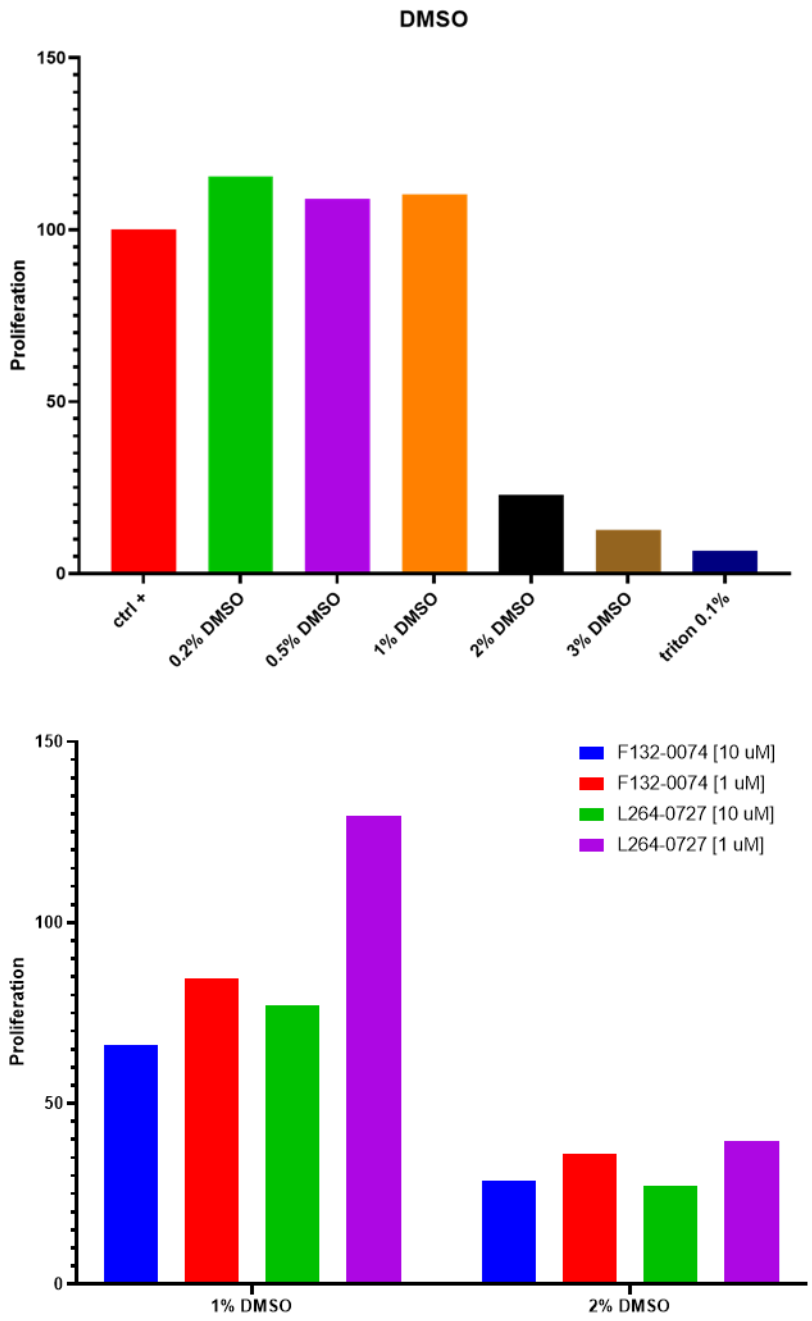


***Figure 69**: **MTT test shows the effect of DMSO on the mCEMIP cell line**. Top: Effect of DMSO on mCEMIP cells. Bottom: Difference in proliferation of two molecules at 1% and 2% DMSO. Molecules should be tested in less than 1% DMSO.*

The first set of molecules received included STL439159, STL441845, STK038497, STK059275, STK175734 and STK496096. None of these molecules inhibited HA depolymerization of the mCEMIP cells at a concentration range around their presumed K_d, between 0.1 µM and 100 µM

(**Figure 70**). A second set of molecules, F132-0074, L264-0727 and d-p-Y, showed no sign of hCEMIP inhibition (**Figure 71**).

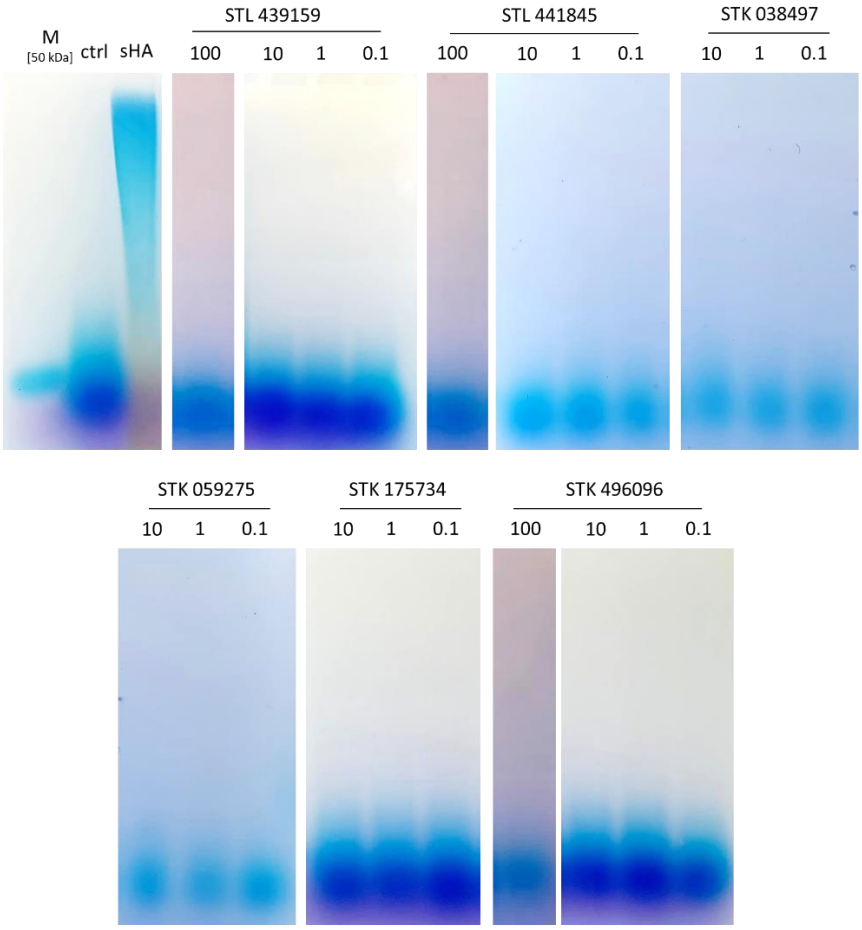


***Figure 70**: **HA assay of molecules STL439159, STL441845, STK038497, STK059275, STK175734 and STK496096**. Control with HEK293T-hCEMIP stable clone, inhibition control with sHA at 0.1 µM.*

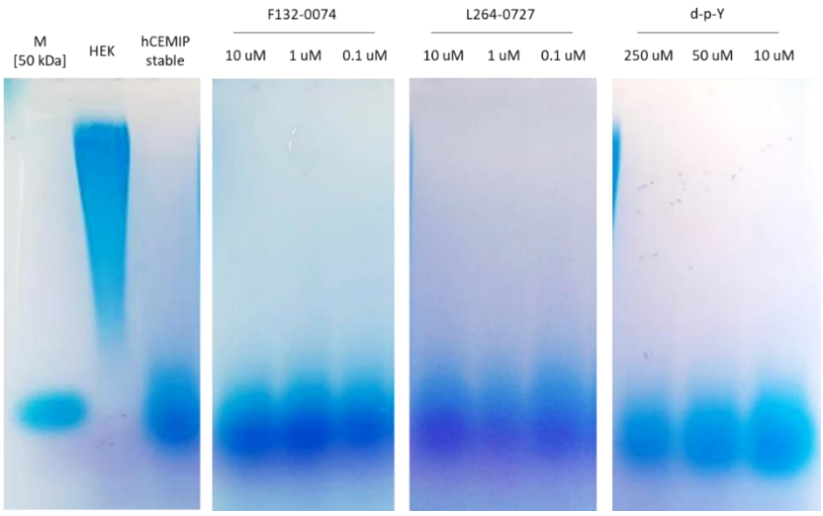


***Figure 71**: **HA assay of molecules F132-0074, L264-0727 and d-p-Y**. Control with HEK293T cells and HEK293T-hCEMIP stable clone.*

Molecule CEMIP-113 didn't influence HA degradation, while molecule CEMIP-119 showed some inhibition at 250 µM (**Figure 72**).

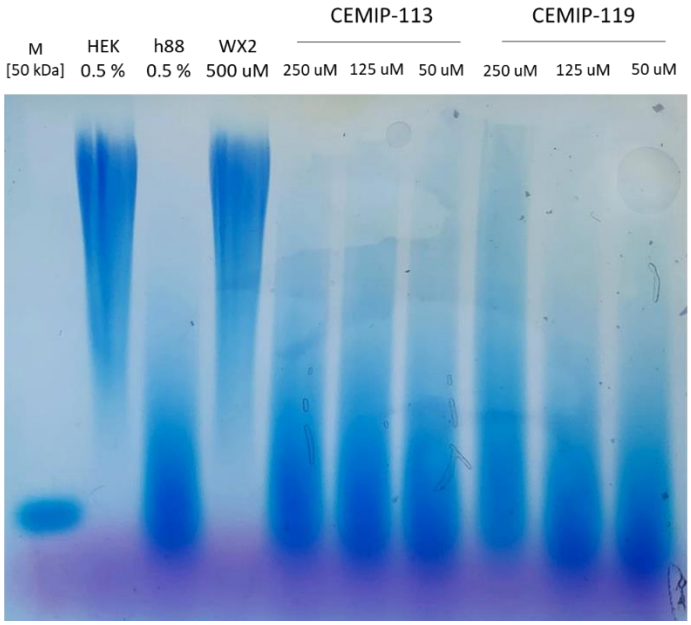


***Figure 72**: **HA assay of molecules CEMIP-113 and CEMIP-119**. Control with HEK293T 0.5% DMSO, HEK293T-hCEMIP stable clone (h88) 0.5% DMSO, inhibition control with molecule WX2 at 500 µM.*

Other inhibitors coming from DEL screening from WuXi were tested, called WX-1, WX-2 and WX-3. WX-2 showed a complete inhibition at 50 µM and 125 µM and a partial inhibition at 1 µM and 5 µM, while WX-1 and WX-3 showed a reduced HA degradation at around 50-125 µM (**Figure 73**).

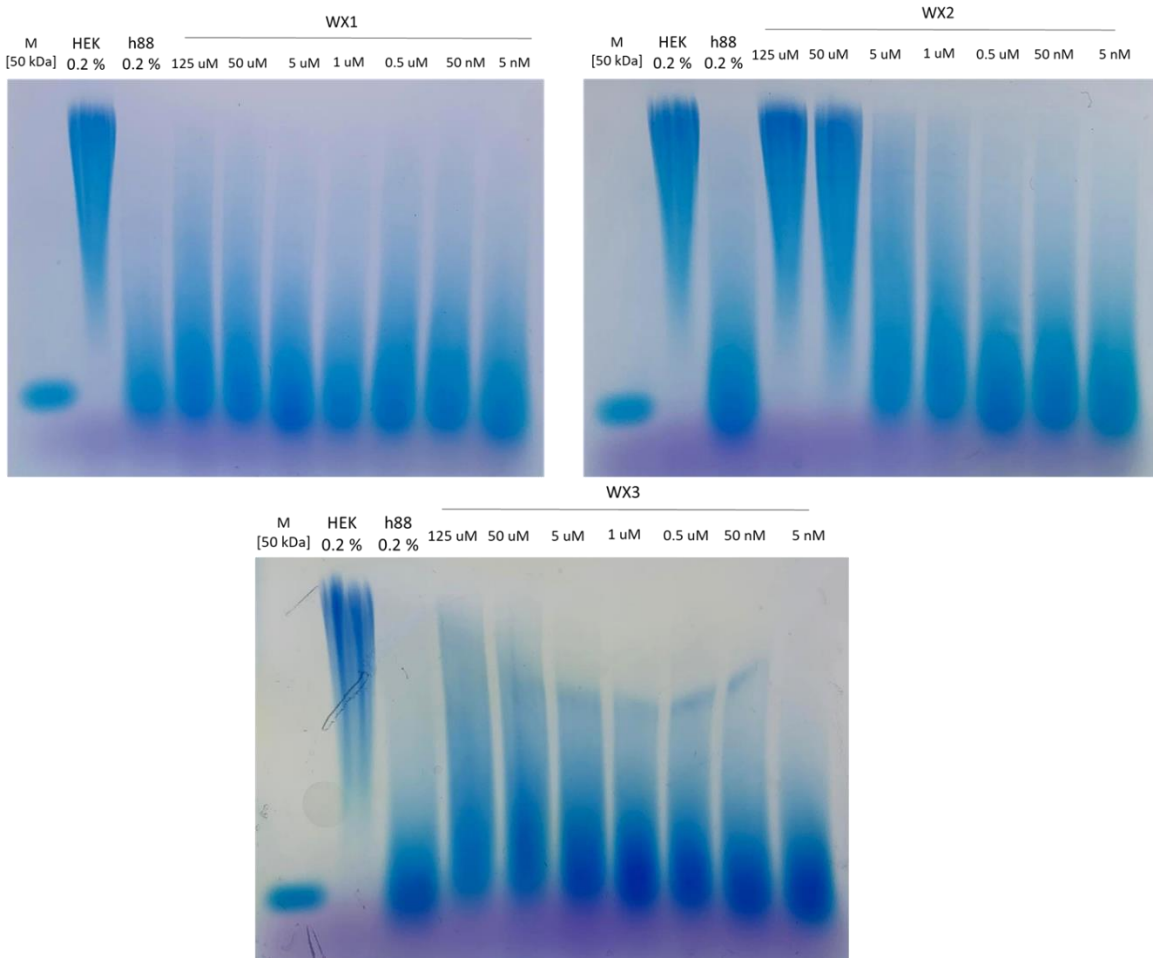


***Figure 73**: **HA assay of molecules WX-1, WX-2 and WX-3**. Control with HEK293T 0.2% DMSO and HEK293T-hCEMIP stable clone (h88) 0.2% DMSO.*

At this point, it became clear that a cell toxicity and cell proliferation test was needed, to be sure that the lack of HA processing was due to CEMIP inhibition and not to a reduced number of cells in the assay. For this purpose, the release of adenylate kinase from damaged cells was measured with a ToxiLight kit from Lonza. Molecule WX-2 exhibited an effect on cell proliferation at 125 µM and 50 µM and showed toxicity at 125 µM (**Figure 74**). Moreover, for molecule WX-2, cell

proliferation was tested by measuring confluency with an Incucyte machine, confirming that cell proliferation is negatively impacted at 50 µM and 125 µM (**Figure 75**).

Nevertheless, WX-2 has proven to be a good candidate as CEMIP inhibitor at 1-50 µM, with only limited influence on cell proliferation at 50 µM.

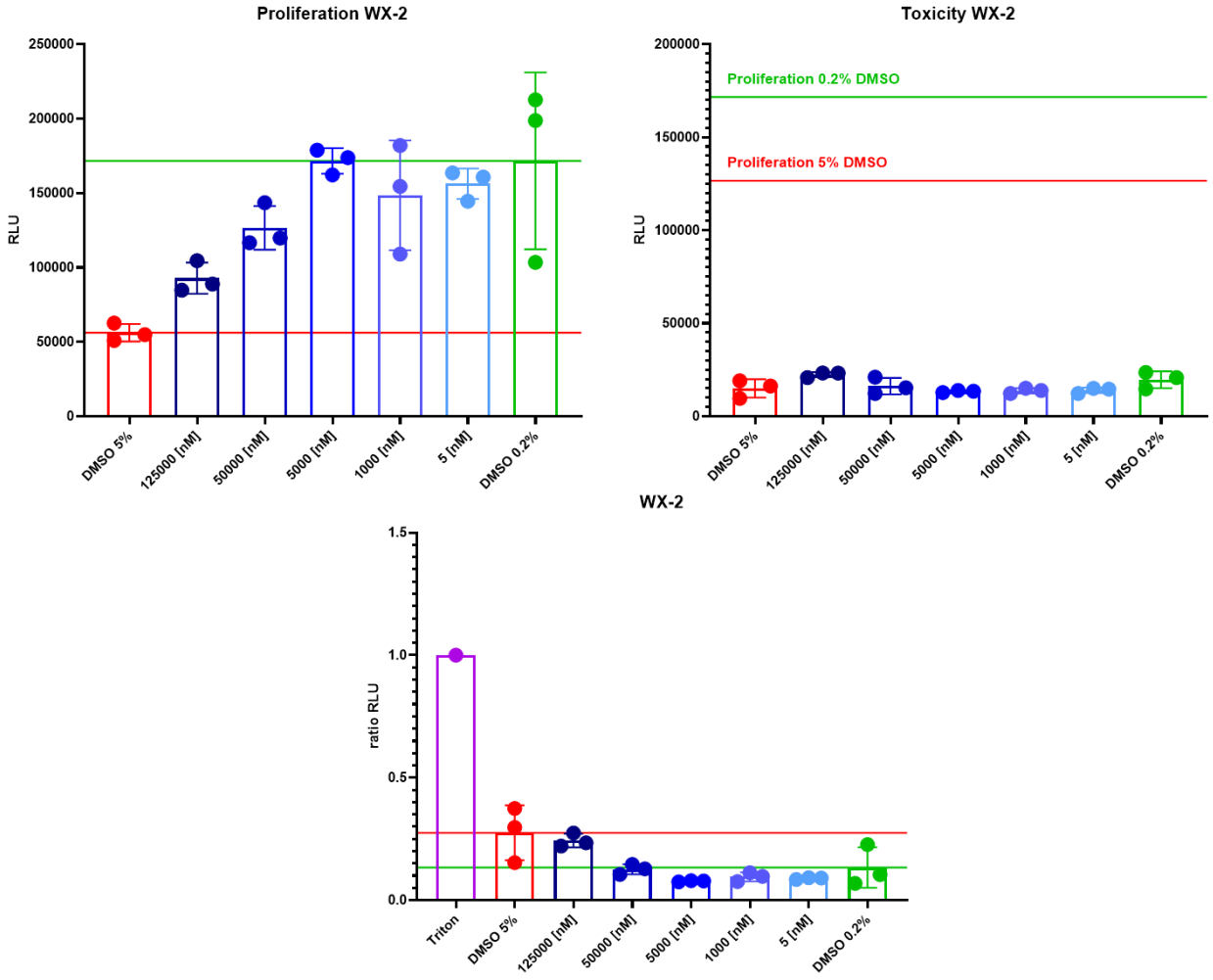


***Figure 74**: **Cell proliferation, toxicity and relative toxicity of molecule WX-2**. Proliferation assay (top left), toxicity assay (top right) and relative toxicity (bottom; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green.*

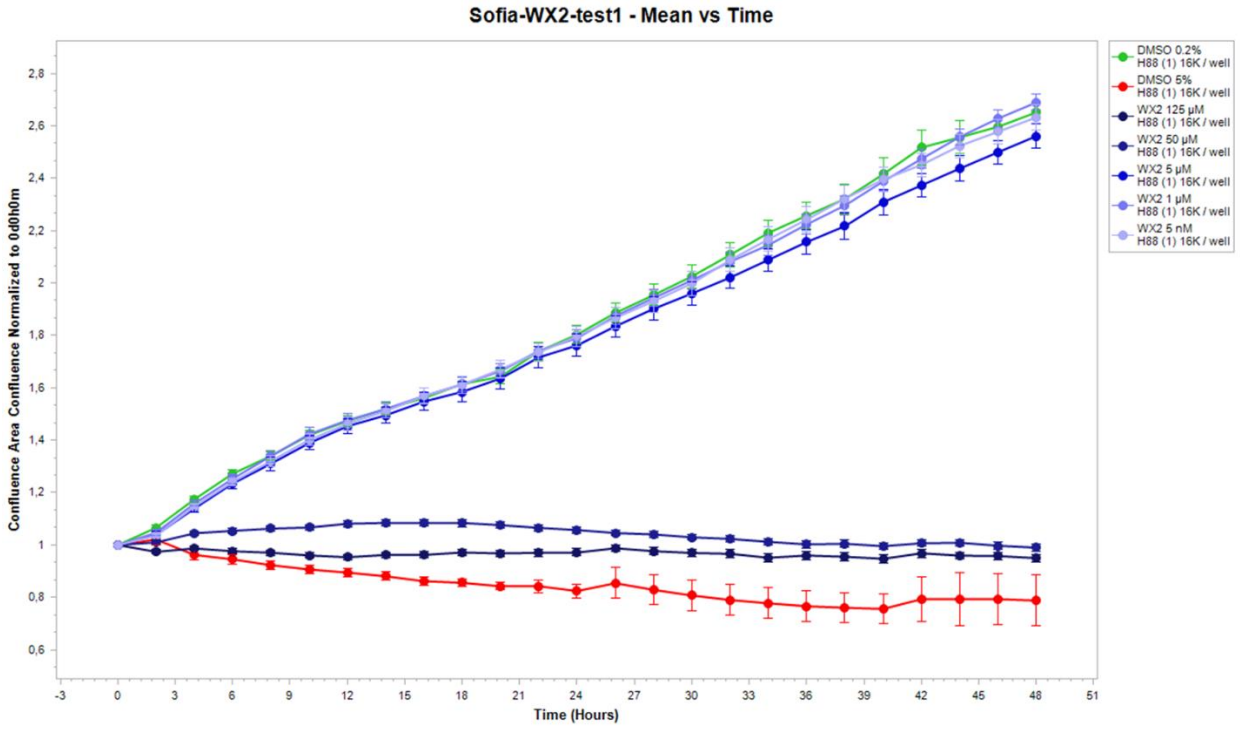


***Figure 75**: **Cell proliferation of molecule WX-2 at different concentrations**. Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. X axis: time in hours; Y axis: confluence area.*

Over the following years, other 11 molecules were tested with an HA assay as well as with cell proliferation and toxicity assay: CEMIP-135, CEMIP-137, FDG-013, FDG-021, FDG-037, FDG-038, FDG-039B, FDG-061, FDG-03, FDG-064 and FDG-065.

Molecules CEMIP-137, FDG-013, FDG-037, FDG-038 and FDG-065 demonstrated solid CEMIP inhibition at 50-125 µM without being toxic or influencing cell proliferation (**Figure 76**, **Figure 77**, **Figure 78**, **Figure 79**, **Figure 80**). On the other hand, CEMIP-135, FDG-021, FDG-039B, FDG-061, FDG-03 and FDG-064 did not show any sign of inhibition (annex **Figure 128**, **Figure 129**, **Figure 130**, **Figure 131**, **Figure 132**, **Figure 133**).

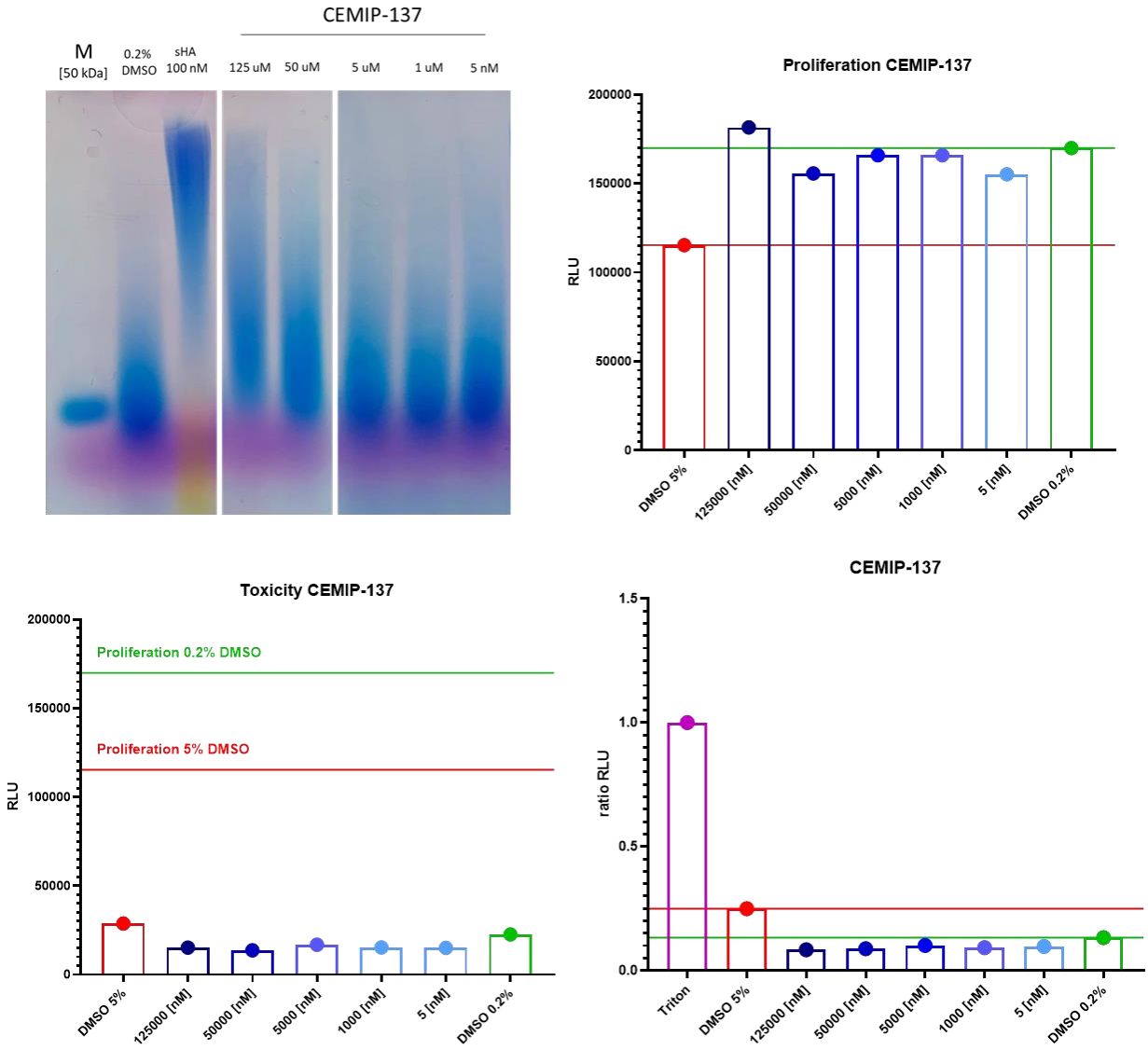


***Figure 76**: **HA assay, cell proliferation, toxicity and relative toxicity of molecule CEMIP-137**. HA assay of the inhibition of CEMIP-137 on h88 cells (top left; control with h88 0.2% DMSO and inhibitions control sHA at 100 nM), proliferation assay (top right), toxicity assay (bottom left) and relative toxicity (bottom right; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. h88 = HEK293T-hCEMIP stable clone.*

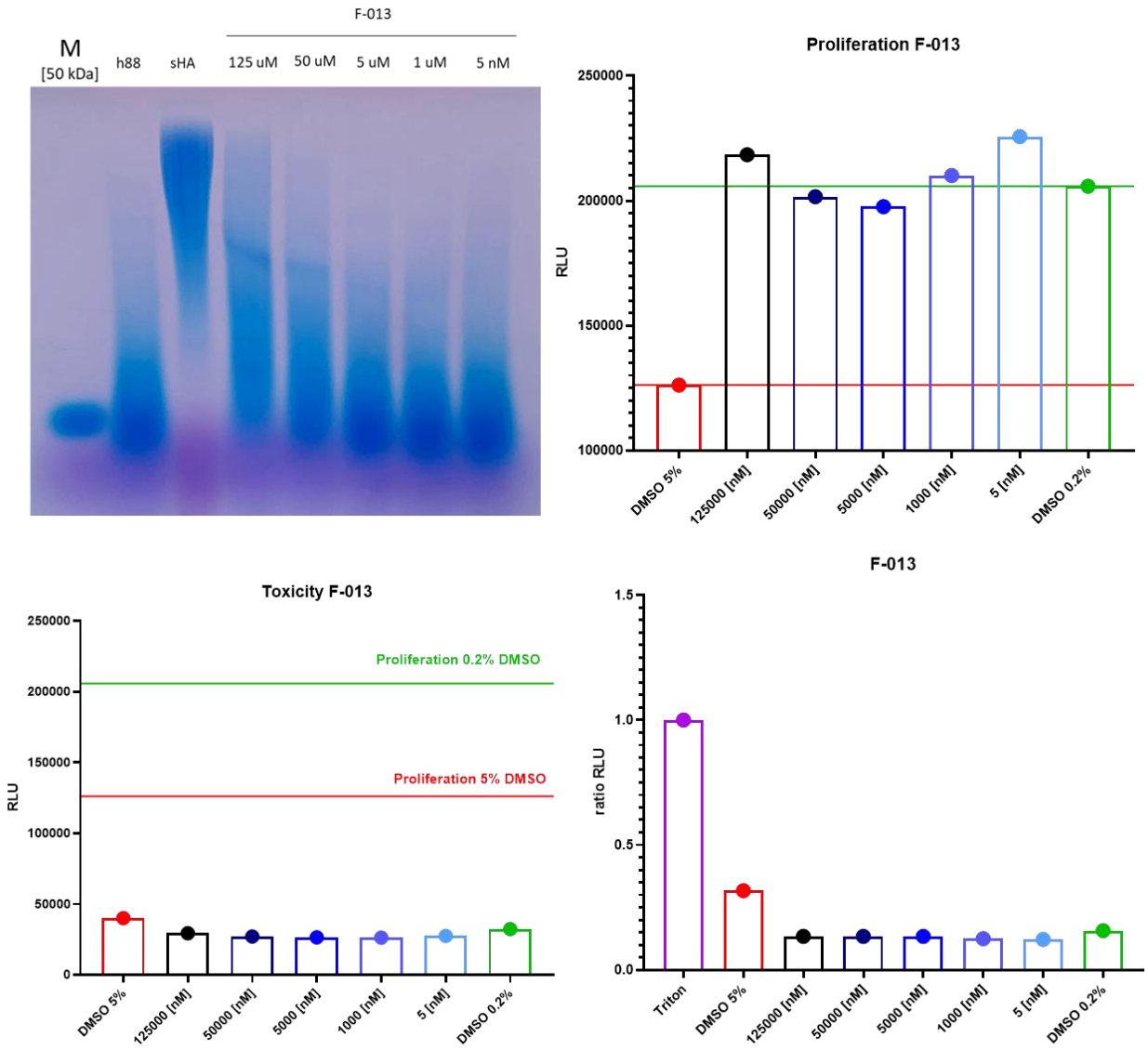


***Figure 77**: **HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-013.** HA assay of the inhibition of FDG-013 on h88 cells (top left; control with h88 0.2% DMSO and inhibitions control sHA at 100 nM), proliferation assay (top right), toxicity assay (bottom left) and relative toxicity (bottom right; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. h88 = HEK293T-hCEMIP stable clone.*

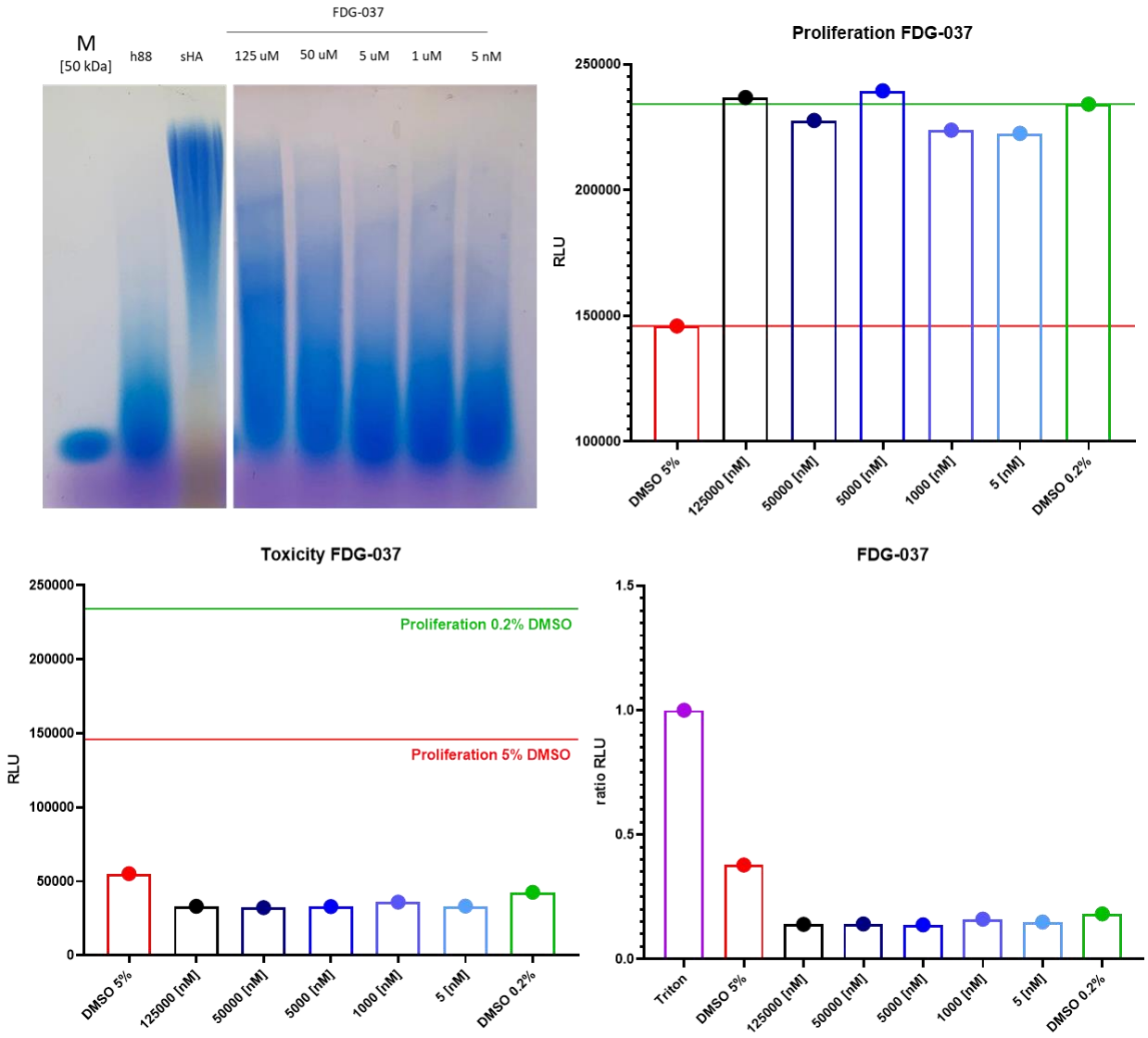


***Figure 78**: **HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-037**. HA assay of the inhibition of FDG-037 on h88 cells (top left; control with h88 0.2% DMSO and inhibitions control sHA at 100 nM), proliferation assay (top right), toxicity assay (bottom left) and relative toxicity (bottom right; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. h88 = HEK293T-hCEMIP stable clone.*

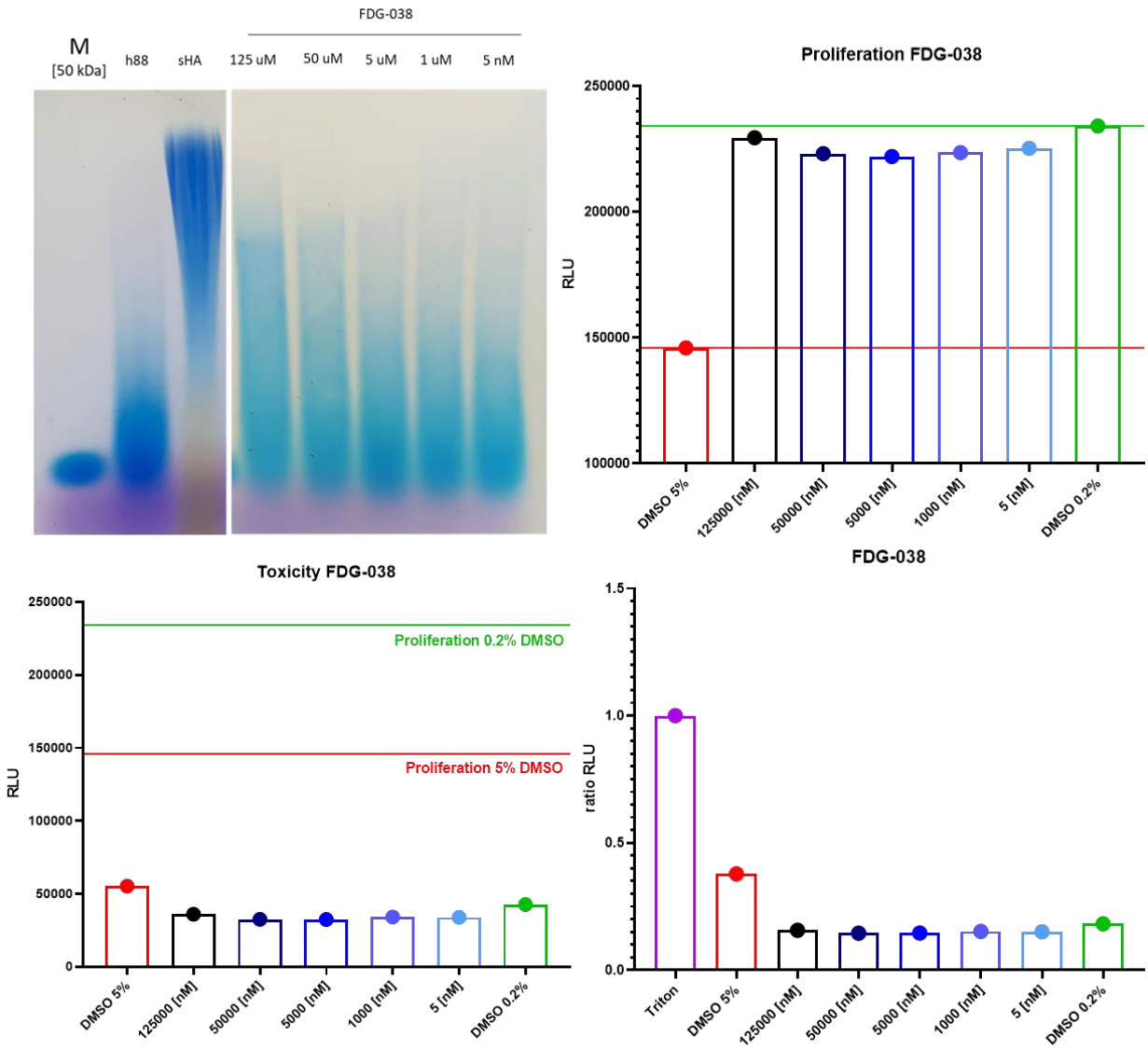


Figure 79: ***HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-038***. *HA assay of the inhibition of FDG-038 on h88 cells (top left; control with h88 0.2% DMSO and inhibitions control sHA at 100 nM), proliferation assay (top right), toxicity assay (bottom left) and relative toxicity (bottom right; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. h88 = HEK293T-hCEMIP stable clone.*

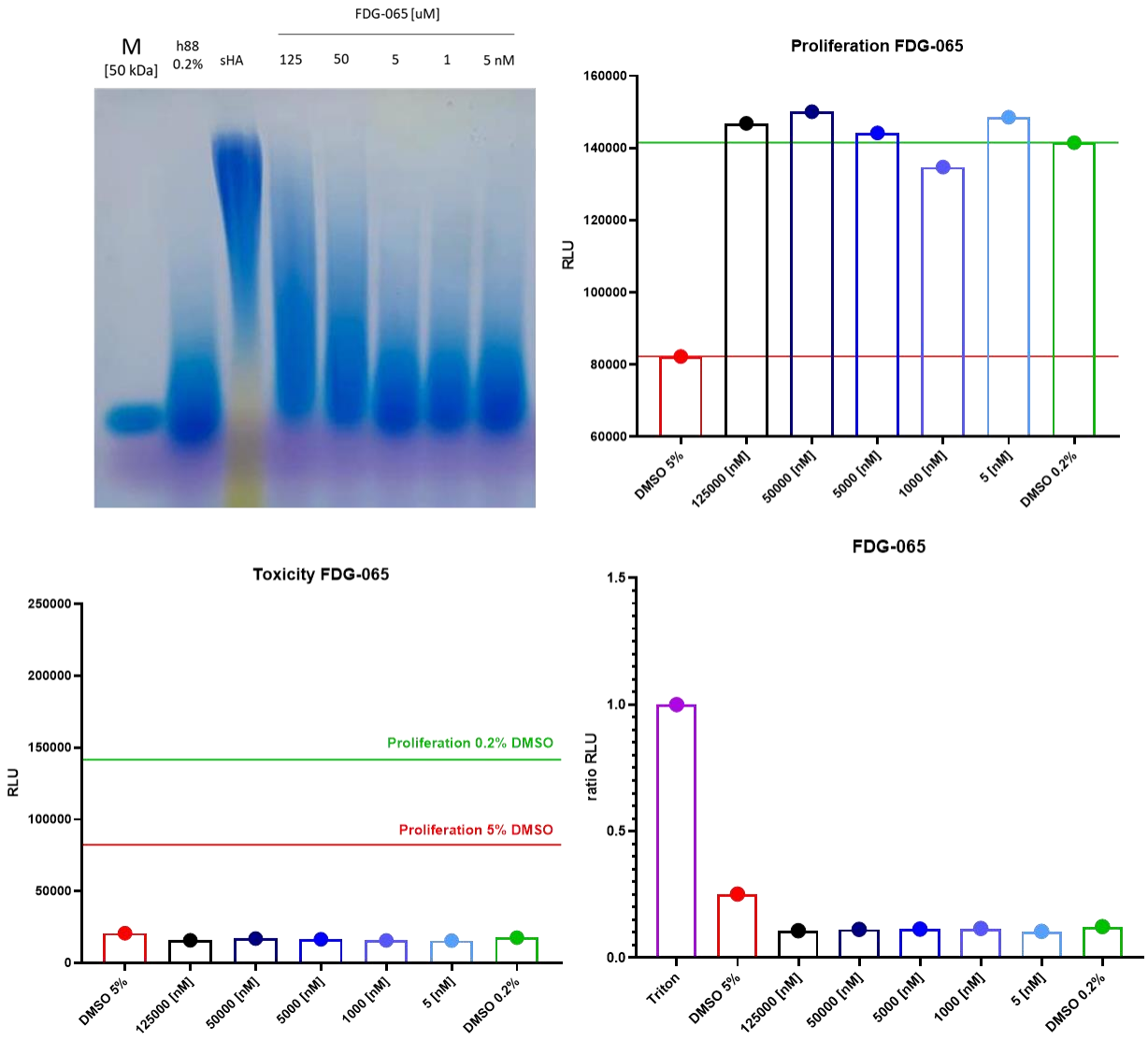


***Figure 80**: **HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-065**. HA assay of the inhibition of FDG-065 on h88 cells (top left; control with h88 0.2% DMSO and inhibitions control sHA at 100 nM), proliferation assay (top right), toxicity assay (bottom left) and relative toxicity (bottom right; toxicity standardized to the total number of cells for each well). Condition that mimics high toxicity with 5% DMSO in red, control condition with 0.2% DMSO in green. h88 = HEK293T-hCEMIP stable clone.*

In order to uncover if a small molecule is binding to the same HA binding site, the first GG domain in this case, I performed a pulldown. If in the presence of a molecule inhibiting CEMIP in the HA assay the binding is lost, it means that the molecule binds to the first GG domain. This was observed for example with sHA in the previous section (**Figure 68**). Moreover, a pulldown in the presence of an inhibitor could also help with discriminating molecules binding to the catalytic site

from the ones binding to the HA binding site. In a pulldown experiment with molecule CEMIP-137 no difference between the pulled-down CEMIP with only DMSO or with the molecule was detected (**Figure 81**). This was also the case for sCEMIP.

This result could mean that CEMIP-137 is binding to the putative catalytic site of CEMIP, or in another unknown region important for the HA processing.

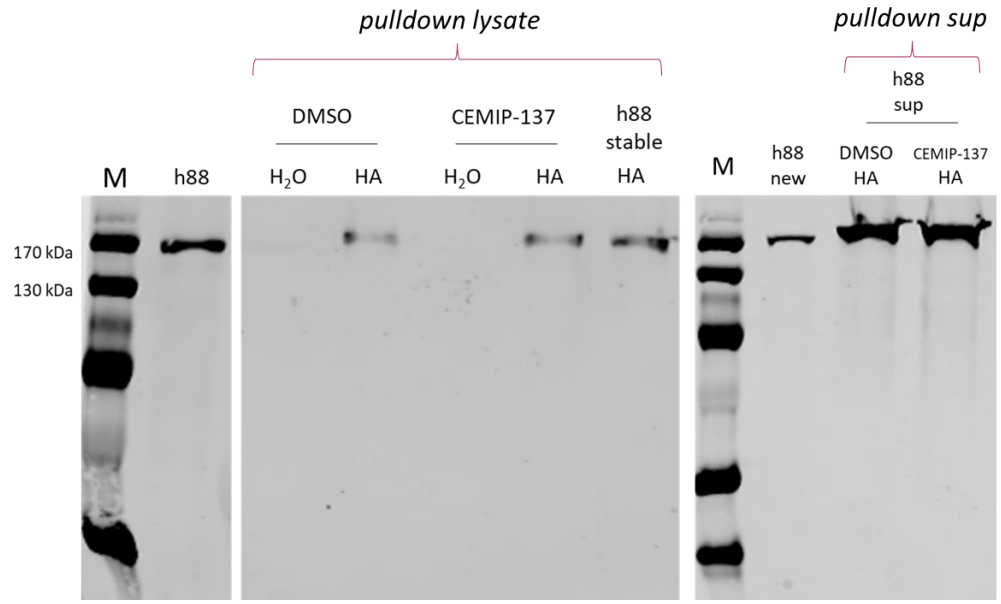


***Figure 81**: **Pulldown of intracellular and extracellular CEMIP with molecule CEMIP-137**. Concentration of CEMIP-137 at ten times its K_d (760 µM). No difference was detected between pulled-down CEMIP or sCEMIP with DMSO or with CEMIP-137. First lane after the marker represents the control h88 (no pulldown). Other lanes represent the pulldown with and without molecule, and water as control. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP stable clone.*

Of all the 25 molecules tested, seven show inhibition of CEMIP between 50 µM and 125 µM: WX-2, WX-3, CEMIP-137, FDG-013, FDG-037, FDG-038 and FDG-065. Excluding WX-2 and WX-3 that come from the DEL screening, the other five molecules share a similarity in structure: a benzene ring with an amine group (aniline, **Figure 82**), that could fit in a negative pocket of CEMIP.

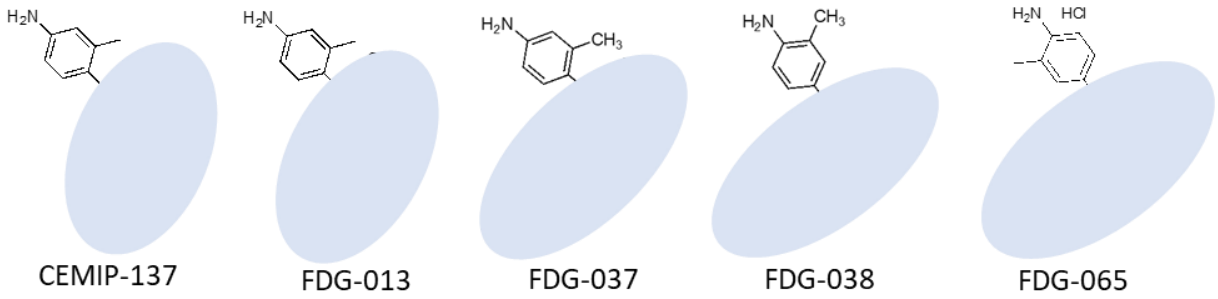


***Figure 82**: **Hit molecules structure**. CEMIP-137, FDG-013, FDG-037, FDG-038 and FDG-065 share a benzene ring with an amine group.*

***Table 9**: Summary of molecules tested in Geneva.*

Name	MW (g/mol)	K_d SPR	HA assay (50-125 µM)
STL 439159	371.420	19.2 µM	NO
STL 441845	302.306	4.5 µM	NO
STK 038497	461.522	Binding at 100 µM	NO
STK 059275	452.535	Binding at 100 µM	NO
STK 175734	439.495	Binding at 100 µM	NO
STK 496096	321.331	Binding at 50 µM	NO
F132-0074	403.461	N/A	NO
L264-0727	393.462	N/A	NO
d-p-Y	261	50 µM	NO
WX-1	483.5	N/A	NO
WX-2	449.5	300 µM	YES
WX-3	519	N/A	YES
CEMIP-113	414.44	84 µM	NO
CEMIP-119	402.39	123 µM	NO
CEMIP-135	386.86	52.8 µM	NO
CEMIP-137	372.40	76 µM	YES
FDG-013	390.398	190 µM	YES
FDG-021	424.892	> 250 µM	NO

FDG-037	386.4	142 µM	YES
FDG-038	406.4	124 µM	YES
FDG-039B	386.43	16 µM	NO
FDG-061	375.38	> 250 µM	NO
FDG-063	391.38	200 µM	NO
FDG-064	432.41	> 250 µM	NO
FDG-065	438.89	> 250 µM	YES

7.2 Development and validation of a selective monoclonal anti-CEMIP antibody

A selective monoclonal antibody against human CEMIP was developed with an epitope in the first GG domain, regions F^{180}ERSWGHRGVIVHV^{193} or V^{189}IVHVIDPKSGTVI^{202}. These two segments contain in fact the two mutations previously mentioned, R187C and D195N respectively, known to reduce hyaluronidase activity of CEMIP (Arg^{187}) and TMEM2 (Arg^{265} and Asp^{273})[37,40]. For this reason, an antibody binding to these epitopes is expected to reduce CEMIP activity. The development and production of the antibody was performed by GenScript initially, and in-house later. After careful selection between hybridomas and characterization of the selected antibody, the latter was proven ineffective as a potential CEMIP blocker, but still represents a great tool for research and investigation of CEMIP in this project.

7.2.1 Hybridoma selection

I received 30 hybridomas supernatant from GenScript and tested them against mCEMIP in Western blot. Six supernatants showed a specific CEMIP band: 2E1, 3F1, 7H5, 8B5, 12G2 and 12G5 (**Figure 83**).

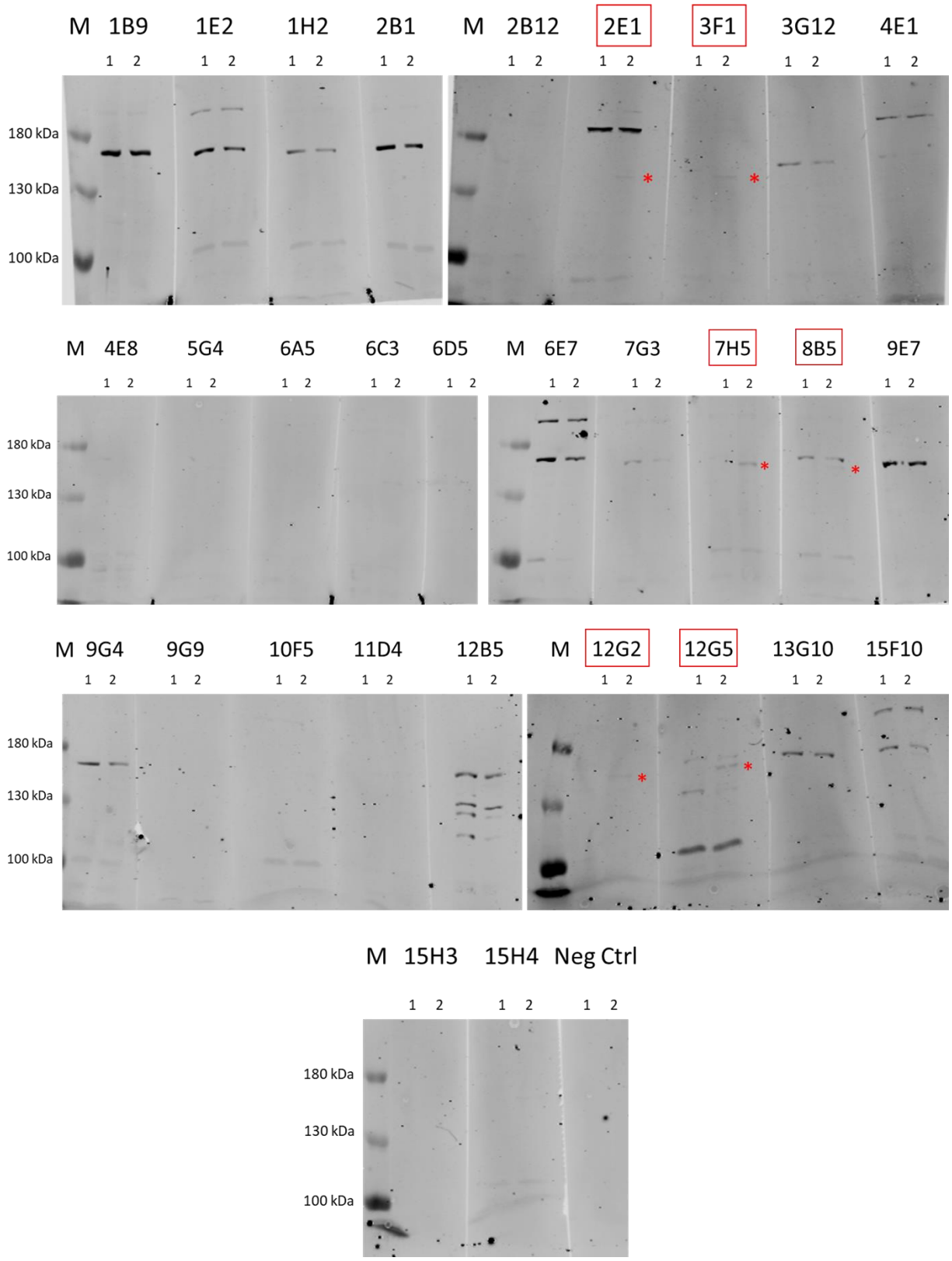


***Figure 83**: **Western blot with different hybridoma supernatants**. Lane 1 = HEK293 WT, lane 2 = HEK293-mCEMIP, M = molecular marker. Star indicates the presence of an mCEMIP band.*

Based on Western Blot results, intracellular staining, quantity, ELISA and EC_{50} (performed by the Cavalli laboratory in IRB, Bellinzona), five hybridoma supernatants were selected for further testing (**Table 10**, in green), and 10 new samples from GenScript were ordered: 3F1-1, 3F1-2, 7H5-1, 7H5-2, 8B5-1, 8B5-2, 12B5-1, 12B5-2, 12G5-1 and 12G5-2.

***Table 10**: Summary of hybridoma characteristics.*

Sample number	Code	Quantity (ng/mL)	EC_{50} (ng/mL)	WB	Intracellular staining
7	3F1	1641,5	21,4	*	x
18	8B5	460,2	54,3	*	x
26	12G5	341,7	71,9	*	x
5	2B12	243,6	24,8		x
11	5G4	399,9	38,8		x
13	6C3	533,1	71,0		x
24	12B5	1695,2	74,9		x
17	7H5	907,2	12,9	*	
1	1B9	141,1	25,3		
6	2E1	64,7	39,0	*	
25	12G2	2914,5	127,5	*	
9	4E1	67,6	18,7		
14	6D5	362,2	25,0		
20	9G4	526,9	27,2		
2	1E2	1043,1	32,6		
22	10F5	739,7	37,3		
19	9E7	412,6	42,7		
4	2B1	5,3	45,4		
29	15H3	62,1	49,8		
3	1H2	71,6	56,7		
8	3G12	36,5	60,8		

16	7G3	212,7	65,9		
23	11D4	318,1	140,5		
27	13G10	29,0	148,8		
12	6A5	267,7	288,9		
21	9G9	696,9	393,4		
28	15F10	89,2	412,7		
15	6E7	53,8	576,4		
30	15H4	101,3	36402,0		
10	4E8	162,9	44931,0		

The 10 new hybridoma supernatants were tested on mCEMIP lysates again. All of them revealed a specific CEMIP band, stronger for hybridomas 12G5-1, 7H5-1, 12B5-2 and 12G5-2 (**Figure 84**).

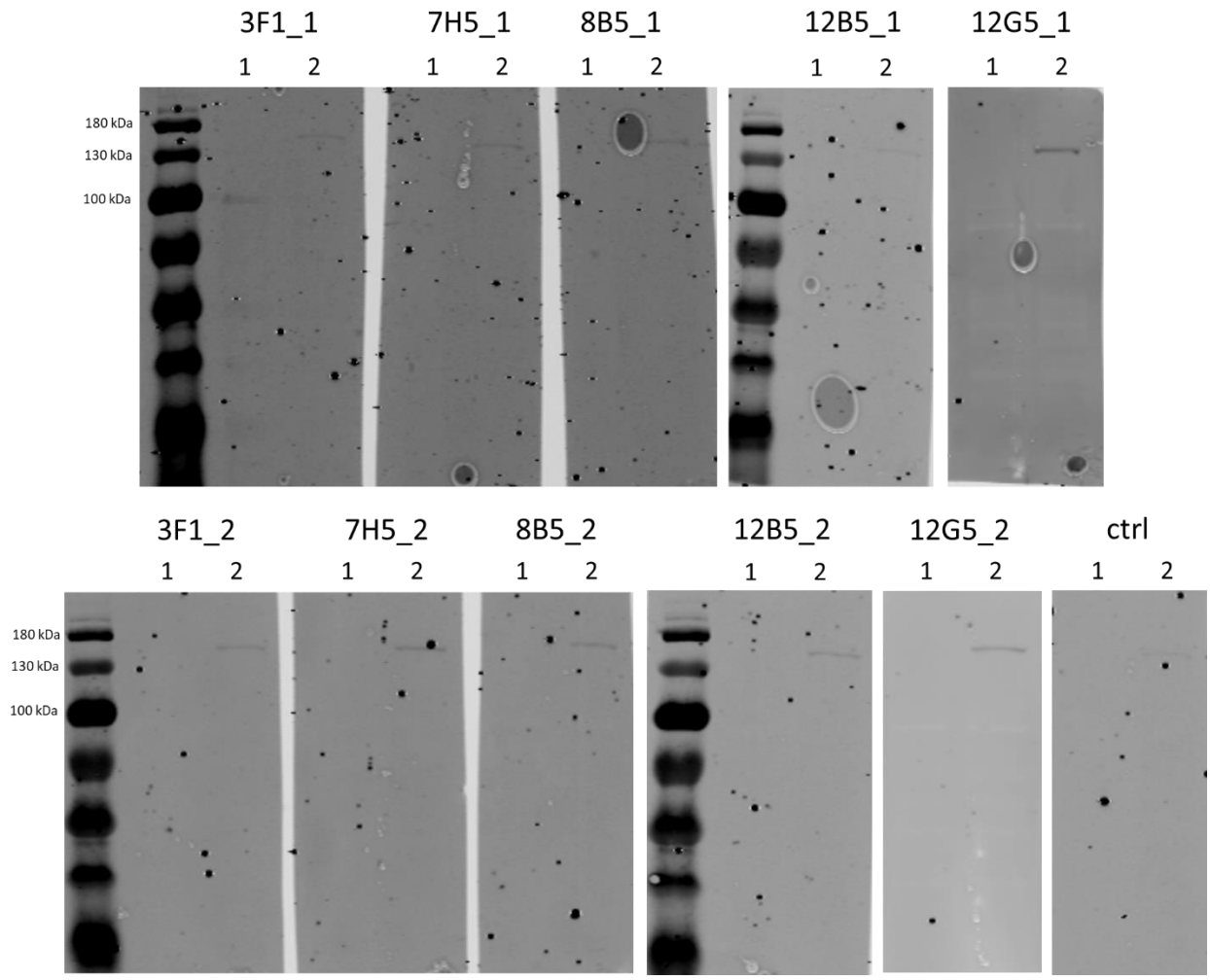


***Figure 84**: **Western blot with different hybridoma supernatants**. Lane 1 = HEK293 WT, lane 2 = HEK293-mCEMIP, M = molecular marker. Star indicates the presence of a mCEMIP band.*

After dialysis in PBS overnight, the ten antibodies were tested in an HA assay to assess the potential murine and human CEMIP inhibition. Unfortunately, no inhibition of HA degradation was detected in multiple experiments and at different concentrations (**Figure 85**, annex **Figure 134** and **Figure 135**).

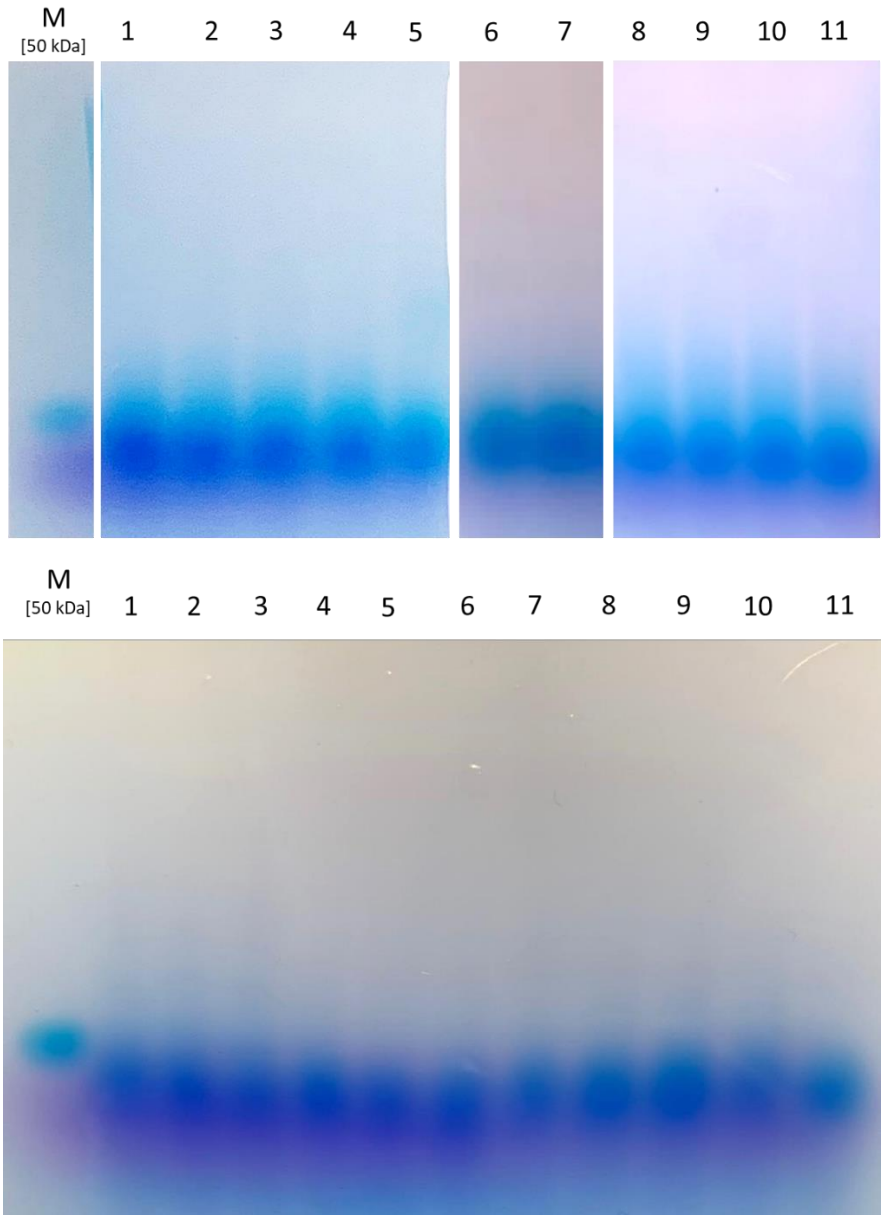


***Figure 85**: **HA assay of selected antibodies**. 1 = 3F1-1, 2 = 7H5-1, 3 = 8B5-1, 4 = 12B5-1, 5 = 12G5-1, 6 = 3F1-2, 7 = 7H5-2, 8 = 8B5-2, 9 = 12B5-2, 10 = 12G5-2, 11 = negative control (another antibody). Top: 10% of antibody on mCEMIP; bottom: 25% of antibody on hCEMIP.*

Regarding the binding ability of these antibodies, as illustrated in **Figure 86** five exhibited binding to the GG domain (12G5-1, 12B5-2, 8B5-2, 7H5-2 and 3F1-1), and two are binding to the full hCEMIP as well (12G5-1 and 12B5-2).

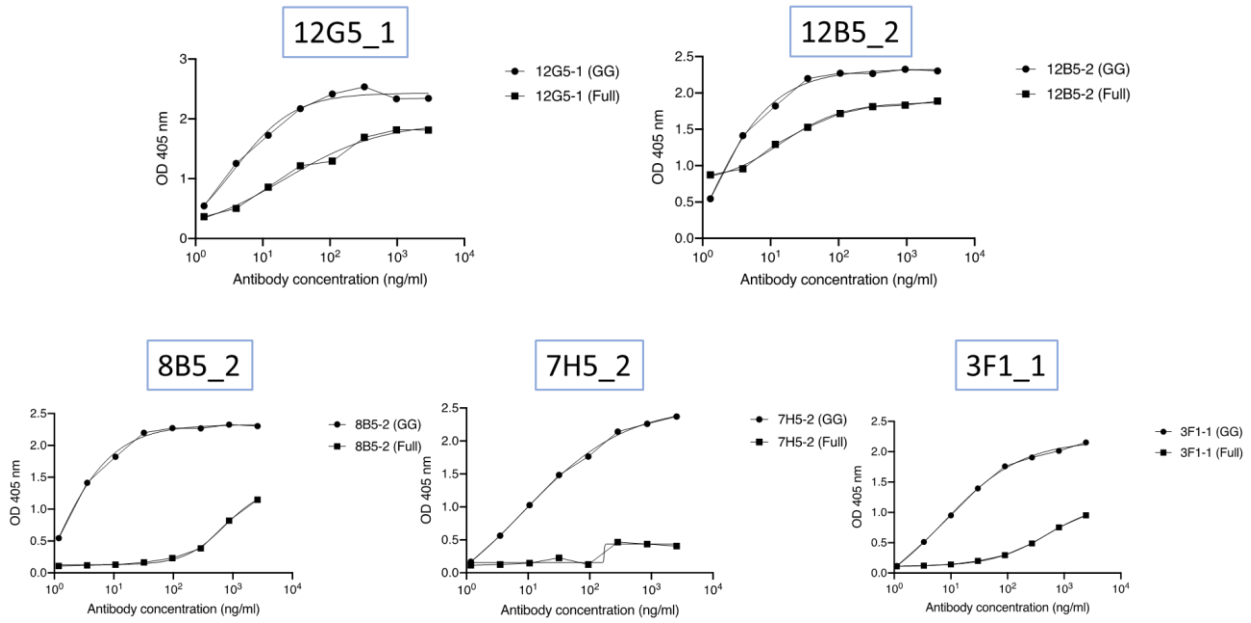


***Figure 86**: **Binding of antibodies to full hCEMIP and GG domain by ELISA.** ELISA results of antibodies 12G5-1, 12B5-2, 8B5-2, 7H5-2 and 3F1-1. X axis: antibody concentration in ng/mL; Y axis: OD 405 nm. Courtesy of Jacopo Sgrignani.*

In an additional ELISA binding assay, hybridoma 12B5-2 confirmed its strong binding to the GG domain and the full CEMIP protein, and a Western blot confirmed binding to hCEMIP in the lysate and in the supernatant (**Figure 87**). Antibody 12B5-2 was selected for further analysis.

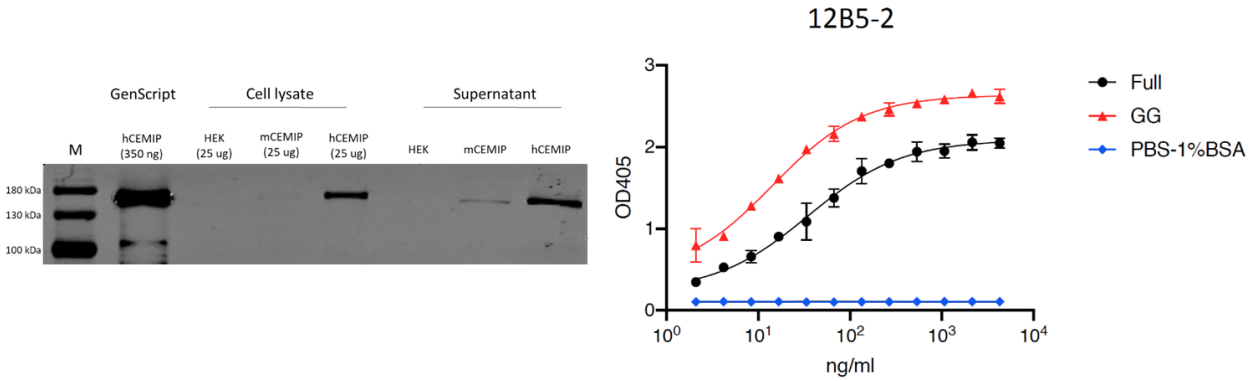


***Figure 87**: **Antibody 12B5-2 binds hCEMIP and the first GG domain**. Antibody 12B5-2 recognizes hCEMIP in Western blot (left) and binds to the GG domain as well as the whole CEMIP protein in ELISA (right). Right: Courtesy of Jacopo Sgrignani.*

7.2.2 Antibody 12B5-2 validation

As mentioned above, the antibody from hybridoma 12B5-2 turned out to be the best candidate in terms of binding in ELISA and Western blot. Yet, a last confirmation of its CEMIP inhibition was tried with an HA assay. 12B5-2 was tested again on cells with a wider concentration range, to be sure about its inhibitory inactivity (from 1/10 to 1/20'000). Unfortunately, depolymerization of HA was observed, closing all further discussion on using the antibody as a potential CEMIP blocker (**Figure 88**).

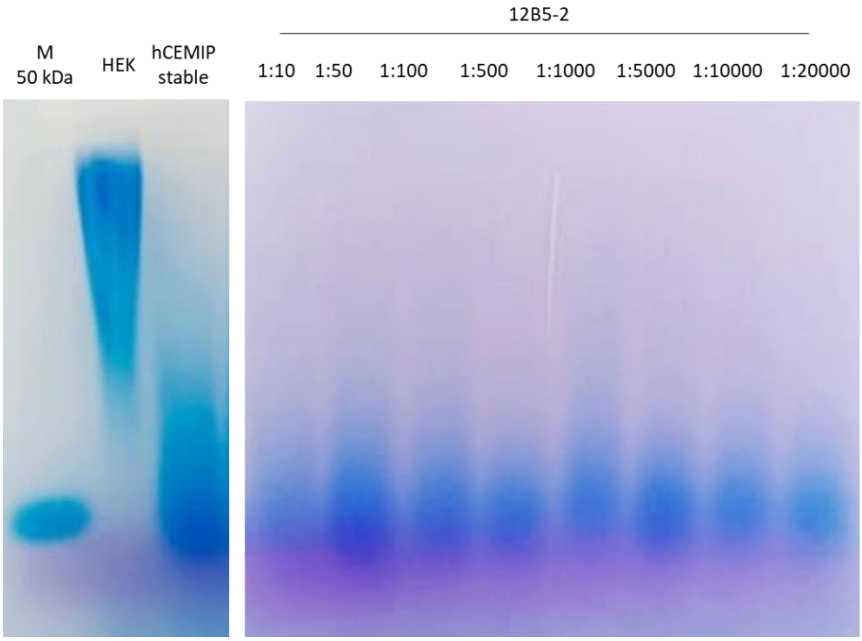


***Figure 88**: **HA assay of antibody 12B5-2 at different dilutions**. Antibody tested on hCEMIP stable clone at different dilutions (1:10 to 1:20'000). Controls with HEK293T and hCEMIP stable clone.*

To further confirm that antibody 12B5-2 cannot act as a CEMIP inhibitor, I performed a pulldown experiment at a maximum concentration (1/10). Not surprisingly, CEMIP is pulled down equally in the presence of PBS or with the antibody (**Figure 89**).

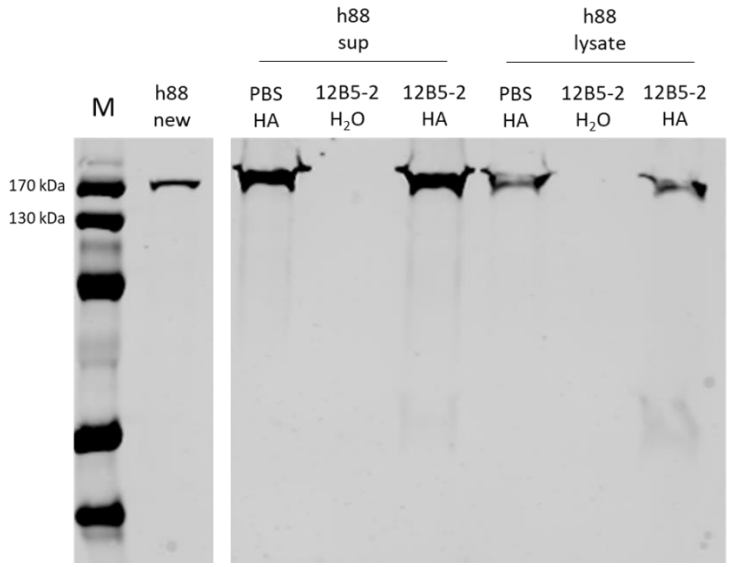


***Figure 89**: **Pulldown of extracellular and intracellular CEMIP with antibody 12B5-2**. 12B5-2 cannot prevent the binding between No difference was detected between pulled-down CEMIP or sCEMIP with PBS or with the antibody. First lane after the marker represents the control h88 (no pulldown). Other lanes represent the pulldown with and without antibody, and water as control. Western blot performed with an anti-CEMIP antibody. h88 = HEK293T-hCEMIP stable clone.*

Either way, this antibody represents a valuable tool for investigating CEMIP in this project and its characterization in Western blot and immunohistochemistry was pursued. A new purified batch was ordered from GenScript and tested with a titration in Western blot, revealing the best working concentration to be between 1/200 and 1/1000 (**Figure 90**).

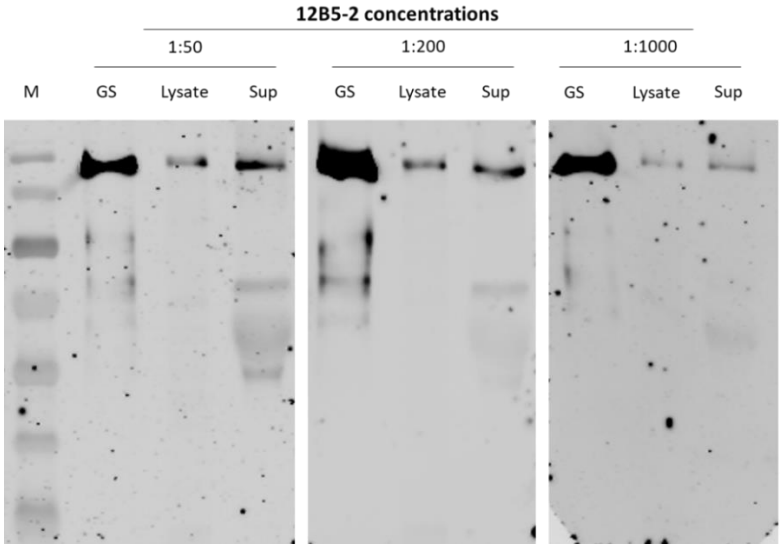


***Figure 90**: **Titration of purified 12B5-2.** Western blot performed with the purified antibody (1:50, 1:200 and 1:1'000) against the purified protein (GS) and the intracellular and extracellular CEMIP. GS = purified CEMIP protein from GenScript.*

Interestingly, when working with the different mutants, I observed that antibody 12B5-2 doesn't bind to mutant D195N (**Figure 91**). This is consistent with the fact that one of the two epitopes was designed around Asp^{195} (V^{189}IVHVIDPKSGTVI^{202}).

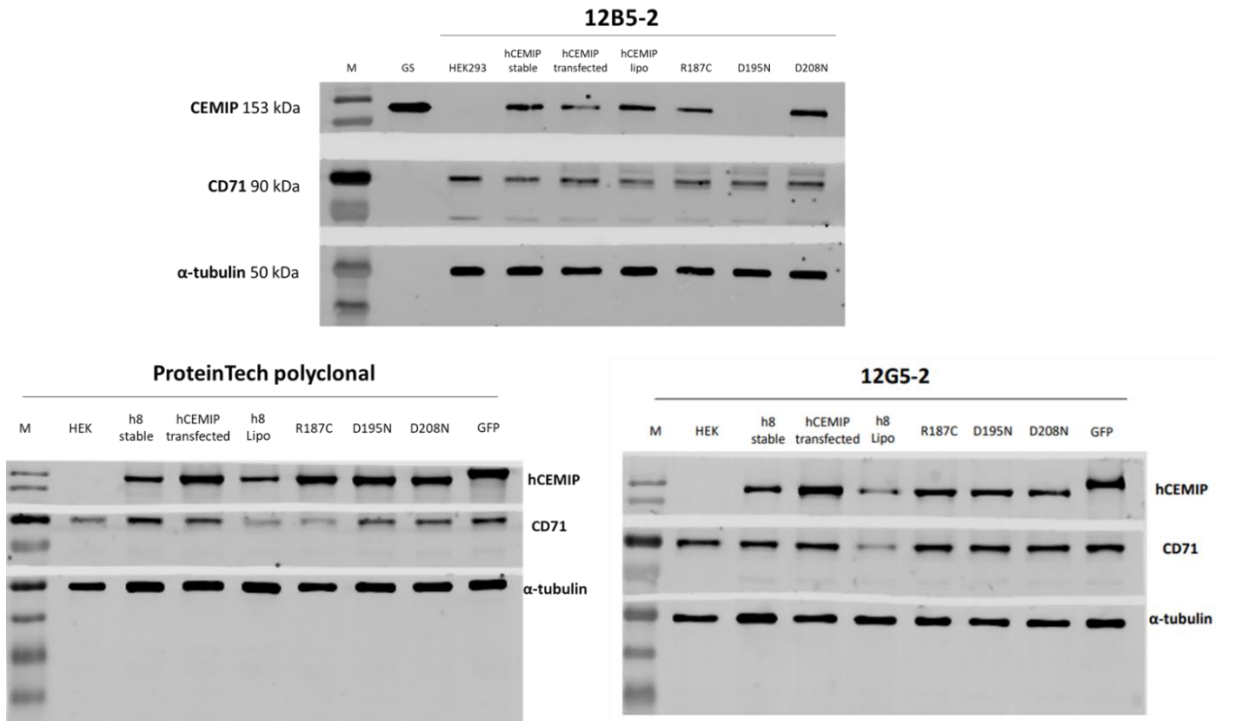


***Figure 91**: **Antibody 12B5-2 is not binding to CEMIP mutant D195N**. Top: Western blot with 12B5-2. Bottom left: Western blot with commercial anti-CEMIP from ProteinTech. Bottom right: Western blot with 12G5-2. Anti-tubulin as loading control, anti-CD71 as membranous protein control. GS = purified CEMIP protein from GenScript.*

Antibody 12B5-2 was also tested on sections of pancreatic tissue with immunohistochemistry (IHC) procedures, revealing a good staining of the islets of Langerhans (**Figure 92**). Although at a higher concentration compared to the antibody developed and kindly provided by Hiroyuki Yoshida, 12B5-2 represents an asset when working on tissue sections of patients. Indeed, Prof. Solange Moll, a key collaborator of this project, will investigate CEMIP expression in tissues from patients affected by various glomerulopathies, renal fibrosis and other fibrotic diseases (systemic sclerosis, idiopathic pulmonary fibrosis, peritoneal fibrosis and in stromal reaction in carcinomas available in the pathology archive) with IHC experiments.

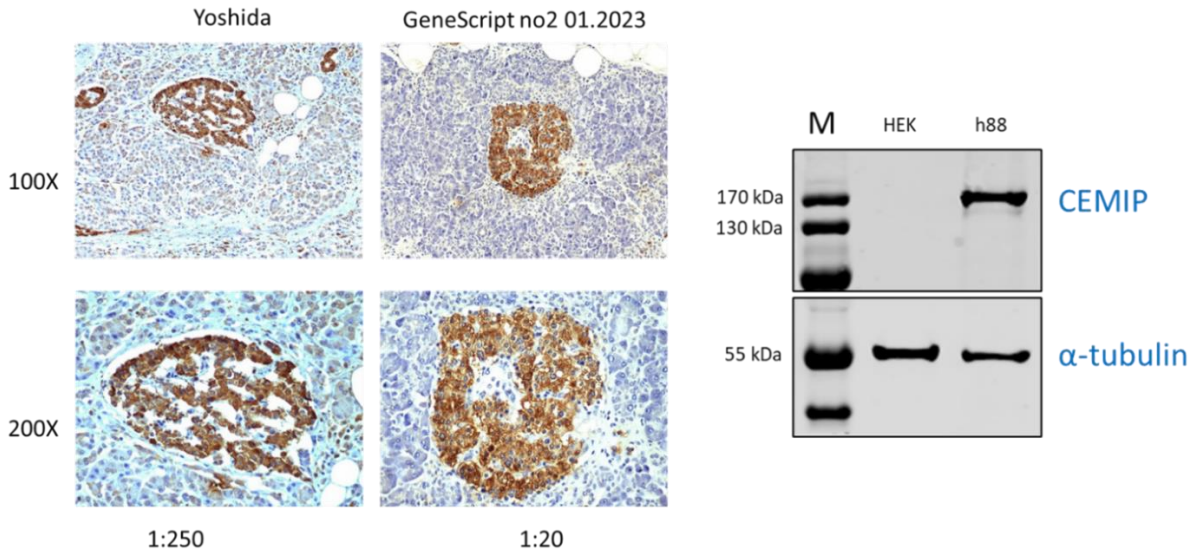


***Figure 92**: **12B5-2 as a tool for IHC and WB**. Left: immunohistochemistry of pancreatic islet of Langerhans with a control antibody provided by Yoshida and the GenScript antibody at 1:250 and 1:20 dilution; courtesy of Thomas Cagarelli and Solange Moll. Right: Western blot with the GenScript antibody and anti-tubulin as loading control. h88 = HEK293T-hCEMIP stable clone.*

7.2.3 Antibodies 12B5-2 and 12G5-2 in-house production

To reduce cost and gain valuable experience in antibody production, I decided to produce 12B5-2 and 12G5-2 in-house with the hybridoma cells provided by GenScript. Cells were first cultivated in DMEM and then passed in OptiMEM for one week before harvesting supernatant. Antibodies in the supernatant were concentrated and run on an SDS gel. 12B5-2 was far less concentrated compared to 12G5-2 in the gel as well as during the purification step (**Figure 93** and **Figure 94**).

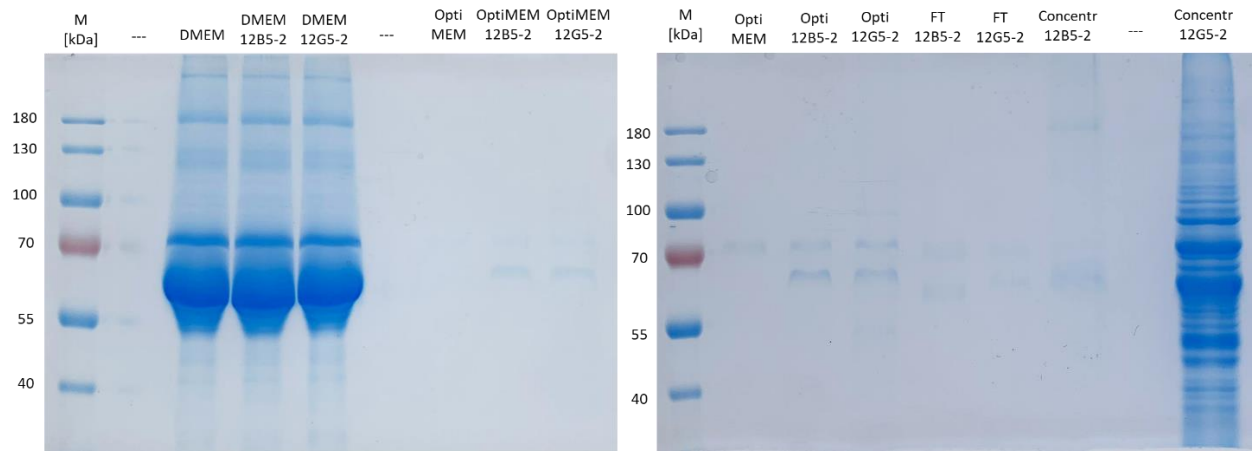


***Figure 93**: **Antibodies 12G5-2 and 12B5-2 production in DMEM and OptiMEM**. Left: antibodies in supernatant; Right: antibodies after concentration. FT = flow-though.*

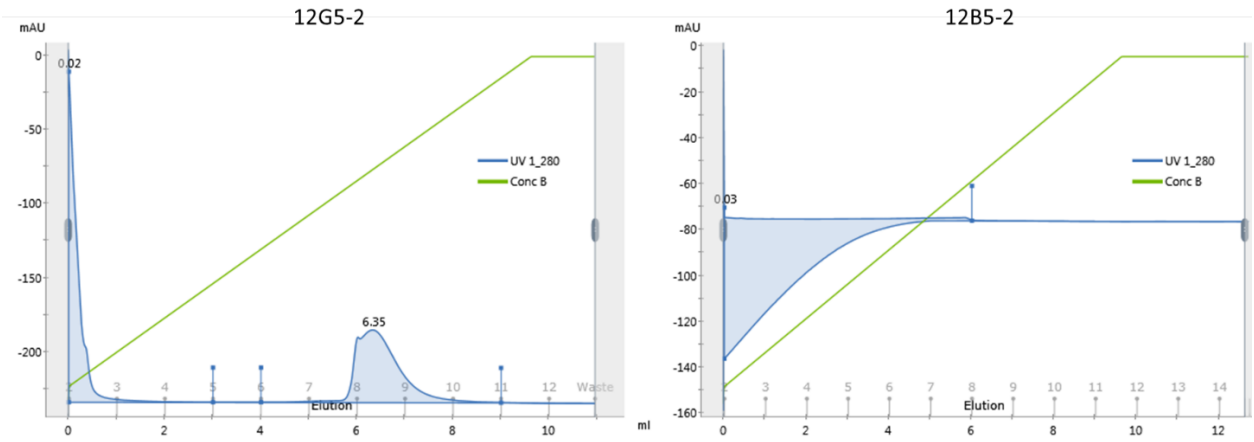


***Figure 94**: **Antibody purification with protein G column**. Antibody 12B5-2 (right) is less concentrated than 12G5-2 (left). X axis: elution fractions; Y axis: UV Absorbance at 280nm.*

Indeed, the quantification by Nanodrop of the combined fractions confirmed that 12G5-2 is at a concentration of 0.764 mg/mL and 12B5-2 at only 0.011 mg/mL (**Table 11**).

***Table 11**: Fractions concentration after purification of 12B5-2 and 12G5-2.*

Hybridoma	Fraction	Concentration [mg/mL]	
12B5-2	7	-1.146	
12B5-2	8	-0.015	➡ 200 uL @ 0.011 mg/mL
12G5-2	7	0.007	
12G5-2	8	**0.125**	➡ 300 uL @ 0.764 mg/mL
12G5-2	9	0.012	

An SDS gel of samples in normal and reducing conditions verified that the purified protein is in fact the antibody (**Figure 95**). The profile of the antibody from GenScript corresponds to the 12G5-2 produced in-house, while for antibody 12B5-2 it is possible to observe that no protein is present in the sample.

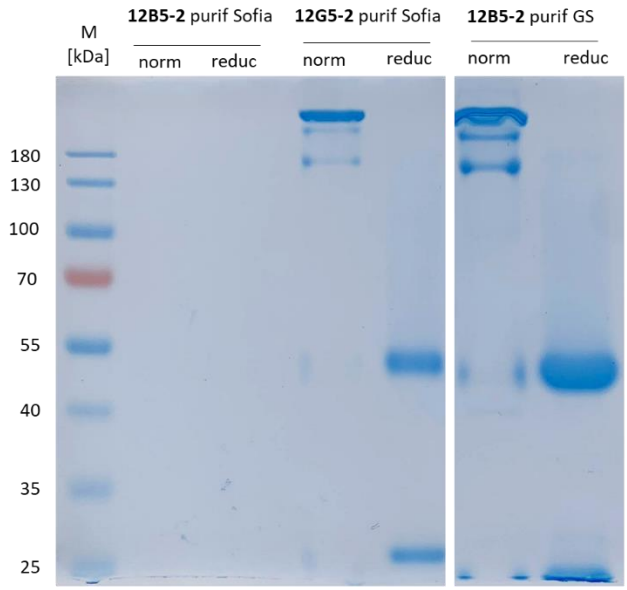


***Figure 95**: **Effect of reduction (DTT) and comparison with 12B5-2 from GenScript**. SDS gel of antibodies purified by me (Sofia) and by GenScript (GS) in normal (norm) and reducing conditions (reduc). Antibody 12G5-2 is well purified and its profile corresponds to the one from GenScript, while antibody 12B5-2 is not present in the sample.*

A Western blot validated the hypothesis that 12B5-2 is less concentrated and almost not present in the sample (**Figure 96**). Interestingly, for WB purposes the purification step was not necessary for a proper result, since the antibody 12G5-2 reveals a CEMIP band also when only concentrated.

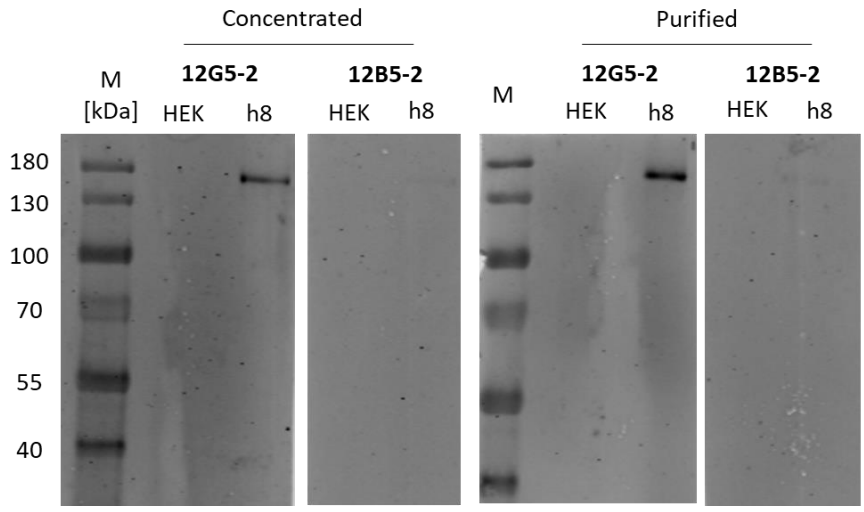


***Figure 96**: **Effect of purification on produced antibody in Western blot**. Western blot with HEK293T WT (HEK) and HEK293T-hCEMIP stable clone (h8) with concentrated and purified 12G5-2 and 12B5-2.*

I then focused on the production of antibody 12B-2 and decided to change the culture medium, from DMEM to RPMI1640. Cells were growing faster and in a better shape, giving hope for better antibody production. The purification was successful and after concentration 200 µL at 3.3 mg/mL were obtained (**Figure 97**).

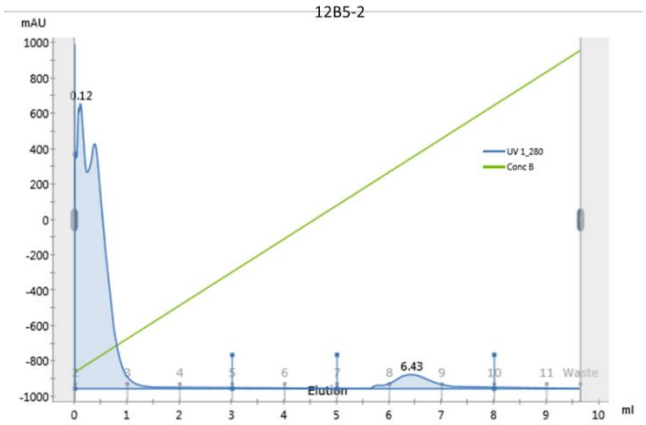


***Figure 97**: **New 12B5-2 purification**. X axis: elution fractions; Y axis: UV Absorbance at 280nm.*

The profile of the newly produced 12B5-2 was identical to the one from GenScript, and in Western blot it was working perfectly also to a dilution of 1/10'000 (**Figure 98**).

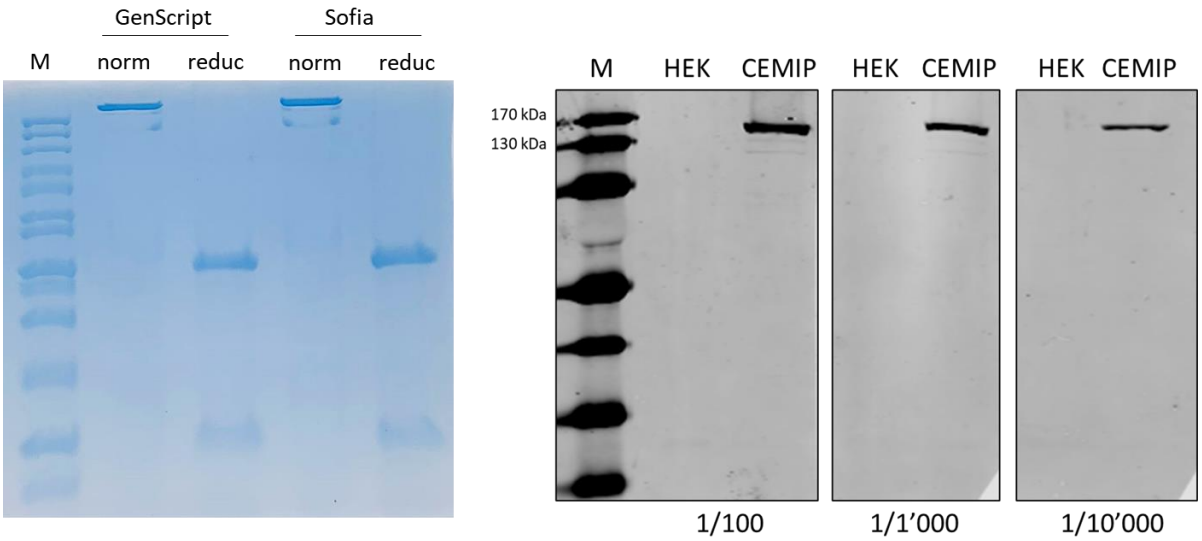


***Figure 98**: **In-house 12B5-2 validation**. Left: in-house 12B5-2 profile (Sofia) compared to 12B5-2 GenScript production in SDS gel in normal (norm) and reducing (reduc) conditions. Right: Western blot with in-house 12B5-2 with HEK293T WT (HEK) and HEK293T-hCEMIP stable clone (CEMIP) at different dilutions (1/100, 1/1'000 and 1/10'000).*

I scaled up the production to 0.5 L of cultured cells and obtained between 1.5 mg and 2 mg of purified 12B5-2 antibody (**Figure 99** and annex **Figure 136**).

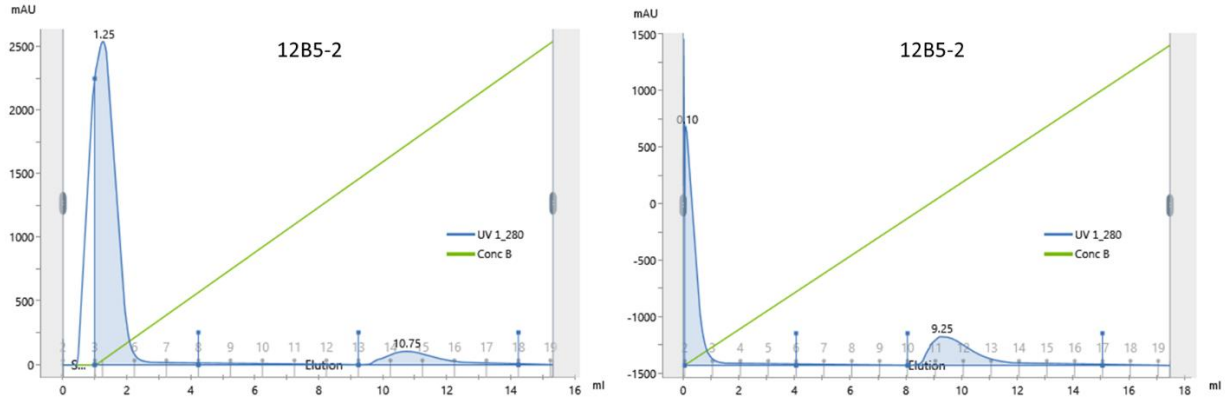


***Figure 99**: **Up-scaled purification of 12B5-2.** X axis: elution fractions; Y axis: UV Absorbance at 280nm.*

In-house 12B5-2 was also tested on pancreas tissue in IHC, but unfortunately, it needed to be far more concentrated to reach Yoshida's result (17-fold, **Figure 100**). To gain insight about some

special procedure I were maybe neglecting, I contacted GenScript, but they assured me that no specific measure was taken in the production of this antibody. Unfortunately, the amount of antibody needed to perform IHC on a larger scale meant a scale-up production not feasible in my laboratory and not worth it in terms of money and time. This is why for IHC experiments the antibody from GenScript was considered the most suited one, while the in-house 12B5-2 was kept for WB analyses.

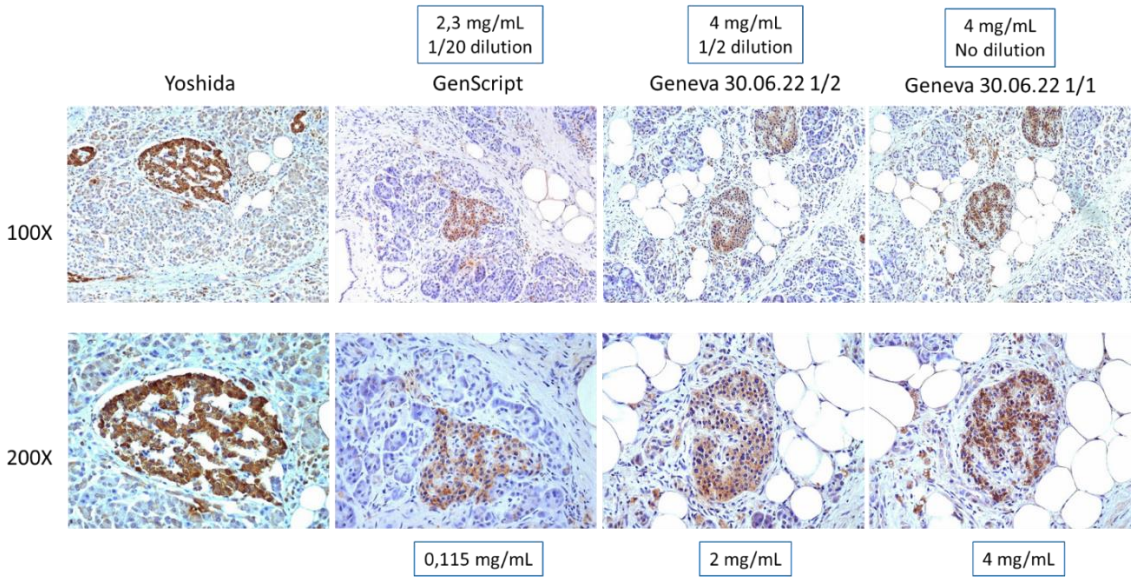


***Figure 100**: **12B5-2 in IHC**. Immunohistochemistry of pancreatic islet of Langerhans with the GenScript and the in-house antibody. Control antibody provided by Yoshida. Tissue control = pancreas. Antibody from GenScript is used at 0.115 mg/mL, while the in-house at 2-4 mg/mL. Courtesy of Thomas Cagarelli and Solange Moll.*

7.2.4 Recombinant 12B5-2 antibody

A recombinant version of 12B5-2 after sequencing and production by the University of Geneva antibody facility was tested in Western blot but revealed a less intensity signal (**Figure 101**). Consequently, 12B5-2 from GenScript (or from in-house production) was used in Western blots and IHC.

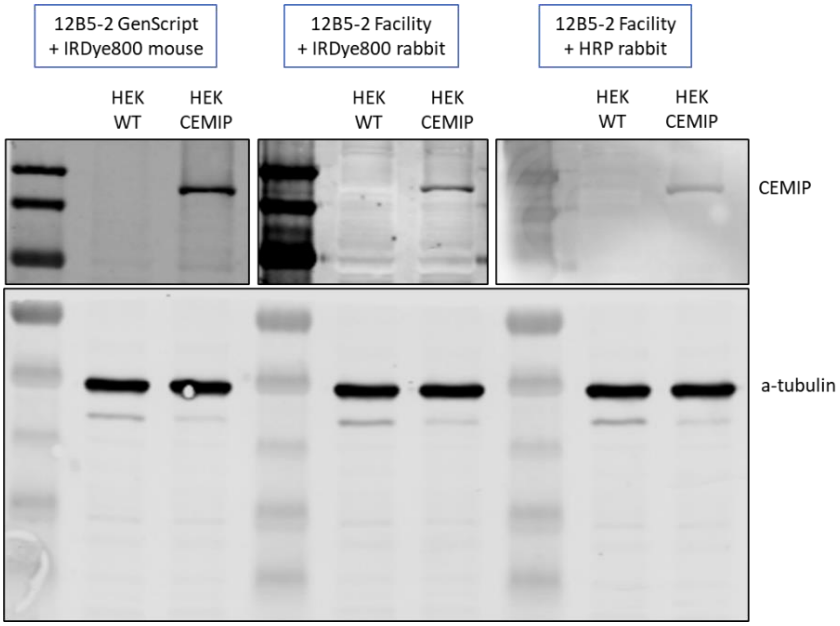


***Figure 101**: **Recombinant 12B5-2**. Western blot with the GenScript 12B5-2 or the recombinant antibody form the facility with HEK293T WT (HEK WT) and HEK293T-hCEMIP stable clone (HEK CEMIP).*

7.3 Discussion

Sulfated hyaluronic acid has been shown to have preventive and therapeutic effects on experimental mesangial proliferative GN[38], indicating that hyaluronidase activity plays a role in GN. Moreover, recent studies confirmed that sulfated HA inhibits CEMIP hyaluronidase activity with an IC_{50} of 5-20 nM[128]. For the first time, I showed that sHA inhibits CEMIP by preventing its binding to hyaluronic acid.

In the present thesis work, attempts have been made to come up with selective CEMIP inhibitors: two groups of compounds were identified: WX-2 and WX-3 from the DEL screening and CEMIP-137, FDG-013, FDG-037, FDG-038 and FDG-065 from subsequent computational design.

Interestingly, molecules from the second group share the same functional chemical group, an aniline. This new finding is a step forward in the drug discovery process, providing new insight for the design of new molecules. Even if the molecules were designed to bind the first GG domain, I don't have any evidence supporting this fact. The pulldown experiment shows no signs of reduced HA binding in the presence of an inhibitory molecule, indicating that the small molecule could

bind in the catalytic region, or in another unknown region important for the HA processing. Nevertheless, the exact binding site of the molecules remains elusive.
Crystallization work of the GG domain, described in the next chapter, will possibly shed light on CEMIP fine molecular structure enabling us to perform a more precise molecular docking and develop a proper structure-activity relationship for the future hits.

In this chapter, I also described the development of a selective monoclonal antibody against human CEMIP. After a careful hybridoma selection, the candidate 12B5-2 was found effective in ELISA, Western blot and immunohistochemistry experiments. Unfortunately, the developed anti-CEMIP antibody showed no signs of blocking CEMIP hyaluronidase activity. Yet, antibody 12B5-2 (and 12G5-2) constitutes a very valuable asset for this research, and their in-house production can be maintained for Western blot and IHC experiments.

7.4 Materials and methods

Cell culture, HA assay, CPC pulldown and Western blot protocols are previously described in section 6.8.

7.4.1 DEL screening

DEL screening was performed by WuXi AppTec.

7.4.2 ToxiLight and Incucyte

Molecules or antibodies were added 2h before the HA in the HA assay (see section 6.8.5). This pre-incubation period allows cell penetration and avoids competition with HA (in case the molecule binds to the same site). In order to test molecules' toxicity on cells, the release of adenylate kinase from damaged cells was measured with a ToxiLight kit (Lonza, LT07-217). Cell proliferation was also tested with an Incucyte instrument (Sartorius). Data were analyzed with GraphPad Prism. A standardization was applied as follows. The same condition was replicated in two wells, with the only difference that in one well all cells were lysed with Triton X-100 for 10 minutes, providing a measure of cell proliferation. In this way, the signal from the adenylate kinase due to the molecule toxicity can be divided by the one coming from the total number of cells.

7.4.3 Antibody production

The development and production of the antibody were performed by GenScript (order ID: U6027EA170-5). Regarding the in-house production, hybridoma cells 12B5-2 and 12G5-2 were cultured in RMPI 1640 (Gibco, 21875) enriched with L-Glutamine, 10% FCS, streptomycin, and penicillin. Cells were amplified until the desired volume reached maximal confluence and transferred into Opti-MEM medium (Gibco, 31985062) for one week. The supernatant was collected and concentrated using an Amicon Ultra-15 100 kDa (Millipore, UFC9100). The antibody was purified with a HiTrap protein G column of 1 mL (Cytiva, 17040401) on an Akta pure machine (GE HealthCare).

8 CHAPTER III – TOWARD SOLVING THE STRUCTURE OF CEMIP

Up to date, the structure of CEMIP is not solved yet. All computational analyses present in this work were performed on a homology model of the GG domain or, after AlphaFold was released, on the AlphaFold prediction. To work on the exact GG domain structure by X-ray crystallography and on the entire CEMIP structure by cryogenic electron microscopy (cryo-EM), I established a collaboration with the laboratory of Prof. Andreas Boland (University of Geneva), with the end goal of gathering information about the putative catalytic site and improve the docking and design of small molecules. Detailed results and information about the material and methods used in this chapter can be found in the doctoral thesis of Lina Poulain (University of Geneva).

8.1 First steps in solving the structure of the GG domain by X-ray crystallography

The GG domain (amino acids 149-294) plays a key role in the activity of CEMIP[47]. In fact, subjects with mutations in position 187 (R187C and R187H) are affected by non-syndromic hearing loss caused by an impaired HA depolymerization activity of the protein[37,47]. In this manuscript, I demonstrated how Arg^{187} has an HA-binding role. Furthermore, mutations in the same domain of the homologous protein TMEM2[40], also invalidate the hyaluronidase activity of the enzyme. For this reason, small molecules and antibodies were designed to bind and block the first GG domain. Since no 3D structure of CEMIP is available, computational analyses were carried out with a GG domain modelling, and having an exact structure of the GG domain remains a key step in small molecule design.

8.1.1 GG domain production

At the beginning of the project, I thought about producing the GG domain in the laboratory in Geneva using suspension HEK293SF-3F6 cells and a vector coding for the GG domain with a 6-His tag. The amount of protein expressed was very low, not visible on an SDS gel, but only in Western blot (**Figure 102**).

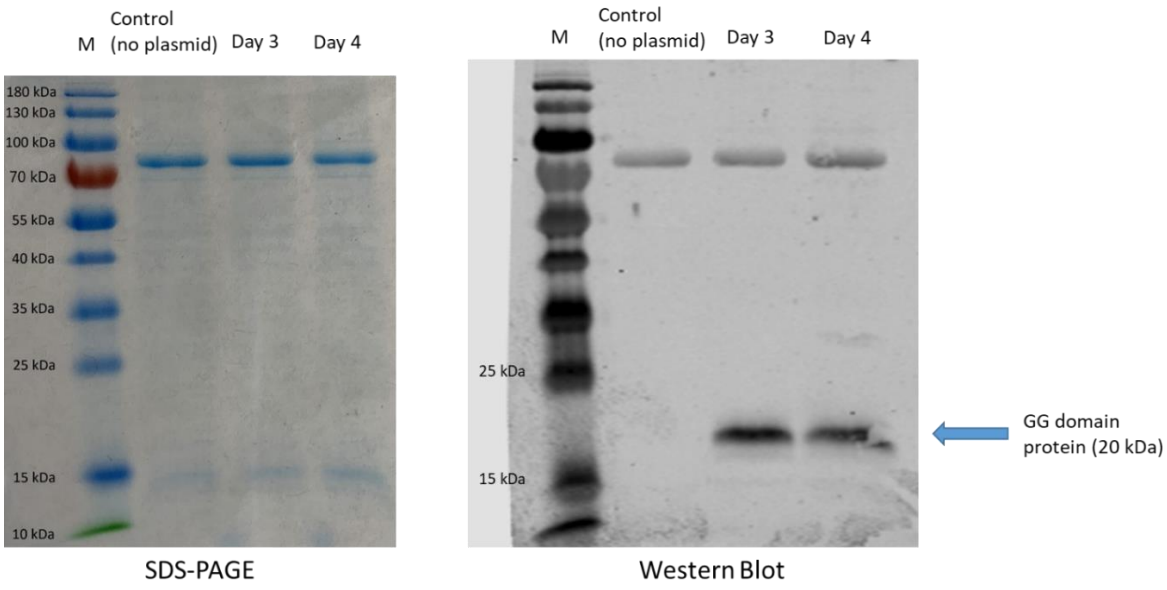


***Figure 102**: **GG domain expression**. SDS gel (left) and Western blot (right) of the supernatant after 3-4 days of GG domain expression.*

After observing a low expression of the GG domain (not enough for crystallization purposes) with PEI transfection, the latter was implemented with other reagents, such as Lipofectamine 3000 and Viromer. Several tests using a GFP-containing vector revealed Lipofectamine to be the best transfection reagent for this type of cells (annex **Figure 137**), but unfortunately, after purification with a nickel column (HisTrap, Cytiva), no protein was detected in the eluted fractions in SDS gel as well as in Western blot (**Figure 103**).

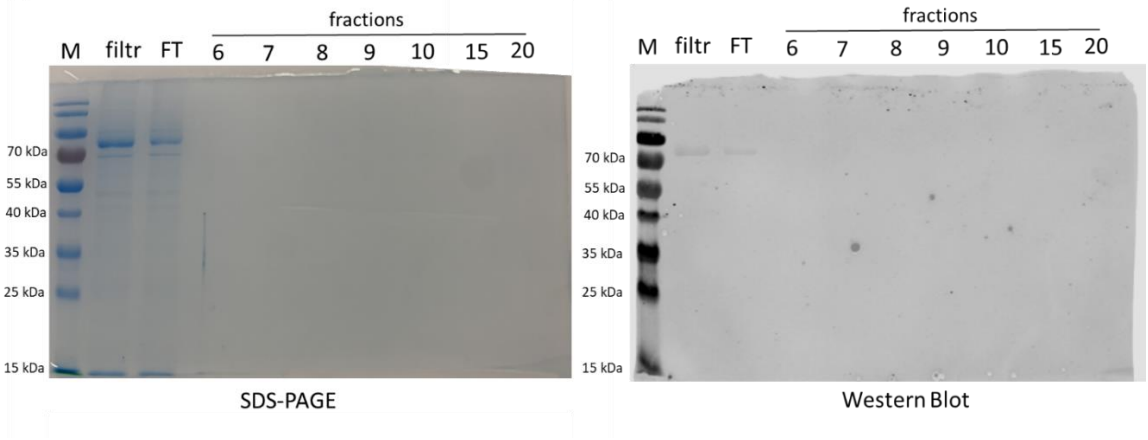


***Figure 103**: **GG domain purification**. SDS gel (left) and Western blot (right) of the fractions after purification of the GG domain. No GG domain protein was detected. Filtr = filtered; FT = flow-through.*

A collaboration was established with the laboratory of Andreas Boland (University of Geneva) to work on the GG domain crystallization in June 2020. Expression was eventually carried out in ExpiCHO cells in EPFL (Protein Production and Structure Core Facility, Lausanne) and the purification in Prof. Boland's laboratory in Geneva (**Figure 104**).

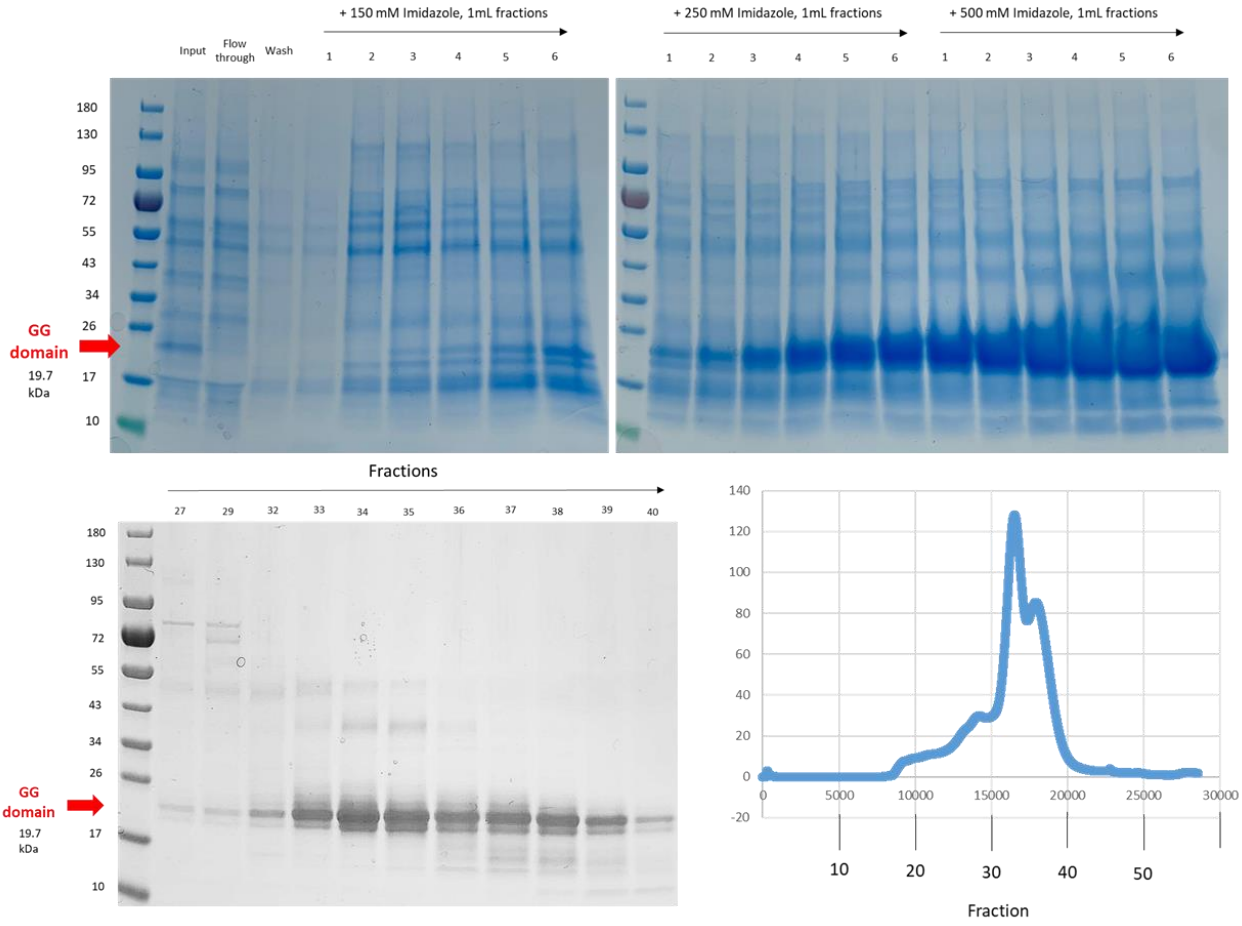


***Figure 104**: **GG domain purification**. Top and bottom left: SDS gel of fractions collected after nickel column purification; bottom right: fractions collected after SEC purification of previous fractions.*

8.1.2 GG domain crystallization

The GG domain was expressed in ExpiCHO cells in EPFL and purified using a His-tag as mentioned in the previous section. Pre-crystallization test plates from different suppliers were set up in order to find a good condition of buffer, salt and polyethylene glycol (PEG). In most of the cases, lots of precipitation was observed (**Figure 105**).

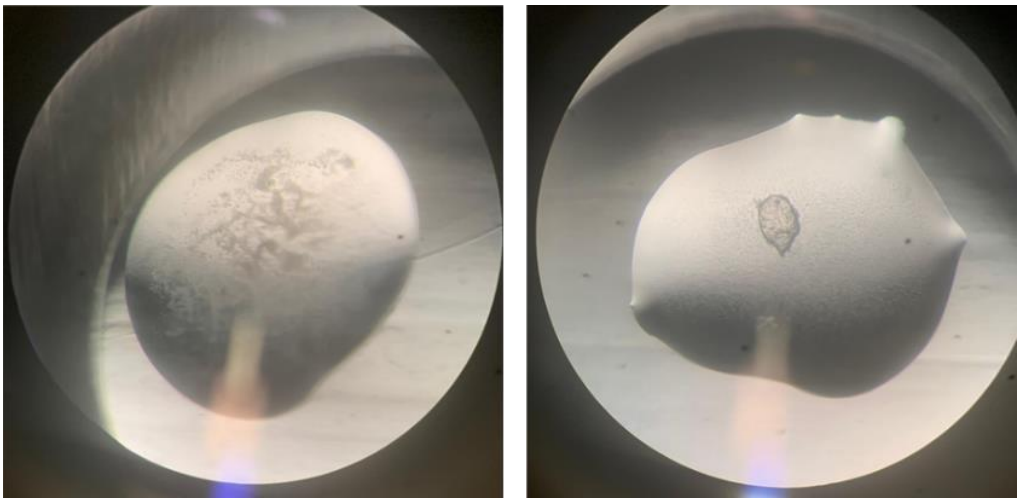


***Figure 105**: **Pre-crystallization hanging drop plate**. Lot of protein precipitation was observed. Magnification 100x.*

A pre-crystallization test with the purified GG domain revealed a condition (0.2 M Zinc acetate dehydrate, 0.1 M imidazole, pH 8.0 and 20% w/v PEG 3000) in which a potent precipitation and some crystals were observed (**Figure 106**). Unfortunately, after seeding and fishing of the crystals, the synchrotron X-ray analysis revealed the diffraction pattern of a salt.

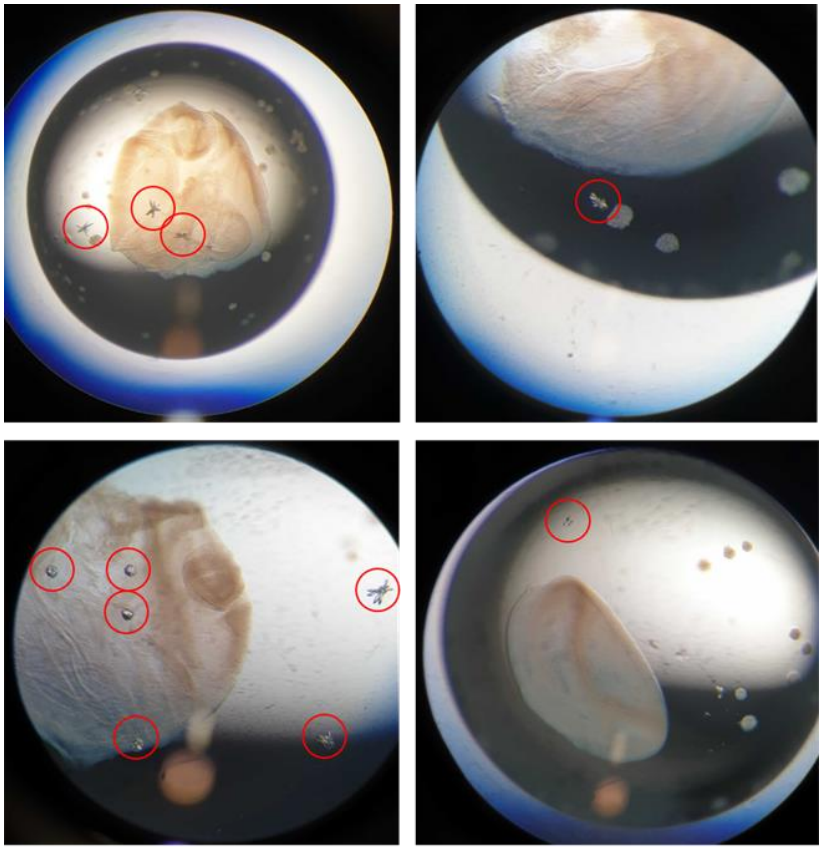


***Figure 106**: **Presumed GG domain crystals (in red circles)**. Magnification 100x.*

Next, Lina and I used limited proteolysis to obtain a more rigid protein increasing our chance of obtaining protein crystals. Using GluC proteinase for 30 minutes at 4°C we obtained a stabilized version of the GG domain (**Figure 107**). This is why we set up crystallization plates with in-situ proteolysis (hanging drop with GluC).

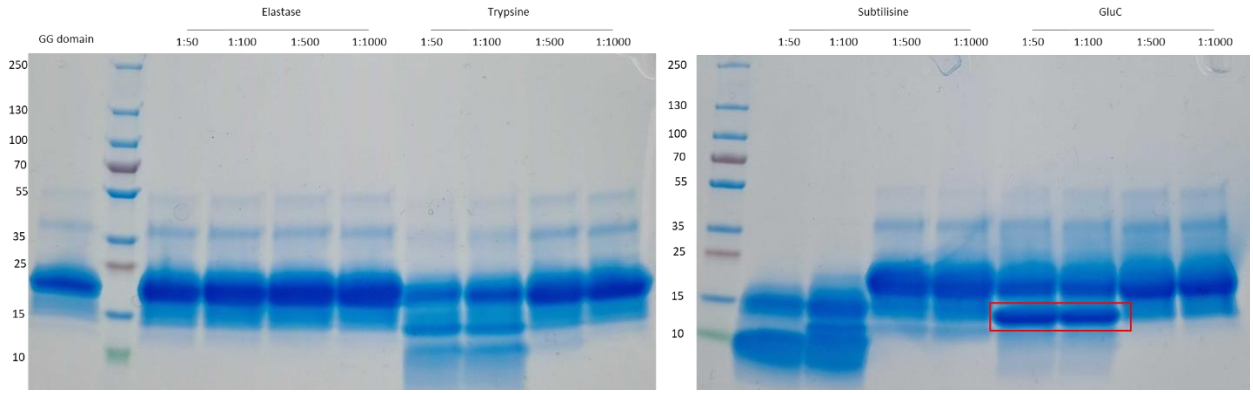


***Figure 107**: **Limited proteolysis of the GG domain with different proteases**. 30 minutes at 4°C. In red, the stabilized version obtained with GluC.*

A small crystal appeared in a new plate in the condition with 30% PEG 5000 monomethyl ether, 200 mM ammonium sulfate and MES buffer pH 6.5 and 0,02 mg/ml GluC protease (**Figure 108**). Even after waiting for some time for the crystal to grow, it was still really hard to fish, and eventually, it was lost during the fishing step.

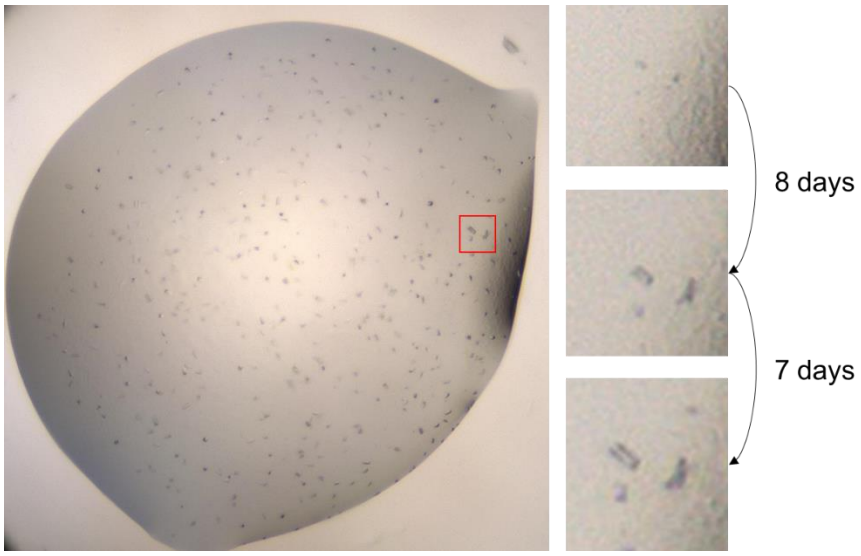


***Figure 108**: **GG domain crystal growing over time**. Magnification 100x (left) and 400x (right).*

New GG domain was produced in order to screen again all the conditions and choose the best ones for further seedings. Moreover, different GG domain constructs were designed to stabilize the GG domain even more, also using antibody 12B5-2 to decrease the freedom of movement and improve crystallization. A differential scanning fluorimetry experiment of the GG domain with or without the 12B5-2 antibody revealed that low pH stabilizes the GG/antibody complex (annex **Figure 138**). Nevertheless, at the time of writing this manuscript, no GG domain crystal was obtained yet, despite a great amount of pre-crystallization trials with different constructs and conditions.

8.2 First steps in solving the structure of CEMIP by cryogenic electron microscopy

Cryo-electron microscopy (cryo-EM) is a powerful method for solving 3D structures for a broad range of biomolecules. Until now, proteins needed to be crystallized in the right conditions and then analyzed with X-ray crystallography. Typically, crystallization can be time-consuming and sometimes doesn't work because some proteins simply don't crystallize. Cryo-EM, which includes

applications like single particle analysis (SPA) and cryo-electron tomography (cryo-ET), has revolutionized how we study proteins, protein complexes, and their dynamics. Moreover, predictions from AlphaFold can be combined with density maps data from cryo-EM to generate molecular models[162]. In this section, I show the first steps taken to reach the tridimensional structure of CEMIP.

8.2.1 hCEMIP production in insect cells

Since the amount of purified CEMIP from GenScript was limited and cryo-EM required a lot more protein, recombinant CEMIP was expressed in insect cells sf9. CEMIP was expressed in the pellet as well as in the supernatant, and elution fractions confirmed the presence of CEMIP together with some chaperone proteins (**Figure 109**).

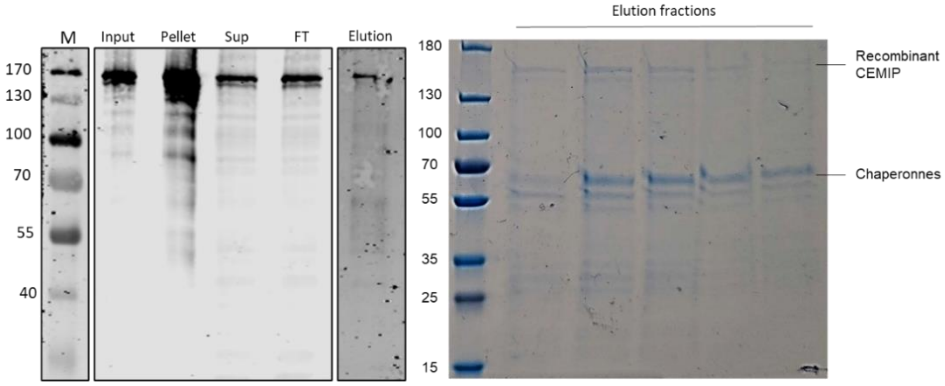


***Figure 109**: **CEMIP expression and purification in insect cells**. Left: Western blot confirms the presence of CEMIP in the pellet and in the supernatant. Right: SDS gel of elution fractions.*

Confirmation about CEMIP expression came also from mass spectrometry analyses, with a 54% coverage (**Figure 110**).

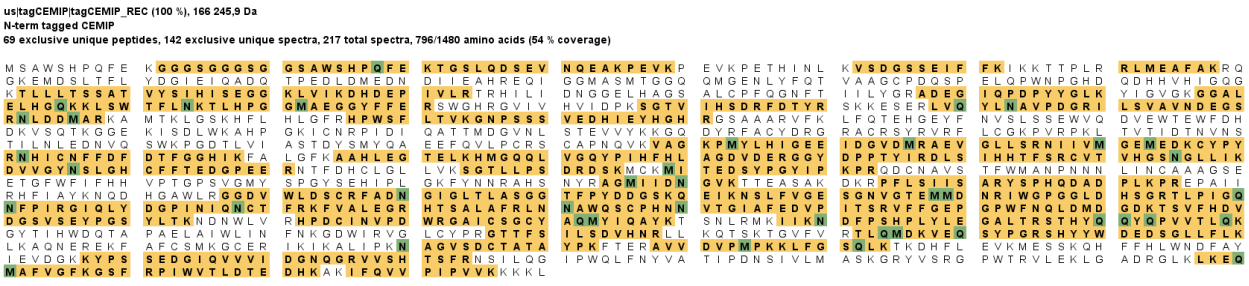


***Figure 110**: **Mass spectrometry of CEMIP**. The presence of CEMIP is confirmed with a 54% coverage (in yellow).*

While working on perfecting the production of CEMIP, Lina and I proceeded with the purified protein from GenScript with some negative stain and cryo-EM experiments to find the perfect grid and condition.

8.2.2 CEMIP cryo-EM

First, negative staining of the sample was made, collecting hundreds of micrographs. Next, all pictures were analyzed with CryoSPARC (Cryo-EM Single Particle Ab-Initio Reconstruction and Classification), a software that allows picking particles and reconstructing 2D classes with different "views" of CEMIP, as shown in **Figure 111**.

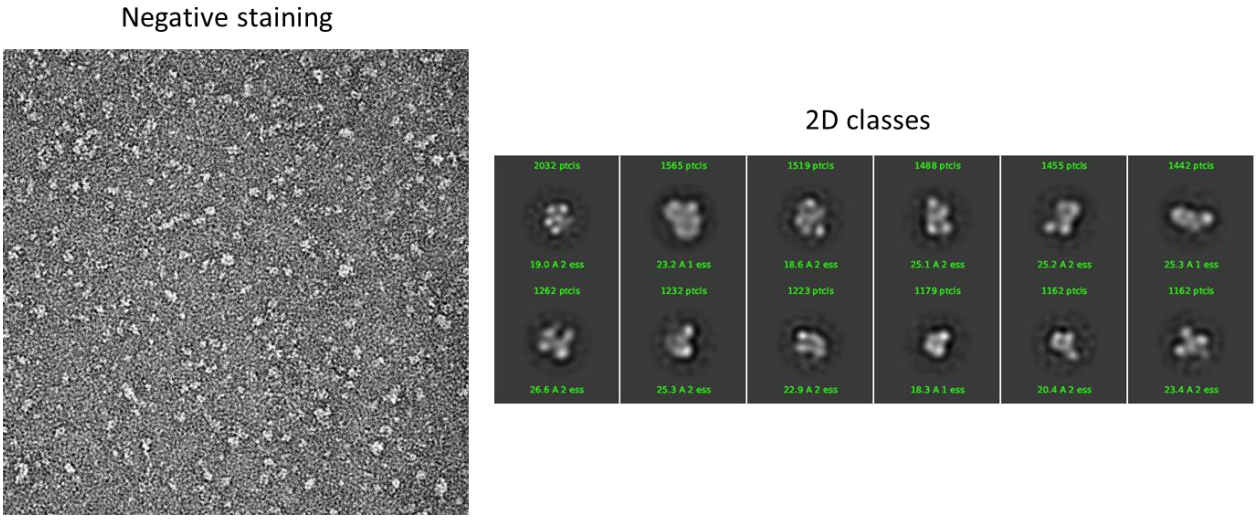


***Figure 111**: **CEMIP negative staining**. Left: micrograph taken with microscope Talos 120 LC (ThermoFisher), pixel size 1.19, 120 kV, 2.7 spherical, 40 electron dose, box size for extraction 320 pixel; right: 2D classes.*

Next, with CryoSPARC we were able to create a 3D *ab-initio* reconstruction of CEMIP (**Figure 112**).

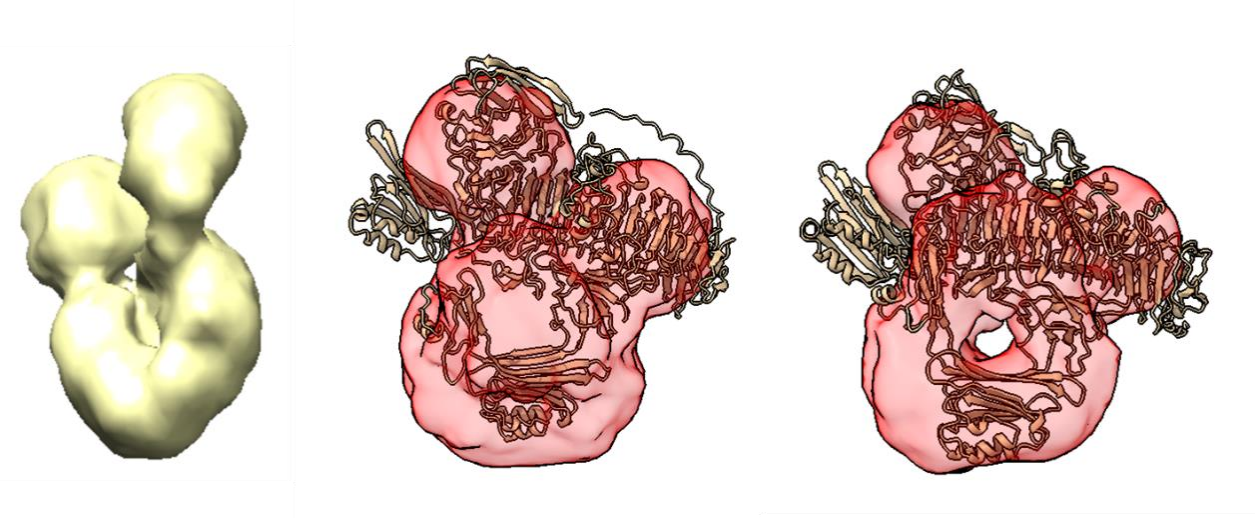


***Figure 112**: **3D ab-initio reconstruction**. 3D ab-initio reconstruction in yellow and red, superposed to the AlphaFold prediction of CEMIP.*

Collection on a Titan Krios 300kV in the Dubochet Center for Imaging (DCI) in Lausanne allowed the collection of 30'000 cryo-EM micrographs. After processing, a first model of CEMIP tridimensional structure was obtained at a 10 Å resolution, more detailed compared to the one reconstructed from negative stain micrographs (**Figure 113**).

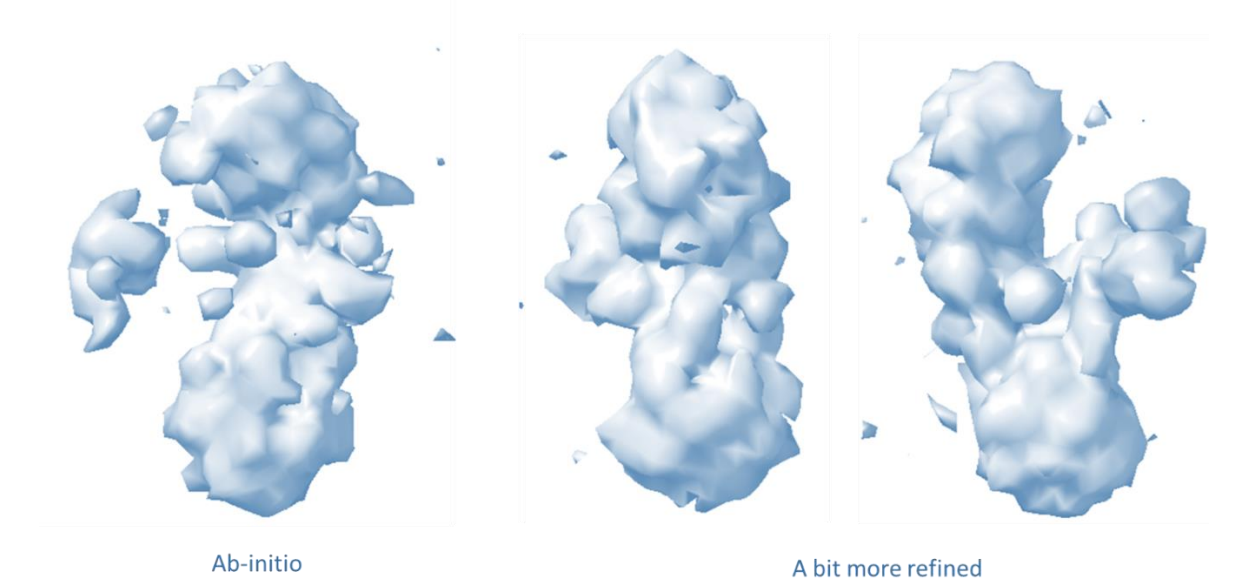


***Figure 113**: **Cryo-EM of CEMIP**. Left: Ab-initio reconstruction. Right: more refined model.*

This result provided solid proof of the value of the pursuance of working on the structure of CEMIP by cryo-EM. At the moment of drafting this manuscript, efforts are still in place to collect data and process them to ensure solving CEMIP as soon as possible.

In the meantime, we looked at CEMIP with sHA in negative stain. Although sHA filaments alone are not visible, when mixed together with CEMIP at the right concentration and with the right grid, sHA filaments are visible in the pictures, with CEMIP clearly attached to them (**Figure 114**).

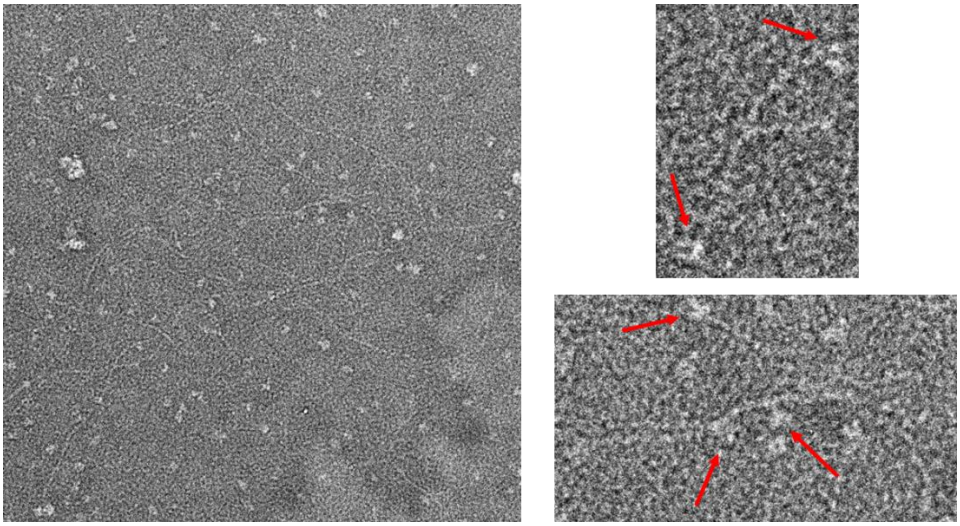


***Figure 114**: **Negative stain of CEMIP and sHA**. Red arrows indicate CEMIP.*

The observed phenomenon was increased when the grid was positively charged (**Figure 115**).

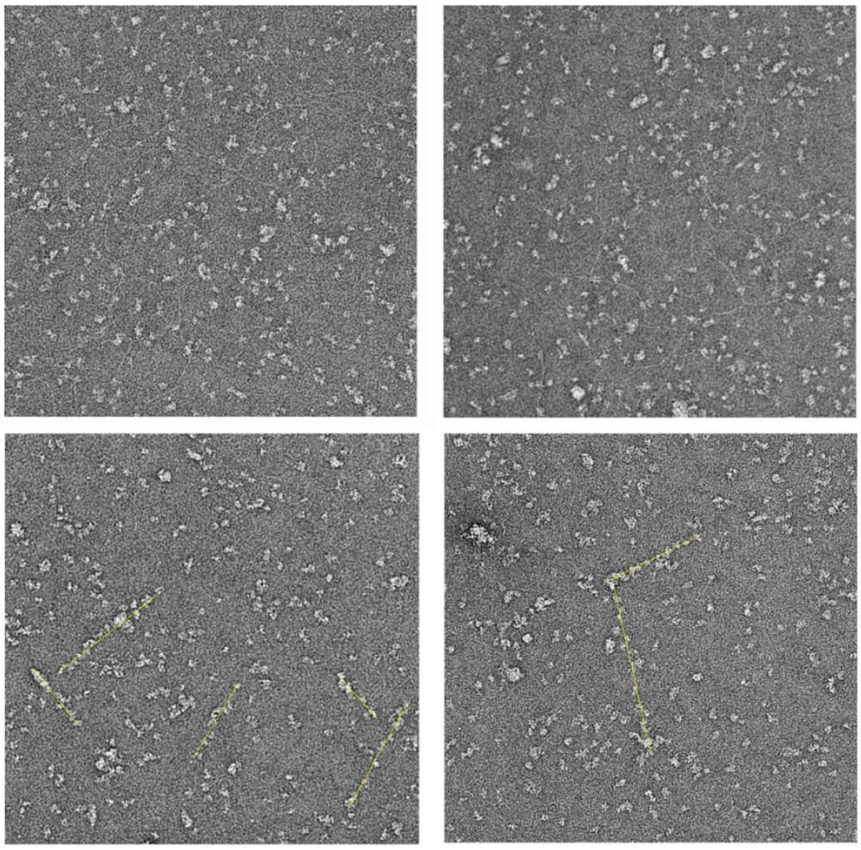


***Figure 115**: **Negative stain of CEMIP and sHA with a positively charged grid**. Bottom: yellow lines indicate CEMIP aligned on sHA filaments.*

8.3 Discussion

In this chapter, I demonstrated that the GG domain of CEMIP can be expressed in mammalian cells and purified. For the expression, the EPFL core facility can be an advantage in terms of reducing time and effort. The availability of pure GG domain allowed the setup of tens of crystallization plates for screening, and the discovery of several conditions of interest. Proteolysis of the GG domain protein, particularly with GluC, emerged as an interesting technique to gain further insight into GG structure.

With the new available GG domain, I plan to establish new plates using *in situ* proteolysis, and screen all the conditions to find a crystal to fish and shoot with X-ray.

On the other hand, the production of the entire CEMIP protein in-house seems quite complex and the best option remains to order new protein from the provider (GenScript) used so far. With the right conditions, I showed that negative stain is a preliminary additional technique already enabling

3D *ab-initio* reconstruction of CEMIP. Once the quality of the sample was confirmed, Lina and I collected data on a more powerful microscope in the Dubochet Center for Imaging (DCI) in Lausanne and obtained the first model of CEMIP tridimensional structure at a 10 Å resolution after processing.

Although AlphaFold predicts relatively well the domains' structure, their exact rearrangement in the space remains elusive, and revealing CEMIP structure by cryoEM could solve this problem. Recently, the structure of soluble TMEM2 was solved using X-ray crystallography, giving us a new input in the material and methods used, that could be partially valid for the crystallization of the GG domain and for CEMIP cryoEM, as well as a mode of comparison from the AlphaFold model[54]. They found that the TMEM2 beta-helix region is predicted correctly, also including the G8 domain[54], which is probably the case for CEMIP too. Moreover, in contrast with the previous hypothesis stating that TMEM2 lacks a second GG domain[40], and in agreement with the AlphaFold prediction, they confirmed that TMEM2 contains a second GG domain. This was also my observation: when doing a structural alignment of the C-terminal portion of TMEM2 (AA 1209-1355) and CEMIP (AA 1209-1353) I underlined the possible presence of a second GG domain in TMEM2[1]. Interestingly, Niu and colleagues observed a high similarity of the two GG domains with the lectin-like domain of POMGNT1 (protein O-linked-mannose β-1,2-N-acetylglucosaminyltransferase 1)[54]. The newly solved TMEM2 structure gives us hope for the continuation of my work on CEMIP. All efforts are now focused on having the structure of CEMIP at a higher resolution (< 5 Å) in the near future.

8.4 Materials and methods

Detailed information about the figures and the material and methods used in this chapter can be found in the doctoral thesis of Lina Poulain (University of Geneva).

9 CONCLUSIONS AND PERSPECTIVES

With this work, I aimed to investigate the DDR1/CEMIP/HA axis to identify new potential therapeutic strategies in the context of glomerular diseases. Specifically, the aim of this thesis was **a)** to elucidate CEMIP mode of action and domain function; **b)** to come up with CEMIP tool compound inhibitors and **c)** to resolve CEMIP 3D structure.

I partially elucidated the mode of action of CEMIP and the role of its domains in the binding and depolymerization of HA (**a**). I showed that the Arg^{187} of CEMIP's first GG domain is binding hyaluronic acid and that the second GG domain has an anchoring function. Asp^{208} seems to be involved in the HA cleavage, potentially as a water activator for the hydrolysis reaction. The tunnel-like structure formed by the G8 domain and the PbH1 repeats is extremely important for both HA binding and HA cleavage. In this region, I identified with selective protein mutagenesis key catalytic residues in the amino acids His560, Asp566 and His611 involved in HA cleavage. These experimental findings induced me to hypothesize the presence of a zinc ion coordinating the hydrolysis of hyaluronic acid. Even if these histidines were predicted to be zinc-binding residues, it cannot be excluded that the ion coordinating the catalysis could also be calcium, especially if we consider that TMEM2's activity is calcium-dependent. Experiments involving chelating agents or a mass spectrometry analysis are required to further elucidate this hypothesis. Furthermore, I showed that HA cleavage by CEMIP occurs only in living cells and that the secreted and membrane-bound CEMIP are inactive. My results are in line with findings from Yoshida *et al.* These authors tested the recombinant CEMIP at pH 4, 5, 6 and 7 detecting no HA degrading activity[37].

Recent unpublished findings reported the generation of two novel antibodies against the G8 domain blocking the interaction with HA (Dr. Stephen Fink, Case Western Reserve University, USA - *personal communication*). The G8 domain blockade using these novel antibodies might be an alternative targeting approach to CEMIP. A collaboration with Dr. Fink is currently established and I plan to upscale the production of these antibodies and test them in functional assay as well as in a pulldown experiment. Based on these observations I will explore the G8 further with a a dedicated mutagenesis study. Mutations in this region could elucidate important amino acids cooperating with the Arg^{187} of the first GG domain in the binding of hyaluronic acid, potentially

providing a more detailed characterization of HA binding to CEMIP. In order to test if the size of the CEMIP-produced HA fragments might be dependent on the length of CEMIP tunnel-like structure, other mutants can be generated, for example by adding one or more PbH1 repeats. Moreover, it might be of interest to additionally investigate the C-terminal part of CEMIP, which preservation I found to be essential for HA degradation.
As we advance in the comprehension of CEMIP mode of action, the next steps of the project will be to elucidate the function of the HA fragments produced by CEMIP and to further explore the CEMIP migration effect on other cells. On another note, preliminary pulldown experiments of different transient transfected mutant sCEMIP (not shown) highlighted the need to create stable mutant cell lines after observing that in the supernatant there is not enough CEMIP to perform the pulldown.

The present thesis work allowed the identification, for the first time, of hit molecules (WX-2, WX-3, CEMIP-137, FDG-013, FDG-037, FDG-038 and FDG-065) inhibiting CEMIP in a 50-125 µM molecular range (**b**). These molecules do not impair cell proliferation and do not show overt toxicity. Specific binding mode of these molecules to CEMIP remains to be elucidated.
The molecular knowledge generated by the mutagenesis screen on CEMIP will be further exploited in the future to specifically target the hypothetical CEMIP catalytic site for instance with a differential cell-DEL screen using WT and mutated CEMIP proteins. Creating stable cell lines of mutants of interest will allow us not only to perform the differential DEL screening but also to perform the pulldown experiments with the mutated secreted CEMIP mentioned above. In the future, additional analogs of these seven molecules will be synthesized and tested.
Initially developed as potential inhibitors, selective monoclonal anti-CEMIP antibodies did not show any inhibitory activity. However, antibody 12B5-2 was further validated as a reliable tool for Western blot and immunohistochemistry experiments, with a ready-to-use in-house production available to supplement ongoing research efforts.

Finally, this project successfully set the first steps toward the realization of the full structure of CEMIP (**c**). In collaboration with Lina Poulain and Andreas Boland, I elucidated CEMIP 3D structure at a medium resolution (10 Å) and I found potentially interesting conditions for the crystallization for CEMIP's first GG domain. CEMIP structure will give an important contribution

to our current knowledge, most importantly in terms of understanding the key amino acids spatial orientation towards hyaluronic acid filaments, and in terms of modelling the docking of small molecules.

The end goal of this project was to define if CEMIP could be considered a potential therapeutic target to develop a treatment for GN conferring renal protection. Translational and preliminary preclinical evidence suggest a potential causal role for the DDR1/CEMIP/HA axis in renal pathologies characterized by a crescentic phase. The role of CEMIP in mouse models of crescentic GN will be further investigated with new set of experiments. CEMIP floxed mice will be bred with Pax8-Cre mice to delete CEMIP specifically in renal epithelial cells, and then these mice will be exposed to nephrotoxic serum (NTS) to induce crescentic glomerulonephritis. We will assess renal function differences by measuring proteinuria and other parameters between these mice and control groups. Histological and immunohistochemical analyses will determine whether the loss of CEMIP significantly reduces the development of crescentic GN. These results will serve as a baseline for a subsequent set of experiments, where the potential protective effect of CEMIP inhibitors will be evaluated in a mouse model of crescentic GN. Since CEMIP is not expressed only in the kidneys, it would be difficult to achieve drug delivery targeted to the kidney and have selectivity in order to avoid off-site effects. For this reason, we can think about using the treatment during acute phases when the patient needs urgent treatment, and the delivery could be intravenous.

To determine the suitability of CEMIP as a therapeutic target for GN, further research will be necessary. A better understanding of CEMIP protein role in GN might also profit other disease areas where CEMIP has been shown to be implicated such as colon cancer, inflammatory bowel disease and osteoarthritis.

10 SCIENTIFIC CONTRIBUTION

Publications in peer-reviewed journals:

- Schmaus A, Rothley M, Schreiber C, Möller S, Roßwag S, Franz S, Garvalov BK, Thiele W, **Spataro S**, Herskind C, Prunotto M, Anderegg U, Schnabelrauch M, Sleeman J. Sulfated hyaluronic acid inhibits the hyaluronidase CEMIP and regulates the HA metabolism, proliferation and differentiation of fibroblasts. Matrix Biol. 2022 May; 109:173-191. doi: 10.1016/j.matbio.2022.04.001.
- **Spataro S**, Guerra C, Cavalli A, Sgrignani J, Sleeman J, Poulain L, Boland A, Scapozza L, Moll S, Prunotto M. CEMIP (HYBID, KIAA1199): structure, function and expression in health and disease. FEBS J. 2022 Aug 23. doi: 10.1111/febs.16600.

Conferences actively attended:

- Swiss Chemical Society (SCS) Fall Meeting 2020, 25th August 2020, online; "Exploring the role of CEMIP in Alport syndrome" poster
- Frontiers in Medicinal Chemistry, 8th March 2021, online: "Exploring the role of CEMIP in Alport syndrome" poster
- International Society for Hyaluronan Sciences (ISHAS) 2021, 14th June 2021: poster and presentation "Exploring the role of CEMIP in Alport syndrome"
- 8th Rare Disease Summer School ITINERARE, 13th July 2022, Zurich; poster and presentation "Exploring the role of CEMIP in Alport syndrome"
- PhD Forum PSLS 2023, 20th June 2023, Geneva; "Exploring the role of CEMIP in Alport syndrome"; presentation

Complementary activities:

- Master thesis supervisions: Filippo Donati (October 2020 - May 2021), Marko Durakovic (May 2021 - January 2022), University of Geneva
- Teaching Assistant for Travaux Pratiques (TP) in « Pharmacochimie: ADME / PK », 4 semesters (2020-2023), University of Geneva
- 8th Rare Disease Summer School ITINERARE, 13th July 2022, Zurich

- MT180 competition training and semi-finalist (December 2022 – March 2023), University of Geneva
- WIPO-UNIGE Intellectual Property Summer School 2023, 19th-30th June 2023, Geneva

Exploring the role of CEMIP in Alport syndrome

sofia.spataro@unige.ch

LinkedIn

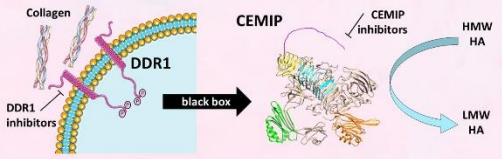
Sofia Spataro[1], Lina Poulain[2], Filippo Donati[1], Jacopo Sgrignani[3], Andrea Cavalli[3], Leonardo Scapozza[1], Marco Prunotto[1]

[1]*School of Pharmaceutical Sciences, University of Geneva, Switzerland*
[2]*Department of Molecular Biology, University of Geneva, Switzerland*
[3]*Institute for Research in Biomedicine, Università della Svizzera Italiana, Switzerland*

Alport syndrome

▻ Alport syndrome (AS) is a genetic rare disease affecting **1 in 10'000** new-borns. It is caused by mutations affecting **type IV collagen**, essential component of the **glomerular basement membrane (GBM)** found in the glomerulus of the kidney, in organ of Corti in the inner ear and in the eye.

▻ This leads to GBM thinning and splitting, which causes injuries in **podocytes** preventing the kidneys from properly filtering the blood. Current treatment aims to control the progression of the disease and treat the symptoms, by reducing the blood pressure and filtration of large molecules (ACE inhibitors, aldosterone inhibitors). However, chronic kidney failure progresses to **end-stage kidney disease**, requiring dialysis or transplantation.



DDR1/CEMIP/HA axis

▻ Podocytes are able to sense the GBM via the **discoidin domain receptor 1 (DDR1)**, which binds collagen, and plays a major role in the onset and progression of AS. After observing reduction of renal damage in DDR1 genetic deletion mouse models, a series of DDR1 inhibitors was developed, but the high concentrations required for clinical efficacy would also inhibit the closely related DDR2 receptor, causing inacceptable side effects.

▻ In order to overcome the limited DDR1 inhibition, we are interested in the mode of action of this receptor to understand the downstream pathway activated by collagen. This led us to show that in presence of collagen DDR1 selectively induces the production of a protein called **CEMIP**, known to be involved in **hyaluronic acid (HA)** depolymerization.

Project aim

*The aim of this project is to explore the biology of the **DDR1/CEMIP/HA axis** in Alport syndrome, and potentially work on CEMIP inhibition as treatment for this rare disease.*

*Is the **DDR1/CEMIP/HA** axis a way to address the unmet medical need in Alport syndrome?*

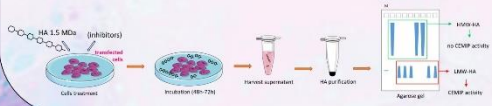
CEMIP binds HA and is involved in its depolymerization

▻ CEMIP = cell migration-inducing and hyaluronan-binding protein (KIAA1199, HYBID)

▻ binds HA and is involved in its depolymerization (Yoshida, 2013)

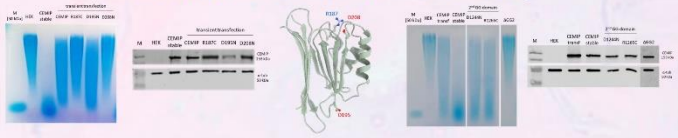
▻ introduce mutations in order to explore CEMIP domains' function

HA assay as a readout of CEMIP hyaluronidase activity

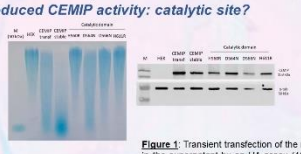


CEMIP domains function

Mutants R187C, D208N and ΔGG2 have reduced CEMIP activity



Mutants H560R, D566N and H611R have reduced CEMIP activity: catalytic site?



Mutants ΔG8 and ΔPbH1 have reduced CEMIP activity

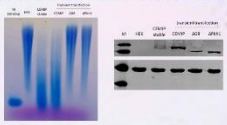


Figure 1: Transient transfection of the mutated hCEMIP plasmids in HEK293T cells followed by analysis of the HA present in the supernatant by an HA assay (48h incubation). Mutants R187C, D208N, ΔGG2, H560R, D566N, H611R, ΔG8 and ΔPbH1 show a reduction in CEMIP activity and a solid CEMIP expression.

CEMIP mechanism and structure

CEMIP is active only in cells that are producing it

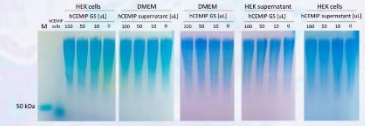


Figure 2: Different conditions of CEMIP and cells or medium was tested in order to select the variables by which CEMIP is active. GS = GenScript protein, DMEM = Dulbecco's Modified Eagle Medium.

Activity of secreted CEMIP is not rescued by ANXA1

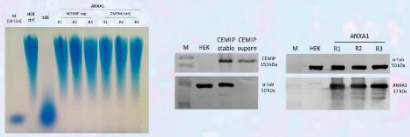
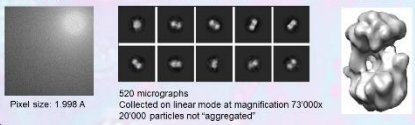


Figure 3: hCEMIP from the supernatant was incubated with HEK293T cells transfected with ANXA1 in order to test whether the activity is rescued. h88 = hCEMIP stable clone, DMEM = Dulbecco's Modified Eagle Medium, sup/supern = supernatant.

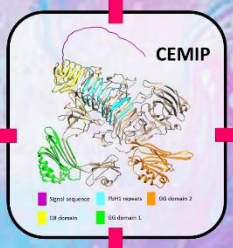
Preliminary structure of CEMIP in cryoEM

Micrograph | *2D classes of picked particles* | *3D ab-initio reconstruction*



Pixel size: 1.998 Å

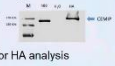
520 micrographs
Collected on linear mode at magnification 73'000x
20'000 particles not "aggregated"



Ongoing

▻ New mutants' generation:

▻ Are mutants binding HA?



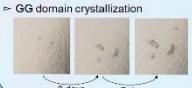
▻ SEC protocol for HA analysis

▻ Anti-CEMIP antibody production

▻ Exploring activation of CEMIP at the membrane

▻ Computational studies: WX-2 and CEMIP-137 modifications

▻ GG domain crystallization



Inhibition of CEMIP

Molecule WX-2 inhibits HA depolymerization

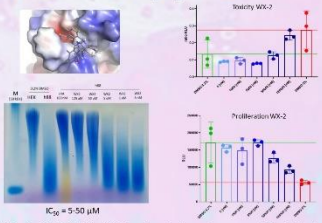


Figure 4: Molecule WX-2 inhibits HA depolymerization at a concentration that doesn't affect proliferation and without toxicity at IC_{20} concentration. WX-2 was incubated for 48h with HEK293T stably expressing CEMIP in the presence of HA. HA in the supernatant was analyzed by an HA assay. Toxicity and cell proliferation was assessed by a ToxiLight assay (right). h88 = hCEMIP stable clone.

Molecule CEMIP-137 inhibits HA depolymerization

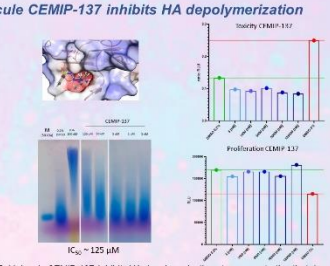


Figure 5: Molecule CEMIP-137 inhibits HA depolymerization at a concentration that doesn't affect proliferation and without toxicity. CEMIP-137 was incubated for 48h with HEK293T stably expressing CEMIP in the presence of HA. HA in the supernatant was analyzed by an HA assay. Toxicity and cell proliferation was assessed by a ToxiLight assay (right). h88 = hCEMIP stable clone.

References

Gross, O., et al., Loss of collagen-receptor DDR1 delays renal fibrosis in hereditary type IV collagen disease. *Matrix Biology*, 2010. 29(5): p. 346-56.
Omachi, K. and J.H. Miner, Alport Syndrome Therapeutics: Ready for Prime-Time Players. *Trends in Pharmacological Sciences*, 2019. 40(11): p. 803-806.
Yamamoto, H., et al., A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. *J Biol Chem*, 2017. 292(18): p. 7304-7313.
Yoshida, H., et al., KIAA1199, a deafness gene of unknown function, is a new hyaluronan binding protein involved in hyaluronan depolymerization. *Proc Natl Acad Sci USA*, 2013. 110(14): p. 5612-7.

11 BIBLIOGRAPHY

1 Spataro, S. *et al.* CEMIP (HYBID, KIAA1199): structure, function and expression in health and disease. *The FEBS journal*, doi:10.1111/febs.16600 (2022).

2 Eymael, J. *et al.* CD44 is required for the pathogenesis of experimental crescentic glomerulonephritis and collapsing focal segmental glomerulosclerosis. *Kidney international* **93**, 626-642, doi:10.1016/j.kint.2017.09.020 (2018).

3 Lazareth, H. *et al.* The tetraspanin CD9 controls migration and proliferation of parietal epithelial cells and glomerular disease progression. *Nat Commun* **10**, 3303, doi:10.1038/s41467-019-11013-2 (2019).

4 Bollée, G. *et al.* Epidermal growth factor receptor promotes glomerular injury and renal failure in rapidly progressive crescentic glomerulonephritis. *Nature medicine* **17**, 1242-1250, doi:10.1038/nm.2491 (2011).

5 Alport, A. C. Hereditary familial congenital haemorrhagic nephritis. *British medical journal* **1**, 504-506, doi:10.1136/bmj.1.3454.504 (1927).

6 Flinter, F. Alport's syndrome. *Journal of medical genetics* **34**, 326-330, doi:10.1136/jmg.34.4.326 (1997).

7 Kruegel, J., Rubel, D. & Gross, O. Alport syndrome--insights from basic and clinical research. *Nature reviews. Nephrology* **9**, 170-178, doi:10.1038/nrneph.2012.259 (2013).

8 Gregorio, V., Caparali, E. B., Shojaei, A., Ricardo, S. & Barua, M. Alport Syndrome: Clinical Spectrum and Therapeutic Advances. *Kidney medicine* **5**, 100631, doi:10.1016/j.xkme.2023.100631 (2023).

9 Savige, J. Alport syndrome: its effects on the glomerular filtration barrier and implications for future treatment. *The Journal of physiology* **592**, 4013-4023, doi:10.1113/jphysiol.2014.274449 (2014).

10 Kashtan, C. E. & Michael, A. F. Alport syndrome. *Kidney international* **50**, 1445-1463, doi:10.1038/ki.1996.459 (1996).

11 Watson, S., Padala, S. A., Hashmi, M. F. & Bush, J. S. in *StatPearls* (StatPearls Publishing
Copyright © 2023, StatPearls Publishing LLC., 2023).

12 Ortiz, A. *et al.* Translational value of animal models of kidney failure. *European journal of pharmacology* **759**, 205-220, doi:10.1016/j.ejphar.2015.03.026 (2015).

13 Gross, O. *et al.* Loss of collagen-receptor DDR1 delays renal fibrosis in hereditary type IV collagen disease. *Matrix Biol* **29**, 346-356, doi:10.1016/j.matbio.2010.03.002 (2010).

14 Cosgrove, D. & Liu, S. Collagen IV diseases: A focus on the glomerular basement membrane in Alport syndrome. *Matrix Biol* **57-58**, 45-54, doi:10.1016/j.matbio.2016.08.005 (2017).

15 Funk, S. D., Lin, M. H. & Miner, J. H. Alport syndrome and Pierson syndrome: Diseases of the glomerular basement membrane. *Matrix Biol* **71-72**, 250-261, doi:10.1016/j.matbio.2018.04.008 (2018).

16 Gross, O. *et al.* Preemptive ramipril therapy delays renal failure and reduces renal fibrosis in COL4A3-knockout mice with Alport syndrome. *Kidney international* **63**, 438-446, doi:10.1046/j.1523-1755.2003.00779.x (2003).

17 Gross, O. *et al.* Early angiotensin-converting enzyme inhibition in Alport syndrome delays renal failure and improves life expectancy. *Kidney international* **81**, 494-501, doi:10.1038/ki.2011.407 (2012).

18 Torra, R. & Furlano, M. New therapeutic options for Alport syndrome. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*, doi:10.1093/ndt/gfz131 (2019).
19 Omachi, K. & Miner, J. H. Alport Syndrome Therapeutics: Ready for Prime-Time Players. *Trends in pharmacological sciences* **40**, 803-806, doi:10.1016/j.tips.2019.07.012 (2019).
20 Zhu, Z. *et al.* Finerenone Added to RAS/SGLT2 Blockade for CKD in Alport Syndrome. Results of a Randomized Controlled Trial with Col4a3-/- Mice. *Journal of the American Society of Nephrology : JASN* **34**, 1513-1520, doi:10.1681/asn.0000000000000186 (2023).
21 Leitinger, B. Discoidin domain receptor functions in physiological and pathological conditions. *International review of cell and molecular biology* **310**, 39-87, doi:10.1016/b978-0-12-800180-6.00002-5 (2014).
22 Abdulhussein, R., McFadden, C., Fuentes-Prior, P. & Vogel, W. F. Exploring the collagen-binding site of the DDR1 tyrosine kinase receptor. *J Biol Chem* **279**, 31462-31470, doi:10.1074/jbc.M400651200 (2004).
23 Curat, C. A. & Vogel, W. F. Discoidin domain receptor 1 controls growth and adhesion of mesangial cells. *Journal of the American Society of Nephrology : JASN* **13**, 2648-2656, doi:10.1097/01.asn.0000032419.13208.0c (2002).
24 Moll, S. *et al.* DDR1 role in fibrosis and its pharmacological targeting. *Biochimica et biophysica acta. Molecular cell research* **1866**, 118474, doi:10.1016/j.bbamcr.2019.04.004 (2019).
25 Guerrot, D. *et al.* Discoidin domain receptor 1 is a major mediator of inflammation and fibrosis in obstructive nephropathy. *Am J Pathol* **179**, 83-91, doi:10.1016/j.ajpath.2011.03.023 (2011).
26 Gross, O. *et al.* DDR1-deficient mice show localized subepithelial GBM thickening with focal loss of slit diaphragms and proteinuria. *Kidney international* **66**, 102-111, doi:10.1111/j.1523-1755.2004.00712.x (2004).
27 Vogel, W. F., Aszódi, A., Alves, F. & Pawson, T. Discoidin domain receptor 1 tyrosine kinase has an essential role in mammary gland development. *Molecular and cellular biology* **21**, 2906-2917, doi:10.1128/mcb.21.8.2906-2917.2001 (2001).
28 Flamant, M. *et al.* Discoidin domain receptor 1 null mice are protected against hypertension-induced renal disease. *Journal of the American Society of Nephrology : JASN* **17**, 3374-3381, doi:10.1681/asn.2006060677 (2006).
29 Chiusa, M. *et al.* The Extracellular Matrix Receptor Discoidin Domain Receptor 1 Regulates Collagen Transcription by Translocating to the Nucleus. *Journal of the American Society of Nephrology : JASN* **30**, 1605-1624, doi:10.1681/asn.2018111160 (2019).
30 Kerroch, M. *et al.* Genetic inhibition of discoidin domain receptor 1 protects mice against crescentic glomerulonephritis. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **26**, 4079-4091, doi:10.1096/fj.11-194902 (2012).
31 Moll, S. *et al.* Selective pharmacological inhibition of DDR1 prevents experimentally-induced glomerulonephritis in prevention and therapeutic regime. *Journal of translational medicine* **16**, 148, doi:10.1186/s12967-018-1524-5 (2018).
32 Richter, H. *et al.* DNA-Encoded Library-Derived DDR1 Inhibitor Prevents Fibrosis and Renal Function Loss in a Genetic Mouse Model of Alport Syndrome. *ACS chemical biology* **14**, 37-49, doi:10.1021/acschembio.8b00866 (2019).

33 Li, Y., Lu, X., Ren, X. & Ding, K. Small molecule discoidin domain receptor kinase inhibitors and potential medical applications. *Journal of medicinal chemistry* **58**, 3287-3301, doi:10.1021/jm5012319 (2015).
34 Liu, L. *et al.* Synthesis and biological evaluation of novel dasatinib analogues as potent DDR1 and DDR2 kinase inhibitors. *Chemical biology & drug design* **89**, 420-427, doi:10.1111/cbdd.12863 (2017).
35 Olaso, E., Arteta, B., Benedicto, A., Crende, O. & Friedman, S. L. Loss of discoidin domain receptor 2 promotes hepatic fibrosis after chronic carbon tetrachloride through altered paracrine interactions between hepatic stellate cells and liver-associated macrophages. *Am J Pathol* **179**, 2894-2904, doi:10.1016/j.ajpath.2011.09.002 (2011).
36 Zhu, T. *et al.* The anti-angiogenic role of discoidin domain receptor 2 (DDR2) in laser-induced choroidal neovascularization. *J Mol Med (Berl)* **93**, 187-198, doi:10.1007/s00109-014-1213-7 (2015).
37 Yoshida, H. *et al.* KIAA1199, a deafness gene of unknown function, is a new hyaluronan binding protein involved in hyaluronan depolymerization. *Proc Natl Acad Sci U S A* **110**, 5612-5617, doi:10.1073/pnas.1215432110 (2013).
38 Ogawa, D. *et al.* Sulfated hyaluronic acid, a potential selectin inhibitor, ameliorates experimentally induced crescentic glomerulonephritis. *Nephron. Experimental nephrology* **99**, e26-32, doi:10.1159/000081795 (2005).
39 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
40 Yamamoto, H. *et al.* A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. *J Biol Chem* **292**, 7304-7313, doi:10.1074/jbc.M116.770149 (2017).
41 Yoshida, H. & Okada, Y. Role of HYBID (Hyaluronan Binding Protein Involved in Hyaluronan Depolymerization), Alias KIAA1199/CEMIP, in Hyaluronan Degradation in Normal and Photoaged Skin. *International Journal of Molecular Sciences* **20**, 5804 (2019).
42 Birkenkamp-Demtroder, K. *et al.* Repression of KIAA1199 attenuates Wnt-signalling and decreases the proliferation of colon cancer cells. *Br J Cancer* **105**, 552-561, doi:10.1038/bjc.2011.268 (2011).
43 He, Q.-y. *et al.* G8: a novel domain associated with polycystic kidney disease and non-syndromic hearing loss. *Bioinformatics* **22**, 2189-2191, doi:10.1093/bioinformatics/btl123 (2006).
44 Bernstein, F. C. *et al.* The protein data bank: A computer-based archival file for macromolecular structures. *J Mol Biol* **112**, 535-542 (1977).
45 Guo, J., Cheng, H., Zhao, S. & Yu, L. GG: a domain involved in phage LTF apparatus and implicated in human MEB and non-syndromic hearing loss diseases. *FEBS Lett* **580**, 581-584, doi:10.1016/j.febslet.2005.12.076 (2006).
46 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).
47 Abe, S., Usami, S. & Nakamura, Y. Mutations in the gene encoding KIAA1199 protein, an inner-ear protein expressed in Deiters' cells and the fibrocytes, as the cause of nonsyndromic hearing loss. *J Hum Genet* **48**, 564-570, doi:10.1007/s10038-003-0079-2 (2003).
48 Garron, M. L. & Cygler, M. Uronic polysaccharide degrading enzymes. *Curr Opin Struct Biol* **28**, 87-95, doi:10.1016/j.sbi.2014.07.012 (2014).

49 Yamaguchi, Y., Yamamoto, H., Tobisawa, Y. & Irie, F. TMEM2: A missing link in hyaluronan catabolism identified? *Matrix biology : journal of the International Society for Matrix Biology* **78-79**, 139-146, doi:10.1016/j.matbio.2018.03.020 (2019).
50 Ward, C. J. *et al.* The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. *Nat Genet* **30**, 259-269, doi:10.1038/ng833 (2002).
51 Muhammed, M. T. & Aki-Yalcin, E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chemical biology & drug design* **93**, 12-20, doi:10.1111/cbdd.13388 (2019).
52 Kuwabara, N. *et al.* Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of α-dystroglycan. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 9280-9285, doi:10.1073/pnas.1525545113 (2016).
53 Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876, doi:10.1126/science.abj8754 (2021).
54 Niu, M. *et al.* Structure of the transmembrane protein 2 (TMEM2) ectodomain and its apparent lack of hyaluronidase activity. *Wellcome open research* **8**, 76, doi:10.12688/wellcomeopenres.18937.2 (2023).
55 Narita, T. *et al.* TMEM2 is a bona fide hyaluronidase possessing intrinsic catalytic activity. *J Biol Chem* **299**, 105120, doi:10.1016/j.jbc.2023.105120 (2023).
56 Vigetti, D., Viola, M., Karousou, E., De Luca, G. & Passi, A. Metabolic control of hyaluronan synthases. *Matrix Biol* **35**, 8-13, doi:10.1016/j.matbio.2013.10.002 (2014).
57 Padhi, A. & Nain, A. S. ECM in Differentiation: A Review of Matrix Structure, Composition and Mechanical Properties. *Ann Biomed Eng* **48**, 1071-1089, doi:10.1007/s10439-019-02337-7 (2020).
58 Fallacara, A., Baldini, E., Manfredini, S. & Vertuani, S. Hyaluronic Acid in the Third Millennium. *Polymers (Basel)* **10**, doi:10.3390/polym10070701 (2018).
59 Bayer, I. S. Hyaluronic Acid and Controlled Release: A Review. *Molecules* **25**, doi:10.3390/molecules25112649 (2020).
60 Yoshida, H. *et al.* Murine homologue of the human KIAA1199 is implicated in hyaluronan binding and depolymerization. *FEBS Open Bio* **3**, 352-356, doi:10.1016/j.fob.2013.08.003 (2013).
61 Milner, C. M., Tongsoongnoen, W., Rugg, M. S. & Day, A. J. The molecular basis of inter-alpha-inhibitor heavy chain transfer on to hyaluronan. *Biochem Soc Trans* **35**, 672-676, doi:10.1042/BST0350672 (2007).
62 Supp, D. M., Hahn, J. M., McFarland, K. L. & Glaser, K. Inhibition of hyaluronan synthase 2 reduces the abnormal migration rate of keloid keratinocytes. *J Burn Care Res* **35**, 84-92, doi:10.1097/BCR.0b013e3182a2a9dd (2014).
63 Volpi, N., Schiller, J., Stern, R. & Soltes, L. Role, metabolism, chemical modifications and applications of hyaluronan. *Curr Med Chem* **16**, 1718-1745, doi:10.2174/092986709788186138 (2009).
64 Zhang, Z. *et al.* Hyaluronan synthase 2 expressed by cancer-associated fibroblasts promotes oral cancer invasion. *J Exp Clin Cancer Res* **35**, 181, doi:10.1186/s13046-016-0458-0 (2016).
65 Li, Y. *et al.* Hyaluronan synthase 2 regulates fibroblast senescence in pulmonary fibrosis. *Matrix Biol* **55**, 35-48, doi:10.1016/j.matbio.2016.03.004 (2016).

66 Zhang, H. *et al.* Hyaluronan synthase 2 is an adverse prognostic marker in androgen receptor-negative breast cancer. *J Clin Pathol* **69**, 1055-1062, doi:10.1136/jclinpath-2016-203617 (2016).
67 Stern, R. Hyaluronidases in cancer biology. *Semin Cancer Biol* **18**, 275-280, doi:10.1016/j.semcancer.2008.03.017 (2008).
68 Kaul, A., Short, W. D., Wang, X. & Keswani, S. G. Hyaluronidases in Human Diseases. *Int J Mol Sci* **22**, doi:10.3390/ijms22063204 (2021).
69 Zadnikova, P. *et al.* The Degradation of Hyaluronan in the Skin. *Biomolecules* **12**, doi:10.3390/biom12020251 (2022).
70 Csoka, A. B., Frost, G. I. & Stern, R. The six hyaluronidase-like genes in the human and mouse genomes. *Matrix Biol* **20**, 499-508, doi:10.1016/s0945-053x(01)00172-x (2001).
71 Garantziotis, S. & Savani, R. C. Hyaluronan biology: A complex balancing act of structure, function, location and context. *Matrix Biol* **78-79**, 1-10, doi:10.1016/j.matbio.2019.02.002 (2019).
72 Jing, W. & DeAngelis, P. L. Synchronized chemoenzymatic synthesis of monodisperse hyaluronan polymers. *J Biol Chem* **279**, 42345-42349, doi:10.1074/jbc.M402744200 (2004).
73 Schmaus, A. *et al.* Accumulation of small hyaluronan oligosaccharides in tumour interstitial fluid correlates with lymphatic invasion and lymph node metastasis. *Br J Cancer* **111**, 559-567, doi:10.1038/bjc.2014.332 (2014).
74 Bauer, J. *et al.* TGFbeta counteracts LYVE-1-mediated induction of lymphangiogenesis by small hyaluronan oligosaccharides. *J Mol Med (Berl)* **96**, 199-209, doi:10.1007/s00109-017-1615-4 (2018).
75 Soroosh, A. *et al.* Crohn's Disease Fibroblasts Overproduce the Novel Protein KIAA1199 to Create Proinflammatory Hyaluronan Fragments. *Cell Mol Gastroenterol Hepatol* **2**, 358-368 e354, doi:10.1016/j.jcmgh.2015.12.007 (2016).
76 Zhang, B. *et al.* INT-HA induces M2-like macrophage differentiation of human monocytes via TLR4-miR-935 pathway. *Cancer Immunol Immunother* **68**, 189-200, doi:10.1007/s00262-018-2261-6 (2019).
77 Kuang, D. M. *et al.* Tumor-derived hyaluronan induces formation of immunosuppressive macrophages through transient early activation of monocytes. *Blood* **110**, 587-595, doi:10.1182/blood-2007-01-068031 (2007).
78 Sokolowska, M. *et al.* Low molecular weight hyaluronan activates cytosolic phospholipase A2alpha and eicosanoid production in monocytes and macrophages. *J Biol Chem* **289**, 4470-4488, doi:10.1074/jbc.M113.515106 (2014).
79 Ghazi, K., Deng-Pichon, U., Warnet, J. M. & Rat, P. Hyaluronan fragments improve wound healing on in vitro cutaneous model through P2X7 purinoreceptor basal activation: role of molecular weight. *PLoS One* **7**, e48351, doi:10.1371/journal.pone.0048351 (2012).
80 Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun* **11**, 5120, doi:10.1038/s41467-020-18794-x (2020).
81 Yoshida, H. *et al.* N-Terminal signal sequence is required for cellular trafficking and hyaluronan-depolymerization of KIAA1199. *FEBS Lett* **588**, 111-116, doi:10.1016/j.febslet.2013.11.017 (2014).

82 Zhang, W. *et al.* Secreted KIAA1199 promotes the progression of rheumatoid arthritis by mediating hyaluronic acid degradation in an ANXA1-dependent manner. *Cell Death Dis* **12**, 102, doi:10.1038/s41419-021-03393-5 (2021).
83 Shostak, K. *et al.* NF-kappaB-induced KIAA1199 promotes survival through EGFR signalling. *Nat Commun* **5**, 5232, doi:10.1038/ncomms6232 (2014).
84 Usami, S. *et al.* The localization of proteins encoded by CRYM, KIAA1199, UBA52, COL9A3, and COL9A1, genes highly expressed in the cochlea. *Neuroscience* **154**, 22-28, doi:10.1016/j.neuroscience.2008.03.018 (2008).
85 Usami, S. *et al.* The responsible genes in Japanese deafness patients and clinical application using Invader assay. *Acta Otolaryngol* **128**, 446-454, doi:10.1080/00016480701785046 (2008).
86 Hosoya, M., Fujioka, M., Okano, H. & Ogawa, K. Distinct Expression Pattern of a Deafness Gene, KIAA1199, in a Primate Cochlea. *Biomed Res Int* **2016**, 1781894, doi:10.1155/2016/1781894 (2016).
87 Marella, M. *et al.* KIAA1199 expression and hyaluronan degradation colocalize in multiple sclerosis lesions. *Glycobiology* **28**, 958-967, doi:10.1093/glycob/cwy064 (2018).
88 Yang, X. *et al.* KIAA1199 as a potential diagnostic biomarker of rheumatoid arthritis related to angiogenesis. *Arthritis Res Ther* **17**, 140, doi:10.1186/s13075-015-0637-y (2015).
89 Nagaoka, A. *et al.* Regulation of Hyaluronan (HA) Metabolism Mediated by HYBID (Hyaluronan-binding Protein Involved in HA Depolymerization, KIAA1199) and HA Synthases in Growth Factor-stimulated Fibroblasts. *J Biol Chem* **290**, 30910-30923, doi:10.1074/jbc.M115.673566 (2015).
90 Shiozawa, J. *et al.* Implication of HYBID (Hyaluronan-Binding Protein Involved in Hyaluronan Depolymerization) in Hyaluronan Degradation by Synovial Fibroblasts in Patients with Knee Osteoarthritis. *Am J Pathol* **190**, 1046-1058, doi:10.1016/j.ajpath.2020.01.003 (2020).
91 Shimoda, M. *et al.* Hyaluronan-Binding Protein Involved in Hyaluronan Depolymerization Controls Endochondral Ossification through Hyaluronan Metabolism. *Am J Pathol* **187**, 1162-1176, doi:10.1016/j.ajpath.2017.01.005 (2017).
92 Shinohara, T. *et al.* Hyaluronan metabolism in overloaded temporomandibular joint. *J Oral Rehabil* **43**, 921-928, doi:10.1111/joor.12443 (2016).
93 Shimizu, H. *et al.* Hyaluronan-Binding Protein Involved in Hyaluronan Depolymerization Is Up-Regulated and Involved in Hyaluronan Degradation in Human Osteoarthritic Cartilage. *Am J Pathol* **188**, 2109-2119, doi:10.1016/j.ajpath.2018.05.012 (2018).
94 Deroyer, C. *et al.* CEMIP (KIAA1199) induces a fibrosis-like process in osteoarthritic chondrocytes. *Cell Death Dis* **10**, 103, doi:10.1038/s41419-019-1377-8 (2019).
95 Sato, S. *et al.* Pro-inflammatory cytokines suppress HYBID (hyaluronan (HA) -binding protein involved in HA depolymerization/KIAA1199/CEMIP) -mediated HA metabolism in human skin fibroblasts. *Biochem Biophys Res Commun* **539**, 77-82, doi:10.1016/j.bbrc.2020.12.082 (2021).
96 Yoshida, H. *et al.* Reduction of hyaluronan and increased expression of HYBID (alias CEMIP and KIAA1199) correlate with clinical symptoms in photoaged skin. *Br J Dermatol* **179**, 136-144, doi:10.1111/bjd.16335 (2018).

97 Yoshida, H. *et al.* HYBID (alias KIAA1199/CEMIP) and hyaluronan synthase coordinately regulate hyaluronan metabolism in histamine-stimulated skin fibroblasts. *J Biol Chem* **295**, 2483-2494, doi:10.1074/jbc.RA119.010457 (2020).
98 Krishnamurthy, V. K. *et al.* Dysregulation of hyaluronan homeostasis during aortic valve disease. *Matrix Biol* **62**, 40-57, doi:10.1016/j.matbio.2016.11.003 (2017).
99 Gaggar, A. & Weathington, N. Bioactive extracellular matrix fragments in lung health and disease. *The Journal of clinical investigation* **126**, 3176-3184, doi:10.1172/jci83147 (2016).
100 Puperi, D. S. *et al.* Hyaluronan Hydrogels for a Biomimetic Spongiosa Layer of Tissue Engineered Heart Valve Scaffolds. *Biomacromolecules* **17**, 1766-1775, doi:10.1021/acs.biomac.6b00180 (2016).
101 Liu, J., Yan, W., Han, P. & Tian, D. The emerging role of KIAA1199 in cancer development and therapy. *Biomed Pharmacother* **138**, 111507, doi:10.1016/j.biopha.2021.111507 (2021).
102 Sabates-Bellver, J. *et al.* Transcriptome profile of human colorectal adenomas. *Mol Cancer Res* **5**, 1263-1275, doi:10.1158/1541-7786.MCR-07-0267 (2007).
103 LaPointe, L. C. *et al.* Discovery and validation of molecular biomarkers for colorectal adenomas and cancer with application to blood testing. *PLoS One* **7**, e29059, doi:10.1371/journal.pone.0029059 (2012).
104 Fink, S. P. *et al.* Induction of KIAA1199/CEMIP is associated with colon cancer phenotype and poor patient survival. *Oncotarget* **6**, 30500-30515, doi:10.18632/oncotarget.5921 (2015).
105 Xu, G. *et al.* CEMIP, acting as a scaffold protein for bridging GRAF1 and MIB1, promotes colorectal cancer metastasis via activating CDC42/MAPK pathway. *Cell Death Dis* **14**, 167, doi:10.1038/s41419-023-05644-z (2023).
106 Suh, H. N. *et al.* Identification of KIAA1199 as a Biomarker for Pancreatic Intraepithelial Neoplasia. *Sci Rep* **6**, 38273, doi:10.1038/srep38273 (2016).
107 Thiery, J. P. & Sleeman, J. P. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol* **7**, 131-142, doi:10.1038/nrm1835 (2006).
108 Santamaria, P. G., Moreno-Bueno, G., Portillo, F. & Cano, A. EMT: Present and future in clinical oncology. *Mol Oncol* **11**, 718-738, doi:10.1002/1878-0261.12091 (2017).
109 Tang, Z. *et al.* KIAA1199 promotes invasion and migration in non-small-cell lung cancer (NSCLC) via PI3K-Akt mediated EMT. *J Mol Med (Berl)* **97**, 127-140, doi:10.1007/s00109-018-1721-y (2019).
110 Jia, S. *et al.* KIAA1199 promotes migration and invasion by Wnt/beta-catenin pathway and MMPs mediated EMT progression and serves as a poor prognosis marker in gastric cancer. *PLoS One* **12**, e0175058, doi:10.1371/journal.pone.0175058 (2017).
111 Jiao, X. *et al.* KIAA1199, a Target of MicoRNA-486-5p, Promotes Papillary Thyroid Cancer Invasion by Influencing Epithelial-Mesenchymal Transition (EMT). *Med Sci Monit* **25**, 6788-6796, doi:10.12659/MSM.918682 (2019).
112 Jiang, Z., Zhai, X., Shi, B., Luo, D. & Jin, B. KIAA1199 overexpression is associated with abnormal expression of EMT markers and is a novel independent prognostic biomarker for hepatocellular carcinoma. *Onco Targets Ther* **11**, 8341-8348, doi:10.2147/OTT.S187389 (2018).

113 Evensen, N. A. *et al.* Hypoxia promotes colon cancer dissemination through up-regulation of cell migration-inducing protein (CEMIP). *Oncotarget* **6**, 20723-20739, doi:10.18632/oncotarget.3978 (2015).
114 Zhai, X. *et al.* Serum KIAA1199 is an advanced-stage prognostic biomarker and metastatic oncogene in cholangiocarcinoma. *Aging (Albany NY)* **12**, 23761-23777, doi:10.18632/aging.103964 (2020).
115 Kikuchi, M. *et al.* Epigenetic regulation of ZEB1-RAB25/ESRP1 axis plays a critical role in phenylbutyrate treatment-resistant breast cancer. *Oncotarget* **7**, 1741-1753, doi:10.18632/oncotarget.6480 (2016).
116 Xu, J., Lamouille, S. & Derynck, R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res* **19**, 156-172, doi:10.1038/cr.2009.5 (2009).
117 Zhang, L., Zhou, F. & ten Dijke, P. Signaling interplay between transforming growth factor-beta receptor and PI3K/AKT pathways in cancer. *Trends Biochem Sci* **38**, 612-620, doi:10.1016/j.tibs.2013.10.001 (2013).
118 Sabbah, D. A., Hajjo, R. & Sweidan, K. Review on Epidermal Growth Factor Receptor (EGFR) Structure, Signaling Pathways, Interactions, and Recent Updates of EGFR Inhibitors. *Curr Top Med Chem* **20**, 815-834, doi:10.2174/1568026620666200303123102 (2020).
119 Wang, A. *et al.* Downregulation of KIAA1199 by miR-486-5p suppresses tumorigenesis in lung cancer. *Cancer Med* **9**, 5570-5586, doi:10.1002/cam4.3210 (2020).
120 Xu, Y. *et al.* KIAA1199 promotes sorafenib tolerance and the metastasis of hepatocellular carcinoma by activating the EGF/EGFR-dependent epithelial-mesenchymal transition program. *Cancer Lett* **454**, 78-89, doi:10.1016/j.canlet.2019.03.049 (2019).
121 Basu, S., Cheriyamundath, S. & Ben-Ze'ev, A. Cell-cell adhesion: linking Wnt/beta-catenin signaling with partial EMT and stemness traits in tumorigenesis. *F1000Res* **7**, doi:10.12688/f1000research.15782.1 (2018).
122 Teeuwssen, M. & Fodde, R. Wnt Signaling in Ovarian Cancer Stemness, EMT, and Therapy Resistance. *J Clin Med* **8**, doi:10.3390/jcm8101658 (2019).
123 Chanmee, T., Ontong, P. & Itano, N. Hyaluronan: A modulator of the tumor microenvironment. *Cancer Lett* **375**, 20-30, doi:10.1016/j.canlet.2016.02.031 (2016).
124 Koh, Y., Kim, D., Jung, W. J., Ahn, K. S. & Yoon, S. S. Revealing Genomic Profile That Underlies Tropism of Myeloma Cells Using Whole Exome Sequencing. *Int J Genomics* **2015**, 675379, doi:10.1155/2015/675379 (2015).
125 Rodrigues, G. *et al.* Tumour exosomal CEMIP protein promotes cancer cell colonization in brain metastasis. *Nat Cell Biol* **21**, 1403-1412, doi:10.1038/s41556-019-0404-4 (2019).
126 Tsuji, S. *et al.* HYBID derived from tumor cells and tumor-associated macrophages contribute to the glioblastoma growth. *Brain Res* **1764**, 147490, doi:10.1016/j.brainres.2021.147490 (2021).
127 Kohi, S., Sato, N., Koga, A., Matayoshi, N. & Hirata, K. KIAA1199 is induced by inflammation and enhances malignant phenotype in pancreatic cancer. *Oncotarget* **8**, 17156-17163, doi:10.18632/oncotarget.15052 (2017).
128 Schmaus, A. *et al.* Sulfated hyaluronic acid inhibits the hyaluronidase CEMIP and regulates the HA metabolism, proliferation and differentiation of fibroblasts. *Matrix Biol* **109**, 173-191, doi:10.1016/j.matbio.2022.04.001 (2022).
129 Yoshino, Y., Goto, M., Hara, H. & Inoue, S. The role and regulation of TMEM2 (transmembrane protein 2) in HYBID (hyaluronan (HA)-binding protein involved in HA

depolymerization/ KIAA1199/CEMIP)-mediated HA depolymerization in human skin fibroblasts. *Biochem Biophys Res Commun* **505**, 74-80, doi:10.1016/j.bbrc.2018.09.097 (2018).
130 Kuscu, C. *et al.* Transcriptional and epigenetic regulation of KIAA1199 gene expression in human breast cancer. *PLoS One* **7**, e44661, doi:10.1371/journal.pone.0044661 (2012).
131 Wang, L. *et al.* The miR-29c-KIAA1199 axis regulates gastric cancer migration by binding with WBP11 and PTP4A3. *Oncogene* **38**, 3134-3150, doi:10.1038/s41388-018-0642-0 (2019).
132 Zhang, D. *et al.* Down-regulation of KIAA1199/CEMIP by miR-216a suppresses tumor invasion and metastasis in colorectal cancer. *Int J Cancer* **140**, 2298-2309, doi:10.1002/ijc.30656 (2017).
133 Zhou, Y. & Mu, T. LncRNA LINC00958 promotes tumor progression through miR-4306/CEMIP axis in osteosarcoma. *Eur Rev Med Pharmacol Sci* **25**, 3182-3199, doi:10.26355/eurrev_202104_25727 (2021).
134 Hsieh, I. Y. *et al.* H3K27me3 loss plays a vital role in CEMIP mediated carcinogenesis and progression of breast cancer with poor prognosis. *Biomed Pharmacother* **123**, 109728, doi:10.1016/j.biopha.2019.109728 (2020).
135 Michishita, E., Garces, G., Barrett, J. C. & Horikawa, I. Upregulation of the KIAA1199 gene is associated with cellular mortality. *Cancer Lett* **239**, 71-77, doi:10.1016/j.canlet.2005.07.028 (2006).
136 Csoka, A. B., Scherer, S. W. & Stern, R. Expression analysis of six paralogous human hyaluronidase genes clustered on chromosomes 3p21 and 7q31. *Genomics* **60**, 356-361, doi:10.1006/geno.1999.5876 (1999).
137 Miyata, S. & Kitagawa, H. Formation and remodeling of the brain extracellular matrix in neural plasticity: Roles of chondroitin sulfate and hyaluronan. *Biochim Biophys Acta Gen Subj* **1861**, 2420-2434, doi:10.1016/j.bbagen.2017.06.010 (2017).
138 Rauch, U. Extracellular matrix components associated with remodeling processes in brain. *Cell Mol Life Sci* **61**, 2031-2045, doi:10.1007/s00018-004-4043-x (2004).
139 Wilson, E., Knudson, W. & Newell-Litwa, K. Hyaluronan regulates synapse formation and function in developing neural networks. *Sci Rep* **10**, 16459, doi:10.1038/s41598-020-73177-y (2020).
140 Yoshino, Y. *et al.* Distribution and function of hyaluronan binding protein involved in hyaluronan depolymerization (HYBID, KIAA1199) in the mouse central nervous system. *Neuroscience* **347**, 1-10, doi:10.1016/j.neuroscience.2017.01.049 (2017).
141 Yoshino, Y. *et al.* Targeted deletion of HYBID (hyaluronan binding protein involved in hyaluronan depolymerization/ KIAA1199/CEMIP) decreases dendritic spine density in the dentate gyrus through hyaluronan accumulation. *Biochem Biophys Res Commun* **503**, 1934-1940, doi:10.1016/j.bbrc.2018.07.138 (2018).
142 Matsuzaki, S. *et al.* Clinicopathologic significance of KIAA1199 overexpression in human gastric cancer. *Ann Surg Oncol* **16**, 2042-2051, doi:10.1245/s10434-009-0469-6 (2009).
143 Liu, J. *et al.* Knockdown of KIAA1199 attenuates growth and metastasis of hepatocellular carcinoma. *Cell Death Discov* **4**, 102, doi:10.1038/s41420-018-0099-5 (2018).
144 Tiwari, A. *et al.* Early insights into the function of KIAA1199, a markedly overexpressed protein in human colorectal tumors. *PLoS One* **8**, e69473, doi:10.1371/journal.pone.0069473 (2013).

145 Evensen, N. A. *et al.* Unraveling the role of KIAA1199, a novel endoplasmic reticulum protein, in cancer cell migration. *J Natl Cancer Inst* **105**, 1402-1416, doi:10.1093/jnci/djt224 (2013).
146 Galarza, T. E., Taquez Delgado, M. A., Mohamad, N. A., Martin, G. A. & Cricco, G. P. Histamine H4 receptor agonists induce epithelial-mesenchymal transition events and enhance mammosphere formation via Src and TGF-beta signaling in breast cancer cells. *Biochem Pharmacol* **180**, 114177, doi:10.1016/j.bcp.2020.114177 (2020).
147 Lee, T. H., Wisniewski, H. G. & Vilcek, J. A novel secretory tumor necrosis factor-inducible protein (TSG-6) is a member of the family of hyaluronate binding proteins, closely related to the adhesion receptor CD44. *The Journal of cell biology* **116**, 545-557, doi:10.1083/jcb.116.2.545 (1992).
148 Sleeman, J. P., Kondo, K., Moll, J., Ponta, H. & Herrlich, P. Variant exons v6 and v7 together expand the repertoire of glycosaminoglycans bound by CD44. *J Biol Chem* **272**, 31837-31844, doi:10.1074/jbc.272.50.31837 (1997).
149 Matsuda, M. *et al.* Therapeutic effect of sulphated hyaluronic acid, a potential selectin-blocking agent, on experimental progressive mesangial proliferative glomerulonephritis. *The Journal of pathology* **198**, 407-414, doi:10.1002/path.1209 (2002).
150 Jordan, A. R. *et al.* Antitumor activity of sulfated hyaluronic acid fragments in pre-clinical models of bladder cancer. *Oncotarget* **8**, 24262-24274, doi:10.18632/oncotarget.10529 (2017).
151 Pandey, M. S., Harris, E. N., Weigel, J. A. & Weigel, P. H. The cytoplasmic domain of the hyaluronan receptor for endocytosis (HARE) contains multiple endocytic motifs targeting coated pit-mediated internalization. *J Biol Chem* **283**, 21453-21461, doi:10.1074/jbc.M800886200 (2008).
152 Knudson, W., Chow, G. & Knudson, C. B. CD44-mediated uptake and degradation of hyaluronan. *Matrix Biol* **21**, 15-23, doi:10.1016/s0945-053x(01)00186-x (2002).
153 Tammi, R. *et al.* Hyaluronan enters keratinocytes by a novel endocytic route for catabolism. *J Biol Chem* **276**, 35111-35122, doi:10.1074/jbc.M103481200 (2001).
154 Shu, N., Zhou, T. & Hovmöller, S. Prediction of zinc-binding sites in proteins from sequence. *Bioinformatics* **24**, 775-782, doi:10.1093/bioinformatics/btm618 (2008).
155 Letunic, I. & Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res* **46**, D493-d496, doi:10.1093/nar/gkx922 (2018).
156 Zhao, L. *et al.* KIAA1199 promotes metastasis of colorectal cancer cells via microtubule destabilization regulated by a PP2A/stathmin pathway. *Oncogene* **38**, 935-949, doi:10.1038/s41388-018-0493-8 (2019).
157 Yoshida, H. *et al.* Inhibition of HYBID (KIAA1199)-mediated hyaluronan degradation and anti-wrinkle effect of Geranium thunbergii extract. *Journal of Cosmetic Dermatology* **18**, 1052-1060, doi:10.1111/jocd.12813 (2019).
158 Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244-248, doi:10.1093/nar/gki408 (2005).
159 Hildebrand, A., Remmert, M., Biegert, A. & Söding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins* **77 Suppl 9**, 128-132, doi:10.1002/prot.22499 (2009).

160 Ovchinnikov, S., Kamisetty, H. & Baker, D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* **3**, e02030, doi:10.7554/eLife.02030 (2014).
161 Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **31**, 455-461, doi:10.1002/jcc.21334 (2010).
162 Hryc, C. F. & Baker, M. L. AlphaFold2 and CryoEM: Revisiting CryoEM modeling in near-atomic resolution density maps. *iScience* **25**, 104496, doi:10.1016/j.isci.2022.104496 (2022).

12 ANNEXES

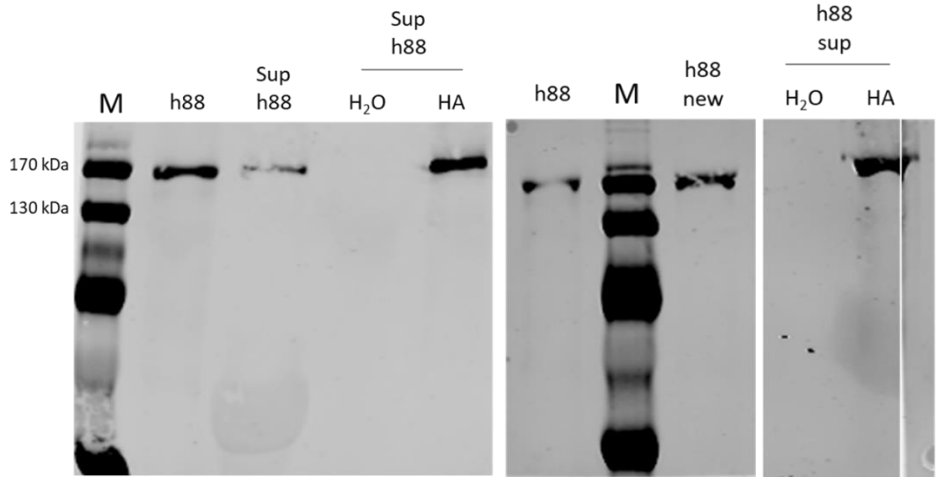


***Figure 116**: Pulldown experiment of sCEMIP with HA 1500 kDa.*

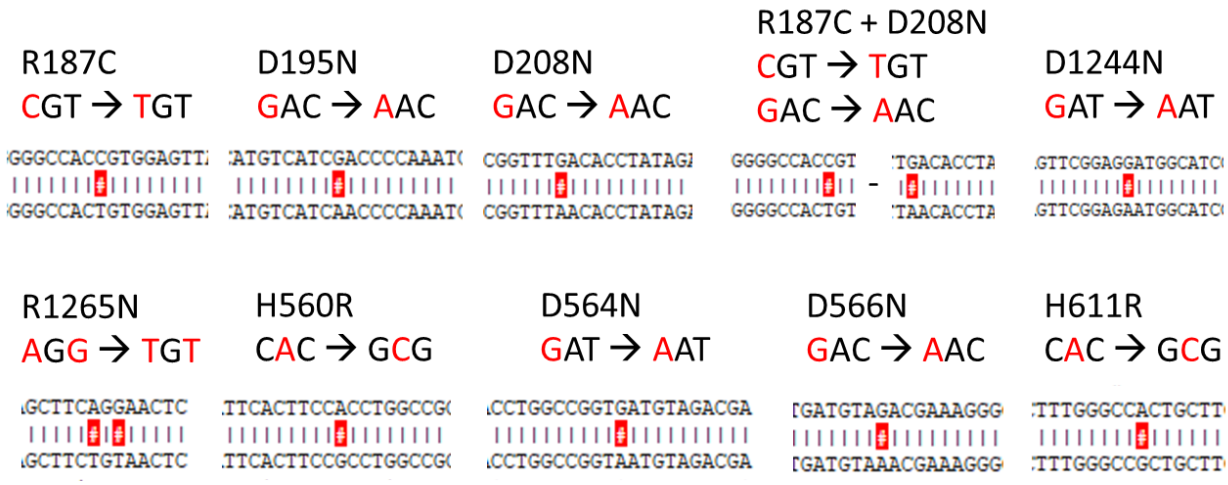


***Figure 117**: Sequencing results of mutants R187C, D195N, D208N, R187C+D208N, D1244N, R1265N, H560R, D564N, D566N and H611R.*

ΔGG2

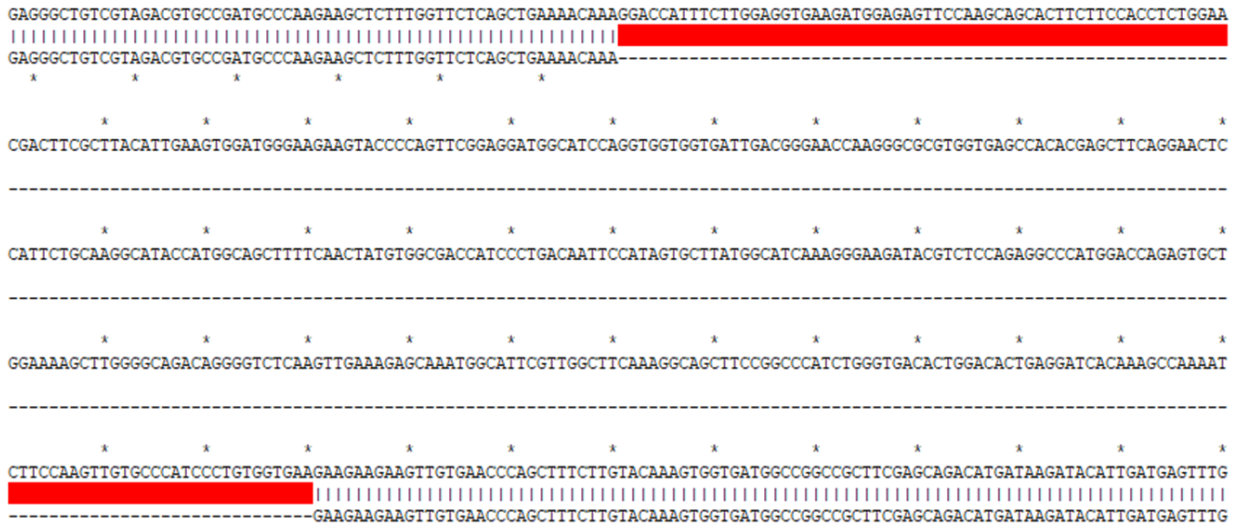


***Figure 118**: Sequencing result of mutant ΔGG2.*

ΔG8

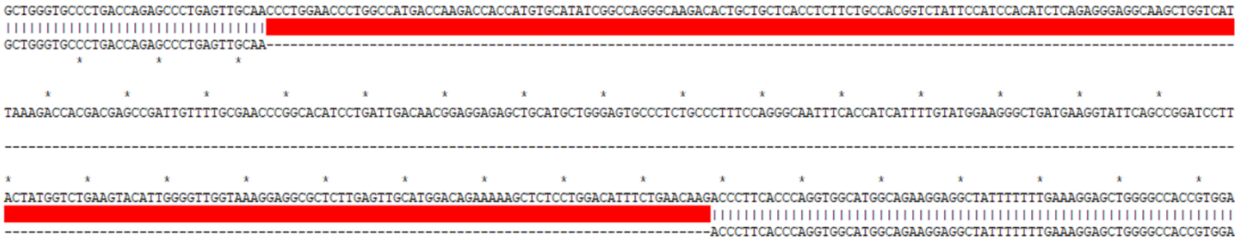


***Figure 119**: Sequencing result of mutant ΔG8.*

ΔPbH1

***Figure 120**: Sequencing result of mutant ΔPbH1.*

ΔPbH1-4

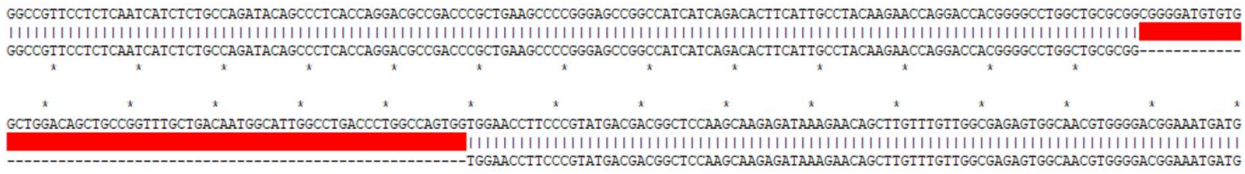


Figure 121*: Sequencing result of mutant ΔPbH1-4.*

ΔPbH1-3-4

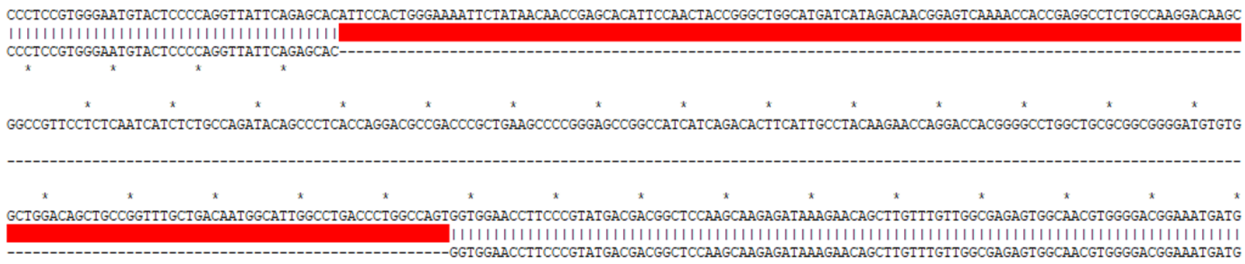
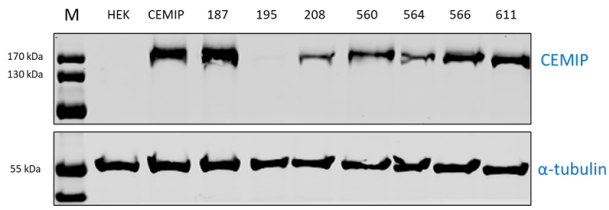
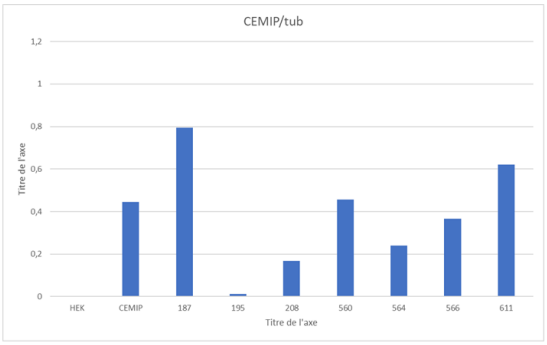


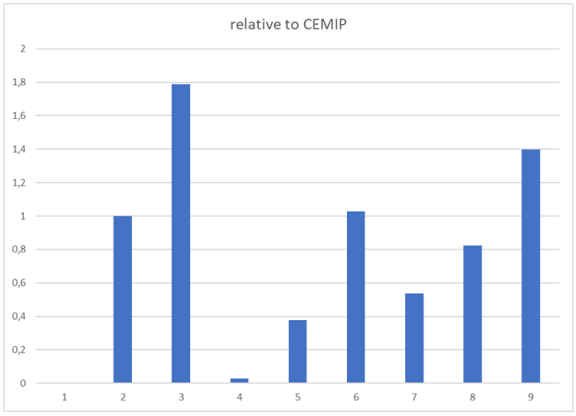
Figure 122*: Sequencing result of mutant ΔPbH1-3-4.*

ΔPbH1-2-3-4

Figure 123*: Sequencing result of mutant ΔPbH1-2-3-4.*



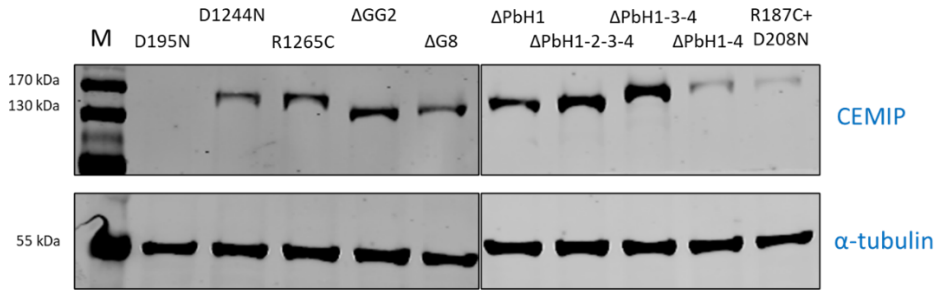


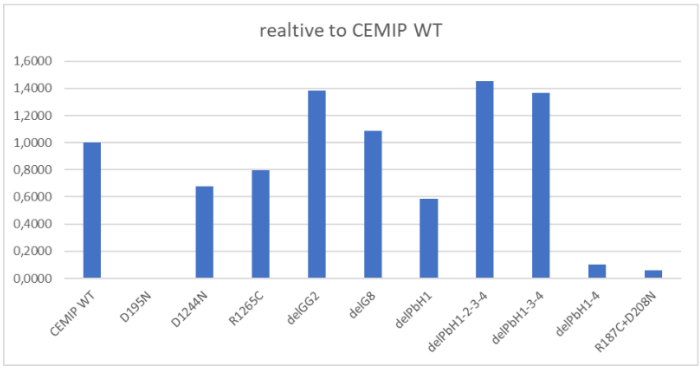


Working concentrations

	uL of lysate	ul water	conc
CEMIP	14,5	120,5	0,50
187	8,1	126,9	0,29
208	38,3	96,7	1,74
560	14,1	120,9	0,62
564	26,9	108,1	1,42
566	17,6	117,4	0,79
611	10,3	124,7	0,50

***Figure 124**: Quantification and normalization of mutants R187C, D208N, H560R, D564N, D566N and H611R.*





Working concentrations

	uL of lysate	ul water	conc
CEMIP WT	14,5	120,5	0,50
~~D195N~~	~~0,0~~	~~135,0~~	~~0,00~~
D1244N	21,4	113,6	0,71
R1265C	18,1	116,9	0,64
delGG2	10,5	124,5	0,37
delG8	13,3	121,7	0,52
delPbH1	24,6	110,4	0,85
delPbH1-2-3-4	9,9	125,1	0,33
delPbH1-3-4	10,6	124,4	0,34
~~delPbH1-4~~	~~140,2~~	~~-5,2~~	~~4,76~~
~~R187C+D208N~~	~~239,0~~	~~-104,0~~	~~8,82~~

***Figure 125**: Quantification and normalization of mutants D1244N, R1265C, ΔGG2, ΔG8, ΔPbH1, ΔPbH1-2-3-4, ΔPbH1-3-4, ΔPbH1-4 and R187C + D208N.*

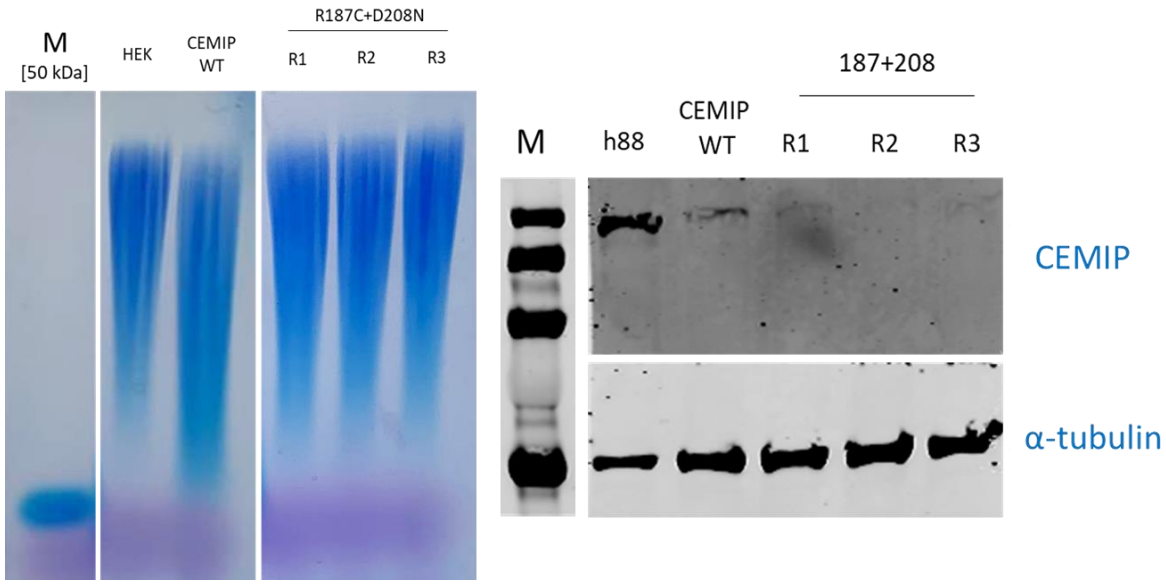


***Figure 126**: HA assay and Western blot of double mutant R187C + D208N (replicates).*

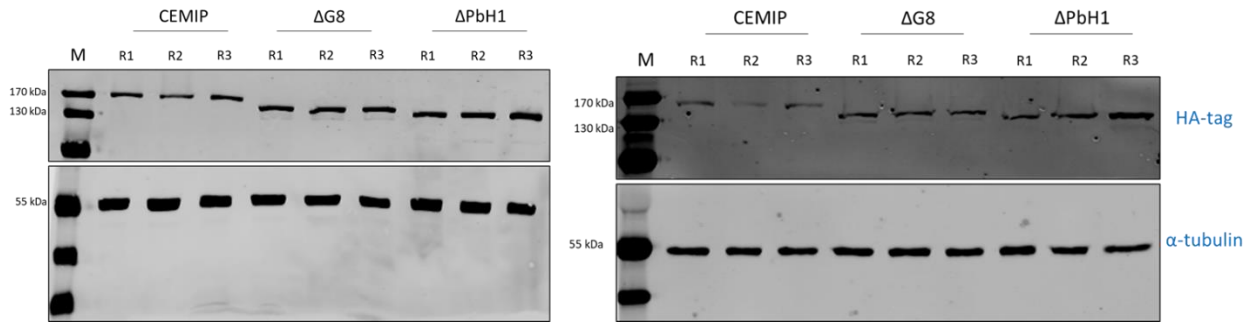


***Figure 127**: HA-tag antibody is less effective than CEMIP antibody. Left: WB of mutants containing HA-tag revealed with anti-CEMIP; right: same samples revealed with anti-HA-tag antibody.*

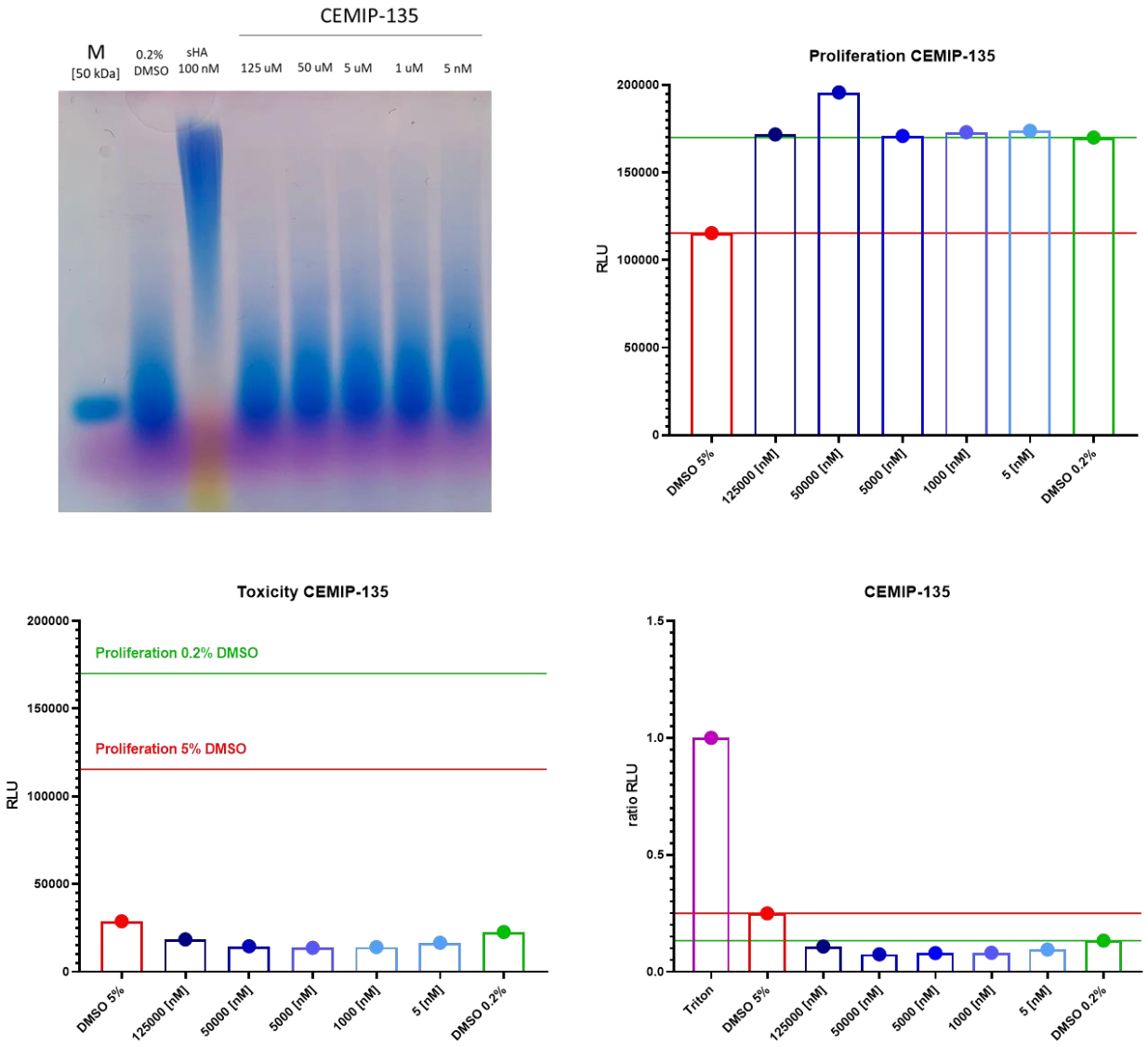


***Figure 128**: HA assay, cell proliferation, toxicity and relative toxicity of molecule CEMIP-135.*

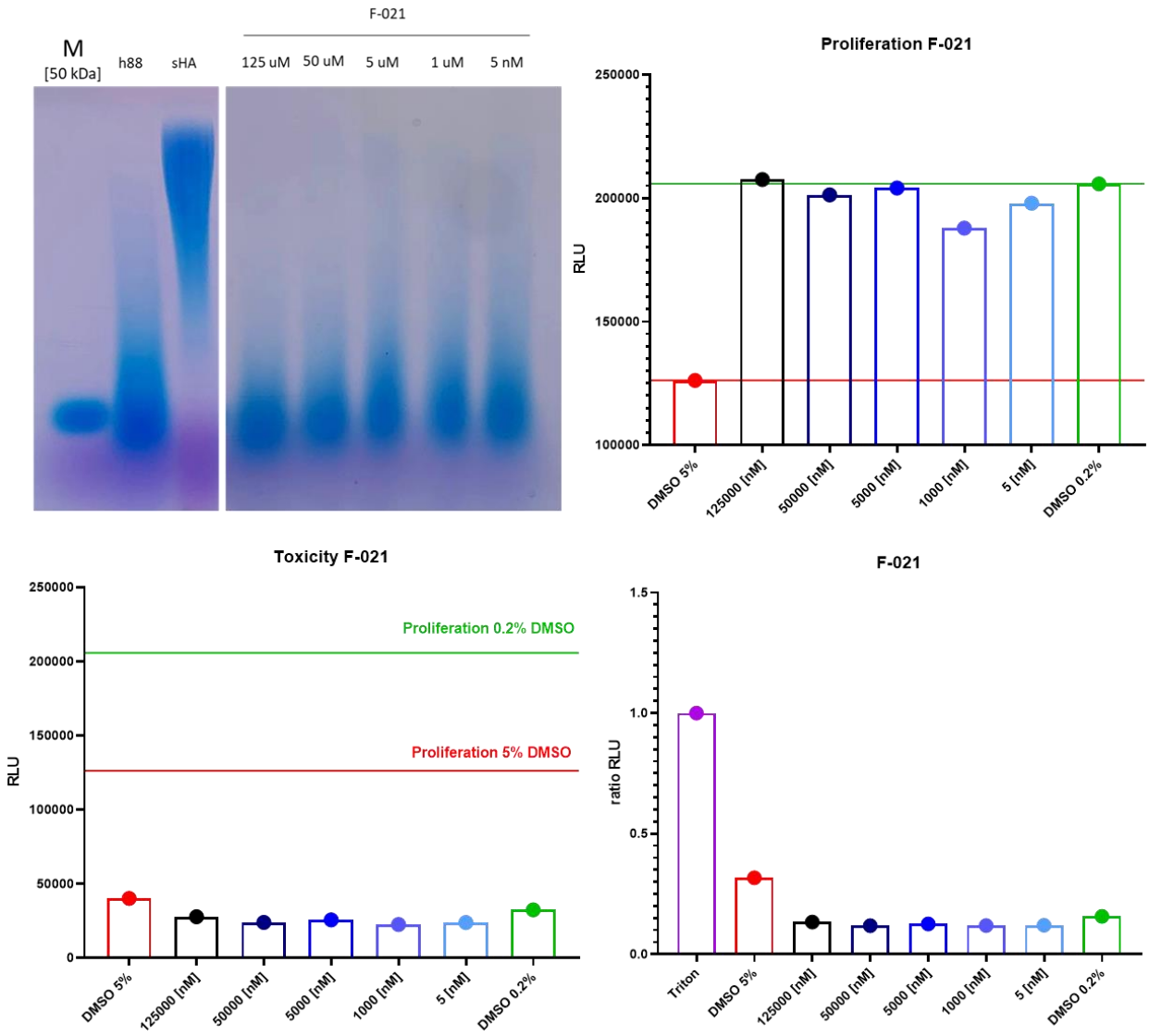


***Figure 129**: HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-021.*

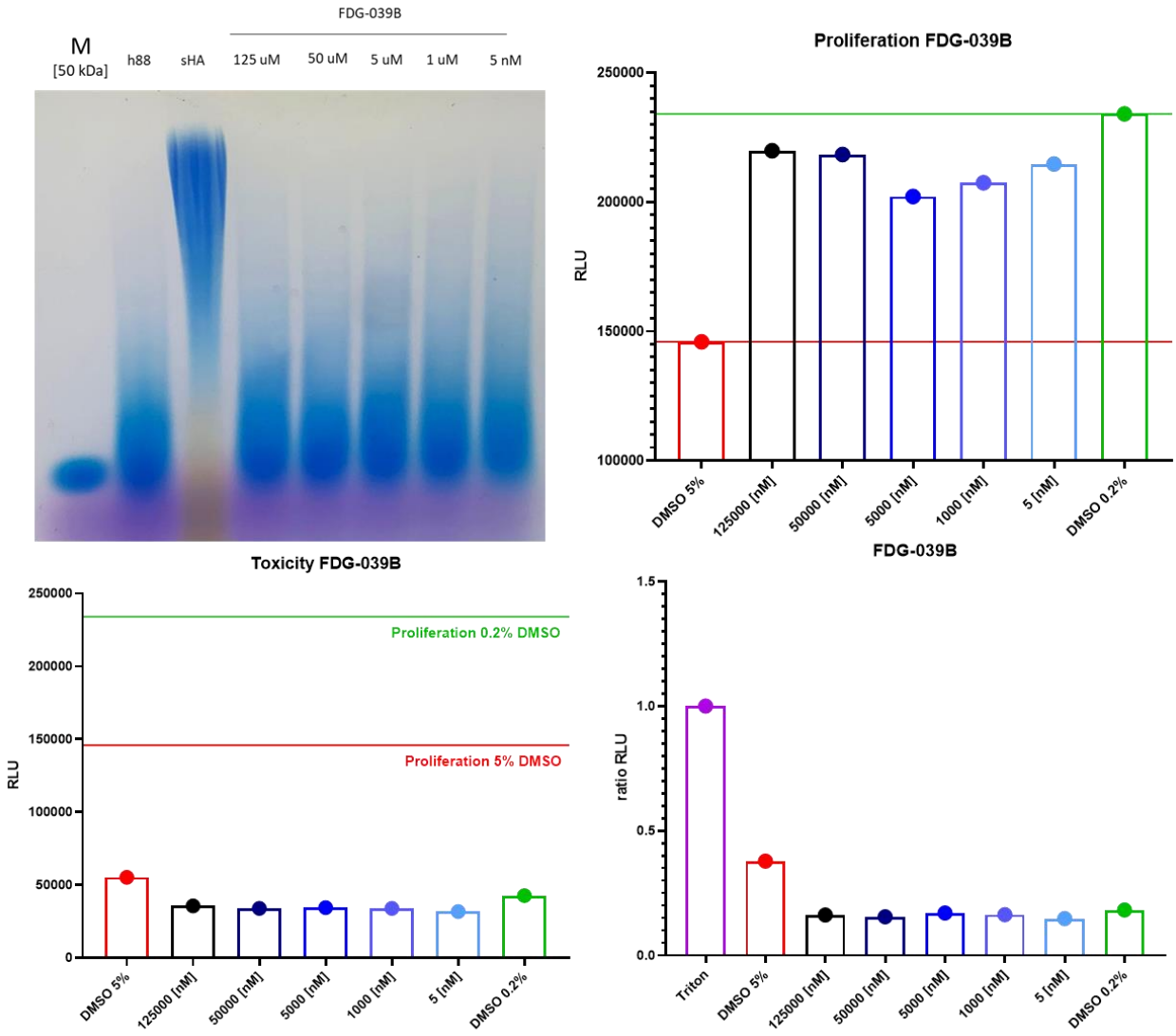


***Figure 130**: HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-039B.*

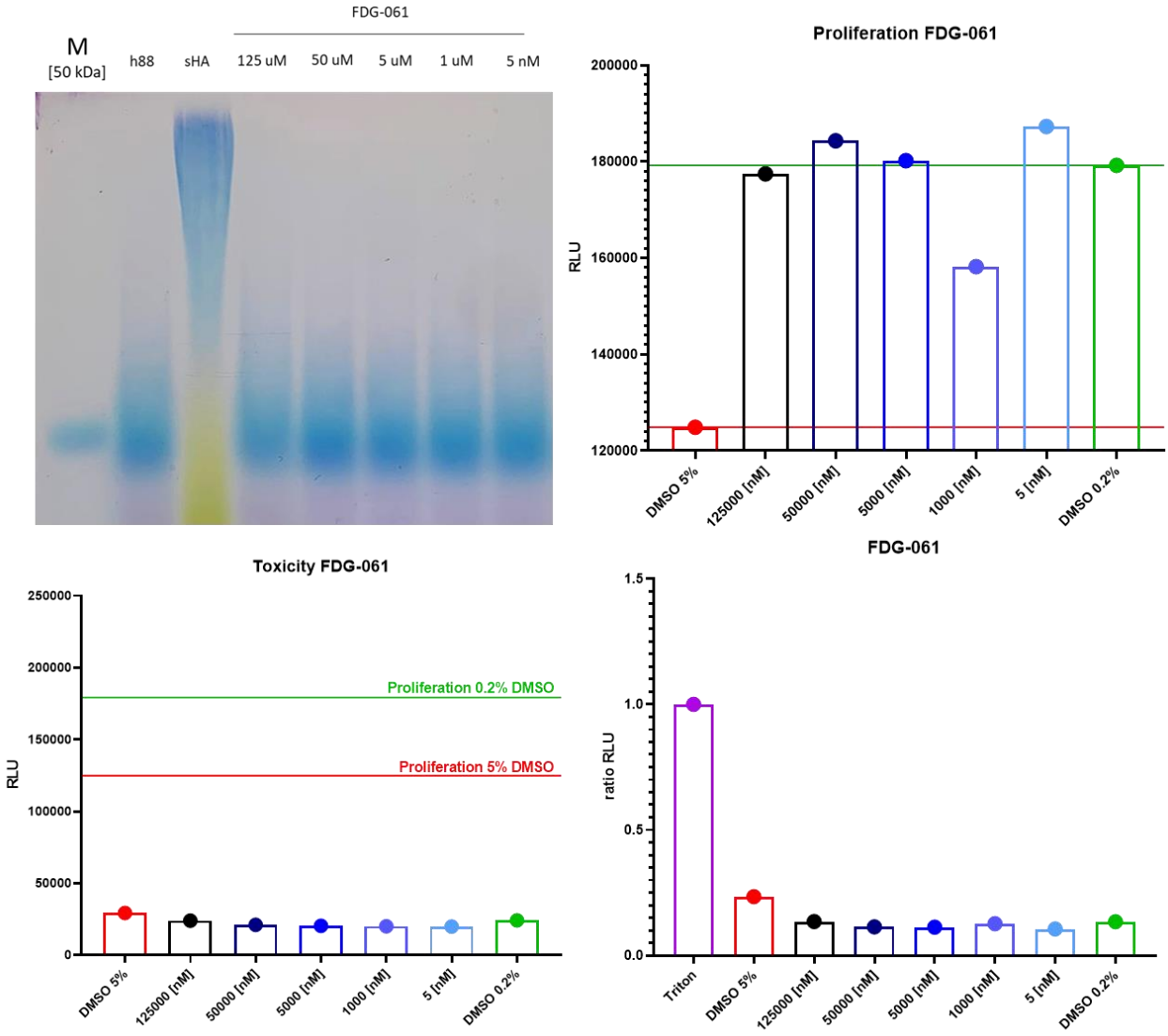


***Figure 131**: HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-061.*

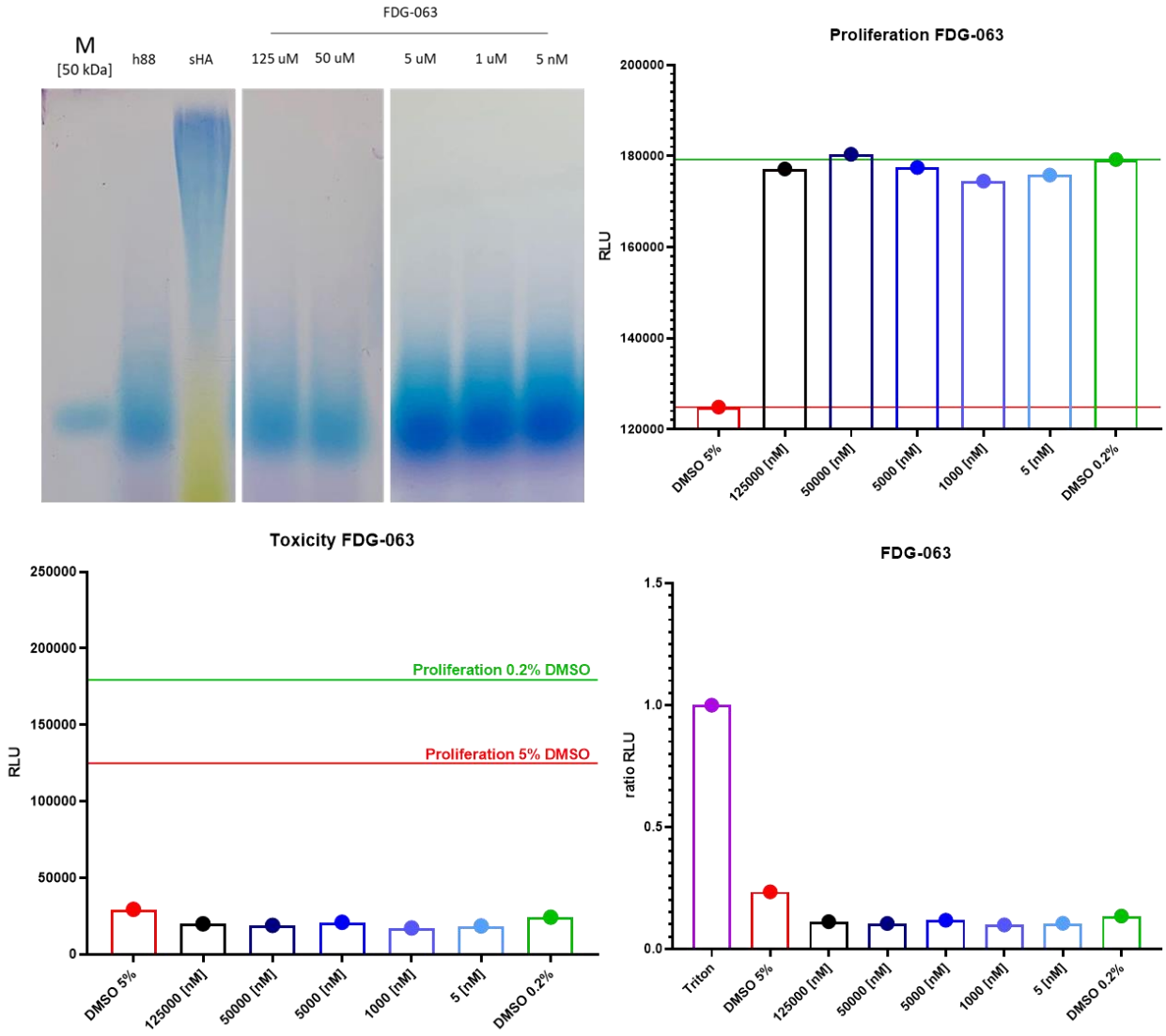


Figure 132: *HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-063.*

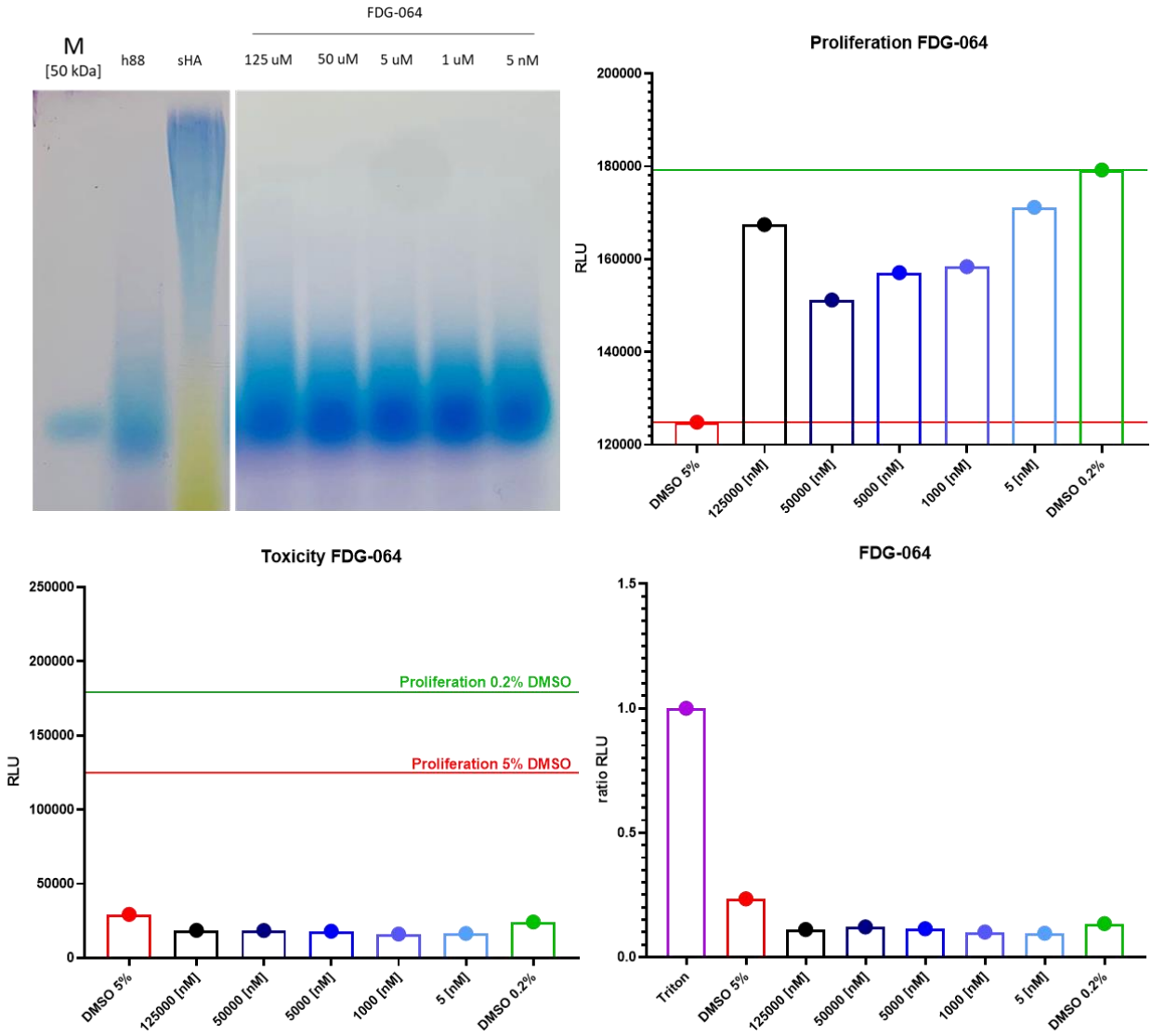


***Figure 133**: HA assay, cell proliferation, toxicity and relative toxicity of molecule FDG-064.*

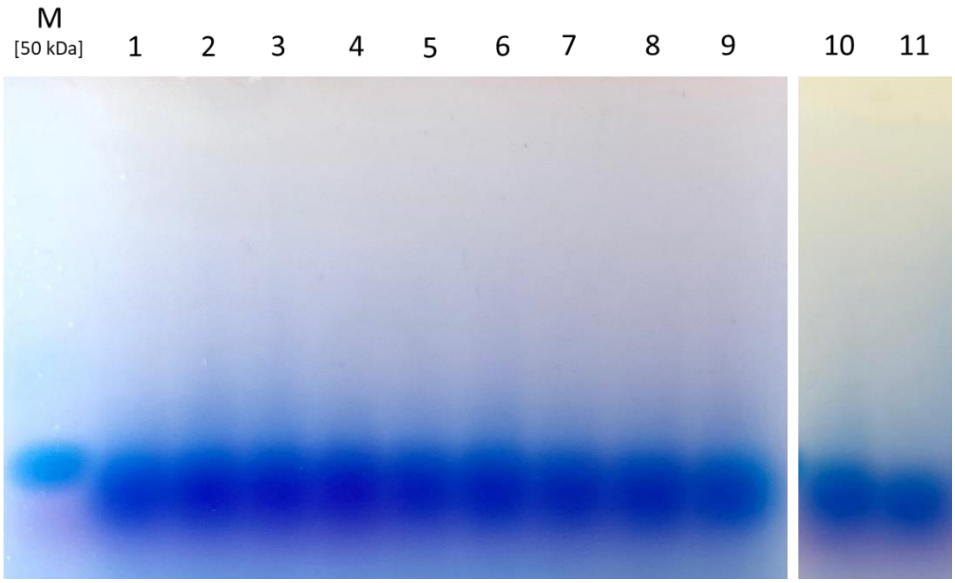


***Figure 134**: HA assay of selected antibodies. 1 = 3F1-1, 2 = 7H5-1, 3 = 8B5-1, 4 = 12B5-1, 5 = 12G5-1, 6 = 3F1-2, 7 = 7H5-2, 8 = 8B5-2, 9 = 12B5-2, 10 = 12G5-2, 11 = negative control (another antibody). 10% of antibody on hCEMIP.*

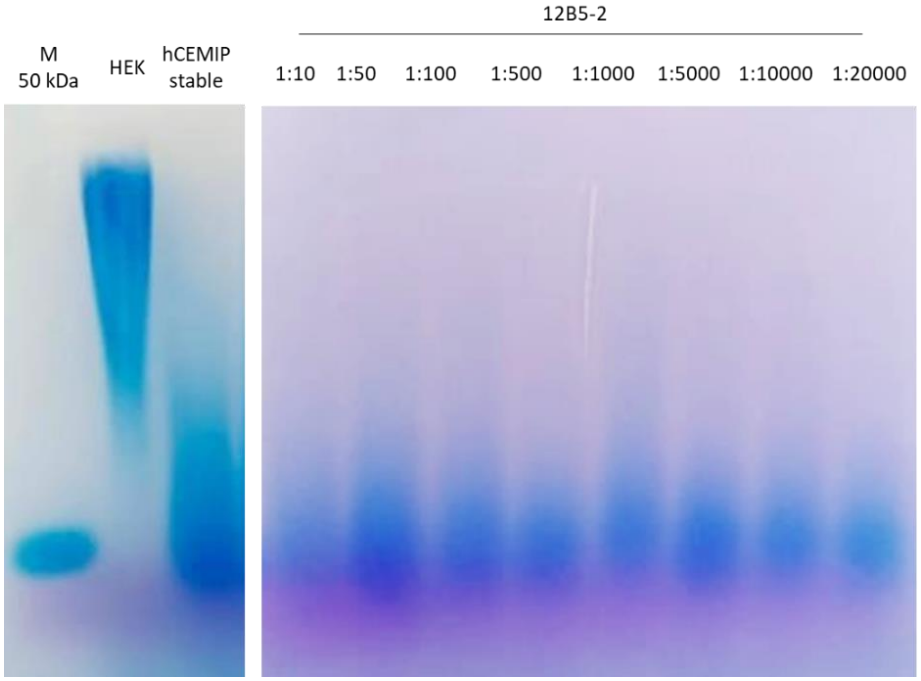


***Figure 135**: HA assay of antibody 12B5-2 on hCEMIP cells at different concentrations.*

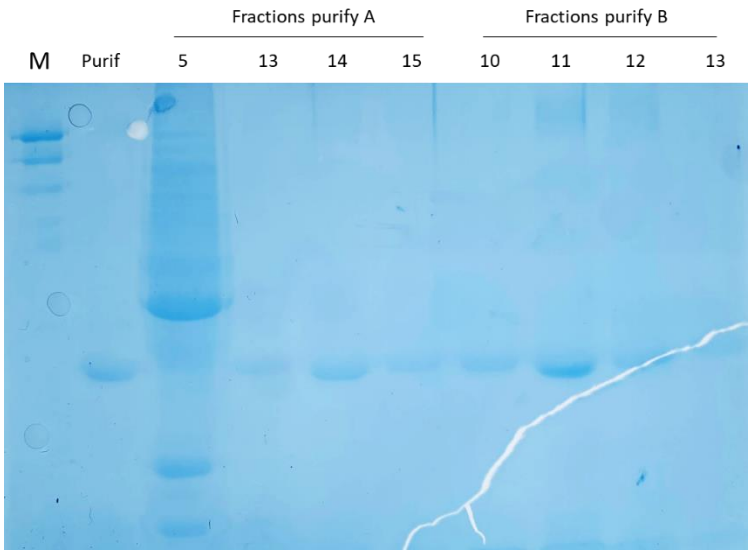


***Figure 136**: Fractions of two scaled-up purifications of 12B5-2.*

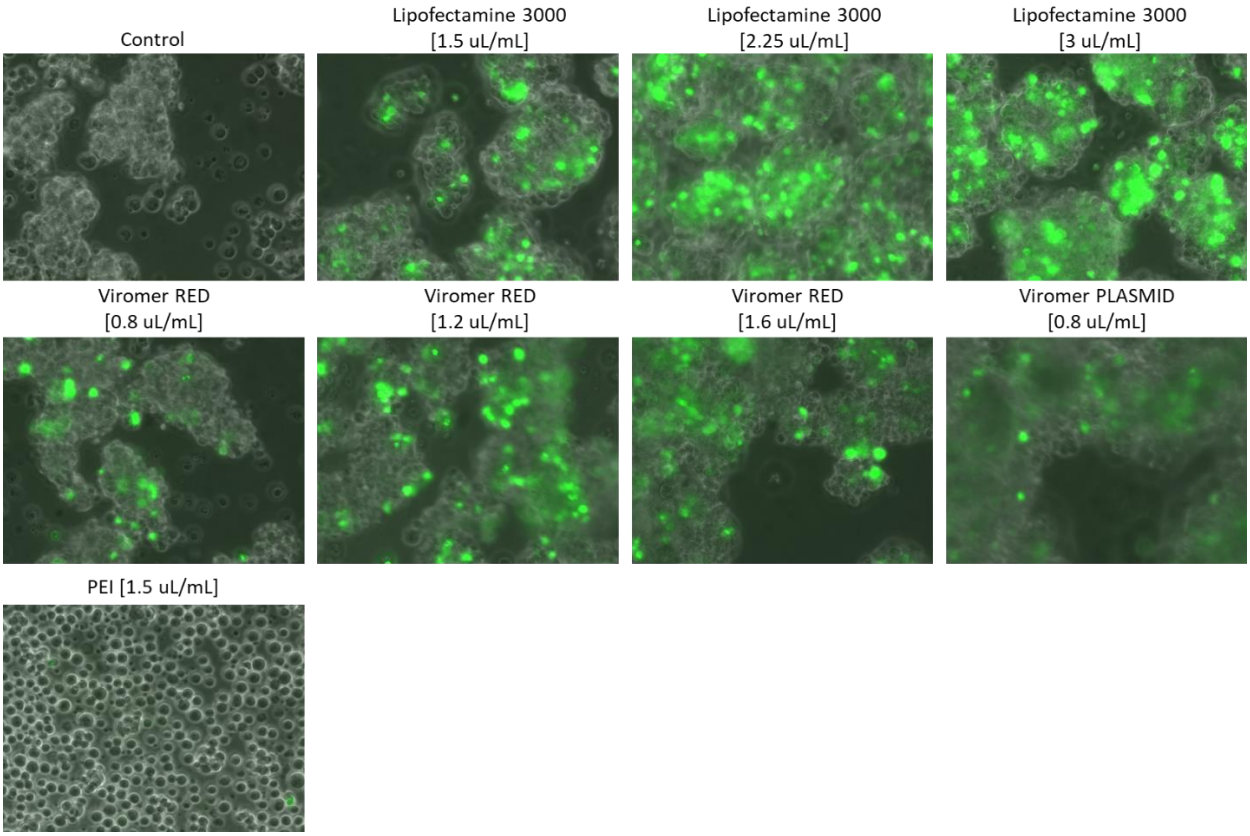


***Figure 137**: Transfection tests in suspension HEK293SF-3F6 with different reagents. Magnification 200x.*

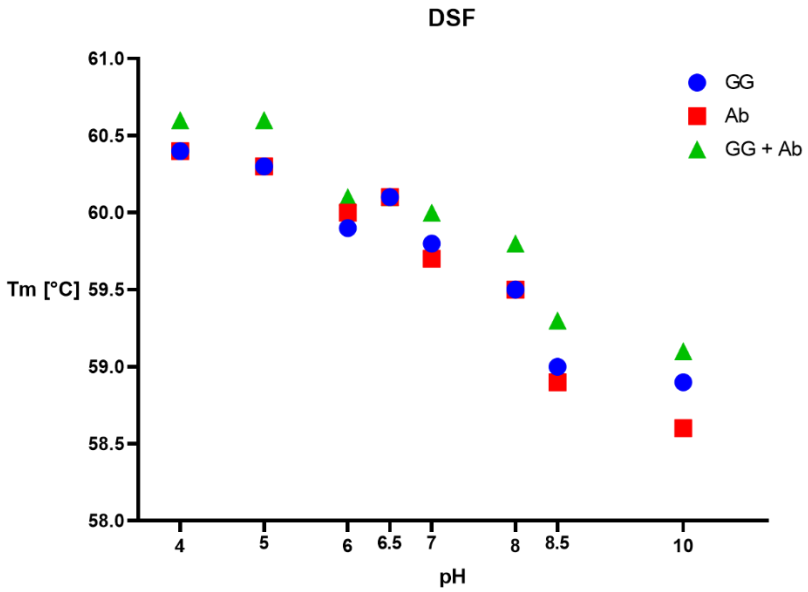


***Figure 138**: DSF of the GG domain and 12B5-2 antibody – Low pH stabilizes the GG/antibody complex.*

13 ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my esteemed thesis director and co-director: Prof. Leonardo Scapozza and Dr. Marco Prunotto. Thank you for the incredible opportunity to pursue my doctoral studies with this amazing project that got me even more passionate about biomedical science. You have been incredible mentors. Leonardo, I am deeply thankful for your invaluable advice, both scientific and personal. Marco, I appreciate your thoughtful and constructive feedback, which has significantly contributed to the refinement of my work. Your mentorship has been a source of inspiration, and I am truly grateful for your unwavering support.

To the jury members, Prof. Yogeshvar Kalia, Prof. Solange Moll, Prof. Jeffrey Miner, Prof. Jonathan Sleeman and Prof. Rafael Fridman, thank you for your time and for accepting this important role. Your feedback on my work is extremely valuable.

Thanks to all the people who collaborated on this project. Lina, your exceptional contributions to the structural biology aspect of this project have been truly commendable. Collaborating with you has not only been a pleasure but has also been an enlightening experience, allowing me to delve into and expand my understanding of an entirely new field. Special thanks to Prof. Andreas Boland and his group, Dr. Yashar Sadian and all the DCI CryoEM facility for the help and support. I would like to thank my collaborators in IRB, Dr. Andrea Cavalli, Dr. Jacopo Sgrignani and Dr. Concetta Guerra, for the constant support and rich discussion during our biweekly meetings. A heartfelt thank you to my collaborators in Germany. Jonathan, your kindness and support have truly made a difference in my PhD journey. Anja, thank you for the initial technical support and for sharing your expertise, as well as for your active participation in the project.

Thank you to all the FABIP group members and all the people in CMU for all the support during these years and for creating an exceptional working environment. Special thanks to Aurélie and Piter, for your patience and guidance. We would be lost without you! Margaux, thank you for your input on the modeling part. Séb, thank you for your input on the chemistry part. Thank you, Filippo, for being an excellent Master student, and eventually a colleague and a friend in the lab. Your hard work during your Master's thesis is a vital component of this project. Thank you to my fellow PhD travelers of the lab: Sherihan, Firdaws, Ghali, and Ece. It was a pleasure to share part of our

wonderful doctoral journey with you. Thank you, Luce, Angelica, Ozlem, Kenza and Bart for all the fantastic moments spent together inside and outside the lab. I would like to also thank Massimo Caine for giving me the possibility to work as a science communicator and senior scientific editor in the journal TheScienceBreaker.

My journey into the academic scientific world wouldn't be the same without the exceptional passion that Dr. Bernard Schneider, Dr. Paola Solanes, Dr. Valérie Vilmont and Dr. Isabel Lopez-Mejia shared with me during my master thesis and pre-master project. I thank you for taking the time to teach me and guide me as I was taking my first steps into research, as well as for the immense opportunities to learn and contribute to the field of gene editing for hearing loss, metabolism of Parkinson's disease, and cancer.

Un grazie di cuore a Livia, Giorgia e Dalila per essere state la migliore *study gang* durante i miei anni a Losanna. Le nostre sessioni di studio, le nostre cene e i nostri viaggi rimarranno per sempre tra i miei migliori ricordi. In particolare vorrei ringraziare Livia, la cui amicizia ci unisce da ben un quarto di secolo. Grazie per tutte le avvenure vissute al tuo fianco, sei come una sorella per me. Un grazie anche alle mie amiche del liceo, Francesca e Julia, con cui ho condiviso anni fantastici, e a Mira, Sofia P., Sophie e Céline.

During my time in CMU, I had the chance to find amazing friends. I know that this is a rare event in life, and I don't take it for granted. First, I would like to thank the members of the *Magic5 (THE conclave)*: Carlotta, Sara, Rebecca and Ilaria. Siete una perla rara. Non pensavo di poter trovare delle amiche vere in questo stadio (avanzato ahahah) della mia vita. E invece la vita a volte ti sorprende, e ti offre regali inaspettati. Carli, grazie per avermi accolta calorosamente nel laboratorio, la tua tenacia e la tua passione mi motivano costantemente a crescere e a migliorarmi. Sari, grazie per i tuoi preziosi consigli durante i momenti difficili, la tua forza e intraprendenza mi ispirano a puntare sempre in alto. Ilariona Cocchiarona, grazie per tutte le risate che ci siamo fatte, la tua positività e energia mi fanno sempre sentire a casa. Rebi, grazie per tutti i momenti in cui hai confidato in me condividendo gioie (e dolori) della tua vita, la tua dolcezza e la tua sensibilità mi ricordano che al mondo persone pure esitono ancora. Grazie a tutte voi per i momenti stupendi passati insieme, a per tutti quelli futuri! Second, I would like to thank the *Gilmore FABIP*, Radhia and Solène, with whom I shared not only the office, but also many many movie nights together.

Radhia, tu es une collègue de laboratoire incroyable. Merci d'avoir toujours été là. Je pense que tu es la personne qui connaît le mieux les hauts et les bas de mon projet, et les défis quotidiens de ma thèse, tout comme moi je connais les tiens (side eye). Solène, tu es arrivée comme une tornade au labo. Merci pour avoir mis une ambiance de ouf au bureau, pour ton énergie et pour tout tes conseils précieux. Les filles, les moments que nous avons partagés m'ont fait sentir vraiment à la maison. And remember, "*on Wednesdays, we wear pink*".

Infinitamente grazie ai miei genitori, Raoul e Nicoletta, per avermi sempre sostenuta con tanto amore. Il vostro supporto incondizionato è stato il pilastro su cui ho potuto costruire il mio percorso accademico, fino alla cima della montagna con questo dottorato. Questa tesi è il risultato di un viaggio lungo e impegnativo, e sono grata di poterlo condividere con voi. Mumi, sei stata la mia prima insegnante e la mia pirma guida. Grazie per esserci sempre per me in qualunque circostanza. Pupi, mi hai trasmesso la passione per la scienza fin da piccola. Grazie per tutti i sacrifici che hai fatto per me, e anche per tutti noi. La vostra influenza positiva nella mia vita è inestimabile, e sono fortunata di avere genitori così straordinari. Vi voglio bene. Grazie a mio fratello Michele per essere stato un alleato in casa (e al telefono) quando quando ne avevo più bisogno. Grazie ai miei nonni eccezionali, Carmen, Suzanne, Enea e Piero, per avere colorato la mia vita. Grazie alle mie zie, zii e cugini: Ines, Oliviero, Lisa, Mattia, Vito, Paola, Leo e Francesco. Merci aussi à toi, petit Red, pour être notre *furry baby* et pour toujours me comprendre sans juger. Tu es un être spécial. I would like to also thank my family in law for making Geneva feel like home and for all their support during these years.

Last but not least, I want to thank the person who has my heart, my fiancé, Francis. Thank you for being by my side in this mad world for more than 8 years. You are the most amazing person and soon-to-be husband in the universe. Your love has been my anchor, providing unwavering support and encouragement since we met. You are the first-row witness of my academic journey, with all the struggles and successes. Your patience, understanding, and belief in me have been a source of strength: from late-night study sessions to celebratory moments, your presence has made every step of this journey more meaningful. I love you to the moon and back. Here's to many more years of love, shared dreams and accomplishments.