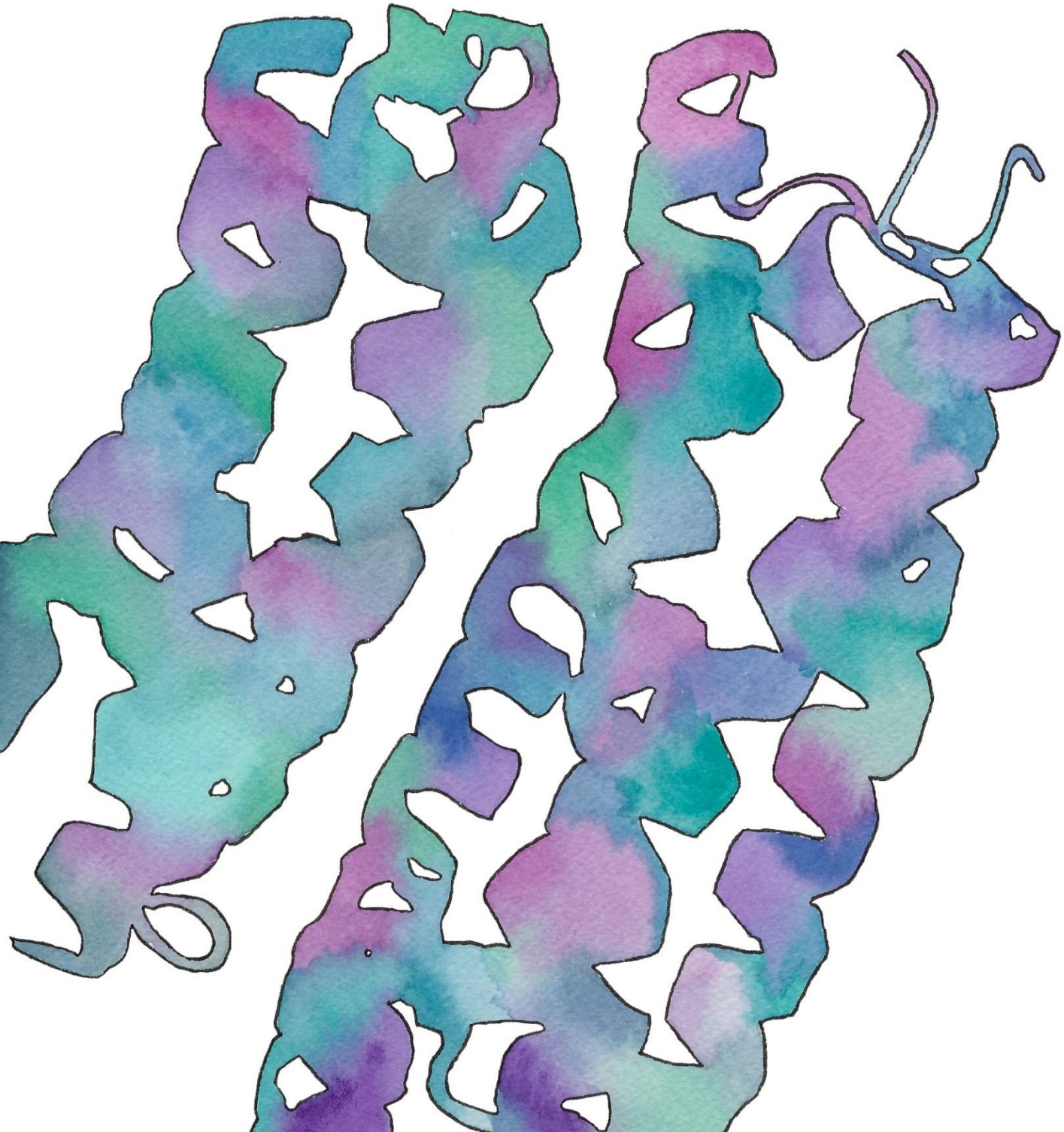
UNIVERSITAT DE BARCELONA

Insights into MLC pathophysiology: a biochemical and structural approach

Laura Ferigle Burgada

INSIGHTS INTO MLC PATHOPHYSIOLOGY: A BIOCHEMICAL AND STRUCTURAL APPROACH

Laura Ferigle Burgada



UNIVERSITAT DE BARCELONA

INSIGHTS INTO MLC PATHOPHYSIOLOGY: A BIOCHEMICAL AND STRUCTURAL APPROACH

Programa de Doctorat en Biomedicina

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
Cofinanciado por la Unión Europea
AGENCIA ESTATAL DE INVESTIGACIÓN

Proyecto PID2021-126246NB-100 financiado por MCIN/AEI/10.13039/50110001033/ y por FEDER, UE.

Tesi doctoral:

Laura Ferigle Burgada

Direcció:

Dr. Raúl Estévez Povedano

Dr. Ekaitz Errasti Murugarren

Departament de Ciències Fisiològiques

Facultat de Medicina. Campus Bellvitge

Universitat de Barcelona, Setembre 2024

PROGRAMA DE DOCTORAT EN BIOMEDICINA

UNIVERSITAT DE BARCELONA

Memòria presentada per optar al grau de doctor per la Universitat de Barcelona, 2024

LAURA FERIGLE BURGADA

L'interessada,

Laura Ferigle Burgada

Els directors de la tesi,

Raúl Estévez Povedano

Ekaitz Errasti Murugarren

Departament de Ciències Fisiològiques, UB

AGRAÏMENTS

Umuntu ngumuntu ngabantu, en zulú, es pot traduir com una persona és una persona gràcies a altres persones. I jo he d'agrair moltíssim a tots aquells que m'han concedit el seu temps i m'han acompanyat durant aquests anys dins i fora el laboratori.

En primer lloc, vull agrair als meus directors de tesi per guiar-me en aquesta aventura. He après moltíssim. Gràcies per estar sempre disponibles per resoldre dubtes i discutir resultats. Espero que quan sigueu grans expliqueu la batalleta de com vam aconseguir nanobodies a últim minut, perquè jo ho faré. A en **Raúl**, gràcies per donar-me l'oportunitat de fer la tesi al teu laboratori i confiar en mi pel projecte d'estructura. Espero que algun dia trobis la funció de MLC1 i vindré per a que m'ho expliquis fent molts esquemes en papers. A l'**Ekaitz**, per pujar-te al carro de MLC1 i donar sempre ànims. Trobaré a faltar perseguir-te pel passadís per poder parlar de ciència. Queda pendent que m'ensenyis a escalar!

Als companys de laboratori, des del primer dia amb pandèmia i restriccions em vaig sentir molt acollida, gràcies per estar sempre disposats a ajudar, el bon ambient i la companyia. A l'**Efren** que em vas ensenyar com funcionava el lab. No oblidaré mai el pic d'adrenalina de pujar a la teva moto portant la proteïna intentant que no es descongelés. A l'**Héctor**, el rei de la molecular i del lab en general, gràcies per interessar-te sempre pel projecte i sobretot, gràcies per tots els acudits dolents, que tothom sap que són els millors. A la **Marta**, em va agradar molt compartir un trosset d'aquest camí amb tu, espero que el teu postdoc sigui un èxit. A l'**Aida**, per l'ajuda i la simpatia, vaig trobar a faltar les tertúlies al patch quan vas marxar. A l'**Ash**, segurament la millor persona d'aquest planeta, gràcies per cuidar-nos i ajudar-nos quan ho hem necessitat. Per apuntar-te sempre a fer cerveses i també, gràcies per tornar. A en **Guillem**, per agafar el relleu i aguantar les meves queixes quan els bacteris no creixien, però sobretot gràcies per assegurar-te que la concentració de glucosa al lab estigui sempre en el rang de la molaritat. Et desitjo molta sort en la teva tesi i que tinguis el mínim de "muelles" possibles. I a l'**Adrià**, per fer un upgrade a la meva vida amb bona literatura, cafè i cervesa. Gràcies per fer-me riure una infinitat de vegades i aconsellar-me.

És sobretot quan tens un projecte difícil i fas balanç dels quatre anys quan t'adones de tota la gent que d'una manera o altra ha estat implicada i ha ajudat a tirar la feina endavant. Sense tots vosaltres, aquesta tesi no existiria.

A **Manuel Palacín**, por abrirme las puertas de su laboratorio y mostrar siempre interés en el proyecto. Gràcies també a tots els membres del grup per acollir-me i ajudar-me quan ho he necessitat. En especial, a en **Josep** i en **Xavi** per haver posat fàcil compartir expressions durant els primers anys.

A **Óscar Llorca** y a **Nayim** por ser los mejores colaboradores que una puede pedir. Mil gracias por estar siempre motivados con el proyecto e intentar sacarlo adelante. Gracias también por acogerme en vuestro laboratorio. Nayim, mi amiguito de la ciencia, fue un verdadero placer trabajar contigo con las megapurificaciones.

A **Beatriz Herguedas** y a **Irene** por dejarnos visitar vuestro laboratorio y ayudarnos con los baculovirus.

One of the best experiences during this PhD was to attend the nanobody workshop in Brussels. Thanks to all the attendees for their kindness and motivational scientific discussion. Also to **Els Pardon** for adapting the deadlines so we could send our samples and to all the technicians that were super helpful and nice, especially to **Eve**, for adopting me when I was clearly very lost.

To MLC Alliance for organizing the first meeting of MLC, it was very inspiring. Here is my grain of sand in contributing to the knowledge of the disease. I hope one day you will find a cure.

Als membres de la Protein Facility, sobretot a la **Mery**, per facilitar-nos les expressions quan no paràvem de demanar litres i litres de cultiu. A la **Marta Taulés,** per fer-nos un forat a últim moment per a fer servir el Biacore. I a l'**Àlex Perálvarez,** per cedir-nos les columnes de l'ÄKTA: n'hem tret profit.

Al grup de Farmaco perquè aquesta tesi no existiria sense vosaltres, literalment: gràcies per deixar ocupar-vos el luminòmetre setmana rere setmana, pel munt de reactius i plàsmids que us he demanat i per estar disposats a ajudar en tot moment. A en **Josep**, a qui més he molestat, gràcies per l'amabilitat, em va agradar molt poder treballar amb tu a cultius. A la **Laura** i la **Glò**, companyes de nanobodies i a en **Marc** (inserir emoticona d'una pistola) pels nervis de sortir del lab amb les mans ocupades.

A la gent del departament i la quarta planta per crear un molt bon ambient, sense el qual seria impossible treballar. Gràcies per les converses a mig passadís, els Carnestoltes, mojitades i cases rurals. En especial, he d'agrair a la **Carmen** (companya de curses) sempre és una alegria trobar-te al passadís i espero seguir trobant-nos fora del lab. I a l'**Ingrid**, pels esmorzars que feien començar bé el dia, ha sigut un plaer compartir el camí de la tesi amb tu.

A los de recepción de tardes, **Mari, Antonio** y **Fer**, gracias por vuestra simpatía y las charlas al salir. I al **personal de neteja**, per a fer veure que no us molestava que trepitgés el terra recent fregat quan arribava molt d'hora al matí (un cop més, ho sento).

A la **Judith**, per la teva professionalitat i tenir més paciència que una pedra.

Als meus amics de pis, és divertidíssim viure amb vosaltres. Gràcies per cuidar-nos, us estimo. A **María**, mi palentina-gironina-barcelonina favorita, gracias por reírte siempre y compartir penas y glorias del doctorado. En nada estamos celebramos tu tesis, chiguita! A la **Marta**, meteoròloga de confiança, sense tu no hagués anat al gimnàs ni la meitat de vegades que hi he anat. Gràcies per apuntar-te sempre a fer pastissos, pintar ventiladors o el que faci falta, perquè al final tot és un catxondeo. I a en **Dani**, sembla mentida que ja faci dotze anys que ens coneixem, gràcies per ser-hi sempre, en els moments de riure (moltíssims) i algun de plorar. Poques coses em fan més il·lusió que veure't dalt d'un escenari, espero seguir celebrant tots els èxits que vindran.

A la **Judit**, gràcies per demanar-me el telèfon un dia de casualitat perquè ara no sé què faria sense tu. Gràcies per escoltar-me sempre i pels cafès que se'ns allarguen hores i hores.

A la comunitat swing de Sabadell-Barcelona, perquè ballant s'arreglen els dies dolents i si és anant de festival encara més. A en **David**, gràcies per acollir-me a Madrid, va ser molt divertit passar aquella setmana amb tu, intentaré fer una nova ocupació al pis d'Alemanya. A en **Joan**, per ser-hi d'una forma o altra i interessar-te per les proteïnes i les microcoses. A la **Mireia** i la **Mercè**, perquè heu sigut un descobriment molt bonic. Espero seguir compartint moments amb vosaltres dins i fora del ball.

A l'**Ari** i en **Dani**, que hi sou des de fa una eternitat. Malgrat que hem pres camins molt diferents, sempre és una alegria seguir-nos trobant, no deixem mai de fer-ho.

A la meva família per cuidar-me sempre i preocupar-vos per mi. Als meus avis **Joan** i **Montse**, va ser una sort poder compartir tants anys amb vosaltres, us trobo a faltar. Als meus avis **Ramon** i **Carme**, que durant aquests anys us heu interessat pel que feia i em preguntàveu si les cèl·lules estaven bé. Avi, m'hagués agradat que veiessis la tesi acabada. Iaia, si acabar el doctorat no em converteix d'una vegada per totes en la néta preferida, jo ja no sé què fer.

A la **Marta**, tinc claríssim que he guanyat la loteria de germanes. Per ser sempre amiguetes aquí, a Sabadell, a Manresa o on faci falta. Als meus **pares**, espero que digueu a tothom que teniu una filla doctora de veritat i una de mentida i tingueu clar qui és qui. Bromes a part, qualsevol forma d'agraïment es queda curta quan m'ho heu donat tot. Espero que estigueu orgullosos de la vostra filla com jo ho estic que sigueu els meus pares. Us estimo molt, molt, molt. Moltíssim.

ABSTRACT

Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare type of leukodystrophy characterized by macrocephaly and white matter vacuolation. The pathogenesis of the disease is suggested to be caused by an impaired water and ionic homeostasis by glial cells. MLC is caused by mutations in *MLC1, GLIALCAM, GPRC5B* and *AQP4. MLC1* and *GLIALCAM* encode for membrane proteins that form a complex in astrocytes, whose exact function remains unclear. Following studies identified the G-protein-coupled receptor GPRC5B as an important member of the GlialCAM/MLC1 interactome and relevant to the regulation of related physiological processes. One of the main objectives of this thesis is to determine the role of GPRC5B and its signalling activity related to the pathophysiology of MLC. We determine that MLC1 has a negative modulatory effect on GPRC5B signalling pathways. Besides, we propose that mutations in *GLIALCAM* that encode for residues located in GlialCAM IgC2 domain are pathogenic due to the increased stability of GlialCAM oligomeric structures. Moreover, we give evidence that GlialCAM endocytosis is mediated by GPRC5B. We also study the recent identified mutations in *GPRC5B* known to cause MLC and observe a resistance to internalization resulting in the same outcome as those GlialCAM IgC2 mutations. Another objective of this thesis is to obtain the tridimensional structure of MLC1. We have achieved a preliminary structure of homo-trimeric MLC1 at a resolution of 7 Å, approximately. At the same time, we have developed nanobodies that specifically recognize MLC1 to help orientate particles during cryo-electron microscopy (cryo-EM).

INDEX

LIST OF FIGURES AND TABLES

ABBREVIATIONS

2xTY: Yeast extract tryptone

3D: Tridimensional

AEBSF: 4-benzenesulfonyl fluoride hydrochloride

AF: Alpha-Fold

APS: ammonium persulfate

AQP: aquaporin

AraC: Cytosine β-D-arabinofuranoside

BBB: Blood-brain barrier

BCA: Bicinchoninic acid assay

Bp: base pair

BRET: Bioluminiscence resonance energy transfer

BSA: Bovine serum albumin

BV: Bed volume

C12E8: Octaethylene glycol monododecyl ether

CAM: cell adhesion molecule

cDNA: complementary deoxyribonucleic acid

CFU: colony forming unit

CHS: Cholesteryl hemisuccinate Tris salt

CMC: Critical micelle concentration

CNS: Central nervous system

Cryo-EM: Cryogenic electron microscopy

DAPI: 4’,6-diamidino-2-phenylindole

DDM: n-dodecyl-β-D-maltoside

DGC: Dystrophin glycoprotein complex

DM: n-decyl-β-D-maltopiranoside

DMEM: Dulbecco’s Modified Eagle Medium

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

DNPP: 4-nitrophenyl phosphate disodium salt hexahydrate

dNTPs: Deoxyribonucleotides

DTT: Dithiothreitol

E. coli: *Escherichia coli*

ECL: Enhanced chemiluminescence substrate

EDTA: Ethylenediaminetetraacetic acid

ER: Endoplasmic reticulum

FACS: Fluorescent activated cell sorting

FBS: Fetal bovine serum

FRET: Fluorescence resonance energy transfer

FSEC: Fluorescence size exclusion chromatography

GFP: Green fluorescent protein

GPCR: G-protein coupled receptor

HPLC: High-performance liquid chromatography

HRP: Horseradish peroxidase

Ig: Immunoglobulin

IPTG: Isopropyl β- d-1-thiogalactopyranosid

K_D: Equilibrium dissociation constant

kDa: Kilodalton

KI: Knock-in

KO: Knock-out

LAPAO: Lauramidopropylamine oxide

LB: Luria broth

LDAO: Lauryldimethlamine oxide

Lg: LargeBiT

LSB: Loading sample buffer

MLC: Megalencephalic leukoencephalopathy with subcortical cysts

MRI: Magnetic resonance imaging

MS: Mass spectrometry

MW: Molecular weight

Nb: Nanobody

Ni-NTA: Nickel-nitrilotriacetic acid resin

NS: Negative staining

O/N: Overnight

OD: Optical density

PAE: Predicted aligned error

PBL: Peripheral blood lymphocytes

PBS: Phosphate buffered saline

PCR: Polymerase chain reaction

pDEST: Destination vector

PEG: Polyethylene glycol

PEI: Polyethylenimine

PFA: Paraformaldehyde

Pfu: Plaque-forming units

PLA: Proximity ligation assay

PMSF: Phenylmethylsulfonyl fluoride

RFU: Relative fluorescent units

RNA: Ribonucleic acid

RNase: Ribonuclease

RT: Room temperature

RT-PCR: Reverse transcription polymerase chain reaction

RU: Resonance units

RVD: Regulatory volume decrease

sAB: Semi-synthetic antibody

SD: Standard deviation

SDS: Sodium dodecyl sulfate

SEC: Size exclusion chromatography

Sm: SmallBiT

SOC: Super optimal broth

SOE: Splicing by overlap extension

SPR: Surface plasmon resonance

SST: Silver stain

TAE: Tris-acetate-EDTA buffer

TBS: Tris saline buffer

TEMED: Tetramethylethylenediamine

TES: Tris-EDTA buffer

TEV: Tobacco etch virus

TM: Transmembrane domain

TTBS: TBS-Tween 20

UV: Ultraviolet

V: Volts

VRAC: Volume-regulated anion channel

WB: Western blot

WT: Wild type

INTRODUCTION

The main purpose of this thesis has been to gain a better understanding into the pathophysiology of Megalencephalic Leukoencephalopathy with Subcortical Cysts disease by improving our knowledge of the physiological role and the regulation of the proteins known to be involved in it. Since our group is mostly focused on studying this disease, several thesis have pursued the same objective. However, this project has particularly focused on the role of GPRC5B on MLC pathogenesis and the structural resolution of MLC1 protein.

We have studied the interplay between GPRC5B and its signalling with MLC1 and GlialCAM proteins. Right before the beginning of this thesis, GPRC5B was identified by our group as a member of the MLC1/GlialCAM interactome. During this project, GPRC5B has also been identified by others as the third gene causing MLC, which has gained interest in understanding its role and regulation of signalling pathways.

In this Introduction, we first sum up MLC and its main features, both at a clinical and at a genetic-biological level. We present MLC1 and GlialCAM extensively, including the available information regarding their role in MLC pathogenesis based on their mutations. The second part of the introduction consists of an overview of G-Protein-Coupled Receptors (GPCRs) and their classification. We focus on orphan receptors GPR37, GPR37L1 and GPRC5 family due to their relevance in the pathophysiology of MLC disease. Lastly, since one of the main objectives of this thesis has been to obtain the tridimensional structure of MLC1 protein, the third chapter of this introduction reviews the study of membrane proteins in structural biology from different methods for protein expression to the strategies followed for achieving a good particle orientation.

1. MLC DISEASE

1.1 CNS OVERVIEW

The Central Nervous System (CNS) is responsible for receiving, processing, and responding to sensory information. It consists of two basic cell types: neurons and glial cells. Neurons are specialized cells in charge of transmitting rapid electric signals in form of action potentials. They connect with each other through synapses that are specialized intracellular adhesion sites. During the neuronal signalling process, the action potential propagates through the axon of the neuron to the pre-synaptic region of the neighbouring neuron, where neurotransmitters are released. These molecules will bind to the receptors from the post-synaptic membrane of a second neuron, which will continue to transmit the electric signal.

Glial cells are the most abundant cells in the brain. They are responsible of maintaining the homeostasis, modulating the synapsis, and mediating immune responses, among other functions. Glial cells in the CNS can be classified into four types based on their morphology, function, and localization: astrocytes, microglia, ependymal cells, and oligodendrocytes (Kettenmann and Verkhratsky, 2011).

In the CNS, astrocytes are the most abundant cell type. They provide trophic and structural support to the neurons. Microglia are known as the resident macrophage of the CNS. They are the principal effector of neuroinflammation and can become promptly activated with different stimuli. Ependymal cells are connected through adherent, gap, and tight junctions forming the cerebral epithelium, which is a fundamental part of the blood-brain barrier (BBB). Oligodendrocytes produce myelin sheaths to wrap neuronal axons. They have a small soma with long branched prolongations.

Myelin is the specialized plasma membrane of oligodendrocytes. It is distributed as concentric layers around the axon, alternating electrodense layers with clear layers. The first ones are the result of fusion between membranes, while the clear ones are due to extracellular matrix fusion. Mature myelin is the laminar structure resulting from the compact layers. Myelin sheaths wrap axons dividing them into segments called internodes, which provide the axons with electric insulation. The spaces between internodal areas are not insulated and they are called the Ranvier nodes. Internodal areas contain dense ion channels and transporters responsible for regenerating action potentials at the end of each internode. The action potential has to ‘jump’ from one node

to the next one in a process called saltatory conduction (Huxley and Stämpfli, 1949). It allows for rapid nervous impulse transmission and minimizes its energy expense.

Myelinated axons conform the white matter. During the first stages of postnatal brain development, myelination occurs (Kamholz, 1996). The cells responsible for this process are oligodendrocyte precursor cells (OPCs) which migrate from germinal zones to unmyelinated axons. Both formation and maintenance of white matter are essential for neuronal network connectivity, which is the basis of brain function (Domingues et al., 2016, Liedtke et al., 1996).

Myelin alterations can lead to a wide variety of neurological diseases. Autoimmune diseases can result in myelin damage, for example, multiple sclerosis. Environmental factors, such as the exposure of organic solvents, can impair the myelin function. Genetic diseases caused by mutations in myelin-related genes are also described. Among the genetic disorders, leukodystrophies are found.

1.2 DEMYELINATING LEUKODYSTROPHIES

The term leukodystrophy has a Greek origin that translates as follows: *leuko* for white, *dys* for abnormal and *trophy* for growth. Leukodystrophies were originally defined as a group of genetic neurodegenerative disorders, primarily affecting the white matter in the CNS. However, with the increasing knowledge on the subject they were redefined as heritable disorders affecting the white matter of the CNS, sharing glial cell or myelin sheath abnormalities. The neuropathology of leukodystrophies is primarily characterized by the involvement of oligodendrocytes, astrocytes and other non-neuronal cell types, although in many disorders the mechanism of disease remains unknown, and in other cases is suspected to include significant axonal pathology (Vanderver et al., 2015).

Leukodystrophies include a wide range of hereditary disorders. Concerning their clinical evolution, they can be progressive, lethal, static, or even ameliorate with time. The onset may be at any age, from prenatal life to senescence. The wide range of genes identified to cause leukodystrophies lead to a subclassification of the disorder. Initially, leukodystrophies were classified into four groups: hypomyelinating (deficient myelin production), demyelinating (myelin degradation), dysmyelinating (biochemically and structurally abnormal myelin production) and myelinolytic (vacuolated myelin) (Naidu, 1999).

Later, Dr Marjo van der Knaap proposed a new classification that included astrocytes, neurons, microglia, and blood vessels defects to the existing criteria of alterations in oligodendrocytes and white matter (van der Knaap and Bugiani, 2017). The designed classification is displayed in Table 1.

MYELIN DISORDERS

Hypomyelination

a. Pelizaeus-Merzbacher disease
b. Peripheral neuropathy, central hypomyelination, Waardengurg-Hirschsprung
c. Cx47-related Pelizaeus-Merzbacher-like disease
d. Hypomyelination on early myelinated structures

Demyelination

a. Metachromatic leukodystrophy
b. Multiple sulfatase deficiency
c. Globoid cell leukodystrophy (Krabbe disease)
d. X-linked adrenoleukodystrophy, cerebral form

Myelin vacuolization

a. Mithocondrial diseases with leukoencephalopathy
b. Phenylketonuria
c. Canavan disease
d. Other selected disorders of amino acid metabolism
e. Cx32-related (X-linked) Charcot-Marie-Tooth disease

Astrocytopathies

a. Alexander disease
b. Megalencephalic leukoencephalopathy with subcortical cysts
c. Clc-2 - related disease
d. Vanishing white matter
e. Aicardi-Goutières syndrome and variants
f. Oculodentodigital dysplasia (Cx43)
g. Giant axonal neuropathy

Leuko-axonopathies

a. Hypomyelination with atrophy of the basal ganglia and cerebellum
b. Hypomyelination with congenital cataract
c. Early-onset neuronal degenerative disorders
    1. Gangliosidosis GM1 and GM2
    2. Infantile neuronal ceroid lipofuscinosis
    3. *AGC1*-related disease
    4. *AIMP1*-related disease
    5. *HSPD1*-related disease
d. Pol III-related leukodystrophies
e. Leukoencephalopathy with brainstem and spinal cord involvement and high lactate
f. Hypomyelination with brainstem and spinal cord involvement and leg spasticity
g. Giant axonal neuropathy

Microgliopathies

a. *CSF1R*-related disorders
    1. Hereditary diffuse leukoencephalopathy with spheroids
    2. Pigmentary ortochromatic leukodystrophy
b. Nasu-Hakola disease

Leuko-vasculopathies

a. Cerebral autosomal dominant arteriopathy with subcortical infracts and leukoencephalopathy
b. Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy
c. Cathepsin A-related arteriopathy with strokes and leukoencephalopathy
d. Leukoencephalopathy with calcifications and cysts

Table 1. Classification of leukodystrophies. The table is meant to give examples and not to be exhaustive. Adapted from (van der Knaap and Bugiani, 2017).

1.3 MEGALENCEPHALIC LEUKOENCEPHALOPATHY WITH SUBCORTICAL CYSTS (MLC)

Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare genetic leukodystrophy disorder classified into the group of astrocytopathies. MLC was first discovered in 1995 by Dr van der Knaap who described eight children with cerebral leukoencephalopathy and megalencephaly (van der Knaap et al., 1995). Independently, the following year Dr Singhal described thirty patients in India presenting the same disease (Singhal et al., 1996). As commonly seen in rare diseases, their prevalence is usually found increased in populations where consanguinity is prevailing (Topcu et al., 1998). In this case, MLC is most prevalent in the Agrawal community in East India (Gorospe et al., 2004), Jewish families in Lebanon and Turkey (Ben-Zeev et al., 2002), and some families in Japan (Shimada et al., 2014) and Italy (Montagna et al., 2006), among others.

MLC diagnosis is made based on characteristic abnormalities observed with brain Magnetic Resonance Imaging (MRI) (Figure 1). The alterations may include:

- Bilateral atrophy and swelling in brain hemispheres.
- Abnormally diffuse white matter.
- White matter in the corpus callosum, internal capsule and brain stem are better preserved. Subcortical regions and periventricular areas in the occipital brain are also less affected.
- White matter in the cerebellum not swollen, although the signal is lightly abnormal.
- Subcortical cysts in the anterior-temporal brain. Cysts can grow in number and size with time. In some cases, they expand to the great majority of white matter.
- Cerebral atrophy replaces inflammation in white matter as the disease progresses. MRI signal can return to normality in some individuals.
- Unaffected grey matter.

The first symptom for most patients is an increased head circumference within the first year of life accompanied by a mildly delayed initial motor development. Other symptoms may include:

- Macrocephaly during the first year of life that can reach 4 to 6 SD above population mean. Head growth is then reduced to normal and the growth curve for head size in MLC patients continues a parallel evolution to healthy children.

- Motor function slow decline, mild cerebellar ataxia and spasticity appearing in late-childhood and adolescence. Patients present difficulty walking and postural instability.
- Dysarthria.
- Epileptic seizures in early stages.
- Temporary exacerbation of signs and symptoms after minor head trauma.
- Cognitive impairment with a mild to late-onset.
- Behavioural problems.
- Autism.

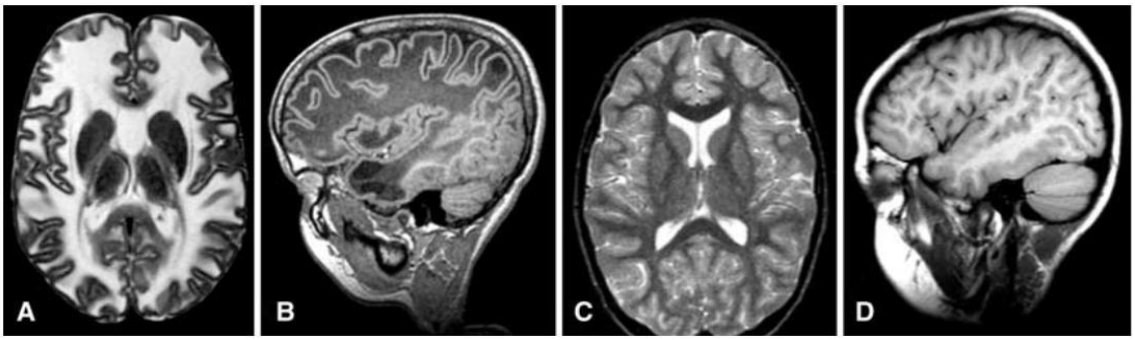


Figure 1. Brain MRI of an MLC patient. Comparison between an MRI from an MLC patient (A, B) and one from a healthy individual (C, D). The transverse T2-weighted images show diffusely abnormal, slightly swollen white matter in the patient (A compared to C). The sagittal T1-weighted images show subcortical cysts in the anterior-temporal and parietal subcortical region (B compared to D). Extracted from (Leegwater et al., 2002).

A wide variability in pace and severity of the disease has been identified, even for patients carrying the same mutation or among relatives (Blattner et al., 2003), suggesting that environmental factors might influence the clinical outcome. Given the fact that MLC has been known for a relatively short time, information about the lifespan is limited. There are documented cases of patients alive in their fifties, while others have died in their teens or twenties.

Furthermore, two phenotypes of MLC disease were described: a progressive and a remitting phenotype. In fact, it was not until 2010 when it was identified a group of patients showing a mild MLC phenotype (van der Knaap et al., 2010). Although those patients presented the typical MLC symptomatology during the first years of life, their symptoms ameliorated with time. Patients showed macrocephaly during the first years, small cognitive impairment, and difficulty in motor function without ataxia nor spasticity. Epileptic seizures were present, but they also improved with age (López-Hernández et al., 2011a). MRI scans from an MLC patient presenting a remitting phenotype are displayed in Figure 2.

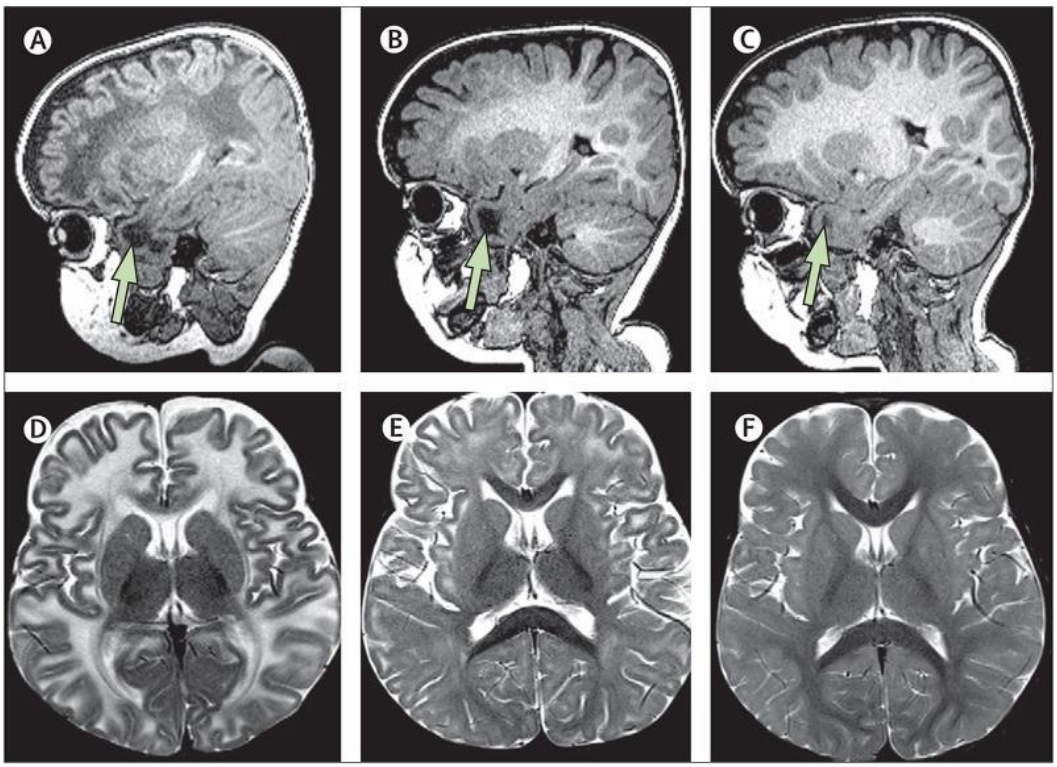


Figure 2. Brain MRI from MLC patients with remitting phenotype. Sagittal T1-weighted (A-C) and axial T2-weighted (D-F) images from a patient with remitting MLC at ages of 9 months (A, D), 1.5 years (B, E) and 3.5 years (C, F). The earlier MRI fulfils the standard criteria for MLC, but the following scans did not: white matter abnormalities had improved and the anterior temporal cyst was reduced (shown in arrows). At 3.5 years, the subcortical white matter was still slightly T2-hypertense (F). Extracted from (van der Knaap et al., 2012).

Thus, according to the clinics, MLC patients can be classified into those presenting a 'classical' phenotype and those who show a remitting phenotype. As it will be explained later in this introduction (*see Section 1.3.1*), genetic differences are responsible for these different variants of the disease.

A decrease in the concentration of various osmolytes in the brain of MLC patients was observed using proton-magnetic resonance spectroscopy. These results suggested that there would be an accumulation of water in the cerebral tissue of the patients (Brockmann et al., 2003, Sener, 2003).

Histopathological studies revealed a spongiform degeneration of white matter and vacuole formation in outer myelin layers. Inner layers, however, were unaffected (van der Knaap et al., 1996). The observation that vacuoles are solely found in outer parts of the myelin suggested that they would not affect dramatically the nervous transmission.

1.3.1 MLC GENETICS

Genetic linkage analysis from affected Turkish families determined that the first gene causing MLC was located on the chromosome 22 (Topçu et al., 2000). Posterior studies identified that the gene *KIAA0027* [MIM604004] located in the locus of chromosomes 22q13.33 was causing MLC. This was the first gene associated with the disease and it was consequently renamed *MLC1* (Leegwater et al., 2001).

MLC1 gene comprises 26 kb and 12 exons, the first of them being non-coding. Two alternative transcripts (NM_015166.3 and NM_139292.2) have been described. Even though they differ in the first exon, they result in the same mRNA that codifies for the same MLC1 protein. In a recent literature review, 151 *MLC1* variants have been found in patients with a confirmed MLC diagnosis. Around 40% of these mutations are missense, 30% are either deletions or insertions and 30% are splicing mutations. Nonsense mutations represent only 2% of the described cases (Passchier et al., 2024).

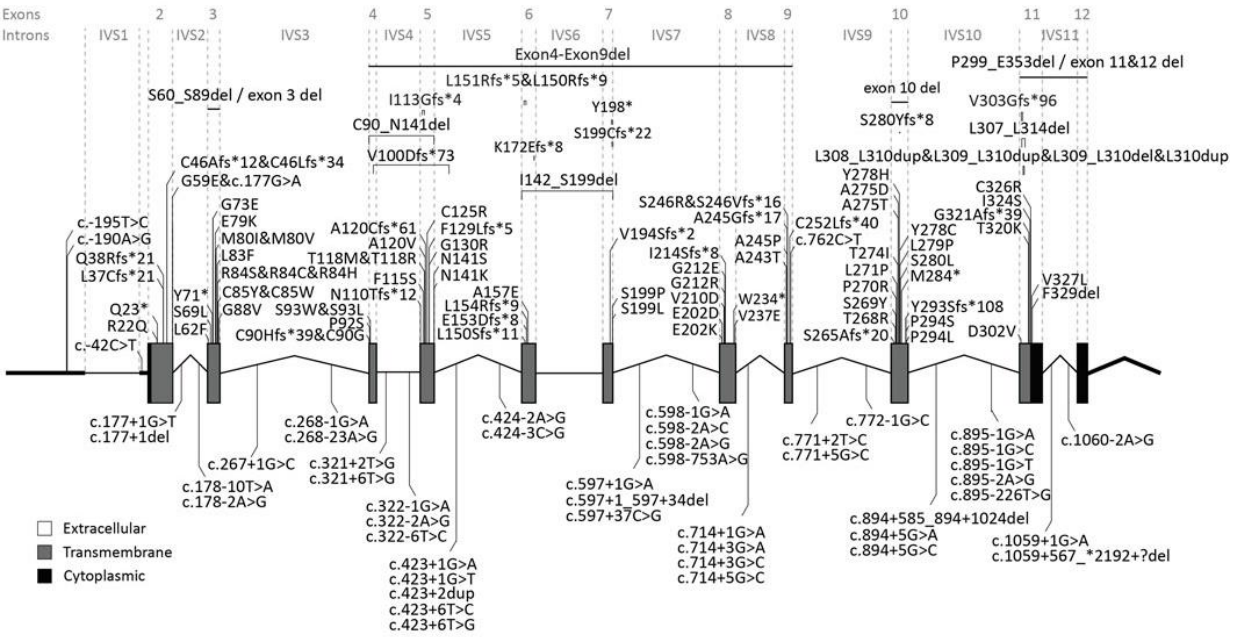


Figure 3. Overview of MLC1 variants found in MLC patients. *MLC1* gene is depicted. Exons are indicated by blocks and introns by lines, both are drawn in scale. All variants are schematically indicated above (exonic) or below (intronic) the gene. Variants are represented approximately in their relative positions. Extracted from (Passchier et al., 2024).

Approximately 76% of MLC patients bear mutations in *MLC1* (Leegwater et al., 2002, Topcu et al., 1998). Their inheritance pattern is autosomal recessive. Most MLC patients carrying mutations in *MLC1* present the classic phenotype. However, two patients with mutations in *MLC1* presenting radiological improvement have been identified (Hamilton et al., 2018). Furthermore, in this thesis, we have also identified and characterized an MLC patient carrying two *MLC1* variants showing radiological improvement (*see Results Chapter 1*). In Figure 3, *MLC1* gene is displayed with all the described mutations.

Around 22% of MLC patients harbour mutations in *HEPACAM* gene [MIM611642]. Although the gene was initially called *HEPACAM*, we prefer the name *GLIALCAM* because of its prominent expression in glial cells in the brain. This gene was linked to MLC by our research group in a proteomic analysis aimed at the identification of MLC1 protein interacting partners. Thanks to this discovery, *GLIALCAM* mutations were identified as the second disease-causing gene (López-Hernández et al., 2011a).

GLIALCAM is located in the locus 11q24.2. It comprises 17 kb and 7 exons with a non-coding region at the 3' end. In a recent literature review, 29 *GLIALCAM* variants have been described in MLC patients. Around 70% of these mutations are missense, 3 are deletions and 4 of them are nonsense mutations. Only one splicing mutation has been identified (Passchier et al., 2024).

Mutations in *GLIALCAM* can cause two distinct phenotypes named MLC2A or MLC2B based on their inheritance pattern. MLC2A corresponds to the classic phenotype and it is inherited in an autosomal recessive manner. Most mutations causing this phenotype are missense. On the contrary, MLC2B conforms the benign or remitting phenotype and it is inherited in an autosomal dominant manner (López-Hernández et al., 2011a, van der Knaap et al., 2010). It is only caused by missense mutations (Arnedo et al., 2014a, van der Knaap et al., 2012). In Figure 4, all *GLIALCAM* variants found in MLC patients are displayed.

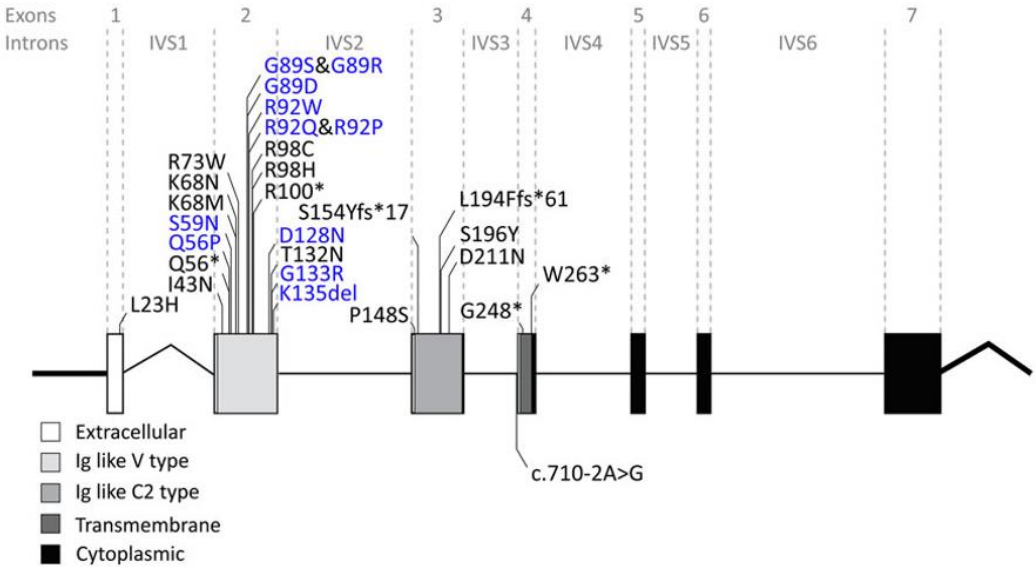


Figure 4. Overview of GLIALCAM variants found in MLC patients. *GLIALCAM* gene is depicted. Exons are indicated by blocks and introns by lines, both are drawn in scale. All variants are schematically indicated above (exonic) or below (intronic) the gene. Variants are represented in their relative positions. Dominant *GLIALCAM* variants are marked in blue, while recessive variants are marked in black. Extracted from (Passchier et al., 2024).

For years, a small percentage of MLC cases (2%) could not be explained by mutations in the previous identified disease-causing genes. It was not until recently that genetic studies performed in five MLC patients not bearing any pathogenic variants in *MLC1* nor *GLIALCAM* discovered two new genes responsible for causing MLC (Passchier et al., 2023). These genes are *GPRC5B* and *AQP4.*

GPRC5B [MIM605948] is located in locus 16p12.3. It comprises 28 kb and 3 exons. Two duplications arose *de novo* were identified in three patients. These mutations are dominant and display a non-remitting phenotype. On the other hand, *AQP4* [MIM600308] is located in locus 18q11.2. It comprises 14 kb and 4 exons. Only two siblings born of consanguineous parents presented a missense mutation in this gene. This recessive mutation cuases a remitting phenotype. In Figure 5, variants found in *GPRC5B* and *AQP4* are represented.

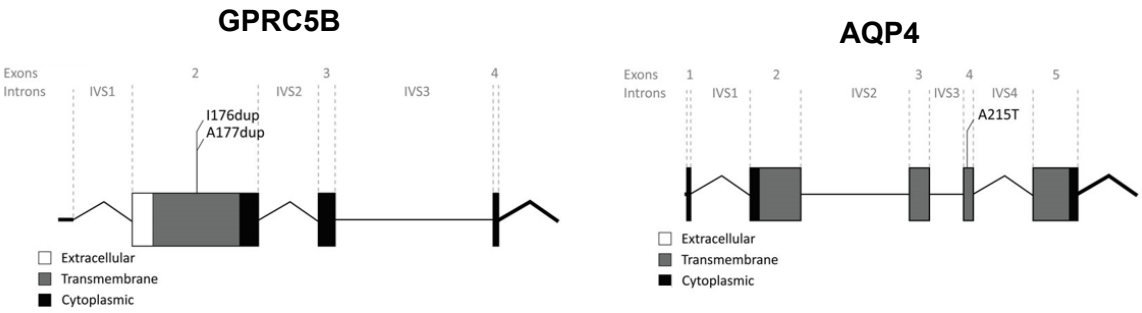


Figure 5. Overview of GPRC5B and AQP4 variants found in MLC patients. On the left, *GPRC5B* gene is depicted. Exons are indicated by blocks and introns by lines, both are drawn in scale. Described variants are schematically indicated above the gene in their relative positions. On the right, *AQP4* gene is depicted. Exons are indication by blocks and introns by lines, both drawn in scale. The single identified variant is schematically indicated above the gene in its relative position. Extracted from (Passchier et al., 2024).

1.3.2 MLC1 PROTEIN

Generalities, Expression, and Localization Studies of MLC1

MLC1 encodes for a membrane protein with the same name. Its length is 377 amino acids and its molecular weight is 41 kDa. *MLC1* orthologues have been identified in chordate families. Its alignment has proved that the protein is highly conserved in myelin-producing vertebrate species, specifically in the putative TM domains and at the end of the C-terminus, whereas the N-terminus present higher variability.

Eight transmembrane (TM) domains were predicted from hydropathy plot analysis. These domains are bound by four extracellular loops and three intracellular loops, being the second one significantly larger. Both N- and C- terminus are intracellular (Boor et al., 2005). Transmembrane topology of MLC1 was validated experimentally by generating truncated MLC1 constructs with a tag and comparing the signal intensity in permeabilizing and non-permeabilizing conditions (Hwang et al., 2021). Besides, AlphaFold MLC1 structural models also suggested that MLC1 conformation consisted on eight TM domains (Jumper et al., 2021, Varadi et al., 2022). In Figure 6, MLC1 putative structure is represented.

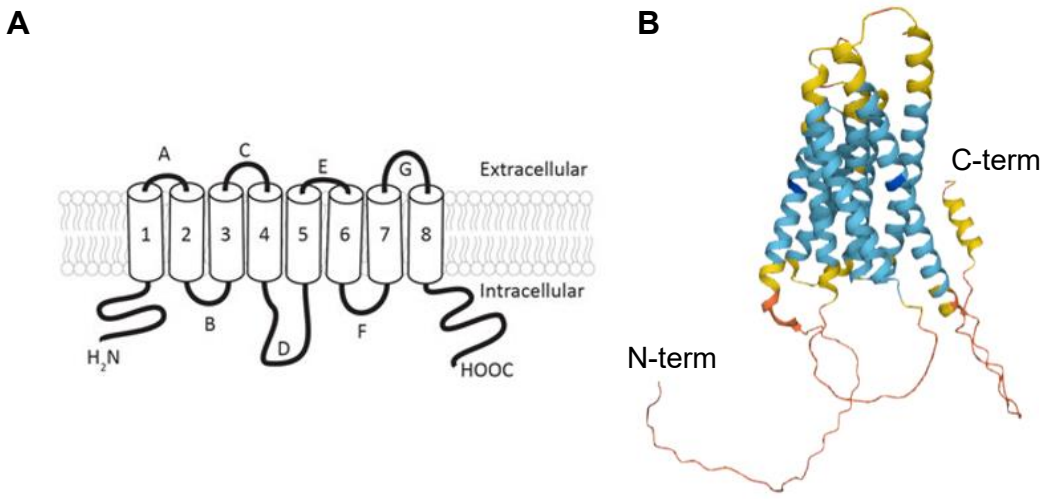


Figure 6. Predicted MLC1 structure. A. Schematic representation of MLC1 in the cell membrane. TM domains are marked in numbers, loops are marked with letters. Extracted from (Passchier et al., 2024). **B.** Predicted AlphaFold structure for MLC1. The coloured pattern corresponds to a confidence score generated by AF. Dark blue: very high confidence, light blue: confident, yellow: low confidence and orange: very low.

MLC1 protein is mainly expressed in the brain, although it has also been detected in leukocytes (Boor et al., 2005). In the brain, MLC1 is found in cerebellar Bergmann glia and in astrocytes surrounding subpial and perivascular spaces. Contrarily to what was initially reported, MLC1 is not expressed in neither oligodendrocytes, microglia nor neurons (Schmitt et al., 2003, Teijido et al., 2004, Boor et al., 2005). High-resolution electron microscopy (immunogold) in mouse and human tissue revealed the localization of MLC1 at astrocyte-astrocyte junctions in the perivascular area (Figure 7A) (Teijido et al., 2007, Duarri et al., 2011).

In humans, MLC1 expression levels are the highest during the first three years of life and then they decreased to stabilize after approximately five years (Dubey et al., 2015). This observation correlates with the myelination process, which is most rapid in the first years of life and reaches near-completion at three to four years. Considering that only species that produce myelin have the *MLC1* gene, these results altogether indicate that MLC1 protein might play an essential role in brain physiology when myelin is present.

Primary cultured astrocytes have been extensively used to study MLC1 at a biochemical level. Initial studies detected endogenous MLC1 protein intracellularly in the endoplasmic reticulum (ER) and in the endosomes (Ambrosini et al., 2008). Posterior studies showed that the treatment with cytosine β-D-arabinofuranoside (AraC), which is a selective inhibitor of DNA synthesis that is used to arrest cells at G0 phase, lead to MLC1 localizing at the plasma membrane and astrocytic junctions in astrocyte cultures (Duarri et al.,

2011). With this, the model replicated *in vitro* the observed MLC1 biochemical behavior seen *in vivo*. Therefore, quiescent cultured astrocytes treated with AraC were established as an MLC study model (Figure 7B).

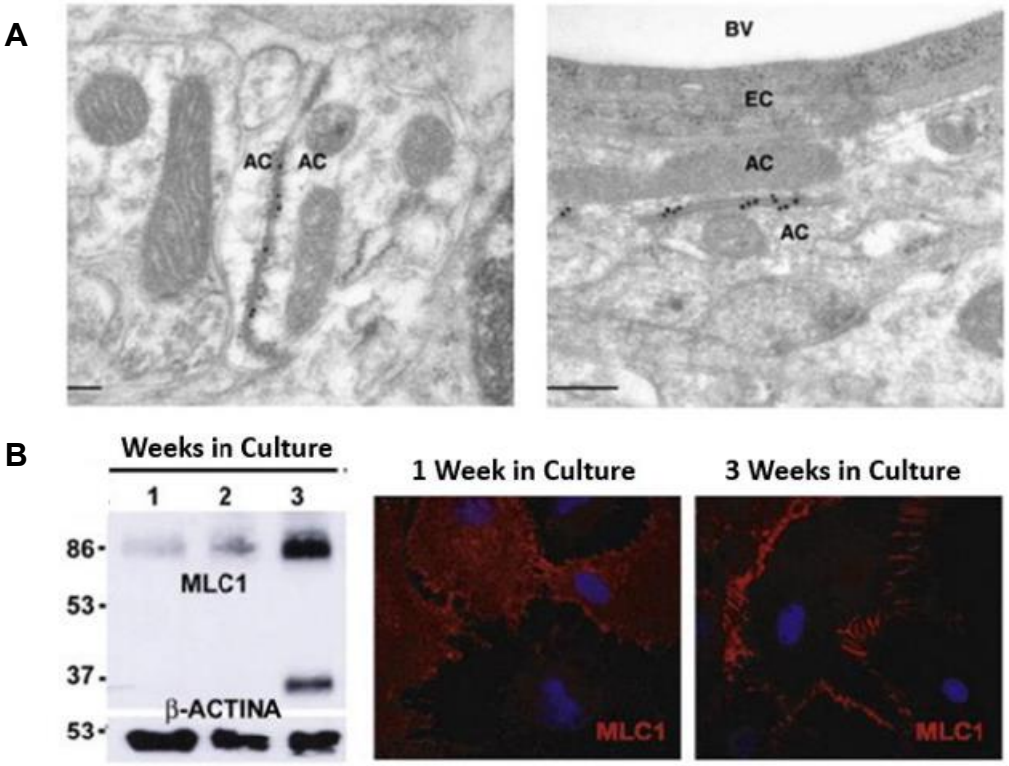


Figure 7. MLC1 localization at tissular level and in primary cultured astrocytes. A. EM images of MLC1 localization at astrocyte-astrocyte junctions, in astroglial processes (left) and perivascular areas (right). AC: astrocyte, BV: blood vessel, EC: endothelial cell. Scale bar 0.25 µm. Modified from (Teijido et al., 2007). **B.** Correlation between MLC1 levels by WB (left) and localization at astrocyte-astrocyte junctions by immunofluorescence assays. After three weeks of treatment with AraC, MLC1 levels increase to its peak expression. Then, the observed phenotype replicates in vivo observations for MLC1 localization.

Thanks to this model, it was described that MLC1 co-localized with tight junctions markers ZO-1 and occludin, as well as with adherent junctions markers β-catenin and N-cadherin (Duarri et al., 2011). Other studies hypothesized that MLC1 could be related to proteins of the dystrophin-associated glycoprotein complex (DGC) which is essential for the anchoring of aquaporin and potassium channels at astrocytic endfeet at the BBB (Ambrosini et al., 2008). Yet, the localization of α- and β-dystroglycan (proteins known to be part of the DGC complex) were not altered in the *Mlc1* knockout (KO) mice (Dubey et al., 2015). Neither was affected the MLC1 expression in the *dystrophin* KO mice (Duarri et al., 2011). In conclusion, although MLC1 is functionally linked to DGC, a clear mechanism of its interaction is yet to be provided.

1.3.2.1 STRUCTURE OF MLC1

To date, the structure of MLC1 protein is not known, neither is its function. Amino acid sequence analysis suggests that *MLC1* originated from the duplication of a smaller gene encoding a protein with 4 TM domains. This hypothesis is sustained on the evidence that the first and the second halves of the protein present domain homology. Furthermore, the fourth and the eighth TM domains contain a segment rich in polyleucine and the intracellular loop between TM 4 and 5 is not conserved among orthologues. Biochemical experiments done in *Xenopus laevis* oocytes by our research group supported this hypothesis (Estévez et al., 2018). The two halves of MLC1 reached the plasma membrane only when they were co-expressed, indicating that the two halves interact and are able to restore the function (Figure 8A).

Other biochemical studies determined that MLC1 homo-oligomerizes. Analysing protein expression by WB, two bands are detected for MLC1 at 37 and 72 kDa, approximately. This is compatible with monomeric and dimeric structures. By split-tobacco etch virus (TEV) assays, the ability to homo-oligomerize was confirmed (Capdevila-Nortes et al., 2012) (Figure 8B). Recently, it was reported that MLC1 forms a homo-trimeric complex in detergent micelles and proteoliposomes (Hwang et al., 2021).

Further characterization of the protein revealed that MLC1 displays a putative glycosylation site between TM 3 and 4, although it is not glycosylated in heterologous expression cell systems (Teijido et al., 2004). Both N- and C- terminal domains are phosphorylated. Serine 27 is phosphorylated by Protein Kinases A and C (PKA and PKC), while Serine 339 is only phosphorylated by PKC (Lanciotti et al., 2010). Posterior studies evidenced the presence of a putative ER retention motif based on arginine (RXR) in the N-terminal region. New potential phosphorylation sites were also identified (Brignone et al., 2015) and a caveolin-binding motif. MLC1 was found present in caveolin lipid rafts from rodent-derived astrocytes (Lanciotti et al., 2010). A summary of the post-translational modifications and motifs is displayed in Figure 8C.

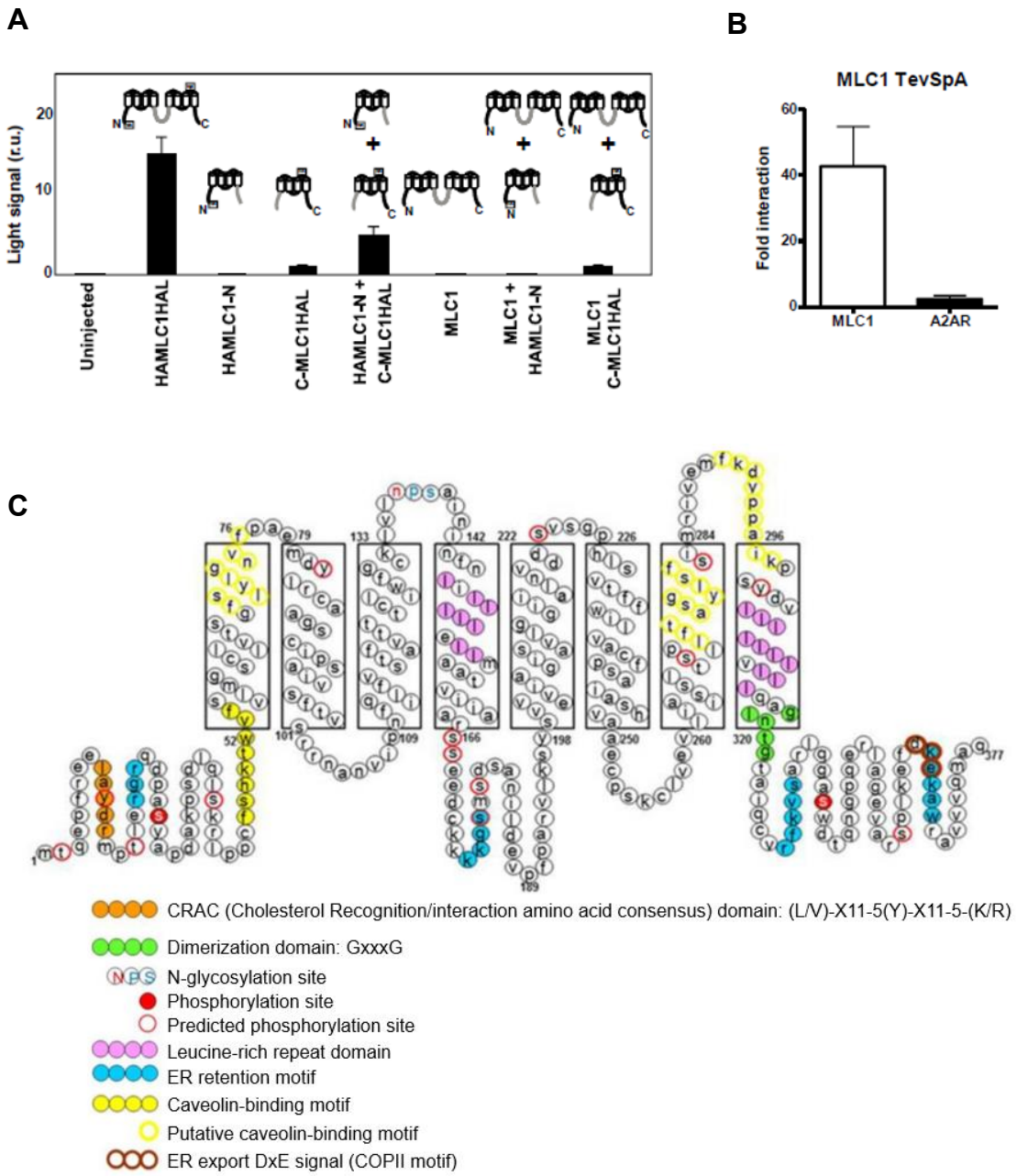


Figure 8. Biochemical features of MLC1 protein. A. Surface expression of MLC1 in *X. laevis* oocytes. MLC1 was expressed either whole or split at the intracellular loop between TM 4 and 5. Splits are HMLC1-N and C-MLC1HAL in the graph. Alone, the splits do not reach the plasma membrane, but, they do so when co-expressed. **B.** Split-TEV analysis of MLC1 homo-oligomerization. A2AR was used as a negative control. Modified from (Estévez et al., 2018). **C.** Schematic representation of specific amino acid consensus motifs found in MLC1 protein. Putative and predictive information is represented in empty circles, confirmed information is plotted in filled circles. Modified from (Brignone et al., 2015).

1.3.2.2 *MLC1* MUTATIONS

MLC1 mutations are found along the gene (Leegwater et al., 2002). Most missense mutations are responsible for a reduction in protein expression levels resulting in a reduction in membrane surface expression (Lanciotti et al., 2010, Montagna et al., 2006, Teijido et al., 2004). The proposed mechanism behind this observation is that the mutation would lead to protein misfolding, resulting in an unstable protein to be degraded by the ER or in lysosomes. Even though some mutated MLC1 forms can reach plasma membrane, they are still targeted for lysosomal degradation (Duarri et al., 2008).

1.3.3 GLIALCAM PROTEIN

Generalities, Expression, and Localization Studies of GlialCAM

GLIALCAM encodes for a membrane protein with the same name. Its length is 416 amino acid and it weighs 72 kDa. It is an adhesion cell molecule belonging to the Ig Superfamily (Moh et al., 2005).

GlialCAM is predicted to have two extracellular domains, a single transmembrane segment and an intracellular C-terminal domain or cytoplasmic tail. In the extracellular domain, two Immunoglobulin-like loops are found, being IgV in N-terminal end and IgC2 next. Six putative N- and O-glycosylation have been identified, as well as a signal peptide (Figure 9) (Moh et al., 2005, Gaudry et al., 2008). In the intracellular domain, there are several putative SRC Homology 3 domain (SH3) and several potential phosphorylation sites by serine/threonine kinases and tyrosine kinases (Moh et al., 2005).

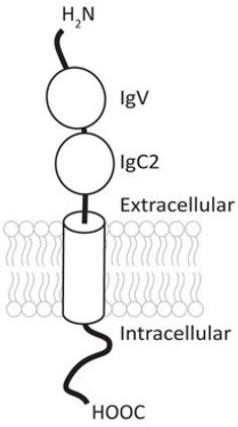


Figure 9. Predicted GlialCAM structure. Schematic representation of GlialCAM at the plasma membrane. The identified secondary structure consists of two Ig-like domains in the extracellular region (IgV and IgC2), a transmembrane domain and a cytoplasmic C-terminal tail. Extracted from (Passchier et al., 2024).

In first studies, GlialCAM was considered a tumour suppressor protein, as it was found that in hepatocellular carcinoma, GlialCAM protein expression levels were lower. They also reported that GlialCAM acted as an antiproliferative agent *in vitro* (Moh et al., 2005). Posterior studies determined that its expression was higher in the nervous system than in the liver, which suggested a predominant physiological role for GlialCAM in the brain (Spiegel et al., 2006). Along with these findings, it was observed that *GlialCAM* $^{-/-}$ mice do not show an increased hepatic tumour frequency (Hoegg-Beiler et al., 2014).

In the CNS, GlialCAM is mostly found in astrocytes and oligodendrocytes. Specifically, it is found in glia-enriched white matter, ependymal cells in ventricular zones, astrocytic junctions and the astrocyte-blood vessel contact areas (Favre-Kontula et al., 2008). Mice studies revealed that GlialCAM expression correlates with Myelin Basic Protein (MBP) expression in the developing brain, suggesting that this protein could be involved in the myelination process. Compatible with this hypothesis, it is known that GlialCAM is present along different stages of oligodendrocyte differentiation *in vitro*. In humans, GlialCAM expression levels are increased during the first three years of life. Then, they decrease until they stabilize approximately at five years (Bugiani et al., 2017).

A recent published article linked autoantibodies to GlialCAM with the pathology of multiple sclerosis. In this work, molecular mimicry between the Epstein-Barr virus nuclear antigen 1 (EBNA1) and GlialCAM is unveiled. The presence of cross-reactive antibodies recognizing GlialCAM is hypothesized to be involved in the development of the disease (Lanz et al., 2022).

Immunofluorescence assays and EM stainings on brain slices revealed that GlialCAM and MLC1 colocalize in the astrocytic junctions (López-Hernández et al., 2011a) (Figure 10). The stability of this colocalization depends on the interaction with the actin cytoskeleton (Moh et al., 2005, Duarri et al., 2011). Remarkably, both MLC1 and GlialCAM are associated with caveolae, which are lipid rafts implicated in the compartmentalization of intracellular signal transduction components, transport processes, endocytosis, and transcytosis (Ambrosini et al., 2008, Lanciotti et al., 2010, Moh et al., 2005, Sowa, 2012).

Slit-TEV assays revealed that GlialCAM can homo-oligomerize in *cis*- through its extracellular domain (López-Hernández et al., 2011b). Such oligomerization is important for the interaction between the cytoplasmic tail of GlialCAM and the actin cytoskeleton, which is essential for the localization of GlialCAM at cell junctions (Capdevila-Nortes et al., 2015). At the same time, the extracellular domain also interacts in *trans*- with GlialCAM molecules from neighbouring cells (Capdevila-Nortes et al., 2015, Hoegg-

Beiler et al., 2014). Furthermore, biochemical studies characterizing patient-derived mutations in GlialCAM determined that for the formation of *cis*- interactions, *trans*-interactions are needed in the first place (López-Hernández et al., 2011b).

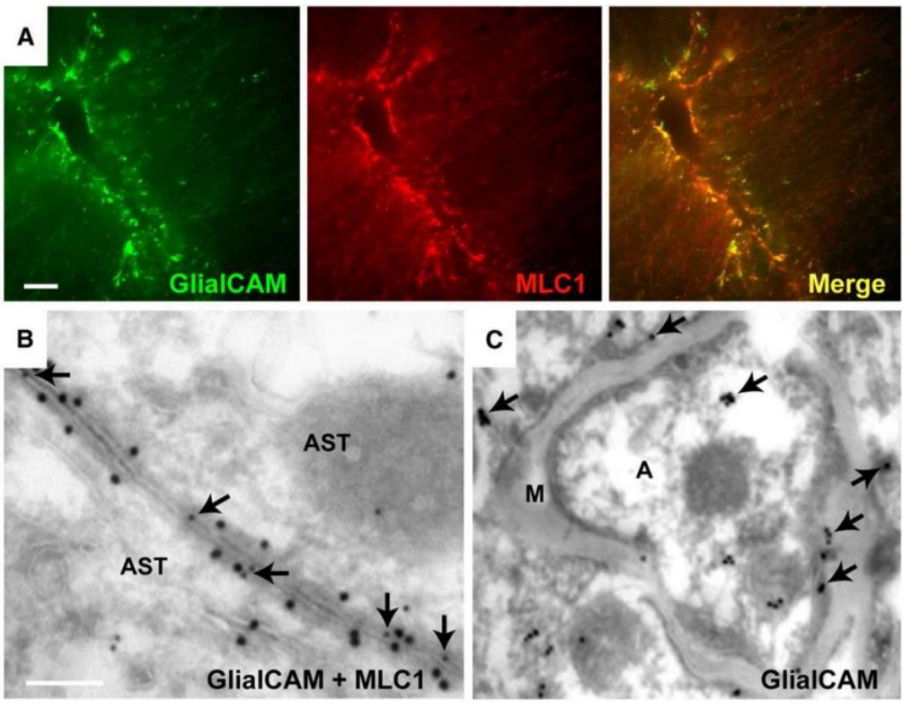


Figure 10. GlialCAM and MLC1 localization in brain tissue. A. Immunofluorescence in human cerebellum tissue. GlialCAM (green, *left*) and MLC1 (red, *centre*) co-localize (yellow, *right*) at the endfeet of perivascular astrocytes. **B.** Double Immunogold EM staining in human brain tissue. GlialCAM and MLC1 co-localize at astrocyte-astrocyte junctions. **C.** Post-embedding Immunogold EM stainings. GlialCAM is localized inside the axons, in contact regions between myelin and axons, and in cells surrounding myelin. Scale bar indicates 50 µm in A and 500 nm in B and C. Modified from (López-Hernández et al., 2011a).

AST: astrocyte, A: axon, M: myelin.

1.3.3.1 STRUCTURE AND FUNCTION OF GLIALCAM PROTEIN

GlialCAM as MLC1 β-subunit

Mass spectrometry (MS) studies using three different antibodies against the N-terminal end of MLC1 on solubilized membranes from brain tissue identified GlialCAM as the second most abundant protein, right behind MLC1 (López-Hernández et al., 2011a). Further studies of MLC1 and GlialCAM interaction, including co-immunoprecipitation, Fluorescence Resonance Energy Transfer (FRET), Bioluminiscence Resonance Energy Transfer (BRET), and Split-TEV confirmed the direct interaction between both proteins. *In vitro* experiments done in HeLa cells showed that the co-expression of MLC1 and

GlialCAM affected the subcellular localization of MLC1. While MLC1 is expressed in the plasma membrane and the ER, the presence of GlialCAM shifted its localization towards an enrichment at cell-cell junctions (Figure 11) (López-Hernández et al., 2011b, López-Hernández et al., 2011a).

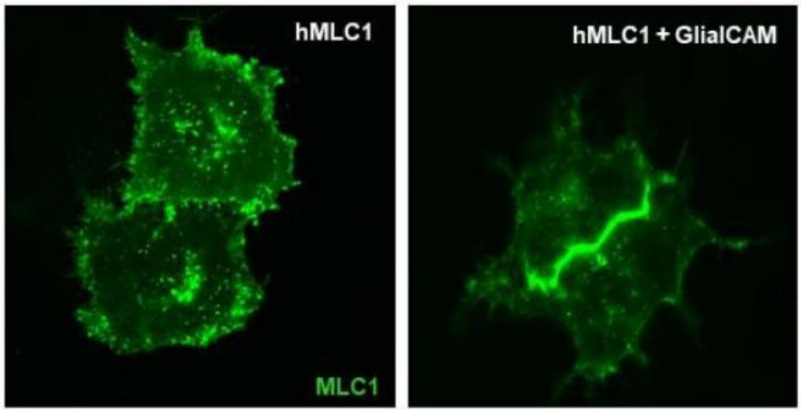


Figure 11. MLC1 changes its localization when co-expressed with GlialCAM. Immunofluorescence assays in transfected HeLa cells. MLC1 is expressed diffusely in the plasma membrane (left), but the presence of GlialCAM targets MLC1 to cell junctions (right). Scale bar: 20 μm. Modified from (López-Hernández et al., 2011b).

Even though the exact function of GlialCAM remains unknown, primary cultured astrocytes identified GlialCAM to act as a β-subunit of MLC1, meaning that GlialCAM is required for the proper traffic of MLC1 from the ER to cell junctions (López-Hernández et al., 2011b). In other words, GlialCAM might act as a chaperone for MLC1 (Capdevila-Nortes et al., 2013). In accordance to this, *Glialcam*$^{-/-}$ mice present lower levels of MLC1 protein and it is mislocalized in Bergmann glia and perivascular astrocytes (Bugiani et al., 2017, Hoegg-Beiler et al., 2014).

GlialCAM subcellular location is independent of MLC1, however, when both proteins are expressed *in vitro*, MLC1 increases the rate of GlialCAM enrichment at cell-cell junctions, suggesting that MLC1 would be needed for the proper localization of GlialCAM at cell junctions (López-Hernández et al., 2011b). *In vivo*, the absence of MLC1 leads to GlialCAM mislocalization, especially in Bergmann glia and perivascular astrocytes (Dubey et al., 2015, Hoegg-Beiler et al., 2014).

In primary astrocyte cultures derived from *Mlc1*$^{-/-}$ animals, GlialCAM is perfectly localized in physiological conditions. Surprisingly, in depolarizing conditions, where there is a high concentration of potassium, GlialCAM is internalized. Studies in MLC patients and in *mlc1* KO zebrafish also exposed GlialCAM mislocalization (Sirisi et al., 2014) concluding that MLC1 and GlialCAM might cooperate for proper expression and localization.

GlialCAM as ClC-2 auxiliary subunit

Considering that GlialCAM is expressed in other cell types where MLC1 is not present, such as oligodendrocytes, it is suggested that GlialCAM might have other functions independently of MLC1. That being the case, proteomic and immunoprecipitation assays on solubilized mouse brain membranes were performed with the objective to find GlialCAM-associated proteins. The chloride channel ClC-2 was identified (Barrallo-Gimeno et al., 2015, Jeworutzki et al., 2012). Through Split-TEV and co-immunoprecipitation assays, the direct interaction between GlialCAM and ClC-2 was validated. Immunofluorescence assays in HeLa cells showed that, similarly to the observation made with MLC1, ClC-2 is located at cell-cell junctions only when it is co-expressed with GlialCAM (Figure 12). In addition, GlialCAM modifies the biophysical properties of the channel by increasing ClC-2- mediated current amplitude and altering its rectification properties. So, in the presence of GlialCAM, ClC-2 is open also at positive membrane potentials (Jeworutzki et al., 2012).

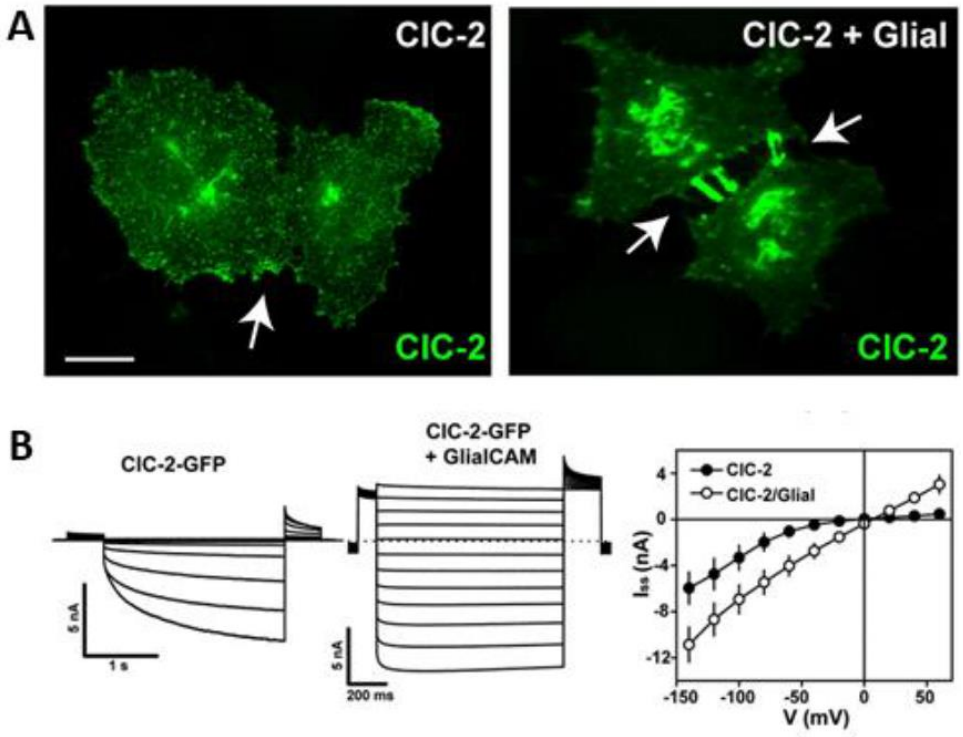


Figure 12. GlialCAM alters ClC-2 localization and function. A. Immunofluorescence assays in transiently transfected HeLa cells. ClC-2 is found enriched at cell-cell junctions in presence of GlialCAM. Scale bar: 10 µm. Adapted from (Jeworutzki et al., 2012). **B.** Functional alterations in the electrophysiological properties of the ClC-2 channel in HEK293-T cells when co-expressed with GlialCAM: the presence of GlialCAM produces the opening of the channel at positive voltages. Modified from (Jeworutzki et al., 2012).

The analysis of ClC-2 activity in presence of MLC1 discovered that such protein does not affect the channel functional properties (López-Hernández et al., 2011a), even so, co-immunoprecipitation studies and Proximity Ligation Assays (PLA) implies that MLC1 might be essential for the interaction of GlialCAM with ClC-2 (Sirisi et al., 2017).

Mlc1$^{-/-}$ and *GlialCAM*$^{-/-}$ animals exhibit ClC-2 mislocalized in Bergmann glia, perivascular astrocytes, and oligodendrocytes (Bugiani et al., 2017, Hoegg-Beiler et al., 2014). Although MLC1 is not expressed in oligodendrocytes, ClC-2 mislocalization in *Mlc1*$^{-/-}$ implies that MLC1/GlialCAM could regulate ClC-2 activity (Hoegg-Beiler et al., 2014).

1.3.3.2 MUTATIONS IN *GLIALCAM* GENE

As previously stated, most *GLIALCAM* mutations causing MLC are missense mutations, and they are located at GlialCAM extracellular region. One exception is p.W263X, which is a nonsense mutation that introduces an early stop codon in the transmembrane domain. Another exception is p.L23H that affects the signal peptide.

MLC-causing missense mutations located at *GLIALCAM* can be either recessive or dominant. Most of them do not compromise protein expression, but present trafficking defects (Figure 13B). Moreover, GlialCAM mutations do not affect the interaction with MLC1 (López-Hernández et al., 2011b). Dominant mutations are located in the IgV domain only, while recessive mutations can be found anywhere from the extracellular region (Figure 13A).

Most *GLIALCAM* dominant mutations show an impaired ability to homo-oligomerize, to locate at cell-cell junctions, and to target either MLC1 or ClC-2 to cell-cell junctions. Despite this phenotype, the interaction of these proteins with GlialCAM is not affected. However, there are some exceptions among dominant mutations. For instance, mutation p.D218N encodes for a protein that does not enrich at cell-cell junctions like the WT protein. Also, p.K135Del mutant does not exhibit any apparent defects (López-Hernández et al., 2011b, Arnedo et al., 2014b).

Regarding *GLIALCAM* recessive mutations, some of them lead to protein variants that display defects in homo-oligomerization and junctional targeting, including GlialCAM with itself as well as associated with MLC1/ClC-2. Nonetheless, there are some described recessive mutants that do not display any of these effects. These mutants are p.R73W, p.P148S, p.S196Y, and p.D211N (Arnedo et al., 2014a, Arnedo et al., 2014b, López-Hernández et al., 2011b).

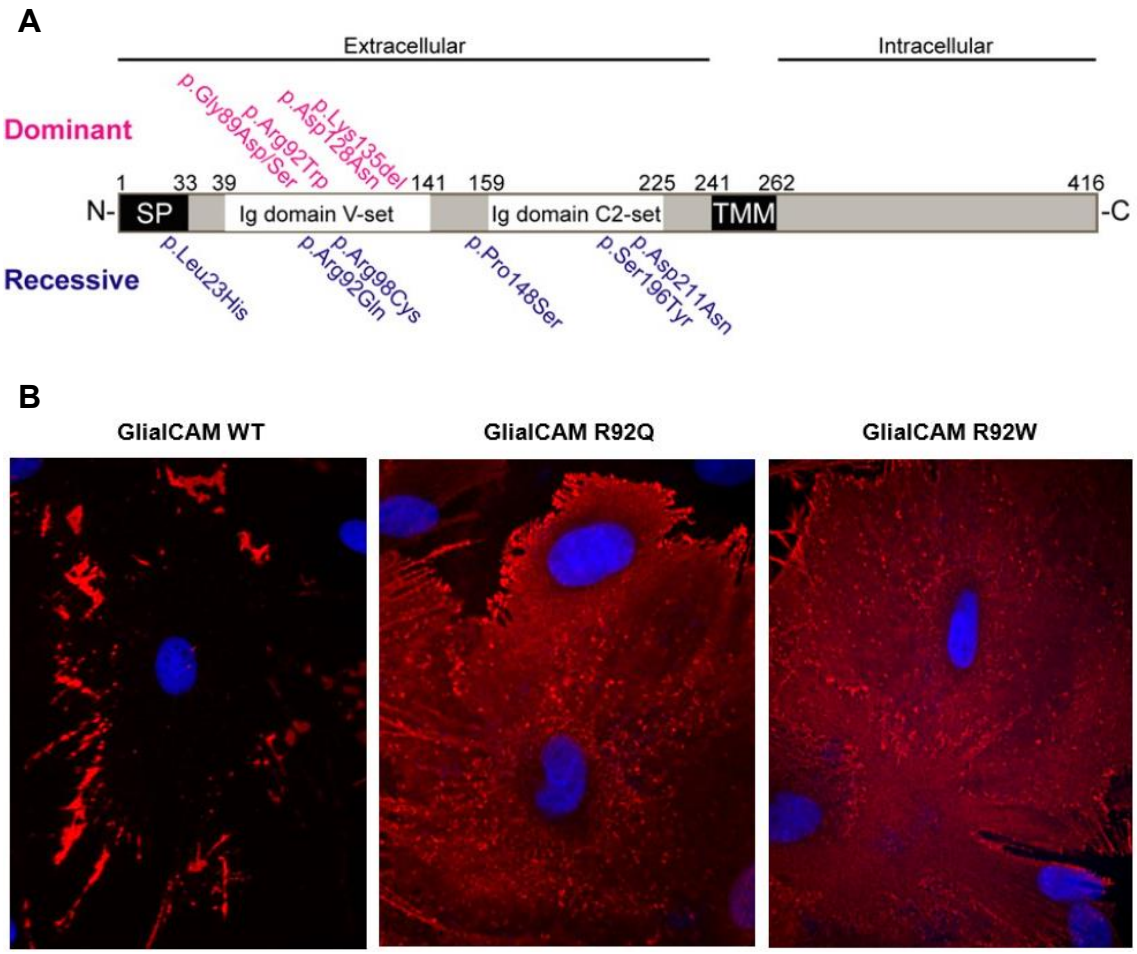


Figure 13. GlialCAM mutations affect protein localization. **A.** Schematic representation of all missense mutations identified for *GLIALCAM,* Dominant mutations affect solely IgV domain are marked in pink. Recessive mutations are written in blue. SP: signal peptide, TMM: transmembrane domain. **B.** Immunofluorescence assays in primary cultured astrocytes infected with adenoviruses expressing GlialCAM WT (left), R92Q recessive mutant (centre) and, R92W dominant mutant (right). GlialCAM is mislocalized in both mutants. Scale bar: 20 µm. Modified from (López-Hernández et al., 2011b).

In a previous work published by our research group, the molecular basis of the dominant or recessive behaviour of GlialCAM variants was investigated using a combination of biochemistry methods together with computer docking (Elorza-Vidal et al., 2020). In this study, *GLIALCAM* variants affecting the IgV domain were classified into three groups. First, mutations found in MLC2A patients that show a reduced ability to interact with GlialCAM WT. Second, a subset of mutants (Dominant, D1) found in MLC2B patients with a reduced ability to interact with GlialCAM WT. Lastly, a subset of mutants (Dominant, D2) found in MLC2B patients which display a normal ability to interact with GlialCAM WT (Figure 14).

From this study, it was shown that the reason why some disease-causing GlialCAM mutations are dominant is that they are located at GlialCAM IgV domain interacting surfaces, where they disrupt either *cis*- (D1 mutations) or *trans*- (D2 mutations) interactions necessary for protein function. *Cis* dimerization is achieved by interactions between two opposing beta-strands of the IgV domain and *trans* interactions occur between salient loops of both IgV domains (Elorza-Vidal et al., 2020).

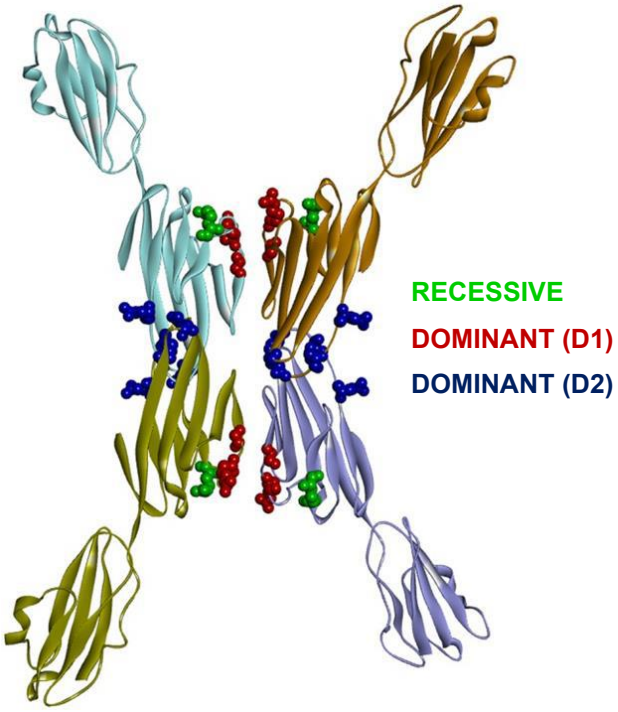


Figure 14. Structural model proposed for GlialCAM homodimers. GlialCAM homodimers would form cis- and trans- interactions through different surfaces of its IgV domain. Residues mutated in MLC2A patients (recessive) are shown in green, D1 residues mutated in MLC2B patients are shown in red and D2 residues mutated in MLC2B patients are shown in blue. Extracted from (Elorza-Vidal et al., 2020).

Other biochemical studies indicated that localization defects in recessive GlialCAM mutant variants are rescued by WT protein. But this defect is not solved in dominant mutations (López-Hernández et al., 2011a, Elorza-Vidal et al., 2020). This *in vitro* observation was validated *in vivo* by the generation of a *GLIALCAM knock-in* (KI) mouse model bearing a dominant mutation. The heterozygous mice were found to also develop a phenotype (Hoegg-Beiler et al., 2014).

An interesting observation is that even though GlialCAM is necessary for correct MLC1 expression and protein levels (Jeworutzki et al., 2012, López-Hernández et al., 2011b, López-Hernández et al., 2011a), the overexpression of either dominant or recessive

GlialCAM mutations is sufficient for proper MLC1 expression levels. Nevertheless, MLC1 is not targeted to cell-cell junctions and VRAC activity is altered (Capdevila-Nortes et al., 2013). Therefore, the problem with these mutants might be at a functional level.

Although with the study of most *GLIALCAM* mutations we were able to gain knowledge on the pathogenicity of these mutants, there was still a group of *GLIALCAM* recessive mutants that display a particular phenotype. These mutations were p.S196Y and p.D211N which are located in the IgC2 domain. It was observed that in both cases, the resulting mutated protein was resistant to depolarization-induced internalization in primary cultured astrocytes derived from *Mlc1*$^{-/-}$ animals. While the incubation with potassium-enriched medium led to GlialCAM WT internalization in these astrocytes, the mutant proteins stayed enriched at cell-cell junctions (Sirisi et al., 2014). Since this phenomenon was not fully understood, a mechanism to explain this phenotype could not been provided.

1.3.4 PATHOPHYSIOLOGY OF MLC DISEASE

The main characteristic observed in MLC patients is an enlarged brain with water accumulation. This lead to hypothesize that MLC1 could function as a water or ion channel, and its defect would cause osmotic alterations. MLC1 sequence identity analysis done by BLAST algorithm revealed that the voltage gated potassium channel Kv1.1 alpha subunit (KCNA1) was the protein with a highest sequence identity. However, the amino acid identity was less than 20% (Teijido et al., 2004). In agreement with this hypothesis, MLC patients may present epileptic seizures, which is a common feature in ion channel diseases, but not in leukodystrophies (Hamilton et al., 2018). Unfortunately, voltage-clamp measurements in *Xenopus laevis* oocytes did not detect any changes in the conductance, neither did patch clamp experiments in HEK293T or HeLa transfected cells. Different pulse protocols and durations were applied without detecting conductivity alterations, neither with the presence of GlialCAM (Estévez et al., 2018).

For the study of MLC, different animal models have been created that have helped understand the function of the proteins implied in the disease. The creation of a *Mlc1* knockout (KO) mouse model was found useful for studying early stages of MLC, as the animals display increased water content in the brain, progressive vacuole formation in the white matter, and morphological changes in perivascular astrocytes (Dubey et al., 2015, Hoegg-Beiler et al., 2014). Similarly, the *GlialCAM* KO mice (Sugio et al., 2017, Dubey et al., 2015, Bugiani et al., 2017, Hoegg-Beiler et al., 2014) and *GlialCAM* KO

zebrafish (Sirisi et al., 2014, Pérez-Rius et al., 2019) present an increased brain water content and cerebellar white matter vacuolation, although no cysts are perceived. In primary astrocyte cultures, the lack of MLC1 causes the appearance of intracellular vacuoles (Duarri et al., 2011, Sirisi et al., 2014). This discovery made reanalyse patient brain biopsies, because until then vacuole formation was believed to be limited to the outer myelin layers (van der Knaap et al., 1996). With this, it was confirmed that vacuole formation also occurred in perivascular astrocytes (Duarri et al., 2011).

Another study done in *Mlc1*$^{-/-}$ mice manifested spontaneous epileptic activity, together with an altered extracellular potassium dynamics and abnormal neuronal network activity, as they present an epileptiform brain activity and a diminished seizure threshold (Dubey et al., 2018). These results are consistent with the hypothesis of MLC1 being functionally linked to water and ion homeostasis. In this regard, extensive proteomic analysis identified several transporters and ion channels interacting with MLC1 and/or GlialCAM (Alonso-Gardón et al., 2021). Some interactors found were the following: the chloride channel ClC-2, the gap junction protein Cx43, the glutamate transporter EAAT1/2, the alpha2 and beta2 subunits of the sodium/potassium ATPase, the sodium bicarbonate transporter NBCe1, the glucose transporter GLUT1, and the sodium calcium exchanger.

Although not appearing as MLC interacting partners, other proteins have also been associated to it. This is the case of volume regulated anion channel (VRAC) or the calcium-permeable channel TRPV4. Primary cultured astrocytes derived from *Mlc*$^{-/-}$ animals present a decrease in VRAC activity and the associated physiological cell volume regulation compensatory mechanism (RVD) (Capdevila-Nortes et al., 2013, Dubey et al., 2015). Contrarily, MLC1 overexpression leads to an increased VRAC activity in primary cultured astrocytes. Still, the interaction between VRAC and MLC1 is yet to be fully understood. Previous studies from the group demonstrated that MLC1 and LRRC8A, which is the VRAC's main subunit, do not co-localize nor directly interact, suggesting an indirect regulation for the channel's activity (Elorza-Vidal et al., 2018).

Lastly, MLC1 overexpression leads to the degradation of Epidermal Growth Factor Receptor (EGFR). Alternatively, the absence of MLC1 in astrocytes increases ERK1/2 phosphorylation (Elorza-Vidal et al., 2018). These findings suggest that MLC1 might participate in intracellular signalling pathways.

Perivascular astrocytes, where MLC1 expression is higher, are essential for the maintenance of the BBB. In *Mlc1* KO mice, the perivascular astrocytic processes' molecular maturation and organization is affected, although the BBB integrity and the

organization of the endothelial network are not altered. This finding leads to hypothesize that MLC1 might also play a role in contractile maturation of vascular smooth muscle cells, arterial perfusion, and neurovascular coupling (Gilbert et al., 2021). Thus, not only the disturbance of ion channels and transporters activity contributes to the increased water content and seizures observed in MLC patients, but it might be that the disruption of gliovascular unit by itself might contribute as well.

Seeing that MLC1 is associated to many different proteins, one of the working hypotheses of our research group is that MLC1 would be a signalling regulator through an unknown mechanism. We also hypothesise that MLC1 could act as a tetraspanin-like molecule. Even though the homology of MLC1 with tetraspanins is low (around 20%), both proteins contain 8 transmembrane domains and form oligomers. Moreover, tetraspanins are known to associate laterally at the plasma membrane with receptors, and also to regulate trafficking or processing of their protein partners (Barreiro et al., 2008, Little et al., 2004, Yang et al., 2020).

2. G-PROTEIN-COUPLED RECEPTORS IN THE CENTRAL NERVOUS SYSTEM

2.1 GENERALITIES OF G-PROTEIN-COUPLED RECEPTORS

Membrane proteins provide communication between the external and internal compartments of the cell. GPCRs constitute a diverse family of membrane proteins involved in cell signalling processes. Its structure consists of seven α-helices that cross the plasma membrane connected through three intracellular and three extracellular loops. The N-terminus is extracellular and the C-terminus is intracellular (Figure 15) (Venkatakrishnan et al., 2013).

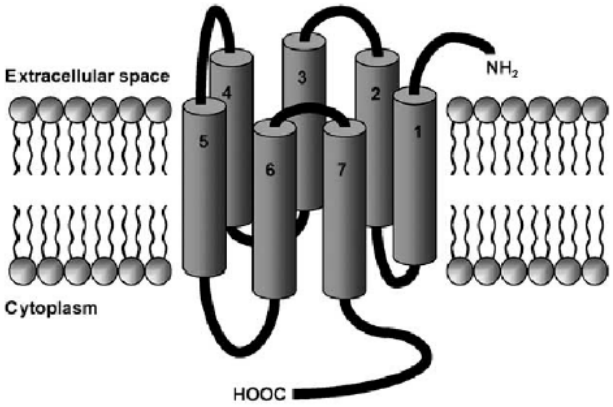


Figure 15. GPCR topology. Schematic representation of the structure of GPCRs. They are formed by 7 TM domains, an extracellular N-terminal domain and a C-terminal cytoplasmic tail. Extracted from (Tikhonova and Costanzi, 2009).

GPCRs are a large superfamily of membrane proteins that perceive extracellular stimuli and, as a result, initiate signalling transduction cascades (Tikhonova and Costanzi, 2009). A wide variety of ligands have been identified to interact with GPCRs, including endogenous and exogenous molecules, such as neurotransmitters, metals, odorant molecules, peptides, fatty acids, photons, ions, small organic molecules, or entire proteins, among others. Through the binding with a GPCR, these molecules can control a vast array of physiological processes, for instance, cell differentiation, changes in blood pressure, immune responses, tumour progression or cell death to mention a few (Rosenbaum et al., 2009). The binding of the ligand with its receptor can occur by three different mechanisms: through a pocket of TM domains, through extracellular loops, or through the N-terminal domain (Jacoby et al., 2006).

The physiological relevance of GPCRs can be noticed by the fact that 800 genes in the human genome, representing 2% of the total human genome, encode for such proteins (Fredriksson et al., 2003, Jacoby et al., 2006). Approximately 100 of these genes do not have an identified ligand that activates the encoded GPCR. Such proteins are called orphan GPCRs (Lagerström and Schiöth, 2008). Seeing that GPCRs participate in a wide variety of physiological mechanisms, it is not surprising that they are one of the most pursued targets for drug development (Fredriksson et al., 2003). In fact, one third of the drugs medically prescribed act on GPCRs (Sriram and Insel, 2018).

For the signal transduction mediated by GPCRs, three main components are needed: the GPCR itself, a heterotrimeric G protein, and an effector protein that will trigger a cellular response through second messengers. The G protein is formed by the subunits α, β and γ and it can be found in two states: inactive or active. When the G protein is inactive, it is bound to GDP, whereas its active state occurs when it binds to GTP.

For the signalling to occur, first, the cytoplasmic tail of the GPCR has to interact with the G protein. The active G protein is unstable and, as a result, the trimer is split into the α subunit and the βγ dimer. This dimer can then interact with the effector protein and second messengers will be recruited, which in its turn will trigger a cellular response. The signalling ends with the hydrolysis of GTP to GDP leading to the inactive state of the G protein with the re-association of the three subunits (Milligan and Kostenis, 2006).

There are many effector proteins that act downstream of the GPCR and that will activate different signalling cascades (Figure 16). One of the most well characterised effectors by Gq protein is Phospholipase C (PLC) which breaks phosphatidylinositol 4,5-biphosphate (PIP_2) into inositol triphosphate (IP_3) and diacylglycerol (DAG) by hydrolysis. IP3 can bind to calcium channels located at intracellular compartments, like the ER, and activates them. As a result, calcium is released to the cytosol. DAG, on the other hand, can interact with protein kinases (Eglen et al., 2007, Hendriks-Balk et al., 2008).

Heterotrimeric G proteins can be classified into four families depending on the structural and functional homology of the Gα subunit. These families are: G_s, $G_{i/0}$, $G_{q/11}$, and $G_{12/13}$ (Hollmann et al., 2005, Offermanns, 2003). Protein G_s stimulates adenylyl cyclase (AC), which produces cAMP at an intracellular level and activates Protein Kinase A (PKA). Protein $G_{q/11}$ increases intracellular calcium concentration. Protein $G_{12/13}$ activates small GTPases such as RhoA and Phospholipase D (PLD) that regulate cell shape and motility by altering the cytoskeleton. Lastly, protein $G_{i/0}$ inhibits AC and regulates potassium channels (Milligan and Kostenis, 2006, Yudin and Rohacs, 2018).

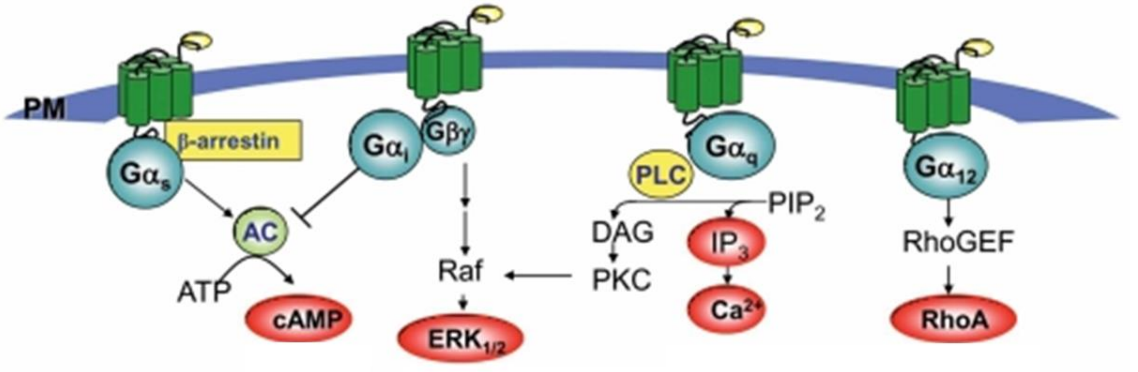


Figure 16. Schematic representation of the main signalling pathways activated by GPCRs. Upon interaction with stimulus, $G_{\alpha s}$ activates AC, while $G_{\alpha s}$ inhibits it. $G_{\beta\gamma}$ activates ERK pathways. $G_{\alpha q}$ activates PLC leading to increased intracellular calcium concentrations and the activation of PKC. $G_{\alpha 12}$ activates GTPases like RhoA. Modified from (Cheng et al., 2010).

Besides the interaction with G proteins, GPCR are found to also interact with β-arrestin. The binding of the ligand to the receptor can induce the recruitment of β-arrestin in a G-protein- independent manner. Then, β-arrestin can activate the signaling pathway Mitogen-activated Protein Kinase (MAPK) pathway, which leads to the activation by phosphorylation of proteins like ERK1/2, RAF-1, or MEK1. This pathway is key for the cell cycle control and the regulation of transcription and apoptosis (DeWire et al., 2007).

As previously mentioned, the ligand that interacts with the GPCR determines the cellular response. Molecules that activate GPCRs can be divided into four groups: agonists, partial agonists, inverse agonists, and antagonists. Agonists activate the receptor after its binding. Partial agonists bind and activate the receptor as well, but they do not elicit a maximum response. Inverse agonists bind to the receptor and decrease its constitutive activity. Lastly, antagonists bind to the receptor and prevent agonist binding (Park et al., 2008).

A key process in GPCR signaling is GPCR desensitization, which aims at avoiding potential negative effects resulted from an excessive cell response. Two types of desensitization have been described: homologous and heterologous. The first one refers to a loss of response solely to agonists that act at a particular GPCR subtype. In this case, once the GPCR is phosphorylated by a GRK, β-arrestin is recruited. As a result, this protein will interdict further G protein activation despite the continued activation of the receptor by agonist by preventing the receptor from exchanging GTP for GDP on the G protein α-subunit (Gainetdinov et al., 2004). On the other hand, heterologous desensitization refers to a more generalized effect involving the simultaneous loss of

agonist responsiveness at multiple GPCR subtypes, even in the absence of agonist. PKA and PKC are responsible of this ligand-independent desensitization (Kelly et al., 2008).

Additionally, β-arrestin activity induces the GPCR internalization. Phosphorylated β-arrestin is recognized by clathrin-coated vesicle formation units leading to receptor endocytosis (Claing et al., 2002). Four types of β-arrestin have been described (β-arrestin 1-4): β-arrestin 1 and 4 are expressed in the retina and they are linked to phototransduction, whereas β-arrestin 2 and 3 are ubiquitous.

Finally, GPCRs can be re-sensitized upon dephosphorylation by protein phosphates and protein turnover to the plasma membrane. However, they can also be targeted for degradation (Métayé et al., 2005).

2.2 GPCR CLASSIFICATION

The first GPCR classification was proposed by Kolakowski and it categorized these proteins into seven groups designated with letters from A to F, and O (Davies et al., 2007, Kolakowski, 1994). Later, it was updated to six classes keeping the nomenclature A-F based on amino acid sequence and functional similarities (Foord et al., 2005). In this system, families A-C comprise the vast majority of GPCRs (Figure 17). Another classification was proposed based on phylogenetics that divided GPCRs into five families named GRAFS (G for Glutamate, R for Rhodopsin, A for Adhesion, F for Frizzled, S for Secretin) (Schiöth and Fredriksson, 2005). However, this system is infrequently used.

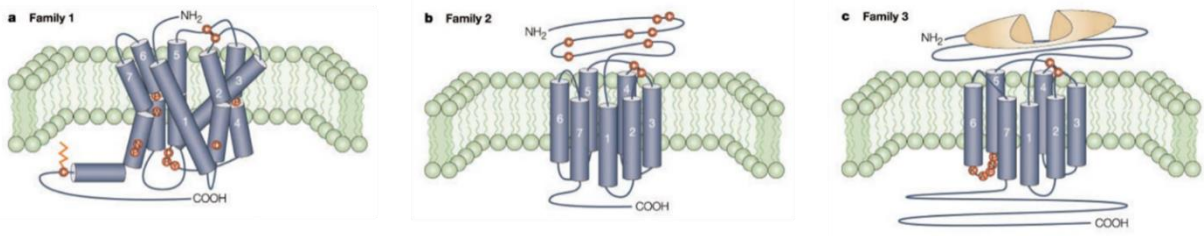


Figure 17. Schematic representation of the main GPCR families. Family 1 corresponds to class A GPCRs; Family 2, class B and Family 3 equals class C. In red, conserved amino acids across family members. In ochre, Venus Flytrap module is marked in class C receptors. Modified from (George et al., 2002).

Class A: Rhodopsin-like receptors

Rhodopsin-like GPCRs constitute the largest GPCR family, comprising around 90% of all GPCRs including odorant receptors and small ligand receptors. Crystallography structure of rhodopsin receptor exposed alterations in the helicoidal domain due to the presence of proline residues that results in TM tilting (Palczewski et al., 2000). These receptors are characterized by the formation of a disulfide bridge connecting the extracellular loops 1 and 2. Furthermore, most of these proteins contain a palmitoylated cysteine residue in the C-terminal end that anchors the receptor to the plasma membrane.

Class B: Secretin family receptors

Secretin family GPCRs include 15 receptors for peptide hormones (Alexander et al., 2013). These receptors contain a large N-terminal extracellular domain with various cysteine residues that form disulfide bonds (Ulrich et al., 1998). The ligands for these GPCRs are hormones of small molecular weight, such as glucagon or calcitonin, that bind the N-terminal ectodomain (Jacoby et al., 2006).

Class C: Metabotropic glutamate receptors

In this family, metabotropic glutamate receptors (mGluRs), $GABA_B$ receptors, Ca^{2+}-sensing receptors, sweet and amino acid taste receptors, pheromone receptors, and several orphan receptors are found (Bräuner-Osborne et al., 2007).

These receptors are composed of an exceptionally large extracellular domain that contains a Venus flytrap (VFT) module and a cysteine rich domain (CRD), except in $GABA_B$ receptor. The orthosteric binding sites reside in the VFT domain, though allosteric sites are located elsewhere. Given the fact that the extracellular domain is larger than in other GPCRs classes, these receptors present more potential allosteric sites than others. The C-terminal tail is highly variable and plays a role in scaffolding and signaling protein coupling. Another unique characteristic of class C GPCRs is their capacity for dimerization, either as homodimers or heterodimers. The allosteric interaction between different dimer domains results in a particularly complicated activation process (Chun et al., 2012).

Finally, class D and E are not found in vertebrates. Class F comprises Frizzled family receptors, which is a minor class involved in the embryonic development, and cell division and polarity (Foord et al., 2002).

2.3 THE ORPHAN RECEPTORS GPR37 AND GPR37L1

A study done by our group identified the members of the GlialCAM interactome (Alonso-Gardón et al., 2021) where both GPR37 and GPR37L1 were discovered to interact with GlialCAM. GPR37 and GPR37L1 proteins are members of class A rhodopsin-like family of GPCRs. They are both considered orphan GPCRs and they are highly expressed in the CNS.

GPR37 is mainly expressed in the cerebellum, corpus callosum, medulla, putamen, caudate nucleus, substantia nigra, and the hippocampus (Donohue et al., 1998, Marazziti et al., 1998, Takahashi and Imai, 2003, Zeng et al., 1997). Regarding cell types, GPR37 is mostly expressed in oligodendrocytes and in a subset of neurons in substantia nigra (Imai et al., 2001). Neuroprotectin D1 (NPD1) has been identified as a ligand for GPR37 and mediates macrophages phagocytosis (Bang et al., 2018).

Otherwise, GPR37L1 is exclusively expressed in glial cells, particularly in Bergmann glia astrocytes in the cerebellum, as well as immature oligodendrocytes (Marazziti et al., 2013). Recently, the proresolving lipid mediator maresin 1 (MaR1) has been identified as a ligand of GPR37L1 (Bang et al., 2024).

MLC1 and GPR37L1 are both expressed in astrocytes and it was established that they co-localize in primary cultured astrocytes. Moreover, both proteins were in proximity in the same cells, as assessed by PLA. Split-TEV assays determined the capacity of GPR37 to directly interact with MLC1 and GlialCAM. By BRET analysis, GPR37L1 was found to also be able to interact with MLC1 and GlialCAM in HEK293T cells (Alonso-Gardón et al., 2021).

Gpr37L1 expression levels gradually increase in Bergmann glia from neonatal stages until adulthood. *Gpr37*$^{-/-}$ mice present precocious cerebellar development, along with premature Bergmann glia and Purkinje cell maturation (Marazziti et al., 2013). In those animals, both MLC1 and GlialCAM expressions are found upregulated without altering MLC1 localization (Alonso-Gardón et al., 2021).

Gpr37$^{-/-}$ mice studies showed that the protein is important for the regulation of oligodendrocyte differentiation and myelin formation as the animals suffer from precocious oligodendrocyte differentiation leading to hypermyelination of the CNS (Yang et al., 2016). No studies have been performed in *Gpr37* KO mice in relationship with MLC1 or GlialCAM.

2.4 GPRC5 FAMILY

2.4.1 GENERALITIES AND CLASSIFICATION

GPRC5 receptors belong to the heterotrimeric orphan class C GPCRs. They were firstly identified as the gene encoding for these proteins was found to be induced by retinoic acid in a context of cancer cells. For this reason, GPRC5 receptors were early named RAIG (Retinoic Acid Inducible Gene) (Cheng and Lotan, 1998). The TM domains from GPRC5 receptors are conserved from invertebrates to mammals. Four GPRC5 subtypes are found in mammals: GPRC5A (RAIG1), GPRC5B (RAIG2), GPRC5C (RAIG3) and GPRC5D (RAIG4) (Robbins et al., 2000).

Although Class C GPCRs typically possess a large N-terminal domain, GPRC5 receptors present an exceptionally short N-terminal domain. This suggests that the ligand binding function of GPRC5 receptors might have been lost during evolution (Kim et al., 2012).

The four members of GPRC5 family display a singular distribution pattern across different tissues (Table 2). GPRC5A is present in the lungs, where it acts as a tumor suppressor protein, and in the colon (Tao et al., 2007). In *Gprc5a* KO mice, NF-κB activation is enhanced in lung epithelial cells leading to a tumor promoting microenvironment (Deng et al., 2010). Furthermore, hypoxia was found to induced GPRC5A activity in cancer cells (Greenhough et al., 2018). GPRC5D can be found in skin and in pancreatic tissue. As for GPRC5D, its *in vivo* functions are still unknown (Inoue et al., 2004). GPRC5B and GPRC5C are expressed ubiquitously, but they are enriched in the CNS, particularly in the cerebellum, but also in placenta and adipose tissue (Tao et al., 2007). GPRC5B mRNA levels are particularly abundant in taste buds, olfactory bulb, and Purkinje cells. Its expression in the cerebellum is developmentally regulated and it continues to increase during aging (Sano et al., 2011). GPRC5C is located at the apical membrane of the proximal tubules and it is essential for kidney homeostasis. In *Gprc5c* KO mice, the acid-base equilibrium is disrupted. Compared to WT animals, their blood pH is lower and their urine pH is higher (Rajkumar et al., 2018).

Although the different families of GPRC5 receptors exhibit different locations and functions, all of them are found glycosylated, and they can be secreted in extracellular exosomes (Kwon et al., 2014).

GPRC5	Tissue distribution	Cellular localization	KO mice	Proposed function and signalling
GPRC5A	Lung, colon, urinary bladder	PM, ER, Golgi Extracellular vesicular exosomes	Lung tumours	Tumour suppressor gene, NF-κB and STAT3 signalling
GPRC5B	Brain, adipose tissue, kidney, mammary grand	PM, Golgi Extracellular vesicular exosomes	Defect in cerebellar motor learning, resistant to obesity, low milk production	SFK-signalling domain formation Pro-inflammatory effects Cardiac fibrosis, breast cancer
GPRC5C	Brain, kidney, liver	PM Extracellular vesicular exosomes	Regulation of hematopoietic system Lower blood pH and higher urine pH	Thromboembolism Kidney liver function
GPRC5D	Hair follicle, pancreas, B cell	PM Extracellular vesicular exosomes	?	Hard-keratinized structures Multiple myeloma antigen

Table 2. Main properties of GPRC5 family members. PM: plasma membrane, ER: endoplasmic reticulum. Adapted from (Hirabayashi and Kim 2020).

2.4.2 GPRC5B RECEPTOR

As previously stated, GPRC5B is ubiquitously expressed but cerebellum-enriched. The receptor participates in motor learning and spontaneous activity in front new environmental stimuli (Sano et al., 2011, Sano et al., 2018). It is also abundant in the adipose tissue where it is believed to play a role in obesity (Tekola-Ayele et al., 2019). At a subcellular level, GPRC5B is located at the plasma membrane, Golgi apparatus, and exosomes. In Madin-Darby canine kidney cells, it was observed that GPRC5B was mediating extracellular vesicle transport and tubule formation (Kwon et al., 2014).

GPRC5B is the only GPRC5 implicated in psychiatric conditions, such as attention deficit hyperactivity disorder (ADHD) (Albayrak et al., 2013), bipolar disorder, and severe depression (Tomita et al., 2013).

According to mice studies, GPRC5B levels are increased in the axons of Purkinje cells in the cerebellum. These neurons are the only output neurons from the cerebellar cortex and they make synaptic contacts with cells from the deep cerebellar nuclei early after birth (Garin and Escher, 2001). For this reason, GPRC5B was involved in the neural circuitry development and in cerebellum-mediated motor control. The absence of this receptor at the SNC or specifically in Purkinje cells leads to an abnormal inflammation in the distal region of Purkinje cells axons, without affecting the dendritic morphology. These axons in *Gprc5b*$^{-/-}$ mice do not develop synapse-associated structures: their mitochondria are dysfunctional and accumulate in the inflamed areas. Due to mitochondrial dysfunction, Reactive Oxygen Species (ROS) are found increased, which

increments ATP production provoking oxidative damage that leads to cellular dysfunction. In turn, synapse formation between Purkinje cells and deep cerebellar nuclei is impaired and this compromises long-term motor learning skills (Sano et al., 2018).

Regarding the adipose tissue, GPRC5B levels are remarkably high (Sano et al., 2011). Even though the exact function of GPRC5B is not yet fully understood, a strong correlation was identified between *GPRC5B* gene expression and Body Mass Index (BMI) in a genome-wide association study (Speliotes et al., 2010). *Gprc5b*$^{-/-}$ mice weighted less than WT animals being fed with a high-fat diet (HFD). In this study, WT animals displayed greater hepatic steatosis caused by lipid accumulation in the liver than KO mice. *Gprc5b*-deficient animals showed reduced insulin resistance, less adipose tissue inflammation as well as smaller adipocyte size. The production of inflammatory cytokines and macrophage infiltration were lowered in the KO. Metabolic parameters, such as circulating glucose and insulin, or leptin concentrations ameliorated in KO animals. All things considered, the *Gprc5b*$^{-/-}$ phenotype might be caused by an increased metabolic rate and enhanced thermogenesis processes (Kim et al., 2012).

At a molecular level, GPRC5B contains multiple sites of phosphorylation. MS analysis identified eight phosphorylated serine residues, two phosphorylated threonine residues at the C-terminal region, and three phosphorylated tyrosine residues (Tyr^{307}, Tyr^{330}, Tyr^{383}) with high conservation across evolution (Figure 18A). Studies *in vitro* demonstrated that the kinase responsible for tyrosine phosphorylation was Fyn, a kinase belonging to the Src family. For the phosphorylation to occur, GPRC5B recruits Fyn through its SH2 domain when it is active. Fyn activity regulates inflammatory responses through NF-κB signaling pathway (Figure 18B-C) (Kim et al., 2012).

In other cell types, such as cardiomyocytes or cardiac fibroblasts, GPRC5B activity has also been linked to NF-κB pathway, and it is suggested to be involved in some pathological conditions (von Samson-Himmelstjerna et al., 2019). Studies done in pancreatic β-cells related GRPC5B activity to TGF-β/Smad 3 pathway and JAK-STAT/IFN-γ pathways, which participate in fibrotic and inflammatory processes, respectively (Atanes et al., 2018).

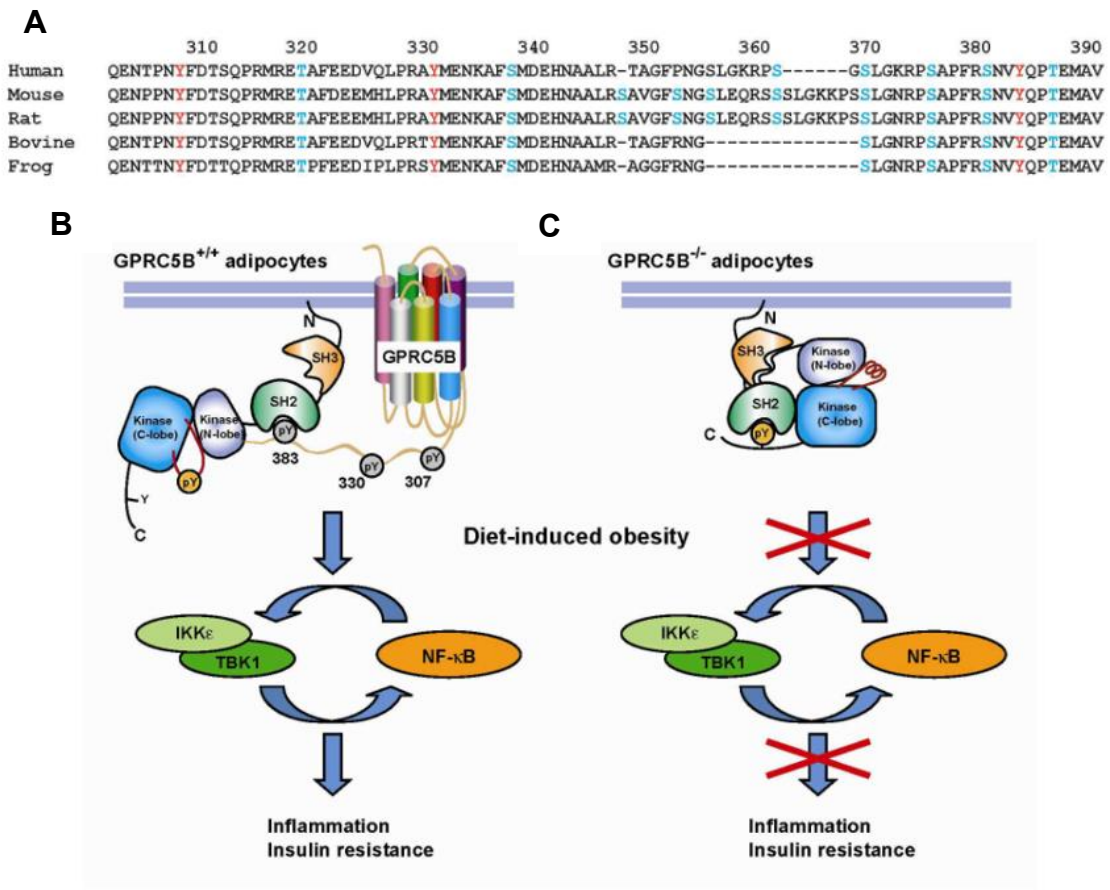


Figure 18. GPRC5B modulation of the inflammatory response in the adipose tissue. A. Sequence alignment for GPRC5B in different species. Conserved phosphorylation sites are concentrated in the C-terminal domain. Phosphotyrosine residues in red and phosphoserine and phosphothreonine residues in blue are highlighted. **B.** In a high-fat diet (HFD), GPRC5B and Fyn-SH2 domain interact causing a locally increased Fyn activity, which will stimulate the positive feedback stablished for NF-κB pathway. This will lead to adipose tissue inflammation and insulin resistance, which are obesity hallmarks. **C.** In the absence of GPRC5B, Fyn activity does not increase and NF-κB signalling pathway is maintained at low levels. Modified from (Kim et al., 2012).

2.4.2.1 GPRC5B AS THE THIRD MLC-CAUSING GENE

One of the latest studies published by our group identified GPRC5B in the brain GlialCAM interactome through an approach based on affinity purifications (Alonso-Gardón et al., 2021). More recently, whole exome sequencing (WES) on genomic DNA from three patients with an MRI-based diagnosis of MLC not showing pathogenic variants in *MLC1* or *GLIALCAM* identified GPRC5B as a disease-causing gene (Passchier et al., 2023). These patients harbored two different duplications arose *de novo* in the fourth transmembrane domain. The variants are p.Ile176dup and p.Ala177dup.

Lymphoblasts from MLC patients with *GPRC5B* variants display a disrupted RVD following swelling induced by hypotonic shock, in a similar manner than lymphoblasts derived from *MLC1* patients. Western blot analysis revealed increased GPRC5B levels in patient lymphoblasts compared to controls, however, this increment was not accompanied by altered MLC1 levels (Passchier et al., 2023).

Previous studies from our group demonstrated that GPRC5B knockdown in astrocytes hampered VRAC activity (Alonso-Gardón et al., 2021). Other groups have reported that GPRC5B overexpression increases VRAC activity, similarly to the increased VRAC activity seen in MLC1 overexpression (Passchier et al., 2023). In regards of the MLC-causing GPRC5B mutations, it could be hypothesized that these variants might be incapable of increasing VRAC activity, since most mutations cause a loss of function. However, the authors observed that the patient-derived mutations were equally effective in increasing VRAC currents than GPRC5B WT. They suggest that an increased GPRC5B expression would cause chronic white matter oedema by disrupting VRAC activity. This would explain how heterozygous variants are sufficient to cause a full-blown MLC phenotype (Passchier et al., 2023). Regardless, further studies are needed to comprehend the role of GPRC5B in the volume regulation and in MLC disease.

3. MEMBRANE PROTEINS IN STRUCTURAL BIOLOGY

3.1 PROTEIN EXPRESSION

One of the first challenging decisions when starting to work with purified target protein is to select the most suitable gene expression system for the expression of the protein of interest. In this sense, a wide variety of systems have been developed each having particular features. Depending on the biological properties of the target protein and the planned applications, one expression system will be more appropriate than others.

Knowing the native localization of the protein of interest, its molecular weight, and post-translational modifications is essential for its correct expression. Moreover, it is also important to identify whether the protein is part of a complex, as in this case, co-expression would probably be required for its stabilization (Schütz et al., 2023).

Commonly, prokaryotic target proteins are produced in *E. coli*. In contrast, for eukaryotic target proteins, different expression systems are used based on multiple factors. Firstly, for simple eukaryotic target proteins, *E. coli* can be considered for its expression. In this host, proteins can be produced intracellularly in the cytoplasm, directed into the periplasm, or secreted to the extracellular milieu. It is noteworthy to mention that *E. coli* cannot provide most of the post-translational modifications and it is unable to fold proteins containing multiple disulfide bonds (Gräslund et al., 2008).

For other eukaryotic target proteins, however, yeast, insect cells, or mammalian cells might be more advisable. One of the key differences between these expression hosts lies in the type of glycosylation they provide.

Yeast can provide certain post-translational modifications such as glycosylation, disulfide bond formation, and proteolytic processing (Patra et al., 2021). *Pichia pastoris*, one of the most used yeast-based gene expression system, is capable to produce N- and O-linked glycosylation (Liu et al., 2022). Still, yeast glycosylation is different than mammalian glycan structures which are known to be more complex. Nevertheless, yeast is popularly used for the production of nanobodies (Matsuzaki et al., 2022), cytokines, or certain vaccines, and enzymes for industrial applications (Baghban et al., 2019).

Baculovirus-mediated gene expression in insect cells has become more popular in recent years, especially to produce membrane proteins, such as GPCRs and ion channels. In insect cells, target proteins can be produced intracellularly or secreted to

the extracellular milieu. They can provide N- and O-glycosylation, but they fail to generate complex type N-glycans (Palmberger et al., 2012).

Mammalian cells are especially used for integral membrane proteins as they offer a cellular environment that closely resembles the native one. The main drawback of this system is that the levels of expression are lower compared to the other methods (Pandey et al., 2016). HEK293 cell lines are commonly used. For obtaining higher volumes, suspension cell cultures are employed such as HEK293-6E, HEK293F, and Expi293F. Target proteins can be expressed transiently in mammalian cells by transfection with a plasmid or by transduction with baculoviruses (BacMam). Alternatively, the generation of a stable cell line expressing the target protein is also an option (Schütz et al., 2023).

Lastly, cell-free expression is an emerging method for fast and efficient production of membrane proteins. This system requires supplementation with exogenous amino acids, cofactors, and energy sources, plus mRNA or DNA templates. They are based on lysates of *E. coli* or mammalian extracts from human cell lines or rabbit reticulocytes (Aleksashin et al., 2023).

3.2 MEMBRANE PROTEIN PURIFICATION

Once the gene expression system has been chosen and the target protein is expressed, it needs to be purified. Focusing on membrane proteins, its purification is challenging due to their low concentration and physiochemical properties. The first step for membrane protein purification is to isolate plasma membranes from cells. One of the most frequently used method for this is to employ differential centrifugation that will separate membranes from other organelles based on their distinct densities (Lin and Guidotti, 2009).

Structural and functional studies often require protein extraction from the plasma membrane followed by its reconstitution into an environment that allows such downstream analyses. This process is called solubilization, where the native phospholipid environment is disrupted while maintaining the protein stable and folded (Ratkeviciute et al., 2021). Solubilization is usually performed by adding detergents, which are amphipathic molecules containing both hydrophobic and hydrophilic groups. They tend to form micelles in water and they bind to hydrophobic parts of the protein. Still, even in the presence of detergents, membrane proteins have a tendency to aggregate (Smith, 2017).

The choice of detergent is critical for a successful solubilization. It depends on the intrinsic properties of the target protein and it should sufficiently solubilize the protein without irreversibly denaturing it (Smith, 2017). When working with detergents, it is important to be aware of their critical micelle concentration (CMC). CMC is the concentration of free detergent at which there is an equilibrium between disperse detergent molecules and micellar structures (Schimerlik, 2001). The detergent's head group influences on the detergent-protein interaction, whilst the length of the alkyl chain determines the CMC. Detergents can be classified into three groups based on the nature of the polar head: ionic, nonionic, and zwitterionic. Ionic detergents, such as sodium dodecyl sulfate (SDS), are efficient solubilizers, but tend to denature the protein. Nonionic detergents usually solubilize without altering protein structure. Some nonionic detergents are digitonin, glucosides, and maltosides. Lastly, zwitterionic detergents, such as lauryldimethylamine oxide (LDAO) are likely to inactivate proteins (Kielkopf et al., 2021).

Several tags can be cloned into the gene of interest to improve the purification of proteins. They can be placed either at the N- or C-terminus of a target protein, however, it should be noticed that placing it on the N-terminal end could improve protein translation. Next, some of the most popular tags are briefly described.

His-tag consists of six or more consecutive histidine residues that bind the immobilized metal ions on resin or beads via the histidine imidazole ring, then, the tagged protein can be easily eluted using elution buffers with imidazole or low pH. Another example is the GST tag that binds to resin immobilized glutathione. GFP can also be used as a tag if bonded to the protein of interest since resins that bind this protein have been developed. Other tags include epitope tags, such as the FLAG-tag which is a short, eight-residue (DYKDDDDK) hydrophilic peptide tag. This sequence can be recognized with high specificity by commercially available antibodies. 3x-FLAG tag is a variant made with three tandem repeats of FLAG sequences. Another example is Strep-tag that exploits the strong interaction of biotin and streptavidin. This system uses a synthetic Strep-tag II peptide (WSHPQFEK) that binds to a recombinant and optimized form of streptavidin called Strep-Tactin. Twin-Strep-tag was also developed and consists of the strep motif twice connected by a linker. It displays higher affinity for Strep-Tactin than the Strep-tag II. Another improvement that has been made is the development of Strep-Tactin XT that shows higher affinity for both Strep-tag II and Twin-Strep-tag. For the elution, desthiobiotin, a biotin analog that competes for the binding of Strep-Tactin, is used. (Malhotra, 2009). One advantage of using desthiobiotin is that the elution does not alter

the pH of the sample, while most of the other tags depend on adding an acidic solution to elute it.

3.3 SOLVING PROTEIN STRUCTURE

In the last years, cryogenic electron microscopy (cryo-EM) has become a powerful method in structural biology to solve protein structures. Compared to other techniques, such as X-ray crystallography and nuclear magnetic resonance (NMR), cryo-EM requires lower sample concentrations and it is compatible with large and heterogeneous macromolecular protein complexes. Unfortunately, high-resolution structure determination of small proteins (<100 kDa) remains difficult for two main reasons. Firstly, small proteins generate a lower signal-to-noise ratio (SNR) especially in the presence of detergent micelles. Secondly, small proteins have few distinctive morphological features that complicates particle alignment (Wentinck et al., 2022).

An approach to overcome this problem is to increase the molecular weight of the target protein. To do so, different strategies can be classified into two groups: target protein binding methods, which result in oligomer formation, and target protein fusion methods that result in chimeras (Wentinck et al., 2022).

3.3.1 TARGET PROTEIN FUSION

Target fusion strategies use small epitope motifs to which high-affinity Fabs are available. For its success, this fusion must be very rigid. The use of Fabs remains restricted as they have to be produced for each individual system. To overcome this problem, an approach that eliminates the need to generate Fabs for each system was developed. For this, fusion proteins are introduced to the exposed loops of membrane proteins and the Fab bind to them. With this end, BRIL system was generated.

3.3.1.1 BRIL

BRIL is an engineered thermostable variant of cytochrome b562 from *E. coli.* It consists of four-helix bundle fold that weights 12 kDa. Its small molecular weight is not enough to become an effective fiducial marker. For this reason, the development of BAG2, a semi-synthetic antibody (sAB), specifically recognizing the BRIL element has served as a tool for the structure determination of BRIL-fused membrane proteins. The addition of a 50

kDa Fab facilitates accurate orientation of the proteins, even when their structure is mostly embedded in the membrane (Mukherjee et al., 2020).

BRIL has been successfully used in the structural resolution of some GPCRs, such as the Frizzled5 receptor (Fzd5) (Tsutsumi et al., 2020) and the adenosine A_{2A} receptor (Zhang et al., 2022). In these cases, BRIL was placed in the intracellular loop 3, between helix V and VI of the GPRCs without disrupting their core structure, and BAG2 was incubated with the target protein.

At the same time, another anti-BRIL antibody (SRP2070Fab) was developed, which was used to determine the serotonin receptor $5HT_{1B}$ and the angiotensin II receptor AT_2R structures through X-ray crystallization (Miyagi et al., 2020). Both SRP2070Fab and BAG2 successfully recognize the BRIL protein fusion, but they bind to different epitopes. BAG2 recognizes helices II and III, whereas SRP2070Fab binds to helices III and IV (Miyagi et al., 2023). Although, SRP2070Fab has not been tested for cryo-EM analyses, it is believed that it could also be useful in this technique (Miyagi et al., 2020).

Beside GPCRs, BRIL has also been used for the solving of several protein structures. Among them there can be found the zinc transporter YiiP (Mukherjee et al., 2020) (Figure 19), human pannexin 2 channel (PANX2) (He et al., 2023), the proton-activated chloride channel TMEM206 in pufferfish (Deng et al., 2021), the heterohexameric LRRC8A:C channel in mice (Kern et al., 2023) or the folate transporter SLC19A1 (Dang et al., 2022).

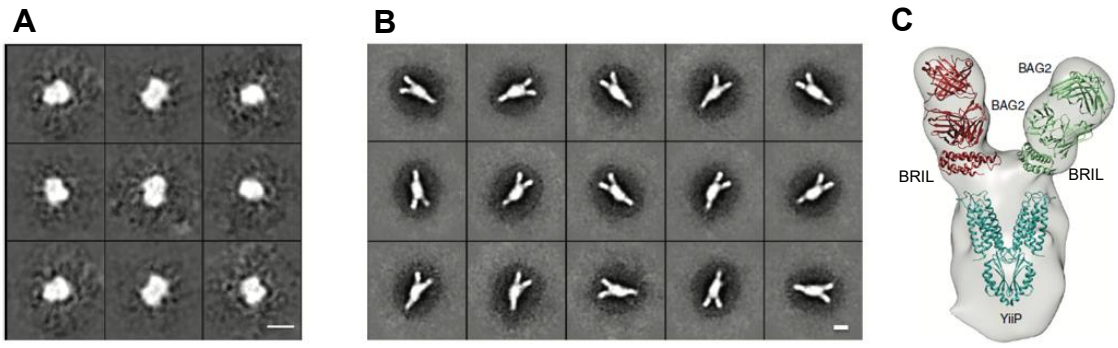


Figure 19. BRIL-fused construct is recognized by BAG2. A. 2D class averages from negative-stain images of the zinc transporter YiiP-BRIL alone. **B.** 2D class averages from negative-stain images of YiiP-BRIL bound to BAG2. **C.** 3D map of YiiP-BRIL-BAG2 complex. Crystal structure of YiiP (in blue, PDB ID: 3H9O) and BRIL-BAG2 (two copies in red and green respectively, PDB ID: 6CBV) were fitted into the low-resolution map. Adapted from (Mukherjee et al., 2020).

3.3.2 TARGET PROTEIN BINDING

Target binding strategies include scaffolding DNA (Aissaoui et al., 2021), thermostabilized antagonists (Zhang et al., 2022), antigen-binding fragments (Fabs) (Wu et al., 2012), and nanobodies (Uchański et al., 2020), among others.

Fabs consist of the antigen binding region of an antibody. They weigh approximately 50 kDa and they have proven to be very successful in solving protein structures, especially for small membrane protein (Wentinck et al., 2022). As an example, a Fab identified from screening a phage display library was used to determine the structure of a 64 kDa glycosyltransferase, ALG6, that is involved in ER-luminal N-glycan synthesis (Bloch et al., 2020). Unfortunately, a stable rigid Fab is not commercially available for every protein, and a single Fab might still be insufficient to add molecular weight for structure determination (Wentinck et al., 2022).

3.3.2.1 NANOBODIES

Another approach for target binding is the use of nanobodies. Nanobodies are the variable domains of heavy-chain only antibodies that naturally occur in camelids (Uchański et al., 2020). Structurally, nanobodies present the immunoglobulin fold, whose core consists of a series of nonconsecutive antiparallel beta strands that form two sheets held together by an intramolecular disulfide bond (Condeminas and Macias, 2024).

Due to their limited size (~ 12 kDa), they do not contribute to increase the target protein molecular weight, but they present several features that makes them attractive for protein structure determination. One singular characteristic is that they can bind to small exposed surfaces, which might not be accessible to conventional antibodies. Furthermore, they can lock the target protein in a specific conformation, thereby reducing target flexibility (Wentinck et al., 2022).

Nanobodies have been successfully used in X-ray crystallography as they bind their target and they induce different contacts resulting in improved crystal packing and resolution (Errasti-Murugarren et al., 2019). In cryo-EM, some protein structures have also been unveiled by the application of nanobodies. For instance, the Na^+-taurocholate cotransporting polypeptide (NTCP) conjugated with two nanobodies revealed two distinct conformations as one nanobody fixed an inward-facing state and the other promoted the opening of the pore (Goutam et al., 2022). In another study, the structure of human

GPR75 in its active state by using an intracellular nanobody was reported (Lv et al., 2022).

Regarding small protein structure resolution in cryo-EM, the use of nanobodies is not common because the small size of this particles does not sufficiently increase the target protein weight. To overcome this limitation, different approaches have emerged that are based in adding extra weight to the nanobody particle.

3.3.3 MEMBRANE MIMETIC SYSTEMS

Structural characterization of membrane proteins remains challenging mainly because of the difficulties in keeping structural and functional integrity when these proteins are extracted from their native environment. As already mentioned, this process is typically achieved with detergents, however, detergent micelles fail to mimic the complexity of the native membrane environment (Autzen et al., 2019). Additionally, the presence of detergents often complicates cryo-EM analysis as they increase the background noise. In this sense, detergents can form empty micelles that can be as large as the micelles containing a small target protein which increases the background noise as well (Nygaard et al., 2020).

For these reasons, several non-detergent alternatives have been developed, including amphipols, nanodiscs, and proteoliposomes, among others (Figure 20). Usually, detergents are preferred for membrane protein extraction, but before sample vitrification for cryo-EM, the detergent is removed and the other membrane mimetic system is added.

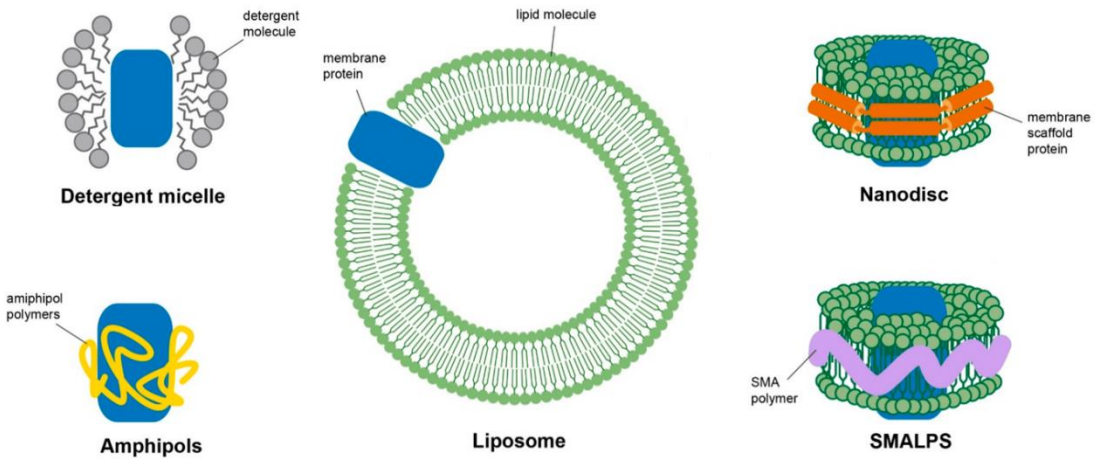


Figure 20. Schematic diagram of different membrane mimetic systems. Membrane proteins can be reconstituted in detergent micelles, amphipols, liposomes and nanodiscs. Membrane protein is represented as a blue block. Modified from (Yeh et al., 2020).

3.3.3.1 AMPHIPOLS

Amphipols are amphipathic polymers formed by hydrophilic backbones and hydrophobic side chains. They can self-assemble into particles that surround membrane proteins and maintain them in their native state. Compared to detergents, they display a greater affinity for the hydrophobic transmembrane domains. Thus, no free amphipol molecules are found in solution, which is an advantage in cryo-EM as they will not influence on the background noise (Autzen et al., 2019). Samples in amphipols produce more uniform ice films on the grids, therefore, a better particle distribution is achieved in comparison to detergents (Flötenmeyer et al., 2007). Furthermore, amphipols can also be useful to stabilize protein complexes and facilitate the analyses of fragile membrane proteins or membrane proteins complexes (Zoonens and Popot, 2014). In few cases, amphipols have been used for protein resolution in cryo-EM, which might be limited for their cost. For instance, in the case of the lysosomal calcium channel TRPML1, three distinct conformational states were identified using amphipols (Zhang et al., 2017a).

3.3.3.2 NANODISCS

Nanodiscs consist of a non-covalent assembly of phospholipid and a genetically engineered membrane scaffold protein (MSP) derived from apolipoprotein. The phospholipid associates as a bilayer domain while two molecules of MSP wrap around the edges of the structure in a belt-like configuration (Bayburt and Sligar, 2010). Its diameter is dictated by the length of the MSP. In nanodiscs, the protein is in a more native like environment, which is typically more stabilizing than the environment created by detergents. Nanodiscs also have a more distinct shape than detergent micelles, which helps align the particles in cryo-EM (Nygaard et al., 2020).

Nanodiscs have been extremely useful in blood clotting research, as the activity of some proteins are dependent on the phospholipid composition (Shaw et al., 2007). They have also been used to test a broad class of GPCR drug targets (Leitz et al., 2006). Regarding protein structures, M2 and β2-adrenergic receptors were reconstituted in nanodiscs to study structural and functional properties in their complexes with β-arrestin (Staus et al., 2020). Another example of cryo-EM resolution is GLUT4, the primary glucose transporter in adipose and skeletal muscle tissue. Its structure was solved bound to a small inhibitor and it was achieved in both detergent micelles and lipid nanodiscs (Yuan et al., 2022).

Nanodiscs can also be made of styrene maleic acid (SMA) copolymers instead of MSP. In this case, SMA wraps around the membrane protein along with its surrounding lipids forming disc-shaped structures referred as SMA lipid particles (SMALPs) or SMA nanodiscs. The main difference is that SMA can only be created directly from intact cells, since the native membrane phospholipids are maintained (Dörr et al., 2016).

SMALPs have been used for the extraction and purification of a human dopamine receptor 1 (DR1) (Bada Juarez et al., 2020) and the structural resolution of glycine receptor (GlyR) bound to glycine, taurine or GABA (Yu et al., 2021).

3.3.3.3 PROTEOLIPOSOMES

Liposomes are vesicles formed by a phospholipid bilayer, which may form spontaneously in aqueous media (Singh, 2019). Proteoliposomes are liposomes to which a protein has been incorporated or inserted (Ciancaglini et al., 2012). Their similarity with the native plasma membrane allows for functional and structural protein studies (Puvanendran et al., 2020).

Although liposomes are widely used for functional protein studies, they are not commonly chosen for structural studies. Recently, a cryo-EM structural analysis workflow for membrane proteins embedded in liposomes was published. In this research, they worked with the well-characterized AcrB *E. coli* transporter as a prototype (Yao et al., 2020).

Later, the structural resolution of other proteins has also been solved in cryo-EM reconstituting the protein in liposomes. This is the case for the bacterial voltage-gated sodium channel NaChBac, which was reconstituted in liposomes under an electrochemical gradient (Chang et al., 2023). Another example is the structural resolution of Respiratory complex I, which is the main entry point for electrons into the mitochondrial respiratory chain. In a recent study, this complex from bovine heart was reconstituted in proteoliposomes. Researchers identified three conformational states for the complex, being two open and one close conformation (Grba et al., 2024).

OBJECTIVES

Megalencephalic Leukoencephalopathy with subcortical cysts (MLC) is a rare genetic disease caused by mutations on either *MLC1, GLIALCAM, GPRC5B* or *AQP4.* To date, the pathophysiological mechanisms underlying the disease are not fully understood.

Our group has extensively studied MLC1 and GlialCAM and has found a wide number of proteins related to them. At the same time, both proteins have been linked to different physiological processes in which they participate. For instance, GPRC5B was found to be part of this network. With this thesis, we aimed to gain knowledge on the molecular mechanisms underlying GPRC5B regulation related to MLC disease, as well as, to gain further understanding of the structure of MLC1 and GlialCAM proteins. The specific objectives are:

1. Characterize an MLC patient carrying two MLC1 mutations showing a remitting phenotype using genetic and biochemical studies.

2. Study the GPRC5B- mediated signalling transduction pathways and the potential implication of MLC1 and GlialCAM in the signalling cascade.

3. Study the pathogenicity of *GLIALCAM* patient-derived mutations located at the IgC2 domain, as well as the *GPRC5B* patient-derived mutations. Determine a potential link of GPRC5B activity on the function of GlialCAM related to its IgC2 domain.

4. Obtain the tridimensional structure of MLC1 via cryo-EM.

METHODOLOGY

1. MOLECULAR TECHNIQUES

In this thesis, different genes of interest were cloned into bacterial or eukaryotic vectors to express them in cellular models. The generation of these vectors was achieved by Polymerase Chain Reaction (PCR) followed by Gateway technology (developed by Invitrogen).

1.1 PCR

PCR is a technique that allows amplifying selectively DNA fragments exponentially.

Material

- KOD Hot Start DNA Polymerase kit (Novagen) containing:
  - Polymerase Buffer 10x
  - dNTPs 2 mM
  - $MgSO_4$ 25 mM
  - KOD polymerase 1 U/µL
- Forward Primer 10 µM
- Reverse Primer 10 µM
- Template DNA 100 ng/µL
- Milli-Q water
- Thermocycler

Protocol

1. Prepare the PCR mix for each reaction by adding the following volumes in each tube:

- Polymerase Buffer 10x	5 µL
- dNTPs 2 mM	5 µL
- $MgSO_4$ 25 mM	3 µL
- KOD polymerase 1U/µL	1 µL
- Forward Primer 10 µM	1.5 µL
- Reverse Primer 10 µM	1.5 µL
- Template DNA 100 ng/µL	1-3 µL
- Milli-Q water	X µL
Final Volume	50 µL

2. Spin shortly the mix.
3. Place the tubes in the thermocycler and set the PCR protocol with the following conditions:

Initiation	2 minutes at 95°C	
Denaturation	30 seconds at 95°C	25 cycles
Annealing	30 seconds at 55°C	
Extension	1 minute and 30 seconds at 70°C	
Ending	10 seconds at 95°C	
Storage	4°C	

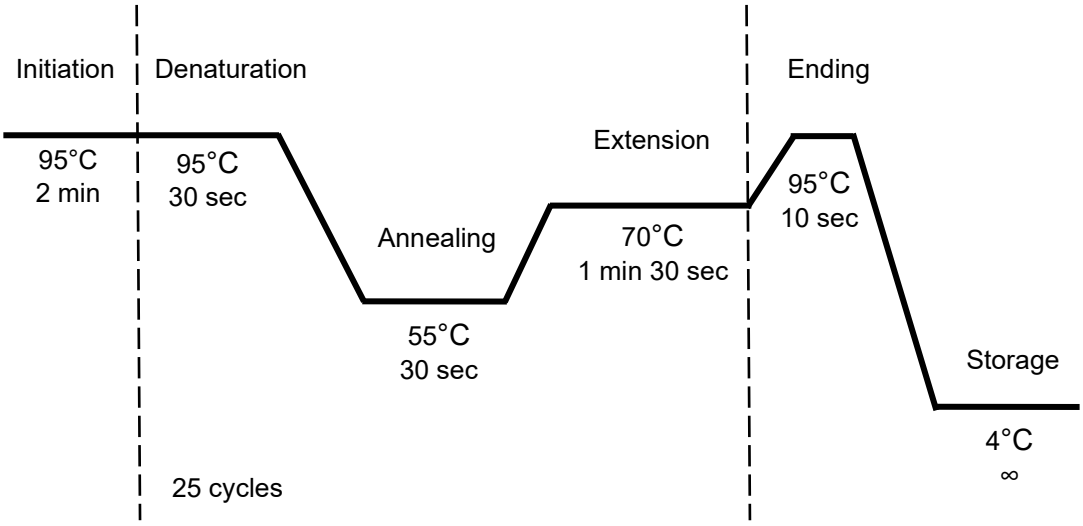


Figure 21. PCR protocol. PCR consists of multiple cycles of denaturation where the DNA strands get separated; annealing where primers bind to the DNA strands; and extension where DNA polymerase extends the 3' end.

* The duration of the extension phase depends on the fragment length, it is considered to be 1 Kb per 1 minute.

** The annealing temperature is based on the primer sequence, and it is calculated with the following formula:

$$Tm = [4(G + C) + 2(A + T)] - 4$$

1.1.1 RECOMBINANT PCR/ SPLICING BY OVERLAP EXTENSION (SOE) PCR

Recombinant PCR or Splicing by Overlap Extension (SOE) PCR is used to induce point mutations, insertions, and/ or deletions on the template DNA. In this technique, two fragments containing the mutation are cloned separately, followed by a second PCR aimed at overlapping them.

Material

- KOD Hot Start DNA Polymerase kit (Novagen) containing:
  - Polymerase Buffer 10x
  - dNTPs 2 mM
  - $MgSO_4$ 25 mM
  - KOD polymerase 1 U/µL
- Forward Primer 10 µM
- Reverse Primer 10 µM
- Template DNA 100 ng/µL
- Milli-Q water
- Thermocycler

Protocol

1. Add the following volumes for each reaction in an Eppendorf tube:

- Polymerase Buffer 10x	5 µL
- dNTPs 2 mM	5 µL
- $MgSO_4$ 25 mM	3 µL
- KOD polymerase 1U/µL	1 µL
- Forward Primer 10 µM	1.5 µL
- Reverse Primer 10 µM	1.5 µL
- Template DNA 100 ng/µL	1-3 µL
- Milli-Q water	X µL
Final Volume	50 µL

2. Put the tubes in the thermal cycler and program it with the following conditions:

Initiation	2 minutes at 95°C	
Denaturation	30 seconds at 95°C	25 cycles
Annealing	30 seconds at 55°C	
Extension	1 minute and 30 seconds at 70°C	

Ending	10 seconds at 95°C
Storage	4°C

1.2 DNA ELECTROPHORESIS AND PURIFICATION

Once the DNA is cloned, the PCR product must be checked. For this, a non-denaturalizing agarose gel electrophoresis is run, where the DNA migrates based on its molecular weight. From this gel, bands can be cut and purified. DNA electrophoresis is also performed when digesting vectors with restriction enzymes.

1.2.1 ELECTROPHORESIS

Material

- TAE 1x: prepared from TAE 50x diluted in distilled water
  TAE 50x: 2 M Tris-base; 1 M acetic acid; 50 mM EDTA pH 8 in distilled water
- Loading buffer 6x (ThermoFisher Scientific)
- GeneRuler (ThermoFisher Scientific) DNA molecular weight ladder
- Agarose
- SYBR Safe DNA Gel Stain (Invitrogen)
- Mould tray
- Well comb
- Electrophoresis bucket (Bio-Rad)
- Power supply
- Trans-illuminator UV light

Protocol

1. Weigh 1 g of agarose in 100 mL of TAE 1x.
2. Heat 2 minutes in the microwave or until the agarose is dissolved.
3. Once it has cooled down, add 10 µL of RedSafe and pour it into the mould tray. Let it cool down to solidify.
4. Prepare samples with loading buffer at a final concentration of 1x.
5. Load DNA samples and DNA ladder.
6. Run the electrophoresis at 80-100 V.
   * The running time will depend on the size of the expected bands.
7. Visualize the gel in a UV trans-illuminator.

8. Cut the bands of interest using a blade and collect them in separate Eppendorf tubes.
   * Bands are not cut on control plasmid digestions with restriction enzymes.
9. Acquire photo of the gel.

1.2.2 DNA PURIFICATION

Material

- High Pure PCR product purification kit (Roche) containing:
  - Centrifuge tubes with coupled columns
  - Binding buffer
  - Washing buffer
- Milli-Q water
- Blades
- Dry bath

Protocol

1. Cut the desired band from the agarose gel with a blade and put it in an Eppendorf tube.
2. Add 500 µL of Binding Buffer.
3. Heat the sample at 60°C in a dry bath until the agarose is melted.
4. Transfer the liquid sample to a column coupled to an Eppendorf tube from the Purification kit.
5. Centrifuge at 13,000 rpm for 1 minute.
6. Discard the supernatant and add 500 µL of Washing buffer.
7. Centrifuge at 13,000 rpm for 1 minute.
8. Discard the supernatant and add 200 µL of Washing buffer.
9. Centrifuge at 13,000 rpm for 1 minute.
10. Discard the supernatant and transfer the column to a new Eppendorf tube.
11. Add 50 µL of milli-Q water, incubate for 1 minute to elute the DNA.
12. Centrifuge at 13,000 rpm for 1 minute.

1.3 GATEWAY SYSTEM

In this thesis, different constructs have been generated to express cloned genes or DNA fragments into living systems with the aim to perform functional assays or express proteins for their purification. Gateway recombination cloning has been used to achieve so.

Gateway technology (Invitrogen) is based on the highly specific integration and excision reactions of lambda phage sequences into the *E. coli* genome. When the phage infects bacteria, the DNA sequence *att*P from the phage recombines with the bacterial DNA sequence *att*B in the presence of integration-specific enzymes. As a result, the phage DNA is integrated into the bacterial genome and the hybrid recombination sites are now called *att*L and *att*R. In the lytic phase, phage DNA can excise itself from bacterial DNA in the presence of a different set of recombination and excision enzymes. These recombinant reactions are called BP (*att*B x *att*P → *att*L + *att*R) and LR (*att*L x *att*R → *att*B + *att*P), they are reversible, and they are the basis of the Gateway cloning system. Different *att* sites facilitate the cloning of each DNA fragment into the vector, for instance, *att*B1 can only recombine with *att*P1 and not *att*P2 (Figure 22).

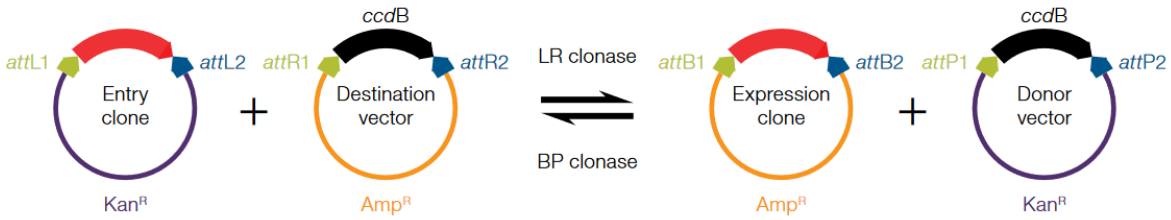


Figure 22. Gateway reactions. Scheme showing the four types of plasmids and enzyme mixes involved in Gateway cloning reactions. Red arrows represent the gene of interest. Adapted from (Katzen, 2007).

1.3.1 GENERATION OF PRIMERS

The first step for Gateway cloning is the addition of *att*B sequences flanking the DNA of interest. The primers used for the PCR are designed to include the *att*B regions, which are 18-25 bp in length. The *att*Bs regions used in this thesis are *att*B1, *att*B2, *att*B5R and *att*B5, where *att*B5R and *att*B5 recombine. Its use will depend on the desired type of construct. In order to introduce our insert in an expression vector without any molecular tags or epitopes, we will flank the insert with *att*B1 (5') and *att*B2 (3'). In contrast, to fuse our insert with a tag or epitope, one sequence will be flanked with *att*B1 (5') and *att*B5R

(3') and the other with *att*B5 (5') and *att*B2 (3'). This variant is called Multisite Gateway two-fragment cloning.

When generating the primers, it is also important to consider that the *att*B1 primer should contain the Kozac sequence (ACC) right before the start codon (ATG) as it is necessary for the recognition of the start codon for gene translation. In the same manner, the *att*B2 should contain the sequence of the stop codon, but it should be removed in the *att*B5R so that the second fragment can be expressed. In Figure 23, the *att*Bs sequences used in this thesis are displayed.

*att*B1 (*forward*): 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTAACC-(GEN)-3'
*att*B1 *Kozak*

*att*B5R (*reverse*): 5'-GGGGACAACTTTTGTATACAAAGTTGT-(GEN)-3'
*att*B5R

*att*B5 (*forward*): 5'-GGGGACAACTTTGTATACAAAAGTTGNN-(GEN)-3'
*att*B5

*att*B2 (*reverse*): 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT TAG-(GEN)-3'
*att*B2 *stop*

Figure 23. Example of attB sequences.

1.3.2 BP CLONING REACTION

For every DNA fragment of interest, an Entry clone needs to be generated by the BP cloning reaction. In this reaction, the PCR product flanked by *att*B sequences is combined with a commercially available donor vector (pDONR221) thanks to the BP clonase enzyme. The donor vector contains *att*P1 and *att*P2 sequences. In the event of cloning two DNA fragments, the first entry clone needs to be flanked by *att*B1 and *att*B5R, whereas the other one by *att*B5 and *att*B2. In this case, the donor vectors will contain *att*P1- *att*P5R and *att*P5 - *att*P2, respectively (Figure 24).

Figure 24. BP recombination scheme. Adapted from ThermoFisher – Gateway Technology Manual.

After an incubation at room temperature, competent *E. coli* cells are transformed and kanamycin resistant colonies will be obtained, since the donor vector contains a kanamycin resistance gene. The efficiency is high due to the specificity of the recombinant reaction and negative selection is conferred by the presence of *ccdB* gene.

Material

- PCR product
- pDONR221
- BP clonase II enzyme mix
- Milli-Q water
- Proteinase K
- Dry bath

Protocol

1. Prepare the reaction mix. The final volume is 5 µL.

*att*B-PCR product (20-50 fmols)	3 µL
pDONR (150 ng/µL)	1 µL
BP Clonase II	1 µL

2. Spin the mix.
3. Incubate at room temperature for 2 hours.
   * BP reaction can also be left O/N.
4. Stop the reaction adding 0.5 µL of Proteinase K and incubate at 37°C for 10 minutes.

After, the BP reaction is transformed into competent bacteria and seeded into kanamycin plates as described in *Section 1.6*. The final vectors are digested by restriction enzymes and sequenced, as described in *Sections 1.9 and 1.10*, respectively.

1.3.3 LR II REACTION

Once the entry clones are created, they need to be introduced into the destination vector. The LR recombination between an entry clone and a destination vector will produce an expression clone. It is performed by the enzyme LR clonase II (Figure 25).

After an incubation at room temperature, competent *E. coli* cells are transformed and ampicillin resistant colonies will be obtained, since the destination vector contains an

ampicillin resistance gene. The efficiency is high due to the specificity of the recombinant reaction and negative selection is conferred by the presence of *ccdB* gene.

Figure 25. LR recombination scheme. Adapted from ThermoFisher – Gateway Technology Manual.

Material

- Entry clone
- Destination vector
- LR clonase II enzyme mix
- Milli-Q water
- Proteinase K
- Dry bath

Protocol

1. Prepare the reaction mix. The final volume is 5 µL.

Entry clone	150 ng (1-3 µL)
Destination vector	150 ng (1-3 µL)
LR Clonase II	1 µL
Milli-Q water	Up to 5 µL

2. Spin the mix.
3. Incubate at room temperature O/N.
4. Stop the reaction by adding 0.5 µL of Proteinase K and incubate at 37°C for 10 minutes.

After, the LR reaction is transformed into competent bacteria and seeded into ampicillin plates as described in *Section 1.6.* The final vectors are digested by restriction enzymes, as described in *Section 1.9*.

1.3.4 LR II + REACTION

In the Multisite Gateway System, molecular tags are fused into our gene of interest. The recombination occurs between two entry clones and the destination vector. The first entry must contain *att*L1 and *att*5R sequences, while the second one needs to have *att*L5 and *att*B2 sequences. The reaction is performed by the enzyme LR clonase II+. The whole Gateway Multisite system for DNA cloning is summarized in Figure 26.

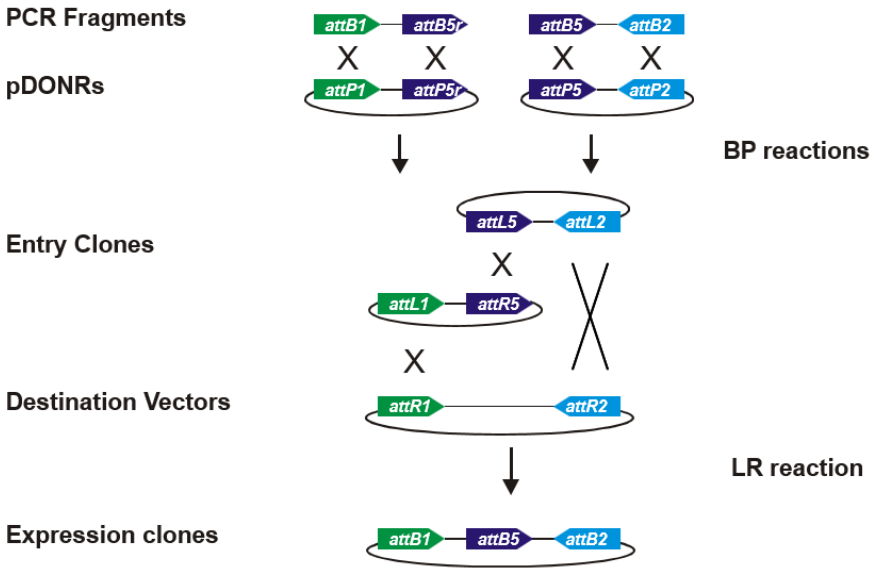


Figure 26. LR Multisite recombination scheme. Adapted from ThermoFisher – Gateway Technology Manual.

Material

- Entry clone
- Destination vector
- LR clonase II enzyme mix
- Milli-Q water
- Proteinase K
- Dry bath

Protocol

1. Prepare the reaction mix. The final volume is 5 µL.

Entry clone 1	10 fmol (≤1 µL)
Entry clone 2	10 fmol (≤1 µL)
Destination vector	20 fmol (≤1 µL)
LR Clonase II+	1 µL
Milli-Q water	Up to 5 µL

   To convert DNA femtomoles (fmols) into nanograms (ng) the following formula is applied:

$$ng = (x\ fmols) \cdot (N) \cdot \left(\frac{660\ fg}{fmols}\right) \cdot \left(\frac{1\ ng}{10^6\ fg}\right)$$

   where *x* is the number of fmols and *N* is the DNA size in base pairs (bp).
2. Spin the mix.
3. Incubate at room temperature O/N.
4. Stop the reaction adding 0.5 µL of Proteinase K and incubate at 37°C for 10 minutes.

After, the LR reaction is transformed into competent bacteria and seeded into ampicillin plates as described in *Section 1.6*. The final vectors are digested by restriction enzymes, as described in *Section 1.9*.

1.4 BACMAM SYSTEM

The BacMam technology is based on using baculoviruses as vehicles to efficiently express genes in mammalian cells. A baculovirus is a double-stranded DNA insect virus that has been modified for enabling transgene expression in mammalian cells. BacMam particles are taken up by endocytosis and released for transcription and expression following migration to the nucleus. BacMam has been used for a growing number of applications, including drug discovery or large-scale protein production for structural studies (Goehring et al., 2014).

For the generation of the BacMam virus, the Bac-to-Bac baculovirus expression system (Invitrogen) was employed. This system uses the Tn7-mediated site-specific transposition reaction to direct the integration of the transfer vector expression cassettes

into a baculovirus backbone vector (bacmid) in the *E. coli* strain DH10Bac (Fornwald et al., 2016).

For the generation of the plasmid containing our gene of interest, also called donor plasmid, restriction enzymes were used for the cloning. The donor plasmid allows the generation of an expression construct containing the gene of interest under the control of a baculovirus-specific promoter. This expression cassette is flanked by Tn7 and it also includes an antibiotic resistance gene and an SV40 polyadenylation signal to form a mini Tn7.

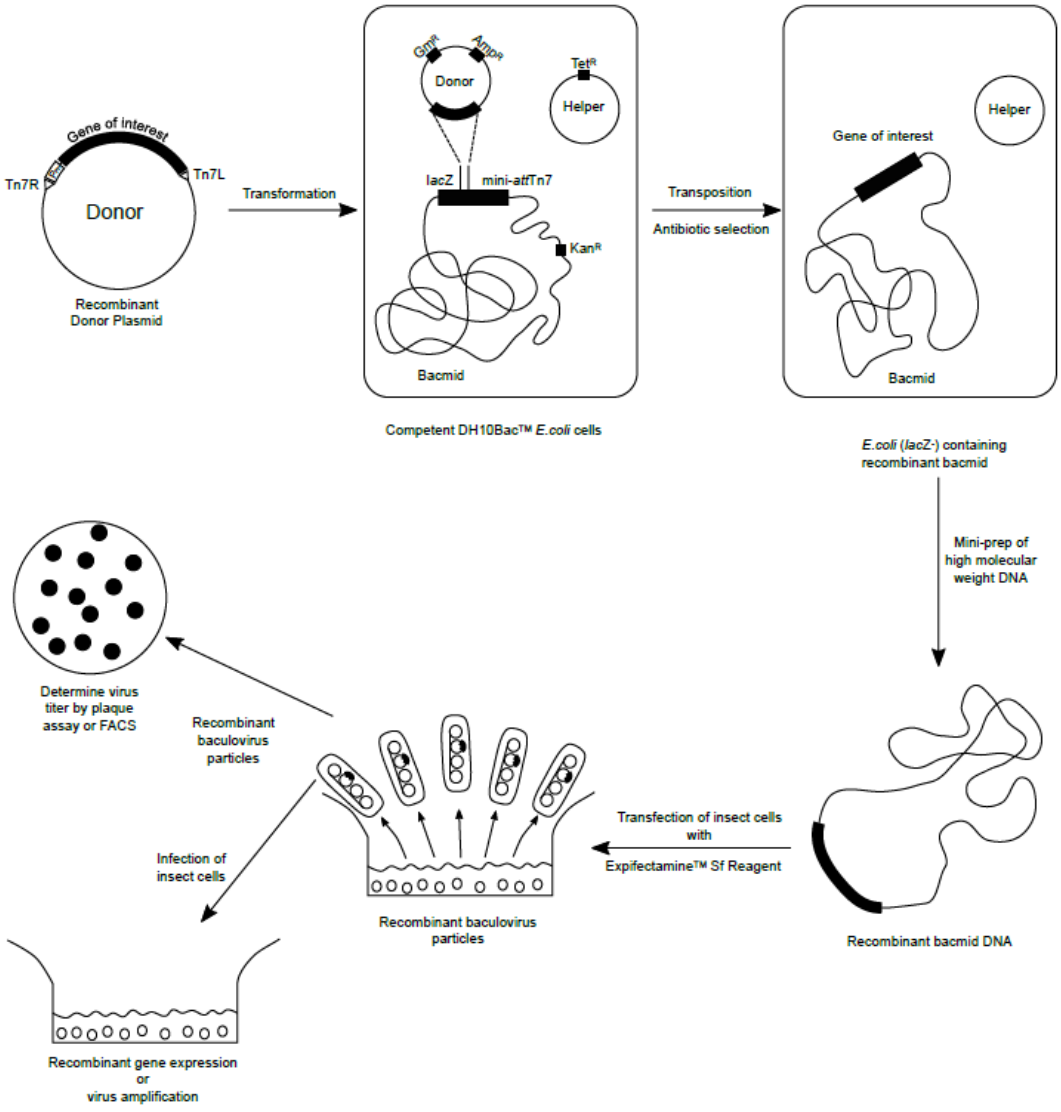


Figure 27. Scheme of the Bac-to-Bac system. Generation of recombinant baculovirus and the expression of a gene of interest using the Bac-to-Bac Baculovirus Expression System. Extracted from: Bac-to-Bac Baculovirus Expression System – User Guide from ThermoFisher Scientific.

Once the donor vector is cloned, it is transformed in DH10Bac *E. coli* strain which is used as the host for the construct containing the gene of interest. These cells enclose a bacmid with a mini-*att*Tn7 target site and a helper plasmid. When the vector is transformed into DH10Bac cells, transposition occurs between the mini-Tn7 element on the donor vector and the mini-*att*Tn7 target site on the bacmid and the recombinant bacmid is produced. To ensure the success of the reaction, the presence of transposition proteins supplied by the helper plasmid is necessary. DH10Bac transformation is detailed in *Section 1.6.4.*

Then, the bacmid is isolated and it is transfected into Sf9 insect cells to generate the recombinant BacMam virus. A scheme of the Bac-to-Bac system is displayed in Figure 27. Once the BacMam virus is generated in insect cells, it will be used to infect mammalian cells. The production of BacMam virus and the mammalian cell infection are further explained in *Section 2.7* and *Section 2.8*, respectively.

1.5 COMPETENT BACTERIA OBTENTION

Constructs generated with the Gateway system can be transformed and amplified, since these plasmids contain a replication origin (ORI) that enables its replication in bacterial cells. This process, called transformation, occurs spontaneously in bacteria called competent, although it happens at a low frequency. However, it can be reproduced in the laboratory for research purposes. In this case, bacterial cell competence is artificially induced, and higher amounts of DNA are obtained.

There are two main methods for bacteria cells to artificially become competent: the chemical or thermal competence and the electrocompetence. Both methods share the basis of generating pores in bacterial cell membranes. In the first case, the chemical competence, cells are treated with a calcium rich environment to prevent the electrostatic repulsion between the plasmid DNA and the bacterial cellular membrane. For their transformation, usually a heat shock is performed, where a sudden increase in temperature creates pores in the membrane that allows for plasmid DNA to enter the cell. In the second method, cells obtain their competence through electricity, meaning that after an electric charge, pores are generated in the bacterial wall whereby DNA can enter the cell.

In this thesis, we worked with three different *E. coli* strains, *DH5α*, WK6, and DH10Bac and we followed different transformation procedures for their acquisition of competence. *DH5α* strain was used to generate and amplify most constructs. WK6 strain was used to

amplify nanobodies because they enable the nanobody to be expressed in the periplasm. Lastly, DH10Bac strain was used for the generation of bacmids.

Plasmid DNA contains an antibiotic resistance gene which is key for selective growth of successfully electroporated bacterial cells in an antibiotic-containing media. Only positive-transformed cells will survive, which then will be seeded on LB-agar Petri dishes to form strains macroscopically visible.

1.5.1 ELECTROCOMPETENT BACTERIA OBTENTION

Material

- 2 Erlenmeyer flasks (1 L capacity)
- 2 Centrifuge bottles (500 mL capacity)
- Spectrophotometer and spectrophotometer compatible cuvettes
- Refrigerated centrifuge and rotor compatible with 500 mL bottles
- Laminar flow hood (for bacteria)
- Liquid nitrogen within a tank
- *E. coli* (DH5α or DB3.1)

  * DB3.1 strain is used to amplify empty vectors. It contains *gyrA462* allele, which makes these bacteria resistant to *ccdB* lethality gene.

  * To seed pre-cultured DB3.1 bacteria transformed with empty vectors, selective LB medium containing chloramphenicol and the antibiotic used for recombined vectors (kanamycin for entry clones, ampicillin for destination vectors) are needed.
- LB sterile media
- Glycerol 10% (v/v) in milli-Q water

Protocol

When working with bacteria, it is strongly recommended to work under the laminar flow hood to maintain sterile conditions and avoid contaminations.

Day 1

1. Pour 10 mL of LB inside 50 mL Falcon tubes and inoculate the tubes with bacterial stock.
2. Incubate at 250 rpm for 12-16 hours at 37°C.

3. Autoclave the flasks containing 500 mL of LB, 1 L of glycerol 10% and the centrifuge bottles.

Day 2

1. Add one starter culture to each flask.
2. Incubate at 250 rpm for 100 minutes at 37°C.
3. Transfer 1 mL of culture into a cuvette to read OD at a 600 nm wavelength.
4. Once the OD_{600} reaches 0.5-0.7; place the flasks on ice to stop bacterial growth.
5. Centrifuge the cultures at 4,000 x g for 15 minutes at 4°C.
6. Discard the supernatant by decantation.
7. Resuspend the pellets with a sterile Pasteur pipette using 2 mL of 10% glycerol.
8. Mix the resuspended pellets in one bottle and add 300 mL of 10% glycerol.
9. Centrifuge at 4,000 x g for 15 minutes at 4°C.
10. Discard the supernatant by decantation.
11. Wash again with 300 mL of 10% glycerol.
12. Centrifuge at 4,000 x g for 15 minutes at 4°C.
13. Wash twice with 150 mL of 10% glycerol.
14. Centrifuge at 4,000 x g for 15 minutes at 4°C.
15. Resuspend the pellet with 2 mL of 10% glycerol and prepare 50 µL aliquots.
16. Freeze immediately after pipetting each volume by placing the tubes in liquid nitrogen and store at -80°C.

1.5.2 COMPETENT BACTERIA THROUGH $CaCl_2$ OBTENTION

For the acquisition of chemical competence in bacteria, cells are treated with $CaCl_2$ during the exponential growth phase to permeabilize the bacterial wall.

Material

- LB sterile
- $CaCl_2$ 50 mM sterile, prepared with milli-Q water
- Spectrophotometer and spectrophotometer compatible cuvettes
- Laminar flow hood (for bacteria)

Protocol

When working with bacteria, it is strongly recommended to work under the laminar flow hood to maintain sterile conditions and avoid contaminations.

1. Inoculate 3 mL of LB without antibiotic with a bacteria colony. Grow overnight at 37°C with shaking.
2. Dilute the preculture 1/100 in LB (250 µL of preculture in 25 mL of LB). Incubate at 37°C with shaking until the DO at 600 nm is between 0.3-0.4 (around 2-3 hours), which corresponds to $5 \cdot 10^7$ cells/ mL.
3. Stop the growing incubating the bacteria 15-20 minutes on ice. From now on, work at 4°C and in sterile conditions.
4. Centrifuge at 4,000 rpm for 5 minutes at 4°C. Discard the supernatant and resuspend the cells in 25 mL of cold $CaCl_2$ 50 mM.
5. Maintain the suspension on ice for 30 minutes.
6. Centrifuge at 4,000 rpm for 5 minutes at 4°C.
7. Discard the supernatant and resuspend cells with 2 mL of $CaCl_2$ 50 mM, aliquot in 100-200 µL and store at -80°C. At this point, cells are competent, and its transformation efficiency increases 4-6 times during the first 12-24 hours if kept at 4°C.

1.6 BACTERIA TRANSFORMATION

1.6.1 BACTERIA TRANSFORMATION THROUGH ELECTROPORATION

Material

- 50 µL of DH5α strain of *E. coli* bacteria
- Plasmid DNA
- Micropulser™ electroporation cuvettes 0.1 cm (Bio-Rad)
- Electroporator (Bio-Rad Micropulser)
- LB medium
- LB-agar Petri dishes with antibiotic for bacterial selection

Protocol

1. Thaw the competent cells on ice.
2. Add 0.5 µL of plasmid DNA into an aliquot of 50 µL electrocompetent bacteria, mix gently and keep on ice.
3. Transfer the volume to a cuvette.
4. Place the cuvette in the electroporator and apply a 375 V electric charge.
5. Collect the electroporated bacteria by adding 250 µL of LB media and transfer the volume to a 15 mL Falcon tube.
6. Incubate 1 hour at 37°C and 250 rpm.
7. Seed 100-200 µL of the bacteria on the antibiotic-containing LB-agar Petri dish.
   * Seed 5 µL if the amplified plasmid is an expression vector.
8. Incubate the Petri dish at 37°C O/N.
9. Select isolated colonies and grow them in LB medium containing the antibiotic of selection at 37°C O/N.

* It is recommended to work within range of a Bunsen burner to maintain sterile conditions.

1.6.2 BACTERIA TRANSFORMATION THROUGH THERMAL SHOCK

Material

- DH5α Competent Cells for Subcloning (#EC0111, ThermoScientific)
- LB or SOC media
- Plasmid DNA
- LB plates with antibiotic

Protocol

1. Thaw competent cells on ice.
2. Gently mix the cells and make 50 µL aliquots of competent cells.
3. Add 1-5 µL of DNA directly into a tube of competent cells. Mix well by gently flicking tube several times.
4. Incubate cells on ice for 30 minutes.
5. Heat-shock the cells for exactly 20 seconds in 42°C water bath.
   * Do not mix or shake the tube.
6. Incubate the cells on ice for 2 minutes.
7. Add 950 µL of pre-warmed growth medium LB or SOC.

8. Incubate the tube at 225 rpm for 1 hour at 37°C.
9. Spread 20 µL of cells from each transformation reaction on separate LB plates containing the appropriate selective antibiotic.
10. Invert the plates and incubate overnight at 37°C.
11. Select isolated colonies and grow them in LB medium containing the antibiotic of selection at 37°C O/N.

* It is recommended to work within range of a Bunsen burner to maintain sterile conditions.

* This kit of competent bacteria was used routinely for the amplification of vectors. Its transformation efficiency is calculated to be around 10^6 CFU/µg. When amplifying BP or LR products, high efficiency competent cells were used (DH5α Competent Cells #EC0112, ThermoScientific), which have a transformation efficiency of 10^9 CFU/µg. In this case, the protocol is the same, except that the heat-shock lasts for 30 seconds instead of 20; and 250 µL of LB or SOC medium is added instead of 950 µL.

1.6.3 BACTERIA TRANSFORMATION THROUGH $CaCl_2$

Material

- Competent cells *E. coli* WK6 100-200 µL
- Plasmid DNA
- LB
- LB plates with antibiotic

Protocol

1. For each aliquot of 100-200 µL competent cells, add 50 ng of the recombinant plasmid.
2. Keep the mix 30-60 minutes on ice.
3. Incubate 2-3 minutes at 42°C.
4. Incubate 2 minutes on ice.
5. Add 800 µL of LB without antibiotic. Incubate 1 hour at 37°C and 250 rpm.
6. Seed on LB plate with the antibiotic the plasmid is resistant to.
7. Invert the plates and incubate overnight at 37°C.
8. Select isolated colonies and grow them in LB medium containing the antibiotic of selection at 37°C O/N.

   * It is recommended to work within range of a Bunsen burner to maintain sterile

1.6.4 DHBAC CELL TRANSFORMATION

Material

- MAX Efficiency DH10Bac Competent Cells (#10361012, Gibco)
- BacMam plasmid DNA
- SOC media
- Plates supplemented with kanamycin (50 µg/mL), gentamycin (10 µg/mL), tetracycline (10 µg/mL), BluoGal (100 µg/mL) and IPTG (40 µg/mL)

Protocol

1. Mix 100 µL of DH10Bac cells with 200 ng of pBacMam plasmid.
2. Incubate on ice for 30 minutes.
3. Heat-shock cells 42 seconds at 42°C.
4. Incubate 2 minutes on ice.
5. Add 950 µL of SOC media and incubate 4 hours at 37ºC with agitation.
6. Plate on agar plates.
7. Invert the plates and incubate overnight at 37°C.
8. Select isolated colonies and grow them in LB medium containing the antibiotics of selection at 37°C O/N.

1.7 PLATES

Material

- LB with agar
- Petri dishes
- Antibiotic
- Distilled water

Protocol

1. Weigh the LB and the agar, dilute with distilled water.
2. Autoclave it.
3. Wait until the solution reaches 65°C, then, under sterile conditions, add the antibiotic. If needed, add BluoGal and IPTG (see Table 3 for working concentrations).
4. Mix and pour 20 mL of liquid into each Petri dish.
5. Wait until it solidifies, and plates reach room temperature.

6. Store at 4°C for its use.

	Stock	Working concentration
Ampicillin	100 mg/mL	100 µg/mL
Kanamycin	25 mg/mL	25 µg/mL
Tetracyclin	5 mg/mL	10 µg/mL
Gentamycin	10 mg/mL	10 µg/mL
BluoGal	20 mg/mL	100 µg/mL
IPTG	200 mg/mL	40 µg/mL

Table 3. Complementation of LB plates. Antibiotic, BluoGal and IPTG concentrations needed for LB plates.

1.8 PLASMID DNA PURIFICATION

Commercial kits have been developed to extract and purify DNA produced by transformed bacteria. In our laboratory, we work with QIAprep kits from Qiagen. Depending on the amount of DNA needed, minipreps, midipreps and maxipreps will be used.

1.8.1 MINIPREPARATION

Material

- 5 mL of transformed bacteria culture
- QIAprep Spin Miniprep Kit containing:
    - Resuspension Buffer P1
    - RNAse A
    - Lysis Buffer P2
    - Neutralization Buffer N3
    - Qiagen ionic columns with coupled collector tubes
    - Washing buffer
- Milli-Q water

Protocol

1. Grow overnight 5 mL of bacterial culture.
2. Centrifuge at 4,000 rpm for 10 minutes and discard the supernatant.
3. Resuspend the bacterial pellet in 250 µL of Resuspension buffer.
4. Add 250 µL of Lysis buffer and mix by inversion.
   * If LyseBlue has been added to the Resuspension buffer the solution should become blue.
5. Add 350 µL of Neutralization buffer and mix by inversion.
   * The solution should become white.
6. Centrifuge at 13,000 rpm for 10 minutes.
7. Transfer the supernatant into the Qiagen ionic column.
8. Centrifuge at 13,000 rpm for 1 minute.
9. Discard the supernatant.
10. Wash the column with 750 µL of Washing buffer.
11. Centrifuge at 13,000 rpm for 1 minute.
12. Place the column in a 1.5 mL Eppendorf tube.
13. Add 50 µL of milli-Q water and incubate for 1 minute at room temperature.
14. Centrifuge at 13,000 rpm for 1 minute.
15. Quantify DNA using Nanodrop
16. Store at -20°C.

1.8.2 MIDIPREPARATION

Material

- 50 mL of transformed bacteria culture
- QIAprep Spin Midiprep Kit containing:
    - Resuspension Buffer
    - RNAse A
    - Lysis Buffer
    - Neutralization Buffer
    - QIAfilter Cartridge, plunger and tube extender
    - Capture Buffer
    - Qiagen ionic columns
    - ETR Wash buffer
    - PE Wash buffer

    - Elution buffer
- Milli-Q water

Protocol

1. Grow overnight 50 mL of bacterial culture.
2. Centrifuge at 5,400 rpm for 20 minutes and discard the supernatant.
3. Resuspend the bacterial pellet in 4 mL of Resuspension buffer.
4. Add 4 mL of Lysis buffer, mix by inversion, and incubate 3 minutes at room temperature.

   * If LyseBlue has been added to the Resuspension buffer the solution should become blue.
5. Add 4 mL of Neutralization buffer and mix by inversion.

   * The solution should become white.
6. Couple the QIAfilter Cartridge on a 50 mL Falcon tube and transfer the bacterial lysate.
7. Incubate 10 minutes at room temperature.
8. Introduce the plunger and press to filter the lysate.
9. Add 2 mL of BB buffer and mix by inversion.
10. Couple the tube extender on the upper part of the column and transfer the lysate through the tube extender into the column.
11. Insert the vacuum in the bottom part of the column.
12. Add 700 µL of ETR Wash buffer and apply the vacuum.
13. Add 700 µL of PE Wash buffer and apply the vacuum.
14. Centrifuge at 17,900 rpm for 1 minute.
15. Couple the column into a new 1.5 mL Eppendorf tube and add 200 µL of EB buffer or milli-Q water to elute.
16. Incubate 1 minute at room temperature.
17. Centrifuge at 17,900 rpm for 1 minute.
18. Quantify DNA using Nanodrop
19. Store at -20°C.

1.8.3 MAXIPREPARATION

Material

- 250 mL of transformed bacteria culture
- QIAprep Spin Maxiprep Kit containing:
    - Resuspension Buffer
    - RNAse A
    - Lysis Buffer
    - Neutralization Buffer
    - QIAfilter Cartridge, plunger and tube extender
    - Capture Buffer
    - Qiagen ionic columns
    - ETR Wash buffer
    - PE Wash buffer
    - Elution buffer
- Isopropanol
- 70% ethanol
- Milli-Q water

Protocol

1. Grow overnight 250 mL of bacterial culture.
2. Centrifuge at 6,000 x g for 15 minutes at 4°C and discard the supernatant.
3. Resuspend the bacterial pellet in 10 mL of Resuspension buffer.
4. Add 10 mL of Lysis buffer, mix by inversion and incubate 5 minutes at room temperature.

   * If LyseBlue reagent has been added to the Resuspension buffer the solution should turn blue.
5. Add 10 mL of prechilled Neutralization buffer, mix by inversion and incubate on ice for 20 minutes.

   * If using LyseBlue reagent, mix the solution until it is colorless.
6. Centrifuge at ≥ 20,000 x g for 30 minutes at 4°C.
7. Meanwhile, equilibrate a Qiagen-tip with 10 mL of Buffer QBT and allow column to empty by gravity flow.
8. Apply the supernatant from step 6 to the Qiagen-tip and allow it to enter the resin by gravity flow.
9. Wash twice with 30 mL of Buffer QC.
10. Elute DNA with 15 mL of Buffer QF into a clean 50 mL falcon tube.

11. Precipitate DNA by adding 10.5 mL (0.7 volumes) room-temperature isopropanol to the eluted DNA and mix.
12. Centrifuge at ≥ 15,000 x g for 30 minutes at 4°C.
13. Wash the DNA pellet with 5 mL of room temperature 70% ethanol.
14. Centrifuge ≥ 15,000 x g for 10 minutes.
15. Air-dry the pellet for 5-10 minutes and redissolve DNA with 200 µL of milli-Q water.
16. Quantify DNA using Nanodrop
17. Store at -20°C.

1.8.4 BACMIPREP

Material

- QIAprep Spin Miniprep Kit containing:
    - Resuspension Buffer P1
    - RNAse A
    - Lysis Buffer P2
    - Neutralization Buffer N3
    - Qiagen ionic columns with coupled collector tubes
    - Washing buffer
- Isopropanol
- 70% ethanol
- Milli-Q water

Protocol

1. Centrifuge cells at 16.000 g for 5 minutes and discard the supernatant.
2. Resuspend the pellet in 400 µL of P1 Buffer.
3. Add 400 µL of P2 Lysis Buffer, invert the tube 10 times and incubate at RT for 5 minutes.
4. Add 400 µL of N3 Neutralization Buffer and invert the tube 10 times, gently.
5. Centrifuge at 16.000 g for 25 minutes at 4°C and transfer the supernatant to a fresh Eppendorf tube.
6. Centrifuge the supernatant again at 16.000 g for 15 minutes at 4°C.
7. Split the supernatant into two tubes and add 700 µL of isopropanol to each tube.
8. Place the tubes on ice for 30 minutes in order to precipitate the bacmid DNA.
9. Centrifuge at 16.000 g for 15 minutes at 4°C.

10. Carefully, remove the supernatant without disturbing the pellet and add 1 mL of 70% ethanol to the tube.
11. Invert several times and centrifuge at 16.000 g for 15 minutes at 4°C.
12. Remove the supernatant and let the pellet dry (for no longer than 5 minutes).
13. Resuspend in 100 µL of TE buffer.
14. Quantify DNA using the Nanodrop and store at 4ºC.

1.9 RESTRICTION ENZYMES DIGESTION

Restriction endonucleases are enzymes that recognize a specific DNA sequence and cleave it. They were first discovered in bacteria, and they are commonly used in research, especially in DNA cloning.

In this thesis, restriction enzymes have been used as control digestion analyses, aimed to validate the specificity of the produced DNA. To do so, firstly a vector map is created using computer software (VectorNTI). Once the map is designed with all its features, a restriction analysis is virtually simulated where we obtained different size bands as the result of the restriction enzyme cleaving our vector. In the laboratory, the results are validated by reproducing the reaction and running and agarose gel for the digested samples.

Material

- Construct DNA
- Digestion enzymes and buffers (New England Biolabs)
- Milli-Q water

Protocol

1. Design the restriction reaction using Vector NTI software
   * When possible, choose restriction enzymes that cut in the middle of the insert and in an insert-free area of the vector.
   * NEB supports an online website to check the compatibility of different restriction enzymes activity, as well as the optimal buffer for the reaction: Double Digest Finder (https://nebcloner.neb.com/#!/redigest)
2. Switch on the thermal dry bath at 37°C.

3. Prepare the mix for each construct to digest
    - 600 ng DNA
    - 0.5 µL per restriction enzyme
    - 1 µL buffer
    - Milli-Q water

      Final Volume 10 µL
4. Incubate 1 hour in the thermal dry bath.
5. Run an agarose gel (as described in *Section 1.2.1*) at 100 V for 45 minutes approximately.

   * The running time will depend on the size of the expected bands.
6. Acquire photo of the gel.

1.10 DNA SEQUENCING

To verify that the generated construct is correct and after the restriction enzyme validation, the construct is sequenced. DNA sequencing has been performed at the Genomics Facility of Centres Científics i Tecnològics de la Universitat de Barcelona (CCIT-UB).

Material

- Primers
- DNA
- Thermal cycler

Protocol

1. Add 150-600 ng of DNA with 4 µL of the primer at 1 µM into a tube.
2. Dry the tubes in the thermal cycler programmed for 20 minutes at 80°C.
3. Send the samples at the Genomics Facility of CCIT-UB.

2. CELL BIOLOGY TECHNIQUES

HEK293 and HeLa cells have been used for different techniques. HeLa cells come from human cervix adenocarcinoma, which is characterised for the presentation of an epithelial phenotype and for the incorporation of human papillomavirus 18 (HPV-18) sequences. In contrast, HEK293 cells derive from human renal epithelia, and they have been transformed with the adenovirus E1A gene (Graham et al., 1977). HEK293T cell line is derived from a HEK293 line that is expressing the antigen T from SV40 virus, which enables the episomal plasmid replication of those plasmids containing an origin of replication and an early SV40 promoter (Alwine, 1985). On the other hand, HEK2936E cell line, used for big scale expressions, expresses the Epstein-Barr Virus Nuclear Antigen-1 (HEK293-EBNA1 or HEK293E) (Tom et al., 2008).

2.1 GENERAL CONDITIONS FOR ADHERENT HELA AND HEK293T CELLS

Manipulation of cell lines is performed under sterile conditions in fully equipped laminar flow cabinets. All material is one-usage or previously sterilized. Cells are grown at 37°C, 5% CO_2 and 90% relative humidity. The cell media used is Dulbecco's Modified Eagle Medium (DMEM, Biological Industries) supplemented with 1% glutamine, 1% penicillin/streptomycin antibiotic solution, and 10% Fetal Bovine Serum (FBS). FBS is previously treated for 30 minutes at 56°C in order to inactivate the complement and antibodies.

2.2 GENERAL CONDITIONS FOR SUSPENSION HEK293-6E CELLS

Manipulation of cell lines is performed under sterile conditions in fully equipped laminar flow cabinets. All material is one-usage or previously sterilized. Cells are grown at 37°C, 5% CO_2, 90% relative humidity and 125 rpm. The cell media used is Freestyle F17 (Life Technologies) supplemented with 2% L-glutamine, 0.1% pluronic F68 (Applichem) and 0.05% G418.

2.3 TRYPSINIZATION

Once cells are confluent, they are detached by the addition of trypsin and seeded at a lower density for their optimal growth. Trypsinizations are routinely done twice per week.

Material

- Trypsin-EDTA 1x (Biological Industries)
- Complemented DMEM media
- PBS 1x sterile
- Neubauer chamber

Protocol

1. Wash cells carefully with 10 mL of sterile PBS.
2. Remove the PBS and add 1 mL of trypsin.
3. Incubate cells at 37°C for a couple of minutes, during this time, check their detachment.
4. Once detached, add 9 mL of complemented media to stop the trypsinization, resuspend up-and-down for a homogeneous solution.
5. Count cells with the Neubauer chamber

2.4 FREEZING

Material

- Complemented DMEM media
- Trypsin
- Sterile dimethyl sulfoxide (DMSO)
- PBS 1x
- Cryotubes

Protocol

1. Trypsinize cells and resuspend them in complemented DMEM media.
2. Centrifuge cells 5 minutes at 200 x g and aspirate the supernatant.
3. Resuspend the pellet with complemented media containing 10% of DMSO.
4. Aliquot 1 mL of cells in each cryotube.
5. Store the tubes at -80°C between 10 and 16 hours.
6. Later, store the tubes at -80°C or a liquid nitrogen tank.

2.5 THAWING

Material

- Complemented DMEM media
- Cryotubes with frozen cells
- 10 cm cell culture plates

Protocol

1. Thaw cells in a 37°C bath.
2. Put cells into a falcon with 10 mL of media.
3. Centrifuge cells 5 minutes at 180 x g and discard the supernatant.
4. Resuspend the pellet with 10 mL of media and seed in a 10 mm plate.
5. Put the plates in the incubator for cells to grow.

2.6 TRANSIENT TRANSFECTION

2.6.1 TRANSIENT TRANSFECTION IN ADHERENT CELL LINES

Material:

- HeLa or HEK293T cells
- DMEM Supplemented
- OptiMEM Glutamax media (Gibco)
- Sterile PBS 1x
- Transfectin Lipidic Reagent (Bio-Rad)
- 6-well culture plates
- cDNAs constructs

Protocol:

1. Trypsinize cells and seed at a 35% confluence (approximately 400,000 cells/well) and keep them for 24 hours in cell incubator to reach a 70-80% confluence for an optimal transfection.
2. Mix 3 µg of cDNAs with 250 µL of OptiMEM.
3. Mix 3 µL of Transfectin Lipidic Reagent with 250 µL of OptiMEM.
   * If transfecting 10 mm Petri plates, mix 11 µg of cDNAs with 1.5 mL of OptiMEM for each dish and 11 µL of transfectin with 1,5 mL of OptiMEM.

4. Mix the DNA-OptiMEM with Transfectin-OptiMEM and incubate for 20 minutes at room temperature.
5. Meanwhile, wash cells with PBS 1x and add 1 mL of OptiMEM in each well.
6. Add the DNA-Transfectin into each well and incubate for 4 hours at 37°C.
7. Discard the media and add 2 mL of DMEM in each well.
8. Incubate cells for 48 hours before performing the experiment.

2.6.2 TRANSIENT TRANSFECTION IN SUSPENSION CELL LINE

Material

- HEK293-6E cells at a concentration of $1.8 \cdot 10^6$ cells/mL
- Trypan Blue
- Counting slides
- Automatic cell counter
- 50 mL Falcon tubes
- F-17 complemented media
- PEI MAX 1 mg/mL
- Sterile flasks
- cDNAs constructs

Protocol

1. Count cells mixing one part of trypan blue with one part of cell culture in a counting slide; use an automatic cell counter.
2. Prepare two falcons, in one put media plus the DNA, and in the other the media plus PEI. The ratio used for PEI: DNA is 4:1.
   For 1 L cell culture, prepare two falcons with 50 mL of media each. In one, add the total amount of 1 mg of DNA and, in the other one, add 4 mL of PEI.
3. Mix the falcons and incubate 3 minutes at RT.
4. Pour the mix into the cell culture.
5. Incubate cells for 48 hours.

Considerations

Do not dilute cells above $0.7 \cdot 10^6$ cells/mL. The maximum concentration of cells in any Erlenmeyer flask is around $3 \cdot 10^6$- $4 \cdot 10^6$ cells/mL. The maximum volume of any flask is 20% of its capacity. The amount of transfected DNA is 1 µg per 1 mL of media.

2.7 PRODUCTION OF BACMAM VIRUS

2.7.1 TRANSFECTION IN SF9 CELLS

Material

- Sf9 cell culture
- Bacmid DNA
- PEI max
- 150 mM NaCl
- FBS

Protocol

1. Prepare solution A and B:
    a. Solution A: dilute 20 µg of Bacmid DNA into 0.5 mL of 150 mM NaCl.
    b. Solution B: dilute 80 µL of PEI into 0.5 mL of 150 mM NaCl.
2. Mix solution A and B and incubate for 30 minutes at RT.
3. In parallel, grow 10 mL of Sf9 cells until they reach a density of 1-1.2 cells/mL.
4. Add the transfection mixture to the cell culture.
5. Check fluorescence and cell viability.
6. After 5-7 days post-transfection there should be a high range of infected cells.
7. Centrifuge cells at 4,000 rpm during 10 minutes.
8. Collect the supernatant.
9. Add 10% of FBS and store at 4°C protected from light. This is the P0 virus.

2.7.2 P1 VIRAL PRODUCTION

Material

- Sf9 cell culture
- P0 virus
- FBS

Protocol

1. Grow 30-35 mL of Sf9 cell culture at $1.2 \cdot 10^6$ cells/mL.
2. Infect the culture with 1% of P0.

3. After 5-7 days, collect the media.
4. Centrifuge at 4,000 rpm during 10 minutes at RT.
5. Collect the supernatant and store it at 4°C with 10% of FBS. This is the P1 virus.

2.7.3 P2 VIRAL PRODUCTION

Material

- Sf9 cell culture
- P1 virus
- FBS

Protocol

1. Grow 500 mL of Sf9 cell culture at $1.2 \cdot 10^6$ cells/mL.
2. Infect the culture with 1% of P1.
3. After 5-7 days, collect the media.
4. Centrifuge at 4,000 rpm during 10 minutes at RT.
5. Collect the supernatant and store at 4°C. This is the P2 virus.

2.8 MAMMALIAN CELL INFECTION

2.8.1 VIRUS PRECIPITATION

Material

- P2 virus
- PEG solution (for a final volume of 500 mL):
    - 120 g PEG6000 + 320 mL H_2O
    - 40 mL NaCl 5M
    - 20 mL HEPES
    - pH 7.4
    - Autoclave

Protocol

1. Centrifuge P2 virus at 4,000 rpm for 10 minutes.
2. Add ¼ volume of 4xPEG to the P2 virus, mix and store at 4°C O/N.

3. Centrifuge the sample at 4,000 rpm for 30 minutes at 4°C.
4. Discard the supernatant. The pellet with the virus should be white/greyish.
5. Discard all PEG solution and resuspend the virus in Freestyle media at 37°C.
   * The stock should have an OD_{600} of 0.3 or higher. To achieve so, usually the pellet is resuspended in 10-15 mL of media.
6. Store the virus at 4°C protected from light.

2.8.2 CELL INFECTION

2.8.2.1 INFECTION IN SUSPENSION CELLS

Material

- HEK293-6E suspension cells
- Freestyle media
- P2 virus at OD_{600} 0.3
- 1M sodium butyrate in PBS

Protocol

1. Cells should be at a density of $2.5 \cdot 10^6$ cells/mL and the incubator temperature should be at 30°C.
2. Infect cells at 1.5% vol/vol.
3. Add 5 mM sodium butyrate at the moment of infection.
4. Depending on the protein expression, collect the cell culture at 36, 48 or 72 hours.

2.8.2.2 INFECTION IN ADHERENT CELLS

Material

- HEK293T cells
- DMEM + 10% FBS
- DMEM without FBS
- P2 virus at OD_{600} 0.3
- 1M sodium butyrate in PBS

Protocol

1. Plate cells the previous day with DMEM 10% FBS.
2. At the moment of infection, change media to DMEM without FBS.

3. Set the incubator temperature at 30°C.
4. Infect cells at 1.5% vol/vol and add 5 mM sodium butyrate.
5. Depending on the protein expression, collect the cell culture at 36, 48 or 72 hours.

2.9 IMMUNOCYTOCHEMISTRY

2.9.1 SEEDING FOR IMMUNOCYTOCHEMISTRY

Material

- HeLa cells in 6-well plates
- Sterile PBS 1x
- DMEM supplemented
- Trypsin
- Neubauer chamber
- Coverslips
- Surgical tweezers
- 48-well plates

Protocol

1. Wash cells with PBS 1x.
2. Tyrpsinize cells with 200 µL of trypsin per well.
3. Add 800 µL of DMEM to stop trypsinization.
4. Count cells with the Neubauer chamber.
5. Seed 100,000 cells into a 48-well plate with a coverslip in each well.
6. Add media for a final volume of 400-500 µL.

2.9.2 IMMUNOCYTOCHEMISTRY

Immunocytochemistry assays are performed to study the subcellular localization of different proteins. In this assay, specific primary antibodies detecting the protein of interest have been used. These antibodies are detected by fluorescence-labelled secondary antibodies. Its fluorescent signal is observed under the microscope.

Material

- Cultured cells on 10 mm coverslips
- 24-well plate
- Paraformaldehyde (PFA) 4%
- PBS 1x
- Blocking solution PBS 1x-FBS 10%
- Blocking and permeabilizing solution PBS 1x- FBS 10%
- Primary antibodies
- Secondary antibodies
- Mounting medium Vectashield (VECTOR) with DAPI 1.5 µg/mL
- Surgical tweezers
- Glass micro slides (26 x 76 mm)
- Nail polish

Protocol

1. Discard the cell media and wash twice with PBS 1x.
2. Fix cells with 250 µL PFA 4% and incubate 20 minutes at room temperature (shaking is optional).
   * PFA is toxic by skin contact and inhalation, working under the hood is strongly recommended.
3. Discard the PFA and wash twice with PBS 1x.
   * Fixed cells can be stored in PBS 1x- sodium azide 0.05% at 4°C.
4. Incubate the cells with blocking solution for 1 hour.
5. Incubate with primary antibodies diluted in blocking solution at RT for 1 hour
   * The incubation can also be left O/N at 4°C.
6. Wash with PBS 1x 3 times for 10 minutes each wash.
7. Incubate with secondary antibodies diluted in blocking solution at RT for 1 hour protected from the light.
8. Wash with PBS 1x 3 times for 10 minutes each wash.
9. Mount the coverslips onto glass micro slides with Vectashield + DAPI mounting medium.
10. Remove the excess mounting media solution carefully and seal the coverslips with nail polish.
11. Once the glass slides are dried, store them at 4°C protected from light.
12. Acquire images using a Cell R Olympus DSU spinning-disk microscope.

2.9.3 DATA ANALYSIS

Immunofluorescence experiments were performed on HeLa cells and analysed manually using ImageJ software. The experiments done in this thesis were to assess protein distribution and enrichment in tight junctions.

To determine whether a protein was enriched in tight junctions, images of two fluorescent cells were acquired. Its analysis is based on the comparative quantification of the fluorescent signal between cell-cell contacts (FC) and the plasma membrane of each cell (F1, F2). When FC > F1 + F2, its concluded that the protein being analysed is enriched at cell-cell junctions (Figure 28).

Material

- Microscope images
- ImageJ software

Protocol

1. Open images in ImageJ software.
2. Select *Straight* box to draw a straight line crossing a pair of cells.
3. Calculate the fluorescence profile following the command: *Analyse* > *Plot Profile*
   If FC > F1 + F2, the immunolabeled protein is enriched at cell-cell junctions.

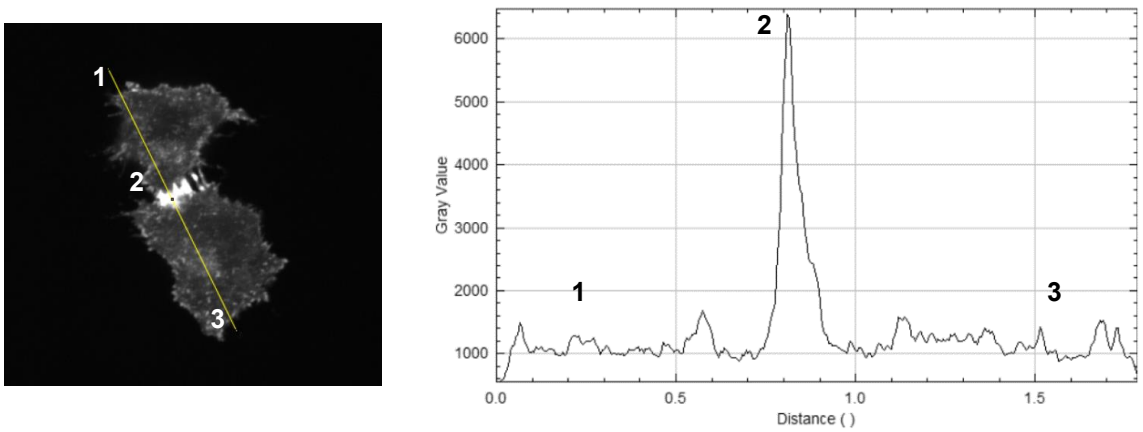


Figure 28. Analysis of cell-cell junction localization using ImageJ command.

2.9.4 COLOCALIZATION

Material

- Microscope images
- ImageJ software
    - The plugin Colocalization Finder is needed

Protocol

1. Open images in ImageJ software.
2. Convert images to 8-bit.
3. Go to Plugin -> Colocalization Finder.
4. Select the cells of interest in the new window, use the freehand selection.
5. Selection area is logged into a results window with a scatter plot where the Pearson's correlation is calculated.

2.10 NANOBIT

NanoBiT System is a complementation technique used to reveal protein-protein interactions. The method is based on the NanoLuc (NLuc) luciferase which is a small enzyme engineered for the function as a luminescent reporter. Luciferases are commonly used to monitor gene expression due to their broad dynamic range and their sensitivity. The enzyme is divided into two fragments: LgBiT and SmBiT (for large and small NanoBiT, respectively). LgBiT fragment weighs 17.6 kDa and SmBiT, 1.3 kDa. Upon interaction, active luciferase is reconstituted in the presence of substrate (Figure 29).

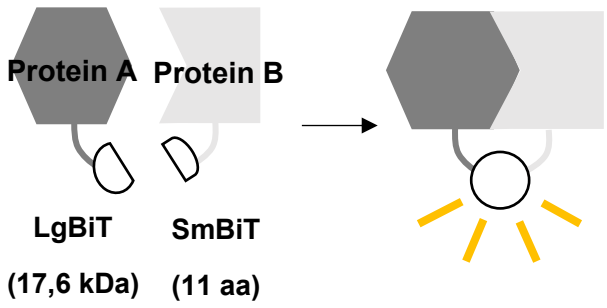


Figure 29. Schematic representation of NanoBiT interaction between two proteins. Each protein is fused to either largeBiT (LgBiT) or smallBiT (SmBiT) fragment. The interaction between the two fusion proteins leads to the complementation of the luciferase, restoring its function and eliciting reporter luminescent signal.

Two NanoBiT assays were performed, the first one analysed the activity of NanoLuc luciferase by adding coelenterazine as the substrate and it was used for the splits experiments. While the second one evaluated the activity of two luciferase enzymes and it was employed for all the other NanoBiT assays.

Material

- Transfected HEK-293T cells
- PBS 1x
- Coelenterazine H
- Luminometer

Protocol

1. Detach cells manually by adding PBS 1x and transfer the volume to Eppendorf tubes.
2. Centrifuge 5 minutes at 200 x g.
3. Wash once with PBS 1x.
4. Centrifuge 5 minutes at 200 x g.
5. Resuspend each tube with 60 µL of PBS 1x.
6. Switch on the luminometer.
7. Add 40 µL of coelenterazine at the moment of the lecture (one by one).

2.10.1 DUAL NANOBIT ASSAY

One of the main limitations of the just described NanoBiT assay was the variability among cell cultures and conditions. With that assay, we were assuming that the transfection would be equal in all conditions, which is hardly ever the case. Since we wanted to use this technique for signalling studies that required the transfection of more constructs, the described variability was a major setback. For this reason, we decided to work with a dual-luciferase assay.

The Nano-Glo Dual- Luciferase Reporter Assay (Promega) was employed. This kit uses the firefly (*Photinus pyralis*) luciferase as the constitutively expressed control reporter and *Renilla* luciferase as the experimental reporter. Dual assays are helpful to normalize the experimental luminescence with the co-transfected control, as they reduce experimental variability arising from transfection efficiency, cell number, cell viability, temperature, and measurement time. The kit allows for sensitive detection of both luciferases in the same sample. First, the firefly signal is read by adding the reagent

needed and, then, the second reagent is added where the firefly signal is quenched while providing the substrate needed for the second luciferase.

Material

- Transfected HEK293-T cells
- PBS 1x
- Nano-Glo Dual Luciferase Assay (Promega)
- 96-well white plates

Protocol

1. Equilibrate plates and solutions at room temperature.
2. Discard the cell media and resuspend cells in 1 mL of PBS 1x.
3. Add 50 µL of the resuspended cells to a 96-well white plate.
   * Per each condition, prepare triplicates on the plate.
4. Add 50 µL of ONE-Glo Ex reagent per well.
5. Incubate the samples for at least 3 minutes to allow cells to lyse.
6. Measure firefly luciferase activity in the luminometer.
7. Add 50 µL of NanoDLR Stop&Go Reagent.
8. Incubate the plate on an orbital shaker at 600-900 rpm for at least 3 minutes.
9. Incubate 10 minutes without shaking.
10. Measure NanoLuc luciferase activity in the luminometer.

2.11 FICOLL SEPARATION OF PBLs

Peripheral blood leukocytes (PBLs) were isolated from blood samples by using the Ficoll cell separation method. Ficoll is a synthetic polysaccharide used to create a density fluid gradient that allows the separation of different cell types (Figure 30).

Material

- PBS 1x sterile
- Ficoll
- Neubauer chamber
- Pasteur pipette
- Falcons 15/ 50 mL
- Cryotubes

Protocol

1. Put the blood sample in a 15 mL Falcon tube and add the same volume of PBS.
2. Homogenise the mixture.
3. In a different falcon, add Ficoll (the same volume as the blood sample) at room temperature.
4. Pour the blood/PBS into the Ficoll. The Falcon tube has to be inclined and the first droplet has to fall slowly into the wall of the tube in order to obtain two differentiated phases.
5. Centrifuge at 2,000 rpm for 20 minutes without break.
6. Collect the mononuclear cells with a Pasteur pipette, without taking the Ficoll.
7. Put the cells into a Falcon with 50 mL of PBS.
8. Homogenise.
9. Count cells with the Neubauer chamber.
10. Centrifuge at 1,500 rpm for 10 minutes without break.
11. Resuspend with 2 mL of sterile PBS.
12. Centrifuge at 3,000 rpm for 5 minutes.
13. Remove the supernatant.
14. Freeze cells in cryotubes.
15. Store cells at -80°C.

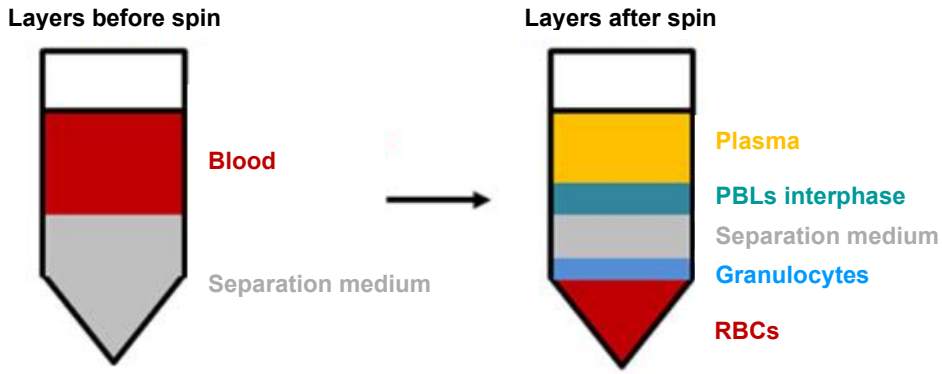


Figure 30. Separation of blood components by density gradient centrifugation. PBLs are enriched in an interphase on top of the separation medium (Ficoll) by density gradient centrifugation. Layers before (left) and after (right) centrifugation are shown. PBL: peripheral blood leukocytes, RBC: red blood cells. Adapted from (Posch et al., 2016).

3. PROTEIN STRUCTURAL TECHNIQUES

3.1 PROTEIN EXTRACTION

Most of the work carried out in this thesis was done in transfected cell lines. To study their protein levels, protein extraction and quantification needed to be performed. The extraction consists of the cell lysis by chemical and mechanical means. Total protein levels are quantified using BCA commercial kit.

Depending on the objective, two different procedures for protein extraction were followed. In adhesion cells, total protein extracts were obtained following standard protocols, whereas in suspension cells only membrane proteins were extracted, and they were used for membrane protein purification.

3.1.1 TOTAL PROTEIN EXTRACTION for adhesion cells

Material

- Cell culture
- PBS 1x
- Lysis buffer: protease inhibitors (Aprotinin and PMSF at 1 mM, Leupeptin and Pepstatin at 2 µM), NaCl 150 mM, 1% Triton X-100, PBS 1x.
- Cell scraper

Protocol

1. Wash twice the cells with PBS 1x.
   * From this point, all the protocol is performed at 4°C to avoid protein degradation.
2. Remove the PBS and add the Lysis buffer (0.8-1 mL in 10 mm plates, 100-120 µL in 6-well plates).
3. Scrape the cells with the cell scraper and resuspend the cell lysate.
4. Incubate the cell lysate in agitation for 1 hour at 4°C.
5. Centrifuge the cell lysate at 14,000 rpm for 10 minutes at 4°C.
6. Recover the supernatant and store it at -80°C.
7. Quantify total protein using the BCA Protein Assay Kit.

3.1.2 MEMBRANE PROTEIN EXTRACTION for suspension cells

Material

- Cell culture
- PBS 1x
- Lysis Buffer: 10 mM Tris pH 8; 10 mM NaCl; Complete mini EDTA-free protease inhibitors (Roche).
  * One tablet of Complete mini EDTA-free per 50 mL of lysis buffer.
- Solubilisation buffer without detergent: 100 mM Tris pH 8; 150 mM NaCl, 10% glycerol; Complete mini EDTA-free protease inhibitors.
  * One tablet of Complete mini EDTA-free per 50 mL of lysis buffer.
- Potter
- Ultracentrifuge
- Ultracentrifuge tubes and rotor

Protocol

1. Count cells with an automatic cell counter to know the concentration and cell viability.
2. Centrifuge the cell culture at 500 x g for 10 minutes at 4°C.
3. Discard the supernatant, wash the pellet with PBS 1x and centrifuge at 500 x g for 10 minutes at 4°C.
4. Repeat the previous step.
5. Resuspend the pellet in Lysis Buffer and homogenize with the Potter 30 times.
6. Centrifuge at 4,000 x g for 15 minutes at 4°C.
7. Ultracentrifuge the supernatant at 200,000 x g for 1 hour at 4°C (50.2Ti rotor at 41,000 rpm or 70Ti at 44,000 rpm) and discard the supernatant.
8. Resuspend the pellet in solubilisation buffer without detergent and homogenize the membranes with the Potter 30 times.
9. Quantify the total membrane protein using the BCA Protein Assay Kit.
10. Freeze membranes in liquid nitrogen and store them at 80°C.

3.2 TOTAL PROTEIN QUANTIFICATION (BCA)

The BCA method is a biochemical assay that uses the bicinchoninic acid (BCA) for colorimetric detection and quantification of total protein. BCA is a stable, water-soluble compound capable of forming an intense purple complex with cuprous ion (Cu^{1+}) in an alkaline environment. In this environment, ion cupric (Cu^{2+}) becomes reduced to Cu^{1+} at complexation sites within protein molecules. Then, the coloured chromophore develops from the union between BCA and Cu^{1+}. The generation of this coloured complex is established between peptides bonds and copper atoms. The colour produced is stable and increases proportionally with protein concentration.

Material

- 96-well plate
- BCA protein assay kit (ThermoFisher Scientific)
- BSA 2 mg/mL
- ELISA microplate reader (Biotek)

Protocol

1. Prepare the standard and the samples.
2. Mix reagent A and B at 1:50 ratio following manufacturer's instructions.
3. Add 200 µL of the reagents mix into each well.
4. Incubate 30 minutes at 37°C.
5. Read the absorbance at 595 nm.
6. Calculate sample protein concentration using the values obtained from the standard line.

3.3 DETERGENT PREPARATION

During this thesis, we have worked with different detergents, being digitonin, DM, and DDM the most employed. In Table 4, all the detergents used are listed.

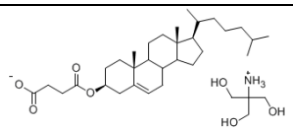
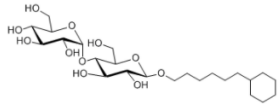
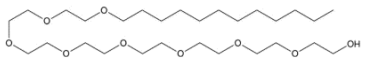
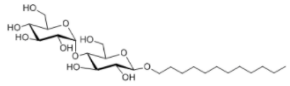
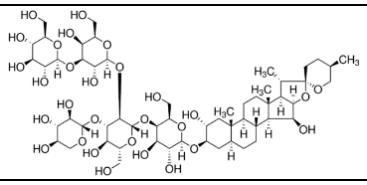
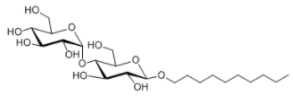
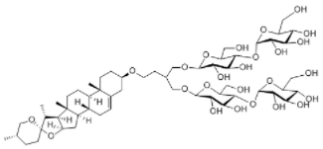
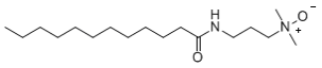
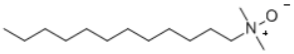
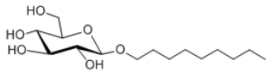
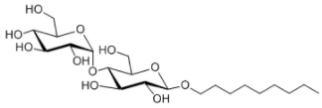
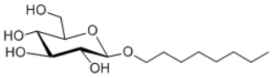
NAME	FULL NAME	C.M.C (%)	CHEMICAL STRUCTURE
CHS	Cholesteryl hemisuccinate		
CY-6	Cymal-6	0.028	
C12E8	Octaethylene Glycol Monododecyl Ether	0.0048	
DDM	n-Dodecyl-β-D-Maltopyranoside	0.0087	
Digitonin	Digitonin	0.02-0.03	
DM	n-Decyl-β-D-Maltopyranoside	0.087	
GDN	Glyco-diosgenin	0.0021	
LAPAO	3-Dodecylamido-N,N'-Dimethylpropyl Amine Oxide	0.047	
LDAO	n-Dodecyl-N,N-Dimethylamine-N-Oxide	0.023	
Nonylglucoside	n-Nonyl-β-D-Glucopyranoside	0.2	
Nonylmaltoside	n-Nonyl-β-D-Maltopyranoside	0.28	
Octylglucoside	n-Octyl-β-D-Glucopyranoside	0.53	

Table 4. Detergents used for detergent screening. Detergent full name, CMC and chemical structure are detailed.

3.3.1 DIGITONIN

Material

- Digitonin, High Purify (Merck)
- Buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water.
- Thermal dry bath
- 0.2 µm filter

Protocol

1. Set the thermal dry bath at 95°C and preheat the buffer 1 minute in the microwave.
2. Weigh the digitonin needed in a laminar flow hood.
   * We typically prepare digitonin at 5% concentration.
3. Add buffer and incubate 20 minutes at 95°C.
4. After this time, incubate on ice for 1 hour minimum.
5. Filter the digitonin and store at 4°C for immediate use or at -20°C.

3.3.2 OTHER DETERGENTS

Material

- Detergent (Anatrace)
- CHS (Anatrace)
- Buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water.
- Rotating wheel
- Sonicator

Protocol

1. Weigh the desired detergent.
   * We typically prepare detergents at 10% concentration.
2. Add buffer and leave it at the rotating wheel at RT until the detergent is dissolved.
3. Weigh and add CHS to a final concentration of 2%.
4. Sonicate until the solution is translucid.
   * Be careful to not heat the detergent. If the temperature increases, stop sonicating or do shorter cycles.
5. Store at 4°C for immediate use or at -20°C.

3.4 FLUORESCENCE SIZE EXCLUSION CHROMATOGRAPHY (FSEC)

FSEC is a technique widely employed for membrane protein expression screening. In general, the protein of interest is fused to GFP. Protein level is analysed by size-exclusion chromatography (SEC) linked to fluorescence detection using a HPLC instrument. In this thesis, FSEC was used to screen in a quick manner the solubilization capacity of different detergents and the putative ability of several nanobodies to specifically bind MLC1.

Material

- Transfected HEK293-6E
- PBS 1x
- Detergents
- Solubilization buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; 10% glycerol; 1% detergent; Complete EDTA-free protease inhibitors (Roche) in milli-Q water.
  * Use 1 tablet of Complete EDTA-free protease inhibitors per 50 mL of buffer.
- Rotating wheel
- Eppendorf ultracentrifuge
- Eppendorf Safe-Lock tubes
- Spin-X centrifuge tube filters, pore size 0.22 µm (Corning)
- Bottle-top vacuum filters, pore size 0.22 µm (Corning)
- ÄKTA buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; double CMC % of DDM in milli-Q water.
- Superose 6 Increase 30/100 GL column (Cytiva)
- ÄKTA machine
- 96-well black plates
- Fluorometer

Protocol

1. Centrifuge the cell culture at 200 x g for 10 minutes.
2. Wash twice with PBS 1x.
3. Resuspend the pellet with solubilization buffer without detergent. Divide the volume in parts.
4. Add the appropriate detergent in each part.
5. Incubate for 1 hour in the rotating wheel at 4°C.
6. Take 100 µL of each sample, put it in a 96-well black plate and read the fluorescence.

7. Centrifuge 1 hour at 200,000 x g for 1 hour at 4°C (Rotor TLA-55 at 55,000 rpm). * Use Eppendorf safe-lock tubes for the centrifugation.
8. Take 100 µL of each sample, put it in a 96-well black plate and read the fluorescence.
9. In the meantime, prepare the equilibration buffer, filter it with a bottle-top vacuum filter and equilibrate the ÄKTA as described in *Section 3.6.1*.
10. Once the ultracentrifuge has finished, filter the samples with centrifuge tube filters by centrifuging the samples at 3,000 x g during 5 minutes at 4°C.
11. Inject each sample to the ÄKTA as described in *Section 3.6.2* and collect the fractions in 96-well black plates.
12. Read the fluorescence.

3.5 PROTEIN PURIFICATION

Protein purification is a fundamental technique for analysing single proteins or protein complexes and identifying interactions with other proteins or performing functional and structural studies. The purification process consists of obtaining the protein alone in their physiological conformation and once it is purified, different assays can be done, such as obtaining its structure in cryo-electron microscopy (Cryo-EM).

The basis of protein purification is that the target protein is separated from other molecules by specifically binding to a ligand present in the resin. Therefore, as the sample is passed through the resin, the protein of interest is bound to it while the other components are washed away. Finally, by changing the buffer conditions, the binding between the ligand and the protein breaks.

To facilitate the purification procedure, affinity tags can be fused to the protein of interest. In this thesis we have worked with three different tags: TwinStrep, Flag and Polyhistidine tag depending on the analysed protein. The resins handled as well as their specific elution are detailed in Table 5.

In this thesis different protein purification protocols have been followed based on the objective of the designed experiment. Adherent cells were used when doing small protein purifications, but for larger volumes we worked with suspension cells. Both protocols are described hereunder.

RESIN	ELUTION
StrepTactin Superflow – (IBA LifeSciences)	25 mM desthiobiotin (IBA LifeSciences)
Pierce Anti- DYKDDDDK (Flag) – (ThermoFisher)	0.2 mg/mL Flag peptide (ThermoFisher)
CaptoL (Cytiva)	LSB 4x and incubate 10 minutes at 65°C
Ni-NTA – (ThermoFisher)	0.3 M imidazole

Table 5. Resins used for protein purification. Three different resins were used: StrepTactin, Flag and Niquel resins to purify proteins that incorporated the TwinStrep, Flag or Histidine tags, respectively.

3.5.1 PROTEIN PURIFICATION IN ADHERENT CELLS

Material

- Transfected HEK293T cells in Petri dish
  * Usually, one Petri dish is transfected per each working condition.
- Trypsin-EDTA 1x (Biological Industries)
- PBS 1x
- Detergent (see *Section 3.3*)
- Supplemented DMEM media
- Solubilization buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; 10% glycerol; 1% detergent; Complete EDTA-free protease inhibitors (Roche) in milli-Q water.
  * Use 2% of detergent for the purification with digitonin.
  * Use 1 tablet of Complete EDTA-free protease inhibitors per 50 mL of buffer.
- Resin of choice (see Table 5)
- Plastic columns
- Rotating wheel
- Equilibration buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl in milli-Q water.
- Wash buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; 2 x CMC detergent in milli-Q water.
- Elution buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; 2 x CMC detergent in milli-Q water plus the elution method of choice depending on the resin (see Table 5)

Protocol

1. Trypsinize cells with 1 mL of trypsin.
2. Once the cells are detached, stop de trypsinization with 1 mL of supplemented DMEM media.
3. Centrifuge 1 minute at 850 x g and discard the supernatant.
4. Wash with 1 mL of PBS 1x.
5. Repeat steps 3 and 4.
6. Discard the supernatant and solubilize the pellet with 2 mL of the solubilisation buffer.
7. Incubate 1 hour at 4°C in the rotating wheel.
   * For the purification of TwinStrep-MLC1 with digitonin, incubate 2 hours at 2% of detergent.
8. Centrifuge 1 hour at maximum speed of a benchtop centrifuge at 4°C.
   * Collect 50 µL of sample before and after the centrifugation in order to obtain the Total Protein and the Solubilization samples. For the Total Protein, centrifuge 5 minutes at maximum speed at 4°C before loading the WB.
9. Meanwhile, add 250 µL of StrepTactin-Superflow resin (considering 125 µL of beads) into the plastic column. Wash the resin three times with milli-Q water and one time with equilibration buffer leaving enough volume for the resin not to dry.
10. Once the centrifuge has finished, incubate the supernatant with the resin for 2 hours at 4°C in the rotating wheel.
11. After this time, put the column into a laboratory stand and collect the flowthrough.
    * To ensure that the protein always remains at 4°C, working inside the cold room is strongly recommended.
12. Wash three times with 250 µL of Washing buffer.
13. Elute five times with 250 µL of Elution buffer.
14. Store the samples at -80°C.
    * For double purifications, once the elution is obtained, it is incubated with the second resin and the same protocol is repeated from the step 10 onwards.
    ** For the CaptoL elution, after the sample is washed normally, LSB 4x is added to the wash buffer and the sample is heated at 65°C at a dry bath for 10 minutes. Then, it is recommended to centrifuge the sample to eliminate the resin.

3.5.2 PROTEIN PURIFICATION IN SUSPENSION CELLS

Material

- Membranes
- Detergent (see *Section 3.3*)
- Rotating wheel
- Ultracentrifuge tubes and rotor
- Ultracentrifuge
- Solubilisation buffer: 100 mM Tris pH 8; 150 mM NaCl; 10% glycerol; 1% detergent, Complete EDTA-free protease inhibitors in milli-Q water.
  * Use 2% of detergent for the purification with digitonin.
  * Use 1 tablet of Complete EDTA-free protease inhibitors per 50 mL of buffer.
- Plastic columns
- Resin of choice (see Table 5)
- Equilibration buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water
- Wash buffer: 100 mM Tris pH 8; 150 mM NaCl; detergent at 2 x CMC in milli-Q water
- Elution buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; 2 x CMC detergent in milli-Q water plus the elution method of choice depending on the used resin (see Table 5)

Protocol

1. Prepare the solubilization buffer.
2. Once the membrane sample is thawed, add the solubilization buffer and incubate 1 hour in the rotating wheel at 4°C.
   * For the purification of TwinStrep-MLC1 with digitonin, incubate 2 hours at 2% of detergent.
3. Ultracentrifuge samples at 200,000 x g for 1 hour at 4°C (Rotor 50.2Ti at 41,000 rpm or Rotor 70.1Ti at 55,000 rpm)
   * Take 50 µL before and after the centrifugation if WB of the process is desired.
4. On the meantime, wash the resin three times with milli-Q water and one time with equilibration buffer leaving enough buffer for the resin not to dry.
5. Once the ultracentrifuge has finished, incubate the supernatant with the resin for 2 hours in the rotating wheel at 4°C.
6. On the meantime, equilibrate the ÄKTA as described below (see *Section 3.6.1*)
7. Pour the sample with the resin into a plastic column. Collect the flowthrough.
8. Wash three times with 2 column volumes (CV) of wash buffer.

9. Elute 7 times with 1 CV of elution buffer.
10. Quantify protein concentration in all elution fractions using the Nanodrop (A280), mix the elution tubes that contain protein.

3.6 SIZE EXCLUSION CHROMATOGRAPHY

Size exclusion chromatography (SEC) is a technique that separates different components based on their molecular size. A column filled with resin, that can be made from different materials, is used and samples will go through it. Depending on its composition, the resin will form different size of pores. The sample is injected into the column and with the help of a flow rate, it will pass through the column. The biggest molecules will not be able to enter between the resin particles, as a result, they will pass in the column laterals and they will reach earlier the end of the column, where the UV detector is placed. In contrast, small molecules will be able to enter between the different resin pores and, accordingly, they will need more time to reach the end of the column (Figure 31).

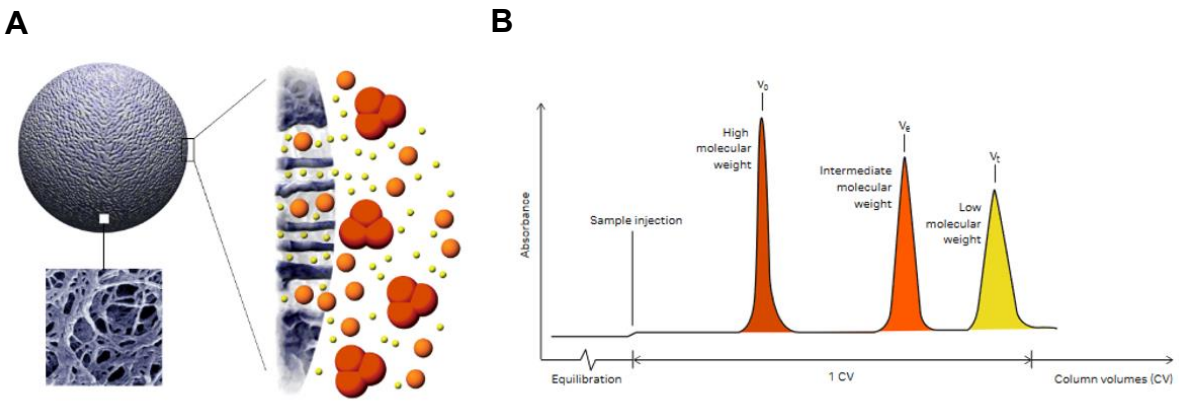


Figure 31. Size exclusion chromatography. A. Schematic picture of a particle with an electron microscopic zoom. Bigger molecules (in red) will not enter the pores of the resin and they will be eluted first, whereas the smaller molecules (in yellow) will be more delayed. **B.** Schematic chromatogram where high molecular weight molecules appear earlier than low molecular weight particles. Adapted from Cytiva – Size Exclusion Chromatography – Principles and methods.

Based on their composition, length and wide, a broad variety of columns exist. Changing the material of the resin or its length can modify the separation of different molecular weights. For this thesis, the following columns have been used: Superdex 200 Increase 10/300 GL, Superose 6 Increase 10/300 GL, and Superose 6 Increase 3.2/300 GL.

Superdex and Superose 6 resins present high reproducibility and low nonspecific interactions. Both have the same high-flow agarose base matrix with an average particle size of 8.6 µm. However, their fractionation range varies resulting in different chromatography profiles. Superdex 200 Increase resin consists of a base matrix made of dextran and agarose. Its fractionation range allows separation of proteins from 10,000 to 600,000 Da. In contrast, Superose 6 Increase resin consists of cross-linked porous agarose particles. It presents a fractionation range between 5,000 and 5,000,000 Da. So, in comparison, Superose 6 Increase displays a better separation between largest molecules than Superdex 200 Increase in which bigger molecules will appear in the void volume.

Column length is also an essential consideration for the success of protein separation. Columns 10/300 GL are larger and they are designed for high-resolution small-scale preparative purification, and protein characterization. In opposition, 3.2/300 GL columns are shorter and they are the first choice for rapid screening, protein characterization and working with very small sample volumes.

3.6.1 ÄKTA PREPARATION

Before working with our samples and inject them into the HPLC machine, some preparation is needed. The equipment is usually stored in 20% ethanol.

Material

- 20% ethanol
- Milli-Q water
- Equilibration buffer: 100 mM Tris pH 8; 150 mM NaCl; detergent at the double CMC in milli-Q water
- Bottle-top vacuum filters, pore size 0.22 µm (Corning)
- 5 mL syringe

Protocol (the day before)

1. Wash the pumps with milli-Q water. To do so:
    - Click on the system scheme *Pump A -> Start pump A wash*

   Once it finishes, press stop.
2. Clean the loop with milli-Q water with the help of a syringe.
3. Connect the column into the system and wash with milli-Q water by running a low flow that can be left overnight. To do so:
    - *System -> Execute Manual Instructions*
        - Set Precolumn Pressure at 2.5 MPa
        - Set Timer -> Volume at 100 mL
        - Set Flow rate at 0.1 mL/min

Protocol (day of the experiment)

1. Stop the flow set the previous day.
2. Prepare the equilibration buffer and filter it.
3. Repeat the cleaning process as described earlier for the pumps and the loop with the equilibration buffer instead of milli-Q water
4. Equilibrate the column by running the sample protocol:
    - *File -> Run -> Protocol*
5. Once the conductivity, the system pressure and the ultraviolet detector are stable, the machine is ready to use. Prepare new collection tubes for each new run.
6. Run a protocol for each sample that is going to be injected as described in *Section 3.6.2.*
7. After finishing the experiment, clean the pump, the loop and the column with milli-Q water as previously described. Run a low flow of milli-Q water O/N.

Protocol (the following day)

1. After finishing the experiment, clean the pump, the loop and the column in 20% ethanol.
2. Store the column at RT.

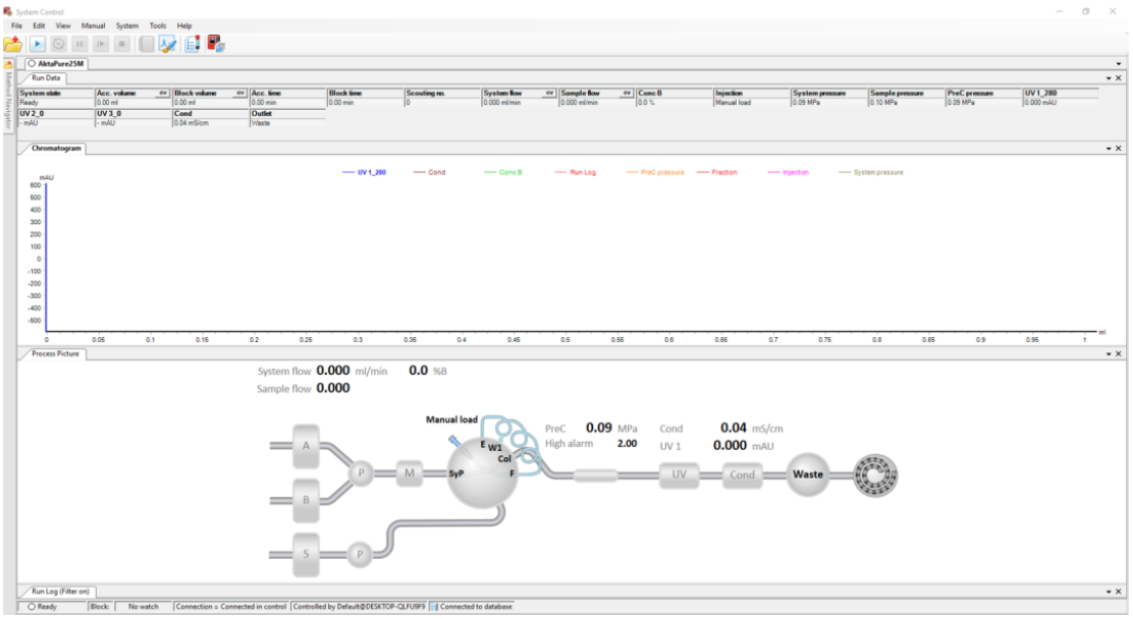


Figure 32. Control screen for the ÄKTA Pure 25M system.

3.6.2 SAMPLE INJECTION

Material

- Column filters Amicon Ultra Centrifugal Filter Devices 100kDa (Millipore)
- Equilibration buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl with milli-Q water
- Column buffer: 100 mM Tris-HCl pH 8; 150 mM NaCl; detergent 2 x CMC with milli-Q water
- Spin-X centrifuge tube filters, pore size 0.22 µm (Corning)
- SEC column
- 20% ethanol
- Milli-Q water

Protocol

1. Equilibrate the filter column with equilibration buffer.
2. Put the sample into the column.
3. Concentrate samples centrifuging 5 minutes at 3,000 x g at 4°C. Repeat as many times as necessary to obtain the desired volume to inject in the ÄKTA.
   * The sample volume should be half the volume of the ÄKTA loop.
4. Quantify protein concentration in the Nanodrop.
5. Filter the sample with Spin-X centrifuge tube filters by centrifuging 5 minutes at 3,000 x g at 4°C.

6. Quantify protein concentration in the Nanodrop.
7. Inject the sample into ÄKTA.

3.7 WESTERN BLOT

Western blot (WB) is one of the most common techniques in molecular biology to detect specific proteins from biological samples. It allows to gain knowledge about protein size and biochemical properties such as expression levels, states of oligomerization or phosphorylation. The technique is divided in three steps. Firstly, an SDS-PAGE electrophoresis separates proteins based on their molecular weight. Next, proteins are transferred from the gel onto a polyvinylidenedifluoride (PVDF) membrane. Lastly, specific antibodies are used to detect specific epitopes of the protein.

3.7.1 ACRYLAMIDE GELS

Material

- Mini-Protean Tetra Cell Casting Stand and Clamps (Bio-Rad)
- 1.5 mm gel glasses
- 10-well or 15-well comb for 1.5 mm gels
- Acrylamide/Bis-acrylamide 40% (Bio-Rad)
- Running gel buffer: 1.5 M Tris-base pH 8.8; 0.1% SDS
- Stacking gel buffer: 0.5 M Tris-base pH 6.8; 0.1% SDS
- Ammonium persulfate (APS) 10% in distilled water
- TEMED
- Isopropanol

Protocol

1. Mount the glasses into the gel support, fill with water and check for leakages.
2. Prepare the running and stacking gels without APS and TEMED (volumes detailed in Table 6).
3. Add APS and TEMED to the running gel, vortex the mix and pour it between the glasses.
4. Pour a little bit of isopropanol on top of the gel.
5. Wait until it polymerizes.
6. Remove the isopropanol.

7. Add APS and TEMED to the stacking gel, vortex, pour it and add the comb.
8. Wait until it polymerizes.

	RUNNING GEL (2 x 1.5 mm gels)					**STACKING GEL (2 x 1.5 mm gels)**
	7.5%	**10%**	**12%**	**15%**		**4%**
Acrylamide	3.8 mL	5 mL	6 mL	7.5 mL	**Acrylamide**	1 mL
Running gel buffer	5.2 mL	5.2 mL	5.2 mL	5.2 mL	**Stacking gel buffer**	2.5 mL
Water	11 mL	9.8 mL	8.8 mL	7.3 mL	**Water**	6.5 mL
APS	200 µL	200 µL	200 µL	200 µL	**APS**	200 µL
TEMED	20 µL	20 µL	20 µL	20 µL	**TEMED**	20 µL

Table 6. Acrylamide gels composition.

3.7.2 ELECTROPHORESIS

Material

- Mini-Protean Trans Blot system (Bio-Rad)
- Electrophoresis buffer 1x: prepared from 10 x diluted in distilled water.
  10x: 250 mM Tris-base; 1.92 M glycine; 1% SDS in distilled water.
- Loading sample buffer (LSB) 4x: 0.4 M Tris-base pH 6.8; 80% glycerol; 8% SDS; 0.005% bromophenol blue
- 10% β-mercaptoethanol or 100 mM DTT
- Page Ruler protein ladder (Bio-Rad)

Protocol

1. Dilute LSB 4x with β-mercaptoethanol or DTT into the protein samples.
2. Heat the samples.
   - At 56°C for 3 minutes when working with membrane proteins.
   - At 95°C for 5 minutes for cytosolic proteins or very stable membrane proteins.
3. Cast the gels into the buffer tank and fill it with electrophoresis buffer 1x.
4. Remove the well comb, load the protein ladder in the first lane and load the samples in the others.
5. Close the tank with the lid and connect it to the power supply.

6. Set the voltage at 100 V for 10 minutes, once the samples have passed the concentrator gel, set the voltage at 120 V until the desired level of protein migration taking the protein ladder as a reference.

3.7.3 TRANSFERENCE

Material

- Trans-Blot Turbo Transfer System (Bio-Rad)
- Whatmann paper 3 mm
- Western blot roller (Bio-Rad)
- Immobilon-P 0.45 µM PVDF transfer membrane (Millipore)
- Anode Buffer (Tris-Base 0.3 M, methanol 20% in distilled water)
- Cathode Buffer (Aminocaproic acid 40 mM, methanol 20% in distilled water)
- Methanol

Protocol

1. Hydrate PVDF membranes with methanol in agitation for 3 minutes.
2. Assemble the blotting cassette:
    a. Place 3 Whatmann papers previously soaked with Anode buffer to the transfer tray.
    b. Place on top the hydrated PVDF membrane previously soaked with Anode buffer.
    c. Place on top the acrylamide gel previously soaked with Cathode buffer.
    d. Place on top 3 Whatmann papers previously soaked with Cathode buffer.
    e. Press the transfer stack with a roller to remove air bubbles.
3. Set the cassette inside the Trans-Blot Turbo blotting instrument. Set the pre-programmed standard protocol for 1 or 2 mini gels: 25 V, 1 A, 30 minutes.

3.7.4 IMMUNODETECTION

Material

- Ponceau staining solution: 5% acetic acid; Ponceau S 0.1% in distilled water
- TTBS 1x washing solution: TBS 1x; 0.1% Triton X-100 in distilled water
- Blocking solution: 5% dry fat-free milk in TTBS 1x.
- Primary antibodies
- Secondary antibodies HRP-conjugated
- ECL solution A: 1 M Tris-base pH 9; 0.2 mM coumaric acid; 250 mM luminol in milli-Q water
- ECL solution B: 30% H_2O_2 in milli-Q water
- AI800 Imaging System (Amersham, GE)

Protocol

1. Stain transferred membranes with Ponceau solution to validate correct protein transference.
2. Wash with TTBS 1x several times, until the membrane is no longer red.
   * All washes are done in agitation.
3. Block membranes with blocking solution for 1 hour at RT.
4. Incubate the membranes with primary antibody diluted in blocking solution O/N at 4°C.
   * Alternatively, 1 hour incubation can also be performed.
5. Wash membranes with TTBS 1x three times for 10 minutes.
6. Incubate the membranes with secondary antibody diluted 1/5,000 in blocking solution 1 hour at RT.
7. Wash membranes with TTBS 1x three times for 10 minutes.
8. Incubate the membranes with a combination of ECL solutions (1 mL solution A + 25 µL solution B).
9. Develop signal using AI800 Imaging System.

3.7.5 DATA ANALYSIS

The images acquired from WB show protein detection and expression levels, although it is not fully quantitative. Using software data analysis (ImageJ), this data can be transformed into quantitative information. Intensity of bands can be measured and compared across samples.

Material

- WB images acquired using AI800
- ImageJ software

Protocol

1. Open the images in ImageJ.
2. Create a rectangular selection around the widest lane within our image.
3. Drag the selection to the first signal band from the membrane.
4. Go to *Analyze -> Gels -> Select First Lane.*
5. Move the selection to the following band.
6. Go to *Analyze -> Gels -> Select Second Lane.*
7. Repeat steps 5 and 6 for all bands.
8. Go to *Analyze -> Gels -> Plot Lanes.*
9. Create a straight line closing the area under the peak for each graph.
10. Select the magical wand menu and click on each graph.
11. Copy data to the numerical data analysis software.

3.8 SILVER STAIN

Silver staining of SDS-polyacrylamide gel electrophoresis is used for detecting proteins present in quantities of even nanograms. The principle of this technique is that selective reduction of silver into metallic silver occurs at the initiation site in the close proximity of protein molecules. The staininig process consists of protein fixation, sensitization, washing, silver impregnation, and image development (Kumar, 2018). In Figure 33 an example of silver stain gel is presented.

Material

- Acrylamide gel
- Pierce Silver Stain Kit (ThermoFisher Scientific)
- Ethanol
- Acetic acid

Protocol

1. Once the electrophoresis is finished, take the acrylamide gel out of the bucket.
2. Wash twice with milli-Q water for 5 minutes.
   * All washes are done in agitation.
3. Fix gel twice with 30% ethanol + 10% acetic acid solution for 15 minutes.
4. Wash twice with 10% ethanol for 5 minutes.
5. Wash twice with milli-Q water for 5 minutes.
6. Prepare Sensitizer Working solution with 50 µL of Sensitizer in 25 mL of milli-Q water.
7. Sensitize gel for 1 minute.
8. Wash twice with milli-Q water for 1 minute.
9. Prepare Stain Working Solution with 0.5 mL of enhancer in 25 mL of Stain Solution.
10. Stain gel for 30 minutes.
11. Prepare Developer Working Solution with 0.5 mL of enhancer with 25 mL of Developer Solution.
12. Wash gel twice with milli-Q water for 20 seconds.
13. Develop gel for 2-3 minutes or until bands appear.
14. Stop with 5% acetic acid for 10 minutes.
15. Scan the gel or take a picture.

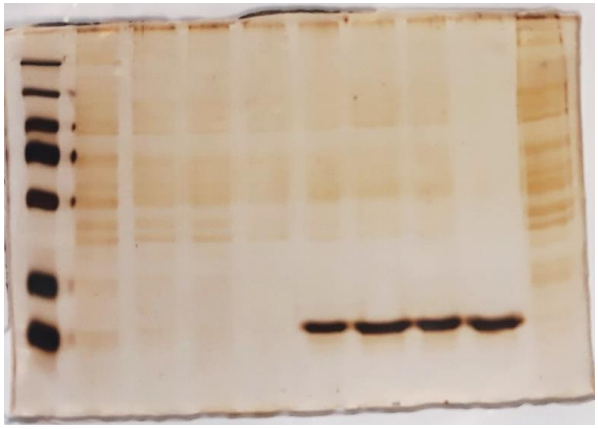


Figure 33. Example of a Silver Stain gel.

3.9 AMPHIPOL RECONSTITUTION

Amphipols are short amphipathic polymers that can substitute detergents to keep integral membrane proteins soluble in water (Tribet et al., 1996). Although they resemble detergents, amphipols can stabilize membrane proteins more efficiently by interacting firmly with the protein (Le Bon et al., 2018).

In this thesis, the work done with MLC1 reconstituted in amphipols had two different objectives. On one hand, an attempt was made to reconstitute the membrane protein in amphipols for the llama immunization to obtain nanobodies. On the other hand, MLC1 was reconstituted in amphipols for its posterior cryo-EM analysis.

Material

- A8-35 amphipols (Affymetrix Ltd)
- Purified protein
- Biobeads SM-2 adsorbents (Bio-Rad)
- Dialysis membrane 100 kDa (Fisher Scientific)
- Buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water

Protocol

1. Once the protein is purified, quantify its concentration using the Nanodrop.
2. Mix the purified protein with amphipols using a ratio 1:5 of protein: amphipols.
3. Incubate 30 minutes on ice.
4. Mix the sample with biobeads, considering 20 g of wet beads per g of detergent.
5. Incubate 1 hour at 4°C in mild agitation.
6. Rinse the dialysis membrane in milli-Q water.
7. Dialyze O/N at 4°C with a substantial volume of buffer.

3.10 PROTEOLIPOSOME RECONSTITUTION

Proteoliposomes mimic lipid membranes to which a protein has been incorporated. Membrane proteins reconstitution into liposomes is a method of choice for the investigation of membrane proteins because it allows both functional and structural studies (Puvanendran et al., 2020). Although liposomes organize spontaneously in aqueous solutions, they tend to form systems with many bilayers, called multilamellar. For the obtention of unilamellar systems similar to cell membranes, a process of sonication or extrusion needs to be performed (Ciancaglini et al., 2012).

Material

- Purified protein
- *E. coli* total lipid extract (Avanti Polar Lipids)
- Methanol
- Chloroform
- CHS
- Buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water
- 10% Octylglucoside (Anatrace) in milli-Q water
- Dialysis membrane 100 kDa (Fisher Scientific)
- Thermal dry bath

Protocol

1. Weigh the quantity of lipids needed for the experiment together with CHS.
   * Ratio of 1:10 CHS: lipids.
2. Add 1 part of methanol and 3 parts of chloroform to the lipids.
3. Evaporate the liquid with N_2 gas and let it dry for 2 hours at RT.
4. Resuspend the lipids with the buffer to a final concentration of 20 mg/mL.
5. Freeze the liposomes with liquid Nitrogen and thaw in a thermal dry bath at 26°C. Do 10 cycles of freezing/thawing to obtain unilamellar liposomes.
6. Add 1% of octylglucoside to the liposomes and incubate 30 minutes at 4°C.
   * Ratio of 1:50 protein: liposomes.
7. Add the purified protein to the liposomes and incubate 1 hour at 4°C.
8. Rinse the dialysis membrane in milli-Q water.
9. Dialyze the sample O/N in agitation at 4°C with enough volume of buffer.

3.11 SUCROSE GRADIENT

Sucrose gradient was carried out to validate the correct reconstitution of MLC1 into liposomes. The density gradient is created by layering sucrose solutions of different concentrations, creating a density gradient from the bottom to the top of the centrifuge tube.

Material

- Sucrose
- PBS 20 mM
- Swinging rotor SW41 and tubes

- Ultracentrifuge
- Buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water

Protocol

1. Prepare a sucrose gradient using serial dilutions of the following concentrations:

   65%, 45%, 35%, 28%, 23%, 18%, 13%, 8% and 3%

2. Gently, place the different concentrations into an ultracentrifuge tube starting with the 65%.
3. Mix 150 µL of 3% sucrose with 150 µL of the protein sample and place it on top of the gradient.
4. Centrifuge at 150,000 x g for 19 hours at 15°C with minimal acceleration and deceleration.
5. Resuspend the pellet with buffer, make aliquots and freeze in liquid nitrogen before storing them at -80°C.

3.12 NANOBODY OBTENTION

Nanobodies are functional, homodimeric heavy chain-only antibodies that are naturally produced by the immune system of Camelidae. Their stability, small size and easiness to produce have convert them into a powerful tool for multiple applications, including solving protein structures (Jin et al., 2023).

Through the immunization of a llama with our purified protein MLC1 in liposomes, the posterior isolation of PBLs, RNA extraction and synthesis of cDNA an immune library was generated which was cloned in phage display vector pMESy4. With these vectors, different rounds of phage display were performed to enrich those nanobodies that were specifically binding our target protein. The whole nanobody obtention process is summarized in Figure 34.

This work was performed at the Workshop Nanobody4Instruct Training Course funded by Instruct-ERIC that took place at VIB-VUB Centre for Structural Biology in Brussels in January 2023.

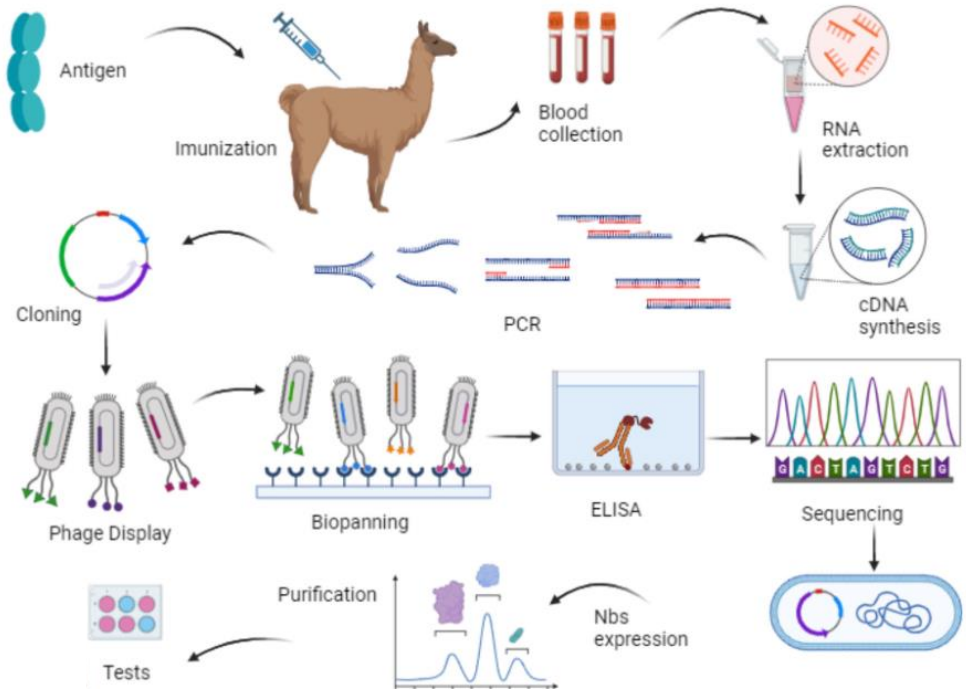


Figure 34. Schematic representation of nanobodies obtention. Adapted from (Minatel et al., 2023).

3.12.1 SELECTION BY PHAGE DISPLAY

Phage display is a method created by G. Smith for presenting polypeptides on the surface of bacteriophages (Smith, 1985). It allows for the creation of antibody libraries containing a great number of phage particles, from which each one encodes and displays a different molecule. This technique is used to study protein interactions with other proteins, peptides, or DNA (Bazan et al., 2012).

The principle of phage display sits on the insertion of a gene encoding a protein of interest into a phage coat protein gene, which causes the phage to display the protein on its capsid. Then, phages with the displayed proteins are screened against other molecules to analyse its interaction. The mostly used bacteriophage is M13 and, usually, antibodies and peptides are cloned into GIII gene, which codifies for the coat protein pIII (Jaroszewicz et al., 2022).

Two elements are typically used in phage display system: the phagemid vector and a helper phage. The first one contains the gene encoding the protein of interest next to GIII, an antibiotic marker for selection and the phage origin of replication. On the other hand, the helper phage contains the complete M13 genome with all the proteins needed for infection, replication, and assembly (Ledsgaard et al., 2018).

Once the library is created, the antigen is presented to the phage library together with the helper phage. Antigen presentation can occur in different forms, such as coating the antigen directly, coating with a capture element (i.e antibody anti-Twinstrep tag or antibody against the target protein) or using beads with a capture system (i.e StrepTactin beads). After the binding, different washes are done to remove non-binding phages and the ones binding are eluted. Then, the eluted phages are amplified so to start a new round of phage display. Each round of phage display panning ends with several eluted infective recombinant phages. When the amount of eluted phage is at least 100-fold higher than the eluted from the negative control, it is recommended to screen individual clones. Contrary, when the enrichment factor is below 100, another selection round is advised. Typically, one or two rounds of panning are sufficient to enrich for target-specific nanobodies (Pardon et al., 2014). In order to monitor progress, titration of phage outputs is performed. In Figure 35, a phage display scheme is displayed.

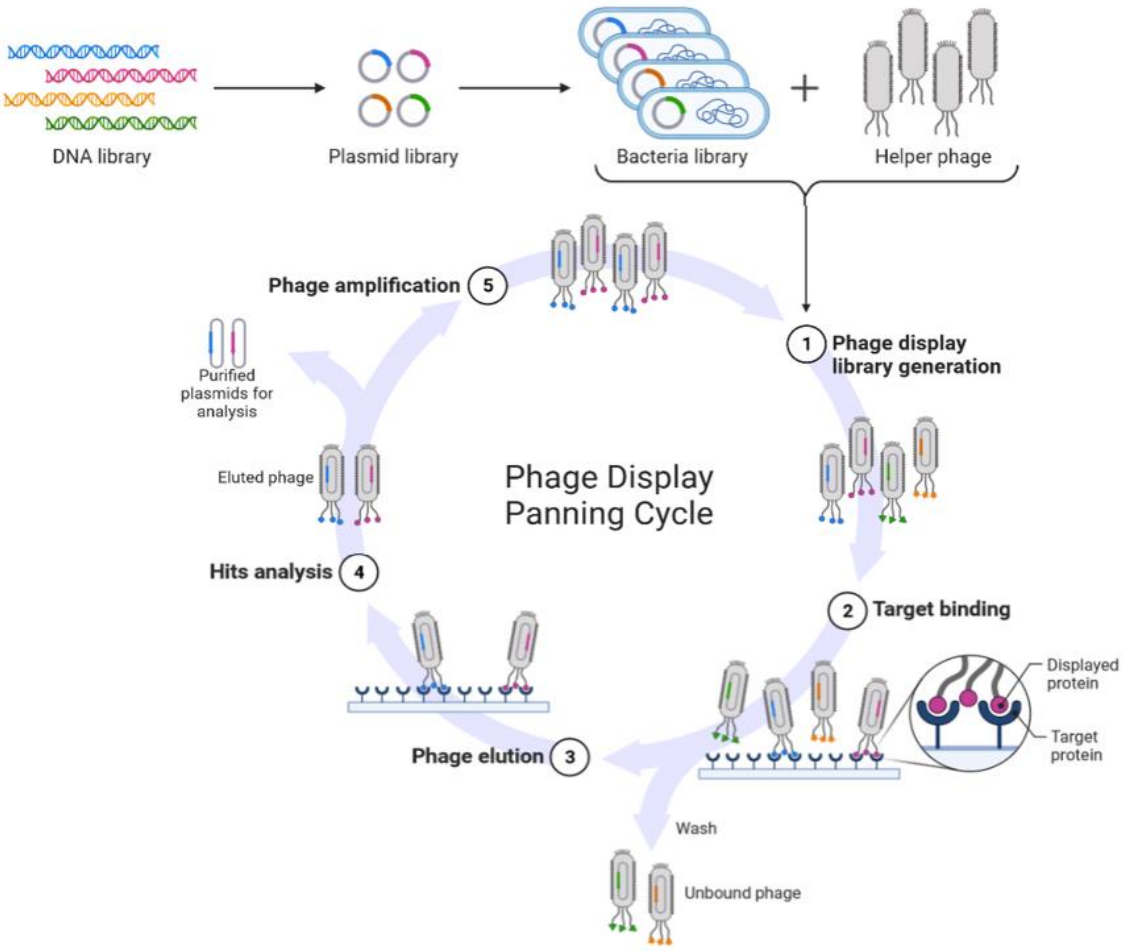


Figure 35. Phage display panning cycle. The process involves several steps. First, the ligand of interest is immobilized into a solid support, to which the phage library is added. After washing, the remaining bound phages are eluted and amplified. Each round of panning cycle increases the proportion of specific binding clones within the phage display library.

3.12.2 RESCUE AND AMPLIFICATION OF PHAGE FROM IMMUNE LIBRARIES

Before each round of panning, phage particles displaying nanobodies are rescued and amplified by adding helper phage. Once they are grown, they are put into contact with TG1 cells for phages to infect them. After an O/N incubation, the phage is isolated and is ready to be used for panning. For an estimation of the number of infective recombinant phage, a titration is done by preparing serial tenfold dilutions in culture plate that is grown O/N and colonies are counted the following day.

3.12.2.1 LIBRARY RESCUE

Material

- 2xTY
- 100 mg/mL ampicillin
- 25 mg/mL kanamycin
- 20% glucose
- VSCM13 helper phage

Protocol

1. Inoculate your library into 45 mL of 2xTY plus 50 µL 100 mg/mL ampicillin and 5 mL of 20% glucose.
2. Grow at 37°C at 150 rpm until OD_{600} is 0.5-0.6.
3. Take 10 mL of culture and add $2.5 \cdot 10^{10}$ pfu VSCM13, mix manually.
4. Incubate 30 minutes at room temperature without shaking.
5. Centrifuge 10 minutes at 2,000 rpm, remove the supernatant.
6. Resuspend the pellet in 60 mL 2xTY plus 60 µL 100 mg/mL ampicillin and 60 µL 25 mg/mL kanamycin.
7. Incubate O/N at 37°C at 180 rpm.

3.12.2.2 PREPARATION OF PHAGE PARTICLES

Material

- 2xTY
- PEG 6000
- 5 M NaCl
- TG1 cells

Protocol

1. Inoculate TG1 cells in 2xTY, grow at 37°C at 200 rpm.
2. Centrifuge the O/N grown culture 30 minutes at 4,000 rpm at 4°C.
3. Transfer 40- 42 mL of supernatant into a new falcon tube.
4. Mix 10 mL of 20% PEG6000/ 2.5 M NaCl with the supernatant.
5. Mix well by inverting the tube and keep on ice for at least 30 minutes.
6. Centrifuge 4,000 rpm for 30 minutes at 4°C.
7. Remove supernatant, dry pellet by keeping the tubes upside down on tissue paper.
8. Resuspend the pellet with 1 mL PBS.
9. Centrifuge the tube with phage 5 minutes at 14,000 rpm.
10. Transfer 1 mL supernatant in a new tube.
11. Precipitate again adding 250 µL of cold PEG/NaCl to the 1 mL phage, mix and put 10-15 minutes on ice.
12. Centrifuge the tube 15 minutes at 14,000 rpm, discard the supernatant and dissolve the pellet in 1 mL PBS.
13. Centrifuge 5 minutes at 14,000 rpm and transfer the supernatant to a new Eppendorf tube.
14. Store the input phage at 4°C.
15. Check OD_{600} of TG1 cells and put them on ice when it reaches 0.6.

3.12.2.3 INPUT PHAGE TITRATION

Material

- Grown TG1 cells
- Input phage
- LB plates 100 mg/mL ampicillin and 2% glucose
- PBS

Protocol

1. Make serial dilutions of input phage (10 µL phage in 90 µL PBS), typically the dilutions used go from 10^{-1} to 10^{-10}.
2. Add 10 µL of phage dilutions to 90 µL of exponentially growing TG1 cells with OD_{600} of 0.6 only on dilutions: 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-9} and 10^{-10}.
3. Infect for 30 minutes at 37°C without shaking.

4. Plate 4 µL of the infected cells on LB plates containing 100 mg/mL ampicillin and 2% glucose.
5. Grow O/N at 37°C.
6. Next day, calculate the number of input phages using the following formula:

$$\#\ of\ colonies \cdot 10 \cdot 250 \cdot dilution\ factor$$

3.12.3 PHAGE SELECTION

Different antigen presentation methods were performed for the phage selection. In general, once the antigen is coated to an ELISA plate or bind to magnetic beads, the phage library is added and incubated with the sample. Then, they are trypsinized and they infect previously grown TG1 cells. After the incubation, a phage titration is done where serial dilutions are plated in a culture plate that is grown O/N. The following day, colonies are counted and the enrichment from the phage library is calculated.

3.12.3.1 SELECTION VIA ANTIGEN COATING

Material

- Coating buffer: 0.1 M $NaHCO_3$ pH 8,2
- Antigen (MLC1 reconstituted in liposomes)
- Sample buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water
- 4% milk in sample buffer
- PBS
- Trypsin
- AEBSF protease inhibitor

Protocol

1. Coat the wells with 100 µL coating buffer and add 2 µg per well of antigen and incubate O/N at 4°C.
2. Wash 5 times with sample buffer.
3. Prepare 4% milk in sample buffer, centrifuge 5 minutes at 4,000 rpm and use the supernatant for the blocking.
4. Block with 200 µL of milk and incubate 2 hours at RT.
5. Wash 5 times with sample buffer.

6. Prepare the mixture of 10 µL of phage input with 1% milk in sample buffer (at least 10 minutes before its use).
7. Add 100 µL of the mixture per well and incubate 2 hours at room temperature.
8. Wash 15 times with sample buffer.
9. Elute with 100 µL of trypsin 250 µg/mL diluted in PBS.
10. Incubate 30 minutes at RT.
11. Transfer the volume into Eppendorf tubes containing 6.25 µL of AEBSF.

3.12.3.2 SELECTION VIA TWINSTREP COATING

Material

- Coating buffer: 0.1 M $NaHCO_3$ pH 8.2
- Anti-TwinStrep Tag (anti-TST)
- Sample buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water
- 4% milk in sample buffer
- PBS
- Trypsin
- AEBSF protease inhibitor

Protocol

1. Coat the wells with 100 µL coating buffer and add 1 µg per well of anti-TST and incubate O/N at 4°C.
2. Wash 5 times with sample buffer.
3. Add 100 µL of capturing antigen at a concentration of 0.7 µg/mL diluted in sample buffer.
4. Incubate 1 hour at RT. Meanwhile prepare 4% milk in sample buffer, centrifuge 5 minutes at 4,000 rpm and use the supernatant for the blocking.
5. Block with 200 µL of milk and incubate 2 hours at RT.
6. Wash 5 times with sample buffer.
7. Prepare the mixture of 10 µL of phage input with 1% milk in sample buffer.
   * Prepare the mix at least 10 minutes before its use.
8. Add 100 µL of the mixture per well and incubate 2 hours at room temperature.
9. Wash 15 times with sample buffer.
10. Elute with 100 µL of trypsin 250 µg/mL diluted in PBS.
11. Incubate 30 minutes at RT.
12. Transfer the volume into Eppendorf tubes containing 6.25 µL of AEBSF.

3.12.3.3 SELECTION VIA STREPTACTIN BEADS

Material

- StrepTactin XT MagStrep type 3 XT beads (IBA LifeSciences)
- Sample buffer: 100 mM Tris pH 8; 150 mM NaCl in milli-Q water
- Magnetic support
- 4% milk in Sample buffer
- PBS
- Trypsin
- AEBSF protease inhibitor
- KingFisher Flex instrument (ThermoFisher)

Protocol

All plates must be prepared manually; the KingFisher instrument will transfer the magnetic beads in the plates by a defined program.

1. Prepare beads
    a. Put the tubes containing the beads on the magnetic support and remove the storing buffer with a pipette.
    b. Remove the tubes from the magnetic support and add 100 µL of sample buffer.
    c. Resuspend the beads, place the beads on the magnet and remove the buffer.
    d. Repeat steps b and c for a second wash.
    e. Resuspend beads in 40 µL of sample buffer and pipette 20 µL to each well in the plate.
2. Add 0.8 µg of antigen per well in sample buffer up to 100 µL.
3. Incubate the plate 1 hour shaking at RT.
4. Prepare 4% milk in sample buffer, centrifuge 5 minutes at 4,000 rpm and use the supernatant for the blocking.
5. Block with 4% milk for 45 minutes
6. Wash once with 500 µL of sample buffer.
7. Add 100 µL of the phage library and incubate for 84 minutes.
8. Wash 8 times with 500 µL of sample buffer.
9. Elute with 250 µg/mL of trypsin in PBS at a finale volume of 100 µL.
10. Transfer the volume into Eppendorf tubes containing 6.25 µL of AEBSF.

3.12.4 OUTPUT PHAGE TITRATION

Material

- TG1 cells exponentially grown
- LB plates 100 mg/mL ampicillin and 2% glucose

Protocol

1. Prepare serial dilutions in a 96-well culture plate (10 µL output phage in 90 µL PBS), typically the dilutions used go from 10^0 to 10^{-5}.
2. Add 10 µL of phage dilutions to 90 µL of exponentially growing TG1 OD_{600} 0.5.
3. Infect for 30 minutes at 37°C without shaking.
4. Plate 4 µL drops in duplicate on LB plates containing ampicillin and glucose.
5. Grow O/N at 37°C.
6. Calculate the number of total output phage using the following formula:

$$\#\ of\ colonies \cdot 10 \cdot 250 \cdot dilution\ factor$$

3.12.5 PREPARATION MASTERPLATE

After panning, 96 individual clones from distinct selection outputs are picked and the nanobodies are expressed in the *E. coli* periplasm.

Material

- Output phage selection
- LB plates 100 mg/mL ampicillin and 2% glucose
- 2xTY 100 mg/mL ampicillin, 2% glucose and 10% glycerol

Protocol

1. Grow separate colonies containing Nb
2. Plate out 40 µL of output phage on LB plates containing ampicillin and glucose.
3. Grow O/N at 37°C.
4. Pick 92 single colonies and inoculate them into 500 µL of 2xTY 100 mg/mL ampicillin, 2% glucose and 10% glycerol.
5. Grow O/N at 37°C with shaking at 250 rpm.
6. Store at -80°C for long term storage.

Expression of Nb in deep-well plate from fresh grown Masterplate:

Material

- 2xTY 100 mg/mL ampicillin; 0.1% glucose
- 10 mM IPTG
- 2xTY

Protocol

1. Take 10 µL of O/N grown masterplate and grow in deep-well plate containing 1 mL of 2xTY 100 mg/mL ampicillin and 0.1% glucose.
2. Cover plate with gas permeable adhesive seal.
3. Grow at 37°C with 250 rpm shaking until OD_{600} is 0.5.
4. Add 100 µL of a 10 mM IPTG solution in 2xTY to the masterplate.
5. Induce and shake for 4 hours at 37°C with 250 rpm shaking.
6. Pellet cells at 4,000 rpm for 10 minutes.
7. Discard SN and freeze pellet at -20°C.

* If expression of Nb is needed from a masterplate stored at -80°C, clones need to grow O/N in a round bottom plate, and proceed as follows:

   a. Make a fresh preculture in 100 µL of 2xTY 100 mg/mL ampicillin and 2% glucose.
   b. Inoculate 5 µL from the stored plate in the 100 µL of preculture.
   c. Take 10 µL of O/N grown masterplate and grow in deepwell plate containing 1 mL of 2xTY 100 mg/mL ampicillin and 0.1% glucose.
   d. Continue as above.

3.12.6 PRODUCTION OF PERIPLASMIC EXTRACTS OF MASTERPLATE

Material

- Masterplate
- Buffer 100 mM Tris pH 8; 150 mM NaCl in milli-Q water

Protocol

1. Thaw deep well plates.
2. Add 100 µL SB to the plates and shake for 30 minutes.
3. Spin at 4,000 rpm for 10 minutes at 4°C and recover 90 µL of supernatant.

2.12.7 SCREENING ELISA

Material

- Coating buffer 0.1 M $NaHCO_3$
- Wash buffer 100 mM Tris pH 8; 150 mM NaCl
- Dry milk
- Capture select C-tag (Life Technologies)
- Anti-Streptavidin alkaline phosphatase (Promega)
- DNPP (Sigma)
- Periplasmic extracts
- ELISA plates

Protocol

1. Coat the plates with 0.1 µg/well of antigen in coating buffer. Incubate O/N at 4°C.
2. Wash five times with Wash buffer.
3. Block the wells with 4% milk in wash buffer for 1 hour at RT.
4. Wash five times with Wash buffer.
5. Add the periplasm extracts to the ELISA plate considering 20 µL/well at 0.4% milk in Wash buffer. Incubate for 1 hour at RT.
6. Wash five times with Wash buffer.
7. Add the mix of Capture Select and strep-AP to the wells. Incubate for 1 hour at RT.
8. Wash five times with Wash buffer.
9. Add 100µL/well of DNPP substrate and incubate for 30 minutes at RT or longer if needed.
10. Read the plate at 405 nm.

3.12.8 KPCR

Material

- Diluted cells
- Buffer 10x
- Primer M57 10 µM
- dNTP mix
- Kapa Taq
- Milli-Q water

Protocol

1. Prepare the master mix of all components:

- Polymerase Buffer 10x	2.5 µL
- dNTPs 10 mM	0.5 µL
- Kapa taq polymerase	0.125 µL
- Primer MP57 10 µM	0.5 µL
- Primer GIII 10 µM	0.5 µL
- Template DNA 100 ng/µL	1-3 µL
- Milli-Q water	X µL
Final Volume	20 µL

2. Add 5 µL of diluted cell suspension.
3. Place the tubes in the thermocycler and set the PCR protocol with the following conditions:

Initiation	2 minutes at 95°C	
Denaturation	30 seconds at 94°C	28 cycles
Annealing	30 seconds at 55°C	
Extension	45 seconds at 72°C	
Ending	7 minutes at 72°C	
Storage	4°C	

3.12.9 PCR CLEANUP

Material

- PCR product
- Exonuclease I (ThermoScientific)
- FastAP (ThermoScientific)
- Milli-Q water

Protocol

1. Mix 0.2 µL of Exonuclease I with 0.4 µL of FastAP in 13.4 µL of milli-Q water.
2. Add 1 µL of PCR product to the mix.

3. Put the samples into the thermocycler:
   15 minutes at 37°C
   15 minutes at 80°C
   Store at 4°C

2.12.10 SEQUENCING

Material

- PCR cleanup product
- Primer MP57 10 µM

Protocol

1. Add 2 µL of 10 µM MP57 primer to the 15 µL of cleanup product.
2. Samples were shipped for sequencing.

3.12.11 ALIGNMENT

Once the sequences were obtained, they were aligned using the program CLC Main Workbench 21. Since the nanobody epitope is usually defined by the CDIII, sequences were aligned based on this region. Based on these sequences, families were defined. A family is determined by more than 80% amino acid identity between the members.

3.13 NANOBODIES EXPRESSION

The most conventional method for producing nanobodies is to promote their secretion in the *E.coli* periplasm. To obtain high yields, it is recommended to express the Nanobody in the original display vector pMESy4 (GenBank KF415192) in the periplasm of *E. coli* strain WK6. The oxidizing conditions of the periplasm favour the formation of the disulfide bonds that stabilize the nanobody structure. With this aim, nanobodies are cloned behind a presignal (PelB signal sequence) that directs them to the periplasmic space. They also carry a His6-tail (and EPEA-tag) at the C-terminus. A map of the vector pMESy4 is shown in Figure 36.

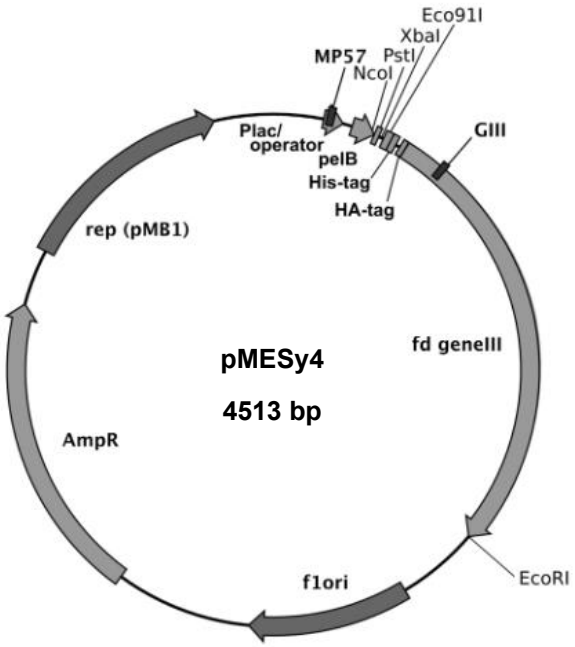


Figure 36. pMESy4 vector scheme. The vector contains a pelB signal after which nanobodies are cloned. Next to the nanobody, a His-tag is present and after the GIII gene is found which codifies for the coat protein pIII.

Material

- LB (autoclaved)
- Ampicillin 1,000x
- Glucose 20% (filtered)
- $MgCl_2$ 2 M (autoclaved)
- IPTG 1 M (filtered)
- TB (autoclaved)
  For 1L: 2.3 g KH_2PO_4; 16.4 g $K_2HPO_4 \cdot 3H_2O$; 12 g Tryptone; 24 g Yeast extract; 4 mL Glycerol 100%.

Protocol

1. Transform WK6 (Su-) cells with plasmid DNA and grow O/N at 37°C on LB ampicillin agar plates with 2% of glucose.
   * Alternatively, streak out a clone from a glycerol stock and grow O/N in the same conditions.
2. Inoculate a colony in 20 mL LB; 20 µL ampicillin 1,000 x; 2 mL glucose 20% and 10 µL $MgCl_2$ 2 M.
3. Shake ON at 37°C, 190 rpm.

4. Grow 3 mL of O/N culture in a 1L flask containing 330 mL TB; 330 µL ampicillin 1,000 x g; 1.5 mL glucose 20% and 330 µL $MgCl_2$ 2 M.
5. Shake at 37°C, 190 rpm until the OD_{600} reaches 0.7 – 1.2.
6. Induce the nanobody expression with 330 µL of IPTG 1 M.
7. Shake 4 hours at 37°C, 190 rpm.
   Alternatively, shake O/N at 28°C, 190 rpm.

3.14 NANOBODIES PURIFICATION

Material

- TES buffer
  0.2 M Tris pH 8; 0.5 mM EDTA; 0.5 M sucrose
- TES/4 buffer
  TES 4 times diluted in H_2O
- Ni/NTA agarose resin
- Phosphate buffer 1
  50 mM Na phosphate; 1 M NaCl pH 7
  For 1 L: 2.6 g $NaH_2PO_4 \cdot H_2O$; 5.4 g Na_2HPO_4; 58 g NaCl.
- Phosphate buffer 2
  50 mM Na phosphate; 1 M NaCl pH 6
  For 1 L: 6 g $NaH_2PO_4 \cdot H_2O$; 1.1 g Na_2HPO_4; 58 g NaCl.
- Elution buffer
  150 mM NaCl with 0.3 M imidazole pH 7
  For 1 L: 2.6 g $NaH_2PO_4 \cdot H_2O$; 5.4 g Na_2HPO_4; 8.7 g NaCl and 20.4 g $C_3H_4N_2$.
- Desalting columns (Cytiva)

Protocol

1. Measure the OD_{600} (it should be around 25).
2. Centrifuge the cultures at 7,000 rpm for 10 minutes.
3. Discard the supernatant.
4. Add 5 mL of TES per flask, resuspend cells.
5. Shake for at least 1 hour at 4°C (it can be left O/N) or resuspend with a magnetic stirrer.
6. Add 10 mL TES/4 and shake for at least 45 minutes at 4°C.
7. Centrifuge at 8,000 rpm for 30 minutes at 4°C.

8. Keep the supernatant.

   * A second osmotic shock can be performed by resuspending the pellet O/N in TES and next day, TES/4 is added. This osmotic shock might yield as much as before but more contaminants will come with it.
9. Use 200 µL of Ni/NTA beads. Spin shortly, remove supernatant and resuspend the beads in 300 µL of phosphate buffer 1.
10. Put the beads on a 50 mL falcon tube and equilibrate with 10 mL phosphate buffer
11. Leave for 10 minutes.
12. Centrifuge at 2,000 rpm for 5 minutes and discard the supernatant.
13. Add the periplasmic fraction to Ni/NTA beads and rotate for 1 hour at room temperature.
14. Bring the solution over in an empty column (take a sample of flowthrough).
15. Wash with 3 mL phosphate buffer 1.
16. Wash with 10 mL phosphate buffer 2 (take a sample of flowthrough).
17. Elute with elution buffer.

    Alternatively, elute with 2-4 mL of acetate buffer 50 mM NaAc; 1 M NaCl at pH 4.5-4.7 in a 15mL falcon tube already containing 1 M Tris pH 7.5.

    * Some nanobodies do not like low pH elution, and imidazole elution is preferred.
18. Measure OD_{280} of the elution.

    * At this point, nanobodies can be stored O/N at 4°C. Do not store at -20°C with imidazole.
19. Pass the eluted sample into a desalting column previously equilibrated to eliminate imidazole.
20. Elute with phosphate buffer 1.
21. Measure OD_{280} of the eluted nanobodies. Store at -80°C for long term storage.

3.15 BIACORE

Surface Plasmon Resonance (SPR) was performed at the Biomolecular analysis unit of Centres Científics i Tecnològics de la Universitat de Barcelona (CCIT-UB) using Biacore^{TM} T-100 device. All the obtained nanobodies against MLC1 were analysed. SPR generates a full kinetic profile that allows the obtention of different constants related to binding: the equilibrium constant (K_D), and the constants of association and dissociation (K_{on} and K_{off}, respectively). It also permits to elucidate the maximum response (R_{max}) that reflects the amount of Nb that the immobilized protein is binding to.

SPR was carried out on Biacore T-100 immobilizing MLC1 purified in digitonin until reaching 600 resonance units (RU) on a CM5 Sensor Chip. Nanobodies were assayed at 3- 0.19 μM in running buffer containing 100 mM Tris pH 8, 150 mM NaCl and 0.005% digitonin.

4. METHODOLOGICAL ANNEX

TABLE OF ANTIBODIES

ANTIBODY	SPECIES
Anti-N4	Rabbit
Anti-flag M2	Mouse monoclonal (Sigma-Aldrich)
Anti-Gprc5b	Rabbit
Anti-His	Mouse (Invitrogen)

Table 7. Antibodies used during this thesis.

TABLE OF PRIMERS

NAME	SEQUENCE	SENSE
attB5 MLC1 Split 2	GGG GAC AAC TTT GTA TAC AAA AGT TGT GGA CTG CAA GAA AAA GAA GGG GTC CAT GTC TGA CAG C	Rv
attB2 MLC1 Split 1	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTA TTA CTC CTC GCT GGA CCG TGC AGC GAT GAT CAC CGT GGC CGC	Rv
attB5R MLC1 Split 1	GGG GAC AAC TTT TGT ATA CAA AGT TGT CTC CTC GCT GGA CCG TGC AGC GAT GAT CAC CGT GGC CGC	Rv
attB1 MLC1 Split 2	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT AAC CAT GGA CTG CAA GAA AAA GAA GGG CTC CAT GTC TGA CAG C	Fw
attB1 TwinStrep MLC1	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT AAC CAT GGC TAG CTG GAG CCA CCC GCA GTT CGA GAA AGG TGG	Fw
MLC1-BRIL	GAG GAG GAC TGC AAG AAA AAG GCC GCT GAT CTT GAA GAC AAT TGG GAA ACC CTC ACC	Fw
MLC1 loop	GAC GAA GTG CCA TTT CCT GCT CGG GTC CTG AAA TCT TAC TCA GTC	Fw
MLC1-BRIL	CCG AGC AGG AAA TGG CAC TTC GTC GAG ATA CTT CTG GAT ATA GGC GTT GCG GGT CGT TTT CAG	Rv
MLC1 loop	CTT TTT CTT GCA GTC CTC CTC GCT GGA CCG TGC AGC	Rv
MLC1 H227A	G GAT GAC TCA GTT TCA GGC CCA GCC CTC TCA GTG ACG TTC TTT TGG	Fw
MLC1 H227A	CCA AAA GAA CGT CAC TGA GAG CCG TGG GCC TGA AAC TGA GTC ATC C	Rv
MLC1 H227R	G GAT GAC TCA GTT TCA GGC CCA CGG CTC TCA GTG ACG TTC TTT TGG	Fw
MLC1 H227R	CCA AAA GAA CGT CAC TGA GAG CCG TGG GCC TGA AAC TGA GTC ATC C	Rv
MLC1 W334A	C CTC TCA GTG ACG TTC TTT GCC ATC CTA GTG GCC TGC	Fw
MLC1 W334A	GCA GGC CAC TAG GAT CCG AAA GAA CGT GAC TGA GAG G	Rv

Table 8. Primers designed during this thesis.

TABLE OF DNA CONSTRUCTS GENERATED

ENTRIES

TwinStrep-hMLC1 P1P2	hMLC1 Split 1 P1P2
TwinStrep-hMLC1 P1P5R	hMLC1 Split 1 P1P5R
TwinStrep-hMLC1-BRIL P1P2	hMLC1 G88V Split 1 P1P5R
hMLC1 H227A P1P2	hMLC1 P92S Split 1 P1P5R
hMLC1 H227R P1P2	hMLC1 Split 2 P1P2
hMLC1 W234A P1P2	hMLC1 Split 2 P5P2

Table 9. Entries generated during this thesis.

EXPRESSION CLONES

TwinStrep-hMLC1 pcDNA3	GlialCAM SmallBiT pcDNA3
TwinStrep-hMLC1-BRIL pcDNA3	Gprc5b SmallBiT pcDNA3
TwinStrep-hMLC1-VFP pcDNA3	Gprc5b LargeBiT pcDNA3
hMLC1 H227A pcDNA3	Gprc5b I176dup SmallBiT pcDNA3
hMLC1 H227R pcDNA3	Gprc5b I176dup LargeBiT pcDNA3
hMLC1 W234A pcDNA3	Gprc5b A177dup SmallBiT pcDNA3
hMLC1 Split 1 pcDNA3	Gprc5b A177dup LargeBiT pcDNA3
hMLC1 Split 1 LargeBiT pcDNA3	
hMLC1 G88V Split 1 LargeBiT pcDNA3	
hMLC1 P92S Split 1 LargeBiT pcDNA3	
hMLC1 Split 2 pcDNA3	
hMLC1 Split 2 SmallBiT pcDNA3	

Table 10. Expression constructs generated during this thesis.

RESULTS

1. CHARACTERIZATION OF AN MLC PATIENT CARRYING TWO MLC1 VARIANTS SHOWING RADIOLOGICAL IMPROVEMENT

As previously described in the Introduction, MLC can be subclassified into two phenotypes: classic and remitting. Classic MLC is caused by recessive variants in *MLC1* gene or recessive variants in *GLIALCAM* gene. In contrast, remitting MLC is caused by dominant variants in *GLIALCAM* gene. Even though in all MLC cases patients develop neurological signs in the first years of age, neurological deterioration does not occur in remitting MLC patients. This group of patients might show initially MRI abnormalities resembling classic MLC, but these alterations improve or normalize with time. Interestingly, two patients with mutations in *MLC1* presenting radiological improvement have been identified, although further characterization is needed (Hamilton et al., 2018).

In this work, we have identified and characterized a novel MLC1 patient that shows a remitting phenotype. For its characterization, we have performed genetic and biochemical studies.

1.1 CASE REPORT

At the time of the study, the patient was an eleven-year-old boy who had evident macrocephaly since the first month of life, being in the percentile >98. Due to a history of macrocephaly present in the father, he was diagnosed with familial macrocephaly.

At 13 months old, he presented a mild simple motor delay along with a delay in cognitive development. The MRI scan showed diffuse signal abnormality and swelling of the cerebral white matter (Figure 37A). An abnormal signal in the posterior limb of the internal capsule was noticed (Figure 37A) as well as minor abnormalities in the cerebellar white matter. At 3 years old, a second MRI was performed where it could be observed a slightly reduction of the swelling. Cystic lesions were found on both temporal lobes, being more visible on the left (Figure 37B), but not on other locations of the brain. Later, at 9 years old, the MRI imaging showed a much more reduced swelling.

Until this study, the patient has never had any epileptic seizures, although it is a common symptomatology of MLC as more than 60% of the published cases presented episodes of epileptic seizures (Yalçinkaya et al., 2003). The patient attends school and follows his education at the same pace as children his age. Regarding his motor level, he is capable of running, jumping, riding a bicycle, and going up and down stairs without difficulty.

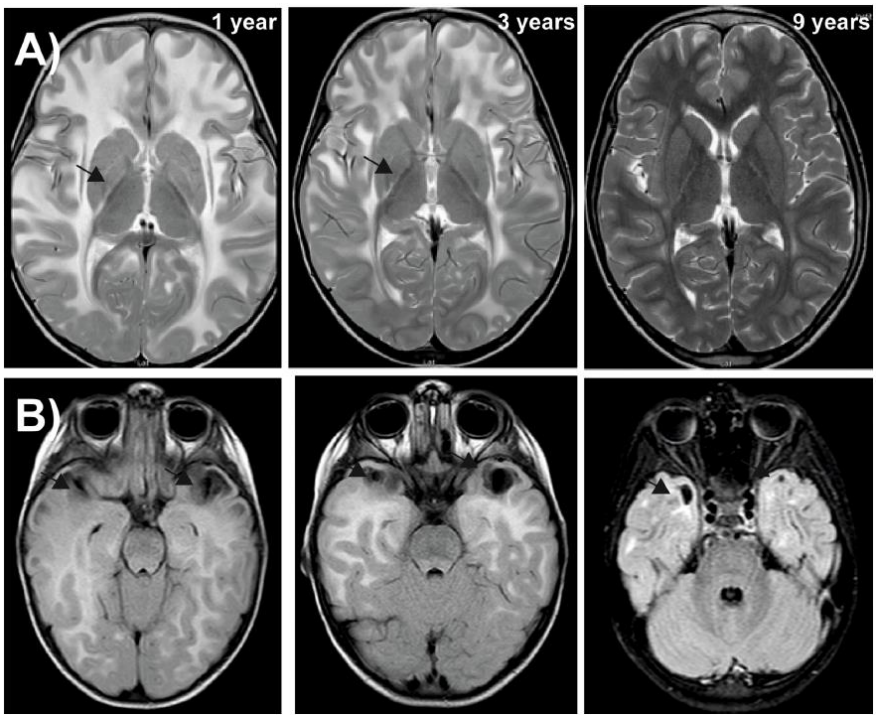


Figure 37. MRI findings in a MLC patient carrying two MLC1 variants. MRI scans done at age 1, 3, and 9 years of age. **A.** T2-weighted axial images of the hemispheres show diffuse white matter abnormalities that improve with time. Marked with an arrow, signal abnormalities observed in the posterior limb of the capsule at age 1 and 3 years are indicated. At 9 years of age, this signal abnormality is reduced. **B.** FLAIR sequences showing the evolution of temporal cysts at different ages. Arrows mark the cysts.

1.2 MOLECULAR STUDIES

Regarding the molecular *MLC1* variants, it was determined that both parents carry two different variants of the gene. The father carried the c.597+37C>G variant, which created a splice acceptor in intron 7. This variant was recently identified in another MLC patient and it was demonstrated that results into the creation of two novel mRNA isoforms (Schlüter et al., 2022).

The mother carried the c.895-1G>T (p.Pro299_Glu353del) variant that also affected splicing. It altered the splice acceptor site of exon 11, which caused a large deletion in the protein sequence and resulted in a shorter RNA sequence (Figure 38). Surprisingly, the father also presented this shorter mRNA sequence, but with lower intensity. In healthy controls this isoform lacking exon 11 was also found, correlating with the presence of the benign polymorphism c.1059+16G>A (Mancini et al., 2012). However, a variant affecting the same nucleotide c.895-1G>C was described in an MLC patient (Patrono et al., 2003). In conclusion, these two variants were pathogenic by affecting splicing.

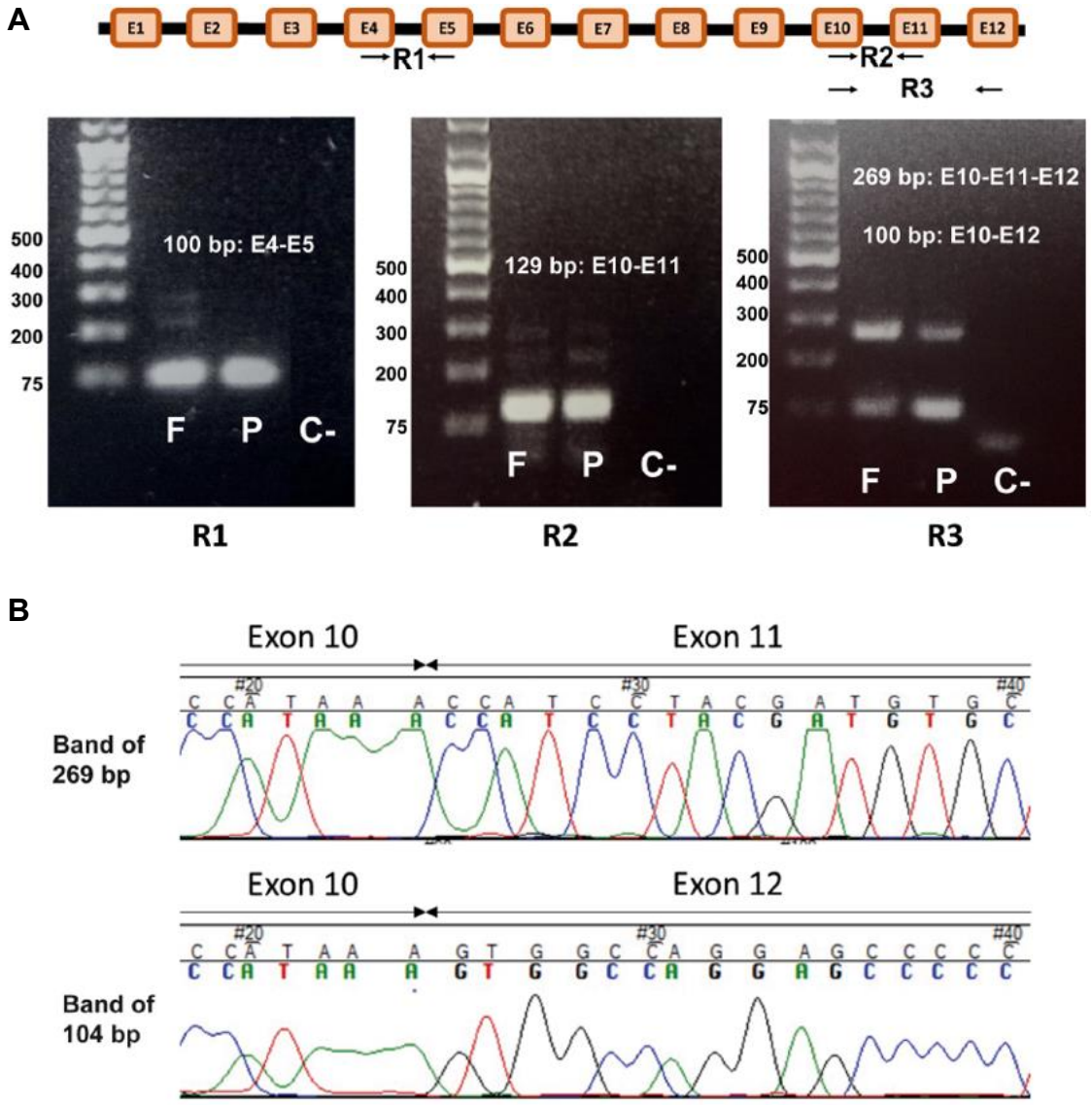


Figure 38. WT *MLC1* mRNA detected in patient and father samples. **A.** Scheme of the strategy followed with the primers used to detect and quantify the WT allele and to confirm exon 11 skipping. cDNA amplification of the PCR products originated with the designed primers (regions R1, R2 and R3). F: father, P: patient, C-: negative control. **B.** Sequence of the two fragments obtained when amplifying exon 10 and 12, confirming the skipping of exon 11 in the shorter fragment.

The improvement observed by MRI together with the improvement at the clinical exploration could be explained by an incomplete penetrance of the splicing variant, which may lead to the expression of some WT MLC1 protein. To test this hypothesis, peripheral blood lymphocytes (PBLs) were isolated from blood samples of both the patient and his father and cDNA was obtained. *MLC1* mRNA was reported to be expressed in PBLs (Nomura et al., 1994). The level of WT transcripts was quantified. For this, a set of primers were designed. A primer pair covered the cDNA sequence from the end of exon 10 and beginning of exon 11 (Figure 38). Since the other *MLC1* allele from the patient contains a variant that induces the skipping of exon 11, the amplification of this sequence assured the specificity for the WT transcript avoiding the interference from the mRNA generated by the allele with the c.895-1G>T variant. To normalize the expression of the WT transcript, a common region of all transcripts (the end of exon 4 and beginning of exon 5) was amplified as well (Figure 38A).

Then, PCR results were quantified. It was observed that the percentage of WT *MLC1* transcripts in the patient was 25%, while the father was 66%, relative to each control *MLC1* gene region (Figure 39A).

To verify that the presence of WT *MLC1* transcript resulted in the presence of MLC1 protein, a solubilized extract from PBLs was obtained from both the patient and his father. Anti-human MLC1 antibodies were used to detect MLC1 protein via Western blot analysis. A band around 30 kDa, corresponding to MLC1 WT protein, was detected in the patient, which was around 10% of the protein detected in the father (Figure 39B). The deletion of p.Pro299_Glu353del might result in an unstable protein that will be degraded by the proteasome system.

In conclusion, we demonstrated that the patient expresses a small amount of MLC1 WT, which might explain its remitting phenotype. We suggest that this patient, together with the other two already described, might belong to a new classification of MLC named MLC1B, in which patients carrying *MLC1* variants display a remitting phenotype.

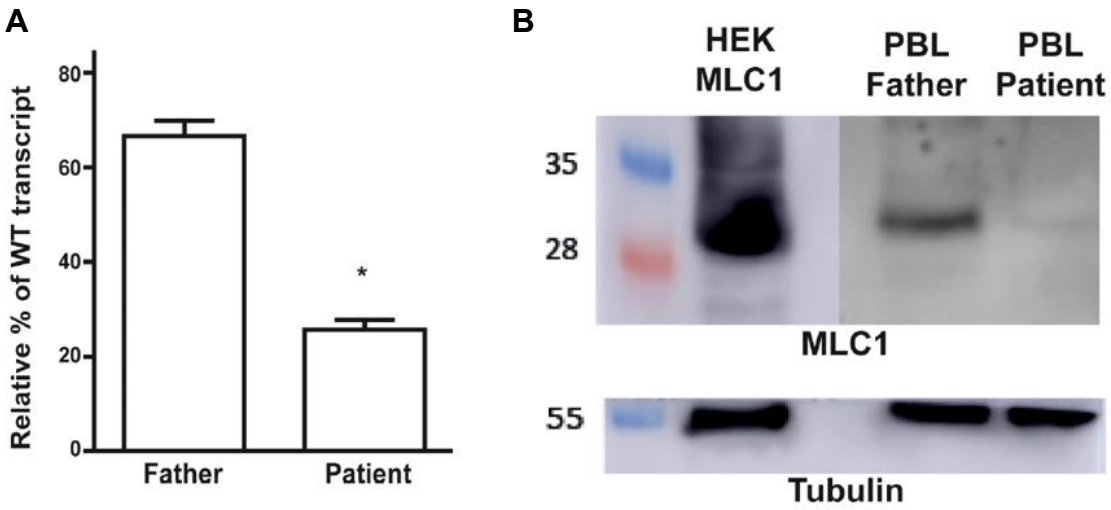


Figure 39. Detection of WT *MLC1* mRNA and protein in PBLs from patient and his father. **A.** Quantitative real time PCR data indicating the presence (at a reduced level) of WT *MLC1* transcript in the patient. T-test *p < 0.05 vs father, n=4. **B.** Detection of MLC1 protein in PBLs from both patient and his father. The detection of human MLC1 in transfected HEK293T cells was used as a control. Tubulin was used as a loading control.

2. STUDIES ON THE ROLE OF GPRC5B SIGNALLING IN THE CONTEXT OF MLC PATHOGENESIS

2.1 GPRC5B SIGNALLING

GPCRs are molecules known to mediate signalling cascades through a variety of mechanisms including G protein-dependent and -independent pathways. Both have the capacity to initiate several intracellular signalling cascades that are essential for many physiological processes. In the case of G protein-dependent signalling, it can be found in the literature that GPRC5B couples with the G12/13 class of heterotrimeric G proteins (Kurabayashi et al., 2013). G12/13 family is known to activate several pathways related with cell growth, cell polarity, migration, and immune and neuronal response (Guo et al., 2022). In regards of the G protein-independent signalling, G-protein coupled receptor kinases (GRKs) interact with the receptor creating binding motifs for arrestin proteins that will induce different signalling pathways (Nürnberg et al., 2024). With the intention of further studying the signalling pathways where GPRC5B is implicated, we set up in our laboratory the NanoBiT *in vitro* technique (*see Methodology section 2.10*).

2.1.1 G-PROTEIN-ALPHA-12 IS INVOLVED IN GPRC5B DOWNSTREAM SIGNALLING

Firstly, we wanted to validate the already described coupling of GPRC5B with the G12 class of G proteins. For this end, HEK293T cells were transfected with the designed NanoBiT constructs necessary for each condition. Minimal G proteins (mini-G) that are solely composed of the GαGTPase domain from G proteins and that closely mimic the pharmacological and structural changes induced in GPCRs were used (Carpenter and Tate, 2016, Carpenter and Tate, 2017). These mini-G had been previously fused to LargeBiT fragments and were found to be useful as vehicles for protein complementation assays (Wan et al., 2018). Forty-eight hours post-transfection, the assay took place and luminescence was read. As depicted in Figure 40, it was confirmed that GPRC5B interacted with the G12 class of G proteins. Unexpectedly, GPRC5B also seemed to interact with Gi. This goes in line with the fact that GPCRs can couple to more than one G protein, giving raise to a variety of signalling for a single receptor based on the relative affinities for different G- proteins (Fonin et al., 2019).

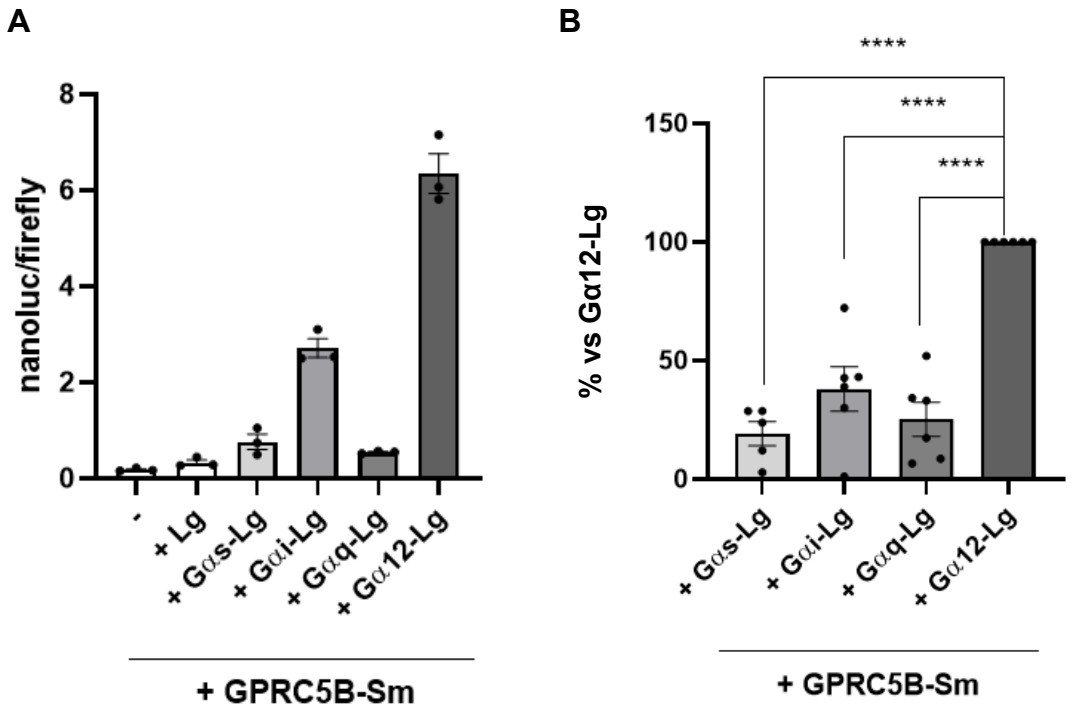


Figure 40. GPRC5B signals via Gα12 family. A. GPRC5B interaction with G proteins. NanoBiT experiments on transfected HEK293T cells. Data comes from one representative experiment with experimental triplicates, the experiment was repeated six times with similar results. Mean ± SEM is represented. **B.** Comparison in percentage of GPRC5B interaction with G proteins giving 100% as an absolute value to Gα12 group. Data comes from five independent experiments. Statistical analysis was assessed using One-way ANOVA analysis (****p-value < 0.0001) with Dunnett's multiple comparisons test: ****p < 0.0001 compared to GPRC5B-SmallBiT + Gα12-LargeBiT group. Mean ± SEM is represented.

Lg = LargeBiT, Sm = SmallBiT.

2.1.2 GPRC5B REGULATION OF VCAM-1 EXPRESSION VIA FYN KINASE

A published article saw evidence of a functional link between GPRC5B constitutive signalling and NF-κB activity in the vascular wall that was dependent of Fyn, a downstream effector of GPRC5B signalling. In a context of inflammation or high glucose concentrations, GPRC5B was found upregulated in cultured immortalized endothelial cells. Also, the overexpression of GPRC5B in the same cells resulted in increased levels of phosphorylation of ERK1/2 and increased activation of NF-κB signalling, which is responsible for increased VCAM-1 protein levels. These effects were confirmed to be Fyn-dependent, since a GPRC5B overexpression accompanied by suppression of Fyn did not show an activation on NF-κB signalling pathways (Freundt et al., 2022). In another study, a GPRC5B mutation (p.Tyr383>Phe, Y383F) was identified in patients suffering from obesity. The authors demonstrated that SH2 domains of Fyn specifically

recognized the Tyrosine residue 383 (Tyr383) for the kinase activation. Thus, the Y383F variant cannot be phosphorylated and, consequently, GPRC5B is not recognized by Fyn (Kim et al., 2012).

Previous studies from the group revealed that the overexpression of GPRC5B in primary astrocyte cultures upregulated VCAM-1 expression. Then, we hypothesized that VCAM-1 induction would be dependent on Fyn kinase activity, so the overexpression of the mutant GPRC5B Y383F, known to not be recognized by Fyn, would not increase VCAM-1 expression. The obtained results are featured in Figure 41. Consistent with our hypothesis, the overexpression of GPRC5B Y383F did not show increased VCAM-1 levels, contrarily to the overexpression of GPRC5B WT. IKB was found phosphorylated in astrocytes overexpressing GPRC5B WT, but not in those overexpressing GPRC5B Y383F. The phosphorylation of IKB is required for the activation of the NF-κB signalling pathway, so the overexpression of GPRC5B WT activates this pathway, but not the GPRC4B Y383F mutant. With this experiment, we concluded that the mechanism by which GPRC5B induces VCAM-1 expression is dependent on Fyn activity.

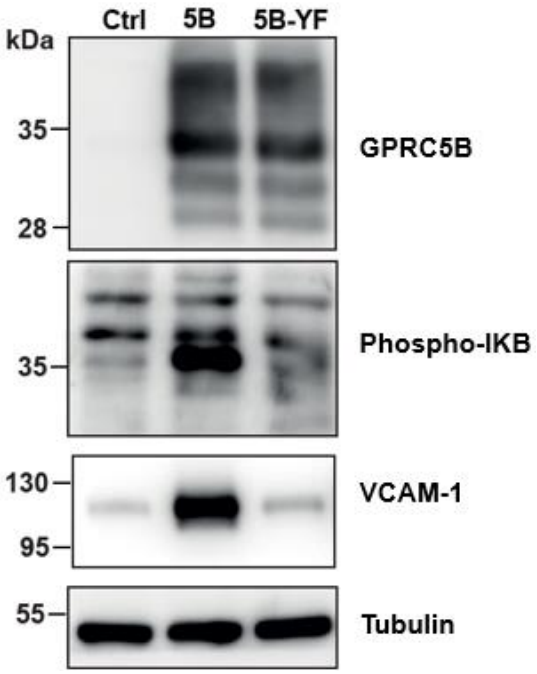


Figure 41. VCAM-1 upregulation induced by GPRC5B is dependent on Fyn activity. Overexpression of GPRC5B Y383F variant using adenoviral particles does not upregulate VCAM-1 compared to GPRC5B WT group. 30 µg of total protein extracts from each group were loaded in 10% acrylamide gels. Data is representative of three independent experiments. Experiments performed by Dr. Adrià Pla Casillanis.

Ctrl: control, 5B: GPRC5B, 5B-YF: GPRC5B Y383F.

2.2 MLC1 NEGATIVELY MODULATES GPRC5B SIGNALLING

Once we had validated that GPRC5B mainly interacted with Gα12 protein, we aimed to assess whether this interaction was altered by the presence of MLC1, GlialCAM or both. For this, HEK293T cells were transfected with the previous NanoBiT constructs plus MLC1, GLIALCAM or GLIALCAM-T2A-MLC1. The last construct enables the expression of both proteins at equimolar concentrations. After 48 hours, NanoBiT assay was performed. As seen in Figure 42, the presence of MLC1 significantly decreased the interaction between GPRC5B and Gα12. In contrast, GlialCAM did not alter it. Interestingly, the co-transfection of both MLC1 and GlialCAM partially recovered the inhibition caused by the presence of MLC1 alone. LRRC8A was added as a control to dismiss a possible competitivity effect. In conclusion, MLC1 inhibited the signalling activity of GPRC5B via Gα12 and this inhibition was partially recovered by the presence of GlialCAM.

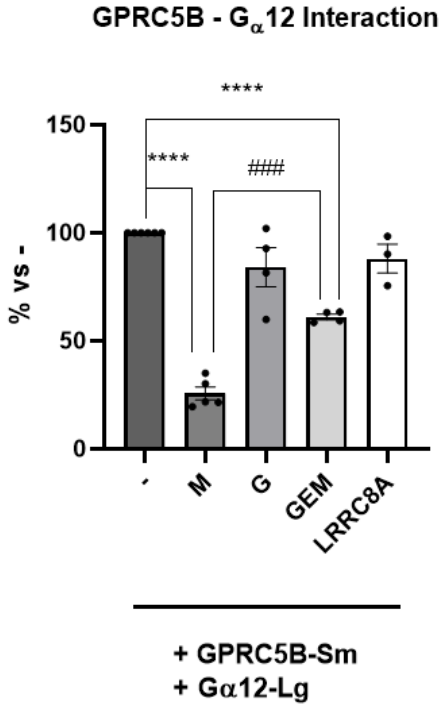


Figure 42. MLC1 negatively modulates GPRC5B-Gα12 signalling. GPRC5B - Gα12 interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from three to six independent experiments depending on the condition. Statistical analysis was determined using One-way ANOVA (****p-value < 0.0001) followed by Tukey's multiple comparisons test. ****p < 0.0001 compared to GPRC5B-SmallBiT + Gα12-LargeBiT group; ### p < 0.001 compared to GPRC5B-SmallBiT + Gα12-LargeBiT + MLC1 group. Mean ± SEM is represented.

Lg = LargeBiT, Sm = SmallBiT, M = MLC1, G = GlialCAM, GEM = GlialCAM T2A MLC1.

Considering that GPRC5B seemed to interact with Gαi and that it might as well bind to β-arrestin, we tested whether the inhibition of MLC1 on GPRC5B – Gα12 interaction could also occur in these two other downstream proteins. NanoBiT experiments were repeated changing the GPRC5B interactor. On the other case, GPRC5B and β-arrestin interaction was significantly decreased by the presence of MLC1 and significantly increased by the presence of GlialCAM. The presence of both MLC1 and GlialCAM did not affect GPRC5B interaction with β-arrestin (Figure 43).

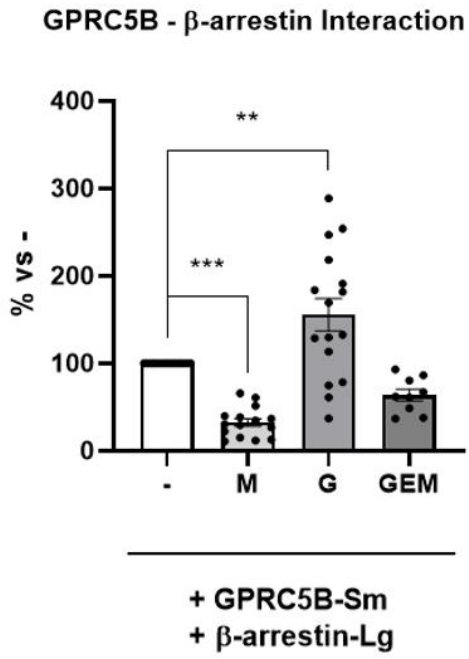


Figure 43. MLC1 negatively modulates GPRC5B interaction with β-arrestin. GPRC5B – β-arrestin interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from nine to sixteen independent experiments depending on the condition. Statistical analysis was determined using One-way ANOVA (****p-value < 0.0001) followed by Tukey's multiple comparisons test. **p < 0.01; ***p < 0.005 compared to GPRC5B-SmallBiT + β-arrestin-LargeBiT group. Mean ± SEM is represented. Experiments done in collaboration with Guillem Pont Espinós.

Lg = LargeBiT, Sm = SmallBiT, M = MLC1, G = GlialCAM, GEM = GlialCAM T2A MLC1.

2.3 MLC1 IMPEDES GPRC5B OLIGOMERIZATION

After seeing that MLC1 disrupted the GPRC5B signalling mediated by G proteins, β-arrestin, we attempted to find a mechanism that would explain how MLC1 would impede this regulation. Since there are no ligands associated with this receptor, no specific mechanisms for its activation have been corroborated. GPRC5B is an orphan receptor classified into the family C of metabotropic glutamate receptors which are known to commonly require dimerization or heteromerization to transduce signalling in response to agonists (Vischer et al., 2015, Kniazeff et al., 2011).

We hypothesized that similarly to some members of class C GPCRs, GPRC5B would need oligomerization to transduce signalling and MLC1 would inhibit this multimerization. To test this hypothesis, NanoBiT assays were performed and the effect of MLC1, GlialCAM or both proteins on GPRC5B interaction with itself was analysed. As shown in Figure 44, GPRC5B oligomerization was significantly decreased by the presence of MLC1, but it was not altered by neither GlialCAM or GlialCAM together with MLC1. With this experiment, we concluded that MLC1 would be impeding the GPRC5B oligomerization, and consequently, the signalling would not occur. Previous studies from the group showed that GPRC5B stabilized MLC1 at the plasma membrane (Alonso-Gardón et al., 2021). This data, together with the obtained results made us propose that the oligomerization might occur at the plasma membrane. So, the signalling at the plasma membrane would be inhibited.

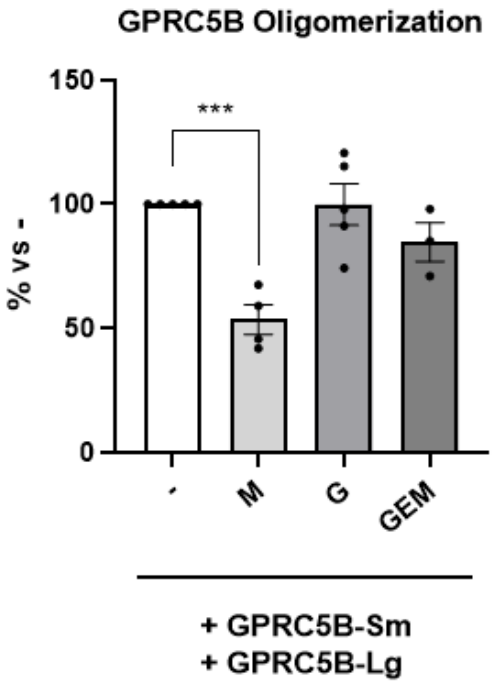


Figure 44. MLC1 impedes GPRC5B homo-oligomerization. GPRC5B Oligomerization modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from three to five independent experiments depending on the condition. Statistical analysis was determined using One-way ANOVA (***p-value < 0.001) followed by Tukey's multiple comparisons test. ***p < 0.001 compared to GPRC5B-SmallBiT + GPRC5B-LargeBiT group. Mean ± SEM is represented.

Lg = LargeBiT, Sm = SmallBiT, M = MLC1, G = GlialCAM, GEM = GlialCAM T2A MLC1.

2.4 GPRC5B CHANGES GLIALCAM LOCALIZATION

As previously reported, GlialCAM potentiated β-arrestin signalling pathway (Figure 43). Since it can be found in the literature that β-arrestin can induce GPCR internalization, we decided to study GPRC5B localization in relationship with GlialCAM. Previous studies from our group described that the expression of GlialCAM was enriched at cell junctions in transfected cells and in primary astrocyte cultures (López-Hernández et al., 2011b).

Co-transfecting HeLa cells with GlialCAM and GPRC5B, immunocytochemistry analysis was done, where GPRC5B polyclonal antibody was utilized and the FLAG tag attached to GlialCAM was detected by using a monoclonal antibody against the tag. Cells were classified based on the dichotomous quantification of whether GlialCAM was enriched or not at cell junctions (*see Methodology section 2.9.3*). It was observed, as depicted in Figure 45, that GlialCAM changed its localization and its expression was reduced at cell-cell junctions in those conditions where GPRC5B was also expressed.

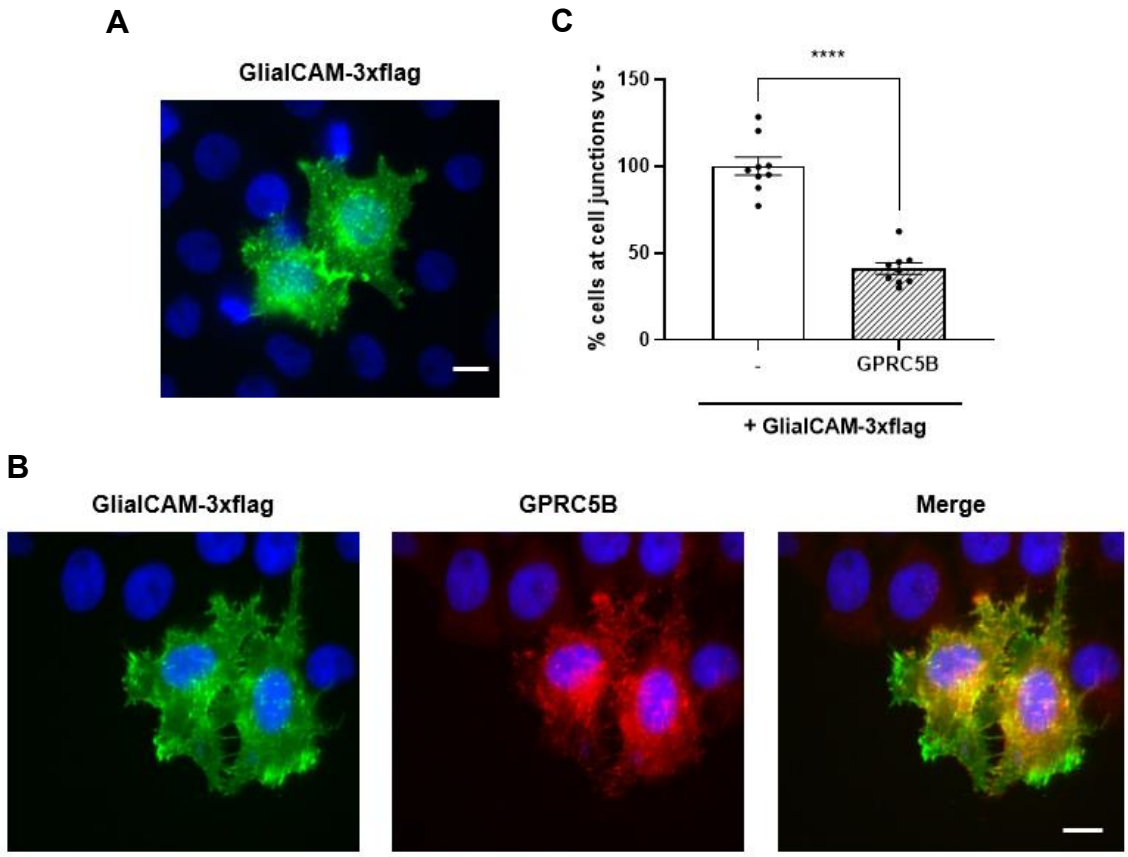


Figure 45. GlialCAM is internalized in presence of GPRC5B. **A.** Immunocytochemistry on HeLa cells transfected with GlialCAM-3xflag. GlialCAM is concentrated at cell junctions. **B.** Immunocytochemistry on HeLa cells co-transfected with GlialCAM-3xflag and GPRC5B. GlialCAM is no longer concentrated at cell junctions. **C.** Quantification of cells where GlialCAM is enriched at cell-cell junctions compared with the condition of transfected GlialCAM alone. Data comes from nine independent experiments. From 50 to 70 cells were analysed per condition. Statistical analysis was determined using Paired t-test (****p-value < 0.0001). Mean ± SEM is represented. Scale bar: 20µm.

Following this observation, we proceeded to determine the role of MLC1 in GlialCAM internalization GPRC5B-dependent. To do so, immunocytochemistry analysis were performed and it was determined that MLC1 protects GlialCAM of GPRC5B-mediated internalization (Figure 46).

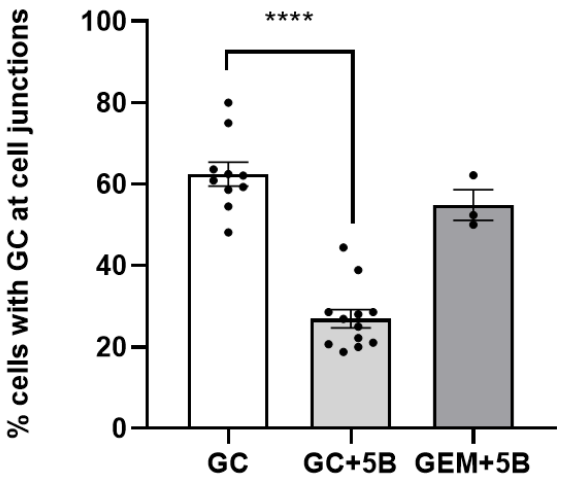


Figure 46. MLC1 protects GlialCAM from GPRC5B internalization. Immunocytochemistry on transfected HeLa cells. GlialCAM is mostly located at cell junctions, but the presence of GPRC5B internalizes it. The addition of MLC1 restores the localization of GlialCAM at cell-cell junctions. Statistical analysis was determined using One-way ANOVA (****$p < 0.0001$) followed by Tukey's multiple comparisons test. ****$p < 0.0001$. Mean ± SEM is represented. Data comes from four to twelve independent experiments depending on the condition. Experiments were performed in collaboration with Dr. Marta Alonso Gardón.

GC: GlialCAM, 5B: GPRC5B, GEM: GlialCAM-T2A-MLC1.

2.5 THE INTERACTION BETWEEN MLC1 AND GLIALCAM INCREASES IN DEPOLARIZING CONDITIONS

As detailed in previous sections, we observed that MLC1 impeded GPRC5B signalling transduction and that the presence of GlialCAM together with MLC1 partially restored it. These results indicated that the interaction between MLC1 and GlialCAM might be regulated somehow. In a Proximity Ligation Assay (PLA) experiment, the interaction with these two proteins was assessed. PLA is based on two labelled secondary antibodies that generate a signal only when detecting two primary antibodies in closer proximity (< 40 nm). Primary astrocyte cultures overexpressing MLC1 and GlialCAM were analysed in physiological and depolarizing conditions. Results are shown in Figure 47, where it can be seen that the association between GlialCAM and MLC1 increases in depolarizing conditions (high potassium concentration in the media).

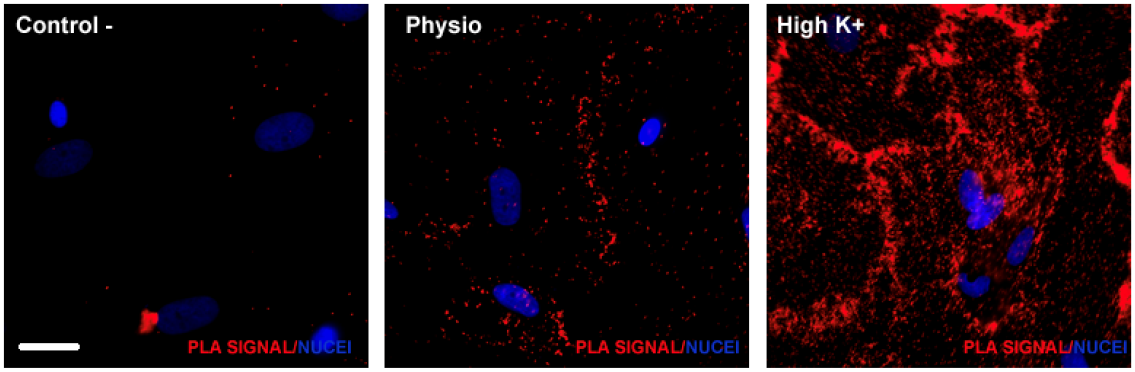


Figure 47. MLC1 and GlialCAM interaction increases in depolarizing conditions. PLA in primary astrocyte culture. Scale bar: 20 µm. Experiment performed by Dr. Xabier Elorza Vidal.

With this result, we postulated that GPRC5B signalling pathway would not be regulated by a ligand, but it would be regulated by the interaction with other proteins. Thus, GPRC5B signalling would be inhibited by MLC1 in physiological conditions, yet in depolarizing conditions, GlialCAM would bind MLC1 and GPRC5B and this interaction would allow GPRC5B signalling to occur (Figure 48).

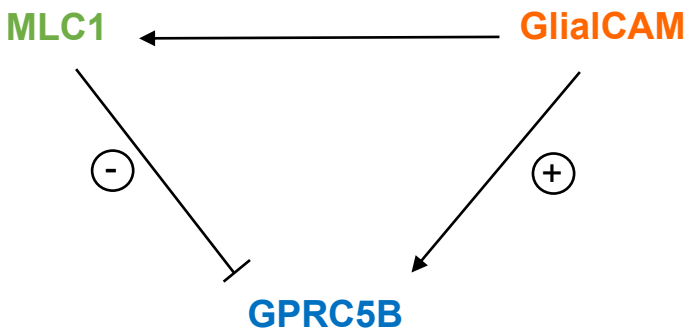


Figure 48. Schematic model of the interplay between MLC1, GlialCAM and GPRC5B. MLC1 would inhibit GPRC5B signalling in physiological conditions. Under depolarizing conditions, GlialCAM would bind to MLC1 and allow GPRC5B signalling.

2.6 GPRC5B MUTANTS BIOCHEMICAL CHARACTERIZATION

Once we had described the GPRC5B signalling pathways and its inhibition through MLC1, as well as the GPRC5B oligomerization which is also inhibited by MLC1, we wondered whether these pathways could be altered in GPRC5B patient-derived mutants. In this section, GPRC5B variants are biochemically described.

2.6.1 GPRC5B-PATIENT DERIVED VARIANTS ARE EXPRESSED

As reported previously in this thesis (*see Introduction Section 2.4.2*), two GPRC5B mutations have been identified of causing MLC: I176dup and A177dup (Passchier et al., 2023). Both mutations are located at the fourth transmembrane domain (Figure 49A). The AlphaFold 3D model predicted these mutants to be found at the interphase of the protein, meaning that they are located to putative interaction sites with itself and other proteins (Figure 49B).

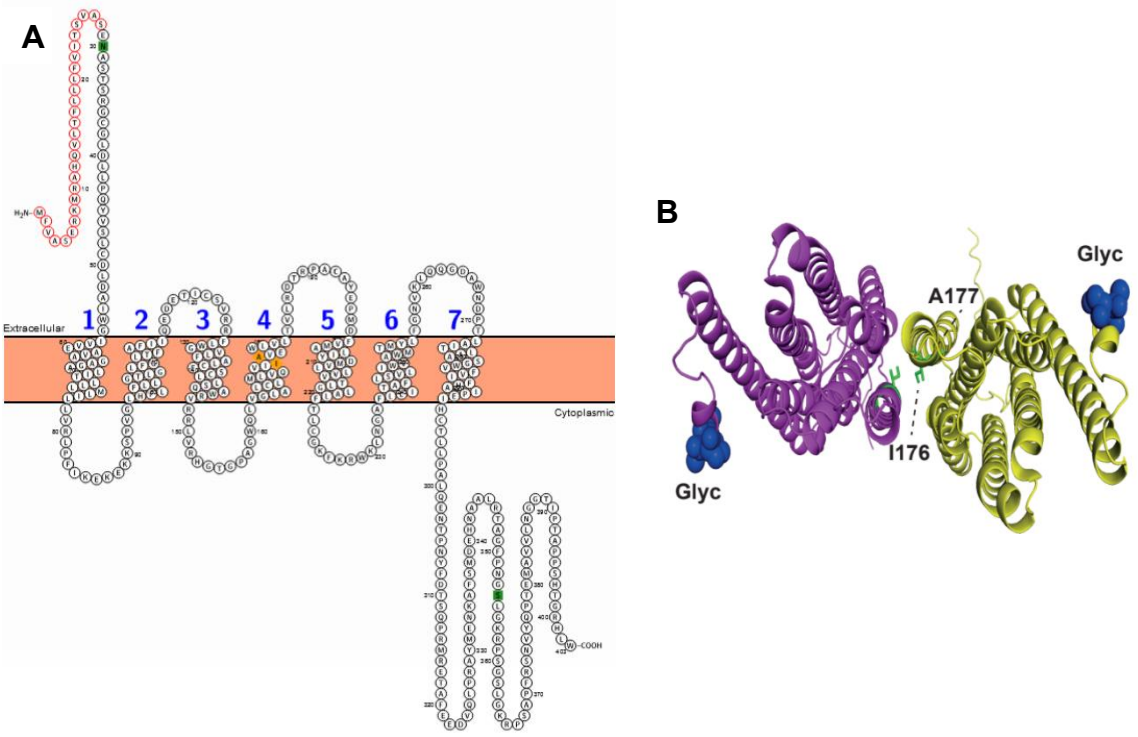
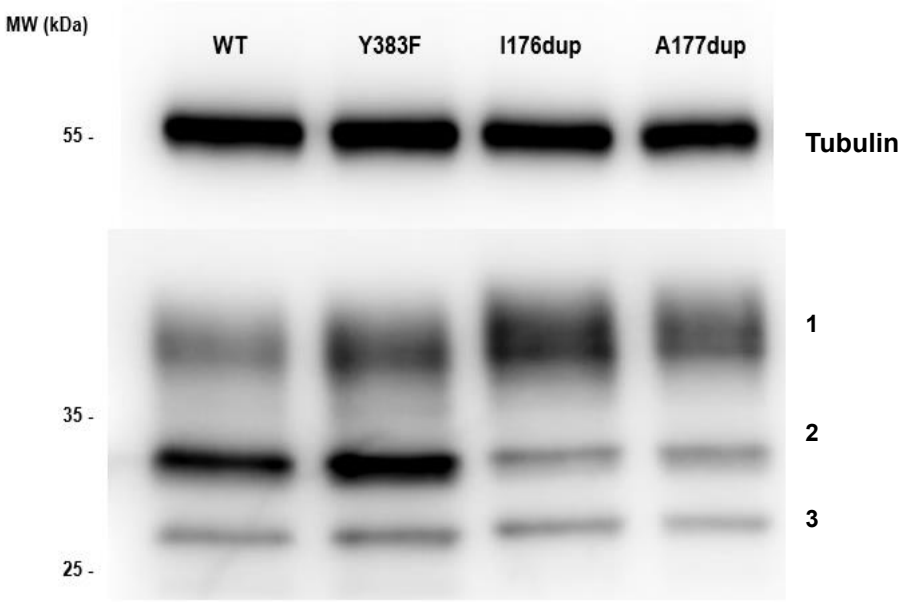


Figure 49. GPRC5B patient-derived variants localization. A. Schematic representation of GPRC5B predicted TM domains. Patient-derived variants are marked in orange. Scheme done in Protter. **B.** 3D representation model of GPRC5B interacting with itself. I176dup and A177dup mutants are detailed in green. Top view.

Since we were studying the GPRC5B protein and its signalling pathways, we wanted to address the pathogenicity of these mutations. First, the protein expression levels were analysed via WB from transfected HEK293T cells lysates. A representative WB membrane is shown in Figure 50A. Although the mutants were expressed to a similar ratio than the WT protein, it can be noticed that the WB bands pattern was distinct compared to the WT. For the WT protein, the most prominent band was the one in the middle (Band 2), which was a little bit smaller than 35 kDa. However, for the mutants, the most distinguishable band was the bigger one (Band 1) which weighted more than 35 kDa. Bands of higher size on a WB could correspond to forms of the protein that have gone through post-translational modifications, such as glycosylation or phosphorylation. Bands were quantified and the ratio between band 1 and band 2 was calculated. As seen in Figure 50B, the ratio Band 1/ Band 2 was significantly increased in the patient-derived mutants compared to the GPRC5B WT.

A



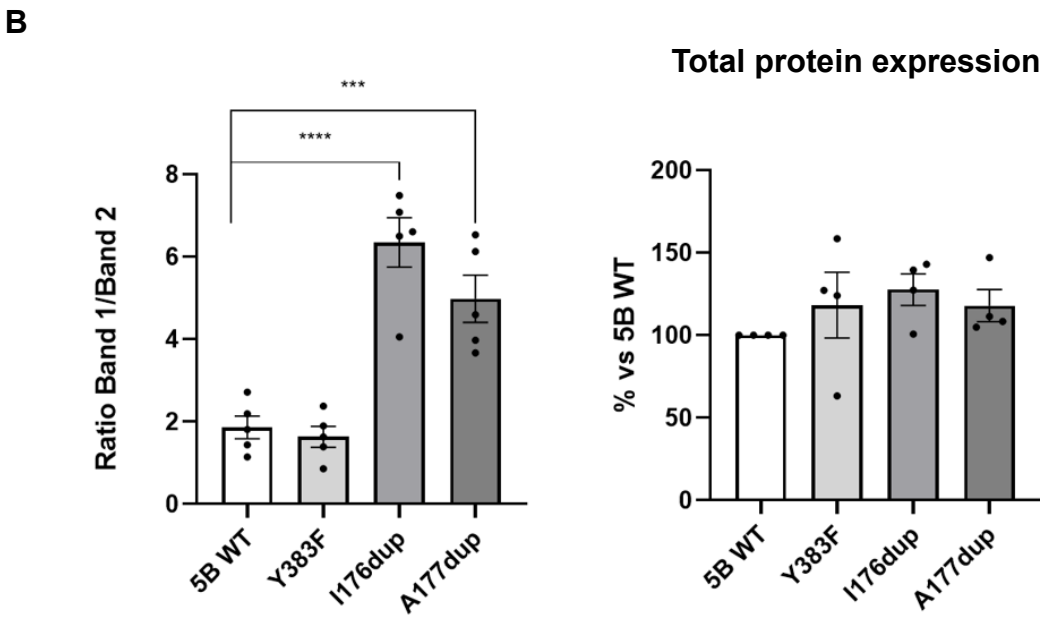


Figure 50. GPRC5B-patient derived variants expression. A. WB detecting GPRC5B from cell lysates previously transfected with GPRC5B WT, I176dup or A177dup. **B.** Ratio bands from the WB (left) and total protein expression (right). Data comes from four independent experiments. Statistical analysis was assessed with One-way ANOVA (***p-value < 0.001) followed by Tukey's multiple comparisons test. ***p < 0.001; ****p < 0.0001 compared to GPRC5B WT. Mean ± SEM is represented. Experiment done by Guillem Pont Espinós.

2.6.2 GPRC5B GLYCOSYLATION PATTERN

To confirm that the WB shifted pattern previous seen in the former section was due to glycosylation modifications, the following experiment was carried out. HEK293T cells were transfected with GPRC5B WT or the variant A177dup and after 48 hours, protein extracts were obtained. Then, the extracts were treated with two different endoglycosidases. Endoglycosidase H (Endo H) is an enzyme that cleaves oligosaccharides while the protein is in the ER or in the early regions of the Golgi complex. On the contrary, Endoglycosidase F (Endo F) can cleave most glycans, including the ones added in the transport from the Golgi complex to the cytoplasm (Freeze and Kranz, 2008).

Comparing protein extracts treated and not treated with Endo H, it was revealed that the band 2 corresponded to glycosylated forms of GPRC5B. The treatment was carried out in protein extracts expressing GPRC5B WT and other extracts expressing GRPC5B A177dup (Figure 51A). However, in the WB a diffuse band 1 can still be seen.

Since Endo H only hydrolyses glycans when the protein is in the ER or in the Golgi complex, we hypothesized that band 1 could correspond to the mature glycosylated form. As an additional proof, Endo F treatment in both WT and A177dup proteins was performed. It was confirmed that the bands with higher molecular weight (bands 1 and 2) belonged to glycosylated forms of the protein because its treatment completely vanished these bands (Figure 51B). This experiment was performed by Guillem Pont Espinós.

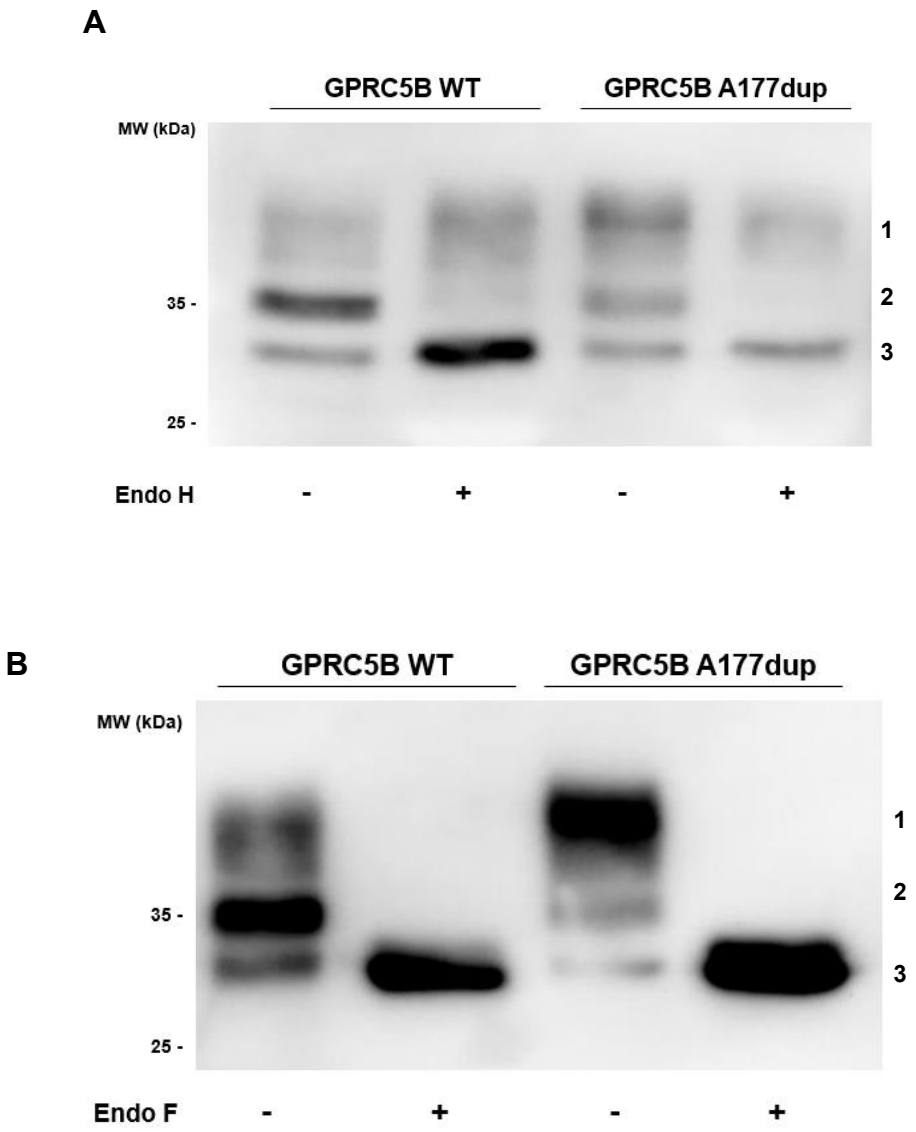


Figure 51. GPRC5B WT and A177dup glycosylation. A. WB detecting GPRC5B from protein extracts expressing GPRC5B WT or GPRC5B A177dup treated or not treated with Endo H. **B.** WB detecting GPRC5B from protein extracts expressing GPRC5B WT or GPRC5B A177dup treated or not treated with Endo F. This experiment was performed once.

2.6.3 GPRC5B MUTANTS CAN OLIGOMERIZE

The multimerization ability of GPRC5B was also studied in the patient-derived GPRC5B mutants. NanoBiT assay was performed and results are represented in Figure 52. It was verified that I176dup and A177dup could perfectly homo-oligomerize.

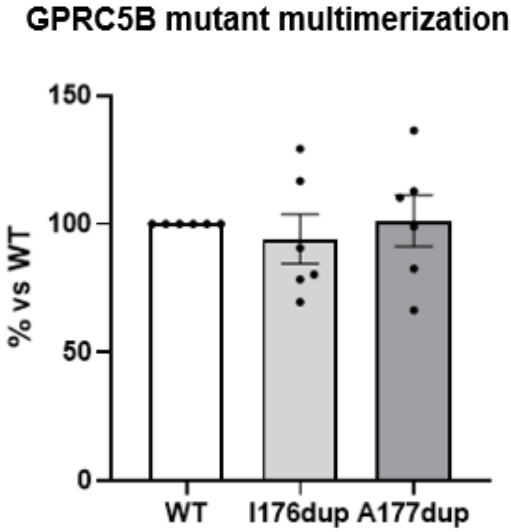


Figure 52. GPRC5B patient-derived mutants can homo-oligomerize. NanoBiT assay on transfected HEK293T cells. Statistical analysis was determined using One-way ANOVA (p-value > 0.05). Mean ± SEM is represented. Data comes from six independent experiments.

2.6.4 GPRC5B-PATIENT DERIVED VARIANTS CAN SIGNAL

After describing an increased interaction of these GPRC5B variants with both MLC1 and GlialCAM, we hypothesized that these mutants would present an impaired signalling. Again, NanoBiT experiments were completed analysing the interaction of Gα12 and β-arrestin with GPRC5B WT and mutants. First, we analysed whether the interaction with the receptor and Gα12 was altered and it was observed that the mutants could interact with Gα12 like the WT (Figure 53A). Considering that GPCR5B can induce signalling cascades mediated by different downstream proteins, other pathways were also studied in the GPRC5B mutants. Taking this into account, the interaction between the mutants and β-arrestin was also assessed. Both duplication mutants could interact with this protein in the same manner than the WT protein (Figure 53B).

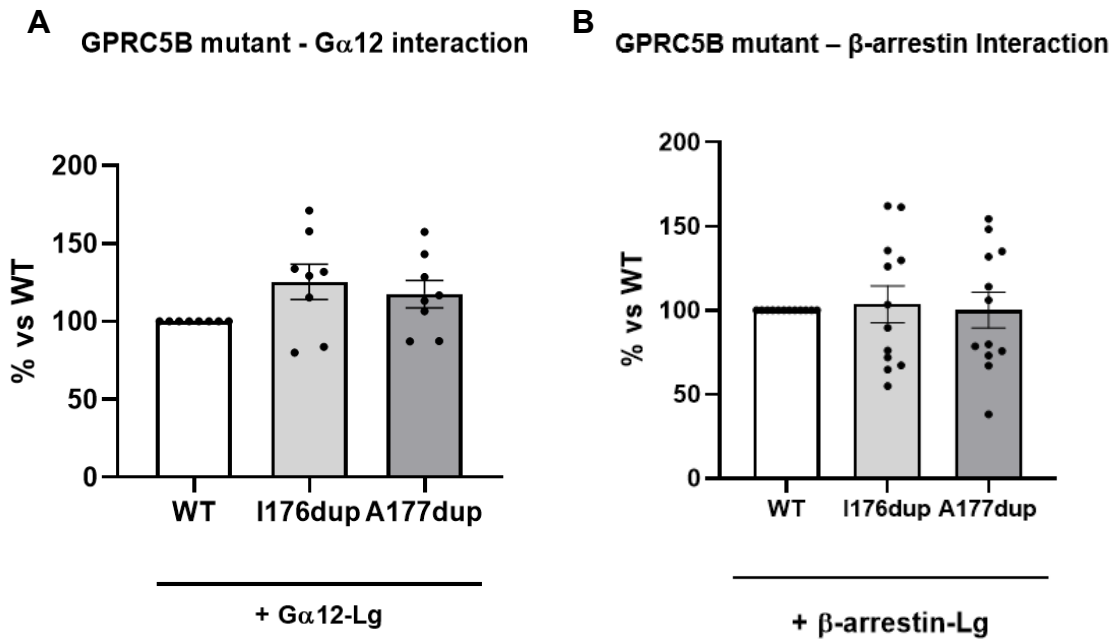


Figure 53. GPRC5B mutant interaction with Gα12 and β-arrestin. A. GPRC5B WT and mutants - Gα12 interaction. NanoBiT experiments on transfected HEK293T cellsStatistical analysis was determined using One-way ANOVA (*p-value < 0.05) followed by Tukey's multiple comparisons test. *p < 0.05 compared to GPRC5B WT-SmBiT + Gα12-LargeBiT. Mean ± SEM is represented. Data comes from eight independent experiments. **B.** GPRC5B WT and mutants – β-arrestin interaction. NanoBiT experiments on transfected HEK293T cells. Statistical analysis was determined using One-way ANOVA (p-value > 0.05). Mean ± SEM is represented. Mean ± SEM is represented. Data comes from twelve independent experiments. Experiments done in collaboration with Guillem Pont Espinós.

Lg = LargeBiT

2.7 GPRC5B MUTANTS RELATIONSHIP WITH MLC1 AND GLIALCAM

2.7.1 MLC1 CAN NEGATIVELY MODULATE GPRC5B PATIENT-DERIVED MUTANTS SIGNALLING

Next, it was hypothesized that although the mutants could signal via Gα12 protein, there could be a deficiency on its regulation through MLC1 and GlialCAM since its interaction with both proteins was altered. To test this, HEK293T cells were transfected with Gα12-Lg, GPRC5B-Sm mutant plus MLC1, GlialCAM or both at equimolar concentrations. Contrarily to our hypothesis, the regulation was not altered because MLC1 could still inhibit the interaction between Gα12 and GPCR5B mutants (Figure 54). Similarly to the interaction observed with GPRC5B WT, the presence of GlialCAM did not affect the

interaction between GPRC5B mutants and Gα12. In this case, the addition of GlialCAM and MLC1 at equimolar concentrations fully recovered the interaction between the GPCR and G protein.

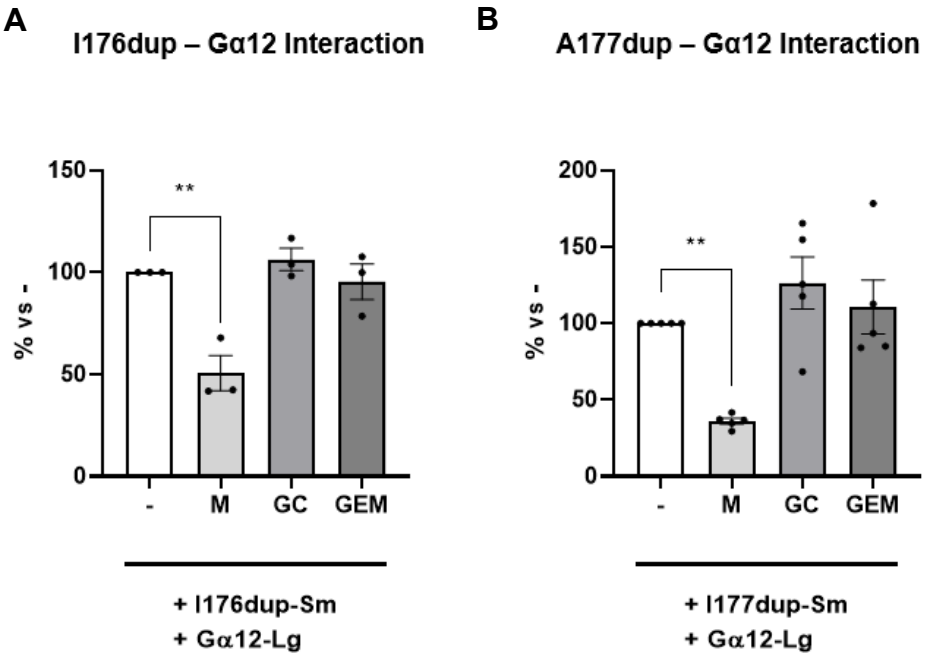


Figure 54. GPRC5B duplication mutants Gα12 signalling. A. GPRC5B I176dup - Gα12 interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from three independent experiments. Statistical analysis was assessed with One-way ANOVA (**p-value < 0.01) followed by Tukey's multiple comparisons test. **p < 0.01 compared to GPRC5B I176dup + Gα12-LargeBiT. Mean ± SEM is represented. **B.** GPRC5B A177dup - Gα12 interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from three independent experiments. Statistical analysis was assessed with One-way ANOVA (**p-value < 0.01) followed by Tukey's multiple comparisons test. **p < 0.01 compared to GPRC5B I176dup + Gα12-LargeBiT. Mean ± SEM is represented.

Lg = LargeBiT, Sm = SmallBiT, M = MLC1, G = GlialCAM, GEM = GlialCAM T2A MLC1.

The same analysis was performed to study the modulation of β-arrestin and GPRC5B mutant interaction (Figure 55). It was revealed that this interaction modulated by MLC1 and GlialCAM behaved similarly than the WT. The presence of MLC1 significantly diminished the interaction with β-arrestin in both mutants, while GlialCAM significantly increased such interaction. In the case of GPRC5B I176dup, the co-expression of MLC1 and GlialCAM at equimolar concentrations also significantly increased the GPCR interaction with β-arrestin.

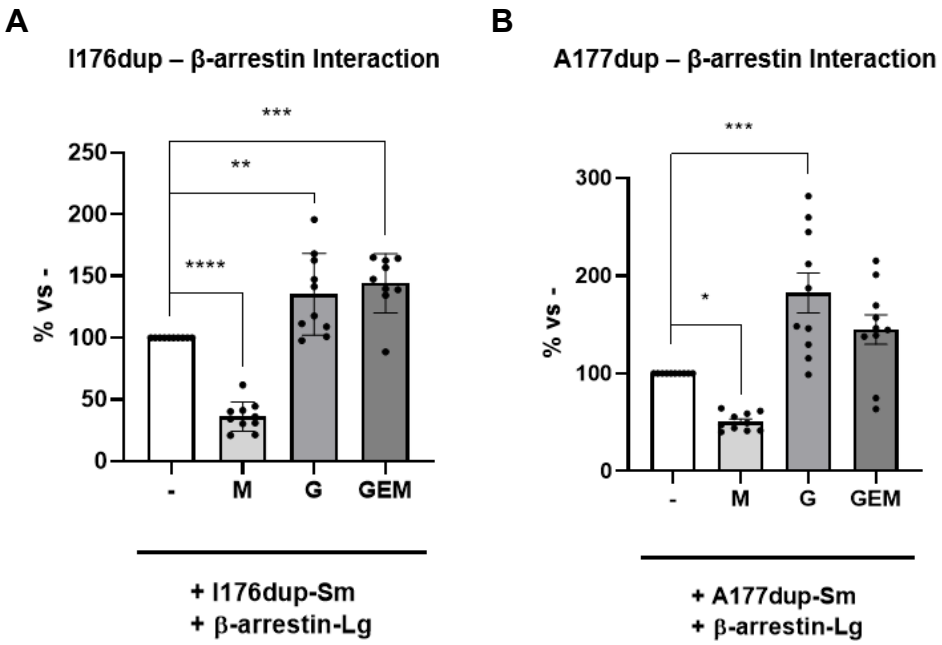


Figure 55. GPRC5B duplication mutants β-arrestin signalling. A. GPRC5B I176dup - β-arrestin interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from ten independent experiments. Statistical analysis was assessed with One-way ANOVA (****p-value < 0.0001) followed by Tukey's multiple comparisons test. **p < 0.01; ***p < 0.001; ****p < 0.001 compared to GPRC5B I176dup + β-arrestin-LargeBiT. Mean ± SEM is represented. **B.** GPRC5B A177dup - β-arrestin interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from ten independent experiments. Statistical analysis was assessed with One-way ANOVA (**p-value < 0.01) followed by Tukey's multiple comparisons test. *p < 0.05; ***p < 0.001 compared to GPRC5B I176dup + β-arrestin-LargeBiT. Mean ± SEM is represented.

Lg = LargeBiT, Sm = SmallBiT, M = MLC1, G = GlialCAM, GEM = GlialCAM T2A MLC1.

Next, we also determined the ability of I176dup and A177dup GPRC5B mutations to activate the NF-κB signalling cascade mediated by Fyn kinase. With this aim, astrocytic primary cultures were infected with adenoviral particles. WB analysis of the protein extracts revealed that both mutants were able to upregulate VCAM-1 expression (Figure 56). It can also be observed that the VCAM-1 band in the mutant conditions is more prominent compared to the GPCR5B WT condition, probably due to the fact that the protein expression of GPCR5B variants seems to be slightly higher than the WT.

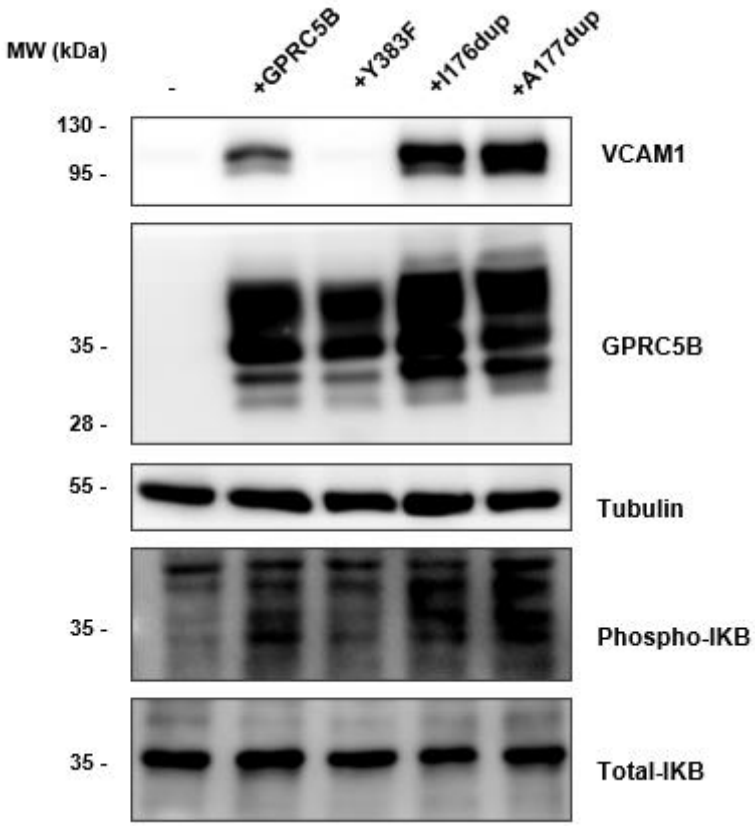


Figure 56. GPRC5B patient-derived variants can induce VCAM-1 expression. Overexpression of GPRC5B WT, and patient-derived mutants I176dup and A177dup using adenoviral particles results in an upregulation of VCAM-1, as revealed by WB analysis. 30 µg of total protein extracts from each group were loaded in acrylamide 10% gels. Data is representative of three independent experiments. Experiments performed by Dr. Adrià Pla Casillanis.

2.7.2 MLC1 CAN INHIBIT GPRC5B MUTANTS OLIGOMERIZATION

The multimerization ability of GPRC5B was also studied in the patient-derived GPRC5B mutants. NanoBiT assay was performed and results are represented in Figure 57. It was verified that I176dup and A177dup could perfectly oligomerize and its multimerization was reduced by the presence of MLC1, in a similar manner to the WT protein. GlialCAM did not affect the oligomerization and the presence of both MLC1 and GlialCAM at equimolar concentrations rescued the decrease seen with MLC1.

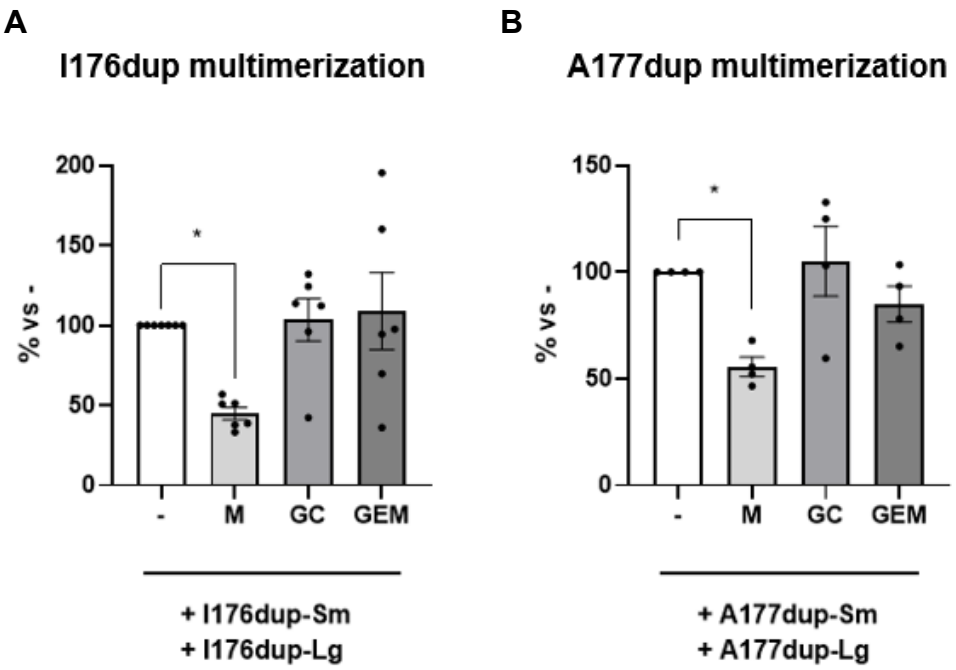


Figure 57. GPRC5B patient-derived mutant oligomerization. A. GPRC5B I176dup homo-interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from six independent experiments. Statistical analysis was assessed with One-way ANOVA (**p-value < 0.01) followed by Tukey's multiple comparisons test. *p <0.05 compared to GPRC5B I176dup-SmallBiT + GPRC5B A177dup-LargeBiT. Mean ± SEM is represented. **B.** GPRC5B I176dup homo-interaction modulated by MLC1, GlialCAM or both. NanoBiT experiments on transfected HEK293T cells. Data comes from four independent experiments. Statistical analysis was assessed with One-way ANOVA (*p-value < 0.05) followed by Tukey's multiple comparisons test. *p < 0.05 compared to GPRC5B A177dup-SmallBiT + GPRC5B A177dup-LargeBiT. Mean ± SEM is represented.

M: MLC1, GC: GlialCAM, GEM: GlialCAM-T2A-MLC1, Sm: SmallBiT, Lg: LargeBiT.

2.7.3 GPRC5B MUTANTS INTERACT WITH MLC1 AND GLIALCAM

Previous published work from our group, described the interaction between MLC1 and GlialCAM with GPRC5B (Alonso-Gardón et al., 2021). In that work, the brain GlialCAM interacting proteome was determined and several proteins, transporters and ion channels were identified. Among them, GPRC5B was found. Bioluminescence resonance energy transfer (BRET) saturation assays were done as a validation of the direct interaction of GPRC5B with both MLC1 and GlialCAM.

Knowing that GPRC5B directly interacts with GlialCAM, we hypothesized that this interaction could be altered in the case of the GPRC5B duplication mutants. With this aim, NanoBiT experiments were performed in transfected HEK293T and it was confirmed

that GPRC5B I176dup and GPRC5B A177dup presented a significantly increased interaction with GlialCAM in comparison with GPRC5B WT (Figure 58A).

We also analysed the interaction of GPRC5B with MLC1 using the same technique. Surprisingly, GPRC5B patient-derived mutants showed a significantly increased interaction with MLC1 in comparison with the WT (Figure 58B). Seeing that the patient-derived GPRC5B mutants interacted more with GlialCAM and MLC1, we proposed that the disrupted regulation of these interactions would confer the pathogenicity to these mutants.

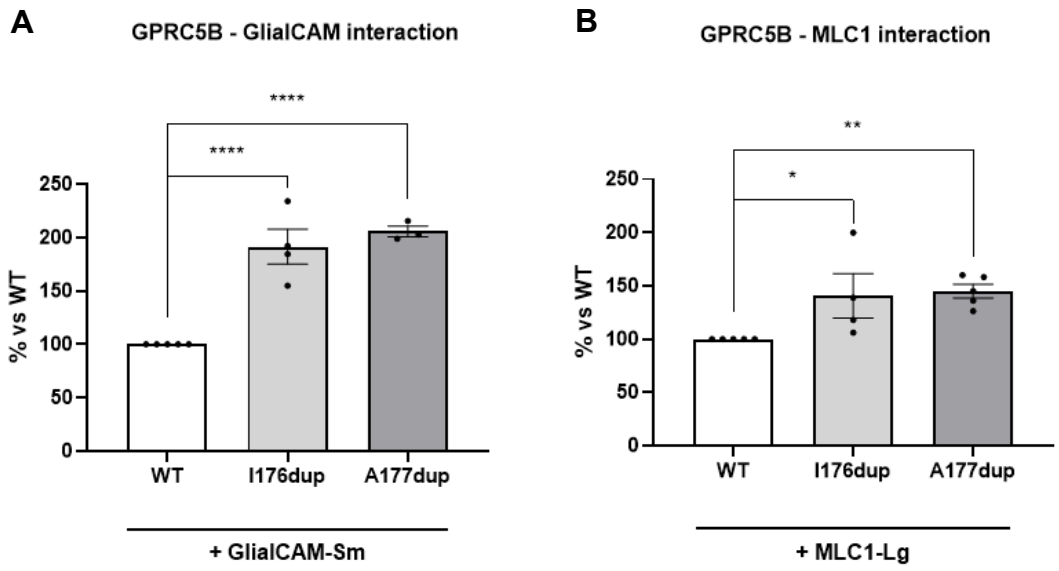


Figure 58. GPRC5B interaction with GlialCAM and MLC1. A. NanoBiT experiments on transfected HEK293T cells. Data comes from three to five independent experiments depending on the condition. Statistical analysis was determined using One-way ANOVA (**** p-value < 0.0001) followed by Tukey's multiple comparisons test. **** $p < 0.0001$ compared to GPRC5B WT- LargeBiT+ GlialCAM - SmallBiT group. Mean ± SEM is represented. **B.** NanoBiT experiments on transfected HEK293T cells. Data comes from five independent experiments. Statistical analysis was determined using One-way ANOVA (****p-value < 0.0001) followed by Tukey's multiple comparisons test. * $p < 0.05$; ** $p < 0.01$ compared to GPRC5B WT + MLC1-LargeBiT. Mean ± SEM is represented. Experiments done in collaboration with Guillem Pont Espinós.

Sm = SmallBiT, Lg = LargeBiT

2.7.4 GPRC5B MUTANTS CANNOT ALTER GLIALCAM LOCALIZATION

Seeing that GPRC5B was found responsible for GlialCAM cellular localization, we hypothesized that the GPRC5B-patient derived variants might not internalize GlialCAM. So, in this case, even in the presence of GPRC5B, GlialCAM would stay at cell-cell junctions. With this aim, HeLa cells were co-transfected with the mutant GPCR together with GlialCAM WT. As it can be seen in Figure 59A, GlialCAM stayed enriched at cell junctions. Interestingly, it was noticeable that GPRC5B variants were also found concentrated at cell junctions, which did not occur in the WT condition.

Again, images obtained in the microscope were classified based on the presence of GlialCAM enriched at cell junctions or not. Cells that had GlialCAM concentrated at cell junctions were compared with those from the condition expressing GlialCAM alone. In both patient-derived mutants the levels of GlialCAM at cell junctions was decreased, but to a lesser extent than GPRC5B WT (Figure 59B). These results suggested that the patient-derived mutants would be more stable at the cell membrane and they would be more resistant to endocytosis.

Since observationally it seemed that both proteins were placed at the same subcellular location, it was decided to quantify their colocalization. As shown in Figure 59C, it was confirmed that GlialCAM colocalized significantly more with GPRC5B mutants compared to the WT suggesting an increased interaction between these proteins, which was already been determined in the NanoBiT assays. The colocalization between GlialCAM and GPRC5B was also assessed in primary astrocytic cultures. *Mlc1*$^{-/-}$ astrocytes were overexpressed with GlialCAM and GPRC5B. Immunocytochemistry analysis and posterior quantification revealed that both I176dup and A177dup co-localized more with GlialCAM compared to GPRC5B WT (Figure 60).

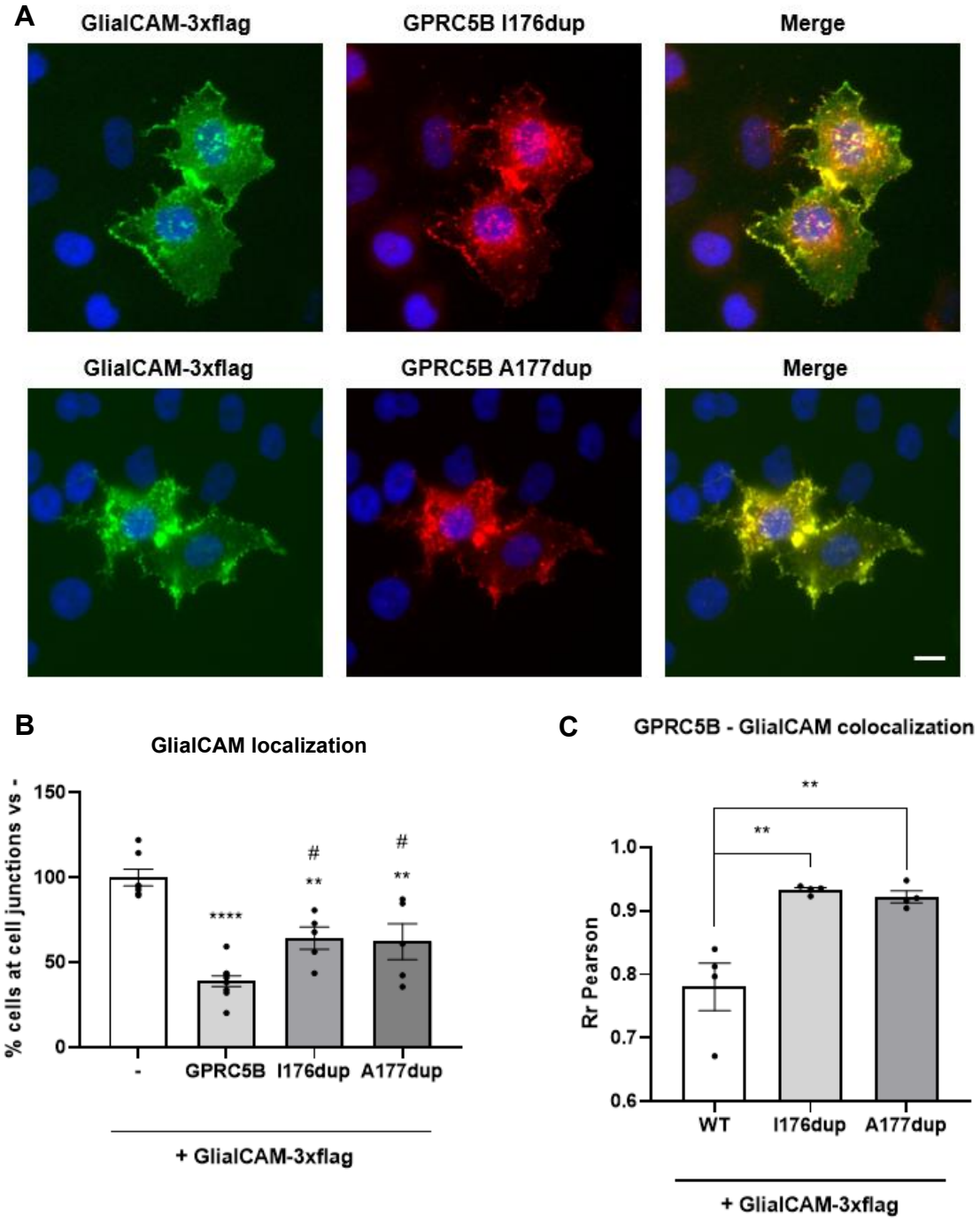


Figure 59. GPRC5B-patient derived mutants maintain GlialCAM enriched at cell junctions. A. Immunocytochemistry on HeLa cells co-transfected with GPRC5B patient-derived mutant and GlialCAM WT. Both mutants GPRC5B I176dup and A177dup stay at cell junctions. Scale bar: 20µm. **B.** Quantification of cells where GlialCAM is enriched at cell-cell junctions compared with the transfection of GlialCAM alone. Data comes from five to ten independent experiments. Between 50 and 70 cells were analysed per condition. Statistical analysis was determined using One-way ANOVA (****p-value < 0.0001) followed by Tukey's multiple comparison test. **p < 0.01; ****p < 0.0001 compared to GlialCAM-3xflag. # p < 0.05 compared to GPRC5B + GlialCAM-3xflag. Mean ± SEM is represented. **C.** Colocalization of GPRC5B WT or mutants with GlialCAM. Statistical analysis was determined using One-way ANOVA (**p < 0.01) followed by Tukey's múltiple comparisons test. **p < 0.01. Data comes from four independent experiments.

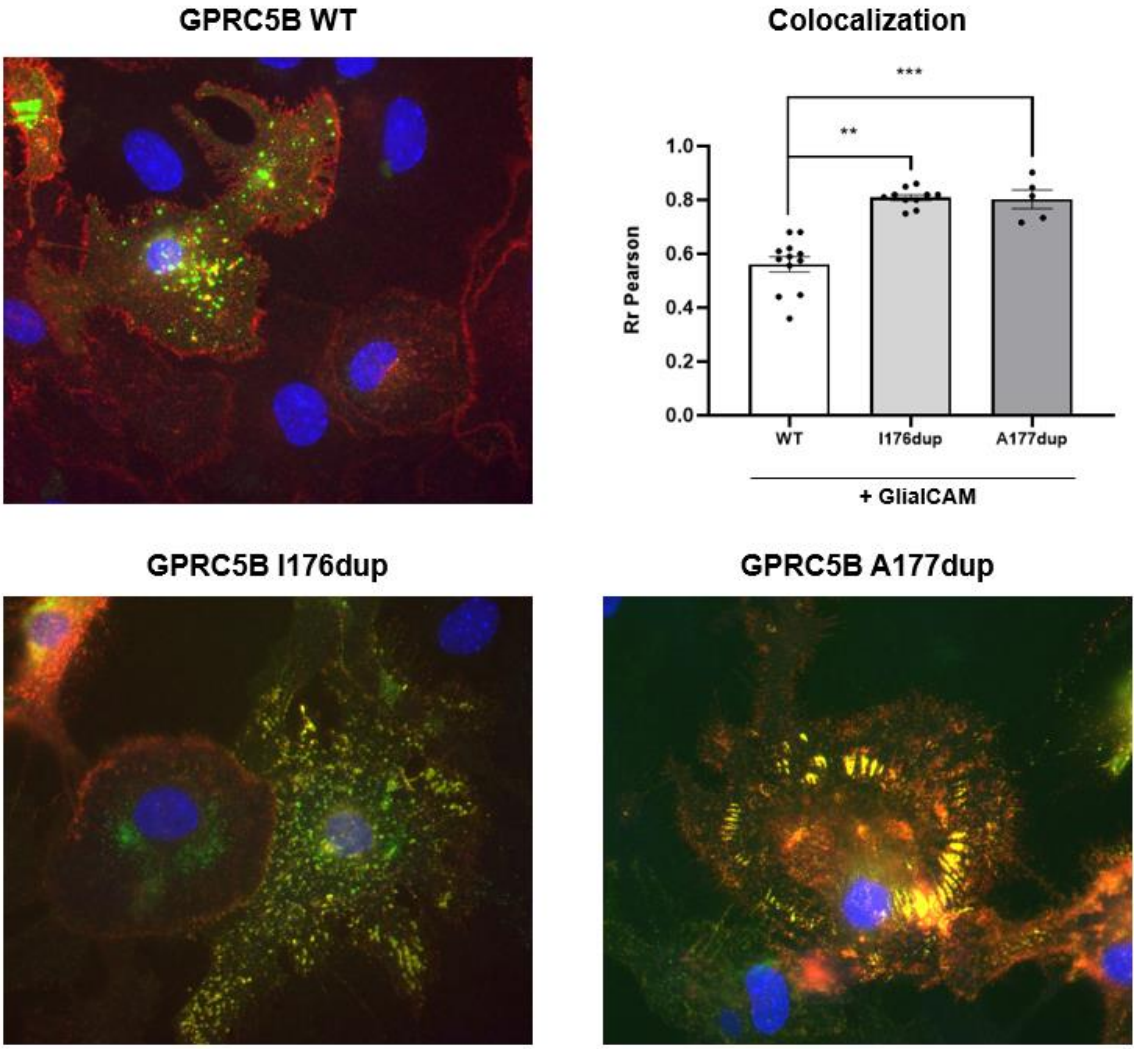


Figure 60. GPRC5B mutants co-localize with GlialCAM in *Mlc1* KO primary astrocytes. Immunocytochemistry on *Mlc1*$^{-/-}$ astrocytes overexpressing GlialCAM and GPRC5B. Both mutants GPRC5B I176dup (bottom left) and A177dup (bottom right) colocalize significantly more with GlialCAM than GPRC5B WT (top left). Colocalization of GPRC5B WT or mutants with GlialLCAM was quantified (top right). Statistical analysis was determined using One-way ANOVA (****p < 0.0001) followed by Tukey's multiple comparisons test. **p < 0.01; ***p < 0.001 vs GPRC5B WT. Mean ± SEM is represented. Data comes from five to twelve independent experiments depending on the condition and they were performed by Dr. Adrià Pla Casillanis.

2.8 A PROPOSED MODEL FOR GPRC5B SIGNALLING

After the study of GPRC5B patient-derived mutants in the context of MLC pathogenesis, we concluded that these variants can perfectly signal. This signalling can also be modulated by MLC1 and GlialCAM, in a similar manner than GPRC5B WT protein. One of the main differences observed in these mutations is that they interact significantly more with both MLC1 and GlialCAM, which could be a result of their increased stability at plasma membrane. Also, GPRC5B patient-derived mutants cannot internalize GlialCAM and they colocalize significantly more at cell-cell junctions.

In addition, considering the PLA assay where it was clearly recognized that the interaction between MLC1 and GlialCAM increases in depolarizing conditions, we propose the following model.

In physiological conditions, GPRC5B and MLC1 would be found interacting at the plasma membrane. MLC1 would be inhibiting GPRC5B homo-oligomerization, hence, the signalling would be inhibited. In a context of hyperpolarization, where the potassium concentration increases in the extracellular media, GlialCAM would bind MLC1 and GPRC5B, thus, MLC1 would no longer be inhibiting GPRC5B. Then, the receptor would homo-oligomerize, it would be internalized, and the signalling would occur. The visual summary of our proposed model is featured in Figure 61.

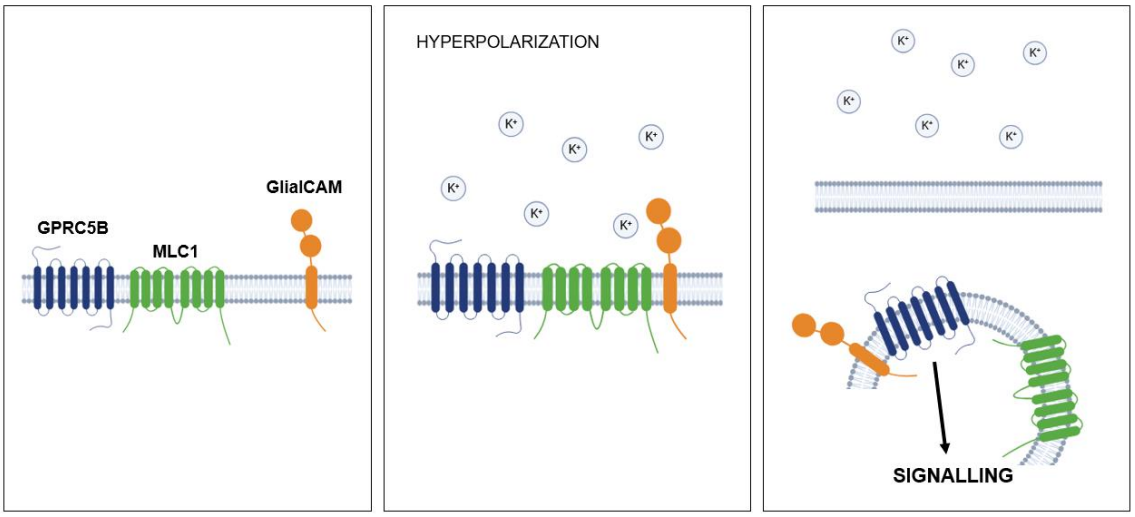


Figure 61. Proposed model for GPRC5B signalling pathway modulation. Our proposed model is that GPRC5B interacts with MLC1 at plasma membrane. In hyperpolarizing conditions, GlialCAM would bind to MLC1-GPRC5B and as a result, GPRC5B would be internalized and the signalling would occur intracellularly.

2.9 GLIALCAM IGC2 MUTANTS

The results obtained in the previous sections, where GPRC5B mutants maintained GlialCAM enriched at cell junctions, recalled previous studies from our group focused on the IgC2 domain of GlialCAM. There have been described two patient mutations located on this domain that cause MLC: p.S196Y and p.D211N (Figure 62A). These two mutants are the only ones described to not show a defect in trafficking (Arnedo et al., 2014b). in this work done in primary knockout MLC1 astrocyte cultures, these two recessive mutations were characterized and it was found that they were not internalized in a high potassium concentration media (Figure 62B). These results suggested that S196Y and D211N mutants might be more stable at the plasma membrane.

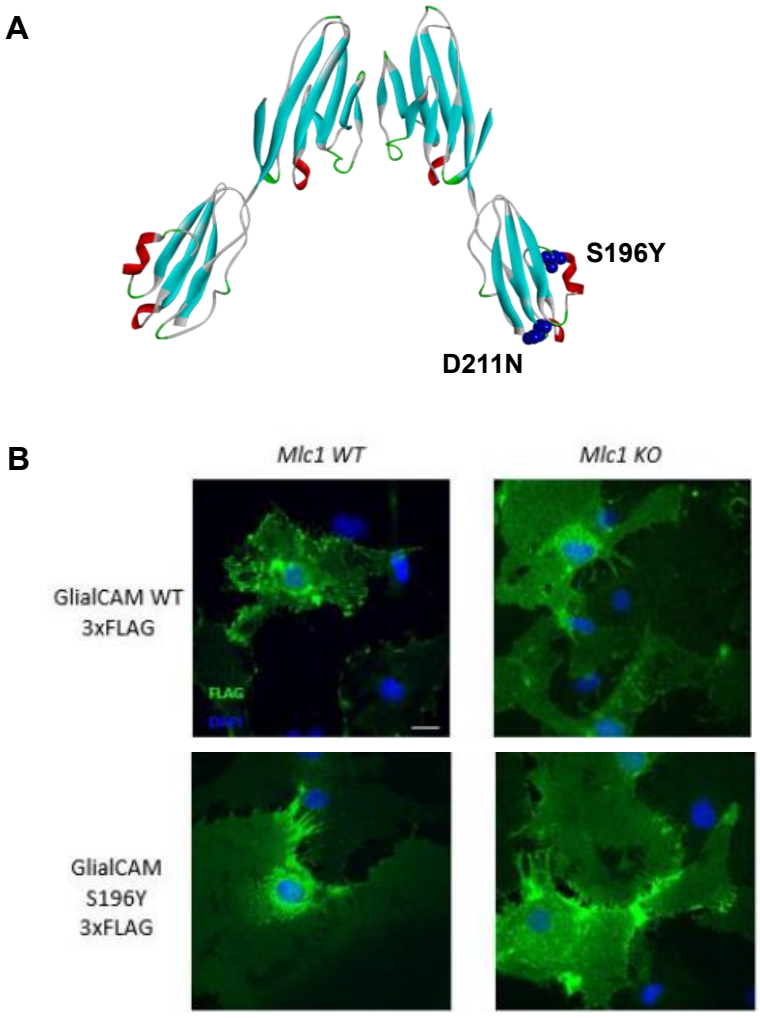


Figure 62. GlialCAM IgC2 patient-derived variants are resistant to endocytosis. A. B. GlialCAM S196Y variant was overexpressed in mouse primary WT astrocytes or *Mlc1* -/- astrocytes. In a context of high potassium cnocentration, both variants were defective for internalization. Images are representative of three independent experiments. Experiments performed by Dr. Adrià Pla Casillanis.

2.9.1 GLIALCAM IGC2 MUTANTS STAY AT CELL JUNCTIONS IN PRESENCE OF GPRC5B

The suggested resistance to endocytosis found in GlialCAM IgC2 patient-derived mutations resembled the decreased internalization observed in GPRC5B mutants when they are co-expressed with GlialCAM, where both proteins stayed enriched at cell-cell junctions.

Having described the internalization of GlialCAM in presence of GPRC5B, we then hypothesized that the patient-derived GlialCAM IgC2 mutations could have a defect of protein localization in cells co-expressed with GPRC5B. Since the mutants were not internalized in the presence of potassium, it was plausible that they would not do so in presence of the GPCR neither.

Both mutants, S196Y and D211N, were expressed in HeLa cells, alone and together with GPRC5B. Results confirmed that IgC2 GlialCAM mutants were enriched at cell-cell junctions (Figure 63). The presence of GPRC5B did not alter the localization of GlialCAM S196Y and D211N as they remained concentrated at cell junctions (Figure 64). These results also validated that all proteins reached the plasma membrane, but the only observed difference in phenotype was their resistance to endocytosis.

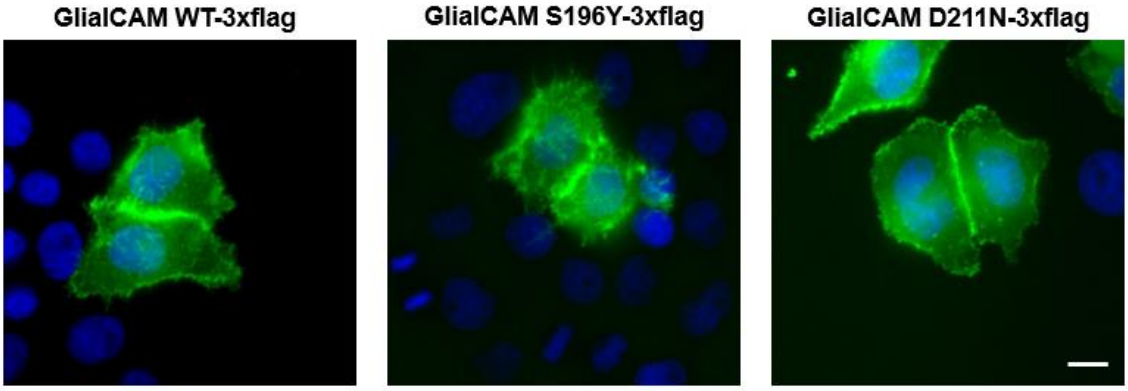


Figure 63. GlialCAM IgC2-patient derived mutants localization. Immunocytochemistry on HeLa cells transfected with GlialCAM WT (left), GlialCAM S196Y (centre), or GlialCAM D211N (right). In all three cases the protein is enriched at cell junctions. Scale bar: 20 μm.

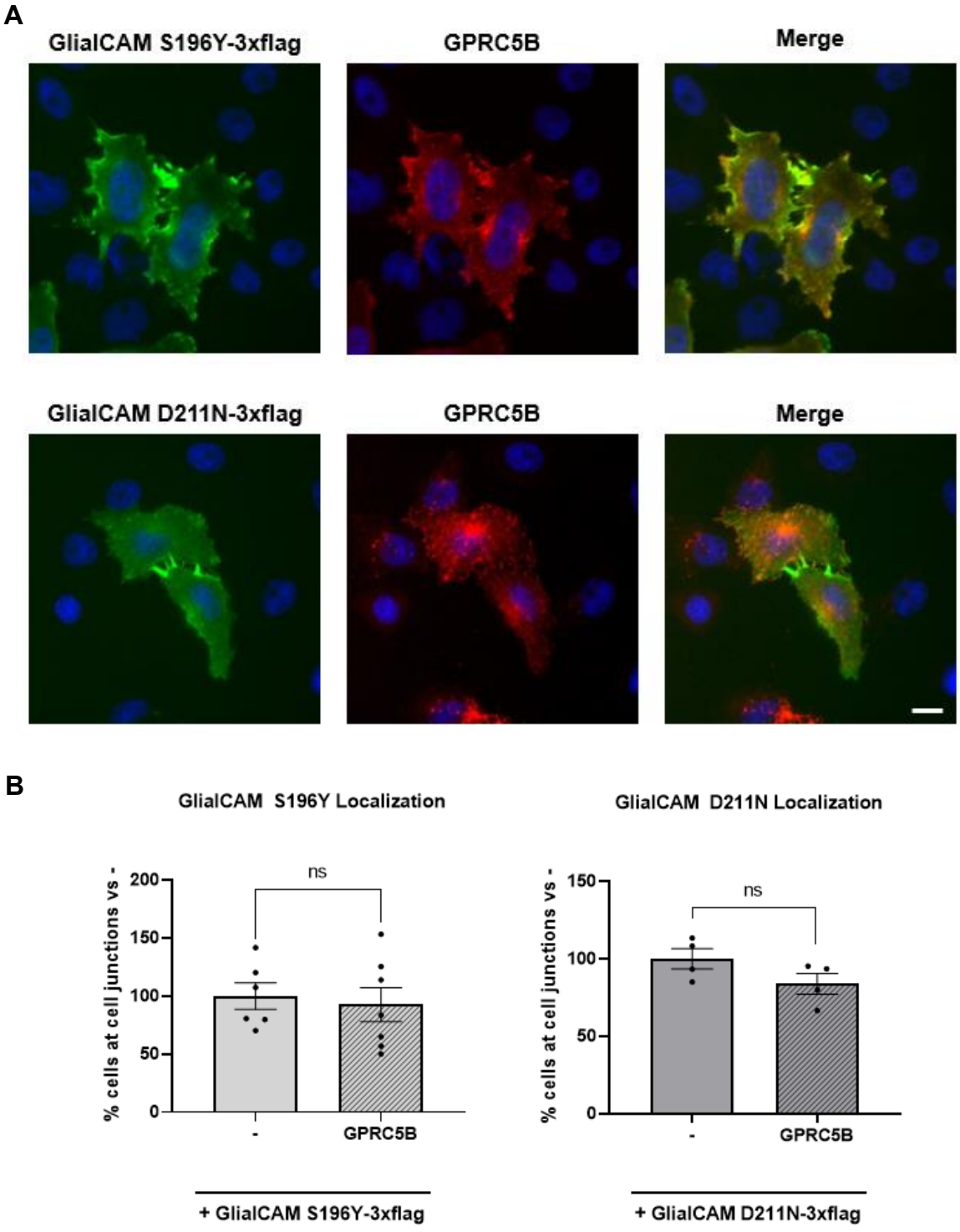


Figure 64. GlialCAM IgC2-patient derived mutants localization with GPRC5B. A. Immunocytochemistry on HeLa cells co-transfected with GlialCAM patient-derived mutant and GPRC5B. Both mutants stay enriched at cell junctions. Scale bar: 20 μm. **B.** Quantification of cells where GlialCAM is enriched at cell-cell junctions compared with the transfection of GlialCAM alone. Data comes from seven independent experiments in the case of S196Y and four, for D211N. Between 50 and 70 cells were analysed per condition. Statistical analysis was determined using Paired t-test (p-value > 0.05). Mean ± SEM is represented. ns: not significant.

Then, the co-localization was also assessed, even though observationally it already seemed that GlialCAM IgC2 variants were not co-localizing notably more than GlialCAM WT. The image quantification confirmed this observation as GlialCAM S196Y and D211N co-localized with GPRC5B in a similar manner than GlialCAM WT (Figure 65).

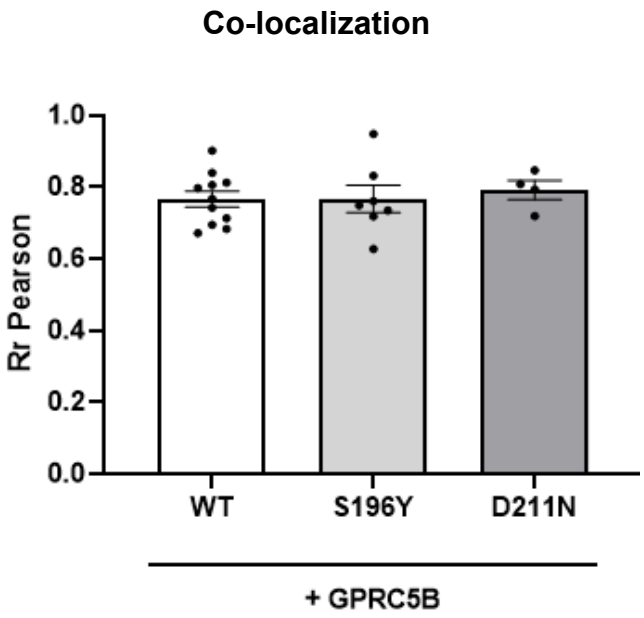


Figure 65. IgC2 GlialCAM variants co-localize with GPRC5B in a similar manner than GlialCAM WT. Colocalization of GlialCAM WT or mutants with GPRC5B. Data comes from four to eleven independent experiments. Statistical analysis was determined using One-way ANOVA (p-value > 0.05). Mean ± SEM is represented.

One possible explanation for these results was that GPRC5B and GlialCAM IgC2 variants did not interact. To confirm the specificity of the obtained information, we analysed their interaction via Split-TEV assays. It was affirmed that GlialCAM IgC2 mutants maintained direct interaction with GPRC5B (Figure 66), indicating that the lack of internalization by GPRC5B is not due to a reduced interaction of the mutant with the GPCR. Therefore, the IgC2 mutants would be resistant to the induction of GPRC5B-mediated internalization.

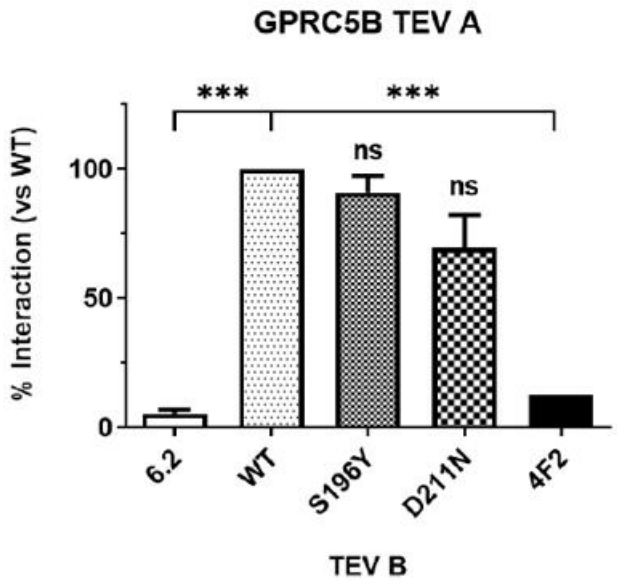


Figure 66. GPRC5B interacts with GlialCAM IgC2 patient-derived mutants. Split-TEV assays analysing the ability of GlialCAM IgC2 mutants to interact with GPRC5B in transfected HeLa cells. Data indicates the average percentage of interaction shown by each protein related to the interaction shown by GlialCAM WT protein. Empty vector pcDNA6.2TK and 4F2 were used as control groups. Data comes from three independent experiments. Statistical significance was analysed by One-way ANOVA with Bonferroni post-test: ***p-value < 0.001. ns: not significant. Experiments performed by Dr. Marta Alonso Gardón.

2.9.2 IGC2 MUTANTS DO NOT POTENTIATE β-ARRESTIN SIGNALLING

Going back to the NanoBiT signalling studies, the only effect observed in GPRC5B signalling caused by the presence of GlialCAM was the potentiation of GPRC5B and β-arrestin interaction (Figure 43). This potentiation was also seen in those cells transfected with I176dup and A177dup GPRC5B (Figure 55). So, we hypothesized that the GlialCAM IgC2 variants might not be able to potentiate GPRC5B – β-arrestin interaction. With this aim, HEK293T cells were transfected with the needed constructs and NanoBiT assays were performed. Results are plotted in Figure 67 where it can be noticed that both mutants fail to potentiate the β-arrestin, in opposite to GlialCAM WT that significantly increases the interaction between GPRC5B and β-arrestin.

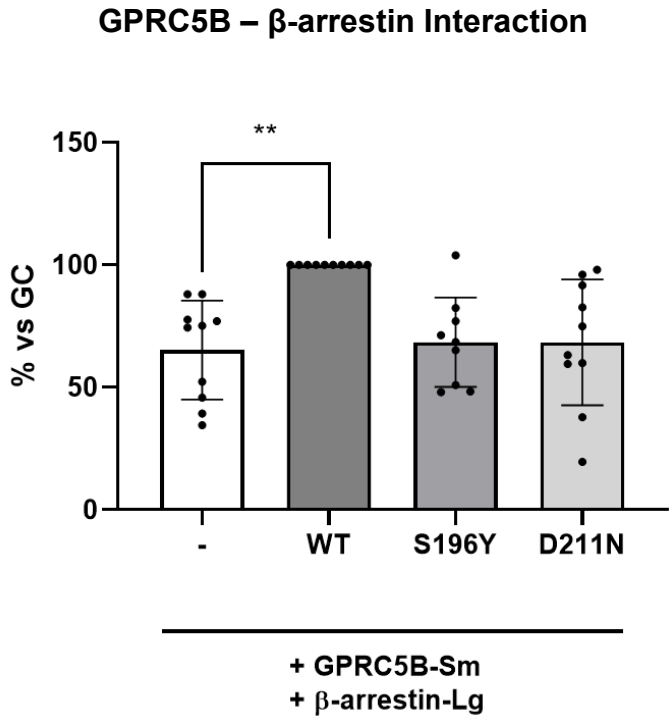


Figure 67. Patient-derived mutants do not potentiate β-arrestin signalling with GPRC5B. NanoBiT assays analysing the ability of GlialCAM IgC2 mutants to affect the interaction between GPRC5B and β-arrestin in HEK293T cells. Data comes from ten independent experiments. Statistical significance was analysed by One-way ANOVA (*** p < 0.005) with Tukey's multiple comparisons post-test. **p-value < 0.01. Experiments performed in collaboration with Guillem Pont Espinós.

2.10 2D MODEL OF GLIALCAM INTERACTIONS

As previously explained (*see Introduction section 1.3.3*), GlialCAM presents two extracellular domains: IgV and IgC2. Most *GLIALCAM* mutations causing MLC are found on the IgV domain, although there are described two mutants located on the IgC2 domain: p.S196Y and p.D211N. These mutants do not show a trafficking defect, meaning that the protein is expressed and it is perfectly localized (Arnedo et al., 2014b). Previous studies done in p.S196Y and D211N concluded that they were resistant to the induction of GPRC5B-mediated internalization. Nevertheless, we wanted to comprehend how these mutations confer resistance to endocytosis to GlialCAM protein.

In collaboration with Dr. Juan Fernández-Recio, a computer-based model for GlialCAM extracellular interactions was obtained. It allowed us to generate a GlialCAM homo-dimerization model that can be found in one of our published works, which highlights the importance of cis- and trans- interactions in GlialCAM (Elorza-Vidal et al., 2020).

Deeply analysis of the 2D model revealed that IgC2 patient-derived mutations were located in the two outward-facing α-helices of this domain. Some authors had previously described the ability of similar IgCAM molecules to form zipper-like structures through lateral interactions between IgC2 domains (Kamiguchi and Lemmon, 2000). Taking this characteristic into consideration and after seeing the relevance of GlialCAM interactions with itself in cis- and trans- (Elorza-Vidal et al., 2020), we hypothesized that IgC2 domains could be relevant for lateral GlialCAM interactions. GlialCAM oligomerization is thought to be crucial for its correct function and we considered that the α-helices motifs of IgC2 may be participating in these putative lateral interactions.

This work was done in collaboration with Dr. Adrià Pla Casillanis and Dr. Efren Xicoy Espaulella. First, mutagenesis studies were performed in which individual amino acids from the α-helices were replaced by cysteines. The objective was to perform cross-linking assays to determine whether the substitute residue would form disulfide bonds with the same residue from the adjacent IgC2 domain.

The first α-helix comprises residues from 191 to 196 and will be referred to as α-helix 1. The second one includes residues from 209 to 211 and will be named α-helix 2. In Table 11, the residues for each helix are enlisted. The proposed computer-based model is shown in Figure 68.

α-helix 1	**α-helix 2**
Serine 191	Methionine 209
Arginine 192	Glutamic acid 210
Methionine 193	Aspartic acid 211
Leucine 194	
Leucine 195	
Serine 196	

Table 11. Residues composing α-helix 1 and 2.

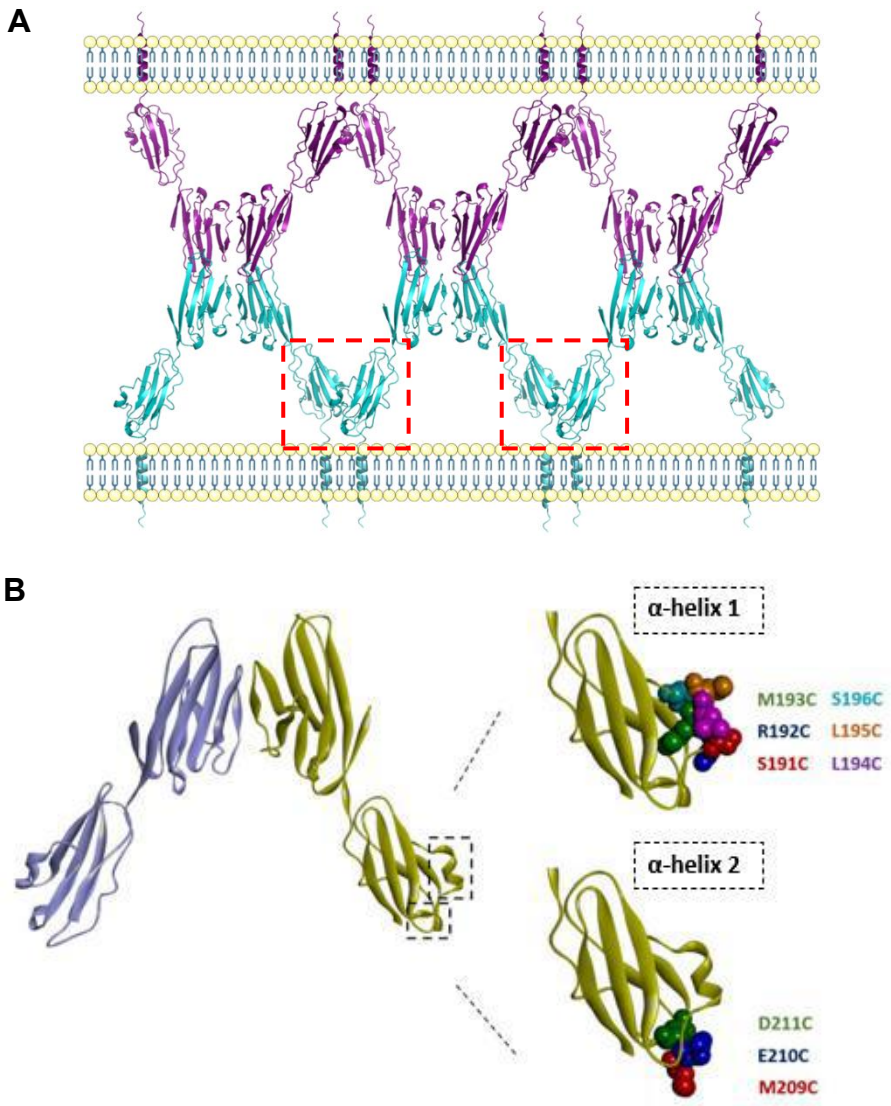


Figure 68. Bioinformatic model of GlialCAM. A. GlialCAM forms interactions in *cis*- and in *trans*- forming a zipper-like structure. In red, the putative lateral interactions by the IgC2 domain are marked. **B.** Zoom of the α-helices with the mutated residues in sphere representation.

For this study, the formation of cross-links through disulfide bonds was analysed via WB. WB experiments were performed in reducing and non-reducing conditions. This was achieved by adding or not DTT, a strong reducing agent in the protein extracts. If the formation of dimers was specific, the WB band corresponding to dimeric GlialCAM should appear in the samples not treated with DTT and disappear in the ones treated. As seen in Figure 69, the WB band corresponding to dimeric GlialCAM (around 150 kDa) appeared in cysteine mutants from both α-helices. The more noticeable dimers were found in mutants S191C and R192C from α-helix 1; and E210C and D211C from α-helix 2. Those protein extracts treated with DTT did not present a band around 150 kDa and only showed a band at 72 kDa, corresponding to monomeric GlialCAM.

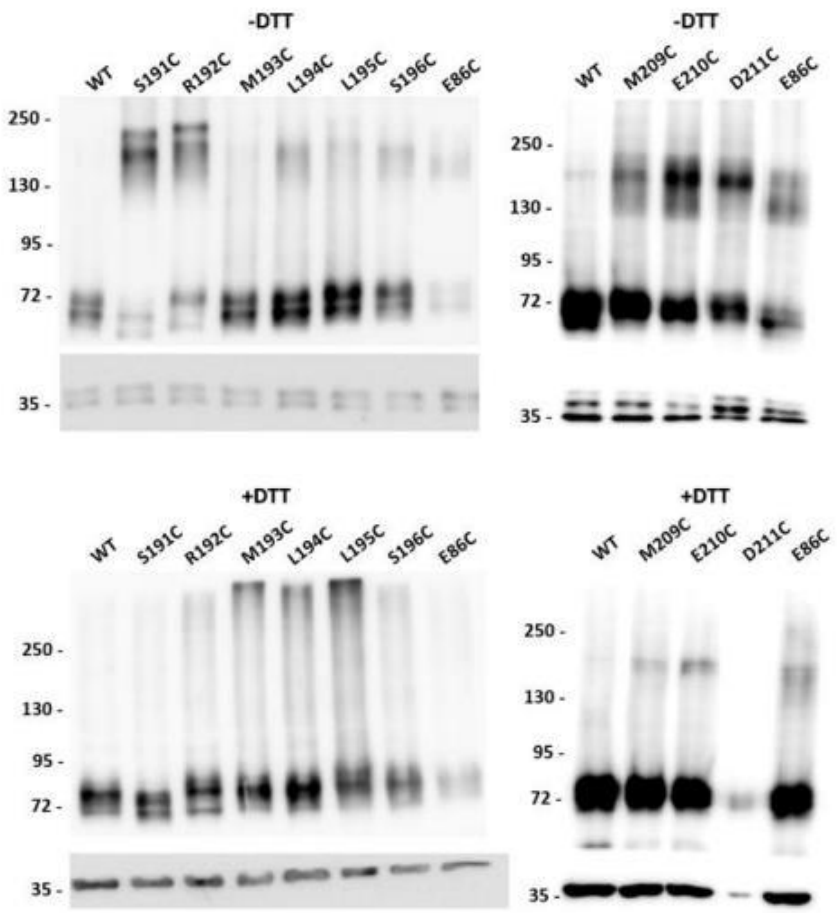


Figure 69. IgC2 GlialCAM variants form cross-links. WB experiments done with protein extracts treated or not treated with DTT 100 mM and heated at 95°C from transfected HeLa cells. 30 µg of protein extract were loaded in a 7.5% acrylamide gel. E86C variant was used as a positive control for specific cross-linking (previously validated in the laboratory). Immunodetection was directed against the FLAG tag of the variants. Actin (36 kDa) was used as a loading control. Data comes from three independent experiments. Experiments done by Dr. Adrià Pla Casillanis and Dr. Efren Xicoy Espaulella.

Immunofluorescence assays were also performed where it was observed that GlialCAM mutants S191C, R192C, E210C and D211C showed an increased number of cells with GlialCAM enriched at cell-cell junctions. This augment was statistically significant only in the S191C mutant, where 90% of cells display GlialCAM concentrated at junctions (Figure 70).

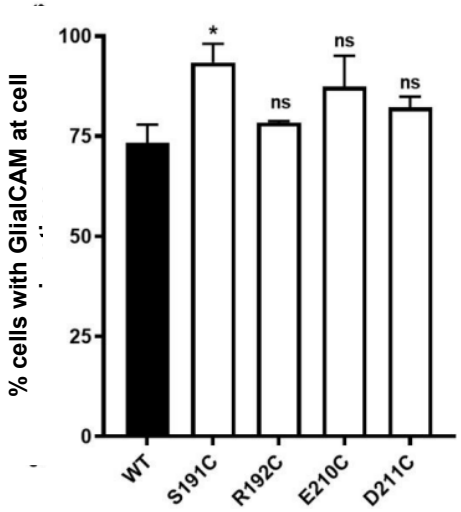


Figure 70. GlialCAM cysteine-mutants are enriched at cell junctions. Quantification of immunocytochemistry of transfected HeLa cells. GlialCAM was detected using an antibody against the FLAG tag that they had incorporated. Data comes from three independent experiments. Statistical analysis was determined by One-way ANOVA with Bonferroni post-test. * $p < 0.05$ compared to GlialCAM WT. ns: not significant. Mean ± SEM is represented. Experiments performed by Dr. Adrià Pla Casillanis and Dr. Efren Xicoy Espaulella.

2.11 GLIALCAM S191C ANALYSIS

Results from WB and immunofluorescence analysis revealed that some cysteine mutants generated disulfide bonds with themselves and they were found enriched at cell junctions. Considering these results, we hypothesized that the mutants would have an altered pattern in a SEC profile as the presence of dimers would be greater than in the WT, whose pattern would be shifted towards monomeric forms. For this experiment, the protein had to be purified. With this aim, a detergent screening was set up for choosing a suitable detergent and GlialCAM WT protein was purified before analysing the S191C mutant.

2.11.1 GLIALCAM DETERGENT SCREENING

For the detergent screening, HEK293-6E suspension cells were transfected with GlialCAM-GFP and after 48 hours of transfection, the pellet was split in parts that were solubilized with different detergents. Fluorescence was read before and after solubilization to obtain the percentage of solubilization for each detergent (Table 12). All the tested detergents were added at 1% together with 0.2% of cholesteryl hemisuccinate

(CHS), except digitonin and GDN that did not have CHS. In most detergents, the protein was hugely solubilized, being octylglucoside, nonylglucoside and nonylmaltoside the less solubilizing detergents for GlialCAM with 54, 64 and 73% of solubilization rate, respectively. Then, each solubilized protein was injected for a gel filtration, fractions were collected and their fluorescence read.

Detergent	% Solubilization
$C_{12}E_8$	100
CY-6	90
DM	100
DDM	79
Digitonin	100
GDN	79
LAPAO	97
LDAO	99
LMNG	75
Nonylglucoside	64
Nonylmaltoside	73
Octylglucoside	54

Table 12. GlialCAM solubilization efficiency for different detergents. Percentage of solubilization was calculated by dividing the fluorescence before and after solubilization. 100 µL for each condition was put in a black 96-well plate and fluorescence was read in a fluorometer. Data comes from one experiment.

In Figure 71, FSEC profiles are presented. It can be clearly seen that most samples presented two fluorescent peaks, being the smaller one (the one that appears later, at a higher volume) more consistent throughout the samples. Based on the approximated weight of those peaks, we determined that the first peak would correspond to GlialCAM and the second one, to GFP alone. From the twelve detergents tested, GlialCAM peaks presented notorious differences. The protein samples solubilized with GDN and LMNG did not show a fluorescent peak for GlialCAM. Taking into consideration both percentages of solubilization and FSEC profiles, octylglucoside, nonylglucoside and nonylmaltoside were also discarded to work with. Some peaks, such as the one for LAPAO and digitonin, were remarkably smaller than the others. From the rest, the peak present in both CY-6 and LDAO conditions was not symmetric. The remaining possible detergent candidates were $C_{12}E_8$, DM and DDM. DM was chosen as the working detergent for GlialCAM solubilization because of the symmetry of the peak and for having the highest fluorescence, hence, the most protein.

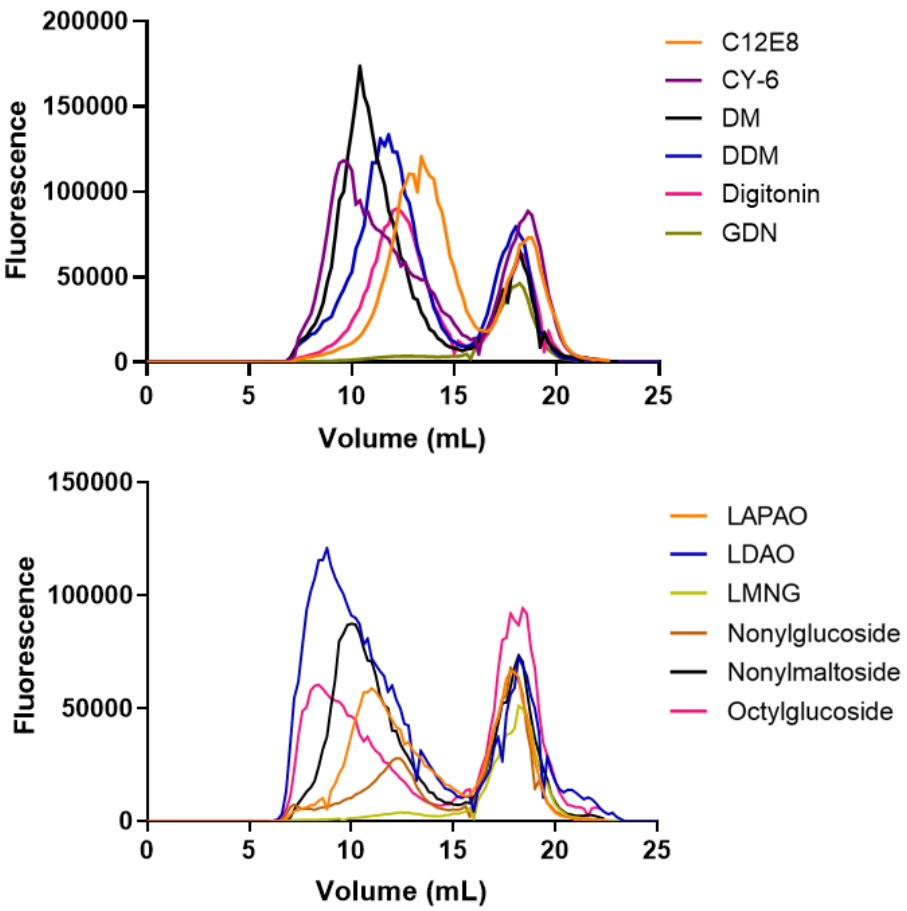


Figure 71. GlialCAM detergent screening. FSEC profiles of GlialCAM solubilized 1 hour with different detergents at 1%. Samples were run in Sepharose 6 Increase 10/300 GL column, at 0.4 mL/min flow rate in Tris/NaCl buffer. Fractions of 200 µL were collected in black 96-well plates and fluorescence was read.

2.11.2 GLIALCAM WT PURIFICATION

Having decided to work with DM as the solubilizing detergent, a first test of GlialCAM purification was carried out. For this, suspension cells were transfected with the construct GlialCAM-3xflag and the pellet was solubilized 1 hour at 1% DM plus 0.2% CHS. GlialCAM was purified using Flag resin and eluted with flag peptide. The eluted output was injected in the ÄKTA for a gel filtration. In Figure 72, the results from GlialCAM purification are displayed.

Mass Spectrometry verified that the most abundant protein present in the silver stain gel was GlialCAM-3xflag. For its analysis, the most prominent band from the second peak in the WT protein was send to IRB Barcelona Mass Spectrometry and Proteomics Core Facility.

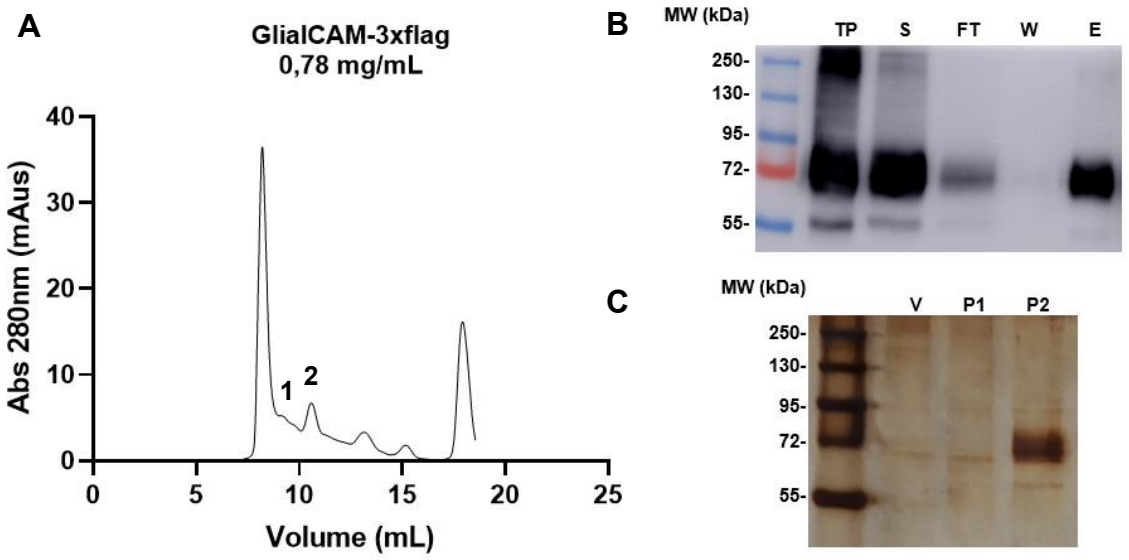


Figure 72. GlialCAM protein purification. A. SEC profile of purified GlialCAM-3xflag. After a large void, two distinguished peaks are observed being the second one more defined. The column used was SuperDex Increase 200 10/300 GL at a flow of 0.3 mL/min in Tris/NaCl buffer. **B.** WB of the purification process. GlialCAM was purified using flag resin and eluted with flag peptide. **C.** Silver Stain of the SEC peaks. Most GlialCAM protein is found on the second peak.

TP: total protein, S: solubilized, FT: flowthrough, W: wash, E: elution, V: void, P1: peak 1, P2: peak 2.

2.11.3 GLIALCAM S191C PURIFICATION

Once the purification conditions were set up, the S191C mutant was tested. For this, two pellets transfected with GlialCAM WT and S191C separately were purified using the Flag resin. Comparing both SEC profiles, a difference in the peaks was observed. As depicted in Figure 73A, the WT protein showed a profile where the second peak is higher than the first one, whereas in the S191C, the first peak was larger than the second one.

To confirm that the mutant established specific disulfide bridges with itself, the peaks from the gel filtration were collected and a Silver Stain was performed in reducing and non-reducing conditions. The peaks from the GlialCAM WT purification showed a band around 72 kDa with and without the presence of the reducing agent, β-mercaptoethanol in this case. Contrarily, when the peaks from the GlialCAM S191C purification were not treated with the reducing agent, a band appeared at 250 kDa. However, when β-mercaptoethanol was added to the samples, a band was present at 72 kDa (Figure 73B). The obtained results indicated that the S191C mutant had a higher tendency to oligomerize than the WT. At the same time, Silver Stain gels revealed that the increased

weight in the S191C mutant was caused by the formation of a cross-link formed by a disulfide bridge.

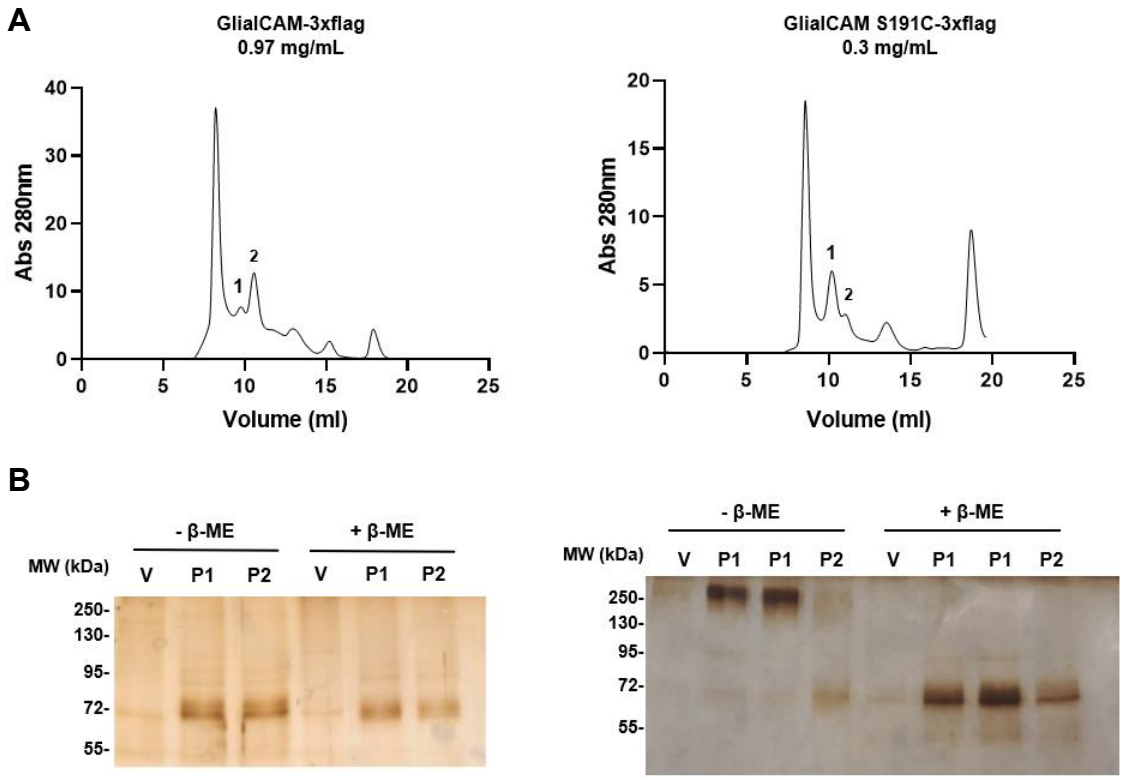


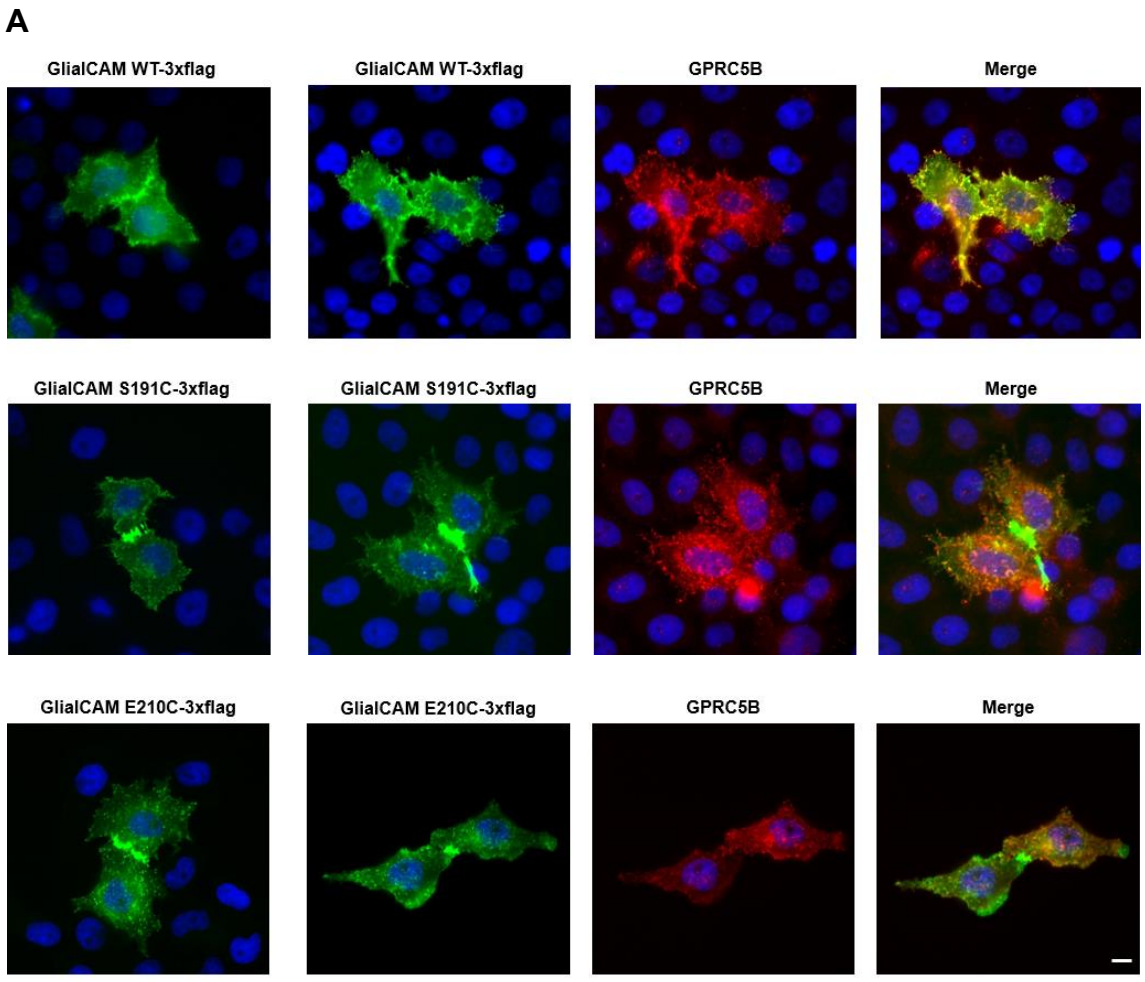
Figure 73. GlialCAM WT and S191C protein purification. A. SEC profiles of purified GlialCAM WT (left) and S191C (right). After a large void, two distinguished peaks are observed in both profiles. The column used was SuperDex Increase 200 10/300 GL at a flow of 0.3 mL/min. **B.** Silver Stain of the SEC peaks of GlialCAM WT purification (left) and S191C mutant (right) with and without β-mercaptoethanol.

V: void, P1: peak 1, P2: peak 2, β-ME: β-mercaptoethanol.

2.12 CYSTEINE IGC2 GLIALCAM MUTANTS ARE RESISTANT TO ENDOCYTOSIS

Afterwards, we wanted to assess whether the cross-link variants showed a defect in protein localization in the presence of GPRC5B as occurred with disease-related GlialCAM variants. We decided to work with the variants S191C from helix-α1 and E210C from helix-α2. We hypothesized that these cysteine variants would stay at cell-cell junctions even with the presence of GPRC5B, as it was described for the S196Y and D211N patient-derived mutations.

The results are shown in Figure 74. As hypothesized, S191C and E210C were found to be perfectly localized at cell junctions in transfected HeLa cells. When these cells where co-transfected with GPRC5B, the expression of GlialCAM protein remained at cell junctions. However, when quantifying the immunofluorescence images and analysing these data statistically, the results showed that the variant E210C significantly decreased its levels at the junctions in the presence of GPRC5B compared to its location when it was expressed alone. Still, the reduction was not as major as the one seen in the WT group. In this case, the localization of GlialCAM WT at cell junctions decreased around 60% in the presence of GPRC5B; whereas for the E210C variant, there was a 25% decline. This significance could be explained by a low variability among the different cell cultures from both conditions expressing GlialCAM E210C with and without GPRC5B compared to the other variants that presented higher variabilities.



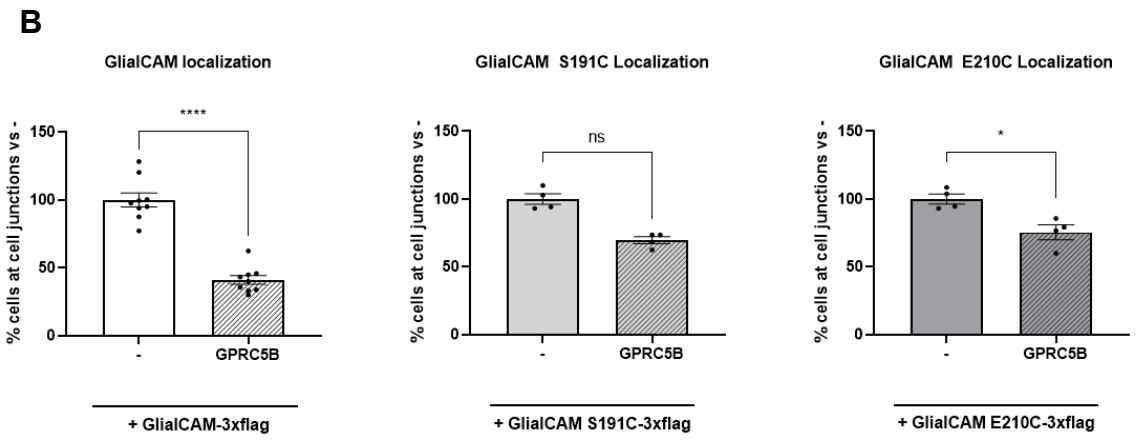


Figure 74. GlialCAM WT, S191C, and E210C protein localization. A. Immunocytochemistry on HeLa cells transfected with GlialCAM WT and IgC2 variants alone (on the left) and with GPRC5B (on the right). Both mutants stay at cell-cell junctions. Scale bar: 20 µm. **B.** Quantification of cells where GlialCAM is enriched at cell-cell junctions compared with the transfection of GlialCAM alone. Data comes from four independent experiments. Between 50 and 70 cells were analysed per condition. Statistical analysis was determined using Paired t-test (ns p-value > 0.05; * p < 0.05; **** p < 0.0001). ns: not significant. Mean ± SEM is represented.

We concluded that GPRC5B- mediated internalization was impaired in S191C and E210C variants, similarly to the observed resistance to endocytosis for the disease-related mutations S196Y and D211N. With these experiments, we confirmed that an altered stabilization of oligomeric structures at the plasma membrane was linked to the pathophysiology of MLC.

2.12.1 IGC2 CYSTEINE MUTANTS DO NOT POTENTIATE β-ARRESTIN SIGNALLING

Seeing that both cysteine mutants presented the same phenotype as the patient-derived IgC2 variants, we tested whether S191C and E210C would be able to potentiate the GPRC5B and β-arrestin interaction. Since S196Y and D211N variants did not potentiate this signalling (Figure 67), we expected that the cysteine mutants would not potentiate it neither. NanoBiT assays confirmed that the signalling potentiation did not occur in the mutants S191C and E210C (Figure 75).

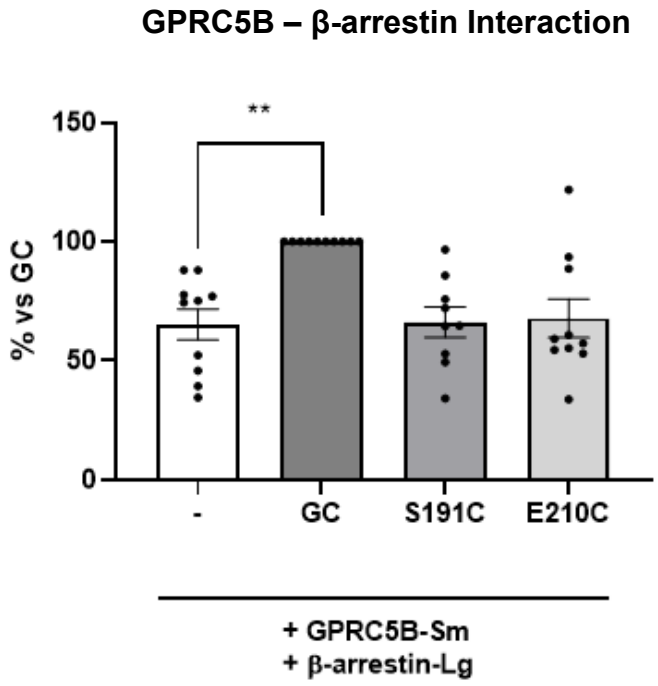


Figure 75. GlialCAM cysteine-mutants do not potentiate β-arrestin signalling with GPRC5B. NanoBiT assays analysing the ability of GlialCAM IgC2 mutants S191C and E210C to affect the interaction between GPRC5B and β-arrestin in HEK293T cells. Data comes from ten independent experiments. Statistical significance was analysed by One-way ANOVA (*** p < 0.001) with Tukey's multiple comparisons post-test. **p-value < 0.01. Experiments performed in collaboration with Guillem Pont Espinós.

2.13 BIOINFORMATIC MODEL FOR GLIALCAM HOMODIMERS INTERACTING THROUGH THE IGC2 DOMAIN

With all the collected information on cysteine-mutants in IgC2, we collaborated again with Dr. Juan Fernández-Recio, who further refined the model of GlialCAM interactions considering the new experimental restrictions. This new model considered IgC2 lateral interactions in addition to IgV cis-and trans- interactions (Figure 76).

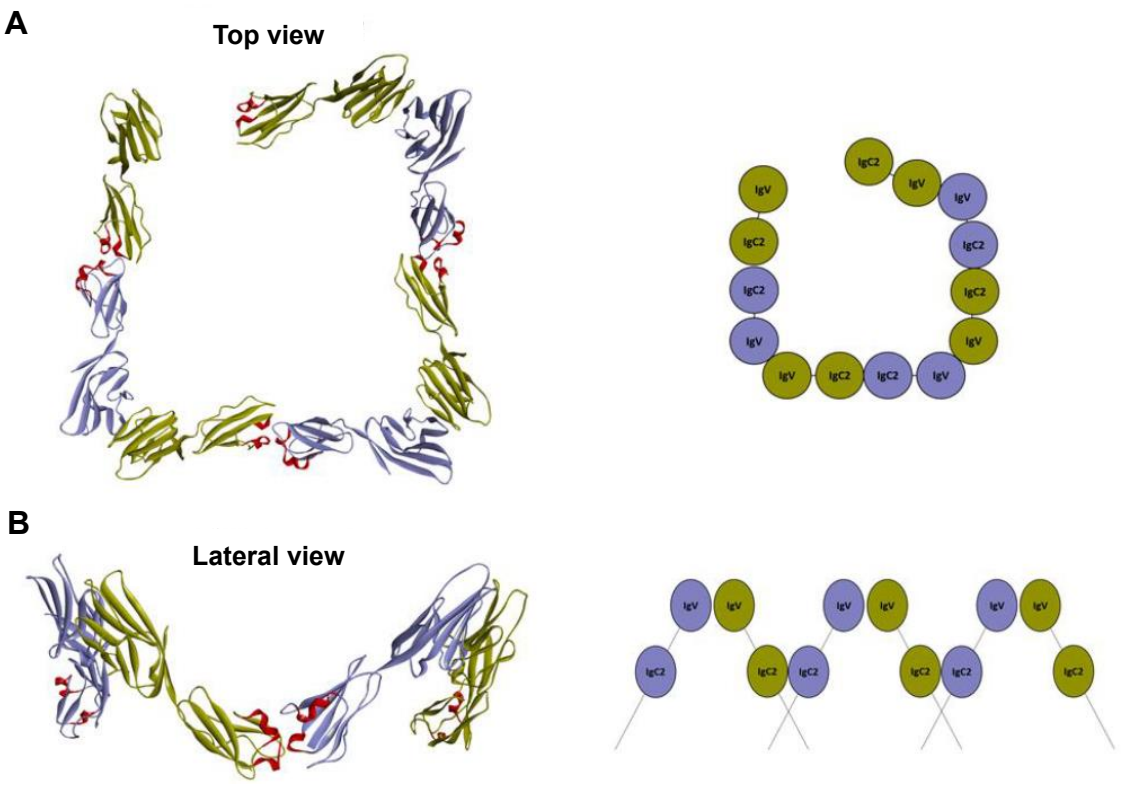


Figure 76. Bioinformatic model of the possible interaction that GlialCAM forms. A. Top view of GlialCAM molecules (in green and purple) interacting through the IgV and the IgC2 domain. In red, the two α-helices from the IgC2 domain are pointed out. **B.** Lateral view of the model (left) and a schematic representation of it (right).

3. STRUCTURAL RESOLUTION OF MLC1

3.1 INTRODUCTION

The obtention of the tridimensional structure of proteins can reveal important indications about their function (Bai et al., 2015). At the same time, it can facilitate the design of drugs to allow therapeutic interventions (Suades et al., 2019). Since the exact function of MLC1 remains unknown, in this thesis we have worked towards solving the 3D structure of this protein. We believe that the obtention of its structure would expand the knowledge about MLC1 function and MLC pathophysiology, while also being crucial for future therapeutic strategies.

This work was started by Dr. Efren Xicoy-Espaulella, who worked with different tags fused to MLC1 for the purpose of finding the best method to purify the protein. After some analysis, it was decided to work with the TwinStrep tag fused at the N-terminus of MLC1. This construct was designed by Dr. Nick Berrow from the Protein Expression Core Facility at IRB.

For a successful homogeneously purified protein, different features must be taken into consideration. One crucial step in membrane protein purification is the solubilization, where proteins are removed from the plasma membrane and incorporated into detergent micelles. During this process, the disruption of the native phospholipid environment is needed whilst, paradoxically, the physiological conformation of the protein must be maintained (Ratkeviciute et al., 2021). For this reason, choosing the right detergent is a critical step for the achievement of the purified protein. After a detergent screening, it was agreed on working with the detergent digitonin. Posterior experiments established that the best solubilization conditions for MLC1 with digitonin were to solubilize 2 hours at 2% of digitonin.

3.2 OPTIMIZATION OF MLC1 PURIFICATION

Once the construct and the detergent were chosen, it was time to optimize other parameters in the purification protocol for MLC1. Our objective was to obtain a homogeneous sample that could be used to solve the 3D structure of our protein in cryo-EM imaging. For the optimization, we evaluated different features such as the co-expression with GlialCAM, the ratio protein:detergent and the proper conditions for protein freezing.

3.2.1 GLIALCAM CO-EXPRESSION

Previous studies from the group demonstrated that GlialCAM protected MLC1 from degradation associated to the endoplasmatic reticulum (Capdevila-Nortes et al., 2013) and it allowed MLC1 localization at cell junctions (López-Hernández et al., 2011b, López-Hernández et al., 2011a). It was also determined that digitonin was the only detergent maintaining the interaction between MLC1 and GlialCAM (López-Hernández et al., 2011a).

Taking all this information into account, we hypothesized that the co-expression of MLC1 with GlialCAM would increase the efficiency of the purification process. Adding GlialCAM also pursued a second objective since its prominent extracellular domain could be useful when analysing particles at cryo-EM. Moreover, it was thought that solving the structure of MLC1 together with GlialCAM would be relevant for understanding their physiological functions and gaining knowledge about their interaction.

To test our hypothesis, HEK293-6E cells were transfected with TwinStrep-MLC1 alone or together with GlialCAM-3xflag. Then, MLC1 was purified in both groups in parallel and the SEC profiles of the purified protein were compared as seen in Figure 77A. It can be observed that in both conditions, the profile looked alike, although the addition of GlialCAM increased the absorbance levels. Moreover, the proportion between peaks was altered because in the condition with MLC1 alone the void and the first peak had the same height, whereas with GlialCAM the first peak was higher than the void. The WB analysis in Figure 77B showed that GlialCAM was eluted in the MLC1 purification.

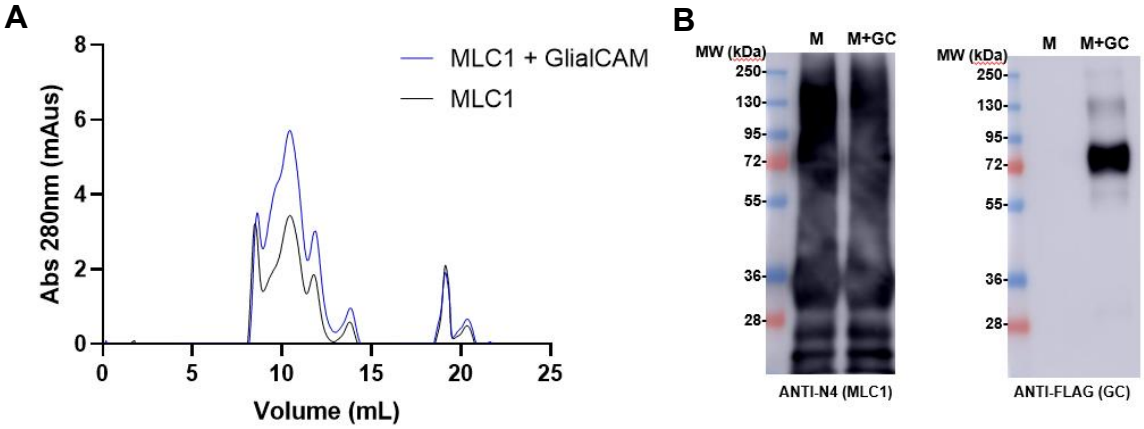


Figure 77. MLC1 protein purification alone or co-expressed with GlialCAM. A. SEC profile of TwinStrep-MLC1 protein purified alone or in the presence of GlialCAM-3xflag. Both samples were solubilized digitonin, incubated with StrepTactin Superflow resin and were run in SuperDex Increase 200 10/300 GL at a flow of 0.3 mL/min in Tris/NaCl buffer plus detergent. **B.** WB of the purified protein detecting MLC1 and GlialCAM.

M: MLC1, GC: GlialCAM.

3.2.2 RATIO PROTEIN: DETERGENT

As previously stated, solubilization is key for good protein purification and it is recommended to be optimized. Apart from detergent selection and concentration, the ratio of protein: detergent needs to be considered as delipidation of the plasma membrane is driven by detergent concentration in purification buffers as well as intrinsic properties of detergent micelles (Urner et al., 2022).

To examine the most optimal protein: detergent ratio for MLC1 purification, the following experiment was set up. A 2 L cell culture was co-expressed with TwinStrep-MLC1 and GlialCAM-3xflag. Then, it was purified separating the volume in three for working with three different solubilization ratios of protein: detergent (w/w). The ratios used were 1:4, 1:8 and 1:12 and the total protein concentration was quantified via BCA. In all three conditions, digitonin was added at 2% of the final volume of the solubilized buffer. SEC profiles are shown in Figure 78. The mobility of peaks did not change among conditions, however, the proportion of peaks from the same conditions did vary. For instance, in the ratio 1:12 group, the proportion between the void and the second peak was greater than in the 1:4 and 1:8 groups.

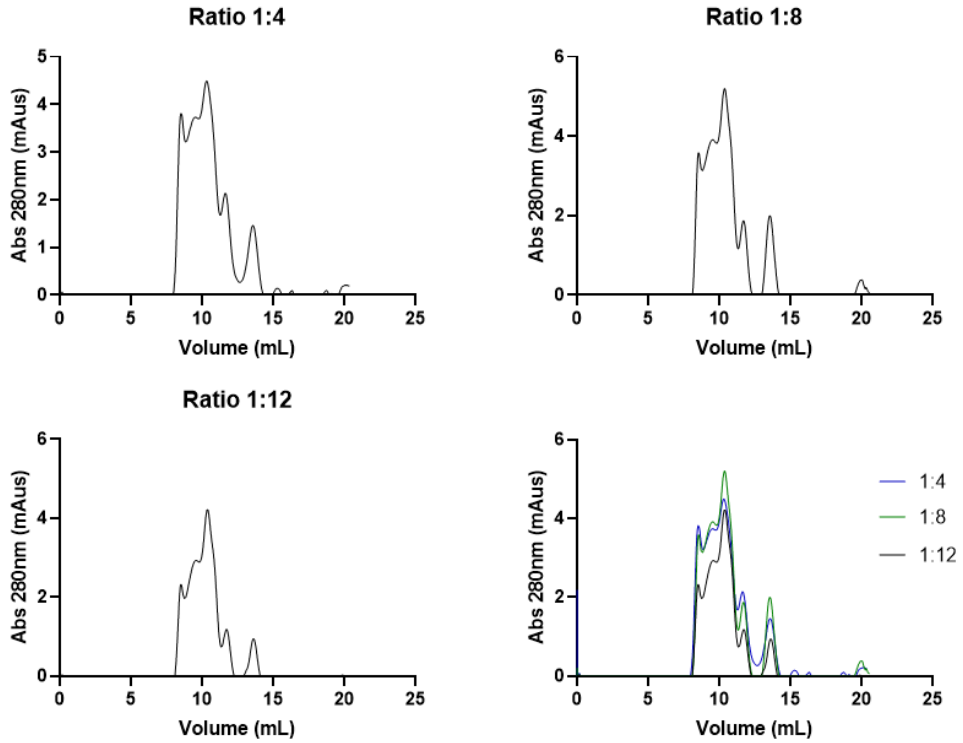


Figure 78. MLC1 protein purification using different protein: detergent ratios. SEC profile of TwinStrep-MLC1 protein coexpressed with GlialCAM-3xflag purified using 1:4, 1:8 and 1:12 ratios of protein: detergent. Samples were solubilized with 2% digitonin, incubated with StrepTactin Superflow resin and were run in SuperDex Increase 200 10/300 GL at a flow of 0.3 mL/min in Tris/NaCl buffer plus digitonin.

3.2.3 PROTEIN FREEZING AND THAWING

The last test required before putting the protein in grids for cryo-EM analysis was to assess whether the protein could be frozen and thawed without aggregating. Considering that our collaborators for cryo-EM were in Madrid, we needed to know if the purified protein could be sent frozen.

With this objective, MLC1 was purified from a 2 L expression of HEK293-6E cells transfected with TwinStrep-MLC1 and GlialCAM-3xflag. Once the SEC profile was obtained, peak 1, 2 and 3 were collected, frozen in liquid nitrogen and stored at -80°C for two days. After this time, peak 1 and 2 were thawed on ice, filtered, and injected again for gel filtration. SEC profiles are presented in Figure 79, where it can be observed that the frozen and thawed peaks did not show mobility differences compared to the sample before freezing. So, the protein could resist freezing and thawing, at least one cycle of it.

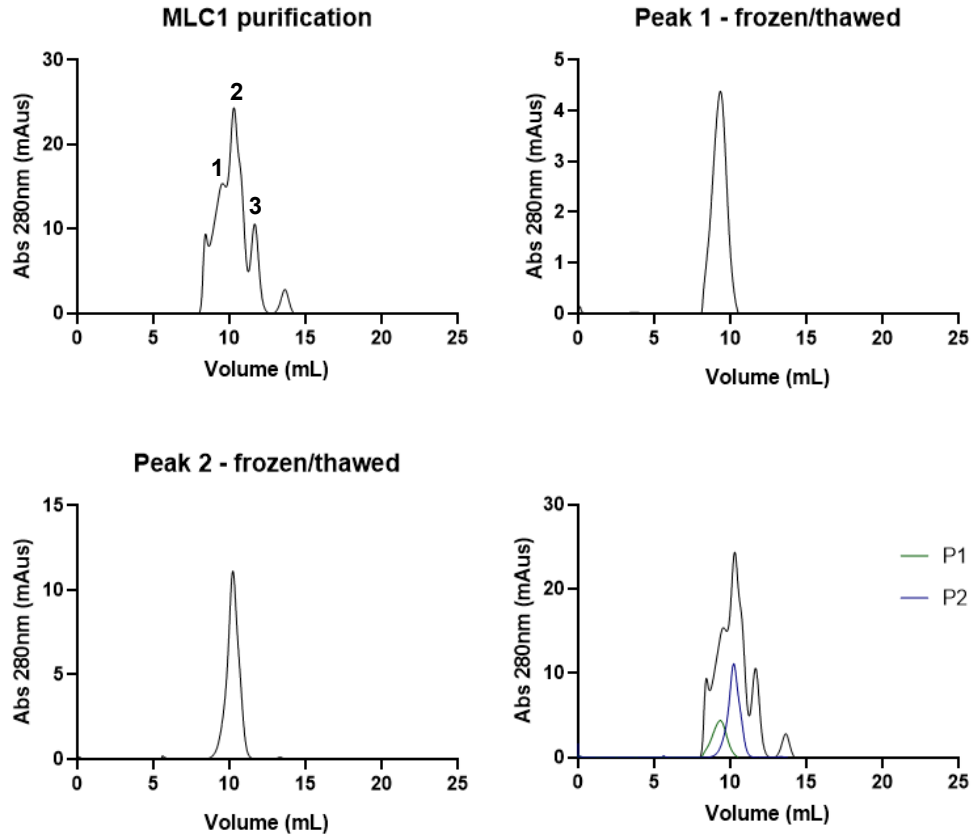


Figure 79. MLC1 protein purification freezing and thawing the peaks. SEC profile of TwinStrep-MLC1 protein coexpressed with GlialCAM-3xflag. Sample was concentrated at 1,3 mg/mL before gel filtration. Peak 1 and 2 were frozen in liquid nitrogen before storing them at -80°C for two days, then, both peaks were thawed and injected to SEC. Samples were solubilized with 2% digitonin, incubated with StrepTactin Superflow resin and were run in SuperDex Increase 200 10/300 GL at a flow of 0.3 mL/min in Tris/NaCl buffer plus digitonin.

P1: peak 1, P2: peak 2.

3.3 NEGATIVE STAINING EM

After all the optimization of the purification protocol and seeing that the protein could be sent frozen, it was time to begin cryo-EM studies. The first thing that had to be done was a negative staining.

Negative staining EM is a technique utilized to assess protein homogeneity and sample concentration before proceeding to cryo-EM. It can be useful to visualize the overall architecture of the protein complex (Gallagher et al., 2019). In negative staining, the protein of interest is embedded to a grid and coated in an electron dense stain, usually uranyl salts. Those areas occupied by the protein are not stained, giving a high contrast visualization of the protein structure, which makes the protein to appear white in a background of dark stain (De Carlo and Harris, 2011).

For the negative staining and cryo-EM we began to collaborate with the group of Dr Óscar Llorca at CNIO (*Centro Nacional de Investigaciones Oncológicas*) in Madrid. The three peaks from the freezing/thawing experiment previously shown in Figure 79 were sent to them and the PhD student Nayim González Rodríguez performed the negative staining.

Analysis of the three peaks showed the presence of enough particles to proceed with the cryo-EM imaging obtention. A representative image of the grids and selected images of single particles from all three peaks are shown in Figure 80, where it can be seen a good distribution of micelle particles among all grids. Images acquired from the first peak revealed an extra density around the micelles, which could correspond to GlialCAM. These results indicated the possibility to obtain micelles with the complex MLC1-GlialCAM in peak 1 and MLC1 alone in peaks 2 and 3.

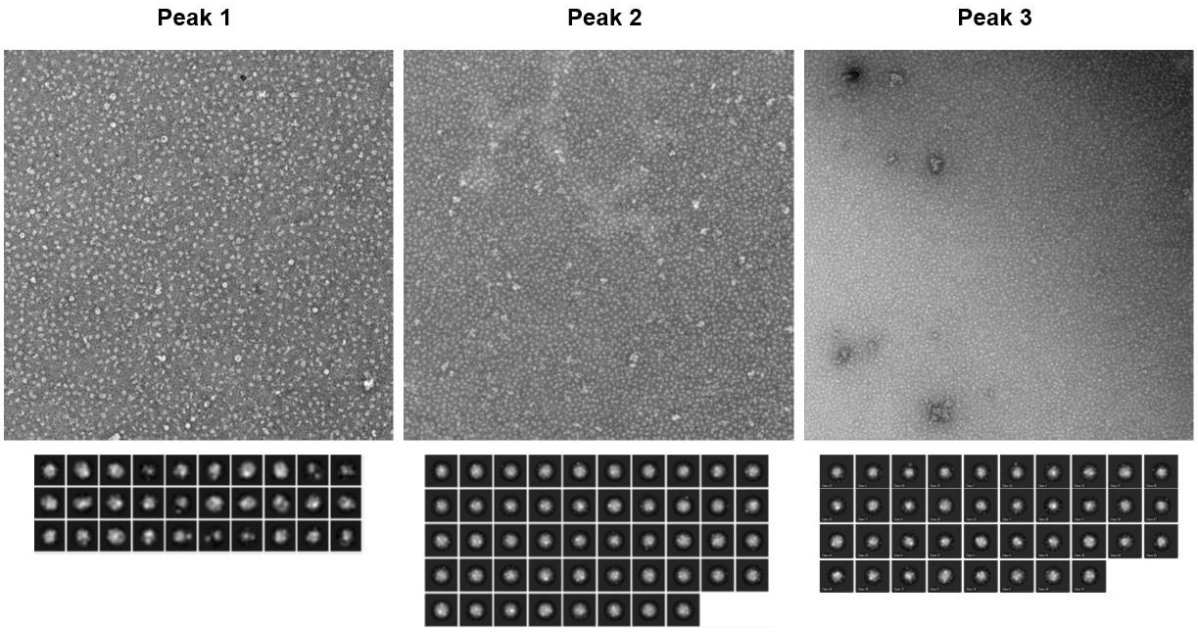


Figure 80. Negative staining for MLC1. The three peaks obtained from the SEC profile of MLC1 purification were analysed. Representative images of the stained grids for the three peaks (top) and a representation of single particles (bottom) are shown.

3.4 PRELIMINARY MLC1 STRUCTURE IN CRYO-EM

Cryo-EM is a structural biological method used to determine the tridimensional structure of biomolecules. Briefly, the purified protein is applied to a glow-discharged grid to create a hydrophobic environment where the protein will be attracted to. After removing the excess buffer, the grid is plunged into liquid ethane which causes a fast-freezing process and the formation of vitreous ice (Cabral et al., 2022). This process of vitrification preserves the sample in its natural state, but also protects it from dehydration (Renaud et al., 2018). Then, images are acquired at the rate of multiple frames per second using a direct detector camera that allows per-pixel quantifications of electron dose. Since the electron beam causes particles to move during image acquisition, the obtained particle images need to be aligned and classified into several 2D class averages based on their similar features and orientations. From this 2D classification, a 3D reconstruction can be generated (Cabral et al., 2022).

Compared to X-ray crystallography, cryo-EM imaging requires less sample and does not involve crystallization processes (Bai et al., 2015). Furthermore, cryo-EM revolutionized the structural biology field since it enabled to obtain the structure of membrane proteins or heterogeneous complexes, that are known to be difficult crystallization targets. For instance, the structure of members of the TRP channel superfamily, that had failed to

crystallize, were determined by cryo-EM in different functional states (Cheng, 2018). For the success of cryo-EM in solving protein structures, samples need to be pure, monodisperse, sufficiently concentrated, show random orientations, and not be disassembled or aggregated (Orlova and Saibil, 2011).

After the first evaluation with negative staining, it was time to begin the analyses of our protein at cryo-EM. For this, an expression of 3,5 L HEK293-6E cells co-transfected with TwinStrep-MLC1 and GlialCAM-3xflag was prepared. After the purification of MLC1, we obtained a sample concentrated at 2 mg/mL. The SEC profile is presented in Figure 81A. Peaks 1, 2 and 3 were collected, concentrated at 1, 1.75 and 1 mg/mL, respectively. Then, they were frozen in liquid nitrogen, stored at -80°C and sent to our collaborators at CNIO.

Before sending the peaks, WB and silver stain were performed to corroborate that MLC1 could be detected, and the sample was not aggregated or contaminated. The silver stain showed a faint band around 72 kDa compatible with GlialCAM and a more prominent band at 36 kDa compatible with MLC1 (Figure 81B). The posterior WB validated that the biggest band corresponded to GlialCAM, whereas the other one, to MLC1 (Figure 81C).

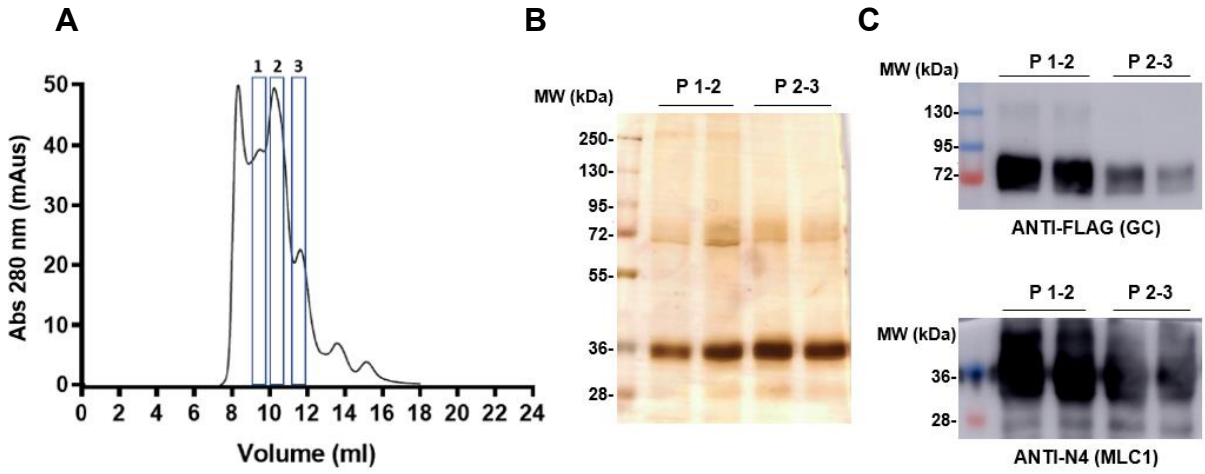


Figure 81. MLC1 protein purification. A. SEC profile after TwinStrep-MLC1 purification coexpressed with GlialCAM-3xflag. Sample was solubilized with 2% digitonin and purified using StrepTactin resin. The eluted protein was concentrated at 2 mg/mL and injected in the Superdex 200 Increase 10/300 GL at a flow of 0.3 mL/min in Tris/NaCl buffer plus digitonin. The three peaks obtained were concentrated at 1, 1.75 and 1 mg/mL, respectively. **B.** SST of the collected peaks **C.** WB of the collected peaks detecting GlialCAM (top) and MLC1 (bottom).

P: peak, GC: GlialCAM.

It was decided to prepare cryo-EM grids with the second peak since it was the most concentrated. The preparation of grids and the posterior processing was done by Nayim González Rodríguez. Grids made of different materials were analysed and Quantifoil Cu 300 mesh 0.6/1 grids that are made of copper were preferentially used. Grids were observed in a 300 kv cryo-EM at LISCB (Leicester Institute of Structural and Chemical Biology). 2D classification identified two population of particles: presenting or not presenting extramicellar densities. Preliminary processing of the population lacking extramicellar densities showed micelles in diverse orientations and faint protein densities inside of them (Figure 82A). Side and top/bottom views could be observed in the 2D averages being indicative of a homogeneous representation of all orientations. Top views looked slightly triangular, being compatible with a trimer of MLC1. The population presenting an extramicellar density consisted of only 10% of total particles and it was compatible with the presence of GlialCAM (Figure 82B). The low number of micelles with GlialCAM correlated with the low amounts of this protein seen in the silver stain gel. The signal that protrudes from the micelle looked blurred, which could be indicating intrinsic flexibility of the molecule.

After the classification and selection of 2D images, the 3D volumes were reconstructed and the heterogeneous particles were classified in six different groups (Figure 83). The grey and purple classes showed densities outside of the micelle, whereas the rest of them seemed to only contain MLC1.

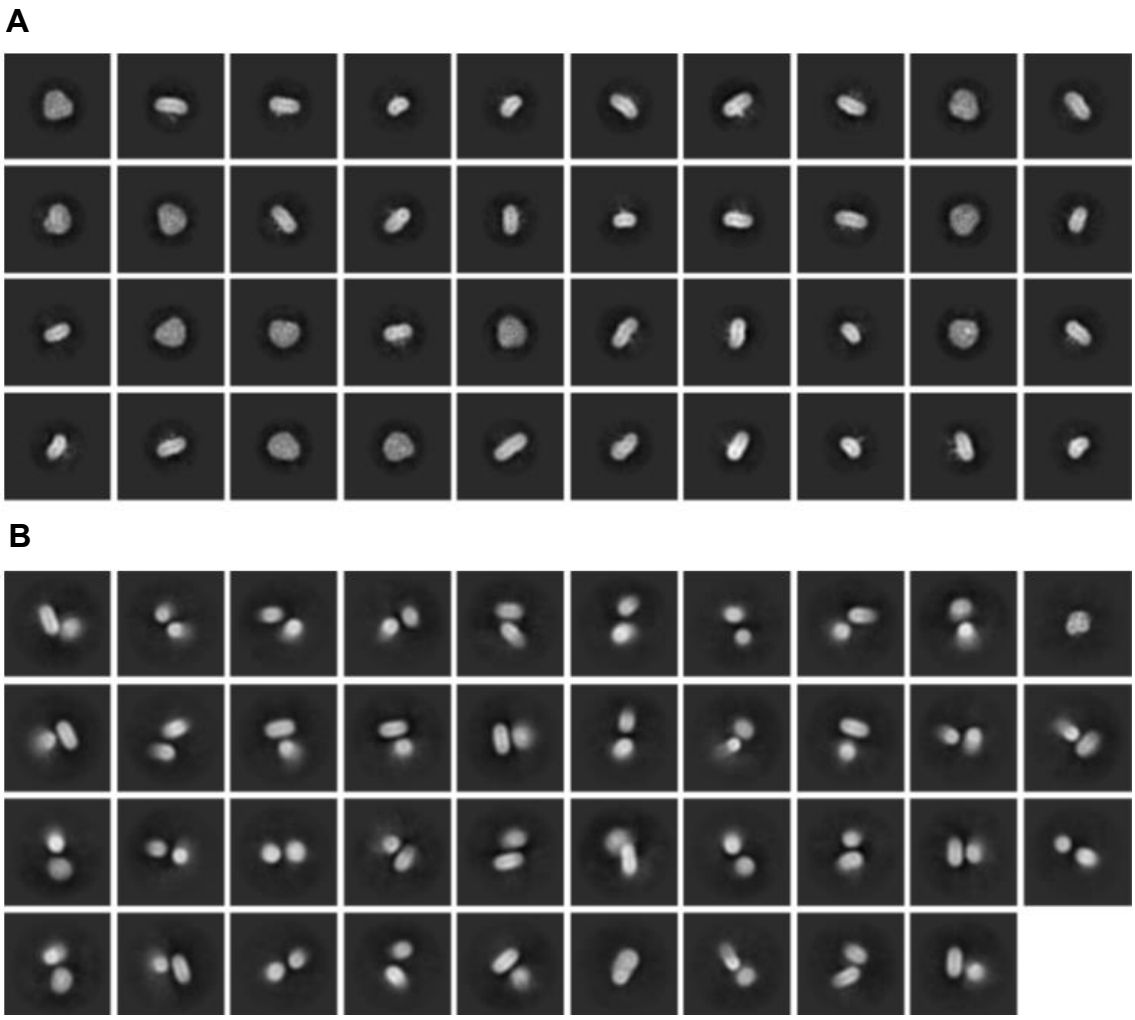


Figure 82. 2D averages from the second peak of MLC1 purification. A. 2D averages of the population of particles that do not present any extramicellar density. **B.** 2D averages of particles presenting extramicellar density.

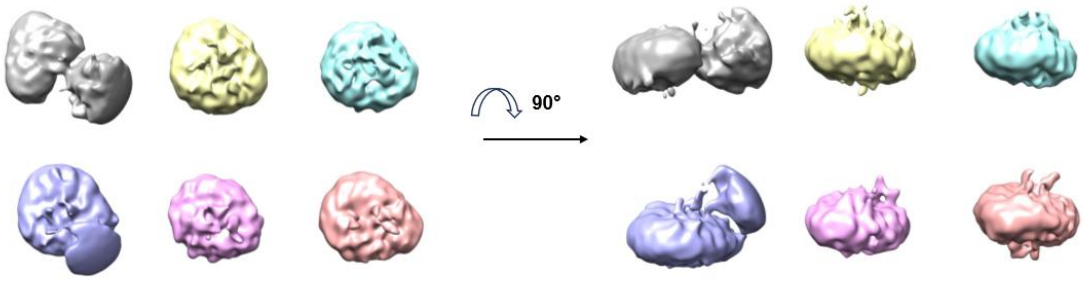


Figure 83. Whole dataset 3D classification. Six classes were found during 3D classification represented in different colors each. Top (left) and side (right) views are represented.

Further subclassification of both populations gave place to more detailed volumes. Slices through the middle of the micelle showed some tubular density compatible with the presence of transmembrane protein segments into the micelle (Figure 84). In both subpopulations, a triangular shape is observed, which would be compatible with MLC1 forming a homotrimer. In the population containing extramicellar densities, the number of particles was more limited and no extra density seemed to arise for the transmembrane region of GlialCAM.

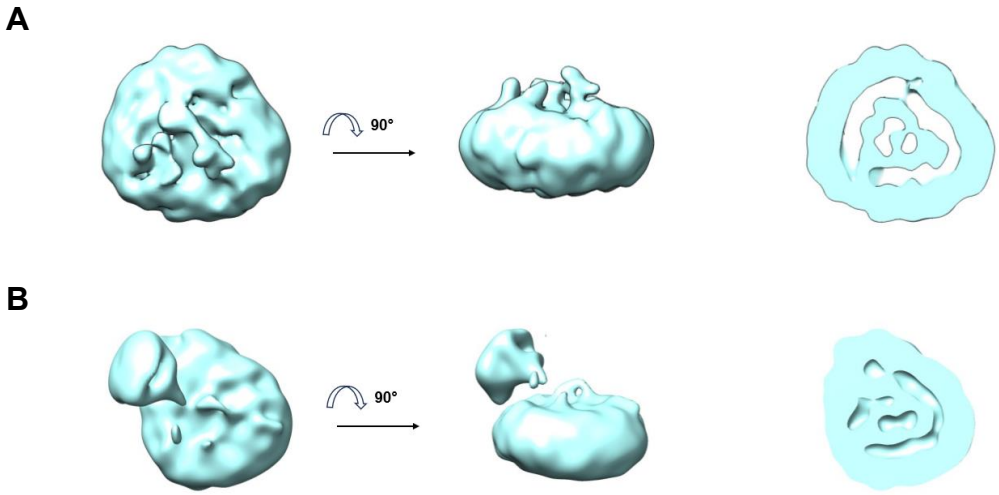


Figure 84. Subclassification of subpopulations. A. Further subclassification of the micelle-only subpopulation. Representative 3D volume after several iterations of 3D classification in top (left) and side (centre) views. On the right, sliced view of the micelle. **B.** Further subclassification of the extramicellar density subpopulation. Representative 3D volume after several iterations of 3D classification in top (left) and side (centre) views. On the right, sliced view of the micelle.

The obtained structures were compatible with resolutions in the range of 12.5-15 Å which were not enough to create a protein model. At that moment, refinement strategies failed to obtain higher resolutions. However, with these results and using predicted structures by AlphaFold2, rough fittings were performed for trimeric MLC1 alone and with GlialCAM. These fittings served for the only purpose of checking whether the dimensions were what we would expect for our protein of interest. The fittings are depicted in Figure 85. It seemed that our preliminary 3D results were compatible with the predicted structure, although further refinement and optimization were needed.

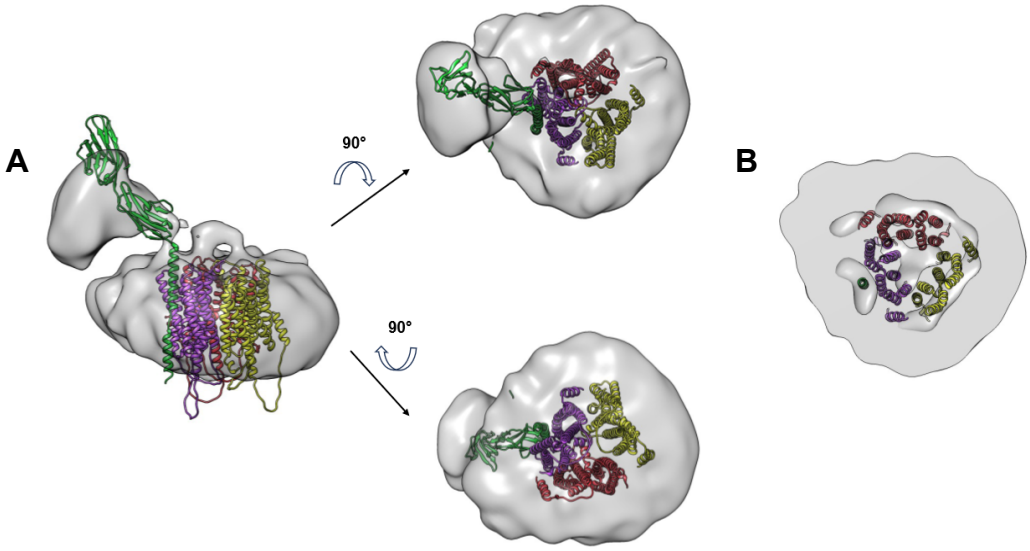


Figure 85. Fitting of cryo-EM volumes with the predicted structures of AlphaFold2. A. Fitting of the trimeric MLC1 and GlialCAM. **B.** Transversal slide of the 3D volume obtained by cryo-EM fitted to the AlphaFold2 predicted structure.

Our collaborators at CNIO continued processing the dataset. Doing an extensive particle cleaning of those particles that did not correctly align, they ended up with a group of 50,000 particles from the initial 4,500,000. Different reasons may be causing a bad alignment, such as bad ice quality of the grid, partially denaturation of the protein or alternative conformations. With this processing, the resolution of the volume obtained was of 7.5 Å. Still, it was not enough to see aminoacidic lateral chains, but at this resolution the transmembrane helices began to appear clear. Using the trimeric MLC1 model generated with AlphaFold, it was observed that MLC1 helices coincided with the tubular densities from our volume (Figure 86).

Figure 86. Fitted model after dataset refinement. In blue the predicted AlphaFold2 model is represented and in grey the experimental cryo-EM volume. Bottom (left), side (centre) and top (right) views of trimeric MLC1 are illustrated.

At that moment, the 3D resolution of the complex was hindered by two main problems: concentration and heterogeneity. In regards of the first one, the low protein concentrations obtained required taking huge datasets. Increasing the number of particles per grid would ease the image acquisition. On the other hand, heterogeneity was caused by the presence of different types of particles in the peak. Not only trimers of MLC1 were found, but also trimers with GlialCAM and elongated micelles compatible with MLC1 dimers.

Focusing on the concentration problem, we hypothesized that preparing cryo-EM grids in fresh would increase the efficiency of the purification. This means that the protein would be purified the same day as the preparation of grids, without freezing it. Our hypothesis was also based on the freeze/thaw experiment previously shown (Figure 79), where it was suggested that part of the purified protein was lost during the freezing protocol.

To test the hypothesis, a 9 L HEK293-6E culture was generated by the Protein Expression Core Facility at IRB and it was co-transfected with TwinStrep-MLC1 and GlialCAM-3xflag. After 48 hours post-transfection, cells were centrifuged and the pellet was frozen. The preparation of membranes, protein purification and grid preparation were done by Nayim González Rodríguez and I at CNIO on the same day. From this big-scale purification, we obtained 8 mg/mL of protein that was injected twice for gel filtration in a volume of 50 and 100 µL. Immediately, fractions were collected and grids prepared for the image acquisition at cryo-EM. The SEC profile from this purification is exposed in Figure 87A. On top view, 2D averages of MLC1 showed clearly the triangular shape of the micelle as it was previously observed and further confirming the formation of a homotrimer (Figure 87B). Only 10% of all particles presented MLC1 with an extra-micellar density, compatible with being GlialCAM (Figure 87D). This density appeared blurry, suggesting that this molecule had some mobility. Again, using the predicted AlphaFold model we could observe that it properly fitted our cryo-EM volume (Figure 87C). Unfortunately, the structural resolution did not improve with this fresh expression and we still had our model at 7 Å.

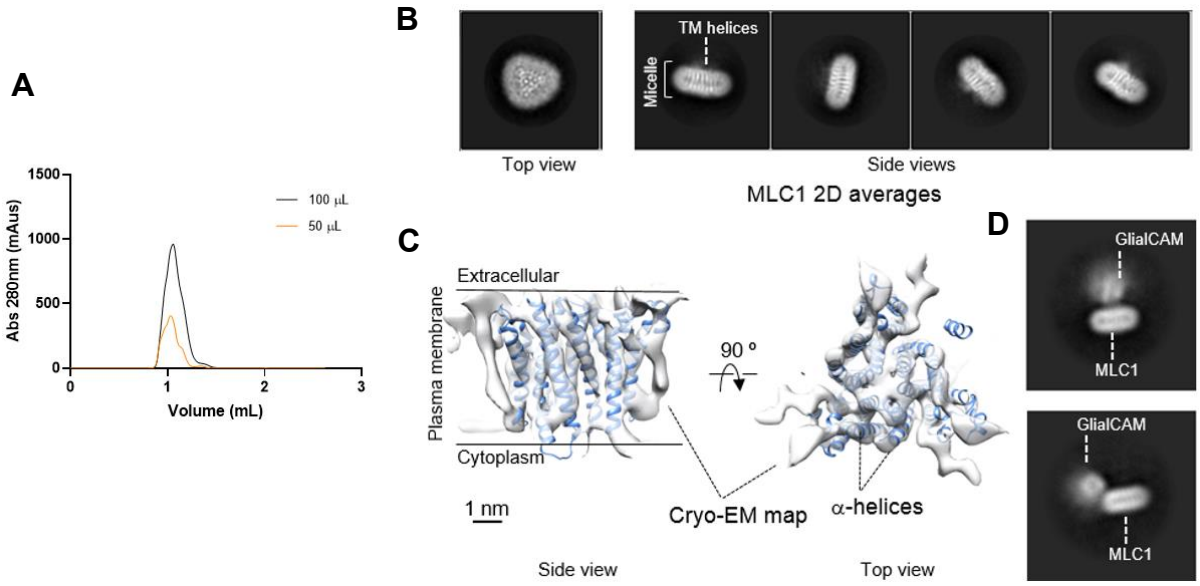


Figure 87. Fresh MLC1 purification. A. SEC profile of TwinStrep-MLC1 purification coexpressed with GlialCAM-3xflag. A 9 L cell expression was solubilized with 2% digitonin and purified using StrepTactin Superflow resin. The elution was concentrated, divided in two volumes (50 and 100 µL) and injected to ÄKTA for gel filtration. The column used was SuperDex 75 Increase 3.2/300 at 0.5 mL/min flow rate in Tris/NaCl buffer plus digitonin. **B.** 2D averages of MLC1 trimer in top and side views. **C.** Fitted model of our cryo-EM data with the AlphaFold predicted structure (in blue). **D.** 2D averages of those particles containing GlialCAM.

3.5 IMPROVING PROTEIN LEVELS

Since the obtention of good amounts of protein was difficult, changing the expression system was studied. Instead of transfecting mammal cell lines (HEK293-6E), we explored using baculovirus-mediated gene expression as it has become more popular in recent years especially to produce membrane proteins. We tested the baculovirus infecting insect cells system and the modified baculovirus (BacMam) to infect mammal cells. The reasoning behind testing both methods was that on one hand, expressing in insect cells would yield higher protein quantities, but the cellular environment would be different than the native one. On the other hand, using BacMam technology, the cellular environment would resemble more the native one, but the quantity of protein obtained would be less than in the insect cells. Still, the efficiency of BacMam was expected to be higher than transfecting cell lines.

3.5.1 BACULOVIRUS FOR MLC1

For the baculovirus expressing MLC1 and GlialCAM, our collaborators at CNIO generated the viruses. Nayim González Rodríguez infected a Sf9 insect cell culture with the baculoviruses and confirmed MLC1 and GlialCAM protein expression via WB (Figure 88).

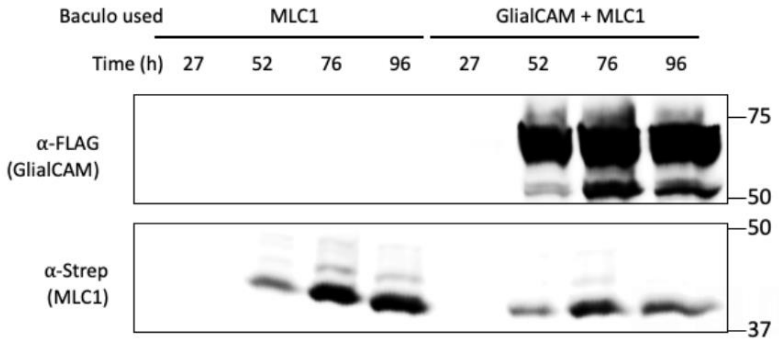


Figure 88. MLC1 and GlialCAM expression in insect cells. Cells were collected at 27, 52, 76, and 96 hours post-infection. Protein extracts were loaded into an acrylamide gel for WB analysis using anti-strep and anti-flag antibodies.

After this first test, a large-scale purification was proceeded where it was determined that MLC1 could be purified using baculovirus as the expression system (Figure 89). Then, another expression was set up to purify MLC1 following the same conditions as before. In this case, the purified protein was reconstituted in amphipols before preparing the cryo-EM grids. With this dataset, a resolution around 6-8 Å was obtained, so, the results were promising since with less material a similar resolution than the one we previously had was achieved. Similarly to the other structural determination, MLC1 was forming a trimeric structure. At that time, another dataset was scheduled and the results should be analysed in the near future.

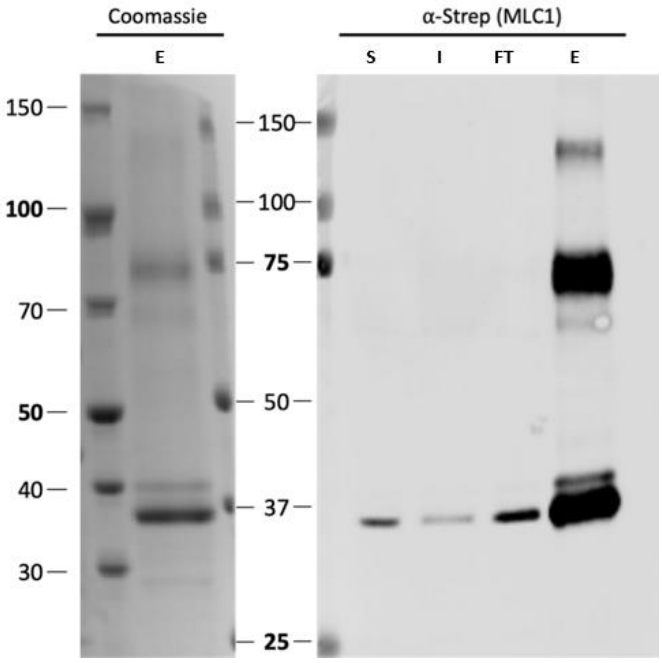


Figure 89. MLC1 purification using baculovirus system expression. Cells were co-infected with MLC1 and GlialCAM baculovirus. After, cells were collected and cell membranes were prepared. Then, the sample was solubilized and afterwards, the protein was purified. Coomassie staining (left) revealed two main bands at 37 kDa approximately and at 75 kDa corresponding to MLC1 and GlialCAM (or dimeric MLC1), respectively. WB analysis using anti-strep antibody established that MLC1 was solubilized and it could be purified.

S: solubilized, I: insolubilized, FT: flowthrough, E: elution.

3.5.2 BACMAM VIRUS FOR MLC1

For the BacMam system set up, we collaborated with the group of Dr. Beatriz Herguedas at BIFI (*Instituto de Biocomputación y Física de Sistemas Complejos*) in Zaragoza. BacMam technology is previously explained in the Materials and Methods of this thesis (*see Section 1.4*).

First, MLC1 and GlialCAM were cloned into two donor vectors that contained GFP and mCherry genes, respectively. For its obtention, traditional cloning using restrictive endonucleases was employed. Once we had the plasmids, they were transfected into HEK293T cells to confirm that the proteins were being expressed before generating the bacmid. 48 hours post-transfection, cells were lysed and protein extracts were obtained for a WB analysis. Results are presented in Figure 90, where it can be observed that both proteins are expressed at higher levels than the expression plasmids used routinely in our laboratory for cell transfection.

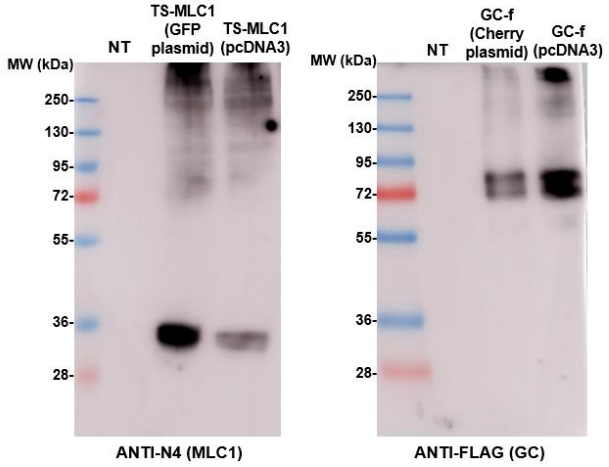


Figure 90. MLC1 and GlialCAM expression in BacMam donor vectors. WB detecting MLC1 (left) and GlialCAM (right) from transfected HEK293T cells.

NT: not transfected, TS: TwinStrep, GC-f: GlialCAM-3xflag.

Once the donor vectors were ready, we transformed the plasmids for the obtention of the bacmid using DH10Bac *E. coli* strain. The grown colonies were picked and the bacmid was isolated. With the generated bacmids we went to the lab of Dr. Beatriz Herguedas for the transfection of Sf9 insect cells in order to produce the recombinant BacMam virus. This work was performed by the PhD student Irene Sánchez Valls. Briefly, after the transfection of Sf9 cells with both MLC1 and GlialCAM plasmids, the fluorescence of the culture was checked as an indicator of transfection rate. Once the transfection was high (after 7 days), cells were centrifuged and the virus was stored. The viral particles were used then for infecting the Sf9 cell culture, after 7 days, the virus was collected again. This process was done twice to obtain the P2 virus. With the P2 MLC1 and GlialCAM virus, we infected adherent and suspension HEK cells in our laboratory. One of the key points for a successful protein expression is to determine the optimal time of expression. So to assess this, we infected cells with the new BacMam virus and we incubated cells for 32, 48 and 72 hours. WB analysis of protein extracts established that the highest protein expression was at 72 hours post-infection (Figure 91).

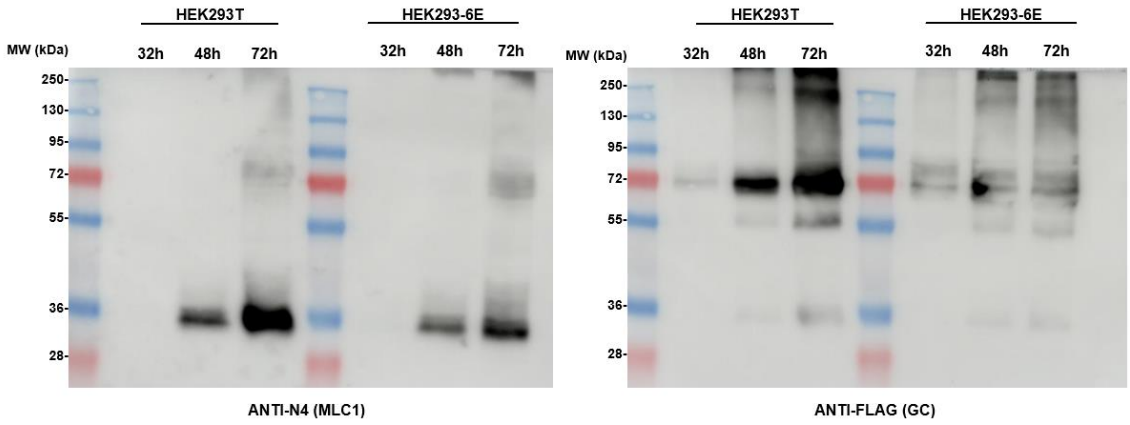


Figure 91. Time course of MLC1 and GlialCAM protein expression in BacMam infected HEK cells. WB detecting MLC1 (left) and GlialCAM (right) from transfected HEK293T and HEK293-6E cells. Cells were collected at 32, 48 and 72 hours.

Having established the time of infection, we decided to carry out a small protein purification with the BacMam system expressed protein. For this experiment, we infected HEK293-6E cells with both MLC1 and GlialCAM encoding virus. After 72 hours, cells were centrifuged and the pellet was split in two parts to do two purifications in parallel using either DM or digitonin. In Figure 92, the WB of the purification process is shown. In both purifications (with DM and digitonin), MLC1 was detected in the total protein, solubilised and eluted fractions concluding that the protein can be expressed in the BacMam system and purified. Nevertheless, in the WB detecting GlialCAM, the protein weighted around 55 kDa, instead of the expected 72 kDa, which suggested that the protein is found in an immature form. Due to lack of time and material, further experiments could not be carried out.

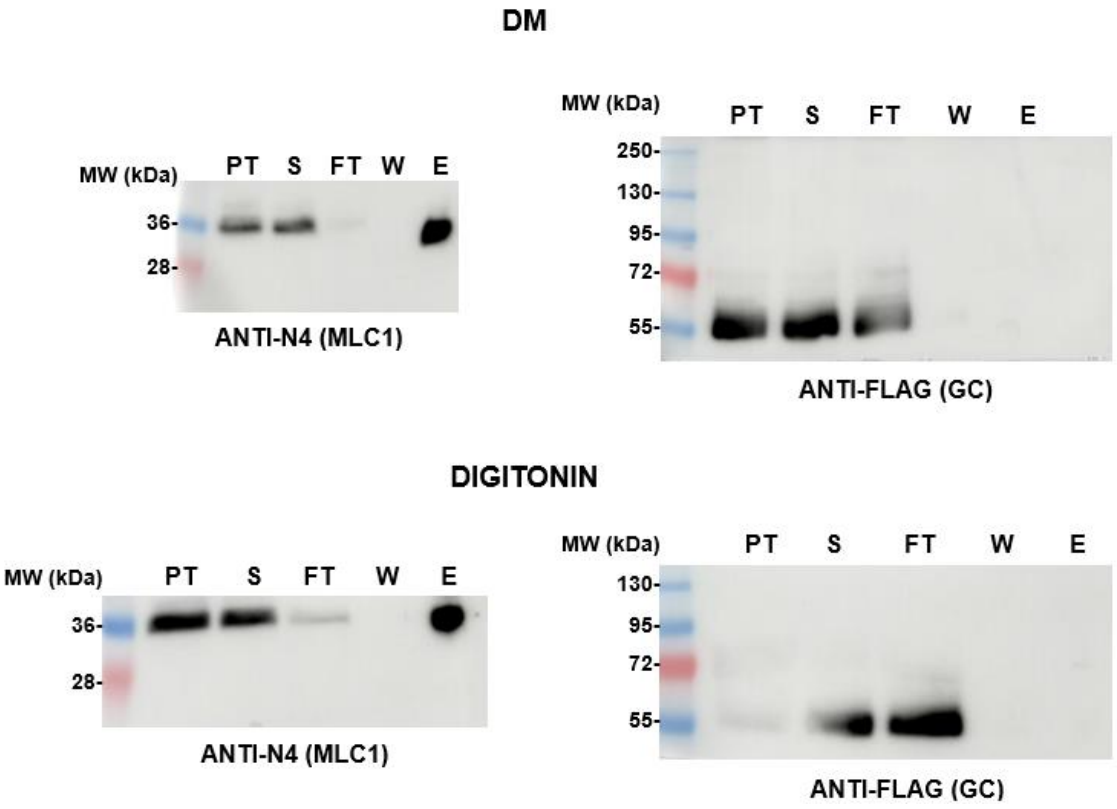


Figure 92. MLC1 protein purification using BacMam system expression. WB detecting MLC1 (left) and GlialCAM (right) from infected HEK293-6E cells. Protein purification was done using DM (top) and digitonin (bottom).

3.6 SOLVING ORIENTATION PROBLEM

After the first cryo-EM analysis, it was concluded that solving MLC1 protein structure would be challenging as this protein does not have any extra-micellar feature that could help orientating the particles. Although MLC1 purification was done in presence of GlialCAM, which has an extracellular domain, the number of particles containing this protein was small. Moreover, GlialCAM showed some flexibility.

Since we had seen that incrementing the amount of protein did not favour the structural resolution, not even working with freshly purified protein, we decided to focus on finding a method to orientate particles. Particle orientation is a common problem when working with membrane proteins because they are embedded in detergent micelles. To solve this problem, antibodies (Fabs) can be used as fiducial marks to facilitate protein orientation (Wu et al., 2012).

In this project, two parallel strategies were followed aiming to solve the orientation challenge. On one hand, we worked with a semi-synthetic antibody (sAB) that recognized the fusion protein BRIL which was fused to our protein of interest. On the other hand, we developed nanobodies against MLC1. Both strategies are based on the same principle of binding a particle (sAB or nanobody) rigidly to our protein of interest for orientation.

3.6.1 BRIL

BRIL is an engineered variant of apocytochrome b562 that weights 12 kDa. It consists of a rigid helical domain that makes it suitable for fusions in the loops connecting α-helices of membrane proteins, but also in the N- and C- termini. The development of sABs against the fusion protein BRIL facilitates the accurate orientation of the proteins. It was demonstrated that BAG2 sAB can bind to BRIL fused into loops or termini of different GPCRs, ion channels, receptors, and transporters without disrupting their structure (Mukherjee et al., 2020).

3.6.1.1 CLONING AND PURIFICATION

In regards of the cloning strategy, we decided to eliminate the third intracellular loop of MLC1, which is the biggest one, and substitute it for the BRIL sequence. Although BRIL has also been successfully added into the N- or C- termini of other proteins, we made the decision to insert it into a loop as we thought it might be more constrained and less flexible this way.

Firstly, the new construct was cloned into an expression vector. We used the TwinStrep-MLC1 plasmid as a template and we cloned the BRIL sequence by removing the third intracellular loop. Afterwards, it had to be assessed whether MLC1 fused to BRIL was expressed and could be purified. WB analysis confirmed that the new protein was being expressed (Figure 93A). In immunocytochemistry assays, it was observed that in transfected cells, MLC1-BRIL was located similarly to MLC1 WT. The co-transfection with GlialCAM showed that MLC1-BRIL was enriched at cell-cell junctions as MLC1 WT (Figure 93B). Using StrepTactin resin and working with the same conditions as the purification of TwinStrep-MLC1, we confirmed that TwinStrep-MLC1-BRIL could also be purified (Figure 93C).

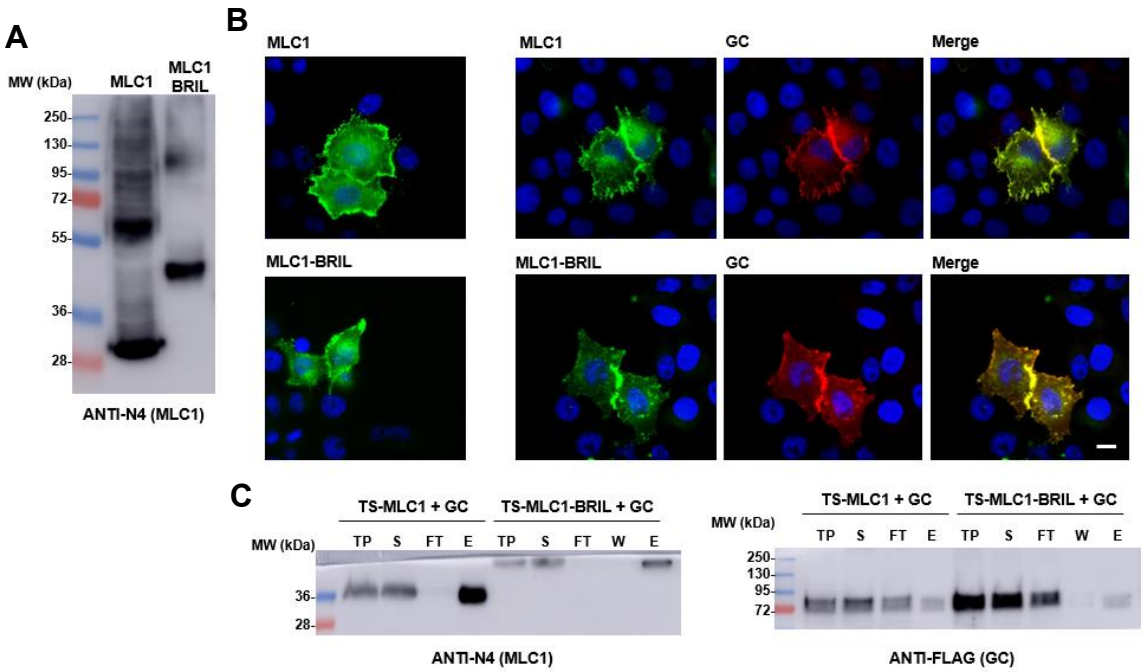


Figure 93. TwinStrep-MLC1-BRIL is expressed and can be purified. A. WB detecting MLC1 of TwinStrep-MLC1-BRIL from transfected HEK293T cells. **B.** Immunocytochemistry of TwinStrep-MLC1-BRIL in HeLa cells showing the same phenotype than MLC1 WT. Scale bar: 20 µm. **C.** WB of the purification of TwinStrep-MLC1-BRIL compared to TwinStrep-MLC1. MLC1 and GlialCAM were both detected using anti-N4 and anti-flag antibodies, respectively.

TS: TwinStrep, GC: GlialCAM, TP: total protein, S: solubilized, FT: flowthrough, W: wash, E: elution

3.6.1.2 BAG2 RECOGNIZES MLC1-BRIL

After the expression assessment, it was time to analyse whether the sAB BAG2 would recognize our protein. For this, the following experiment was designed. HEK293T cells were transfected with TwinStrep-MLC1-BRIL plus GlialCAM-3xflag. After the purification of TwinStrep-MLC1-BRIL, the protein was incubated with BAG2 for 1 hour. Later, the sample was incubated with CaptoL resin. CaptoL has a strong affinity for the variable region of antibody kappa light chain, which makes this resin suitable for the capture of a wide range of antibody fragments. Once it was time to elute, a sample of the flowthrough and elution fractions were collected and loaded into an acrylamide gel for detecting the sAB and MLC1 via WB. If BAG2 binded MLC1-BRIL, the protein should be detected on the elution of the CaptoL resin. On the contrary, if the sAB could not bind to it, MLC1 should be detected on the flowthrough. On Figure 94A a scheme of the experimental approach is shown. The obtained results are presented in Figure 94B-C where it is seen that MLC1 was found in the elution fraction, meaning that BAG2 was binding to MLC1-BRIL.

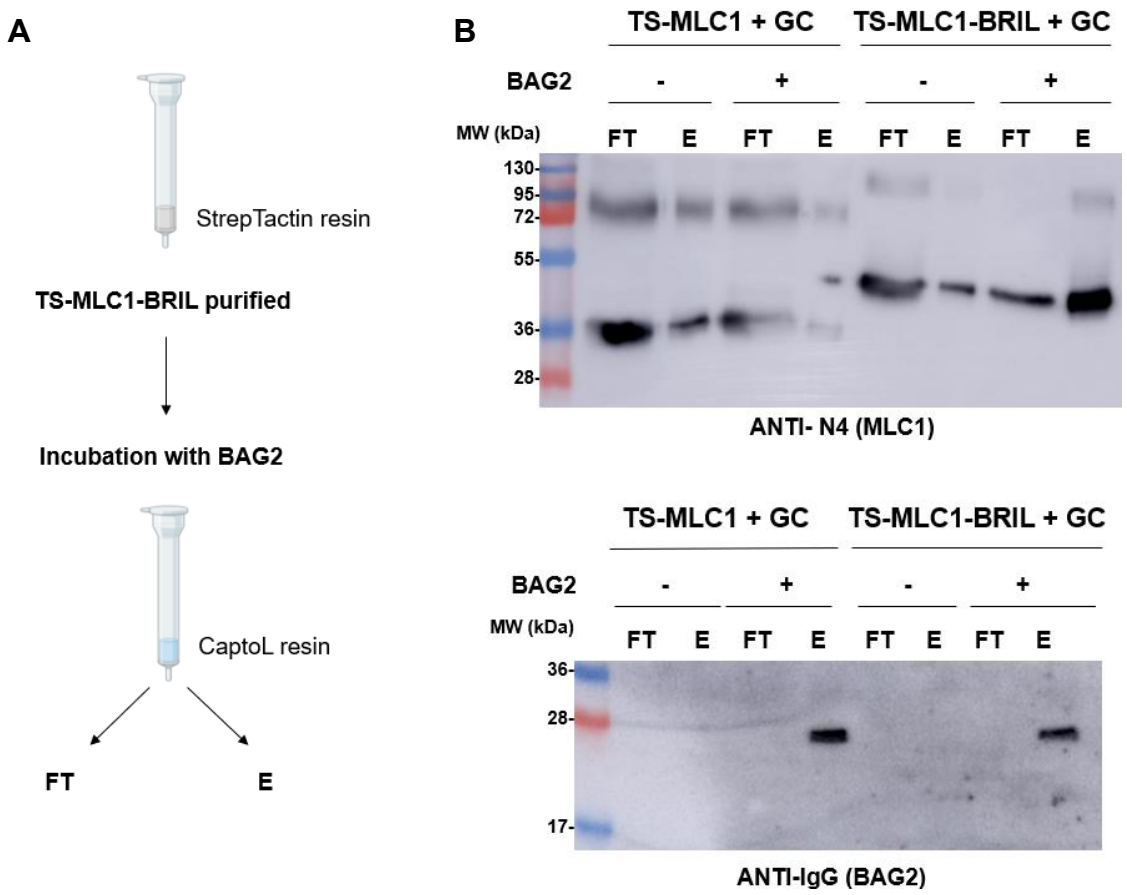


Figure 94. TwinStrep-MLC1-BRIL purification incubated with BAG2. A. Schematic representation of the purification process. Once MLC1 was purified with our standardized protocol, the sample was incubated with BAG2 for 1 hour. Then, it was incubated with CaptoL resin. A fraction of the flowthrough and elution were kept for posterior WB analysis. **B.** WB of the different fractions incubated with anti-N4 (top) and anti-IgG (bottom) to detect MLC1 and BAG2, respectively.

TS: TwinStrep, GC: GlialCAM, FT: flowthrough, E: elution.

To further confirm this result, another experiment with BAG2 was set up. Again, HEK293T cells were co-transfected with TwinStrep-MLC1-BRIL and GlialCAM-3xflag. MLC1-BRIL was purified, incubated with BAG2, and this time, a gel filtration was done. A sample containing only BAG2 was also analysed for SEC. Knowing that in both samples BAG2 was present at the same concentration, a decrease of the peak corresponding to the sAB in the protein sample would mean that there was part of it that was binding to the protein. Affirmatively, a reduced peak was observed and it was confirmed via WB that BAG2 was present in the same peak as MLC1-BRIL (Figure 95). In conclusion, it was determined that BAG2 did recognize the protein TwinStrep-MLC1-BRIL. Afterwards, we collected the last fraction from the first peak seen in the WB and sent it to our collaborators at CNIO. The choice of this fraction was based on the fact that it contained both MLC1 and BAG2 and there was the highest concentration of MLC1.

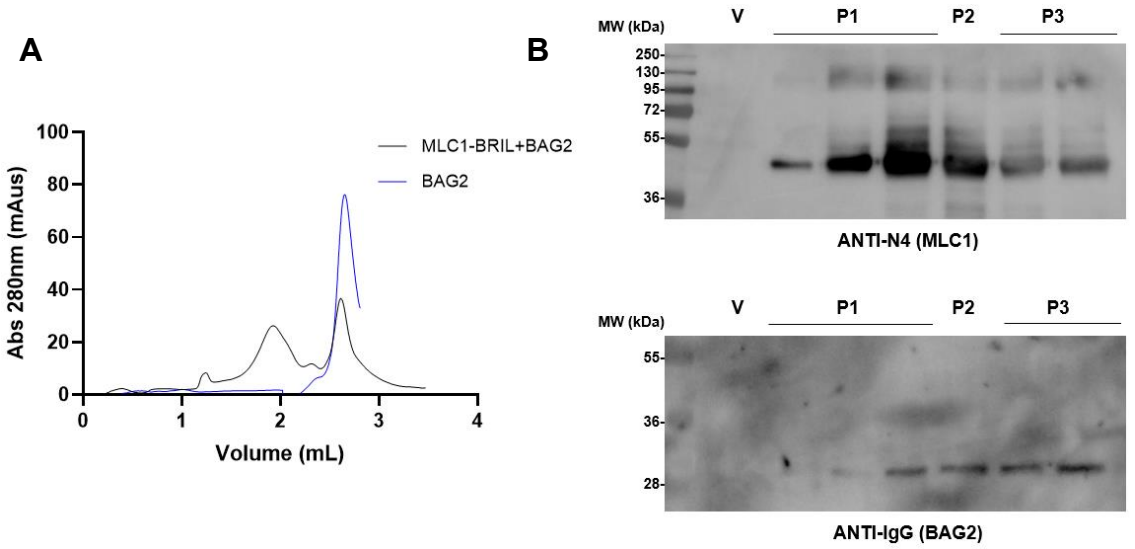


Figure 95. BAG2 recognizes MLC1-BRIL. A. SEC profile of BAG2 alone or with purified TwinStrep-MLC1-BRIL. A 1L cell culture expression was solubilized with 2% digitonin and purified using StrepTactin Superflow resin. The elution was concentrated to 1 mg/mL in a volume of 55 µL, incubated 1 hour with BAG2 at a ratio of 1:1.2 protein: antibody, and injected to ÄKTA for gel filtration. The column used was Superose 6 Increase 5/150 at 0.3 mL/min flow rate in Tris/NaCl buffer plus digitonin. **B.** WB of the collected fractions from the SEC detecting MLC1 (top) and BAG2 (bottom). Both molecules are present in the same fractions from the second and third peak.

V: void, P1: peak 1, P2: peak 2: P3: peak 3.

3.6.1.3 NEGATIVE STAINING

Before putting the protein in grids, Nayim González Rodríguez performed a Coomassie staining to assess the purity of the purified protein. As seen in Figure 96A, the gel appeared clear with two prominent bands. One band between 20-30 kDa corresponded to BAG2 and another one around 45 kDa consistent with MLC1-BRIL. At 70 kDa approximately, a light band can be seen which might correspond to GlialCAM. Since the sample had good purity, the negative staining was executed.

The grids presented a good distribution of particles and many contained spikes coming out of a round shape, consistent with BAG2 binding MLC1-BRIL. In the 2D classification, different stoichiometries were observed from one to three subunits of BAG2 (Figure 96B). These results were consistent with previous datasets showing different sizes of particles consistent with various oligomeric states of MLC1.

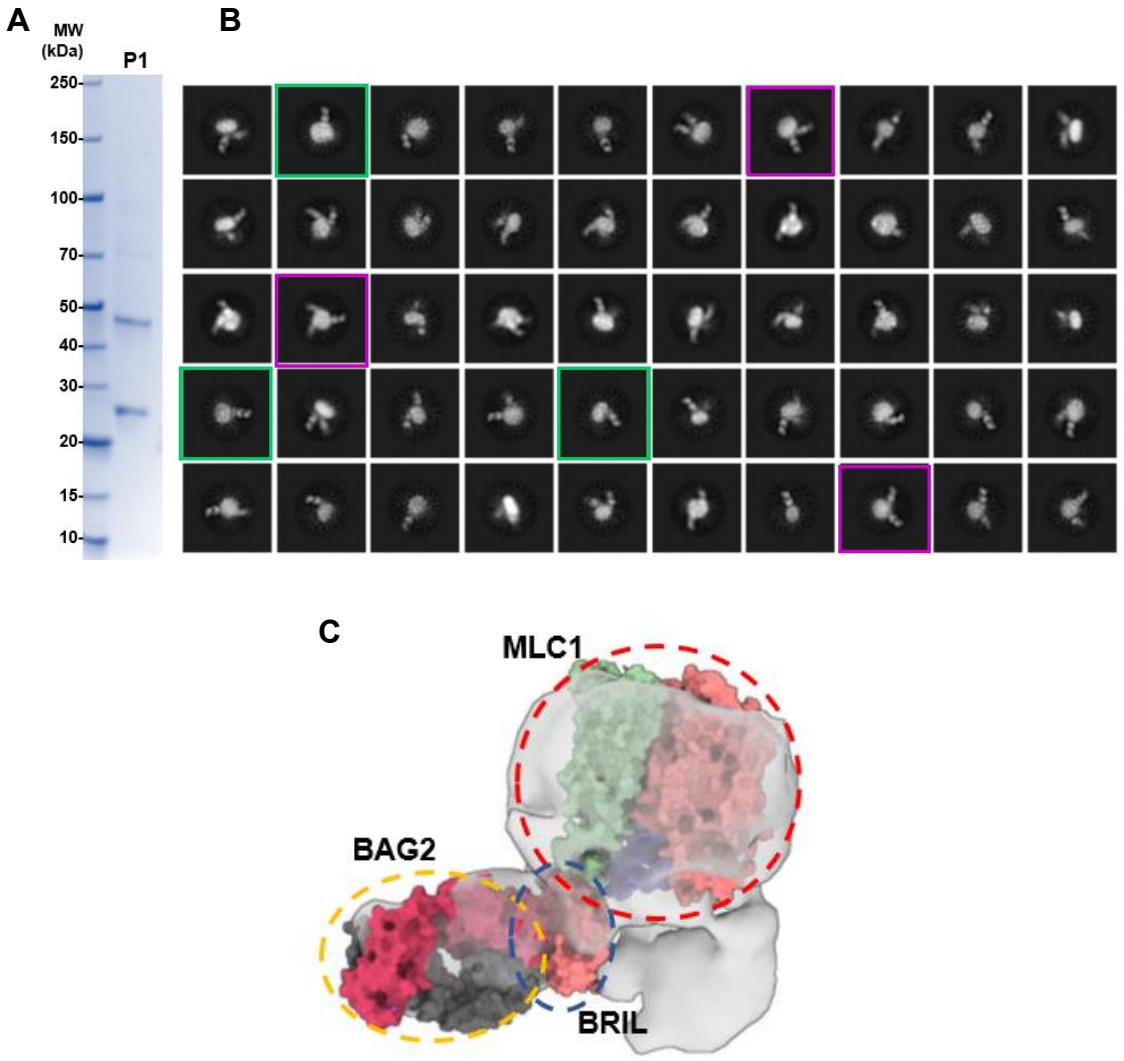


Figure 96. Negative staining for MLC1-BRIL with BAG2. A. Coomassie staining of a SEC fraction containing both MLC1-BRIL and BAG2 as seen previously in a WB. **B.** 2D classification of MLC1-BRIL with BAG2. In purple, some particles presenting two or three BAG2. In green, some particles with only one BAG2. **C.** Preliminary 3D reconstruction of MLC1-BRIL and BAG2.

3.6.1.4 CRYO-EM ANALYSIS

Since the images from the negative staining looked promising and keeping in mind the need of high concentration material for cryo-EM imaging, it was decided to express larger volumes of MLC1-BRIL. These large purifications were done by Nayim González Rodríguez and I at CNIO.

Two big-scale purifications of 3 and 5 L, respectively, were solubilized with 2% digitonin and purified using StrepTactin resin. Then, the samples were concentrated, incubated

with BAG2 and injected for gel filtration. From this, the most concentrated samples were collected and a thorough grid screening followed.

Analysing the cryo-EM dataset, BAG2 was clearly observed, but there was heterogeneity regarding the number of sABs binding the protein. Thus, instead of solving the alignment problem, BRIL strategy had worsened it. It seemed that the BRIL tag was flexible which was complicating the alignment. Seeing that the BRIL fused strategy did not work, we decided to focus on finding a nanobody that would specifically recognize MLC1.

3.6.2 NANOBODIES OBTENTION

Nanobodies are small (15 kDa) and stable single-domain fragments of the heavy chain-only antibodies that naturally occur in the immune system of camelids. They have access to cavities or clefts on the surface of proteins that are often inaccessible to conventional antibodies (Uchański et al., 2020). Nanobodies present a series of characteristic that make them suitable for structural biology field. They are small, easily produced and exhibit superior stability compared to Fabs. They are protease resistant and thermos-resistant, although they show a propensity to aggregate under thermal stress (Muyldermans, 2021). Besides being a powerful tool in protein research, they are also studied for diagnostic and therapeutical applications (Jin et al., 2023).

One of the main limitations of cryo-EM imaging is that it is difficult for small proteins (< 100 kDa) because its membrane-embedded nature impedes the alignment for high-resolution reconstructions. However, the development of nanobody technology has partially overcome this problem since they help orientate the target protein. Furthermore, they serve as fiducial markers for particle localization and alignment (Wentinck et al., 2022). To this aim, many protein structures have been solved using these particles. For instance, the structure of different inactive-state family A GPCRs has been solved at high resolution by using nanobodies binding an intracellular loop (ICL3) of the GPCR (Robertson et al., 2022).

In another study, a nanobody was developed to stabilize the agonist-bound β2 adrenergic receptor to Gs by inhibiting their dissociation. With this, the crystal structure of the complex was obtained (Rasmussen et al., 2011). Later, the same nanobody was used to solve the cryo-EM structures of the adenosine A_{2A} (García-Nafría et al., 2018), GLP1 (Liang et al., 2018, Zhang et al., 2017b) and PTH1 (Zhao et al., 2019) receptors coupled to either Gs alone or together with β-arrestin. Apart from solving protein

structures, the development of nanobodies can also serve the purpose of inhibiting the protein function by binding to specific domains (Irobalieva et al., 2023).

3.6.2.1 SAMPLE PREPARATION FOR IMMUNIZATION

For the llama immunization, purified protein is needed. It is crucial that the protein maintains its native conformation throughout the immunization procedure. For this reason, reconstitution of proteins is typically performed under a lipid environment: phospholipid vesicles or a very tight binding detergent. Three different procedures were done trying to obtain enough purified protein. The first one was to use the detergent DDM for the purification. The other options contemplated purifying with digitonin, but afterwards transferring the protein into amphipols or liposomes, respectively.

DDM PROTEIN PURIFICATION

The choice of DDM was based on the fact that this detergent has a very low CMC (0.0087%), meaning that the micelles are able to form at very low concentrations. This reduces the detergent- protein complex disintegration once the sample is injected into the animal. Moreover, DDM is a non-ionic detergent and it creates large micelles which might help to maintain the stable native state of many proteins (Stetsenko, 2017).

For this purification, 500 mL of HEK293-6E cell culture transfected with TwinStrep-MLC1 was solubilized 1 hour with 1% DDM plus 0.2% CHS. Then, the sample was incubated with StrepTactin resin and eluted using desthiobiotin. The purified protein was concentrated and injected for SEC.

Analysing the SEC profile, it was observed that the protein was aggregated because it was present in the column void (Figure 97). We discarded this option for the immunization. These purification conditions were only tried once, it might be possible to purify MLC1 using DDM with an optimization of the protocol. Furthermore, Dr. Efren Xicoy-Espaulella could purify MLC1 with DDM when he was screening the detergent to work with. Although in his case, the construct analysed was not TwinStrep-MLC1, but MLC1-3C-GFP-10His, which might change the optimal purification conditions.

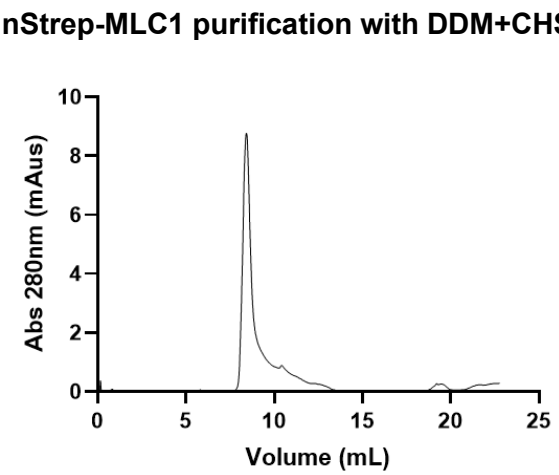


Figure 97. SEC profile of TwinStrep-MLC1 purification. Sample was concentrated at 0.2 mg/mL in a volume of 250 µL before gel filtration and run in SuperDex Increase 200 10/300 GL at a flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

AMPHIPOLS

Amphipols are amphipathic polymers that are able to keep individual membrane proteins soluble in their native state under the form of small complexes (Popot et al., 2011). They have been used to complex and keep membrane proteins soluble. In some cases, membrane proteins in amphipols are more stable than in detergent solution, which has led to large variety of structural and functional studies (Perlmutter et al., 2014). A8-35 is the mostly used amphipol, it is highly soluble in water and it self-associates to form small, compact, hydrated particles with an average weight of 40 kDa (Gohon et al., 2006).

We decided to reconstitute MLC1 in amphipols as an option for llama immunization. To do so, 1 L of HEK293-6E transfected with TwinStrep-MLC1 was purified. Then, amphipols were added to the purified protein in a ratio of 1:5 protein: amphipol and incubated on ice. Polystyrene beads were added to capture the remaining detergent and the sample was dialyzed O/N. After this time, the protein was injected for gel filtration. The SEC profile presented a symmetric and well-defined peak, which was confirmed to be MLC1 via WB (Figure 98A-B). Silver staining was also performed to assess the purity of the purified protein (Figure 98B). A band of 36 kDa approximately was detected in the gel corresponding to MLC1. A very light band was also observed around 72 kDa which could be MLC1 forming a dimer. Apart from these bands, the sample did not contain any other proteins. With this experiment, it was concluded that MLC1 could be reconstituted in amphipols.

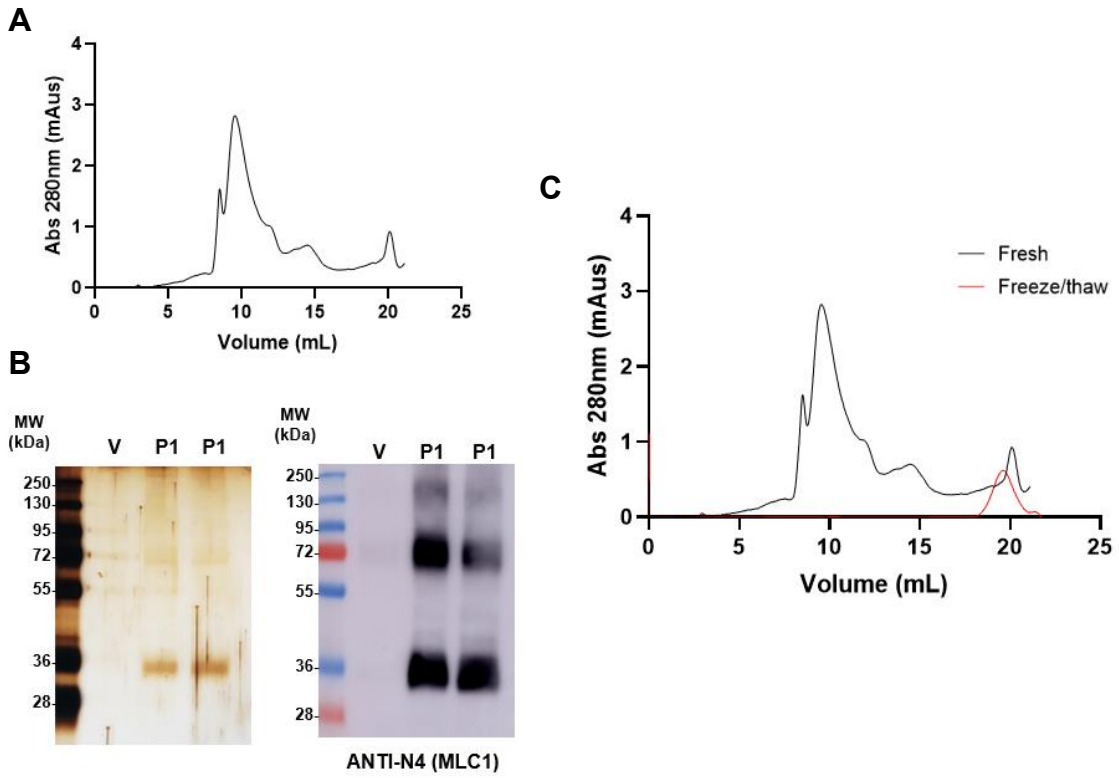


Figure 98. MLC1 reconstitution in amphipols. A. SEC profile of MLC1 reconstituted in A8-35 amphipols. After protein purification, amphipols were added in a ratio of 1:5 protein: amphipols and incubated 30 minutes on ice. Then, bio-beads SM-2 adsorbent were added considering 20 g of wet beads for 1 g of protein and incubated 1 hour. O/N dialysis against 100 mM Tris- 150 mM NaCl buffer was performed before injecting the sample for gel filtration using column SuperDex Increase 200 10/300 at a flow of 0.3 mL/min in Tris/NaCl buffer. **B.** SST (left) and WB detecting MLC1 (right) of the collected fractions from the SEC profile. **C.** SEC profile comparing the fresh sample with the protein that had been frozen and thawed. Samples were run in SuperDex Increase 200 10/300 at a flow of 0.3 mL/min in Tris/NaCl buffer.

P1: peak 1.

Nevertheless, considering that the purified protein had to be shipped, it was necessary to determine whether cycles of freezing and thawing affected the viability of the protein. With this aim, the previous sample of purified MLC1 reconstituted in amphipols was frozen in liquid nitrogen and stored at -80°C. After a couple days, the sample was thawed, filtered, and injected for gel filtration (Figure 98C). In this case, the protein peak previously seen was not visible, implying that the protein had denatured. Accordingly, we had to discard amphipols for the llama immunization. It is worth noting that most proteins need glycerol when freezing/thawing and we did not add it. It might be possible that, by adding glycerol before freezing, MLC1 reconstituted in amphipols can be frozen and thawed without degrading. Due to limited protein material more tests were not carried out.

PROTEOLIPOSOMES

Lastly, the reconstitution of MLC1 in liposomes was attempted as the only option available once the other two had failed. Liposomes are small sphere-shaped vesicles composed of a phospholipid bilayer, which may form spontaneously in aqueous media (Singh, 2019, Mangala Rao and Carl, 2000). Its similarity to cell membranes makes them useful for the study of membrane proteins inserted into them. Proteoliposomes are liposomes to which a protein has been incorporated or inserted (Ciancaglini et al., 2012).

The incorporation of proteins into liposomes can provoke inward and outward-facing protein conformation in a variable percentage. For this reason, there is a need to verify that the reconstitution procedure has yielded the correct orientation of the protein with the analysis of its functional activity (Ciancaglini et al., 2012). In our case, we did not have a functional assay for MLC1 since we did not know its physiological function. Thus, to assess the proper proteoliposome reconstitution, a sucrose gradient was done where the liposomes were separated depending on its size (Sánchez-López et al., 2009). This method enables to distinguish between aggregated and reconstituted protein, since the aggregated will appear in the pellet. The density sucrose gradient was made by piling progressively sucrose layers of different concentrations into an ultracentrifuge tube, from the densest to the less dense. Previously purified protein was used for this experiment. After the ultracentrifugation, the fractions were collected and analysed via WB where MLC1 was detected in different fractions, but not in the pellet, meaning that the protein was not aggregated (Figure 99).

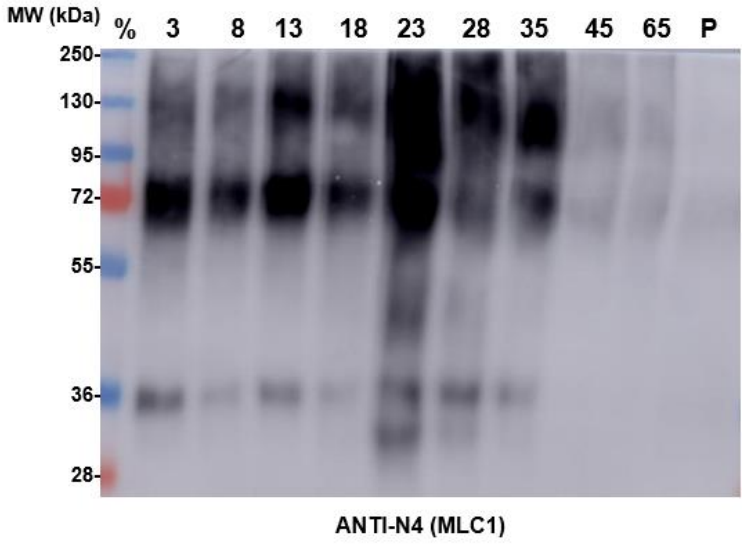


Figure 99. MLC1 reconstitution in liposomes. WB of sucrose gradient fractions detecting MLC1. P: pellet.

3.6.2.2 PHAGE DISPLAY PANNING

Once we had successfully sent our protein reconstituted in liposomes, our collaborators at VIB-VUB Structural Biology Research Centre in Brussels immunized a llama with it. The immunogen administration lasted for six weeks, at the interval of one injection per week. After the last immunization, blood was collected from which PBLs were isolated. RNA was extracted and synthesized to cDNA. From the cDNA, the variable domains of immunoglobulin heavy chains were amplified via PCR and cloned into a phage display vector.

For the phage display panning, I assisted to the Workshop Nanobody4Instruct Training Course funded by Instruct-ERIC at VIB-VUB Centre for Structural Biology in Brussels. Phage display methodology is previously explained in the Methodology of this thesis (*see Section 3.12*). Briefly, the phage library with all cloned nanobodies was put into contact with the antigen. Different antigen presentation methods were tried for the phage selection, including coating with the antigen, coating with the tag and capturing the antigen with magnetic beads.

COATING WITH MLC1

MLC1 reconstituted in liposomes was coated on ELISA plates for selection by phage display. Later, the phage library was added and after 2 hours of incubation, the wells were trypsinized and TG1 cells were infected with the trypsinized product. Then, the output was titrated to estimate the enrichment of the first phage display round. In Figure 100A, the results from the phage titration are presented. It can be observed that the samples coated with MLC1 presented more colonies than the control (without coating). Counting colonies from each group, the number of total output phage was calculated correcting for the dilution factor and for the output volume. In this case, the number of total output phage was 350 times greater in the coating conditions than the control, this is called enrichment and it is a value for the phage specificity.

COATING WITH ANTI-TWINSTREP

In this case, plates were coated with anti-TwinStrep and later incubated with TwinStrep-MLC1 reconstituted in liposomes. Similarly to the MLC1 coating, TG1 cells were infected with the output phage and a titration was done. In Figure 100B the titration is shown. An enrichment of 30 was obtained.

CAPTURE OF MLC1 WITH BEADS

StrepTactin magnetic beads were put in contact with MLC1 proteoliposomes, then, input phages were added and after that they infected TG1 cells. Output phage titration is depicted in Figure 100C. In this case, two rounds of phage display were performed increasing the enrichment of the output phage. The enrichment was 3 on the first round and 83 on the second one.

After testing three different antigen capture for phage display, it was observed that the enrichments were variable depending on the methodology used. Since the highest enrichment was obtained coating with MLC1, its output was employed for the following steps.

A

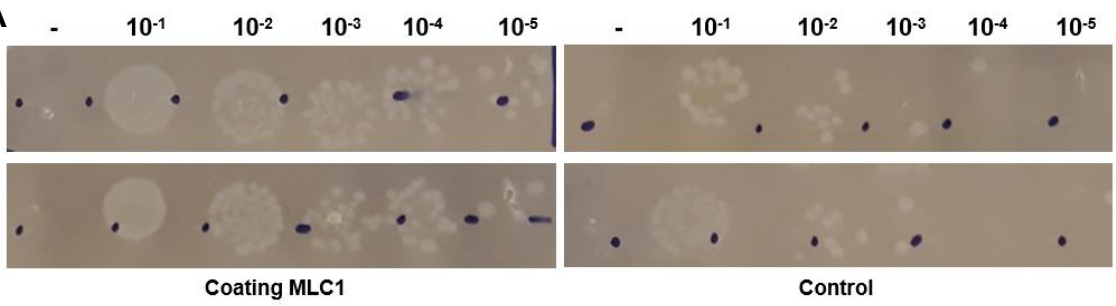


Plate well	0	-1	-2	-3	-4	-5	# eluted Φ	enrichment
Antigen						7	1.75E+08	350
Control				2			5E+05	

B

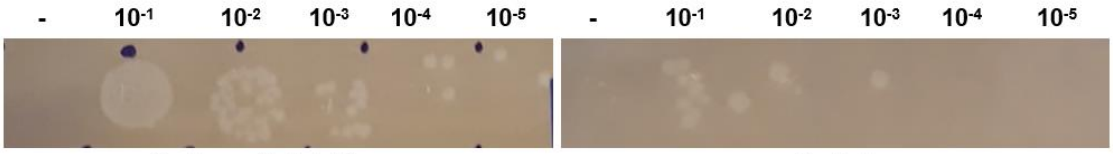


Plate well	0	-1	-2	-3	-4	-5	# eluted Φ	enrichment
Antigen					3		7.50E+06	30
Control				1			2.50E+05	

C

1st ROUND

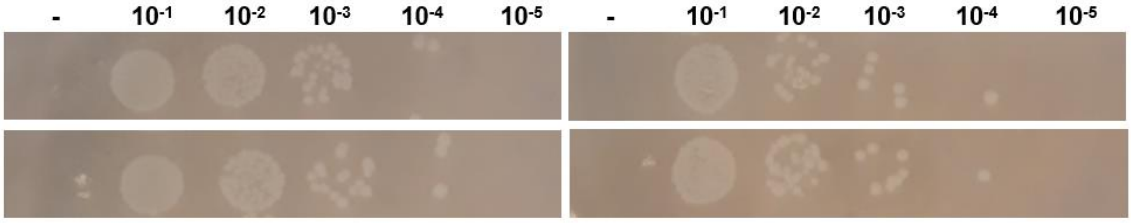


Plate well	0	-1	-2	-3	-4	-5	# eluted Φ	enrichment
Antigen					3		7.50E+06	3
Control					1		2.50E+06	

2nd ROUND

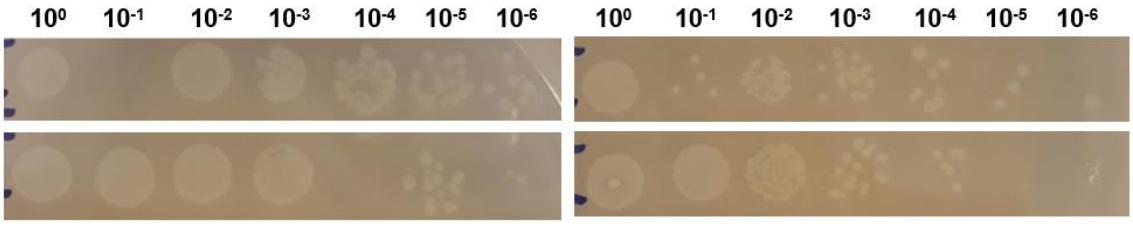


Plate well	0	-1	-2	-3	-4	-5	-6	# eluted Φ	enrichment
Antigen							5	1.25E+09	83
Control					6			1.50E+07	

Figure 100. Antigen presentation in phage display panning for MLC1. A. Titration of phage display selection coating with TwinStrep-MLC1 reconstituted in liposomes. **B.** Titration of phage display selection coating with anti-TwinStrep antibody. **C.** Titration of phage display selection with StrepTactin magnetic beads, two phage display rounds were done. Φ: phages.

3.6.2.3 ELISA SCREENING

After panning, 96 individual clones from the selection output were picked and expressed in *E. coli.* An ELISA assay was performed to screen for target specificity. The nanobody binding was detected by incubating with an anti-mouse-AP conjugate. Results from the spectrophotometer reader after an O/N incubation are presented in Table 13. In dark green, values higher than 2.5-fold compared to the control group (no antigen coating) are marked. These nanobodies were considered positive for binding MLC1.

3.6.2.4 SEQUENCE DETERMINATION

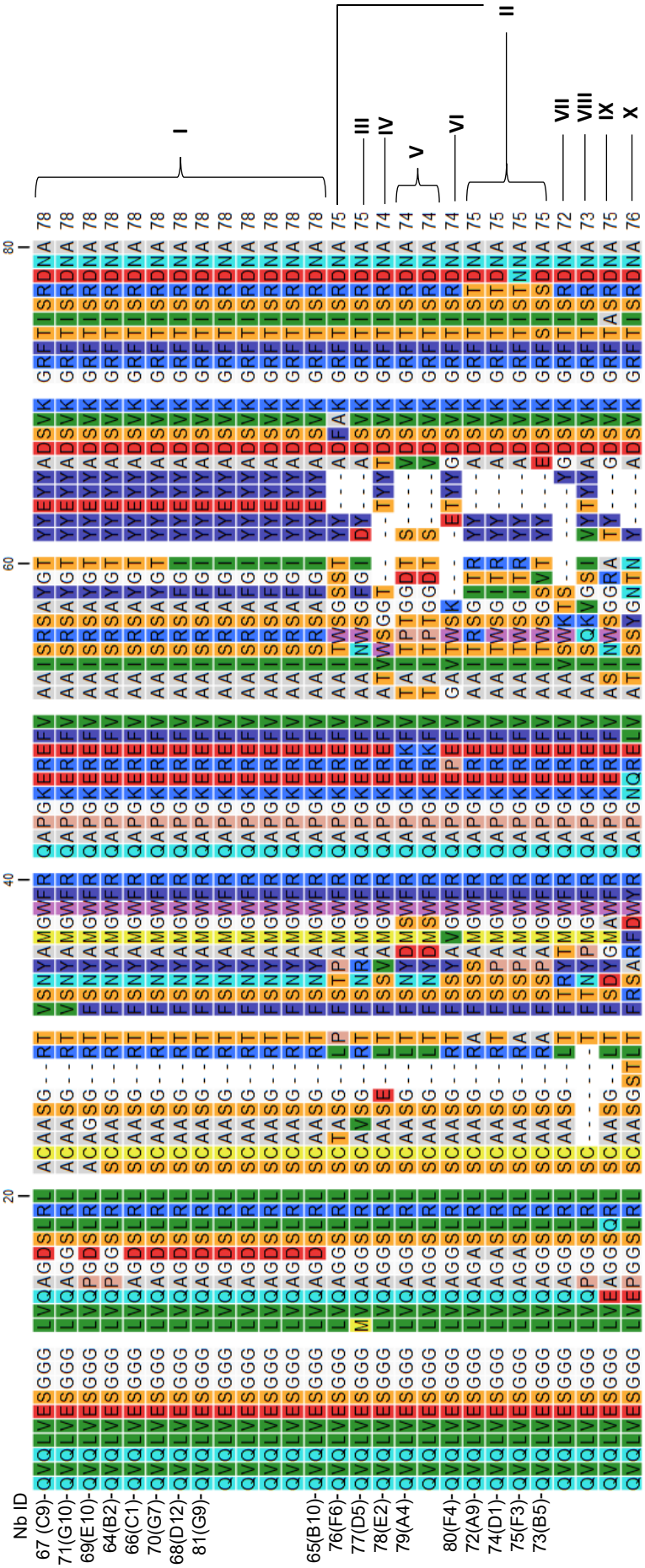
Those positive nanobodies were sequenced via PCR using the primer MP57. PCR products were analysed via gel electrophoresis and a fragment of approximately 700 bp was observed corresponding to the nanobody. Afterwards, a PCR cleanup protocol was done and samples were sent for sequencing. Once we had obtained the sequencing results, a DNA alignment was performed with the objective of classifying nanobodies into sequence families. Nanobodies are considered of belonging to the same sequence family when they present a high similarity in their CDR3 sequence, considering an identic length and > 80% of sequence identity. Nanobodies from the same family derive from the same B-cell lineage and bind to the same epitope on the target (Pardon et al., 2014). Our sequence alignment is presented on Figure 101. From the 27 positive nanobodies obtained in the ELISA screening, and after classifying them into families, we ended up with 21 different nanobodies from 10 different families. The chosen 21 nanobodies were labelled with a two-digit code and were sent to our laboratory for the posterior expression, purification, and testing.

Well	Ag	No Ag
A1	0.32	0.15
B1	0.30	0.17
C1	0.41	0.16
D1	0.40	0.15
E1	0.34	0.16
F1	0.34	0.15
G1	0.31	0.14
H1	0.33	0.15
A2	0.29	0.16
B2	0.43	0.14
C2	0.30	0.14
D2	0.33	0.15
E2	0.54	0.15
F2	0.27	0.14
G2	0.28	0.17
H2	0.47	0.15
A3	0.31	0.18
B3	0.30	0.15
C3	0.28	0.13
D3	0.31	0.17
E3	0.29	0.16
F3	0.79	0.14
G3	0.29	0.14
H3	0.27	0.16
A4	0.43	0.15
B4	0.28	0.14
C4	0.27	0.14
D4	0.27	0.14
E4	0.27	0.14
F4	0.45	0.14
G4	0.31	0.13
H4	0.28	0.16
A5	0.26	0.14
B5	0.63	0.14
C5	0.28	0.14
D5	0.52	0.16
E5	0.32	0.13
F5	0.37	0.14
G5	0.28	0.14
H5	0.25	0.14
A6	0.31	0.13
B6	0.29	0.14
C6	0.28	0.14
D6	0.27	0.20
E6	0.27	0.13
F6	0.58	0.14
C	0.26	0.18
BLK	0.24	0.13

Well	Ag	No Ag
A7	0.26	0.16
B7	0.29	0.15
C7	0.26	0.14
D7	0.29	0.14
E7	0.28	0.15
F7	0.89	0.14
G7	0.75	0.14
H7	0.30	0.16
A8	0.29	0.15
B8	0.56	0.14
C8	0.29	0.14
D8	0.27	0.14
E8	0.30	0.14
F8	0.30	0.14
G8	0.27	0.15
H8	0.27	0.13
A9	0.37	0.14
B9	0.50	0.14
C9	0.73	0.15
D9	0.68	0.16
E9	0.41	0.15
F9	0.75	0.16
G9	0.67	0.15
H9	0.28	0.14
A10	0.27	0.16
B10	0.67	0.14
C10	0.26	0.14
D10	0.28	0.14
E10	0.67	0.14
F10	0.33	0.14
G10	0.63	0.14
H10	0.70	0.14
A11	0.83	0.16
B11	0.32	0.18
C11	0.27	0.14
D11	0.28	0.15
E11	0.29	0.15
F11	0.29	0.15
G11	0.26	0.14
H11	0.25	0.14
A12	0.32	0.15
B12	0.30	0.15
C12	0.28	0.17
D12	0.83	0.17
E12	0.28	0.15
F12	0.28	0.14
C	0.28	0.14
BLK	0.24	0.15

Table 13. ELISA target specificity assay. 96 single clones from output selection were grown and screened via ELISA for their putative binding to the target. MLC1 (antigen) was coated into 96-well ELISA plates and periplasm extracts were incubated later. A biotin anti c-tag conjugate was pre-incubated with streptavidin and added to the samples. Lastly, DNPP was added to detect the interaction between the antigen and the nanobody. The plate was read after 45, 90 and 140 minutes without any signal (not shown). Results shown were obtained after an O/N incubation protected from light at RT. In light green, the cells with 1.5-fold versus the control are marked. In green and in dark green, the values with 2-fold and 2.5-fold, respectively are marked.

Ag: Antigen, C: control, BLK: blank.


I
II
III
IV
V
VI
VII
VIII
IX
X
Nb ID
67 (C9)-QVQLVESGGG
71(G10)-QVQLVESGGG
69(E10)-QVQLVESGGG
64(B2)-QVQLVESGGG
66(C1)-QVQLVESGGG
70(G7)-QVQLVESGGG
68(D12)-QVQLVESGGG
81(G9)-QVQLVESGGG
65(B10)-QVQLVESGGG
76(F6)-QVQLVESGGG
77(D5)-QVQLVESGGG
78(E2)-QVQLVESGGG
79(A4)-QVQLVESGGG
80(F4)-QVQLVESGGG
72(A9)-QVQLVESGGG
74(D1)-QVQLVESGGG
75(F3)-QVQLVESGGG
73(B5)-QVQLVESGGG

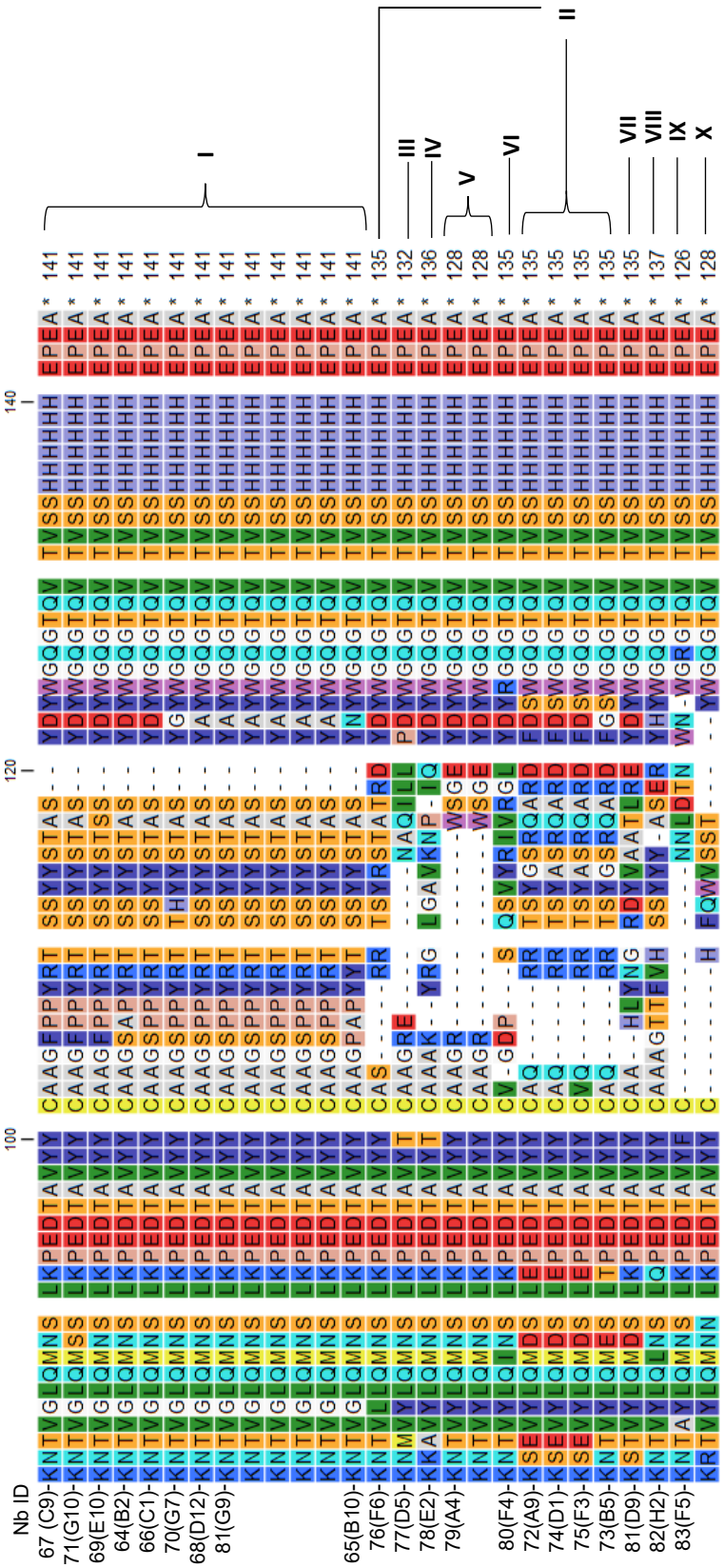


Figure 101. Nanobodies sequence alignment. Twenty-seven different nanobodies were classified into ten different families (Roman numerals).

3.6.2.5 NANOBODY EXPRESSION AND PURIFICATION

To begin working with the nanobodies, firstly they had to be expressed and purified. Therefore, nanobodies were transformed into *E. coli* WK6 strain and grown in TB supplemented with glucose, $MgCl_2$ and ampicillin. Once bacteria had reached an OD_{600} of 0.7, the nanobody expression was induced with IPTG for 4 hours. Afterwards, bacteria were centrifuged and from the cell pellet, the nanobody was purified. To this end, Ni/NTA resin was utilized as nanobodies present a His-tag.

After the purification of all nanobodies, a WB and SST were performed to validate that the purified product contained the nanobody and that it was not contaminated with other molecules. WB and SST for all nanobodies are presented in Figure 102. We detected a band at 28 kDa corresponding to the nanobody. In the silver stain, in the vast majority of nanobodies some thin upper bands could be detected which might correspond to other proteins binding non-specifically to the resin.

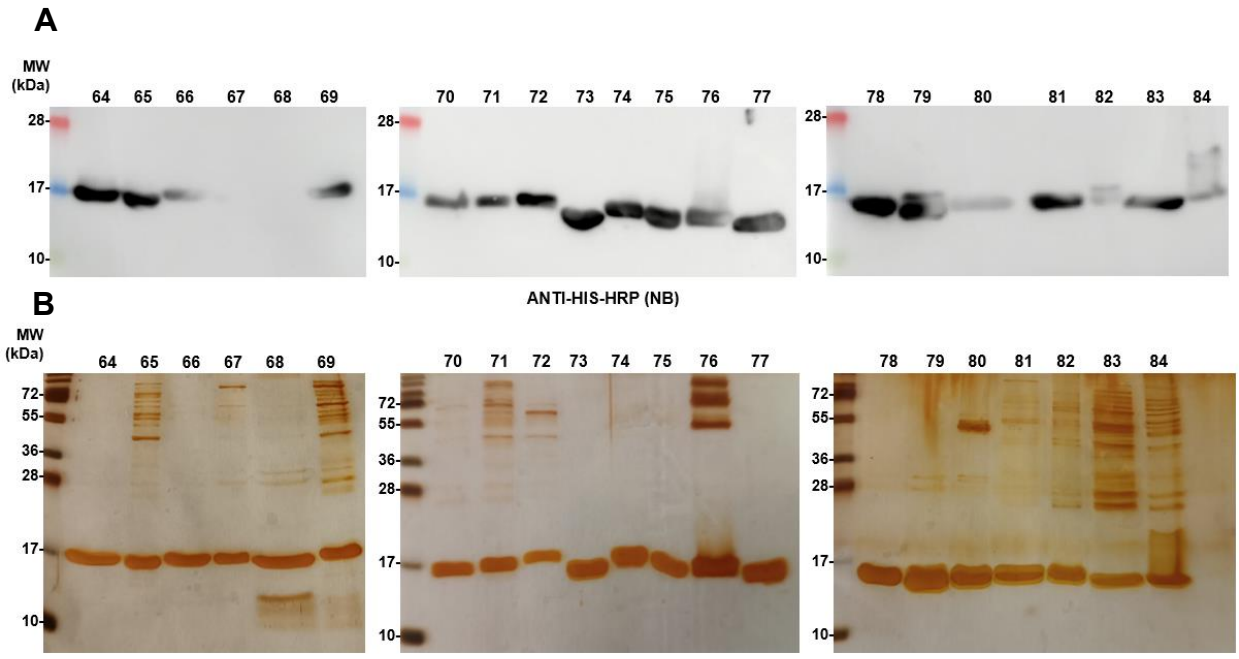


Figure 102. Nanobodies purification. A. WB of the 21 purified nanobodies, the antibody used was anti-His conjugated with HRP. **B.** SST of the 21 purified nanobodies.

NB: nanobody

SEC was also performed to assess the purity of the obtained nanobodies. After their purification, nanobodies were injected for gel filtration in a Superose 6 Increase 5/150 at 0.4 mL/min. The obtained profiles are shown in Figures 103-104-105. Nanobodies were injected at different concentration, depending on the concentration obtained during purification. The injected concentration varies from 0.56 to 3.7 mg/mL and it is written on top of all SEC graphs. All of them presented well defined peaks. Some of them, for instance nanobodies 65, 68 and 70 showed two separated peaks which might be produced by the formation of homodimers. Interestingly, not all nanobodies weighted the same since the peaks appeared at different volumes. In most of them, the main peak was localized before the 3 mL, but in some of them, for example, in nanobodies 67 or 71, the main peak appeared after the 3 mL volume.

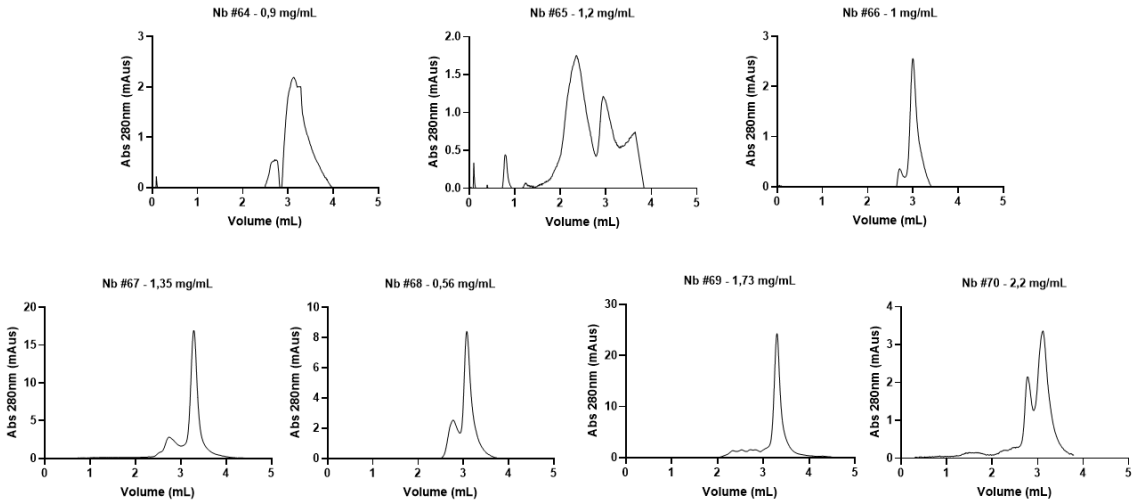


Figure 103. SEC profiles of Nanobodies purification. SEC profile of nanobodies 64, 65, 66, 67, 68, 69 and 70. Samples were run in Superose 6 Increase 5/150 at flow rate of 0.4 mL/min in Tris/NaCl buffer.

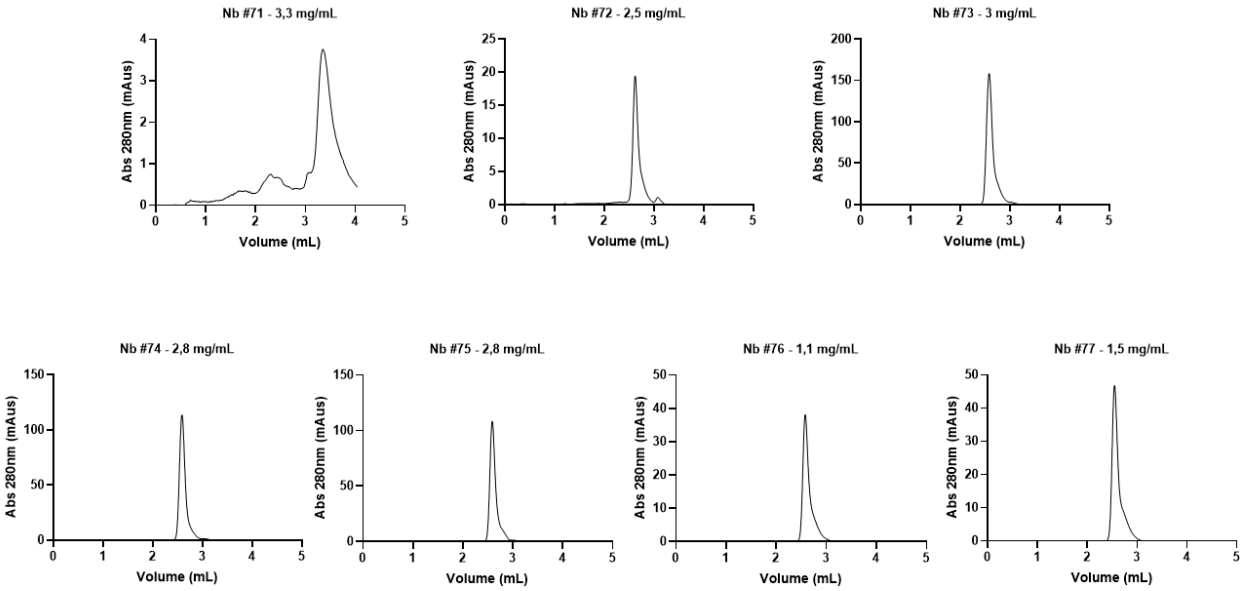


Figure 104. SEC profiles of Nanobodies purification. SEC profile of nanobodies 71, 72, 73, 74, 75, 76 and 77. Samples were run in Superose 6 Increase 5/150 at flow rate of 0.4 mL/min in Tris/NaCl buffer.

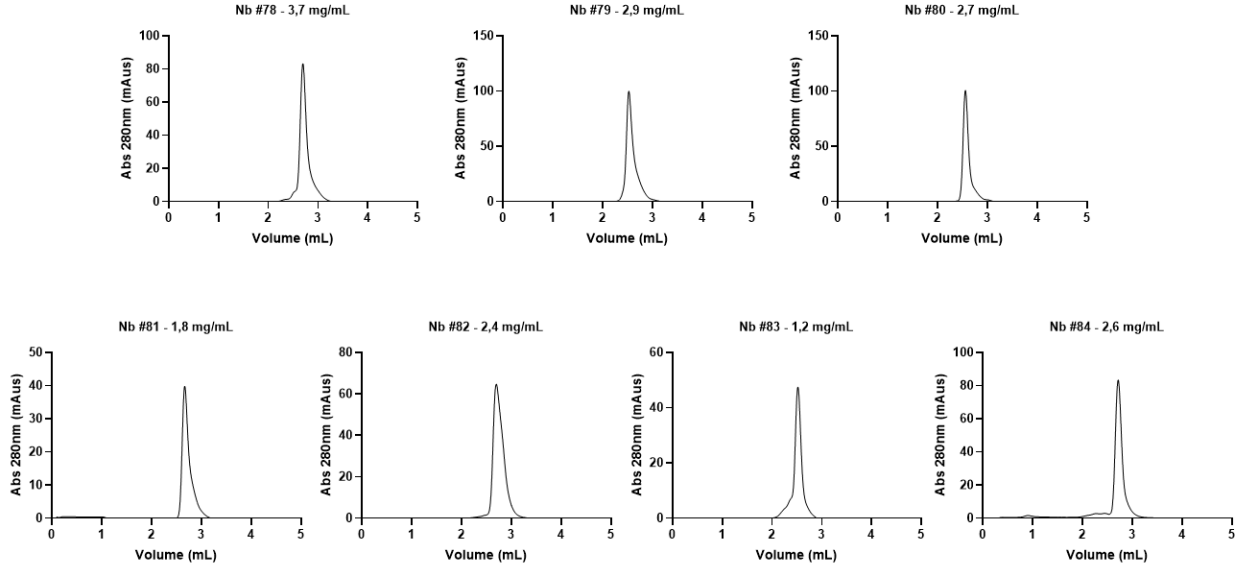


Figure 105. SEC profiles of Nanobodies purification. SEC profile of nanobodies 78, 79, 80, 81, 82, 83 and 84. Samples were run in Superose 6 Increase 5/150 at flow rate of 0.4 mL/min in Tris/NaCl buffer.

3.6.3 NANOBODY BINDING TO MLC1

After the expression and purification of all nanobodies, our objective was to detect whether they would specifically recognize MLC1. It is important to notice that the immunization procedure and the selection of nanobodies were done with the protein reconstituted in liposomes and we were using detergents for protein purification. It was possible that some nanobodies could bind to MLC1 in proteoliposomes, but not in detergent micelles as the protein areas exposed may differ. Since we had 21 different nanobodies and all of them had to be tested, we needed a simple experiment that could be easily reproducible without requiring huge quantities of protein.

Two different strategies were followed. Firstly, we set up a double purification where MLC1 was purified as usual, but then the protein was incubated with a nanobody and the sample was put into contact with Ni-NTA resin and eluted. This resin contains nickel that has affinity for the His-tag fused to the nanobodies. If the nanobody can recognize MLC1, they should be both detected on the elution fraction, whereas if the nanobody does not bind to it, MLC1 should be detected on the flowthrough fraction. For protein detection on the different fractions, WB analysis was performed.

The other strategy was based in generating the construct TwinStrep-MLC1-VFP that would allow us to read fluorescence levels of the protein. The main advantage of using a fluorescent tag is that FSECs require small quantities of protein and it allows for a rapidly analysis of expression levels and changes in the profile. In this case, after the protein purification, the peak obtained for FSEC was collected and incubated with nanobodies. Then, samples were injected again for gel filtration. FSECs were compared trying to detect shifts on the profiles compatible with an increased weight due to the binding of a nanobody. A scheme of both experimental designs is represented in Figure 106.

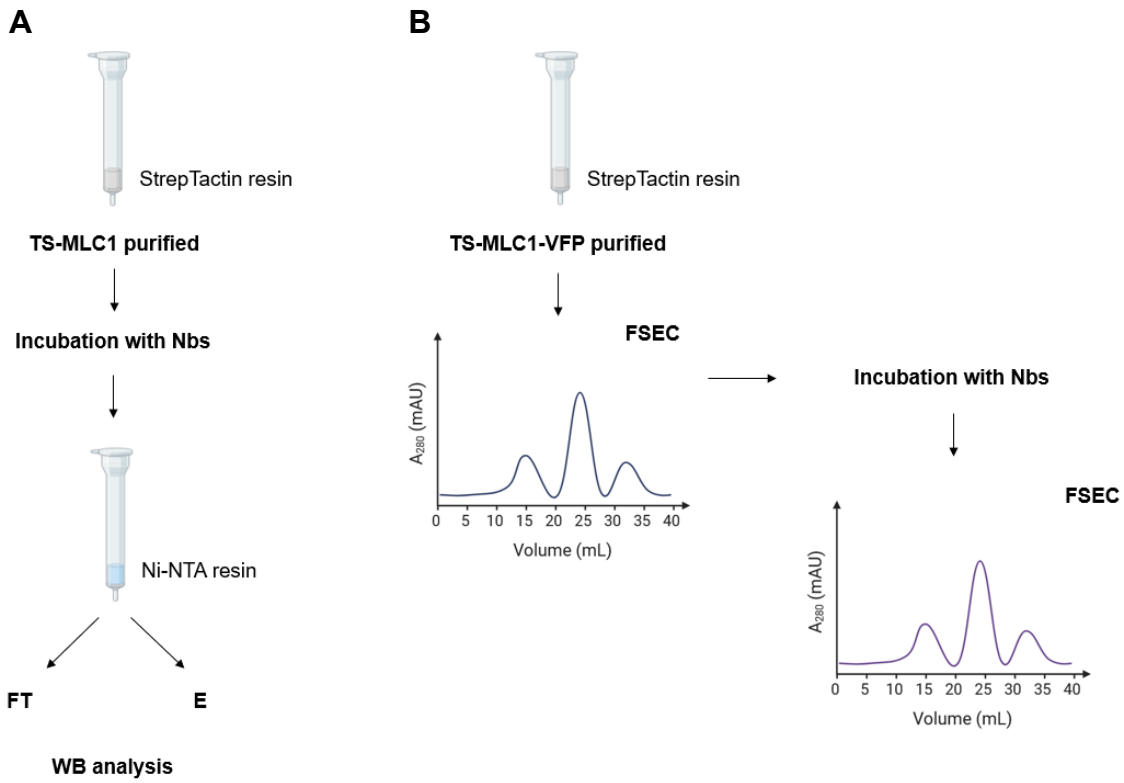


Figure 106. Scheme of the experimental designs followed for testing nanobodies. A. Purification of MLC1 followed by Nb purification. **B.** Purification of MLC1-VFP and posterior FSEC analysis.

TS: TwinStrep, Nbs: nanobodies, FT: flowthrough, E: elution, WB: Western blot, FSEC: Fluorescent size exclusion chromatography.

3.6.3.1 DOUBLE PURIFICATION

HEK293T cells were transfected with TwinStrep-MLC1 and GlialCAM-3xflag. 48 hours post-transfection, cells were centrifuged and solubilized with digitonin. After centrifugation and resin binding, the sample was eluted. The purified protein was split in equal parts and incubated one hour with different nanobodies each. Then, the sample was incubated with Ni-NTA resin, washed, and eluted with imidazole. A fraction of flowthrough and elution were loaded into an acrylamide gel for detecting both MLC1 and the nanobody via WB.

This experiment was repeated with different Nb. In Figure 107A, the tested nanobodies were 67, 68 and 69. MLC1 was detected on the elution fraction of all nanobodies. In the negative control without nanobody incubation (-), MLC1 was solely detected on the flowthrough. As expected, nanobodies were detected on the elution fraction as well.

Since MLC1 in all the Nb condition was present in the elution fraction, we could conclude that these nanobodies did recognize the protein.

Later, nanobodies 70, 71 and 72 were also tested as depicted in Figure 107B. In this case, MLC1 was identified in all fractions, mostly on the flowthrough in all conditions. However, as the signal on the elution fraction on the control is lower than the signal from the elution fractions of nanobodies 70 and 71, we concluded that these nanobodies might as well recognize MLC1. Both nanobodies were not detected on the WB using anti-histidine antibody, probably due to a problem on the WB technique. The repetition with the nanobody 71 confirmed that it did recognize MLC1 (Figure 107C). On this occasion, nanobody 77 clearly did not recognize MLC1 since the protein appeared on the flowthrough fraction. As for nanobody 76, MLC1 was detected in both flowthrough and elution confirming that it was recognized by the nanobody.

Even though with these experiments we could classify nanobodies based on whether they recognized MLC1 or not, in some cases that was difficult to determine. For instance, results obtained with Nb 71 and 72 on Figure 107B were not easily interpreted and we had to repeat Nb 71 to confirm that it was indeed binding to MLC1. At the same time, with these tests, we were working with small protein quantities that were not enough for obtaining a SEC profile. At that moment, it was decided to change the experimental design although not all nanobodies were tried using this approach.

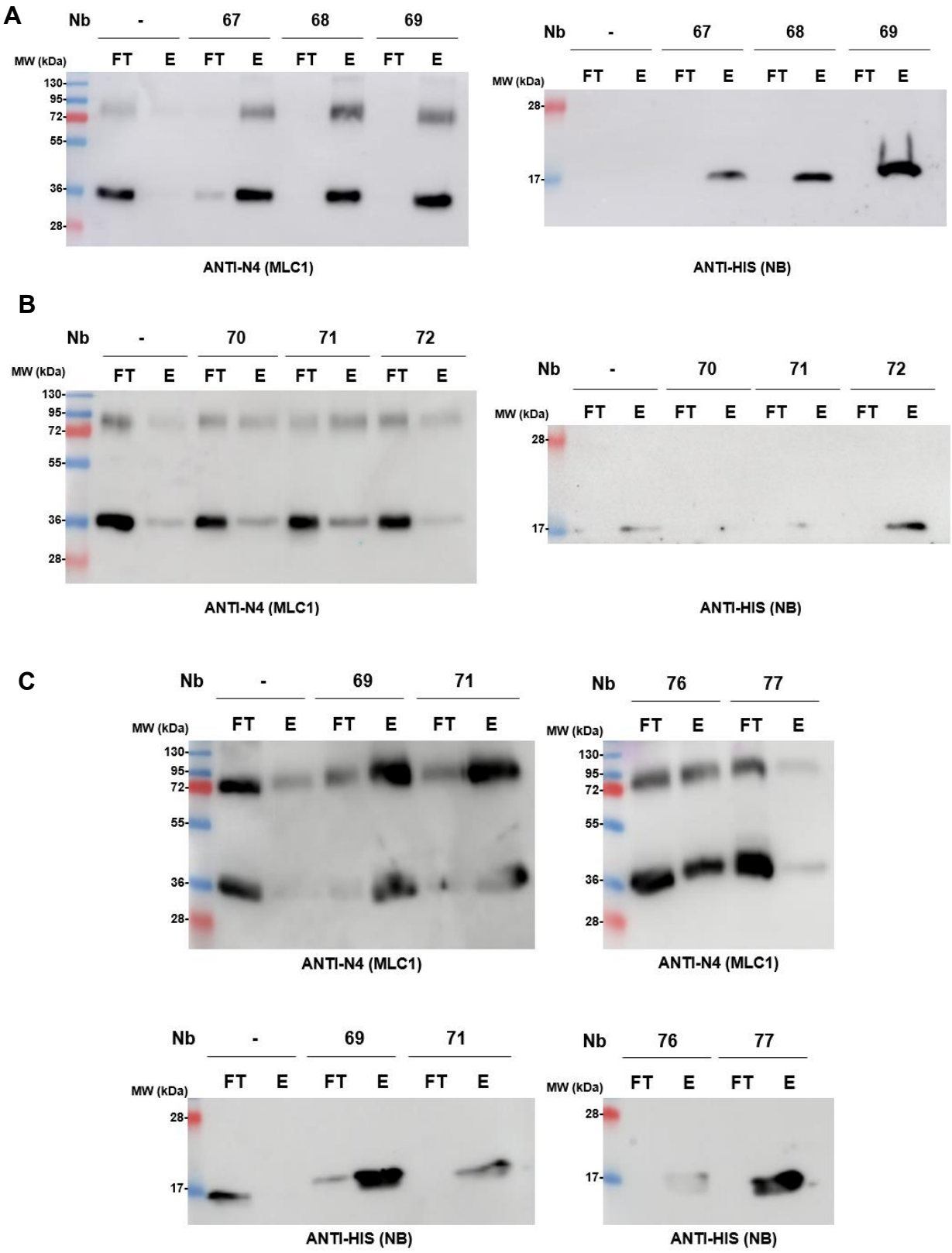


Figure 107. Double purifications of MLC1 incubated with nanobodies. A. WB of the fractions obtained from the double purification with StrepTactin resin and Ni-NTA resin. Nanobodies 67, 68 and 69 were tested. MLC1 and nanobody were detected using specific antibodies. **B.** WB of the fractions obtained from the double purification with nanobodies 70, 71 and 72. **C.** WB of the fractions from the purification with nanobodies 69, 71, 75 and 76.

Nb: nanobody, FT: flowthrough, E: elution.

3.6.3.2 TWINSTREP-MLC1-VFP PURIFICATION

DETERGENT SCREENING

For this experiment, the construct TwinStrep-MLC1-VFP was generated in the first place. The addition of VFP to our protein of interest would allow us to analyse shifts on the FSEC profile by reading fluorescence levels. Although the protein of interest remained the same, the addition of VFP might alter its properties, so the new construct was treated as if it was a completely new protein. Going back to the choice of digitonin for the purification of MLC1 with GlialCAM, we had learnt that the number of particles in cryo-EM containing both proteins was scarce. One of the main reasons for choosing digitonin for the solubilization lied in the fact that it was the only detergent capable of maintaining the union between MLC1 and GlialCAM, but we saw that the resolution of the complex was not feasible. Considering that maintaining the interaction was not essential and that we were working with a new construct, the detergent screening was performed with the objective to select a detergent that would give us the highest protein concentration.

For the screening, 42 mL of HEK293-6E cell culture was co-transfected with TwinStrep-MLC1-VFP together with GlialCAM-3xflag because it was thought that the presence of GlialCAM would increase MLC1 expression. Five different detergents were tested, being DM, DDM, digitonin, LDAO and LMNG. The cell lysate was divided into five groups and solubilized with different detergents. All of them, except digitonin, where prepared adding 0.2% CHS. The solubilization occurred for 1 hour at 2% of detergent, except digitonin which lasted for 2 hours. After the solubilization, samples were centrifuged and filtered for SEC. Fluorescence was read before and after the centrifugation to obtain the percentage of solubilization for each detergent. In Table 14, solubilization rates are listed. Besides LMNG, all the other detergents presented a high rate of solubilization. Fractions from the SEC were collected and read in a fluorometer. FSEC profile is depicted in Figure 108. LMNG group did not show a fluorescent peak, which correlated with the low solubilization previously observed. LDAO and digitonin presented a single and well-defined peak but it was lower than the peak seen for DM and DDM. Since we wanted to obtain the highest quantity of protein possible, we had to decide between DM or DDM. We chose DM because the peak looked more symmetric. Interestingly, DM peak appeared earlier than DDM, meaning that the size of the particles was greater on the first case. However, DM is known to form micelles smaller than DDM (Lal et al., 2023), so the oligomeric state of MLC1 might vary depending on the type of detergent it is embedded to. In conclusion, DM was chosen as the best detergent for MLC1-VFP purification.

Detergent	% Solubilization
DM	95,00
DDM	94,32
LDAO	97,97
LMNG	27,91
Digitonin	89,10

Table 14. Solubilization rates of detergent screening.

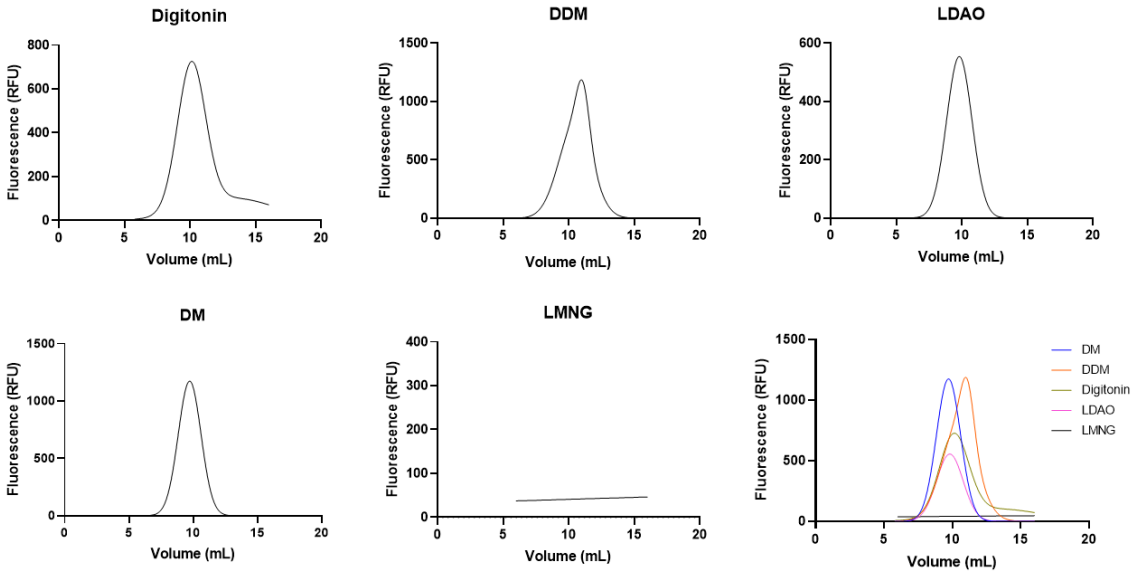


Figure 108. FSEC for detergent screening. Samples were injected for gel filtration using SuperDex 200 Increase 10/300 column with a 0.3 mL/min flow rate in Tris/NaCl plus DDM and CHS. Fractions were collected and fluorescence read. Gaussian fitting was performed on all the FSEC profiles, except for LMNG where a linear fitting was done.

MLC1-VFP PURIFICATION WITH DM

After detergent screening, we proceeded to purify a larger volume of MLC1-VFP to analyse its SEC profile. For this, 0.5 L of cell culture was co-transfected with TwinStrep-MLC1-VFP and GlialCAM-3xflag. After the obtention of membranes, they were solubilized with DM plus CHS and MLC1 was purified using StrepTactin resin and eluted with desthiobiotin. The SEC profile is shown in Figure 109A, where a symmetrical, well-defined peak was observed. The main peak that appeared at approximately 11 mL was surrounded by a smaller peak on the right (13-14 mL) and a “shoulder” or a not very defined small peak on the left, after the void volume. MLC1 was detected by WB analysis during the purification process, in the total protein, solubilized and eluted fractions (Figure 109B). It was also detected through WB in the different peaks seen at SEC (Figure 109C).

The main peak was collected, frozen in liquid nitrogen and stored at -80°C for two days. Then, the sample was thawed and injected again for gel filtration. With this, we wanted to assess the possibility of freezing the purified protein. It was observed, as shown in Figure 109D, that the mAus of the peak were greatly decreased after freezing. In conclusion, MLC1-VFP could not be frozen and the future studies with the nanobodies had to be done on the same day as the protein purification.

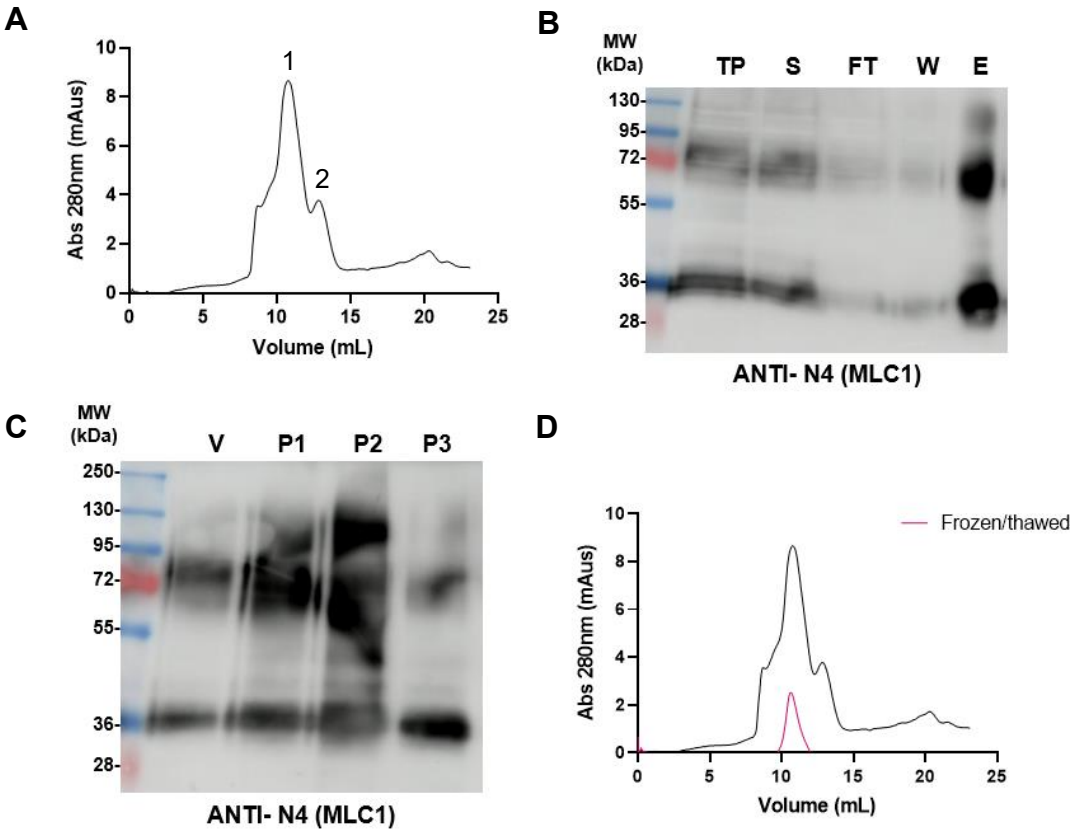


Figure 109. MLC1-VFP purification with DM. A. SEC profile of TwinStrep-MLC1-VFP purified protein with DM. Sample was run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent. **B.** WB of the purification process detecting MLC1. **C.** WB of the SEC peaks detecting MLC1. **D.** Comparison of SEC profiles before (black) and after freezing (red).

TP: total protein, S: solubilized, FT: flowthrough, W: wash, E: elution, V: void, P: peak.

3.6.3.3 NANOBODIES BINDING TO MLC1

Once the purification conditions for MLC-VFP had been stated, we started assessing whether the nanobodies could recognize the purified protein. The experimental design was the following. Firstly, the protein was purified and injected for gel filtration. Based on the SEC profile, the main peak was collected and split in parts which were incubated each with a nanobody O/N. The following day, all the samples were injected again for gel filtration and SEC profiles were compared to see mobility shifts. Besides, fluorescence levels from the SEC fractions were read to analyse possible mobility shifts as well. Results are shown in Figures 110, 111, 112, 113, 114 and 115.

In Figure 110 results from the first experiment using the aforementioned design are represented. The obtained SEC profile (top left) showed a high peak corresponding to aggregates that do not have fluorescence. Then, a wide peak is observed at 11 mL approximately. The collection of this peak and its fluorescent reading (top center) revealed that the SEC peak was fluorescent, meaning that it corresponded to our protein of interest VFP-tagged. The mentioned peak was split in three equal parts that were incubated O/N with nanobodies 64, 76 and 77. After, samples were injected again for gel filtration. The three SEC profiles (bottom row) presented two peaks, the first one corresponding to MLC1 and the second one that appeared later corresponding to the nanobody alone. Comparing them with the profile previously obtained in the absence of nanobody, it was observed that all the peaks appeared at around 11 mL, approximately. Hence, a mobility shift was not observed in any of the tested conditions. The FSEC profiles did not present a displacement neither (top right). Based on this experiment, it could be concluded that these three nanobodies did not recognize MLC1. However, in the double purification experiment it was established that nanobody 76 did recognize the protein, but 64 and 77 failed to bind to it. It could be possible that changing the detergent affected the nanobody recognition, as the protein areas exposed depend on the size of the micelle, which is variable among detergents. This would explain why the results obtained in the double purification with digitonin were different that the ones obtained now with DM.

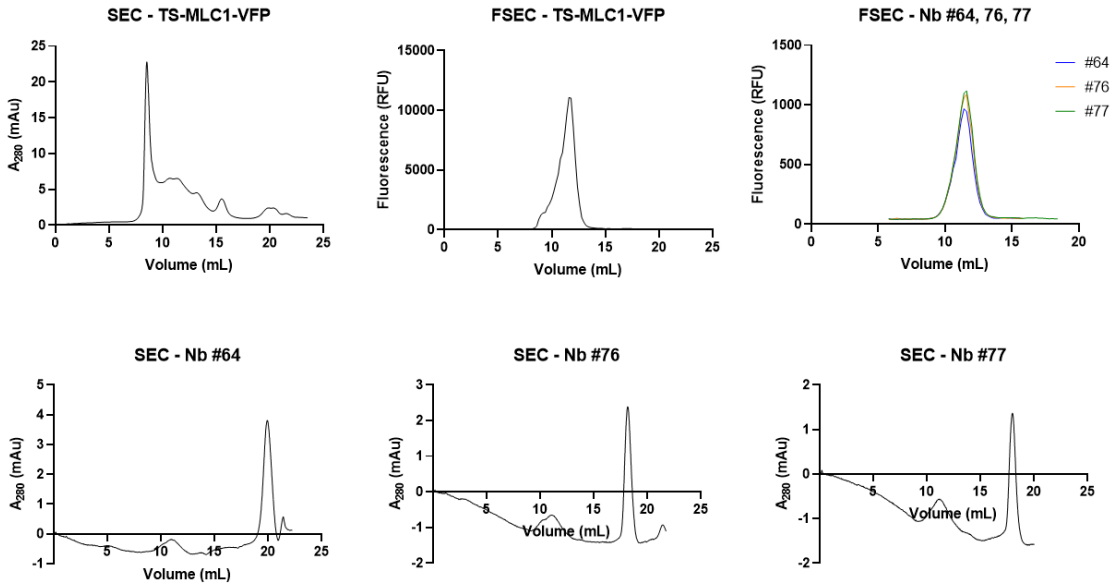


Figure 110. Nanobodies 64, 76 and 77- MLC1 recognition. SEC (top left) and FSEC (top center) profiles of the purified TwinStrep-MLC1-VFP protein. After the incubation with nanobodies 64, 76 and 77, FSEC (top right) and SEC (bottom row) profiles were analyzed. Samples were run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

The same experiment was repeated working with other nanobodies (Figure 111). Purifying MLC1 in same conditions, a similar SEC profile was obtained (top left). The peak after the void was collected and fluorescence was detected (top centre). Then, it was split in five equal parts and incubated with nanobodies 65, 67, 70, 75 and 80. The analysis of their SEC profile revealed that nanobodies 65, 67, 70 and 80 did not generate changes in the profiles (middle and bottom row), neither in the FSEC profiles (top right). In these cases, the first peak appeared around 11 mL similarly to the sample before nanobody incubation and later, a second peak corresponding to the nanobody alone was visible.

Surprisingly, that was not the case for nanobody 75. Its SEC profile revealed a peak around 15 mL after the peak corresponding to MLC1 (~11 mL). A very small peak around 20 mL could also be observed which would correspond to the nanobody. Since the expected nanobody peak (at around 20 mL) had few mAU, we hypothesized that the nanobody had bind to MLC1 and that binding would correspond to the peak observed at approximately 16 mL. However, the addition of a nanobody to our protein of interest would increase the mass and the expected peak should appear before the peak corresponding to MLC1 alone, which was not the case. A possible explanation for this would be that the binding of the nanobody to MLC1 would be only on dimers or monomers of the protein.

Then, the peak compatible with containing MLC1 and nanobody 75 was collected, concentrated, and injected again for gel filtration. Curiously, this second injection displayed a different pattern than the previously observed (bottom row centre). In this case, two separated, well-defined peaks were seen around 13 and 17 mL, which did not correspond to the previously represented peak at 15 mL, approximately. A plausible interpretation was that the concentration of the peak had provoked changes in the stoichiometry of the complex which would have been disassembled. Still, if that was the case, the second peak should appear at the same volume that the one corresponding to the nanobody alone (at 20 mL). Since it was bigger, it could mean that the nanobody was bound to a dimeric MLC1 form and the concentration caused the disassembling of the dimer, but remaining bound to a monomeric form. Nevertheless, this would not explain why the first peak from the second injection (13 mL) had appeared before the concentrated peak.

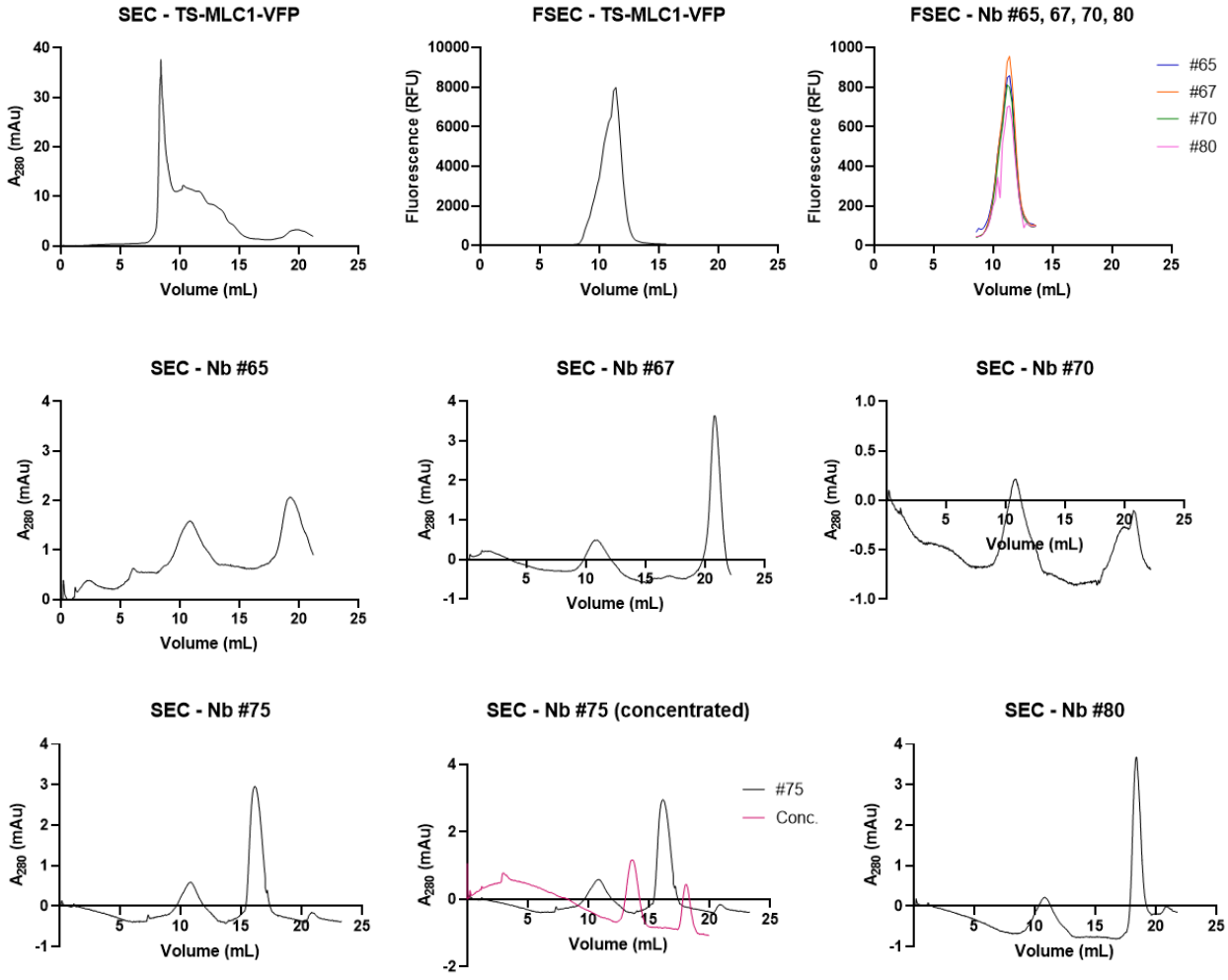


Figure 111. Nanobodies 65, 67, 70, 75 and 80- MLC1 recognition. SEC (top left) and FSEC (top center) profiles of the purified TwinStrep-MLC1-VFP protein. After the incubation with nanobodies 65, 67, 70, 75 and 80, FSEC (top right) and SEC (middle and bottom rows) profiles were analyzed. Samples were run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

It should be considered that many stoichiometries could be present in this case. We knew that MLC1 could be trimeric, dimeric, or monomeric and the binding with nanobodies can give raise to many complex possibilities. For instance, one single molecule of a nanobody could bind the MLC1 trimer, but also two or three nanobodies could bind to it, which would result in subtle changes on the SEC profile. Although we did not have a clear explanation for the behaviour of nanobody 75, we concluded that it recognized MLC1.

The testing with nanobodies 66, 73, 74 and 83 did not show shifted patterns for SEC nor FSEC profiles (Figure 112). The profiles obtained from the purification of TwinStrep-MLC1-VFP showed the same pattern as the previously seen, with a fluorescent peak around 11 mL. The posterior incubation with nanobodies did not alter the SEC profiles, neither it affected the FSEC ones.

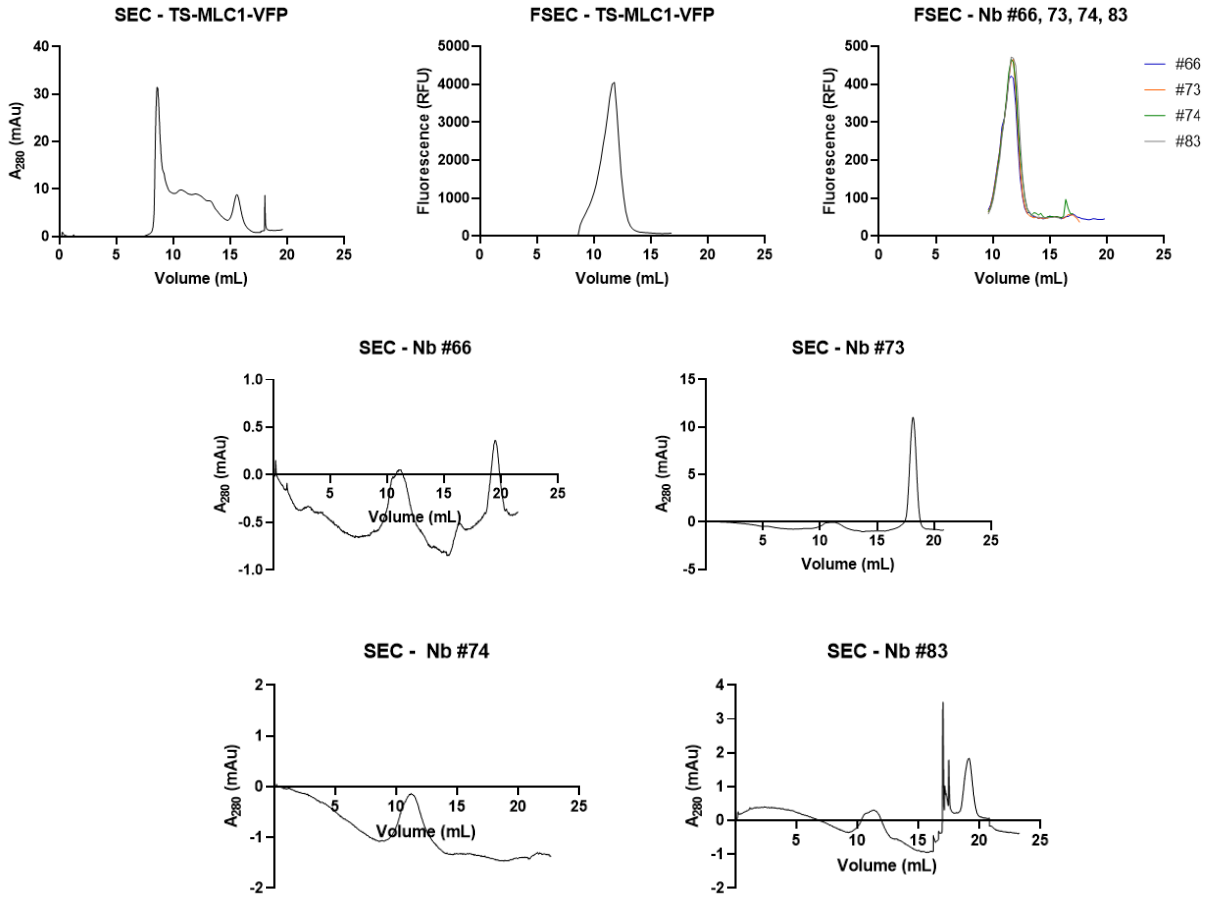


Figure 112. Nanobodies 66, 73, 74 and 83- MLC1 recognition. SEC (top left) and FSEC (top center) profiles of the purified TwinStrep-MLC1-VFP protein. After the incubation with nanobodies 66, 73, 74 and 83, FSEC (top right) and SEC (middle and bottom rows) profiles were analysed. Samples were run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

For the nanobodies 68, 79, 81 and 84 similar results were found (Figure 113). In this case, the SEC and FSEC profiles were analogous to the previously obtained. The profiles after the nanobody incubation showed two distinguished peaks compatible with MLC1 alone on the first peak and the nanobody alone on the second one. For nanobody 79, a small peak was noticed between the mentioned peaks, similarly to the seen case for nanobody 75. Due to limited amounts of protein, no further experiments could be performed.

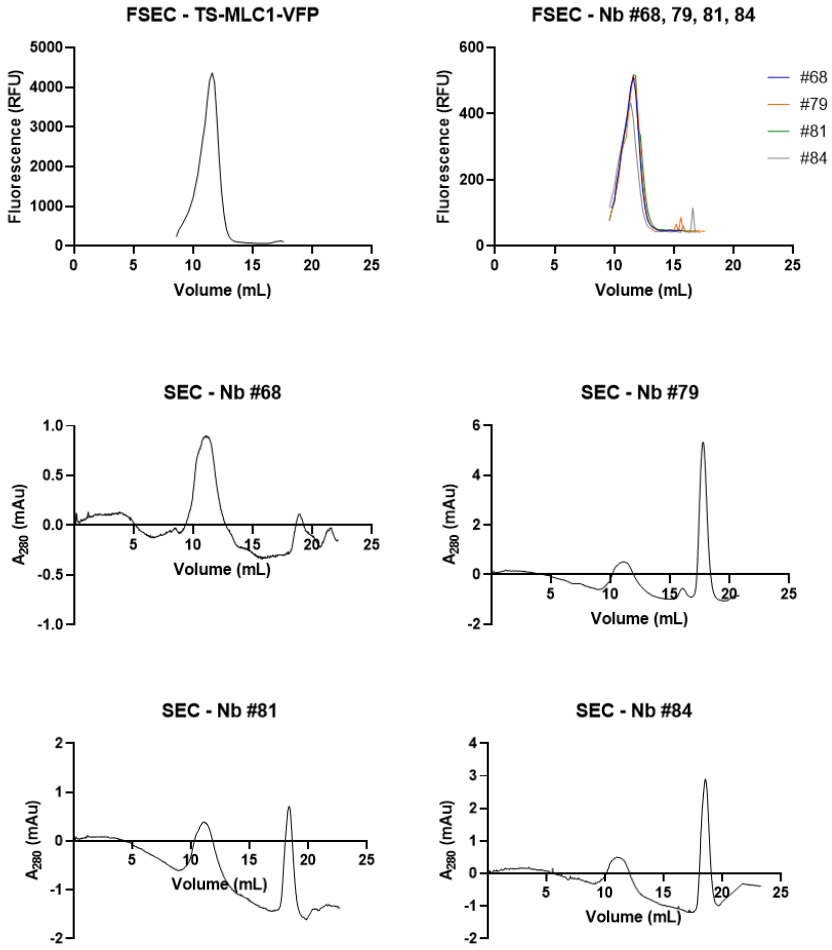


Figure 113. Nanobodies 68, 79, 81 and 84- MLC1 recognition. FSEC (top left) profile of the purified TwinStrep-MLC1-VFP protein. After the incubation with nanobodies 68, 79, 81 and 84, FSEC (top right) and SEC (middle and bottom rows) profiles were analyzed. Samples were run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

For nanobodies 69, 71, 72 and 82 similar results were obtained (Figure 114). Mobility shifts were not observed in none of the tested nanobodies in this experiment. Lastly, the remaining nanobody (78) was also tested with this experimental design. Results are depicted in Figure 115. The addition of this nanobody did not produce any shift on SEC profile.

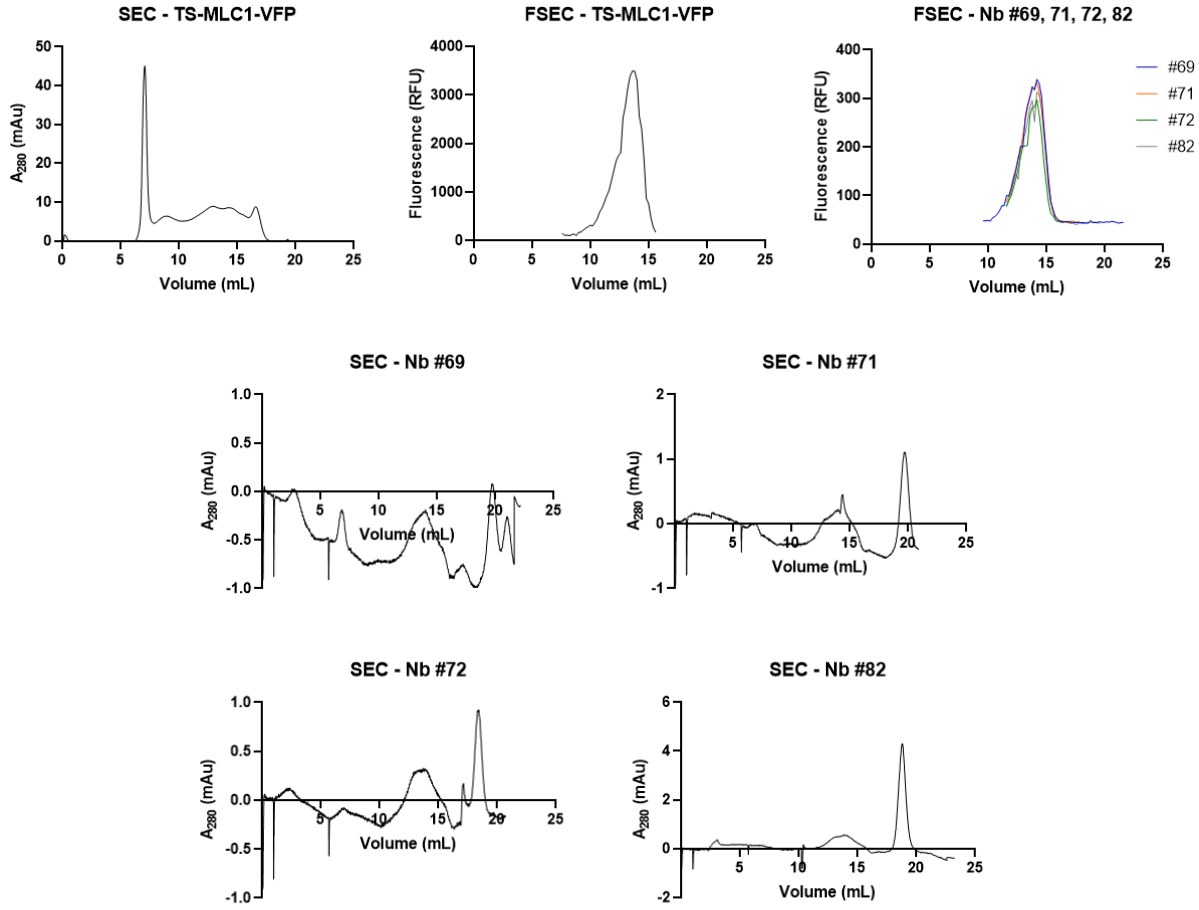


Figure 114. Nanobodies 69, 71, 72 and 82- MLC1 recognition. SEC (top left) and FSEC (top center) profiles of the purified TwinStrep-MLC1-VFP protein. After the incubation with nanobodies 69, 71, 72 and 82, FSEC (top right) and SEC (middle and bottom rows) profiles were analyzed. Samples were run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

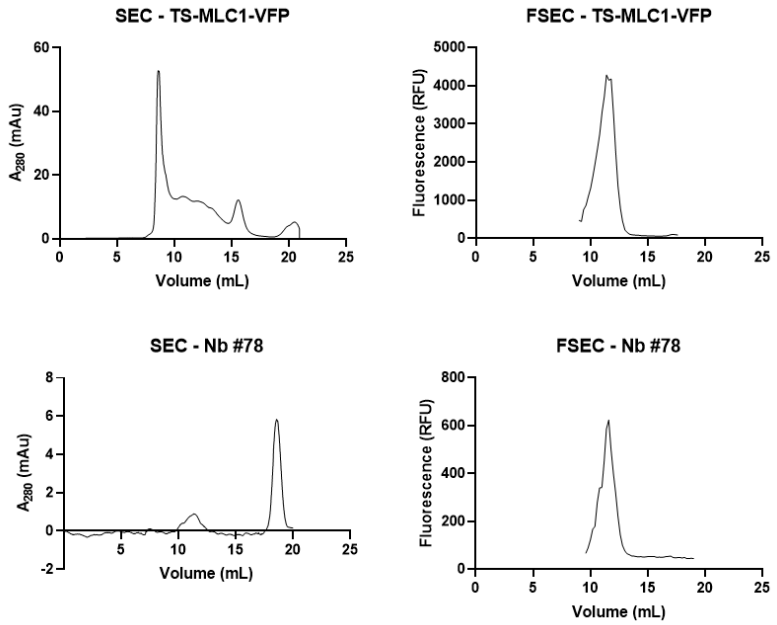


Figure 115. Nanobody 68- MLC1 recognition. SEC (top left) and FSEC (top right) profiles of the purified TwinStrep-MLC1-VFP protein. After the incubation with nanobody 78, FSEC (bottom right) and SEC (bottom left) profiles were analyzed. Samples were run in SuperDex 200 Increase 10/300 at flow rate of 0.4 mL/min in Tris/NaCl buffer plus detergent.

3.6.3.5 BIACORE

Experimentally, the binding affinity of all nanobodies with the protein of interest was tested through Biacore. Surface plasmon resonance (SPR) biosensors, such as Biacore systems have become standard tools for protein interaction studies. They are typically used for measuring binding affinities, as well as kinetic rate constants. SPR is an optical phenomenon that occurs as a result of the internal reflection of light at a metal film-liquid interface. Biacore biosensors measure the shift in SPR angle due to mass changes at the surface of the chip. This allows for the detection and measurement of protein-protein interactions in real time. It usually requires the immobilization of one interactant on the sensor chip surface followed by the injection of the second interactant (Leonard et al., 2017).

Biacore experiment was performed at the Biomolecular analysis unit of CCIT-UB. In our case, MLC1 purified protein in digitonin micelles was immobilized into a chip containing a matrix of carboxymethylated dextran. After a first screen, it was decided to work at pH 5 in Tris/NaCl buffer with 0.005% digitonin. A multicycle approach was followed where all nanobodies were tested.

For equilibrium analysis, saturation values (RU) were plotted against nanobody concentration (M). Only three nanobodies (69, 71, and 83) were fitted to an exponential curve (Figure 116). The rest of them were fitted in a linear curve meaning that they do not bind with high affinity to the protein. In Table 16, the obtained K_D rates are shown. Nanobody 79 was analysed in a previous single cycle study and the K_D obtained was 1.1E-1 M. Since the K_D was high, it was decided to not include this nanobody on the following multicycle study. Calculating the equilibrium constant, it was observed that the two best nanobodies were 69 and 71, which had a K_D value around 800 nM. In general, equilibrium dissociation constants in the nanomolar or picomolar range are considered ideal (Muyldermans, 2013).

In conclusion, all our nanobodies had low affinity for the target protein and they presented a fast kinetics, since they were quickly dissociated, for this reason the dissociation rates were not calculated.

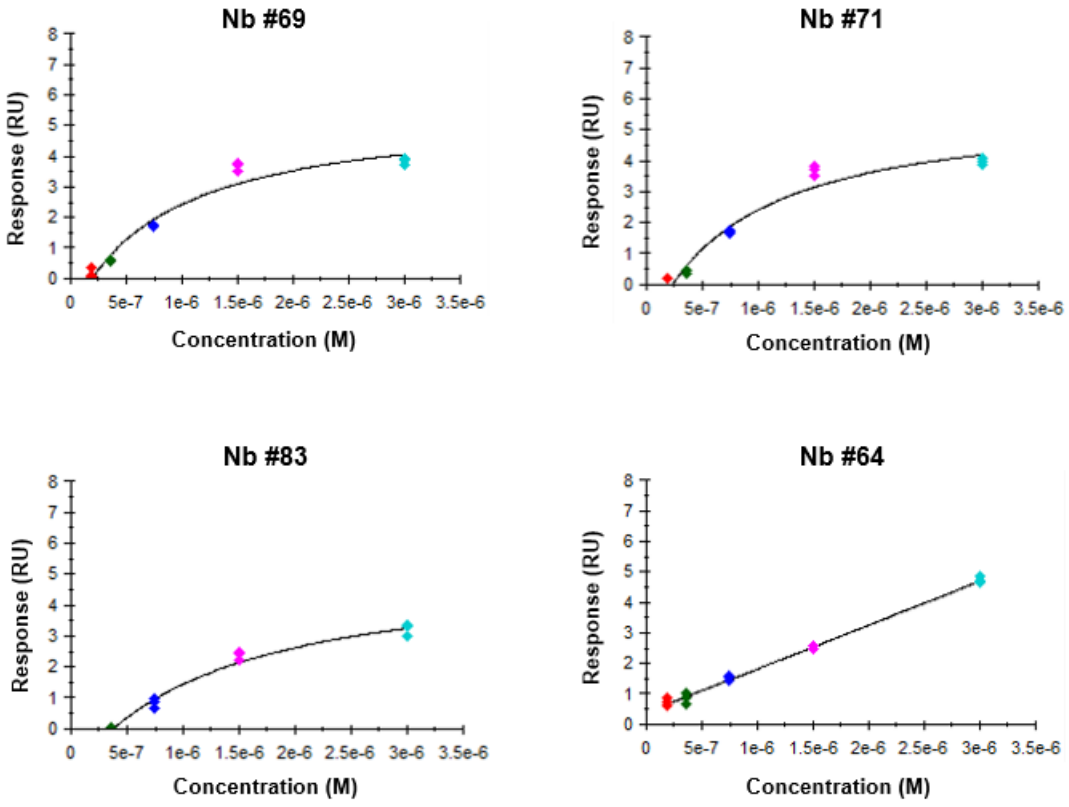


Figure 116. Surface plasmon resonance analysis. Equilibrium fits for the interaction of MLC1 and nanobodies 69, 71, 82, and 64. Nanobody 64 is plotted as a representative nanobody showing a linear equilibrium fit. Each concentration had experimental triplicates.

NB	KD (M)	Rmax (RU)	Chi² (RU²)
64	2.11E-02	30292.2	0.016
65	1.86E-01	257622.3	0.0353
66	3.48E-06	11.3	0.00672
67	5.56E-06	6.4	0.461
68	2.00E-06	11.1	0.0775
69	8.44E-07	6.9	0.133
70	1.56E-06	6.2	0.552
71	8.57E-07	7.5	0.134
72	8.47E-04	1323.2	0.00569
73	2.83E-01	239751.2	0.0104
74	1.05E-05	6.3	0.00376
75	2.56E-01	252100.5	0.0107
76	2.36E-01	289898	0.0136
77	1.58E-05	26.6	0.0049
78	6.95E-02	59550.1	0.006
80	9.42E-03	7607.1	0.00396
81	1.28E-05	23.6	0.0226
82	1.10E-01	145960.2	0.0254
83	1.89E-06	7.8	0.0324
84	1.85E-01	206509.4	0.0273

Table 15. Equilibrium constant results. In terms of KD, the nanobodies showing more affinity for MLC1 are Nb 69 and 71.

Then, a quick study using AF3 prediction models was done to analyse the interaction between trimeric MLC1 and nanobodies. This revealed that all nanobodies bound to one or two MLC1 monomers that constituted the trimer. In the case of the two best SPR candidates (Nb 69 and 71), these nanobodies could bind two monomeric proteins from the trimer. A deeper analysis was performed with these two in order to determine which specific residues were necessary for the interaction between the nanobody and MLC1. The identification of the amino acids implied in the binding could be important in the future as the modification of these residues could be designed to improve the affinity of the interaction. Since these two nanobodies corresponded to the same family, their sequence its very similar. Thus, as observed in Figure 117, they recognize similar MLC1 residues. Specifically, for nanobody 69 AF predicted model, two monomers of MLC1 would be recognized by the nanobody. It might bind to two asparagine residues from MLC1 TM4 domain. Also, a tyrosine residue from the nanobody might bind to an aspartic acid residue from TM8 domain of the adjacent MLC1 (Figure 117A). In the case of nanobody 71 AF predicted model, the identified residues were also T58 and Y60 that bind to two asparagine residues from MLC1 TM4 domain. In addition, an arginine residue from the nanobody could bind to an aspartic acid residue from the extracellular loop

between TM 7 and 8 (Figure 117B). Further research is needed to determine the exact binding sites of the obtained nanobodies.

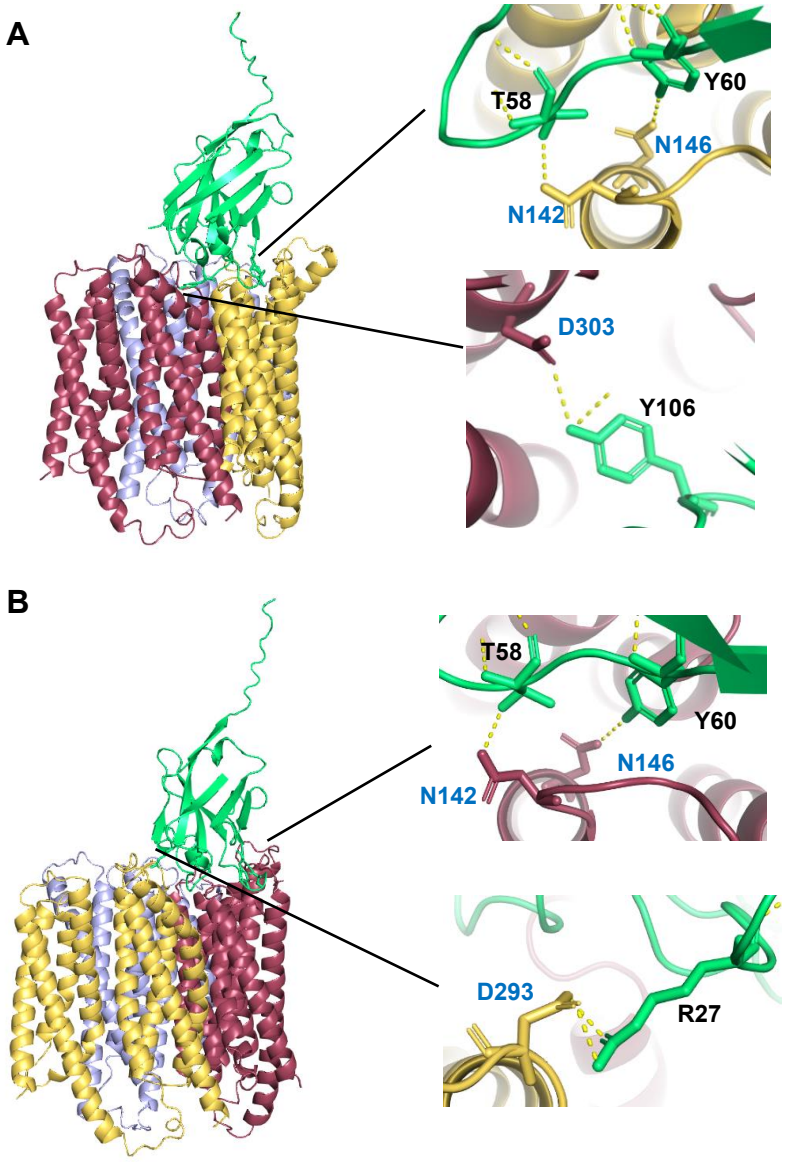


Figure 117. Nanobody-MLC1 binding model. A. Nanobody 69 putative binding with MLC1 homo-trimer. **B.** Nanobody 71 putative binding with MLC1 homo-trimer. MLC1 monomers are coloured in raspberry, yellow orange, and light blue. Nanobodies are coloured in lime green. Residues in black correspond to nanobody amino acids, while the residues in blue correspond to MLC1.

3.7 STRUCTURE ANALYSIS

Although the obtention of a high-resolution protein structure had failed, we were able to further study MLC1 3D structure based on the preliminary model achieved together with AlphaFold predicted models. Our study of MLC1 structure was divided in three parts. Firstly, we started by comparing MLC1 with other protein structures trying to infer on the function of our protein of interest. Secondly, we located the missense patient-derived mutations on the 3D structure and we focused on the interaction between MLC1 transmembrane domains. Lastly, we studied the MLC1 trimer.

3.7.1 INSIGHTS INTO MLC1 FUNCTION BASED ON ITS STRUCTURE

As previously stated in the Introduction of this thesis, MLC1 is known to share low sequence identity (only 20%) with the voltage gated potassium channel Kv1.1 alpha subunit (KCNA1) (Teijido et al., 2004). Considering that similar sequences adopt similar folds and that the protein folding determines the protein function, it was suggested that MLC1 could act as an ion channel. In accordance to this, MLC patients may present epileptic seizures which is a symptom commonly found in channelopathies, but not in leukodystrophies (Hamilton et al., 2018).

A first analysis of amino acid sequence suggested that *MLC1* gene could have evolved from a gene duplication of the first 4 transmembrane domains (TM), resulting in two big domains TM1-4 and TM5-8 connected through an intracellular loop (Figure 118A-B). Using the PDBeFold algorithm, we identified proteins similarly folded to MLC1. Analysis of the TM1-4 MLC1 identified the prokaryotic soluble copper storage protein (PDB ID 6Q6B) from *Streptomyces lividans* (Figure 118C) (Pla-Casillanis et al., 2022). This protein forms homotetramers where each monomer comprises four α-helices arranged to form a four-helix bundle motif. It has the capacity to bind between 13 and 20 cuprous ions per four-helix bundle (Straw et al., 2018). Analysing the whole MLC1 sequence, the algorithm identified the phosphate-specific transport system accessory protein (PhoU) from *Pseudomonas aeruginosa* (PDB ID 4Q25) (Figure 118D). This protein consists of two structurally similar three-helix bundles. Its function is still unknown, but it is believed to play a key role in phosphate homeostasis (Lee et al., 2014). Even though all data suggests that MLC1 may function as an ion channel, to date no conductivity has been detected for this protein (Estévez et al., 2018).

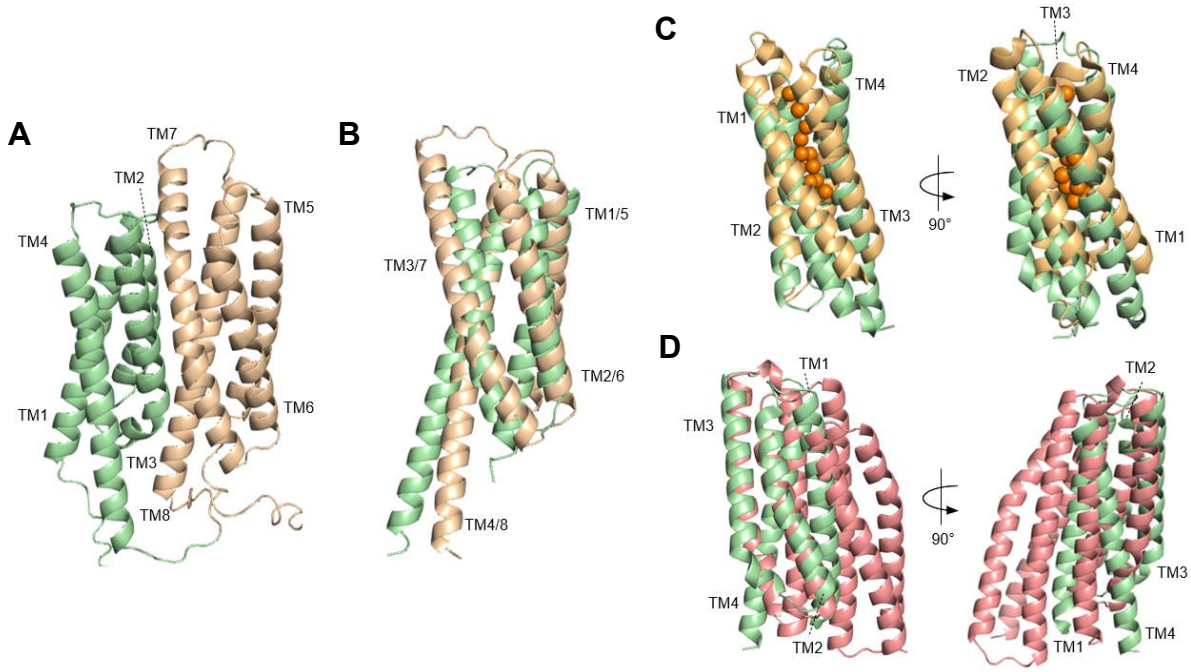


Figure 118. Structural models of MLC1. A. AlphaFold structural model for monomeric MLC1. TMs 1-4 and 5-8 are coloured in pale green and wheat, respectively. **B.** Structural alignment of the two MLC1 domains, composed of TMs 1-4 (pale green) and 5-8 (wheat). **C.** Structural alignment of MLC1 TM 1-4 with a prokaryotic soluble copper storage protein from *Streptomyces lividans* (PDB ID 6Q6B) (light orange). Two different views rotated by 90° are shown. **D.** Structural alignment of MLC1 with the PhoU protein from *Pseudomonas aeruginosa* (PBD ID 4Q25) (in salmon). Two different views rotated by 90° are shown.

3.7.2 MLC1 MUTATIONS

Mutations in MLC patients are spread through all the sequence (Leegwater et al., 2001). However, when these residues are identified in the 3D structure of MLC1, they are mostly located in the interphase of TMs 1-4 and 5-8 (Figure 119). Previous studies from the group co-expressing the two halves of MLC1 (TMs 1-4 and TMs 5-8) in *Xenopus* oocytes demonstrated that the halves were able to reach the plasma membrane, thus indicating that they would interact with each other (Estévez et al., 2018).

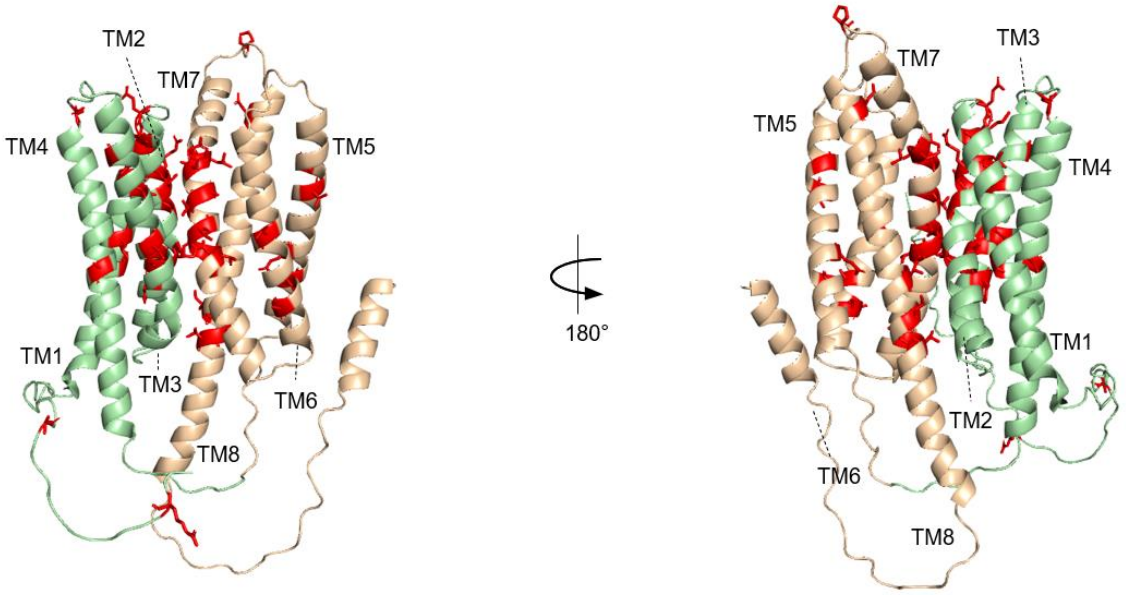


Figure 119. Location of MLC mutants in MLC1 structural model. MLC1 patient-derived mutations are shown as red sticks. Two different views rotated by 180° are shown.

To further validate the interaction between the two halves of MLC1, the following NanoBiT experiment was set up. We cloned into an expression vector the two halves of MLC1 that we called Split 1 and Split 2, and they were fused to the SmallBiT or LargeBiT, respectively. Then, we chose two patient-derived mutations that were hypothesized to be important for the interaction of the splits. These mutants were G88V and P92S, both located on the second transmembrane domain of MLC1 (Figure 120A). Split 1 constructs including these mutants were also generated.

Analysing the interaction between Split 1 and Split 2, it was observed that this interaction was significantly decreased in both mutants compared to the WT (Figure 120B). However, a reduced interaction could also mean that the mutant splits were less expressed than the WT. Thus, we then decided to verify whether the observed decrease was indeed due to a reduction of interaction instead of an altered expression. With this aim, we fused the Gaussia luciferase to Split 1. After adding the substrate coelenterazine, the luminescence was read. As seen in Figure 68C, the luminescence was slightly decreased in the mutated Split 1, being statistically significant in the G88V. Nevertheless, the signal reduction was not as large as the decrease observed in the interaction between splits, so we could confirm that residues G88 and P92 were important for the split's interaction. In the case of G88V mutant, the replacement of glycine, a polar amino acid residue, with valine, a hydrophobic amino acid residue, might affect the interaction between the splits by charges repulsion. For the P92S mutant, proline is known as a

potent breaker of the α-helix, so its substitution for a serine would alter the orientation of the TM2, which might not impede its interaction with the TM7 from Split 2.

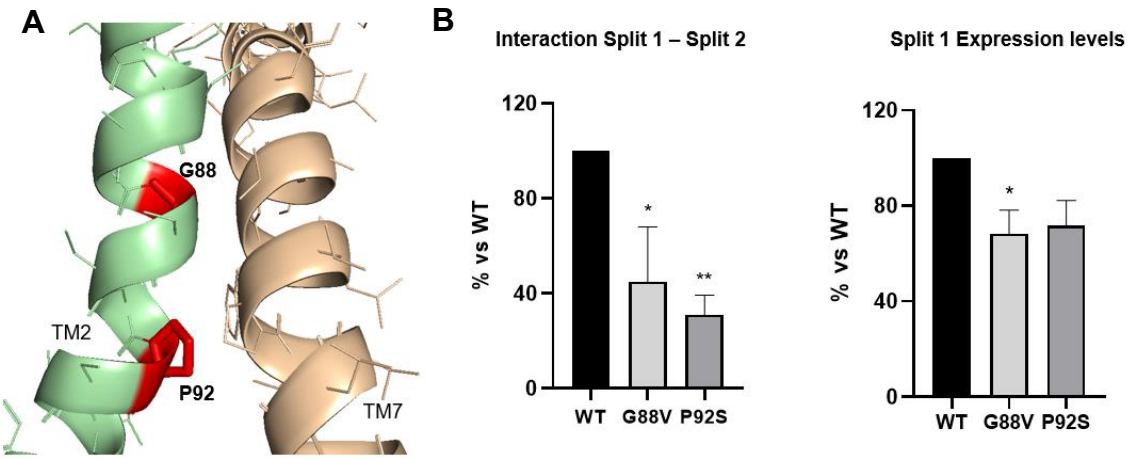


Figure 120. Splits assays. A. AlphaFold representation of TM2 and TM7 where the residues G88 and P92, found mutated in MLC, are marked in red. **B.** Splits interaction. NanoBiT experiments on transfected HEK293T cells. Data comes from five independent experiments. Statistical analysis was assessed using One-way ANOVA analysis (** p-value < 0.01) with Dunnett's multiple comparisons test. * $p < 0.05$; ** $p < 0.01$ compared to WT. Mean ± SEM is represented. **C.** Split 1 expression levels. Gluc was cloned into Split 1 to determine its levels of expression. Statistical analysis was assessed using One-way ANOVA analysis (* p-value < 0.05) with Dunnett's multiple comparisons test. * $p < 0.05$ compared to WT. Mean ± SEM is represented. Data comes from five independent experiments.

3.7.3 PORE

Our low-resolution structure revealed that MLC1 was forming a homo-trimer. Similarly, it was recently reported that MLC1 constitutes a homo-trimeric complex in detergent micelles and proteoliposomes (Hwang et al., 2021). Analysing the trimer, top and bottom views exposed the formation of a pore where three histidines and three tryptophans, one from each monomer, were faced towards the pore (Figure 121). This arrangement resembled the proton-selective ion channel M2 from Influenza virus. M2 channel is gated by pH and it is essential for acidifying the interior of virions during virus uncoating in the lumen of endosomes. When the pH is high, the channel is closed because the histidine located facing the pore is not charged, therefore, the tryptophan obstructs the pore near its cytoplasmic end. In contrast, when the pH is low, the histidine is charged, and as a result, the tryptophan rotates to a conformation parallel to the pore's axis, which permits H^+ to flow (Tang et al., 2002). Taking this channel into account, we tested whether MLC1

could act as a proton channel. Dr. Héctor Gaitán Peñas injected MLC1 cRNA to *Xenopus* oocytes, electric currents were registered with an acidic pH extracellular media. Unfortunately, no currents were detected for MLC1.

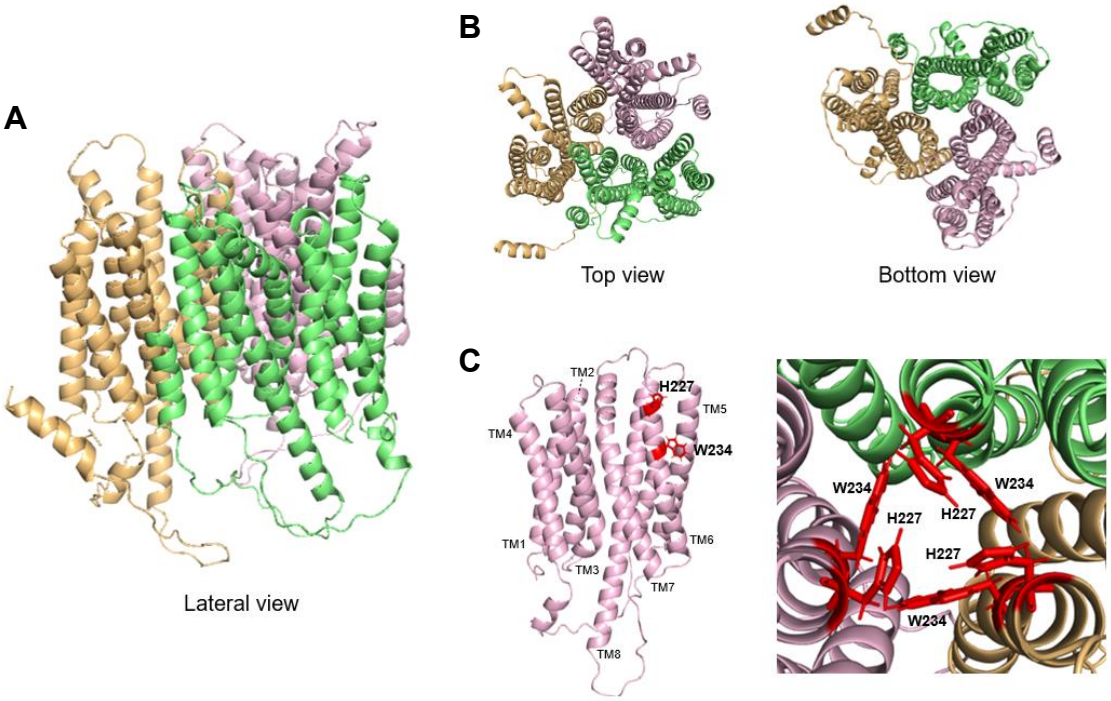


Figure 121. Trimeric MLC1 structure. **A.** Trimeric structure of MLC1. Monomers are painted in wheat, green and light purple. **B.** Top (left) and bottom (right) view of MLC1 trimer. **C.** MLC1 monomer with the pore residues marked in red (left). Top view of the pore formed by MLC1 trimer (right).

Then, to further study the pore, both histidine and tryptophan were mutated. We generated the variants H227A, H227R and W234A. Firstly, its expression and location were tested in transfected cell lines. WB and immunocytochemistry are shown in Figure 122, where it can be seen that the mutants are expressed and have the same location as the WT MLC1.

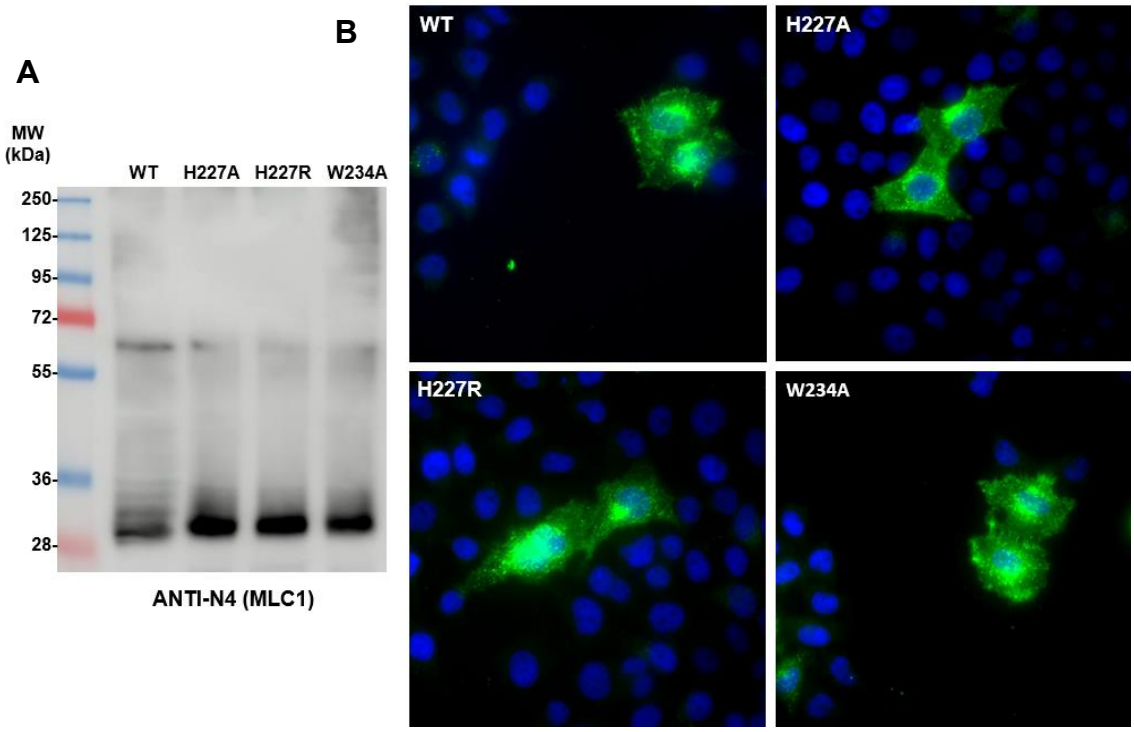


Figure 122. MLC1 pore mutants. A. WB of the generated MLC1 pore mutants H227A, H227R and W234A. **B.** Immunocytochemistry of MLC1 WT and pore mutants. HeLa cells were transfected with MLC1, after 48 hours, they were fixed and incubated with antibody N4 and Alexa 488 anti-rabbit afterwards.

3.7.4 PELE

Following the hypothesis that MLC1 could be implicated in ion transport, either by being a channel or a sensor, Protein Energy Landscape Exploration (PELE) assays were performed. PELE studies were performed by our collaborator Dr. Víctor Guallar at the Barcelona Supercomputing Centre. They studied the possible binding of chloride to MLC1 protein.

PELE consists of the study of substrates and how they can bind to the protein. It combines a Monte Carlo stochastic approach with protein structure predictions. Briefly, PELE involves the consecutive iteration of the following procedure. After an energy calculation for the initial structure, the assay begins with the generation of a perturbation in the system, for example a rotation of the ligand. Next, the energy of all side chains is computed before and after each perturbation and then the minimal interaction energy is identified as the most probable one (Madadkar-Sobhani and Guallar, 2013).

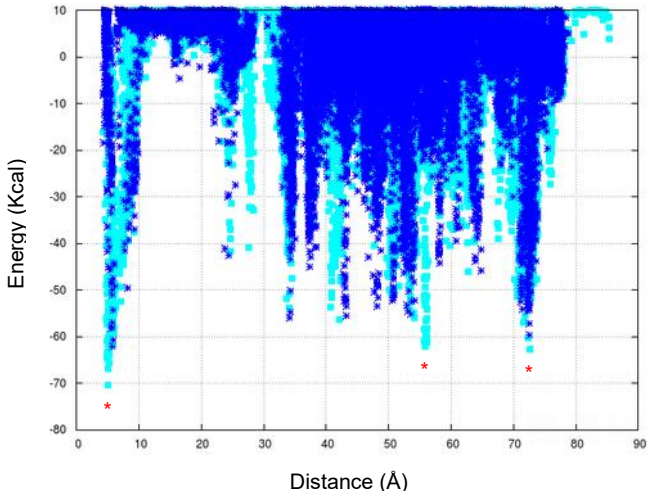


Figure 123. PELE energy profile for negative ions. Two different simulations are represented in cyan and blue showing similar energetic minima. In red stars, the mínima that were analysed.

PELE analysis was performed for positive ions, which did not retrieve any results, and for negative ions. In this last case, the energetic profile revealed some energetic minimums at different positions (Figure 123). It was decided to further study three that had the lowest energetic minima. The three minima studied for negative ions were two intracellular poses and one extracellular for the MLC1 homotrimer. The first intracellular pose (-70 Kcal, 5Å) showed chloride in the N-terminus of MLC1, before the TM1. In this location, Ser48 might bind to the chloride ion through a hydrogen bond, creating a binding pocket with Cys46 and Phe47 (Figure 124A). For the second intracellular pose (-62 Kcal, 55Å), the same Ser48 might form a hydrogen bond with the chloride ion. In this case, a binding site is created formed by Phe47 and Ser48 (Figure 124B). For the extracellular pose (-62 Kcal, 72Å), chloride was found in the second extracellular loop (between the TM3 and 4). In this position, a binding site formed by Asn136, Ser138 and Ala139 is created where chloride forms a hydrogen bond with the serine residue (Figure 124C).

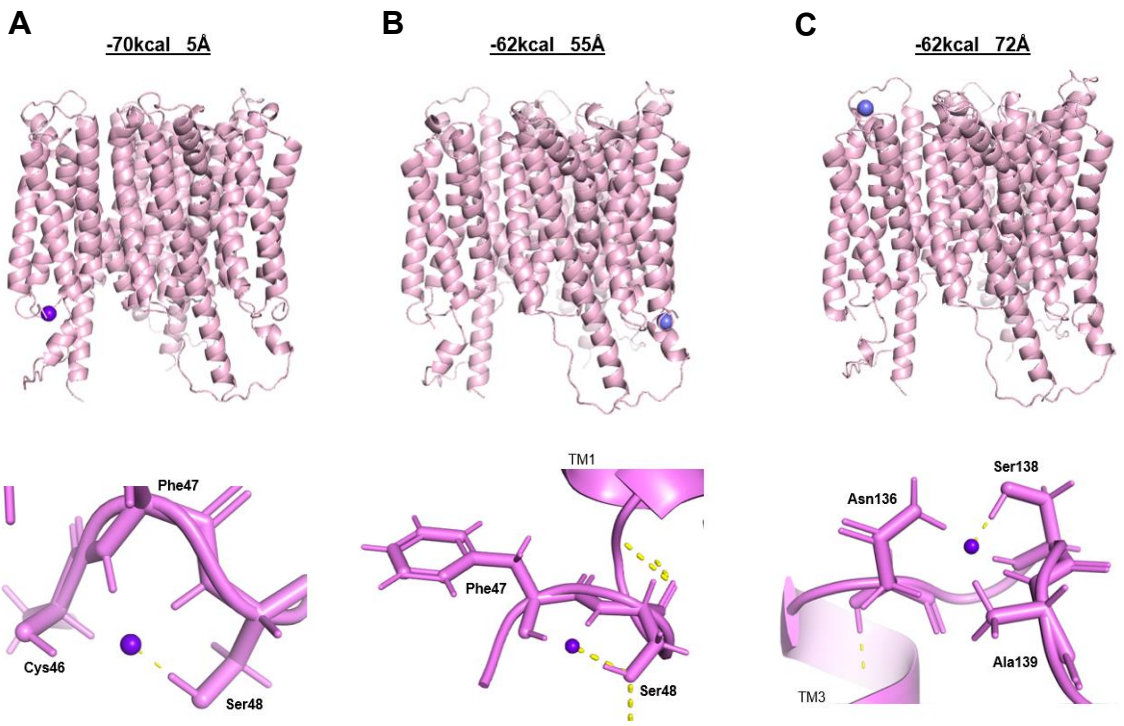


Figure 124. PELE simulations with chloride. A. Intracellular pose for chloride in MLC1 (-70 Kcal, 5Å). **B.** Intracellular pose for chloride in MLC1 (-62 Kcal, 55Å). **C.** Extracellular pose for chloride in MLC1. Chloride ion is represented in purple. Hydrogen bonds are represented in yellow lines.

How chloride binds to proteins is poorly understood, however, chloride-binding sites can be classified into two structural categories. The first one is typified by the ClC family of chloride channels, which binds chloride through backbone-amide and hydrophobic interactions. Other members of this category are the atrial naturetic peptide receptor (ANPR) or the lysine deficient protein kinase 1 (WNK1). The second class binds chloride through hydrophobic and positively charged amino acids. Some examples of proteins with these features are the serotonin transporter (SERT) and the angiotensin converting enzyme (ACE) (Piala et al., 2014). Analysing the putative binding sites for MLC1, they do not resemble any of the two structural classes described. Finding an energetic minimum does not mean that it is a binding site, but it is a site where the ion can possibly bind. In all three cases, a serine residue might form a hydrogen bound with the ion. Interestingly, the residues constituting the putative chloride binding sites in MLC1 are highly conserved across species.

DISCUSSION

Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare type of leukodystrophy caused by mutations in either *MLC1*, *GLIALCAM*, *GPRC5B* or *AQP4* genes (Leegwater et al., 2002, López-Hernández et al., 2011a, Passchier et al., 2023). Although the disease was characterized more than 20 years ago, little is known about its pathophysiological mechanisms. To date, the exact physiological functions of MLC1, GlialCAM and GPRC5B are not fully understood.

MLC1/GlialCAM complex might regulate the activity of different transporters and ion channels through phosphorylation of ERK signalling transduction cascade (Elorza-Vidal et al., 2018). In the absence of MLC1/GlialCAM proteins, the activity of ClC-2 is decreased during the hyperpolarization phase of a neuronal action potential (Hoegg-Beiler et al., 2014). VRAC activity is also diminished, which leads to an impaired regulatory volume decrease (RVD) response (Elorza-Vidal et al., 2018). Cx43 localization is affected as well in the absence of MLC1/GlialCAM as the protein is internalized and no longer located at cell-cell junctions (Baldwin et al., 2021) and Na^+/K^+ ATPase activity is altered (Brignone et al., 2011). In Figure 125, the main physiological processes where MLC1/GlialCAM complex plays a role are represented.

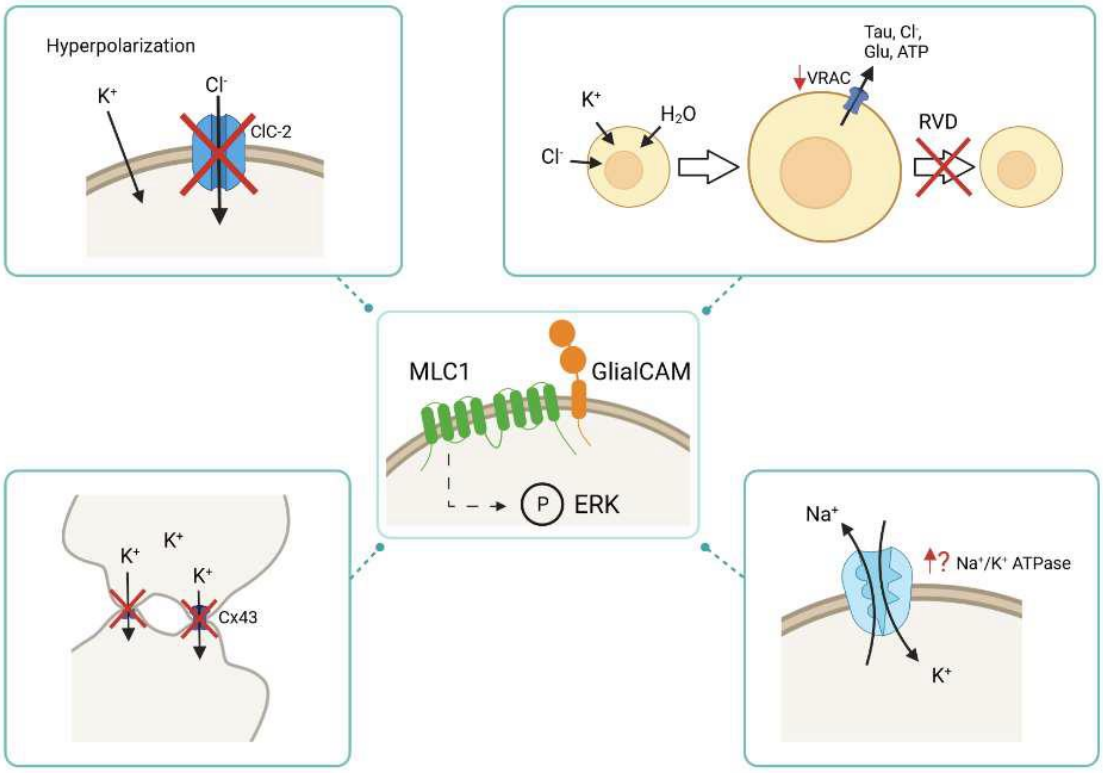


Figure 125. Physiological processes involving the activity of GlialCAM/MLC1 complex. Extracted from (Pla-Casillanis et al., 2022).

1. GPRC5B SIGNALLING IN MLC

The results obtained in this thesis have increased the knowledge regarding MLC pathophysiology. For the first time, we have established MLC1 as a negative modulator for GPRC5B signalling pathways being the first modulator known for this receptor. This inhibition occurs in the interaction of GPRC5B with different downstream proteins, such as Gα12 and β-arrestin.

Considering previous results from the group, GPRC5B depletion resulted in reduced GlialCAM levels at the plasma membrane in *Mlc1* KO astrocytes. This finding suggested that GPRC5B may stabilize GlialCAM at the plasma membrane (Alonso-Gardón et al., 2021). Contrarily, it has been reported in this thesis that GPRC5B internalizes GlialCAM in HeLa cells. In parallel, we have described that GPRC5B multimerization would be needed for its signalling transduction and MLC1 would inhibit it as well. So, we suggest that once the GPRC5B multimerization has occurred, GlialCAM would participate in GPRC5B endocytosis and then, the signalling would occur intracellularly.

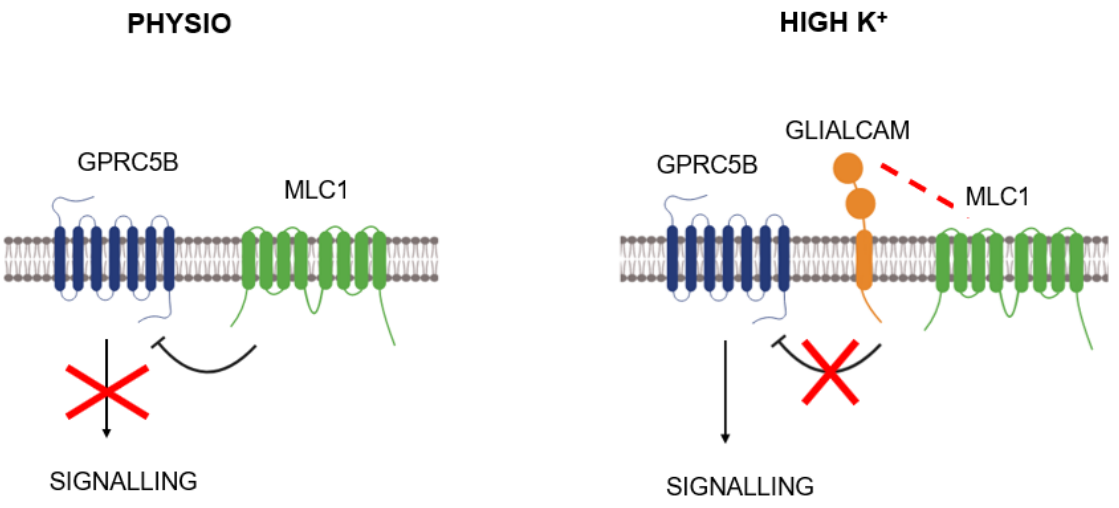


Figure 126. Model of GPRC5B interaction with MLC1 and GlialCAM in physiological and depolarizing conditions.

We propose that in physiological conditions, MLC1 would be inhibiting GPRC5B internalization. Hence, the GPCR would not interact with its downstream effectors and the signalling pathway would not be initiated. Yet, in depolarizing conditions where the potassium concentration is found increased, GlialCAM would interact with MLC1 as it has been observed by PLA experiments. Then, GPRC5B could be internalized and it could signal through different effectors, such as Gα12 or β-arrestin (Figure 126).

In the case of IgC2 GlialCAM variants, its pathogenicity could be explained by their inability to be internalized. Their higher stabilization at the plasma membrane would not allow for GPRC5B internalization, therefore, the signalling would not be possible. Regarding the GPRC5B patient-derived variants, the outcome is the same. In this situation, GPRC5B and GlialCAM co-localize more at the plasma membrane, so the internalization does not occur.

Concerning GPRC5B patient-derived mutants, here we report that their signal transduction might be slightly increased compared to the WT protein. This could be explained by the increased interaction of GPRC5B with GlialCAM. However, we expected that a higher interaction would lead to a higher internalization rate of GlialCAM and what we observed was that both proteins tend to remain at cell junctions and GlialCAM is not internalized at the same rate that in presence of GPRC5B WT. This could be due to the increased interaction of GPRC5B mutant with MLC1 as well.

Thus, seeing that both GlialCAM and GPRC5B mutants can lead to the same outcome, we conclude that for the signalling cascade to occur GlialCAM should not stay at cell junctions. It exists a correlation between protein localization and signal alteration, where the localization of both GlialCAM and GPRC5B needs to be at cell junctions to later be internalized.

Whether the signalling occurs intracellularly and not in the plasma membrane, there should be a regulation to finish it. Considering that MCL1 function is unknown, but it is hypothesized to be related to ion channels or transporters, we propose that MLC1 could act as a proton transporter. Then, in the internalized vesicles, MLC1 could alter the pH and based on that GPRC5B could be recycled to the plasma membrane again or go to the endosomes. In Figure 127, a scheme of the proposed model is displayed.

In conclusion, we believe that the role of GPRC5B is central in MLC disease and that its activity would be regulated by MLC1 and GlialCAM. Our data indicates that GPRC5B, MLC1 and GlialCAM interactions need to be dynamic and they are subjected to different stimuli. Further studies are needed to uncover the regulation of the GPRC5B signalling pathways.

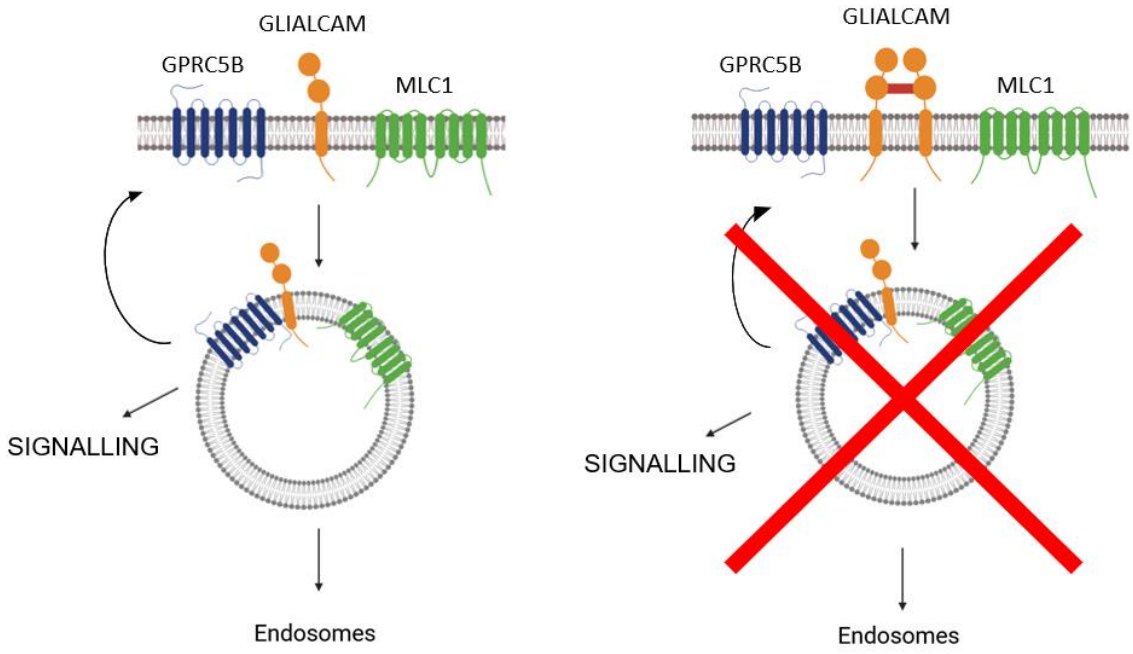


Figure 127. Model of GPRC5B signalling. MLC-causing GlialCAM mutations encoding for residues located at IgC2 domain abnormally stabilize GlialCAM oligomeric structures. Our proposed model is that GPRC5B interacts with GlialCAM at the plasma membrane of astrocytes, leading to endocytosis (left). GlialCAM mutant variants would be abnormally stabilized at the plasma membrane preventing the dissociation of the oligomeric structure. Thus, no GlialCAM would be released to interact and bind GPRC5B. hence, potential downstream signalling of this complex regulation homeostatic processes would not occur, leading to disease.

2. MLC1 STRUCTURE

In regards of MLC1 structure determination, although we have not obtained a high structural resolution, significant advances have been made and we have achieved a resolution around 7 Å in homo-trimeric MLC1. Once the first cryo-EM analyses revealed that particle orientation was challenging, we followed different approaches trying to overcome this problem. Seeing that GlialCAM, which could have helped orientate the particles, was only present in around 10% of them, we first attempted to find a way to bind MLC1 with GlialCAM. For this, we explored NanoBiT as a tethering strategy. A published study showed the structural determination of the vasoactive intestinal polypeptide receptor (VIP1R) using this approach. VIP1R is a widely expressed class B GPCR that was genetically fused to LgBiT and co-expressed with Gs heterotrimer fused to SmBiT. The reconstitution of the luciferase brought the two proteins in close proximity and it improved the stability and homogeneity of the complex (Duan et al., 2020).

Following this approach, we cloned the LgBiT and SmBiT to TwinStrep-MLC1 and GlialCAM-3xflag, respectively. Unfortunately, the addition of a luciferase fragment to MLC1 did not enable protein expression.

Then, we attempted a fusion protein strategy using BRIL. BRIL has been found useful for structure determination of different proteins, especially for GPCRs. Thus, BRIL was used for the structure determination of the human frizzled5 receptor (Fzd5) (Tsutsumi et al., 2020) and the adenosine A_{2A} receptor (Zhang et al., 2022). By binding BRIL to GPCRs through two continuous α-helices, the resulting chimeric receptors should maintain a rigid structure. Even though the BRIL approach has been used for multiple proteins, for MLC1 the strategy was unsuccessful due to the mobility of the BRIL element. Mobility has become the main limitation of this system. For instance, in the case of Fzd5, the acquired structure had a resolution of 3.7 Å. The obtention of a higher resolution was hampered by the mobility of the BRIL (Tsutsumi et al., 2020). To overcome this issue, multiple constructs should be tested, however, the design principles required for such successful reconstructions remain unclear. In other cases, the addition of BRIL made the protein unstable. As a solution, a new version of BRIL was engineered that replaced the original loop with a short GS linker. The modified BRIL was named mBRIL and it is less flexible. Its development enabled the structural identification of β2-adrenergic receptor (Guo et al., 2024).

Lastly, the obtention of nanobodies that bind specifically to MLC1 was a promising alternative. Nanobodies, the variable domains of heavy-chain only antibodies, recognize conformational epitopes and they are becoming popular as versatile binding partners of target proteins. In this thesis, the immunization of a llama with MLC1 reconstituted in liposomes and posterior phage display panning resulted in the identification of 28 different nanobodies that specifically bind to our target protein. It is worth mentioning that GlialCAM might be present in the reconstituted proteoliposomes, so it is possible that some of the obtained nanobodies are actually recognizing GlialCAM, instead of MLC1.

Nanobodies have been extensively used to determine X-ray structures, but their small size limits their cryo-EM application. Nevertheless, this problem can be overcome increasing their size with the attachment of a large scaffold, creating megabodies or legobodies. Megabodies are built by grafting single-domain antigen-binding constructs, such as nanobodies, onto larger scaffold proteins to produce stable, rigid, and efficiently folding monomeric chimeras. By increasing the target protein size, they overcome one of the major limitations in cryo-EM. Moreover, they also help randomize the orientation of particles (Uchański et al., 2021). The first megabodies were used to solve the high

resolution cryo-EM structure of the human synaptic $GABA_AR$ bound to small-molecule modulators (Laverty et al., 2019). Later, megabodies targeting homopentameric GABA receptors were further improved to increase the bulkiness of samples and reduce the biased distribution of orientation of membrane proteins (Uchański et al., 2021).

One limitation of using megabodies for particle alignment is that the β-strands linker between the nanobody and the scaffold can generate some flexibility. For this, another method for allowing cryo-EM analysis of small proteins was described: legobodies. Legobodies consist of the target-specific nanobody, a nanobody-binding Fab, and the maltose-binding protein (MBP)-based fusion protein. With this, the target protein increases its size by adding approximately 120 kDa. Legobodies have been used to determine the structure of KDELR protein and the receptor binding domain (RBD) of SARS-COV2 spike protein (Wu and Rapoport, 2021).

2.1 INTERACTION MODELS FOR MLC PROTEINS

As already mentioned, the appearance of AlphaFold AI has become very helpful during this project. The development of AF3 that enables structure determination of protein-protein interactions has been used for the analysis of the interaction of MLC1, GlialCAM and GPRC5B with each other.

Firstly, the interaction between MLC1 and GlialCAM was assessed using AF3. It was observed that the IgC2 domain of GlialCAM was predicted to be close of MLC1 extracellular part. Analysing the amino acid residues, two MLC1 residues were identified to form a putative hydrogen bond with two GlialCAM residues. Both MLC1 residues (V75 and E79) are located between the first and second transmembrane domain and they are predicted to interact with S154 and S161 GlialCAM residues, respectively (Figure 128). Remarkably, an MLC patient has been described to carry the mutation E79K in *MLC1* (Yüzbaşioğlu et al., 2011). Although there is few information on this MLC-causing mutant, we hypothesize that its pathogenicity would come from an altered interaction between MLC1 and GlialCAM. The substitution of a negatively charged amino acid residue by a positively charged amino acid might disrupt the bound with the GlialCAM serine residue. The residue E79 is located at the beginning of TM2, that is one of the TM domains located at the interphase of MLC1 halves. Although it could be reasoned that this residue would be important for the interactions of the two halves of the protein, based on the AF predictions, E79 is not facing this interphase, suggesting that it might interact with other residues such as GlialCAM IgC2 domain.

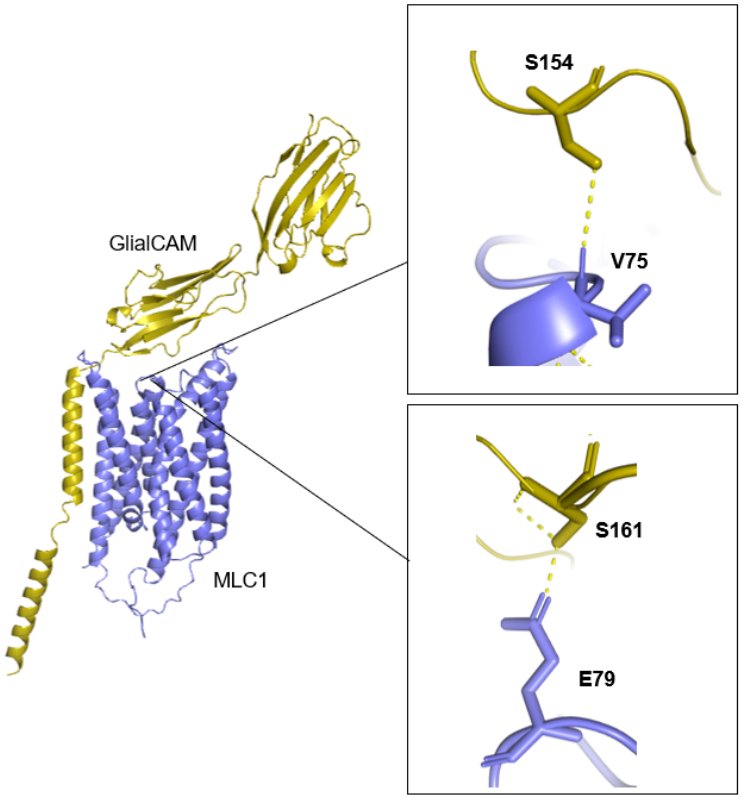


Figure 128. Interaction model for GlialCAM and MLC1. AF3 prediction model for the interaction between MLC1(slate blue) and GlialCAM (olive).

Regarding MLC1 and GPRC5B interaction, the analysis of the AF3 predicted model did not reveal any putative polar contacts between residues from the transmembrane domains of both proteins. However, it could be noted that the residues I176 and A177 from GPRC5B were in close proximity with TM4 of MLC1. The named residues are found duplicated in the MLC patients harbouring mutations in *GPRC5B* (Figure 129). This observation would suggest that these mutations would cause a shift in the following residues that would alter its location in the α-helix and, consequently, it might alter the interaction with MLC1. Accordingly, in this thesis we have determined that both variants present an altered interaction with MLC proteins, as they interact significantly more with MLC1 and GlialCAM. Although this increased interaction could be explained by the fact that the GPRC5B variants are more stable at the plasma membrane.

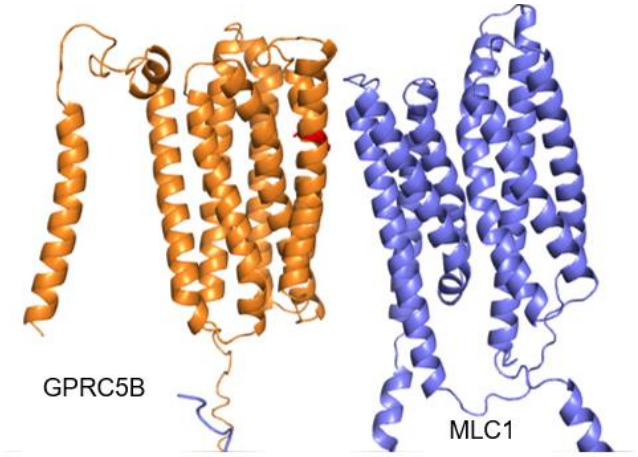


Figure 129. Interaction model for GPRC5B and MLC1. AF3 prediction model for the interaction between GPRC5B (orange) and MLC1 (slate blue). GPRC5B patient-derived residues (I176 and A177) are marked in red.

Finally, the interaction between GlialCAM and GPRC5B was also examined. Similarly to the interaction of GPRC5B with MLC1, the GPRC5B patient-derived mutations present a significantly increased interaction with GlialCAM. The GPRC5B TM4, where the patient-derived mutations are located, is found to be in close proximity with the GlialCAM TM domain. Using AF3, a prediction model was generated with GPRC5B WT and another with the patient-derived duplication I176dup. Besides the expected alteration in the TM4 residue location by the addition of an extra isoleucine residue, it was observed that the position of all GPRC5B TM domains was shifted as well as the position of GlialCAM (Figure 130). This shift implies that the residues that could form polar contacts with GlialCAM residues might now be too far to interact, and the other way round, residues that were not close enough to bind GlialCAM residues, with the duplication, might now be able to interact.

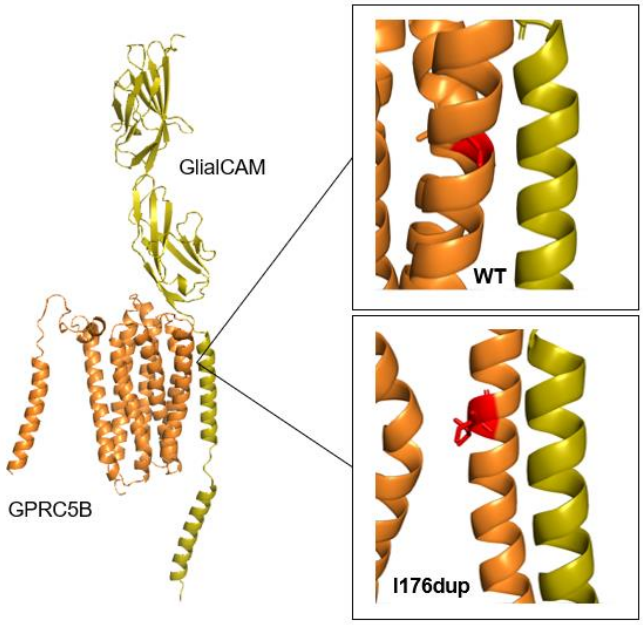


Figure 130. Interaction model for GPRC5B and GlialCAM. AF3 prediction model for the interaction between GPRC5B (orange) and GlialCAM (olive). GPRC5B patient-derived residue duplication I176dup is marked in red. A close-up of the putative interaction between GPRC5B and GlialCAM through their TM domains is represented by showing the GPRC5B WT (top right) and the GPRC5B I176dup (bottom right).

3. TOWARDS A TREATMENT FOR MLC

To date, no effective therapy is available for MLC patients. Based on the clinical phenotype, the disease can be classified into two groups: the classic phenotype, in which patients show a deterioration of the symptomatology, and the remitting phenotype, in which patients present an improvement. During this thesis, we reported a new MLC patient harbouring two variants in *MLC1* showing radiological improvement (Mayayo-Vallverdú et al., 2023). The patient was heterozygous for the variants c.597+37C>G and c.895-1G>T both affecting mRNA splicing, and the latest causing the deletion of p.Pro299_Glu353del.

We demonstrated that the patient had a small amount expression of WT *MLC1* mRNA and protein, suggesting that few quantities of MLC1 protein would be sufficient to improve the clinical phenotype. The molecular basis of this reversibility could be explained by the fact that MLC1 protein expression in humans reaches its high at one year of age. Then, it progressively decreases until 5 years of age, where it stabilized (Dubey et al., 2015). Therefore, we hypothesize that a minimal amount of protein might be enough after 5 years of age, but not at earlier stages. Similarly, in *Mlc1* KO mice

(Hoegg-Beiler et al., 2014), the re-expression of minimal amounts of MLC1, even in adult animals, can abolish the vacuolation observed in the KO (Sánchez et al., 2020).

These results open the door to the development of therapies aimed at administering low doses of MLC1. In this sense, gene therapy could be an option for MLC patients (Bosch and Estévez, 2020). The use of antisense oligonucleotides or RNA editing could be beneficial to ameliorate the symptoms observed in patients with MLC splicing variants (Booth et al., 2023, Hill and Meisler, 2021). Moreover, it was previously demonstrated that the use of an antisense morpholinated oligonucleotide restored in vitro normal splicing of one MLC variant (Mancini et al., 2012).

Regarding the identification of *GPRC5B* as the third gene causing MLC and after the studies that showed that MLC1 and GlialCAM proteins might regulate GPRC5B activity, we believe that the regulation of GPRC5B signalling might be crucial in the pathogenesis of MLC. Thus, GPRC5B might be a proper therapeutic target to develop new therapies for the disease. However, since GPRC5B is expressed ubiquitously, therapeutic strategies should be directed to the brain to avoid significant side effects.

CONCLUSIONS

1. The characterization of an MLC patient carrying two splicing mutations in MLC1 and presenting a remitting phenotype concluded that the presence of residual amounts of MLC1 protein is sufficient to improve the clinical phenotype.
2. In the CNS, we propose that GPRC5B signalling is regulated negatively by MLC1 and positively by GlialCAM, forming an MLC signalling complex. Based on our results, we propose a mechanistic model of the MLC signalling complex, where depolarization causes GlialCAM/GPRC5B internalization, and signalling in intracellular compartments.
3. We have found that GPRC5B patient-derived MLC mutants are more stable and interact significantly more with both MLC1 and GlialCAM, and they are able to signal. However, we propose that its increased stability impedes its internalization by GlialCAM, leading to a defect in intracellular signalling.
4. We suggest that GlialCAM IgC2 domain forms lateral interactions with adjacent GlialCAM molecules. We propose that the molecular mechanism of pathogenesis of GlialCAM containing MLC mutations in its IgC2 domain is that these mutations provoke a gain of stability. This gain leads to a lack of internalization by GPRC5B resulting in an impaired intracellular signalling.
5. We have improved the resolution of the structure of MLC1 by cryo-EM to around a 7 Å resolution.
6. We have been able to obtain nanobodies that are able to recognize the human MLC1 protein with low affinity. Further characterization of these nanobodies is needed for their use in different applications.

BIBLIOGRAPHY

AISSAOUI, N., LAI-KEE-HIM, J., MILLS, A., DECLERCK, N., MORICHAUD, Z., BRODOLIN, K., BACONNAIS, S., LE CAM, E., CHARBONNIER, J. B., SOUNIER, R., GRANIER, S., ROPARS, V., BRON, P. & BELLOT, G. 2021. Modular Imaging Scaffold for Single-Particle Electron Microscopy. *ACS Nano,* 15**,** 4186-4196.

ALBAYRAK, Ö., PÜTTER, C., VOLCKMAR, A. L., CICHON, S., HOFFMANN, P., NÖTHEN, M. M., JÖCKEL, K. H., SCHREIBER, S., WICHMANN, H. E., FARAONE, S. V., NEALE, B. M., HERPERTZ-DAHLMANN, B., LEHMKUHL, G., SINZIG, J., RENNER, T. J., ROMANOS, M., WARNKE, A., LESCH, K. P., REIF, A., SCHIMMELMANN, B. G., SCHERAG, A., HEBEBRAND, J. & HINNEY, A. 2013. Common obesity risk alleles in childhood attention-deficit/hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet,* 162b**,** 295-305.

ALEKSASHIN, N. A., CHANG, S. T. & CATE, J. H. D. 2023. A highly efficient human cell-free translation system. *Rna,* 29**,** 1960-1972.

ALEXANDER, S. P., BENSON, H. E., FACCENDA, E., PAWSON, A. J., SHARMAN, J. L., SPEDDING, M., PETERS, J. A. & HARMAR, A. J. 2013. The Concise Guide to PHARMACOLOGY 2013/14: G protein-coupled receptors. *Br J Pharmacol,* 170**,** 1459-581.

ALONSO-GARDÓN, M., ELORZA-VIDAL, X., CASTELLANOS, A., LA SALA, G., ARMAND-UGON, M., GILBERT, A., DI PIETRO, C., PLA-CASILLANIS, A., CIRUELA, F., GASULL, X., NUNES, V., MARTÍNEZ, A., SCHULTE, U., COHEN-SALMON, M., MARAZZITI, D. & ESTÉVEZ, R. 2021. Identification of the GlialCAM interactome: the G protein-coupled receptors GPRC5B and GPR37L1 modulate megalencephalic leukoencephalopathy proteins. *Hum Mol Genet,* 30**,** 1649-1665.

ALWINE, J. C. 1985. Transient gene expression control: effects of transfected DNA stability and trans-activation by viral early proteins. *Mol Cell Biol,* 5**,** 1034-42.

AMBROSINI, E., SERAFINI, B., LANCIOTTI, A., TOSINI, F., SCIALPI, F., PSAILA, R., RAGGI, C., DI GIROLAMO, F., PETRUCCI, T. C. & ALOISI, F. 2008. Biochemical characterization of MLC1 protein in astrocytes and its association with the dystrophin-glycoprotein complex. *Mol Cell Neurosci,* 37**,** 480-93.

ARNEDO, T., AIELLO, C., JEWORUTZKI, E., DENTICI, M. L., UZIEL, G., SIMONATI, A., PUSCH, M., BERTINI, E. & ESTÉVEZ, R. 2014a. Expanding the spectrum of megalencephalic leukoencephalopathy with subcortical cysts in two patients with GLIALCAM mutations. *Neurogenetics,* 15**,** 41-8.

ARNEDO, T., LÓPEZ-HERNÁNDEZ, T., JEWORUTZKI, E., CAPDEVILA-NORTES, X., SIRISI, S., PUSCH, M. & ESTÉVEZ, R. 2014b. Functional analyses of mutations in HEPACAM causing megalencephalic leukoencephalopathy. *Hum Mutat,* 35**,** 1175-8.

ATANES, P., RUZ-MALDONADO, I., HAWKES, R., LIU, B., PERSAUD, S. J. & AMISTEN, S. 2018. Identifying Signalling Pathways Regulated by GPRC5B in β-Cells by CRISPR-Cas9-Mediated Genome Editing. *Cell Physiol Biochem,* 45**,** 656-666.

AUTZEN, H. E., JULIUS, D. & CHENG, Y. 2019. Membrane mimetic systems in CryoEM: keeping membrane proteins in their native environment. *Curr Opin Struct Biol,* 58**,** 259-268.

BADA JUAREZ, J. F., MUÑOZ-GARCÍA, J. C., INÁCIO DOS REIS, R., HENRY, A., MCMILLAN, D., KRIEK, M., WOOD, M., VANDENPLAS, C., SANDS, Z., CASTRO, L., TAYLOR, R. & WATTS, A. 2020. Detergent-free extraction of a functional low-expressing GPCR from a human cell line. *Biochim Biophys Acta Biomembr,* 1862**,** 183152.

BAGHBAN, R., FARAJNIA, S., RAJABIBAZL, M., GHASEMI, Y., MAFI, A., HOSEINPOOR, R., RAHBARNIA, L. & ARIA, M. 2019. Yeast Expression Systems: Overview and Recent Advances. *Mol Biotechnol,* 61**,** 365-384.

BAI, X. C., MCMULLAN, G. & SCHERES, S. H. 2015. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci,* 40**,** 49-57.

BALDWIN, K. T., TAN, C. X., STRADER, S. T., JIANG, C., SAVAGE, J. T., ELORZA-VIDAL, X., CONTRERAS, X., RÜLICKE, T., HIPPENMEYER, S., ESTÉVEZ, R., JI, R. R. & EROGLU, C. 2021. HepaCAM controls astrocyte self-organization and coupling. *Neuron,* 109**,** 2427-2442.e10.

BANG, S., JIANG, C., XU, J., CHANDRA, S., MCGINNIS, A., LUO, X., HE, Q., LI, Y., WANG, Z., AO, X., PARISIEN, M., FERNANDES DE ARAUJO, L. O., JAHANGIRI ESFAHANI, S., ZHANG, Q., TONELLO, R., BERTA, T., DIATCHENKO, L. & JI, R. R. 2024. Satellite glial GPR37L1 and its ligand maresin 1 regulate potassium channel signaling and pain homeostasis. *J Clin Invest,* 134.

BANG, S., XIE, Y. K., ZHANG, Z. J., WANG, Z., XU, Z. Z. & JI, R. R. 2018. GPR37 regulates macrophage phagocytosis and resolution of inflammatory pain. *J Clin Invest,* 128**,** 3568-3582.

BARRALLO-GIMENO, A., GRADOGNA, A., ZANARDI, I., PUSCH, M. & ESTÉVEZ, R. 2015. Regulatory-auxiliary subunits of CLC chloride channel-transport proteins. *J Physiol,* 593**,** 4111-27.

BARREIRO, O., ZAMAI, M., YÁÑEZ-MÓ, M., TEJERA, E., LÓPEZ-ROMERO, P., MONK, P. N., GRATTON, E., CAIOLFA, V. R. & SÁNCHEZ-MADRID, F. 2008. Endothelial adhesion receptors are recruited to adherent leukocytes by inclusion in preformed tetraspanin nanoplatforms. *J Cell Biol,* 183**,** 527-42.

BAYBURT, T. H. & SLIGAR, S. G. 2010. Membrane protein assembly into Nanodiscs. *FEBS Lett,* 584**,** 1721-7.

BAZAN, J., CAŁKOSIŃSKI, I. & GAMIAN, A. 2012. Phage display--a powerful technique for immunotherapy: 1. Introduction and potential of therapeutic applications. *Hum Vaccin Immunother,* 8**,** 1817-28.

BEN-ZEEV, B., LEVY-NISSENBAUM, E., LAHAT, H., ANIKSTER, Y., SHINAR, Y., BRAND, N., GROSS-TZUR, V., MACGREGOR, D., SIDI, R., KLETA, R., FRYDMAN, M. & PRAS, E. 2002. Megalencephalic leukoencephalopathy with subcortical cysts; a founder effect in Israeli patients and a higher than expected carrier rate among Libyan Jews. *Hum Genet,* 111**,** 214-8.

BLATTNER, R., VON MOERS, A., LEEGWATER, P. A., HANEFELD, F. A., VAN DER KNAAP, M. S. & KÖHLER, W. 2003. Clinical and genetic heterogeneity in megalencephalic leukoencephalopathy with subcortical cysts (MLC). *Neuropediatrics,* 34**,** 215-8.

BLOCH, J. S., PESCIULLESI, G., BOILEVIN, J., NOSOL, K., IROBALIEVA, R. N., DARBRE, T., AEBI, M., KOSSIAKOFF, A. A., REYMOND, J. L. & LOCHER, K. P. 2020. Structure and mechanism of the ER-based glucosyltransferase ALG6. *Nature,* 579**,** 443-447.

BOOR, P. K., DE GROOT, K., WAISFISZ, Q., KAMPHORST, W., OUDEJANS, C. B., POWERS, J. M., PRONK, J. C., SCHEPER, G. C. & VAN DER KNAAP, M. S. 2005. MLC1: a novel protein in distal astroglial processes. *J Neuropathol Exp Neurol,* 64**,** 412-9.

BOOTH, B. J., NOURREDDINE, S., KATREKAR, D., SAVVA, Y., BOSE, D., LONG, T. J., HUSS, D. J. & MALI, P. 2023. RNA editing: Expanding the potential of RNA therapeutics. *Mol Ther,* 31**,** 1533-1549.

BOSCH, A. & ESTÉVEZ, R. 2020. Megalencephalic Leukoencephalopathy: Insights Into Pathophysiology and Perspectives for Therapy. *Front Cell Neurosci,* 14**,** 627887.

BRIGNONE, M. S., LANCIOTTI, A., CAMERINI, S., DE NUCCIO, C., PETRUCCI, T. C., VISENTIN, S. & AMBROSINI, E. 2015. MLC1 protein: a likely link between leukodystrophies and brain channelopathies. *Front Cell Neurosci,* 9**,** 66.

BRIGNONE, M. S., LANCIOTTI, A., MACIOCE, P., MACCHIA, G., GAETANI, M., ALOISI, F., PETRUCCI, T. C. & AMBROSINI, E. 2011. The beta1 subunit of the Na,K-ATPase pump interacts with megalencephalic leucoencephalopathy with

subcortical cysts protein 1 (MLC1) in brain astrocytes: new insights into MLC pathogenesis. *Hum Mol Genet,* 20**,** 90-103.

BROCKMANN, K., FINSTERBUSCH, J., TERWEY, B., FRAHM, J. & HANEFELD, F. 2003. Megalencephalic leukoencephalopathy with subcortical cysts in an adult: quantitative proton MR spectroscopy and diffusion tensor MRI. *Neuroradiology,* 45**,** 137-42.

BRÄUNER-OSBORNE, H., WELLENDORPH, P. & JENSEN, A. A. 2007. Structure, pharmacology and therapeutic prospects of family C G-protein coupled receptors. *Curr Drug Targets,* 8**,** 169-84.

BUGIANI, M., DUBEY, M., BREUR, M., POSTMA, N. L., DEKKER, M. P., TER BRAAK, T., BOSCHERT, U., ABBINK, T. E. M., MANSVELDER, H. D., MIN, R., VAN WEERING, J. R. T. & VAN DER KNAAP, M. S. 2017. Megalencephalic leukoencephalopathy with cysts: the Glialcam-null mouse model. *Ann Clin Transl Neurol,* 4**,** 450-465.

CABRAL, A., CABRAL, J. E. & MCNULTY, R. 2022. Cryo-EM for Small Molecules. *Curr Protoc,* 2**,** e632.

CAPDEVILA-NORTES, X., JEWORUTZKI, E., ELORZA-VIDAL, X., BARRALLO-GIMENO, A., PUSCH, M. & ESTÉVEZ, R. 2015. Structural determinants of interaction, trafficking and function in the ClC-2/MLC1 subunit GlialCAM involved in leukodystrophy. *J Physiol,* 593**,** 4165-80.

CAPDEVILA-NORTES, X., LÓPEZ-HERNÁNDEZ, T., APAJA, P. M., LÓPEZ DE HEREDIA, M., SIRISI, S., CALLEJO, G., ARNEDO, T., NUNES, V., LUKACS, G. L., GASULL, X. & ESTÉVEZ, R. 2013. Insights into MLC pathogenesis: GlialCAM is an MLC1 chaperone required for proper activation of volume-regulated anion currents. *Hum Mol Genet,* 22**,** 4405-16.

CAPDEVILA-NORTES, X., LÓPEZ-HERNÁNDEZ, T., CIRUELA, F. & ESTÉVEZ, R. 2012. A modification of the split-tobacco etch virus method for monitoring interactions between membrane proteins in mammalian cells. *Anal Biochem,* 423**,** 109-18.

CARPENTER, B. & TATE, C. G. 2016. Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation. *Protein Eng Des Sel,* 29**,** 583-594.

CARPENTER, B. & TATE, C. G. 2017. Active state structures of G protein-coupled receptors highlight the similarities and differences in the G protein and arrestin coupling interfaces. *Curr Opin Struct Biol,* 45**,** 124-132.

CHANG, S. S., DIJKMAN, P. M., WIESSING, S. A. & KUDRYASHEV, M. 2023. Determining the structure of the bacterial voltage-gated sodium channel NaChBac embedded in liposomes by cryo electron tomography and subtomogram averaging. *Sci Rep,* 13**,** 11523.

CHENG, Y. 2018. Single-particle cryo-EM-How did it get here and where will it go. *Science,* 361**,** 876-880.

CHENG, Y. & LOTAN, R. 1998. Molecular cloning and characterization of a novel retinoic acid-inducible gene that encodes a putative G protein-coupled receptor. *J Biol Chem,* 273**,** 35008-15.

CHENG, Z., GARVIN, D., PAGUIO, A., STECHA, P., WOOD, K. & FAN, F. 2010. Luciferase Reporter Assay System for Deciphering GPCR Pathways. *Curr Chem Genomics,* 4**,** 84-91.

CHUN, L., ZHANG, W. H. & LIU, J. F. 2012. Structure and ligand recognition of class C GPCRs. *Acta Pharmacol Sin,* 33**,** 312-23.

CIANCAGLINI, P., SIMÃO, A. M. S., BOLEAN, M., MILLÁN, J. L., RIGOS, C. F., YONEDA, J. S., COLHONE, M. C. & STABELI, R. G. 2012. Proteoliposomes in nanobiotechnology. *Biophys Rev,* 4**,** 67-81.

CLAING, A., LAPORTE, S. A., CARON, M. G. & LEFKOWITZ, R. J. 2002. Endocytosis of G protein-coupled receptors: roles of G protein-coupled receptor kinases and beta-arrestin proteins. *Prog Neurobiol,* 66**,** 61-79.

CONDEMINAS, M. & MACIAS, M. J. 2024. Overcoming challenges in structural biology with integrative approaches and nanobody-derived technologies. *Curr Opin Struct Biol,* 84**,** 102764.

DANG, Y., ZHOU, D., DU, X., ZHAO, H., LEE, C. H., YANG, J., WANG, Y., QIN, C., GUO, Z. & ZHANG, Z. 2022. Molecular mechanism of substrate recognition by folate transporter SLC19A1. *Cell Discov,* 8**,** 141.

DAVIES, M. N., SECKER, A., FREITAS, A. A., MENDAO, M., TIMMIS, J. & FLOWER, D. R. 2007. On the hierarchical classification of G protein-coupled receptors. *Bioinformatics,* 23**,** 3113-8.

DE CARLO, S. & HARRIS, J. R. 2011. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron,* 42**,** 117-31.

DENG, J., FUJIMOTO, J., YE, X. F., MEN, T. Y., VAN PELT, C. S., CHEN, Y. L., LIN, X. F., KADARA, H., TAO, Q., LOTAN, D. & LOTAN, R. 2010. Knockout of the tumor suppressor gene Gprc5a in mice leads to NF-kappaB activation in airway epithelium and promotes lung inflammation and tumorigenesis. *Cancer Prev Res (Phila),* 3**,** 424-37.

DENG, Z., ZHAO, Y., FENG, J., ZHANG, J., ZHAO, H., RAU, M. J., FITZPATRICK, J. A. J., HU, H. & YUAN, P. 2021. Cryo-EM structure of a proton-activated chloride channel TMEM206. *Sci Adv,* 7.

DEWIRE, S. M., AHN, S., LEFKOWITZ, R. J. & SHENOY, S. K. 2007. Beta-arrestins and cell signaling. *Annu Rev Physiol,* 69**,** 483-510.

DOMINGUES, H. S., PORTUGAL, C. C., SOCODATO, R. & RELVAS, J. B. 2016. Oligodendrocyte, Astrocyte, and Microglia Crosstalk in Myelin Development, Damage, and Repair. *Front Cell Dev Biol,* 4**,** 71.

DONOHUE, P. J., SHAPIRA, H., MANTEY, S. A., HAMPTON, L. L., JENSEN, R. T. & BATTEY, J. F. 1998. A human gene encodes a putative G protein-coupled receptor highly expressed in the central nervous system. *Brain Res Mol Brain Res,* 54**,** 152-60.

DUAN, J., SHEN, D. D., ZHOU, X. E., BI, P., LIU, Q. F., TAN, Y. X., ZHUANG, Y. W., ZHANG, H. B., XU, P. Y., HUANG, S. J., MA, S. S., HE, X. H., MELCHER, K., ZHANG, Y., XU, H. E. & JIANG, Y. 2020. Cryo-EM structure of an activated VIP1 receptor-G protein complex revealed by a NanoBiT tethering strategy. *Nat Commun,* 11**,** 4121.

DUARRI, A., LOPEZ DE HEREDIA, M., CAPDEVILA-NORTES, X., RIDDER, M. C., MONTOLIO, M., LÓPEZ-HERNÁNDEZ, T., BOOR, I., LIEN, C. F., HAGEMANN, T., MESSING, A., GORECKI, D. C., SCHEPER, G. C., MARTÍNEZ, A., NUNES, V., VAN DER KNAAP, M. S. & ESTÉVEZ, R. 2011. Knockdown of MLC1 in primary astrocytes causes cell vacuolation: a MLC disease cell model. *Neurobiol Dis,* 43**,** 228-38.

DUARRI, A., TEIJIDO, O., LÓPEZ-HERNÁNDEZ, T., SCHEPER, G. C., BARRIERE, H., BOOR, I., AGUADO, F., ZORZANO, A., PALACÍN, M., MARTÍNEZ, A., LUKACS, G. L., VAN DER KNAAP, M. S., NUNES, V. & ESTÉVEZ, R. 2008. Molecular pathogenesis of megalencephalic leukoencephalopathy with subcortical cysts: mutations in MLC1 cause folding defects. *Hum Mol Genet,* 17**,** 3728-39.

DUBEY, M., BROUWERS, E., HAMILTON, E. M. C., STIEDL, O., BUGIANI, M., KOCH, H., KOLE, M. H. P., BOSCHERT, U., WYKES, R. C., MANSVELDER, H. D., VAN DER KNAAP, M. S. & MIN, R. 2018. Seizures and disturbed brain potassium dynamics in the leukodystrophy megalencephalic leukoencephalopathy with subcortical cysts. *Ann Neurol,* 83**,** 636-649.

DUBEY, M., BUGIANI, M., RIDDER, M. C., POSTMA, N. L., BROUWERS, E., POLDER, E., JACOBS, J. G., BAAYEN, J. C., KLOOSTER, J., KAMERMANS, M., AARDSE, R., DE KOCK, C. P., DEKKER, M. P., VAN WEERING, J. R., HEINE, V. M., ABBINK, T. E., SCHEPER, G. C., BOOR, I., LODDER, J. C., MANSVELDER, H. D. & VAN DER KNAAP, M. S. 2015. Mice with

megalencephalic leukoencephalopathy with cysts: a developmental angle. *Ann Neurol,* 77**,** 114-31.

DÖRR, J. M., SCHEIDELAAR, S., KOORENGEVEL, M. C., DOMINGUEZ, J. J., SCHÄFER, M., VAN WALREE, C. A. & KILLIAN, J. A. 2016. The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J,* 45**,** 3-21.

EGLEN, R. M., BOSSE, R. & REISINE, T. 2007. Emerging concepts of guanine nucleotide-binding protein-coupled receptor (GPCR) function and implications for high throughput screening. *Assay Drug Dev Technol,* 5**,** 425-51.

ELORZA-VIDAL, X., SIRISI, S., GAITÁN-PEÑAS, H., PÉREZ-RIUS, C., ALONSO-GARDÓN, M., ARMAND-UGÓN, M., LANCIOTTI, A., BRIGNONE, M. S., PRAT, E., NUNES, V., AMBROSINI, E., GASULL, X. & ESTÉVEZ, R. 2018. GlialCAM/MLC1 modulates LRRC8/VRAC currents in an indirect manner: Implications for megalencephalic leukoencephalopathy. *Neurobiol Dis,* 119**,** 88-99.

ELORZA-VIDAL, X., XICOY-ESPAULELLA, E., PLA-CASILLANIS, A., ALONSO-GARDÓN, M., GAITÁN-PEÑAS, H., ENGEL-PIZCUETA, C., FERNÁNDEZ-RECIO, J. & ESTÉVEZ, R. 2020. Structural basis for the dominant or recessive character of GLIALCAM mutations found in leukodystrophies. *Hum Mol Genet,* 29**,** 1107-1120.

ERRASTI-MURUGARREN, E., FORT, J., BARTOCCIONI, P., DÍAZ, L., PARDON, E., CARPENA, X., ESPINO-GUARCH, M., ZORZANO, A., ZIEGLER, C., STEYAERT, J., FERNÁNDEZ-RECIO, J., FITA, I. & PALACÍN, M. 2019. L amino acid transporter structure and molecular bases for the asymmetry of substrate interaction. *Nat Commun,* 10**,** 1807.

ESTÉVEZ, R., ELORZA-VIDAL, X., GAITÁN-PEÑAS, H., PÉREZ-RIUS, C., ARMAND-UGÓN, M., ALONSO-GARDÓN, M., XICOY-ESPAULELLA, E., SIRISI, S., ARNEDO, T., CAPDEVILA-NORTES, X., LÓPEZ-HERNÁNDEZ, T., MONTOLIO, M., DUARRI, A., TEIJIDO, O., BARRALLO-GIMENO, A., PALACÍN, M. & NUNES, V. 2018. Megalencephalic leukoencephalopathy with subcortical cysts: A personal biochemical retrospective. *Eur J Med Genet,* 61**,** 50-60.

FAVRE-KONTULA, L., ROLLAND, A., BERNASCONI, L., KARMIRANTZOU, M., POWER, C., ANTONSSON, B. & BOSCHERT, U. 2008. GlialCAM, an immunoglobulin-like cell adhesion molecule is expressed in glial cells of the central nervous system. *Glia,* 56**,** 633-45.

FLÖTENMEYER, M., WEISS, H., TRIBET, C., POPOT, J. L. & LEONARD, K. 2007. The use of amphipathic polymers for cryo electron microscopy of NADH:ubiquinone oxidoreductase (complex I). *J Microsc,* 227**,** 229-35.

FONIN, A. V., DARLING, A. L., KUZNETSOVA, I. M., TUROVEROV, K. K. & UVERSKY, V. N. 2019. Multi-functionality of proteins involved in GPCR and G protein signaling: making sense of structure-function continuum with intrinsic disorder-based proteoforms. *Cell Mol Life Sci,* 76**,** 4461-4492.

FOORD, S. M., BONNER, T. I., NEUBIG, R. R., ROSSER, E. M., PIN, J. P., DAVENPORT, A. P., SPEDDING, M. & HARMAR, A. J. 2005. International Union of Pharmacology. XLVI. G protein-coupled receptor list. *Pharmacol Rev,* 57**,** 279-88.

FOORD, S. M., JUPE, S. & HOLBROOK, J. 2002. Bioinformatics and type II G-protein-coupled receptors. *Biochem Soc Trans,* 30**,** 473-9.

FORNWALD, J. A., LU, Q., BOYCE, F. M. & AMES, R. S. 2016. Gene Expression in Mammalian Cells Using BacMam, a Modified Baculovirus System. *Methods Mol Biol,* 1350**,** 95-116.

FREDRIKSSON, R., LAGERSTRÖM, M. C., LUNDIN, L. G. & SCHIÖTH, H. B. 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol,* 63**,** 1256-72.

FREEZE, H. H. & KRANZ, C. 2008. Endoglycosidase and glycoamidase release of N-linked glycans. *Curr Protoc Immunol,* Chapter 8**,** 8.15.1-8.15.26.

FREUNDT, G. V., VON SAMSON-HIMMELSTJERNA, F. A., NITZ, J. T., LUEDDE, M., WALTENBERGER, J., WIELAND, T., FREY, N., PREUSCH, M. & HIPPE, H. J. 2022. The orphan receptor GPRC5B activates pro-inflammatory signaling in the vascular wall via Fyn and NFκB. *Biochem Biophys Res Commun,* 592**,** 60-66.

GAINETDINOV, R. R., PREMONT, R. T., BOHN, L. M., LEFKOWITZ, R. J. & CARON, M. G. 2004. Desensitization of G protein-coupled receptors and neuronal functions. *Annu Rev Neurosci,* 27**,** 107-44.

GALLAGHER, J. R., KIM, A. J., GULATI, N. M. & HARRIS, A. K. 2019. Negative-Stain Transmission Electron Microscopy of Molecular Complexes for Image Analysis by 2D Class Averaging. *Curr Protoc Microbiol,* 54**,** e90.

GARCÍA-NAFRÍA, J., LEE, Y., BAI, X., CARPENTER, B. & TATE, C. G. 2018. Cryo-EM structure of the adenosine A(2A) receptor coupled to an engineered heterotrimeric G protein. *Elife,* 7.

GARIN, N. & ESCHER, G. 2001. The development of inhibitory synaptic specializations in the mouse deep cerebellar nuclei. *Neuroscience,* 105**,** 431-41.

GAUDRY, J. P., AROD, C., SAUVAGE, C., BUSSO, S., DUPRAZ, P., PANKIEWICZ, R. & ANTONSSON, B. 2008. Purification of the extracellular domain of the membrane protein GlialCAM expressed in HEK and CHO cells and comparison of the glycosylation. *Protein Expr Purif,* 58**,** 94-102.

GEORGE, S. R., O'DOWD, B. F. & LEE, S. P. 2002. G-protein-coupled receptor oligomerization and its potential for drug discovery. *Nat Rev Drug Discov,* 1**,** 808-20.

GILBERT, A., ELORZA-VIDAL, X., RANCILLAC, A., CHAGNOT, A., YETIM, M., HINGOT, V., DEFFIEUX, T., BOULAY, A. C., ALVEAR-PEREZ, R., CISTERNINO, S., MARTIN, S., TAÏB, S., GELOT, A., MIGNON, V., FAVIER, M., BRUNET, I., DECLÈVES, X., TANTER, M., ESTEVEZ, R., VIVIEN, D., SAUBAMÉA, B. & COHEN-SALMON, M. 2021. Megalencephalic leukoencephalopathy with subcortical cysts is a developmental disorder of the gliovascular unit. *Elife,* 10.

GOEHRING, A., LEE, C. H., WANG, K. H., MICHEL, J. C., CLAXTON, D. P., BACONGUIS, I., ALTHOFF, T., FISCHER, S., GARCIA, K. C. & GOUAUX, E. 2014. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat Protoc,* 9**,** 2574-85.

GOHON, Y., GIUSTI, F., PRATA, C., CHARVOLIN, D., TIMMINS, P., EBEL, C., TRIBET, C. & POPOT, J. L. 2006. Well-defined nanoparticles formed by hydrophobic assembly of a short and polydisperse random terpolymer, amphipol A8-35. *Langmuir,* 22**,** 1281-90.

GOROSPE, J. R., SINGHAL, B. S., KAINU, T., WU, F., STEPHAN, D., TRENT, J., HOFFMAN, E. P. & NAIDU, S. 2004. Indian Agarwal megalencephalic leukodystrophy with cysts is caused by a common MLC1 mutation. *Neurology,* 62**,** 878-82.

GOUTAM, K., IELASI, F. S., PARDON, E., STEYAERT, J. & REYES, N. 2022. Structural basis of sodium-dependent bile salt uptake into the liver. *Nature,* 606**,** 1015-1020.

GRBA, D. N., WRIGHT, J. J., YIN, Z., FISHER, W. & HIRST, J. 2024. Molecular mechanism of the ischemia-induced regulatory switch in mammalian complex I. *Science,* 384**,** 1247-1253.

GREENHOUGH, A., BAGLEY, C., HEESOM, K. J., GUREVICH, D. B., GAY, D., BOND, M., COLLARD, T. J., PARASKEVA, C., MARTIN, P., SANSOM, O. J., MALIK, K. & WILLIAMS, A. C. 2018. Cancer cell adaptation to hypoxia involves a HIF-GPRC5A-YAP axis. *EMBO Mol Med,* 10.

GRÄSLUND, S., NORDLUND, P., WEIGELT, J., HALLBERG, B. M., BRAY, J., GILEADI, O., KNAPP, S., OPPERMANN, U., ARROWSMITH, C., HUI, R., MING, J., DHE-PAGANON, S., PARK, H. W., SAVCHENKO, A., YEE, A., EDWARDS, A., VINCENTELLI, R., CAMBILLAU, C., KIM, R., KIM, S. H., RAO, Z., SHI, Y.,

TERWILLIGER, T. C., KIM, C. Y., HUNG, L. W., WALDO, G. S., PELEG, Y., ALBECK, S., UNGER, T., DYM, O., PRILUSKY, J., SUSSMAN, J. L., STEVENS, R. C., LESLEY, S. A., WILSON, I. A., JOACHIMIAK, A., COLLART, F., DEMENTIEVA, I., DONNELLY, M. I., ESCHENFELDT, W. H., KIM, Y., STOLS, L., WU, R., ZHOU, M., BURLEY, S. K., EMTAGE, J. S., SAUDER, J. M., THOMPSON, D., BAIN, K., LUZ, J., GHEYI, T., ZHANG, F., ATWELL, S., ALMO, S. C., BONANNO, J. B., FISER, A., SWAMINATHAN, S., STUDIER, F. W., CHANCE, M. R., SALI, A., ACTON, T. B., XIAO, R., ZHAO, L., MA, L. C., HUNT, J. F., TONG, L., CUNNINGHAM, K., INOUYE, M., ANDERSON, S., JANJUA, H., SHASTRY, R., HO, C. K., WANG, D., WANG, H., JIANG, M., MONTELIONE, G. T., STUART, D. I., OWENS, R. J., DAENKE, S., SCHÜTZ, A., HEINEMANN, U., YOKOYAMA, S., BÜSSOW, K. & GUNSALUS, K. C. 2008. Protein production and purification. *Nat Methods,* 5**,** 135-46.

GUO, P., TAI, Y., WANG, M., SUN, H., ZHANG, L., WEI, W., XIANG, Y. K. & WANG, Q. 2022. Gα(12) and Gα(13): Versatility in Physiology and Pathology. *Front Cell Dev Biol,* 10**,** 809425.

GUO, Q., HE, B., ZHONG, Y., JIAO, H., REN, Y., WANG, Q., GE, Q., GAO, Y., LIU, X., DU, Y., HU, H. & TAO, Y. 2024. A method for structure determination of GPCRs in various states. *Nat Chem Biol,* 20**,** 74-82.

HAMILTON, E. M. C., TEKTURK, P., CIALDELLA, F., VAN RAPPARD, D. F., WOLF, N. I., YALCINKAYA, C., ÇETINÇELIK, Ü., RAJAEE, A., KARIMINEJAD, A., PAPROCKA, J., YAPICI, Z., BOŠNJAK, V. M. & VAN DER KNAAP, M. S. 2018. Megalencephalic leukoencephalopathy with subcortical cysts: Characterization of disease variants. *Neurology,* 90**,** e1395-e1403.

HE, Z., ZHAO, Y., RAU, M. J., FITZPATRICK, J. A. J., SAH, R., HU, H. & YUAN, P. 2023. Structural and functional analysis of human pannexin 2 channel. *Nat Commun,* 14**,** 1712.

HENDRIKS-BALK, M. C., PETERS, S. L., MICHEL, M. C. & ALEWIJNSE, A. E. 2008. Regulation of G protein-coupled receptor signalling: focus on the cardiovascular system and regulator of G protein signalling proteins. *Eur J Pharmacol,* 585**,** 278-91.

HILL, S. F. & MEISLER, M. H. 2021. Antisense Oligonucleotide Therapy for Neurodevelopmental Disorders. *Dev Neurosci,* 43**,** 247-252.

HOEGG-BEILER, M. B., SIRISI, S., OROZCO, I. J., FERRER, I., HOHENSEE, S., AUBERSON, M., GÖDDE, K., VILCHES, C., DE HEREDIA, M. L., NUNES, V., ESTÉVEZ, R. & JENTSCH, T. J. 2014. Disrupting MLC1 and GlialCAM and ClC-2 interactions in leukodystrophy entails glial chloride channel dysfunction. *Nat Commun,* 5**,** 3475.

HOLLMANN, M. W., STRUMPER, D., HERROEDER, S. & DURIEUX, M. E. 2005. Receptors, G proteins, and their interactions. *Anesthesiology,* 103**,** 1066-78.

HUXLEY, A. F. & STÄMPFLI, R. 1949. Evidence for saltatory conduction in peripheral myelinated nerve fibres. *J Physiol,* 108**,** 315-39.

HWANG, J., PARK, K., LEE, G. Y., YOON, B. Y., KIM, H., ROH, S. H., LEE, B. C., KIM, K. & LIM, H. H. 2021. Transmembrane topology and oligomeric nature of an astrocytic membrane protein, MLC1. *Open Biol,* 11**,** 210103.

IMAI, Y., SODA, M., INOUE, H., HATTORI, N., MIZUNO, Y. & TAKAHASHI, R. 2001. An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of Parkin. *Cell,* 105**,** 891-902.

INOUE, S., NAMBU, T. & SHIMOMURA, T. 2004. The RAIG family member, GPRC5D, is associated with hard-keratinized structures. *J Invest Dermatol,* 122**,** 565-73.

IROBALIEVA, R. N., MANOLARIDIS, I., JACKSON, S. M., NI, D., PARDON, E., STAHLBERG, H., STEYAERT, J. & LOCHER, K. P. 2023. Structural Basis of the Allosteric Inhibition of Human ABCG2 by Nanobodies. *J Mol Biol,* 435**,** 168234.

JACOBY, E., BOUHELAL, R., GERSPACHER, M. & SEUWEN, K. 2006. The 7 TM G-protein-coupled receptor target family. *ChemMedChem,* 1**,** 761-82.

JAROSZEWICZ, W., MORCINEK-ORŁOWSKA, J., PIERZYNOWSKA, K., GAFFKE, L. & WĘGRZYN, G. 2022. Phage display and other peptide display technologies. *FEMS Microbiol Rev,* 46.

JEWORUTZKI, E., LÓPEZ-HERNÁNDEZ, T., CAPDEVILA-NORTES, X., SIRISI, S., BENGTSSON, L., MONTOLIO, M., ZIFARELLI, G., ARNEDO, T., MÜLLER, C. S., SCHULTE, U., NUNES, V., MARTÍNEZ, A., JENTSCH, T. J., GASULL, X., PUSCH, M. & ESTÉVEZ, R. 2012. GlialCAM, a protein defective in a leukodystrophy, serves as a ClC-2 Cl(-) channel auxiliary subunit. *Neuron,* 73**,** 951-61.

JIN, B. K., ODONGO, S., RADWANSKA, M. & MAGEZ, S. 2023. NANOBODIES®: A Review of Diagnostic and Therapeutic Applications. *Int J Mol Sci,* 24.

JUMPER, J., EVANS, R., PRITZEL, A., GREEN, T., FIGURNOV, M., RONNEBERGER, O., TUNYASUVUNAKOOL, K., BATES, R., ŽÍDEK, A., POTAPENKO, A., BRIDGLAND, A., MEYER, C., KOHL, S. A. A., BALLARD, A. J., COWIE, A., ROMERA-PAREDES, B., NIKOLOV, S., JAIN, R., ADLER, J., BACK, T., PETERSEN, S., REIMAN, D., CLANCY, E., ZIELINSKI, M., STEINEGGER, M., PACHOLSKA, M., BERGHAMMER, T., BODENSTEIN, S., SILVER, D., VINYALS, O., SENIOR, A. W., KAVUKCUOGLU, K., KOHLI, P. & HASSABIS, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature,* 596**,** 583-589.

KAMHOLZ, J. A. 1996. Regulation of myelin development. *Mult Scler,* 2**,** 236-40.

KAMIGUCHI, H. & LEMMON, V. 2000. IgCAMs: bidirectional signals underlying neurite growth. *Curr Opin Cell Biol,* 12**,** 598-605.

KATZEN, F. 2007. Gateway(®) recombinational cloning: a biological operating system. *Expert Opin Drug Discov,* 2**,** 571-89.

KELLY, E., BAILEY, C. P. & HENDERSON, G. 2008. Agonist-selective mechanisms of GPCR desensitization. *Br J Pharmacol,* 153 Suppl 1**,** S379-88.

KERN, D. M., BLEIER, J., MUKHERJEE, S., HILL, J. M., KOSSIAKOFF, A. A., ISACOFF, E. Y. & BROHAWN, S. G. 2023. Structural basis for assembly and lipid-mediated gating of LRRC8A:C volume-regulated anion channels. *Nat Struct Mol Biol,* 30**,** 841-852.

KETTENMANN, H. & VERKHRATSKY, A. 2011. [Neuroglia--living nerve glue]. *Fortschr Neurol Psychiatr,* 79**,** 588-97.

KIELKOPF, C. L., BAUER, W. & URBATSCH, I. L. 2021. Considerations for Membrane Protein Purification. *Cold Spring Harb Protoc,* 2021.

KIM, Y. J., SANO, T., NABETANI, T., ASANO, Y. & HIRABAYASHI, Y. 2012. GPRC5B activates obesity-associated inflammatory signaling in adipocytes. *Sci Signal,* 5**,** ra85.

KNIAZEFF, J., PRÉZEAU, L., RONDARD, P., PIN, J. P. & GOUDET, C. 2011. Dimers and beyond: The functional puzzles of class C GPCRs. *Pharmacol Ther,* 130**,** 9-25.

KOLAKOWSKI, L. F., JR. 1994. GCRDb: a G-protein-coupled receptor database. *Recept Channels,* 2**,** 1-7.

KUMAR, G. 2018. Principle and Method of Silver Staining of Proteins Separated by Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. *Methods Mol Biol,* 1853**,** 231-236.

KURABAYASHI, N., NGUYEN, M. D. & SANADA, K. 2013. The G protein-coupled receptor GPRC5B contributes to neurogenesis in the developing mouse neocortex. *Development,* 140**,** 4335-46.

KWON, S. H., LIU, K. D. & MOSTOV, K. E. 2014. Intercellular transfer of GPRC5B via exosomes drives HGF-mediated outward growth. *Curr Biol,* 24**,** 199-204.

LAGERSTRÖM, M. C. & SCHIÖTH, H. B. 2008. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov,* 7**,** 339-57.

LAL, M., WACHTEL, E., PATI, S., NAMBOOTHIRI, I. N. N. & PATCHORNIK, G. 2023. His(1)-tagged DM or DDM detergent micelles are reversibly conjugated by nickel ions. *Sci Rep,* 13**,** 17138.

LANCIOTTI, A., BRIGNONE, M. S., CAMERINI, S., SERAFINI, B., MACCHIA, G., RAGGI, C., MOLINARI, P., CRESCENZI, M., MUSUMECI, M., SARGIACOMO, M., ALOISI, F., PETRUCCI, T. C. & AMBROSINI, E. 2010. MLC1 trafficking and membrane expression in astrocytes: role of caveolin-1 and phosphorylation. *Neurobiol Dis,* 37**,** 581-95.

LANZ, T. V., BREWER, R. C., HO, P. P., MOON, J. S., JUDE, K. M., FERNANDEZ, D., FERNANDES, R. A., GOMEZ, A. M., NADJ, G. S., BARTLEY, C. M., SCHUBERT, R. D., HAWES, I. A., VAZQUEZ, S. E., IYER, M., ZUCHERO, J. B., TEEGEN, B., DUNN, J. E., LOCK, C. B., KIPP, L. B., COTHAM, V. C., UEBERHEIDE, B. M., AFTAB, B. T., ANDERSON, M. S., DERISI, J. L., WILSON, M. R., BASHFORD-ROGERS, R. J. M., PLATTEN, M., GARCIA, K. C., STEINMAN, L. & ROBINSON, W. H. 2022. Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature,* 603**,** 321-327.

LAVERTY, D., DESAI, R., UCHAŃSKI, T., MASIULIS, S., STEC, W. J., MALINAUSKAS, T., ZIVANOV, J., PARDON, E., STEYAERT, J., MILLER, K. W. & ARICESCU, A. R. 2019. Cryo-EM structure of the human α1β3γ2 GABA(A) receptor in a lipid bilayer. *Nature,* 565**,** 516-520.

LE BON, C., MARCONNET, A., MASSCHELEYN, S., POPOT, J. L. & ZOONENS, M. 2018. Folding and stabilizing membrane proteins in amphipol A8-35. *Methods,* 147**,** 95-105.

LEDSGAARD, L., KILSTRUP, M., KARATT-VELLATT, A., MCCAFFERTY, J. & LAUSTSEN, A. H. 2018. Basics of Antibody Phage Display Technology. *Toxins (Basel),* 10.

LEE, S. J., PARK, Y. S., KIM, S. J., LEE, B. J. & SUH, S. W. 2014. Crystal structure of PhoU from Pseudomonas aeruginosa, a negative regulator of the Pho regulon. *J Struct Biol,* 188**,** 22-9.

LEEGWATER, P. A., BOOR, P. K., YUAN, B. Q., VAN DER STEEN, J., VISSER, A., KÖNST, A. A., OUDEJANS, C. B., SCHUTGENS, R. B., PRONK, J. C. & VAN DER KNAAP, M. S. 2002. Identification of novel mutations in MLC1 responsible for megalencephalic leukoencephalopathy with subcortical cysts. *Hum Genet,* 110**,** 279-83.

LEEGWATER, P. A., YUAN, B. Q., VAN DER STEEN, J., MULDERS, J., KÖNST, A. A., BOOR, P. K., MEJASKI-BOSNJAK, V., VAN DER MAAREL, S. M., FRANTS, R. R., OUDEJANS, C. B., SCHUTGENS, R. B., PRONK, J. C. & VAN DER KNAAP, M. S. 2001. Mutations of MLC1 (KIAA0027), encoding a putative membrane protein, cause megalencephalic leukoencephalopathy with subcortical cysts. *Am J Hum Genet,* 68**,** 831-8.

LEITZ, A. J., BAYBURT, T. H., BARNAKOV, A. N., SPRINGER, B. A. & SLIGAR, S. G. 2006. Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology. *Biotechniques,* 40**,** 601-2, 604, 606, passim.

LEONARD, P., HEARTY, S., MA, H. & O'KENNEDY, R. 2017. Measuring Protein-Protein Interactions Using Biacore. *Methods Mol Biol,* 1485**,** 339-354.

LIANG, Y. L., KHOSHOUEI, M., GLUKHOVA, A., FURNESS, S. G. B., ZHAO, P., CLYDESDALE, L., KOOLE, C., TRUONG, T. T., THAL, D. M., LEI, S., RADJAINIA, M., DANEV, R., BAUMEISTER, W., WANG, M. W., MILLER, L. J., CHRISTOPOULOS, A., SEXTON, P. M. & WOOTTEN, D. 2018. Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1 receptor-Gs complex. *Nature,* 555**,** 121-125.

LIEDTKE, W., EDELMANN, W., BIERI, P. L., CHIU, F. C., COWAN, N. J., KUCHERLAPATI, R. & RAINE, C. S. 1996. GFAP is necessary for the integrity of CNS white matter architecture and long-term maintenance of myelination. *Neuron,* 17**,** 607-15.

LIN, S. H. & GUIDOTTI, G. 2009. Purification of membrane proteins. *Methods Enzymol,* 463**,** 619-29.

LITTLE, K. D., HEMLER, M. E. & STIPP, C. S. 2004. Dynamic regulation of a GPCR-tetraspanin-G protein complex on intact cells: central role of CD81 in facilitating GPR56-Galpha q/11 association. *Mol Biol Cell,* 15**,** 2375-87.

LIU, C., GONG, J. S., SU, C., LI, H., RAO, Z. M., XU, Z. H. & SHI, J. S. 2022. Pathway engineering facilitates efficient protein expression in Pichia pastoris. *Appl Microbiol Biotechnol,* 106**,** 5893-5912.

LV, Z., HE, Y., XIANG, Y., LI, J., MENG, F., LAN, B., GUO, H., HE, D., WANG, Y., ZAHO, H., ZHUO, W., LIU, Y., LIU, X., NI, X. & HENG, J. 2022. Cryo-EM complex structure of active GPR75 with a nanobody. *bioRxiv*.

LÓPEZ-HERNÁNDEZ, T., RIDDER, M. C., MONTOLIO, M., CAPDEVILA-NORTES, X., POLDER, E., SIRISI, S., DUARRI, A., SCHULTE, U., FAKLER, B., NUNES, V., SCHEPER, G. C., MARTÍNEZ, A., ESTÉVEZ, R. & VAN DER KNAAP, M. S. 2011a. Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. *Am J Hum Genet,* 88**,** 422-32.

LÓPEZ-HERNÁNDEZ, T., SIRISI, S., CAPDEVILA-NORTES, X., MONTOLIO, M., FERNÁNDEZ-DUEÑAS, V., SCHEPER, G. C., VAN DER KNAAP, M. S., CASQUERO, P., CIRUELA, F., FERRER, I., NUNES, V. & ESTÉVEZ, R. 2011b. Molecular mechanisms of MLC1 and GLIALCAM mutations in megalencephalic leukoencephalopathy with subcortical cysts. *Hum Mol Genet,* 20**,** 3266-77.

MADADKAR-SOBHANI, A. & GUALLAR, V. 2013. PELE web server: atomistic study of biomolecular systems at your fingertips. *Nucleic Acids Res,* 41**,** W322-8.

MALHOTRA, A. 2009. Tagging for protein expression. *Methods Enzymol,* 463**,** 239-58.

MANCINI, C., VAULA, G., SCALZITTI, L., CAVALIERI, S., BERTINI, E., AIELLO, C., LUCCHINI, C., GATTI, R. A., BRUSSINO, A. & BRUSCO, A. 2012. Megalencephalic leukoencephalopathy with subcortical cysts type 1 (MLC1) due to a homozygous deep intronic splicing mutation (c.895-226T>G) abrogated in vitro using an antisense morpholino oligonucleotide. *Neurogenetics,* 13**,** 205-14.

MANGALA RAO AND CARL, R. A. 2000. Delivery of lipids and liposomal proteins to the cytoplasm and Golgi of antigen-presenting cells. *Advanced Drug Delivery Reviews,* 41**,** 171-188.

MARAZZITI, D., DI PIETRO, C., GOLINI, E., MANDILLO, S., LA SALA, G., MATTEONI, R. & TOCCHINI-VALENTINI, G. P. 2013. Precocious cerebellum development and improved motor functions in mice lacking the astrocyte cilium-, patched 1-associated Gpr37l1 receptor. *Proc Natl Acad Sci U S A,* 110**,** 16486-91.

MARAZZITI, D., GALLO, A., GOLINI, E., MATTEONI, R. & TOCCHINI-VALENTINI, G. P. 1998. Molecular cloning and chromosomal localization of the mouse Gpr37 gene encoding an orphan G-protein-coupled peptide receptor expressed in brain and testis. *Genomics,* 53**,** 315-24.

MATSUZAKI, Y., KAJIWARA, K., AOKI, W. & UEDA, M. 2022. Production of Single-Domain Antibodies in Pichia pastoris. *Methods Mol Biol,* 2446**,** 181-203.

MAYAYO-VALLVERDÚ, C., FERIGLE, L., VECINO-PÉREZ, M., LARA, J., NUNES, V. & ESTÉVEZ, R. 2023. Characterization of an MLC patient carrying two MLC1 variants showing radiological improvement.

MILLIGAN, G. & KOSTENIS, E. 2006. Heterotrimeric G-proteins: a short history. *Br J Pharmacol,* 147 Suppl 1**,** S46-55.

MINATEL, V. M., PRUDENCIO, C. R., BARRAVIERA, B. & FERREIRA, R. S., JR. 2023. Nanobodies: a promising approach to treatment of viral diseases. *Front Immunol,* 14**,** 1303353.

MIYAGI, H., ASADA, H., SUZUKI, M., TAKAHASHI, Y., YASUNAGA, M., SUNO, C., IWATA, S. & SAITO, J. I. 2020. The discovery of a new antibody for BRIL-fused GPCR structure determination. *Sci Rep,* 10**,** 11669.

MIYAGI, H., SUZUKI, M., YASUNAGA, M., ASADA, H., IWATA, S. & SAITO, J. I. 2023. Structural insight into an anti-BRIL Fab as a G-protein-coupled receptor crystallization chaperone. *Acta Crystallogr D Struct Biol,* 79**,** 435-441.

MOH, M. C., ZHANG, C., LUO, C., LEE, L. H. & SHEN, S. 2005. Structural and functional analyses of a novel ig-like cell adhesion molecule, hepaCAM, in the human breast carcinoma MCF7 cells. *J Biol Chem,* 280**,** 27366-74.

MONTAGNA, G., TEIJIDO, O., EYMARD-PIERRE, E., MURAKI, K., COHEN, B., LOIZZO, A., GROSSO, P., TEDESCHI, G., PALACÍN, M., BOESPFLUG-TANGUY, O., BERTINI, E., SANTORELLI, F. M. & ESTÉVEZ, R. 2006. Vacuolating megalencephalic leukoencephalopathy with subcortical cysts: functional studies of novel variants in MLC1. *Hum Mutat,* 27**,** 292.

MUKHERJEE, S., ERRAMILLI, S. K., AMMIRATI, M., ALVAREZ, F. J. D., FENNELL, K. F., PURDY, M. D., SKROBEK, B. M., RADZIWON, K., COUKOS, J., KANG, Y., DUTKA, P., GAO, X., QIU, X., YEAGER, M., ERIC XU, H., HAN, S. & KOSSIAKOFF, A. A. 2020. Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. *Nat Commun,* 11**,** 1598.

MUYLDERMANS, S. 2013. Nanobodies: natural single-domain antibodies. *Annu Rev Biochem,* 82**,** 775-97.

MUYLDERMANS, S. 2021. Applications of Nanobodies. *Annu Rev Anim Biosci,* 9**,** 401-421.

MÉTAYÉ, T., GIBELIN, H., PERDRISOT, R. & KRAIMPS, J. L. 2005. Pathophysiological roles of G-protein-coupled receptor kinases. *Cell Signal,* 17**,** 917-28.

NAIDU, S. 1999. Clinical delineation of leukodystrophies. *J Mol Neurosci,* 12**,** 185-192.

NOMURA, N., MIYAJIMA, N., SAZUKA, T., TANAKA, A., KAWARABAYASI, Y., SATO, S., NAGASE, T., SEKI, N., ISHIKAWA, K. & TABATA, S. 1994. Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1. *DNA Res,* 1**,** 27-35.

NYGAARD, R., KIM, J. & MANCIA, F. 2020. Cryo-electron microscopy analysis of small membrane proteins. *Curr Opin Struct Biol,* 64**,** 26-33.

NÜRNBERG, B., BEER-HAMMER, S., REISINGER, E. & LEISS, V. 2024. Non-canonical G protein signaling. *Pharmacol Ther,* 255**,** 108589.

OFFERMANNS, S. 2003. G-proteins as transducers in transmembrane signalling. *Prog Biophys Mol Biol,* 83**,** 101-30.

ORLOVA, E. V. & SAIBIL, H. R. 2011. Structural analysis of macromolecular assemblies by electron microscopy. *Chem Rev,* 111**,** 7710-48.

PALCZEWSKI, K., KUMASAKA, T., HORI, T., BEHNKE, C. A., MOTOSHIMA, H., FOX, B. A., LE TRONG, I., TELLER, D. C., OKADA, T., STENKAMP, R. E., YAMAMOTO, M. & MIYANO, M. 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science,* 289**,** 739-45.

PALMBERGER, D., WILSON, I. B., BERGER, I., GRABHERR, R. & RENDIC, D. 2012. SweetBac: a new approach for the production of mammalianised glycoproteins in insect cells. *PLoS One,* 7**,** e34226.

PANDEY, A., SHIN, K., PATTERSON, R. E., LIU, X. Q. & RAINEY, J. K. 2016. Current strategies for protein production and purification enabling membrane protein structural biology. *Biochem Cell Biol,* 94**,** 507-527.

PARDON, E., LAEREMANS, T., TRIEST, S., RASMUSSEN, S. G., WOHLKÖNIG, A., RUF, A., MUYLDERMANS, S., HOL, W. G., KOBILKA, B. K. & STEYAERT, J. 2014. A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc,* 9**,** 674-93.

PARK, P. S., LODOWSKI, D. T. & PALCZEWSKI, K. 2008. Activation of G protein-coupled receptors: beyond two-state models and tertiary conformational changes. *Annu Rev Pharmacol Toxicol,* 48**,** 107-41.

PASSCHIER, E. M. J., BISSELING, Q., HELMAN, G., VAN SPAENDONK, R. M. L., SIMONS, C., OLSTHOORN, R. C. L., VAN DER VEEN, H., ABBINK, T. E. M., VAN DER KNAAP, M. S. & MIN, R. 2024. Megalencephalic leukoencephalopathy with subcortical cysts: a variant update and review of the literature. *Front Genet,* 15**,** 1352947.

PASSCHIER, E. M. J., KERST, S., BROUWERS, E., HAMILTON, E. M. C., BISSELING, Q., BUGIANI, M., WAISFISZ, Q., KITCHEN, P., UNGER, L., BREUR, M., HOOGTERP, L., DE VRIES, S. I., ABBINK, T. E. M., KOLE, M. H. P., LEURS, R., VISCHER, H. F., BRIGNONE, M. S., AMBROSINI, E., FEILLET, F., BORN, A. P., EPSTEIN, L. G., MANSVELDER, H. D., MIN, R. & VAN DER KNAAP, M. S. 2023. Aquaporin-4 and GPRC5B: old and new players in controlling brain oedema. *Brain,* 146**,** 3444-3454.

PATRA, P., DAS, M., KUNDU, P. & GHOSH, A. 2021. Recent advances in systems and synthetic biology approaches for developing novel cell-factories in non-conventional yeasts. *Biotechnol Adv,* 47**,** 107695.

PATRONO, C., DI GIACINTO, G., EYMARD-PIERRE, E., SANTORELLI, F. M., RODRIGUEZ, D., DE STEFANO, N., FEDERICO, A., GATTI, R., BENIGNO, V., MEGARBANÉ, A., TABARKI, B., BOESPFLUG-TANGUY, O. & BERTINI, E. 2003. Genetic heterogeneity of megalencephalic leukoencephalopathy and subcortical cysts. *Neurology,* 61**,** 534-7.

PERLMUTTER, J. D., POPOT, J. L. & SACHS, J. N. 2014. Molecular dynamics simulations of a membrane protein/amphipol complex. *J Membr Biol,* 247**,** 883-95.

PIALA, A. T., MOON, T. M., AKELLA, R., HE, H., COBB, M. H. & GOLDSMITH, E. J. 2014. Chloride sensing by WNK1 involves inhibition of autophosphorylation. *Sci Signal,* 7**,** ra41.

PLA-CASILLANIS, A., FERIGLE, L., ALONSO-GARDÓN, M., XICOY-ESPAULELLA, E., ERRASTI-MURUGARREN, E., MARAZZITI, D. & ESTÉVEZ, R. 2022. GPR37 Receptors and Megalencephalic Leukoencephalopathy with Subcortical Cysts. *Int J Mol Sci,* 23.

POPOT, J. L., ALTHOFF, T., BAGNARD, D., BANÈRES, J. L., BAZZACCO, P., BILLON-DENIS, E., CATOIRE, L. J., CHAMPEIL, P., CHARVOLIN, D., COCCO, M. J., CRÉMEL, G., DAHMANE, T., DE LA MAZA, L. M., EBEL, C., GABEL, F., GIUSTI, F., GOHON, Y., GOORMAGHTIGH, E., GUITTET, E., KLEINSCHMIDT, J. H., KÜHLBRANDT, W., LE BON, C., MARTINEZ, K. L., PICARD, M., PUCCI, B., SACHS, J. N., TRIBET, C., VAN HEIJENOORT, C., WIEN, F., ZITO, F. & ZOONENS, M. 2011. Amphipols from A to Z. *Annu Rev Biophys,* 40**,** 379-408.

POSCH, W., LASS-FLÖRL, C. & WILFLINGSEDER, D. 2016. Generation of Human Monocyte-derived Dendritic Cells from Whole Blood. *J Vis Exp*.

PUVANENDRAN, D., SOUABNI, H., SALVADOR, D., LAMBERT, O., CECE, Q. & PICARD, M. 2020. Rationale for the Quantitative Reconstitution of Membrane Proteins into Proteoliposomes. *Methods Mol Biol,* 2168**,** 63-72.

PÉREZ-RIUS, C., FOLGUEIRA, M., ELORZA-VIDAL, X., ALIA, A., HOEGG-BEILER, M. B., EEZA, M. N. H., DÍAZ, M. L., NUNES, V., BARRALLO-GIMENO, A. & ESTÉVEZ, R. 2019. Comparison of zebrafish and mice knockouts for Megalencephalic Leukoencephalopathy proteins indicates that GlialCAM/MLC1 forms a functional unit. *Orphanet J Rare Dis,* 14**,** 268.

RAJKUMAR, P., CHA, B., YIN, J., AREND, L. J., PĂUNESCU, T. G., HIRABAYASHI, Y., DONOWITZ, M. & PLUZNICK, J. L. 2018. Identifying the localization and exploring a functional role for Gprc5c in the kidney. *Faseb j,* 32**,** 2046-2059.

RASMUSSEN, S. G., DEVREE, B. T., ZOU, Y., KRUSE, A. C., CHUNG, K. Y., KOBILKA, T. S., THIAN, F. S., CHAE, P. S., PARDON, E., CALINSKI, D., MATHIESEN, J. M., SHAH, S. T., LYONS, J. A., CAFFREY, M., GELLMAN, S. H., STEYAERT, J., SKINIOTIS, G., WEIS, W. I., SUNAHARA, R. K. & KOBILKA, B. K. 2011. Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature,* 477**,** 549-55.

RATKEVICIUTE, G., COOPER, B. F. & KNOWLES, T. J. 2021. Methods for the solubilisation of membrane proteins: the micelle-aneous world of membrane protein solubilisation. *Biochem Soc Trans,* 49**,** 1763-1777.

RENAUD, J. P., CHARI, A., CIFERRI, C., LIU, W. T., RÉMIGY, H. W., STARK, H. & WIESMANN, C. 2018. Cryo-EM in drug discovery: achievements, limitations and prospects. *Nat Rev Drug Discov,* 17**,** 471-492.

ROBBINS, M. J., MICHALOVICH, D., HILL, J., CALVER, A. R., MEDHURST, A. D., GLOGER, I., SIMS, M., MIDDLEMISS, D. N. & PANGALOS, M. N. 2000. Molecular cloning and characterization of two novel retinoic acid-inducible orphan G-protein-coupled receptors (GPRC5B and GPRC5C). *Genomics,* 67**,** 8-18.

ROBERTSON, M. J., PAPASERGI-SCOTT, M. M., HE, F., SEVEN, A. B., MEYEROWITZ, J. G., PANOVA, O., PEROTO, M. C., CHE, T. & SKINIOTIS, G. 2022. Structure determination of inactive-state GPCRs with a universal nanobody. *Nat Struct Mol Biol,* 29**,** 1188-1195.

ROSENBAUM, D. M., RASMUSSEN, S. G. & KOBILKA, B. K. 2009. The structure and function of G-protein-coupled receptors. *Nature,* 459**,** 356-63.

SANO, T., KIM, Y. J., OSHIMA, E., SHIMIZU, C., KIYONARI, H., ABE, T., HIGASHI, H., YAMADA, K. & HIRABAYASHI, Y. 2011. Comparative characterization of GPRC5B and GPRC5CLacZ knockin mice; behavioral abnormalities in GPRC5B-deficient mice. *Biochem Biophys Res Commun,* 412**,** 460-5.

SANO, T., KOHYAMA-KOGANEYA, A., KINOSHITA, M. O., TATSUKAWA, T., SHIMIZU, C., OSHIMA, E., YAMADA, K., LE, T. D., AKAGI, T., TOHYAMA, K., NAGAO, S. & HIRABAYASHI, Y. 2018. Loss of GPRC5B impairs synapse formation of Purkinje cells with cerebellar nuclear neurons and disrupts cerebellar synaptic plasticity and motor learning. *Neurosci Res,* 136**,** 33-47.

SCHIMERLIK, M. I. 2001. Overview of membrane protein solubilization. *Curr Protoc Neurosci,* Chapter 5**,** Unit 5.9.

SCHIÖTH, H. B. & FREDRIKSSON, R. 2005. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen Comp Endocrinol,* 142**,** 94-101.

SCHLÜTER, A., RODRÍGUEZ-PALMERO, A., VERDURA, E., VÉLEZ-SANTAMARÍA, V., RUIZ, M., FOURCADE, S., PLANAS-SERRA, L., MARTÍNEZ, J. J., GUILERA, C., GIRÓS, M., ARTUCH, R., YOLDI, M. E., O'CALLAGHAN, M., GARCÍA-CAZORLA, A., ARMSTRONG, J., MARTI, I., MONDRAGÓN REZOLA, E., REDIN, C., MANDEL, J. L., CONEJO, D., SIERRA-CÓRCOLES, C., BELTRÁN, S., GUT, M., VÁZQUEZ, E., DEL TORO, M., TRONCOSO, M., PÉREZ-JURADO, L. A., GUTIÉRREZ-SOLANA, L. G., LÓPEZ DE MUNAIN, A., CASASNOVAS, C., AGUILERA-ALBESA, S., MACAYA, A. & PUJOL, A. 2022. Diagnosis of Genetic White Matter Disorders by Singleton Whole-Exome and Genome Sequencing Using Interactome-Driven Prioritization. *Neurology,* 98**,** e912-e923.

SCHMITT, A., GOFFERJE, V., WEBER, M., MEYER, J., MÖSSNER, R. & LESCH, K. P. 2003. The brain-specific protein MLC1 implicated in megalencephalic leukoencephalopathy with subcortical cysts is expressed in glial cells in the murine brain. *Glia,* 44**,** 283-95.

SCHÜTZ, A., BERNHARD, F., BERROW, N., BUYEL, J. F., FERREIRA-DA-SILVA, F., HAUSTRAETE, J., VAN DEN HEUVEL, J., HOFFMANN, J. E., DE MARCO, A., PELEG, Y., SUPPMANN, S., UNGER, T., VANHOUCKE, M., WITT, S. & REMANS, K. 2023. A concise guide to choosing suitable gene expression systems for recombinant protein production. *STAR Protoc,* 4**,** 102572.

SENER, R. N. 2003. Proton MR spectroscopy demonstration of taurine peaks in megalencephalic leukoencephalopathy with cysts. *Comput Med Imaging Graph,* 27**,** 23-6.

SHAW, A. W., PUREZA, V. S., SLIGAR, S. G. & MORRISSEY, J. H. 2007. The local phospholipid environment modulates the activation of blood clotting. *J Biol Chem,* 282**,** 6556-63.

SHIMADA, S., SHIMOJIMA, K., MASUDA, T., NAKAYAMA, Y., KOHJI, T., TSUKAMOTO, H., MATSUBASA, T., OKA, A. & YAMAMOTO, T. 2014. MLC1 mutations in Japanese patients with megalencephalic leukoencephalopathy with subcortical cysts. *Hum Genome Var,* 1**,** 14019.

SINGH, N. A. K. P. A. A. U. A. A. M. 2019. Proliposomes: An Approach for the Development of Stable Liposome. *Ars Pharmaceutica,* 60**,** 231-240.

SINGHAL, B. S., GURSAHANI, R. D., UDANI, V. P. & BINIWALE, A. A. 1996. Megalencephalic leukodystrophy in an Asian Indian ethnic group. *Pediatr Neurol,* 14**,** 291-6.

SIRISI, S., ELORZA-VIDAL, X., ARNEDO, T., ARMAND-UGÓN, M., CALLEJO, G., CAPDEVILA-NORTES, X., LÓPEZ-HERNÁNDEZ, T., SCHULTE, U., BARRALLO-GIMENO, A., NUNES, V., GASULL, X. & ESTÉVEZ, R. 2017. Depolarization causes the formation of a ternary complex between GlialCAM, MLC1 and ClC-2 in astrocytes: implications in megalencephalic leukoencephalopathy. *Hum Mol Genet,* 26**,** 2436-2450.

SIRISI, S., FOLGUEIRA, M., LÓPEZ-HERNÁNDEZ, T., MINIERI, L., PÉREZ-RIUS, C., GAITÁN-PEÑAS, H., ZANG, J., MARTÍNEZ, A., CAPDEVILA-NORTES, X., DE LA VILLA, P., ROY, U., ALIA, A., NEUHAUSS, S., FERRONI, S., NUNES, V., ESTÉVEZ, R. & BARRALLO-GIMENO, A. 2014. Megalencephalic leukoencephalopathy with subcortical cysts protein 1 regulates glial surface localization of GLIALCAM from fish to humans. *Hum Mol Genet,* 23**,** 5069-86.

SMITH, G. P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science,* 228**,** 1315-7.

SMITH, S. M. 2017. Strategies for the Purification of Membrane Proteins. *Methods Mol Biol,* 1485**,** 389-400.

SOWA, G. 2012. Caveolae, caveolins, cavins, and endothelial cell function: new insights. *Front Physiol,* 2**,** 120.

SPELIOTES, E. K., WILLER, C. J., BERNDT, S. I., MONDA, K. L., THORLEIFSSON, G., JACKSON, A. U., LANGO ALLEN, H., LINDGREN, C. M., LUAN, J., MÄGI, R., RANDALL, J. C., VEDANTAM, S., WINKLER, T. W., QI, L., WORKALEMAHU, T., HEID, I. M., STEINTHORSDOTTIR, V., STRINGHAM, H. M., WEEDON, M. N., WHEELER, E., WOOD, A. R., FERREIRA, T., WEYANT, R. J., SEGRÈ, A. V., ESTRADA, K., LIANG, L., NEMESH, J., PARK, J. H., GUSTAFSSON, S., KILPELÄINEN, T. O., YANG, J., BOUATIA-NAJI, N., ESKO, T., FEITOSA, M. F., KUTALIK, Z., MANGINO, M., RAYCHAUDHURI, S., SCHERAG, A., SMITH, A. V., WELCH, R., ZHAO, J. H., ABEN, K. K., ABSHER, D. M., AMIN, N., DIXON, A. L., FISHER, E., GLAZER, N. L., GODDARD, M. E., HEARD-COSTA, N. L., HOESEL, V., HOTTENGA, J. J., JOHANSSON, A., JOHNSON, T., KETKAR, S., LAMINA, C., LI, S., MOFFATT, M. F., MYERS, R. H., NARISU, N., PERRY, J. R., PETERS, M. J., PREUSS, M., RIPATTI, S., RIVADENEIRA, F., SANDHOLT, C., SCOTT, L. J., TIMPSON, N. J., TYRER, J. P., VAN WINGERDEN, S., WATANABE, R. M., WHITE, C. C., WIKLUND, F., BARLASSINA, C., CHASMAN, D. I., COOPER, M. N., JANSSON, J. O., LAWRENCE, R. W., PELLIKKA, N., PROKOPENKO, I., SHI, J., THIERING, E., ALAVERE, H., ALIBRANDI, M. T., ALMGREN, P., ARNOLD, A. M., ASPELUND, T., ATWOOD, L. D., BALKAU, B., BALMFORTH, A. J., BENNETT, A. J., BEN-SHLOMO, Y., BERGMAN, R. N., BERGMANN, S., BIEBERMANN, H., BLAKEMORE, A. I., BOES, T., BONNYCASTLE, L. L., BORNSTEIN, S. R., BROWN, M. J., BUCHANAN, T. A., et al. 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet,* 42**,** 937-48.

SPIEGEL, I., ADAMSKY, K., EISENBACH, M., ESHED, Y., SPIEGEL, A., MIRSKY, R., SCHERER, S. S. & PELES, E. 2006. Identification of novel cell-adhesion molecules in peripheral nerves using a signal-sequence trap. *Neuron Glia Biol,* 2**,** 27-38.

SRIRAM, K. & INSEL, P. A. 2018. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol Pharmacol,* 93**,** 251-258.

STAUS, D. P., HU, H., ROBERTSON, M. J., KLEINHENZ, A. L. W., WINGLER, L. M., CAPEL, W. D., LATORRACA, N. R., LEFKOWITZ, R. J. & SKINIOTIS, G. 2020. Structure of the M2 muscarinic receptor-β-arrestin complex in a lipid nanodisc. *Nature,* 579**,** 297-302.

STETSENKO, A. A. G. A. 2017. An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. *Crystals,* 7.

STRAW, M. L., CHAPLIN, A. K., HOUGH, M. A., PAPS, J., BAVRO, V. N., WILSON, M. T., VIJGENBOOM, E. & WORRALL, J. A. R. 2018. A cytosolic copper storage protein provides a second level of copper tolerance in Streptomyces lividans. *Metallomics,* 10**,** 180-193.

SUADES, A., ALCARAZ, A., CRUZ, E., ÁLVAREZ-MARIMON, E., WHITELEGGE, J. P., MANYOSA, J., CLADERA, J. & PERÁLVAREZ-MARÍN, A. 2019. Structural biology workflow for the expression and characterization of functional human sodium glucose transporter type 1 in Pichia pastoris. *Sci Rep,* 9**,** 1203.

SUGIO, S., TOHYAMA, K., OKU, S., FUJIYOSHI, K., YOSHIMURA, T., HIKISHIMA, K., YANO, R., FUKUDA, T., NAKAMURA, M., OKANO, H., WATANABE, M., FUKATA, M., IKENAKA, K. & TANAKA, K. F. 2017. Astrocyte-mediated infantile-onset leukoencephalopathy mouse model. *Glia,* 65**,** 150-168.

SÁNCHEZ, A., GARCÍA-LAREU, B., PUIG, M., PRAT, E., RUBERTE, J., CHILLÓN, M., NUNES, V., ESTÉVEZ, R. & BOSCH, A. 2020. Cerebellar Astrocyte Transduction as Gene Therapy for Megalencephalic Leukoencephalopathy. *Neurotherapeutics,* 17**,** 2041-2053.

SÁNCHEZ-LÓPEZ, V., FERNÁNDEZ-ROMERO, J. M. & GÓMEZ-HENS, A. 2009. Evaluation of liposome populations using a sucrose density gradient centrifugation approach coupled to a continuous flow system. *Anal Chim Acta,* 645**,** 79-85.

TAKAHASHI, R. & IMAI, Y. 2003. Pael receptor, endoplasmic reticulum stress, and Parkinson's disease. *J Neurol,* 250 Suppl 3**,** Iii25-9.

TANG, Y., ZAITSEVA, F., LAMB, R. A. & PINTO, L. H. 2002. The gate of the influenza virus M2 proton channel is formed by a single tryptophan residue. *J Biol Chem,* 277**,** 39880-6.

TAO, Q., FUJIMOTO, J., MEN, T., YE, X., DENG, J., LACROIX, L., CLIFFORD, J. L., MAO, L., VAN PELT, C. S., LEE, J. J., LOTAN, D. & LOTAN, R. 2007. Identification of the retinoic acid-inducible Gprc5a as a new lung tumor suppressor gene. *J Natl Cancer Inst,* 99**,** 1668-82.

TEIJIDO, O., CASAROLI-MARANO, R., KHARKOVETS, T., AGUADO, F., ZORZANO, A., PALACÍN, M., SORIANO, E., MARTÍNEZ, A. & ESTÉVEZ, R. 2007. Expression patterns of MLC1 protein in the central and peripheral nervous systems. *Neurobiol Dis,* 26**,** 532-45.

TEIJIDO, O., MARTÍNEZ, A., PUSCH, M., ZORZANO, A., SORIANO, E., DEL RÍO, J. A., PALACÍN, M. & ESTÉVEZ, R. 2004. Localization and functional analyses of the MLC1 protein involved in megalencephalic leukoencephalopathy with subcortical cysts. *Hum Mol Genet,* 13**,** 2581-94.

TEKOLA-AYELE, F., LEE, A., WORKALEMAHU, T. & SÁNCHEZ-POZOS, K. 2019. Shared genetic underpinnings of childhood obesity and adult cardiometabolic diseases. *Hum Genomics,* 13**,** 17.

TIKHONOVA, I. G. & COSTANZI, S. 2009. Unraveling the structure and function of G protein-coupled receptors through NMR spectroscopy. *Curr Pharm Des,* 15**,** 4003-16.

TOM, R., BISSON, L. & DUROCHER, Y. 2008. Transfection of HEK293-EBNA1 Cells in Suspension with 293fectin for Production of Recombinant Proteins. *CSH Protoc,* 2008**,** pdb.prot4979.

TOMITA, H., ZIEGLER, M. E., KIM, H. B., EVANS, S. J., CHOUDARY, P. V., LI, J. Z., MENG, F., DAI, M., MYERS, R. M., NEAL, C. R., SPEED, T. P., BARCHAS, J. D., SCHATZBERG, A. F., WATSON, S. J., AKIL, H., JONES, E. G., BUNNEY, W. E. & VAWTER, M. P. 2013. G protein-linked signaling pathways in bipolar and major depressive disorders. *Front Genet,* 4**,** 297.

TOPCU, M., SAATCI, I., TOPCUOGLU, M. A., KOSE, G. & KUNAK, B. 1998. Megalencephaly and leukodystrophy with mild clinical course: a report on 12 new cases. *Brain Dev,* 20**,** 142-53.

TOPÇU, M., GARTIOUX, C., RIBIERRE, F., YALÇINKAYA, C., TOKUS, E., OZTEKIN, N., BECKMANN, J. S., OZGUC, M. & SEBOUN, E. 2000. Vacuoliting megalencephalic leukoencephalopathy with subcortical cysts, mapped to chromosome 22qtel. *Am J Hum Genet,* 66**,** 733-9.

TRIBET, C., AUDEBERT, R. & POPOT, J. L. 1996. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A,* 93**,** 15047-50.

TSUTSUMI, N., MUKHERJEE, S., WAGHRAY, D., JANDA, C. Y., JUDE, K. M., MIAO, Y., BURG, J. S., ADURI, N. G., KOSSIAKOFF, A. A., GATI, C. & GARCIA, K. C. 2020. Structure of human Frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical Wnt signaling. *Elife,* 9.

UCHAŃSKI, T., MASIULIS, S., FISCHER, B., KALICHUK, V., LÓPEZ-SÁNCHEZ, U., ZARKADAS, E., WECKENER, M., SENTE, A., WARD, P., WOHLKÖNIG, A., ZÖGG, T., REMAUT, H., NAISMITH, J. H., NURY, H., VRANKEN, W., ARICESCU, A. R., PARDON, E. & STEYAERT, J. 2021. Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nat Methods,* 18**,** 60-68.

UCHAŃSKI, T., PARDON, E. & STEYAERT, J. 2020. Nanobodies to study protein conformational states. *Curr Opin Struct Biol,* 60**,** 117-123.

ULRICH, C. D., 2ND, HOLTMANN, M. & MILLER, L. J. 1998. Secretin and vasoactive intestinal peptide receptors: members of a unique family of G protein-coupled receptors. *Gastroenterology,* 114**,** 382-97.

URNER, L. H., LIKO, I., PAGEL, K., HAAG, R. & ROBINSON, C. V. 2022. Non-ionic hybrid detergents for protein delipidation. *Biochim Biophys Acta Biomembr,* 1864**,** 183958.

VAN DER KNAAP, M. S., BARTH, P. G., STROINK, H., VAN NIEUWENHUIZEN, O., ARTS, W. F., HOOGENRAAD, F. & VALK, J. 1995. Leukoencephalopathy with swelling and a discrepantly mild clinical course in eight children. *Ann Neurol,* 37**,** 324-34.

VAN DER KNAAP, M. S., BARTH, P. G., VRENSEN, G. F. & VALK, J. 1996. Histopathology of an infantile-onset spongiform leukoencephalopathy with a discrepantly mild clinical course. *Acta Neuropathol,* 92**,** 206-12.

VAN DER KNAAP, M. S., BOOR, I. & ESTÉVEZ, R. 2012. Megalencephalic leukoencephalopathy with subcortical cysts: chronic white matter oedema due to a defect in brain ion and water homoeostasis. *Lancet Neurol,* 11**,** 973-85.

VAN DER KNAAP, M. S. & BUGIANI, M. 2017. Leukodystrophies: a proposed classification system based on pathological changes and pathogenetic mechanisms. *Acta Neuropathol,* 134**,** 351-382.

VAN DER KNAAP, M. S., LAI, V., KÖHLER, W., SALIH, M. A., FONSECA, M. J., BENKE, T. A., WILSON, C., JAYAKAR, P., AINE, M. R., DOM, L., LYNCH, B., KÁLMÁNCHEY, R., PIETSCH, P., ERRAMI, A. & SCHEPER, G. C. 2010. Megalencephalic leukoencephalopathy with cysts without MLC1 defect. *Ann Neurol,* 67**,** 834-7.

VANDERVER, A., PRUST, M., TONDUTI, D., MOCHEL, F., HUSSEY, H. M., HELMAN, G., GARBERN, J., EICHLER, F., LABAUGE, P., AUBOURG, P., RODRIGUEZ, D., PATTERSON, M. C., VAN HOVE, J. L., SCHMIDT, J., WOLF, N. I., BOESPFLUG-TANGUY, O., SCHIFFMANN, R. & VAN DER KNAAP, M. S. 2015.

Case definition and classification of leukodystrophies and leukoencephalopathies. *Mol Genet Metab,* 114**,** 494-500.

VARADI, M., ANYANGO, S., DESHPANDE, M., NAIR, S., NATASSIA, C., YORDANOVA, G., YUAN, D., STROE, O., WOOD, G., LAYDON, A., ŽÍDEK, A., GREEN, T., TUNYASUVUNAKOOL, K., PETERSEN, S., JUMPER, J., CLANCY, E., GREEN, R., VORA, A., LUTFI, M., FIGURNOV, M., COWIE, A., HOBBS, N., KOHLI, P., KLEYWEGT, G., BIRNEY, E., HASSABIS, D. & VELANKAR, S. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res,* 50**,** D439-d444.

VENKATAKRISHNAN, A. J., DEUPI, X., LEBON, G., TATE, C. G., SCHERTLER, G. F. & BABU, M. M. 2013. Molecular signatures of G-protein-coupled receptors. *Nature,* 494**,** 185-94.

VISCHER, H. F., CASTRO, M. & PIN, J. P. 2015. G Protein-Coupled Receptor Multimers: A Question Still Open Despite the Use of Novel Approaches. *Mol Pharmacol,* 88**,** 561-71.

VON SAMSON-HIMMELSTJERNA, F. A., FREUNDT, G., NITZ, J. T., STELTER, F., LUEDDE, M., WIELAND, T., FREY, N. & HIPPE, H. J. 2019. The orphan receptor GPRC5B modulates inflammatory and fibrotic pathways in cardiac fibroblasts and mice hearts. *Biochem Biophys Res Commun,* 514**,** 1198-1203.

WAN, Q., OKASHAH, N., INOUE, A., NEHMÉ, R., CARPENTER, B., TATE, C. G. & LAMBERT, N. A. 2018. Mini G protein probes for active G protein-coupled receptors (GPCRs) in live cells. *J Biol Chem,* 293**,** 7466-7473.

WENTINCK, K., GOGOU, C. & MEIJER, D. H. 2022. Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins. *Curr Res Struct Biol,* 4**,** 332-337.

WU, S., AVILA-SAKAR, A., KIM, J., BOOTH, D. S., GREENBERG, C. H., ROSSI, A., LIAO, M., LI, X., ALIAN, A., GRINER, S. L., JUGE, N., YU, Y., MERGEL, C. M., CHAPARRO-RIGGERS, J., STROP, P., TAMPÉ, R., EDWARDS, R. H., STROUD, R. M., CRAIK, C. S. & CHENG, Y. 2012. Fabs enable single particle cryoEM studies of small proteins. *Structure,* 20**,** 582-92.

WU, X. & RAPOPORT, T. A. 2021. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc Natl Acad Sci U S A,* 118.

YALÇINKAYA, C., YÜKSEL, A., COMU, S., KILIÇ, G., COKAR, O. & DERVENT, A. 2003. Epilepsy in vacuolating megalencephalic leukoencephalopathy with subcortical cysts. *Seizure,* 12**,** 388-96.

YANG, H. J., VAINSHTEIN, A., MAIK-RACHLINE, G. & PELES, E. 2016. G protein-coupled receptor 37 is a negative regulator of oligodendrocyte differentiation and myelination. *Nat Commun,* 7**,** 10884.

YANG, Y., LIU, X. R., GREENBERG, Z. J., ZHOU, F., HE, P., FAN, L., LIU, S., SHEN, G., EGAWA, T., GROSS, M. L., SCHUETTPELZ, L. G. & LI, W. 2020. Open conformation of tetraspanins shapes interaction partner networks on cell membranes. *Embo j,* 39**,** e105246.

YAO, X., FAN, X. & YAN, N. 2020. Cryo-EM analysis of a membrane protein embedded in the liposome. *Proc Natl Acad Sci U S A,* 117**,** 18497-18503.

YEH, V., GOODE, A. & BONEV, B. B. 2020. Membrane Protein Structure Determination and Characterisation by Solution and Solid-State NMR. *Biology (Basel),* 9.

YU, J., ZHU, H., LAPE, R., GREINER, T., DU, J., LÜ, W., SIVILOTTI, L. & GOUAUX, E. 2021. Mechanism of gating and partial agonist action in the glycine receptor. *Cell,* 184**,** 957-968.e21.

YUAN, Y., KONG, F., XU, H., ZHU, A., YAN, N. & YAN, C. 2022. Cryo-EM structure of human glucose transporter GLUT4. *Nat Commun,* 13**,** 2671.

YUDIN, Y. & ROHACS, T. 2018. Inhibitory G(i/O)-coupled receptors in somatosensory neurons: Potential therapeutic targets for novel analgesics. *Mol Pain,* 14**,** 1744806918763646.

YÜZBAŞIOĞLU, A., TOPÇU, M., CETIN KOCAEFE, Y. & OZGÜÇ, M. 2011. Novel mutations of the MLC1 gene in Turkish patients. *Eur J Med Genet,* 54**,** 281-3.

ZENG, Z., SU, K., KYAW, H. & LI, Y. 1997. A novel endothelin receptor type-B-like gene enriched in the brain. *Biochem Biophys Res Commun,* 233**,** 559-67.

ZHANG, K., WU, H., HOPPE, N., MANGLIK, A. & CHENG, Y. 2022. Fusion protein strategies for cryo-EM study of G protein-coupled receptors. *Nat Commun,* 13**,** 4366.

ZHANG, S., LI, N., ZENG, W., GAO, N. & YANG, M. 2017a. Cryo-EM structures of the mammalian endo-lysosomal TRPML1 channel elucidate the combined regulation mechanism. *Protein Cell,* 8**,** 834-847.

ZHANG, Y., SUN, B., FENG, D., HU, H., CHU, M., QU, Q., TARRASCH, J. T., LI, S., SUN KOBILKA, T., KOBILKA, B. K. & SKINIOTIS, G. 2017b. Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature,* 546**,** 248-253.

ZHAO, L. H., MA, S., SUTKEVICIUTE, I., SHEN, D. D., ZHOU, X. E., DE WAAL, P. W., LI, C. Y., KANG, Y., CLARK, L. J., JEAN-ALPHONSE, F. G., WHITE, A. D., YANG, D., DAI, A., CAI, X., CHEN, J., LI, C., JIANG, Y., WATANABE, T., GARDELLA, T. J., MELCHER, K., WANG, M. W., VILARDAGA, J. P., XU, H. E. & ZHANG, Y. 2019. Structure and dynamics of the active human parathyroid hormone receptor-1. *Science,* 364**,** 148-153.

ZOONENS, M. & POPOT, J. L. 2014. Amphipols for each season. *J Membr Biol,* 247**,** 759-96.

ANNEX:

PUBLICATIONS

International Journal of
Molecular Sciences

MDPI

Review

GPR37 Receptors and Megalencephalic Leukoencephalopathy with Subcortical Cysts

Adrià Pla-Casillanis [1,†], Laura Ferigle [1,†], Marta Alonso-Gardón [1], Efren Xicoy-Espaulella [1], Ekaitz Errasti-Murugarren [2,3], Daniela Marazziti [4] and Raúl Estévez [1,3,*]

[1] Unitat de Fisiologia, Departament de Ciències Fisiològiques, Genes Disease and Therapy Program IDIBELL-Institute of Neurosciences, Universitat de Barcelona, L'Hospitalet de Llobregat, 08907 Barcelona, Spain; adria.pla9@gmail.com (A.P.-C.); laura.ferigle@ub.edu (L.F.); malonsga@gmail.com (M.A.-G.); efrenxicoyespaulella@gmail.com (E.X.-E.)
[2] Unitat de Genètica, Departament de Ciències Fisiològiques, Genes Disease and Therapy Program IDIBELL-Institute of Neurosciences, Universitat de Barcelona, L'Hospitalet de Llobregat, 08907 Barcelona, Spain; ekaitz_errasti@ub.edu
[3] Centro de Investigación en Red de Enfermedades Raras (CIBERER), ISCIII, 28029 Madrid, Spain
[4] Institute of Biochemistry and Cell Biology, Italian National Research Council (CNR), I-00015 Rome, Italy; daniela.marazziti@cnr.it
* Correspondence: restevez@ub.edu; Tel.: +34-944-039-781; Fax: +34-934-024-268
† These authors contributed equally to this work.

Abstract: Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare type of vacuolating leukodystrophy (white matter disorder), which is mainly caused by defects in MLC1 or glial cell adhesion molecule (GlialCAM) proteins. In addition, autoantibodies to GlialCAM are involved in the pathology of multiple sclerosis. *MLC1* and *GLIALCAM* genes encode for membrane proteins of unknown function, which has been linked to the regulation of different ion channels and transporters, such as the chloride channel VRAC (volume regulated anion channel), ClC-2 (chloride channel 2), and connexin 43 or the Na^+/K^+-ATPase pump. However, the mechanisms by which MLC proteins regulate these ion channels and transporters, as well as the exact function of MLC proteins remain obscure. It has been suggested that MLC proteins might regulate signalling pathways, but the mechanisms involved are, at present, unknown. With the aim of answering these questions, we have recently described the brain GlialCAM interactome. Within the identified proteins, we could validate the interaction with several G protein-coupled receptors (GPCRs), including the orphan GPRC5B and the proposed prosaposin receptors GPR37L1 and GPR37. In this review, we summarize new aspects of the pathophysiology of MLC disease and key aspects of the interaction between GPR37 receptors and MLC proteins.

Keywords: megalencephalic leukoencephalopathy with subcortical cysts; GlialCAM; MLC1; GPRC5B; GPR37L1; GPR37; glia; ion; water homeostasis

check for updates

Citation: Pla-Casillanis, A.; Ferigle, L.; Alonso-Gardón, M.; Xicoy-Espaulella, E.; Errasti-Murugarren, E.; Marazziti, D.; Estévez, R. GPR37 Receptors and Megalencephalic Leukoencephalopathy with Subcortical Cysts. *Int. J. Mol. Sci.* **2022**, *23*, 5528. https://doi.org/10.3390/ijms23105528

Academic Editor: Ross Bathgate

Received: 21 April 2022
Accepted: 13 May 2022
Published: 16 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare genetic type of leukodystrophy (OMIM 604004). It is a childhood-onset hereditary disease characterized by white matter vacuolation and macrocephaly, which is developed during the first year of life. Most MLC patients present a progressive loss of motor functions with ataxia and spasticity, cognitive decline and epileptic seizures [1]. Symptomatology often worsens after fever or mild head trauma [2].

Magnetic Resonance Imaging (MRI), which is used for diagnostics, shows that patients display diffuse signal abnormality and swelling of the cerebral white matter together with the presence of subcortical cysts, mainly in the anterior temporal regions [3]. Furthermore, MRI together with the histopathology of the brain of an MLC patient revealed

that myelin contains water-filled vacuoles [4]. Up to date, MLC has no cure. Treatment is symptomatic in combination with supportive care.

Two different phenotypes of MLC disease have been described: a classical and a remitting phenotype [5]. The classical phenotype is the most commonly found in patients. It is caused by autosomal recessive mutations in either the *MLC1* or the hepatic cell adhesion molecule (*HEPACAM*) genes, resulting in two disease subtypes, namely MLC1 or MLC2A. The remitting phenotype, named MLC2B, is caused by dominant mutations in *HEPACAM* [6]. Approximately 76% of patients present mutations in *MLC1*, 22% in *HEPACAM* and 2% of MLC cases cannot be explained by mutations in these two genes, suggesting that others might be implied in the disease [2].

MLC1 encodes for a membrane protein with eight predicted transmembrane domains whose function remains unknown. It is expressed exclusively in the brain in astrocytes surrounding blood vessels and Bergmann glia in the cerebellum [7]. More than 50 mutations have been described for *MLC1*, including missense, deletions, insertions and nonsense mutations [1].

On the other hand, *HEPACAM* encodes for a cell adhesion molecule of the immunoglobulin (Ig) family named GlialCAM, which is expressed predominantly in neurons, astrocytes and oligodendrocytes [6]. GlialCAM was first identified in hepatic cancer, where it was downregulated, but it is predominantly expressed in glial cells [8]. GlialCAM can form interactions with other GlialCAM molecules in *cis* (within the same cell) or *trans* (between different cells) [9]. GlialCAM acts as an endoplasmic reticulum (ER) chaperone for MLC1 [10] and it also helps MLC1 to reach astrocyte-astrocyte junctions where both proteins co-localize [6,11]. Apart for being involved in MLC, recently, it has been described that autoantibodies recognizing GlialCAM might be involved in the development of multiple sclerosis [12].

The pathophysiological mechanisms leading to MLC are still unclear [13]. Even though the function of the MLC1/GlialCAM complex is unknown, it has been hypothesized that it may have a role in the regulation of ion/water homeostasis, as it interacts with different transporters and ion channels. Thus, it has been shown that the complex interacts directly with the chloride channel 2 (ClC-2) [14], the gap junction alpha 1 protein (connexin 43, Cx43) [15–17], Na^+/K^+-ATPase [18,19], and it is thought to regulate indirectly the activity of volume-regulated anion channel (VRAC) [10,20,21] or the calcium-permeable channel TRPV4 [22]. Recent bioinformatics developments have suggested new potential functions for the GlialCAM/MLC1 complex (see Section 2).

2. Novel Insights into GlialCAM/MLC1 Function by Alphafold Structural Models

Powerful sequence-alignment methodologies have been used to identify possible functional MLC1 homologues. The underlying assumption of these sequence-based methods is that proteins with similar sequences adopt a similar fold. As protein folding determines function, the detection of evolutionary relationships between proteins can be used to predict functions of non-annotated protein sequences. Thus, MLC1 sequence identity analysis by BLAST/PSI-BLAST algorithm identified the voltage gated potassium channel Kv1.1 alpha subunit (KCNA1) as the protein with the highest sequence identity (less than 20% amino acid identity) [23], suggesting that MLC1 could act as an ion channel. In agreement with this hypothesis, it is known that MLC1 can form homo-oligomers, a characteristic found in many ion channels proteins [24]. Moreover, MLC patients may present epileptic seizures, which is a common feature in ion channel diseases, but not in leukodystrophies [25]. Taking all these data into consideration, the first hypothesis regarding MLC1 function was that it could act as an ion channel. Nevertheless, voltage-clamp measurements in *Xenopus* oocytes did not detect any changes in the conductance and neither did patch clamp measurements in HEK293T or HeLa transfected cells. For those experiments, different pulse protocols and various pulse durations were applied and no conductivity of MLC1 was detected, even in the presence of GlialCAM [26].

The functional characterization of MLC1 is necessary to understand its physiological relevance. As the 3D structure of a protein is believed to be responsible for its biological function, protein structure can provide a better insight into which protein fragments contribute most to the functionality of a protein compared to the primary sequence. In fact, residues located far apart in primary sequence may be close in the 3D structure. Recently reported MLC1 structural models [27,28] suggest a protein fold based on a 4 + 4 structural repeat, involving transmembrane regions (TMs) 1–4 and 5–8 (Figure 1A,B). Interestingly, although mutations found in MLC patients are spread through all the protein sequence, a significant number of identified missense MLC1 mutations are mainly localized on the interphase between MLC1 internal repeats (Figure 1C). As MLC1 mutations have been reported to cause protein instability, which consequently causes its degradation at the endoplasmic reticulum or lysosomes [29–31], we propose that proper interaction between MLC1 internal repeats is of key relevance for proper protein folding and endoplasmic reticulum sorting. Similarly, recently reported MLC1 homo-trimeric complex in detergent micelles and proteoliposomes [24], suggests that mutations affecting MLC1 monomers interaction would also affect protein stability. However, the lack of a 3D structure at atomic resolution precludes the identification of putative protein-protein interaction regions.

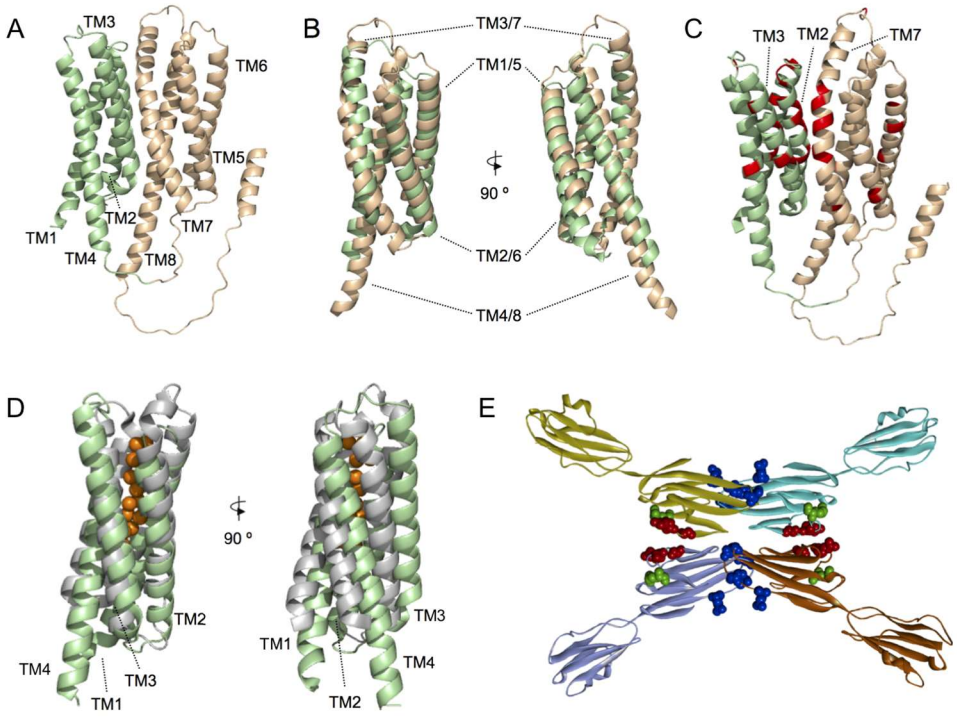


Figure 1. Structural models of MLC1 and GlialCAM proteins. (**A**) Alphafold structural model of monomeric MLC1. Helices 1 to 4 and 5 to 8 are coloured pale green and wheat, respectively. (**B**) Structural superimposition of the two MLC1 subdomains, comprised of TMs 1 to 4 (pale green) and 5 to 8 (wheat), respectively. Two different views, rotated by 90°, are shown. (**C**) MLC causing mutations (red) depicted in the MLC1 structural model. Preferential localization of mutated residues is located in the interaction surface defined by the two MLC1 subdomains (TMs 1 to 4 (pale green) and 5 to 8 (wheat)). (**D**) Structural superimposition of MLC1 structural model TM1 to 4 (pale green) and a prokaryotic soluble copper storage protein from *Streptomyces lividans* (light grey) (PDB ID 6Q6B). Two different views, rotated by 90°, are shown. (**E**) Summary of the structural model proposed for GlialCAM homodimers forming *cis* and *trans* interaction through different surfaces of its IgV domain. *Cis* dimerization is achieved by interactions between two opposing beta-strands of the IgV domain and *trans* interactions occur between salient loops of both IgV domains. Residues mutated in MLC2A patients (recessive) are shown in green, Residues mutated in MLC2B patients are shown in red or blue (affecting blue the interactions in *trans* only).

The growing number of known protein structures and structural models has motivated the interest into developing approaches that aim to identify the protein functions from the structure. Among them, methods comparing the global fold of a query protein with other template proteins of known function are commonly used. Identification of proteins similarly folded to whole MLC1 and TM1-4 MLC1, using the PDBeFold algorithm [32], revealed both prokaryotic and eukaryotic proteins with a similar 3D fold. Particularly, PDBeFold analysis of the TM1-4 MLC1 subdomain resulted in the identification of a prokaryotic soluble copper storage protein from *Streptomyces lividans* (PDB ID 6Q6B) (Figure 1D), suggesting that maybe MLC1 would be involved in ion sensing. However, further research should be conducted to solve this issue.

In addition, based on the identification of the interactome of GlialCAM and MLC1, we recently proposed that MLC1 could act as a tetraspanin-like molecule [16]. Tetraspanins are cell-surface proteins with four transmembrane domains, which can homo- and hetero-oligomerize. They participate in a wide variety of cellular processes such as adhesion, differentiation and cell activation [33]. In many cases, they also form a tight complex with proteins belonging to the Ig superfamily, as it happens with MLC1 and GlialCAM.

Our recent work based on 3D models combined with biochemical analyses indicated that GlialCAM forms interactions in *cis* through an interaction surface comprising residues Glutamate 86 to Arginine 92 in the IgV domain [9] (Figure 1E). Furthermore, two other loops might also be involved in the formation of *trans*- interactions, which are blocked by a nanobody generated against GlialCAM. Residues causing dominant MLC are located in these two interaction surfaces (Figure 1E). It remains to be determined whether the IgC2 domain contributes to the formation of lateral interactions as it has been observed in other Ig proteins, or if it also mediates interaction with the MLC1 protein.

3. Physiological Processes Affected in MLC Deduced from Animal Models of MLC

The main characteristic observed in MLC patients is that the brain is enlarged with increased water accumulation. Water is accumulated in subcortical cysts but also, in a general manner, in the outer part of myelin [4] and in astrocytes surrounding blood vessels [34]. Similarly, the lack of MLC1 or GlialCAM in knockout (KO) mice [18,35–37] or KO zebrafish [7,38] leads to an increased brain water content and cerebellar white matter vacuolation, although no cysts are observed. Moreover, in humans, the oedema is localized mainly in the subcortical white matter, while vacuoles are found in the cerebellum of KO mice. Although vacuoles of water in myelin sheaths that enwrap axons of central neurons can be observed in both patient biopsies and mice samples, still the developmental myelination process is not altered. Considering astrocytes, histological studies revealed that *Mlc1* KO mice presented abnormal astrocytes with thicker cell processes, swollen cell bodies and enlarged end-feet, although the length and number of these cells were not altered [34]. Primary astrocytes from *Mlc1* KO showed an increase in the number of vacuoles [7]. In summary, although there are differences between human and animal models, they show in common an increase in the content of water in the brain, which is accumulated in the form of vacuoles in myelin or in astrocytes.

Why does the lack of MLC1 lead to an increased water content in the brain? The first hypothesis was that MLC1 functions as a water or ion channel, and its defect causes water accumulation due to osmotic alterations linked to neuronal activity [23]. However, this activity has not been detected after expressing the MLC1 protein alone or together with GlialCAM in different cell systems [26]. Thus, a search for other proteins interacting with MLC1 and/or GlialCAM to find a connection with the patient's brain phenotype was developed. Unsurprisingly, several transporters and ion channels were identified by different groups (for review [13]). In an extensive proteomic analysis [16], we have recently identified the following interactor proteins: the chloride channel ClC-2, the gap junction protein Cx43, the glutamate transporter EAAT1/2, the alpha2 and beta2 subunits of the sodium/potassium ATPase, the sodium bicarbonate transporter NBCe1, the glucose transporter GLUT1, and the sodium calcium exchanger. Furthermore, although not appearing as MLC interacting

proteins, other activities have been shown to be affected by the lack of MLC1 such as the volume regulated anion channel (VRAC) or the calcium-permeable channel TRPV4. It has been clearly demonstrated that the activity of the chloride channels ClC-2 [37] and VRAC [21,35] and of the gap junction protein Cx43 [15] are affected in vivo. Thus, one possible hypothesis is that the lack of activity of some of these proteins is the reason that explains myelin vacuolization. In agreement with these hypotheses, a *Clcn2* KO mouse model displayed widespread vacuolization, including the cerebellum [37]. Similarly, defects of connexins have also been linked to myelin vacuolization [39]. As LRRC8 proteins of VRAC channels are involved in cell volume regulation and their absence cause cell vacuolation in several tissues, it will be important to analyse whether a conditional (because the complete KO is deleterious [40]) astrocyte KO of LRRC8A [41] (the main constituent of the VRAC channel [42,43]) also displays astrocyte and myelin vacuolization.

In contrast to other leukodystrophies, MLC patients can suffer from seizures [25]; still the underlying mechanism that links MLC to epilepsy is not known. In this sense, both *Glialcam* and *Mlc1* KO mice present an abnormal extracellular K^+ dynamics and neuronal network activity, as they had an epileptiform brain activity and a lowered seizure threshold [44]. As described previously, defects in either the MLC1 or GlialCAM interacting proteins' function might also affect extracellular potassium dynamics. Astrocyte uptake is mainly mediated by the Na^+/K^+ ATPase pump [45], an interactor of MLC1 [16,18,19]. In *Mlc1* KO mice, there is a reduced expression of the inward rectifier potassium channel Kir4.1 involved in potassium clearance [35], whose absence is known to lead to hyperexcitability and epilepsy [46]. The interaction of GlialCAM with the ClC-2 chloride channel increases total current amplitudes and abolishes the rectification of ClC-2, which is thus opened at positive voltages [14]. In addition, this interaction opens the ClC-2 common gate, which is closed by acidic pH. This reduction in the inward rectification will allow the influx of chloride, which may be needed to maintain the electroneutrality after potassium intake in glial cells [47]. Connexins are also essential to disperse local high potassium concentrations through a glial syncytium. It must be considered that disturbed astrocyte regulation of water homeostasis in MLC affecting the VRAC channel might also cause hyperexcitability of neuronal networks and seizures.

Importantly, perivascular astrocytes by itself, where MLC1 expression is higher and can even be considered a marker of these cells [48], are essential to the maintenance of an adequate blood brain barrier, which is also important for the clearance of potassium [49]. In this sense, it has recently been shown that the lack of MLC1 affects the perivascular astrocytic processes' molecular maturation and organization. In *Mlc1* KO mice, an accumulation of fluid in the brain occurs, although this does not alter the blood-brain barrier integrity and neither the organization of the endothelial network. It has been determined that MLC1 might play a role in contractile maturation of vascular smooth muscle cells, arterial perfusion, and neurovascular coupling. Its absence disturbs the postnatal acquisition of contractile properties by vascular smooth muscle cells and disrupts blood perfusion, vessel diameter and neurovascular coupling [50]. Thus, not only affecting different ion channels and transporters in astrocytes can cause water increase and it might be that disruption of gliovascular unit by itself might contribute also to the increase in water content and seizures observed in MLC patients.

Another important aspect to consider in the pathophysiology of MLC is that MLC1 is expressed in astrocytes, whereas its auxiliary subunit GlialCAM is expressed in astrocytes, oligodendrocytes, and neurons [15,37]. Importantly, GlialCAM regulates astrocyte competition for territory and morphological complexity in the developing mouse cortex [15]. It has been shown that the lack of MLC1 causes GlialCAM mislocalization by an unknown mechanism in astrocytes [7], but also in oligodendrocytes [37], possibly because astrocytic GlialCAM interacts in *trans* through the IgV domain with oligodendrocytic GlialCAM. Moreover, the mislocalization of GlialCAM also affect to the localization of ClC-2 in oligodendrocytes, as shown by immunofluorescence and electrophysiological measurements of ClC-2 activity in oligodendrocytes on cerebellar slices [37]. In line with these studies, it has

been recently shown that the lack of GlialCAM in astrocytes in vivo decreases synaptic inhibition and, therefore, increases excitation, which may also explain seizures [15]. As ClC-2 and GlialCAM are also expressed in inhibitory synapses, this change in synaptic function might also be related to disorganization of GlialCAM (astrocyte)-GlialCAM (neuron) trans contacts. It could also be related to the mislocalization of other interacting partners such as Cx43, which in turn might also interact with connexins present in oligodendrocytes (connexin 47) [51] and/or neurons.

As previously stated in this review, the functional role of the GlialCAM/MLC1 complex is still unknown. Nevertheless, it has been described that several proteins and processes related to brain homeostasis are affected in a GlialCAM or MLC1-dependent manner and it is not clear how GlialCAM and MLC1 exert this effect on the activity of different ion channels and transporters. A suggested hypothesis argues that they might influence signalling cascades by yet undefined mechanisms [52], which may regulate these channels or transporters. In this regard, GlialCAM and MLC1 have been related to signal transduction changes. For instance, it has been described that the overexpression of human MLC1 in astrocytes decreases the phosphorylation of extracellular signal-regulated kinases (ERK), whereas primary astrocytes lacking MLC1 show an increase in phosphorylation [21]. Nonetheless, the mechanisms involved in this process remain unresolved.

In summary, GlialCAM and MLC1 seem to regulate the activity of many different transporters and ion channels in different cell types (Figure 2), possibly by regulating phosphorylation of these proteins. An anomalous activity of these proteins might contribute to the defects observed in the regulation of the extracellular water and ionic homeostasis, which could explain the increased water content and seizure susceptibility of MLC patients.

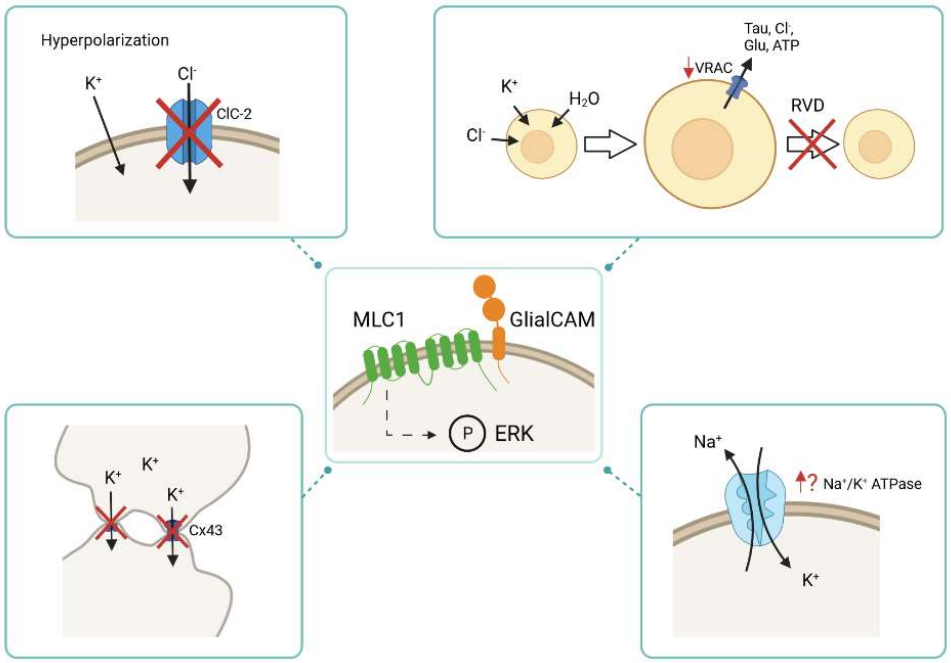


Figure 2. Physiological alterations caused by the lack of MLC1 and/or GlialCAM. MLC1/GlialCAM might regulate the activity of different transporters and ion channels through phosphorylation of ERK signalling transduction cascade. In the absence of MLC proteins, the activity of different transporters and ion channels is altered, such as the activity of ClC-2, VRAC, Cx43 and Na^+/K^+ ATPase. During the hyperpolarization phase of a neuronal action potential, the activity of the chloride channel ClC-2 is decreased in the absence of MLC1/GlialCAM. The lack of MLC proteins diminishes VRAC activity, which leads to an impaired regulatory volume decrease (RVD) response: an important mechanism to shrink cells after cell swelling. The Cx43 localization in the absence of MLC1/GlialCAM is affected as it is internalized and it is no longer located at cell-cell junctions. It is thought that the activity of the Na^+/K^+ ATPase is altered when MLC proteins are not present.

4. Regulation Mechanism of Different Transporters and Ion Channels by MLC Proteins: Is a GPCR Link the Answer?

Recently, our research group identified the GlialCAM interactome through an approach based on affinity purifications (APs) [16]. Four different antibodies specific for GlialCAM were used on samples consisting of membrane fractions prepared from whole brains from adult animals. These included wild type (WT) rats and mice, as well as *Glialcam* KO mice. The previously validated interactors MLC1, ClC-2 and GlialCAM itself were retained in all APs with high efficiency, reinforcing the robustness of this approach. We could identify as many as 21 proteins within the GlialCAM interactome in the rodent brain, some of them already linked to MLC as mentioned above. Within the proteins identified as part of the network of GlialCAM, there were three specific G protein-coupled receptors (GPCRs). Specifically, we retrieved the orphan receptors GPRC5B [53], which we suggested that mutations in this gene could be found in MLC patients without mutations in *MLC1* and *HEPACAM*, and the proposed binders of prosaposin GPR37 and GPR37L1 [54,55]. It is interesting to note that the latter two GPCRs belong to the same protein family. Therefore, we proceeded to determine the potential interaction and the relationship between these GPCRs and MLC-related proteins (GlialCAM/MLC1 and ClC-2). As the purpose of this review is to update the knowledge of the GPR37 family, we will focus on these two proteins.

The GPR37 and GPR37L1 proteins are part of class A rhodopsin-like family of GPCRs, which comprises 80% of all identified GPCRs [56]. Both are considered orphan receptors as no ligand has conclusively been linked to them in vitro [55]. The two GPCRs are widely expressed in the CNS [57]. GPR37 is mainly expressed in the cerebellum, corpus callosum, medulla, putamen, caudate nucleus, substantia nigra and the hippocampus [58–61]. Specifically, oligodendrocytes are the cell type displaying a higher expression of GPR37 together with certain subsets of neurons like dopaminergic neurons in the substantia nigra [62]. On the other hand, GPR37L1 is exclusively expressed in glial cells within the brain, in particular Bergmann glia astrocytes in the cerebellum [63], as well as immature oligodendrocytes [64].

Studies carried out in *Gpr37* KO mice revealed that GPR37 acts as a negative regulator of oligodendrocyte differentiation and myelination. The lack of GPR37 leads to premature differentiation of pre-myelinating oligodendrocytes to myelin producing cells. This alteration is a cause of central nervous system hypermyelination in these mice, already at a very young age and up until adult stages [64]. Furthermore, experiments performed both in primary oligodendrocytes and in brain-derived samples obtained from *Gpr37* KO mice show an increase in ERK1/2 phosphorylation. Pharmacological inhibition of MEK1/2 and ERK1/2 seemed to stop premature differentiation of oligodendrocytes in the absence of GPR37. This inhibition, however, did not affect normal cell proliferation [65]. In addition, adenylyl cyclase inhibition resulted in the impairment of ERK1/2 translocation to the nucleus. Taken together, these data suggested that this pathway is indeed responsible for GPR37 activation during oligodendrocyte differentiation.

Regarding GPR37L1, several studies have highlighted its relevance in the developing brain. One study showed that the lack of Gpr37l1 led to altered cerebellar development in mice [63]. The animals displayed a reduction in neuronal granule cell precursors together with premature maturation of Bergmann glia and Purkinje neurons. Cerebellar layer formation was also altered. However, the authors observed that KO mice seemed to perform better in motor tasks. Motor learning was improved both in juvenile and adult stages, while the adult animals also showed better coordination skills. Furthermore, it has been suggested that Gpr37l1 could play a role in recovery after ischemic injuries, possibly by modulating glutamate transporters [66], which also form part of the GlialCAM interactome [16].

5. Biochemical Relationship between the GPR37/GPR37L1 and MLC Proteins

As detailed above, these two GPCRs, that were identified as members of the GlialCAM interactome, play an important role in mediating a variety of processes in the central

nervous system. Hence, our research group was particularly interested in the analysis of the potential role of these GPCRs in MLC pathophysiology. The first experiments that were carried out aimed to validate the proteomics data regarding the interaction between these proteins and GlialCAM/MLC1 [16]. As MLC1 and GPR37L1 are both expressed in astrocytes, we initially focused our studies on this GPCR. We could establish that there was co-localization between GPR37L1 and MLC1 by immunofluorescence in mouse primary astrocytes. Furthermore, both proteins were in proximity in the same cells, as assessed by Proximity Ligation Assay (PLA) [16]. Moreover, the ability of each of the two GPCRs to directly interact with either MLC1 or GlialCAM was monitored by split-tobacco etch virus (TEV) assays for GPR37 (Figure 3) and bioluminescence resonance energy transfer (BRET) studies in HEK293T cells for GPR37L1 [16].

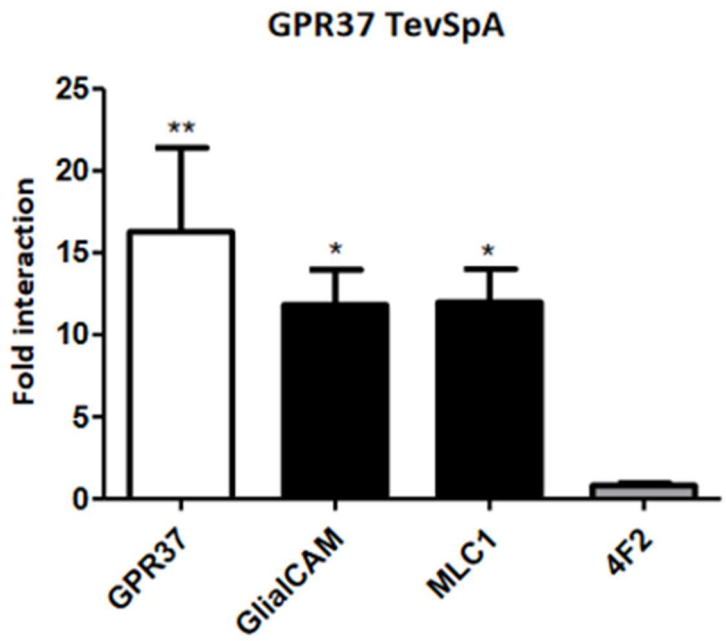


Figure 3. Direct Interaction between GPR37 and GlialCAM/MLC1. Results of split-TEV interaction assays in HeLa cell using GPR37 C-terminally tagged with the N-terminal part of the TEV protease (TevSpA). They were co-transfected with different constructs (GPR37, GlialCAM, MLC1 and 4F2 as a negative control) C-terminally tagged with the C-terminal part of the TEV protease. Plotted data combine the results from three independent experiments. Statistical significance was obtained comparing the interaction within each group to the interaction with the negative control in Bonferroni's multiple comparison test (* $p < 0.05$, ** $p < 0.01$ in the test versus 4F2). Results show that GPR37 is able to homo-oligomerize and hetero-oligomerize with GlialCAM and MLC1, in the same manner as shown for GPR37L1 by BRET.

The above summarized results support the formation of complexes between the GPCRs and MLC-related proteins in living cells. More work was carried out to further characterize the nature of the relationship between these proteins and the physiological role of these complexes. In collaboration with the group led by Daniela Marazziti, we started to address these questions using the *Gpr37l1* constitutive KO mice [63]. As the animals showed no alteration of adult cerebellar layer cytoanatomy and organization and there was no apparent sign of gliosis, it was considered that the analysis of MLC proteins in *Gpr37l1* KO mice could elucidate direct effects of GPR37L1 on MLC protein expression and function.

Biochemically, the first step was to analyse the consequences of the lack of GPR37L1 on MLC1 and GlialCAM protein levels [16]. Western blot experiments of cerebellar membrane fractions indicated that both proteins were upregulated in the *Gpr37l1* KO samples. Likewise, ClC-2 protein levels were also increased. Similarly, in immunofluorescence la-

belling experiments of MLC1 and GlialCAM in cerebellar Bergmann glia samples, an increased signal was observed. These in tissue results were consistent with what was observed in primary astrocytes derived from these mice, in which an increased signal for MLC-related proteins was revealed. However, the MLC proteins and ClC-2 showed a more dotted pattern compared to the WT signal. To determine whether MLC1 subcellular localization was altered in *Gpr37l1* KO mice, we proceeded to detect MLC1 by electron microscopy (EM) immunogold experiments. These experiments showed that the localization of MLC1 in Bergmann glia or in perivascular astrocytic processes was not affected. In summary, we concluded that the lack of GPR37L1 in mice upregulates MLC protein levels without altering their localization [16]. No studies have so far been performed in *Gpr37* KO mice.

6. Possible Role of GPR37 and GPR37L1 in MLC Pathophysiology

As the lack of GPR37L1 increases MLC proteins levels, it was suggested that GPR37L1 might be a negative regulator of MLC proteins in astrocytes (Figure 4). Interestingly, previous studies [67] have suggested that GPR37 interacts and negatively regulates dopamine transporters by regulating their endocytosis and trafficking. As stated previously, GPR37 has been described to act as a negative regulator of myelin formation [64]. As the proper expression of GlialCAM/MLC1 is necessary for myelin homeostasis, we could hypothesize that GPR37L1 would exert a similar negative effect regulating the GlialCAM/MLC1 complex during development (Figure 4).

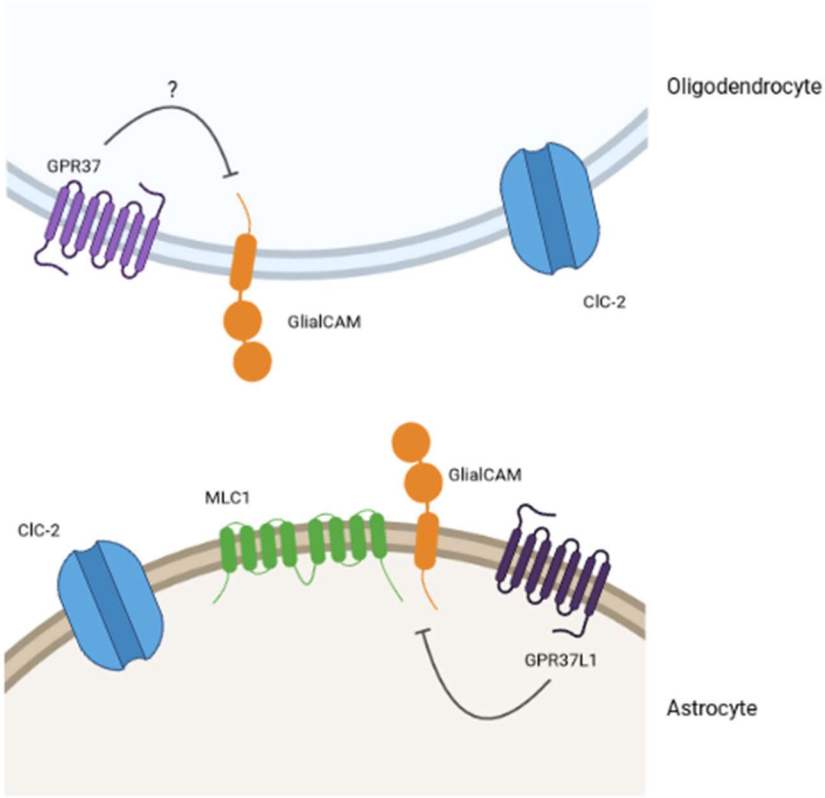


Figure 4. Model for the interplay between GPCRs and MLC-related proteins. The existing pieces of evidence in the literature and the results from our lab lead us to think that GPR37 and GPR37L1 would be part of an interplay with MLC-related proteins. Specifically, we hypothesize (question mark for GPR37 in oligodendrocytes) that these GPCRs would be negative regulators of the physiological MLC1/GlialCAM complex activity.

Previous studies indicated that the interaction between GlialCAM/MLC1 and ClC-2 in primary cultured astrocytes was dynamically regulated, and it was observed only

in depolarizing conditions [47]. Then, we addressed whether the interaction between GPR37L1 and MLC1 was also regulated in a dynamic manner, and compared the interaction between GPR37L1 with MLC1 in physiological versus depolarizing conditions in primary astrocyte cultures [16]. The PLA assays indicated that the interaction between GPR37L1 and MLC1 was decreased, whereas the interaction with the orphan GPCR5B was increased. As GPRC5B interacts more with MLC1 in depolarizing and hypotonic conditions, the activity of GPRC5B might be important in metabolic processes related to changes in the ionic composition. In contrast, we hypothesize that signalling through GPR37L1 might be related to cell differentiation processes. In line with this hypothesis, as mentioned previously, the lack of GPR37L1 resulted in an increase of phospho-ERK1/2, which has also been seen in *Mlc1* KO cells [21].

Another similar correlation has been found recently between MLC proteins and GPR37L1. As mentioned previously, MLC patients [25] and *Mlc1* KO mice [44] show increased seizure susceptibility. A homozygous GPR37L1 variant (c.1047G > T [Lys349Asp]) has been found in a patient with myoclonus epilepsy [68]. Furthermore, both *Gpr37* and *Gpr37l1* KO mice showed an increase in seizure susceptibility [68]. These results are consistent with a possible functional link between these GPCRs and MLC disease. However, it is important to state that future research is needed to understand how GlialCAM and MLC1 modulate GPCR-associated signalling processes.

Funding: This research was funded by MICINN (RTI2018-093493-B-I00).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Van der Knaap, M.S.; Boor, I.; Estevez, R. Megalencephalic leukoencephalopathy with subcortical cysts: Chronic white matter oedema due to a defect in brain ion and water homoeostasis. *Lancet Neurol.* **2012**, *11*, 973–985. [CrossRef]
2. Knaap, M.S.; van der Abbink, T.E.M.; Min, R. Megalencephalic Leukoencephalopathy with Subcortical Cysts. In *Brain Imaging with MRI and CT: An Image Pattern Approach*; Cambridge University Press: Cambridge, UK, 2018; pp. 47–48, ISBN 9780521119.
3. Singhal, B.S.; Gursahani, R.D.; Udani, V.P.; Biniwale, A.A. Megalencephalic leukodystrophy in an Asian Indian ethnic group. *Pediatric. Neurol.* **1996**, *14*, 291–296. [CrossRef]
4. Van der Knaap, M.S.; Barth, P.G.; Vrensen, G.F.; Valk, J. Histopathology of an infantile-onset spongiform leukoencephalopathy with a discrepantly mild clinical course. *Acta Neuropathol.* **1996**, *92*, 206–212. [CrossRef] [PubMed]
5. Van der Knaap, M.S.; Lai, V.; Kohler, W.; Salih, M.A.; Fonseca, M.J.; Benke, T.A.; Wilson, C.; Jayakar, P.; Aine, M.R.; Dom, L.; et al. Megalencephalic leukoencephalopathy with cysts without MLC1 defect. *Ann. Neurol.* **2010**, *67*, 834–837. [PubMed]
6. López-Hernández, T.; Ridder, M.C.; Montolio, M.; Capdevila-Nortes, X.; Polder, E.; Sirisi, S.; Duarri, A.; Schulte, U.; Fakler, B.; Nunes, V.; et al. Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. *Am. J. Hum. Genet.* **2011**, *88*, 422–432. [CrossRef] [PubMed]
7. Sirisi, S.; Folgueira, M.; López-Hernández, T.; Minieri, L.; Pérez-Rius, C.; Gaitán-Peñas, H.; Zang, J.; Martínez, A.; Capdevila-Nortes, X.; De La Villa, P.; et al. Megalencephalic leukoencephalopathy with subcortical cysts protein 1 regulates glial surface localization of GLIALCAM from fish to humans. *Hum. Mol. Genet.* **2014**, *23*, 5069–5086. [CrossRef] [PubMed]
8. Favre-Kontula, L.; Rolland, A.; Bernasconi, L.; Karmirantzou, M.; Power, C.; Antonsson, B.; Boschert, U. GlialCAM, an immunoglobulin-like cell adhesion molecule is expressed in glial cells of the central nervous system. *Glia* **2008**, *56*, 633–645. [CrossRef]
9. Elorza-Vidal, X.; Xicoy-Espaulella, E.; Pla-Casillanis, A.; Alonso-Gardón, M.; Gaitán-Peñas, H.; Engel-Pizcueta, C.; Fernández-Recio, J.; Estévez, R. Structural basis for the dominant or recessive character of GLIALCAM mutations found in leukodystrophies. *Hum. Mol. Genet.* **2020**, *29*, 1107–1120. [CrossRef]
10. Capdevila-Nortes, X.; López-Hernández, T.; Apaja, P.M.; de Heredia, M.L.; Sirisi, S.; Callejo, G.; Arnedo, T.; Nunes, V.; Lukacs, G.L.; Gasull, X.; et al. Insights into MLC pathogenesis: GlialCAM is an MLC1 chaperone required for proper activation of volume-regulated anion currents. *Hum. Mol. Genet.* **2013**, *22*, 4405–4416. [CrossRef]
11. Xu, H.; Isenmann, S.; López-Hernández, T.; Estévez, R.; Lukacs, G.L.; Apaja, P.M. Control of membrane protein homeostasis by a chaperone-like glial cell adhesion molecule at multiple subcellular locations. *Sci. Rep.* **2021**, *11*, 18435. [CrossRef]

12. Lanz, T.V.; Brewer, R.C.; Ho, P.P.; Moon, J.S.; Jude, K.M.; Fernandez, D.; Fernandes, R.A.; Gomez, A.M.; Nadj, G.S.; Bartley, C.M.; et al. Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature* **2022**, *603*, 321–327. [CrossRef] [PubMed]
13. Bosch, A.; Estévez, R. Megalencephalic Leukoencephalopathy: Insights Into Pathophysiology and Perspectives for Therapy. *Front. Cell Neurosci.* **2021**, *22*, 14. [CrossRef] [PubMed]
14. Jeworutzki, E.; Lopez-Hernandez, T.; Capdevila-Nortes, X.; Sirisi, S.; Bengtsson, L.; Montolio, M.; Zifarelli, G.; Arnedo, T.; Muller, C.S.; Schulte, U.; et al. GlialCAM, a protein defective in a leukodystrophy, serves as a ClC-2 Cl^- channel auxiliary subunit. *Neuron* **2012**, *73*, 951–961. [CrossRef]
15. Baldwin, K.T.; Tan, C.X.; Strader, S.T.; Jiang, C.; Savage, J.T.; Elorza-Vidal, X.; Contreras, X.; Rülicke, T.; Hippenmeyer, S.; Estévez, R.; et al. HepaCAM controls astrocyte self-organization and coupling. *Neuron* **2021**, *109*, 2427–2442. [CrossRef] [PubMed]
16. Alonso-Gardón, M.; Elorza-Vidal, X.; Castellanos, A.; La Sala, G.; Armand-Ugon, M.; Gilbert, A.; Di Pietro, C.; Pla-Casillanis, A.; Ciruela, F.; Gasull, X.; et al. Identification of the GlialCAM interactome: The G protein-coupled receptors GPRC5B and GPR37L1 modulate megalencephalic leukoencephalopathy proteins. *Hum. Mol. Genet.* **2021**, *30*, 1649–1665. [CrossRef] [PubMed]
17. Wu, M.; Moh, M.C.; Schwarz, H. HepaCAM associates with connexin 43 and enhances its localization in cellular junctions. *Sci. Rep.* **2016**, *6*, 36218. [CrossRef] [PubMed]
18. Sugio, S.; Tohyama, K.; Oku, S.; Fujiyoshi, K.; Yoshimura, T.; Hikishima, K.; Yano, R.; Fukuda, T.; Nakamura, M.; Okano, H.; et al. Astrocyte-mediated infantile-onset leukoencephalopathy mouse model. *Glia* **2017**, *65*, 150–168. [CrossRef] [PubMed]
19. Brignone, M.S.; Lanciotti, A.; Macioce, P.; Macchia, G.; Gaetani, M.; Aloisi, F.; Petrucci, T.C.; Ambrosini, E. The beta1 subunit of the Na,K-ATPase pump interacts with megalencephalic leucoencephalopathy with subcortical cysts protein 1 (MLC1) in brain astrocytes: New insights into MLC pathogenesis. *Hum. Mol. Genet.* **2011**, *20*, 90–103. [CrossRef]
20. Ridder, M.C.; Boor, I.; Lodder, J.C.; Postma, N.L.; Capdevila-Nortes, X.; Duarri, A.; Brussaard, A.B.; Estévez, R.; Scheper, G.C.; Mansvelder, H.D.; et al. Megalencephalic leucoencephalopathy with cysts: Defect in chloride currents and cell volume regulation. *Brain* **2011**, *134*, 3342–3354. [CrossRef]
21. Elorza-Vidal, X.; Sirisi, S.; Gaitán-Peñas, H.; Pérez-Rius, C.; Alonso-Gardón, M.; Armand-Ugón, M.; Lanciotti, A.; Brignone, M.S.; Prat, E.; Nunes, V.; et al. GlialCAM/MLC1 modulates LRRC8/VRAC currents in an indirect manner: Implications for megalencephalic leukoencephalopathy. *Neurobiol. Dis.* **2018**, *119*, 88–99. [CrossRef]
22. Lanciotti, A.; Brignone, M.S.; Molinari, P.; Visentin, S.; De Nuccio, C.; Macchia, G.; Aiello, C.; Bertini, E.; Aloisi, F.; Petrucci, T.C.; et al. Megalencephalic leukoencephalopathy with subcortical cysts protein 1 functionally cooperates with the TRPV4 cation channel to activate the response of astrocytes to osmotic stress: Dysregulation by pathological mutations. *Hum. Mol. Genet.* **2012**, *21*, 2166–2180. [CrossRef] [PubMed]
23. Teijido, O.; Martínez, A.; Pusch, M.; Zorzano, A.; Soriano, E.; del Río, J.A.; Palacín, M.; Estévez, R. Localization and functional analyses of the MLC1 protein involved in megalencephalic leukoencephalopathy with subcortical cysts. *Hum. Mol. Genet.* **2004**, *13*, 2581–2594. [CrossRef] [PubMed]
24. Hwang, J.; Park, K.; Lee, G.-Y.; Yoon, B.Y.; Kim, H.; Roh, S.H.; Lee, B.-C.; Kim, K.; Lim, H.-H. Transmembrane topology and oligomeric nature of an astrocytic membrane protein, MLC1. *Open Biol.* **2021**, *11*, 210103. [CrossRef] [PubMed]
25. Hamilton, E.M.C.; Tekturk, P.; Cialdella, F.; Van Rappard, D.F.; Wolf, N.I.; Yalcinkaya, C.; Çetinçelik, Ü.; Rajaee, A.; Kariminejad, A.; Paprocka, J.; et al. Megalencephalic leukoencephalopathy with subcortical cysts: Characterization of disease variants. *Neurology* **2018**, *90*, E1395–E1403. [CrossRef]
26. Estévez, R.; Elorza-Vidal, X.; Gaitán-Peñas, H.; Pérez-Rius, C.; Armand-Ugón, M.; Alonso-Gardón, M.; Xicoy-Espaulella, E.; Sirisi, S.; Arnedo, T.; Capdevila-Nortes, X.; et al. Megalencephalic leukoencephalopathy with subcortical cysts: A personal biochemical retrospective. *Eur. J. Med. Genet.* **2018**, *61*, 50–60. [CrossRef]
27. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef]
28. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [CrossRef]
29. Duarri, A.; Teijido, O.; López-Hernández, T.; Scheper, G.C.; Barriere, H.; Boor, I.; Aguado, F.; Zorzano, A.; Palacín, M.; Martínez, A.; et al. Molecular pathogenesis of megalencephalic leukoencephalopathy with subcortical cysts: Mutations in MLC1 cause folding defects. *Hum. Mol. Genet.* **2008**, *17*, 3728–3739. [CrossRef]
30. Lanciotti, A.; Brignone, M.S.; Camerini, S.; Serafini, B.; Macchia, G.; Raggi, C.; Molinari, P.; Crescenzi, M.; Musumeci, M.; Sargiacomo, M.; et al. MLC1 trafficking and membrane expression in astrocytes: Role of caveolin-1 and phosphorylation. *Neurobiol. Dis.* **2010**, *37*, 581–595. [CrossRef]
31. Petrini, S.; Minnone, G.; Coccetti, M.; Frank, C.; Aiello, C.; Cutarelli, A.; Ambrosini, E.; Lanciotti, A.; Brignone, M.S.; D'Oria, V.; et al. Monocytes and macrophages as biomarkers for the diagnosis of megalencephalic leukoencephalopathy with subcortical cysts. *Mol. Cell Neurosci.* **2013**, *56*, 307–321. [CrossRef]
32. Mary Rajathei, D.; Selvaraj, S. Analysis of sequence repeats of proteins in the PDB. *Comput. Biol. Chem.* **2013**, *47*, 156–166. [CrossRef] [PubMed]
33. Pols, M.S.; Klumperman, J. Trafficking and function of the tetraspanin CD63. *Exp. Cell Res.* **2009**, *315*, 1584–1592. [CrossRef] [PubMed]

34. Duarri, A.; Lopez de Heredia, M.; Capdevila-Nortes, X.; Ridder, M.C.; Montolio, M.; López-Hernández, T.; Boor, I.; Lien, C.F.; Hagemann, T.; Messing, A.; et al. Knockdown of MLC1 in primary astrocytes causes cell vacuolation: A MLC disease cell model. *Neurobiol. Dis.* **2011**, *43*, 228–238. [CrossRef] [PubMed]
35. Dubey, M.; Bugiani, M.; Ridder, M.C.; Postma, N.L.; Brouwers, E.; Polder, E.; Jacobs, J.G.; Baayen, J.C.; Klooster, J.; Kamermans, M.; et al. Mice with megalencephalic leukoencephalopathy with cysts: A developmental angle. *Ann. Neurol.* **2015**, *77*, 114–131. [CrossRef]
36. Bugiani, M.; Dubey, M.; Breur, M.; Postma, N.L.; Dekker, M.P.; ter Braak, T.; Boschert, U.; Abbink, T.E.M.; Mansvelder, H.D.; Min, R.; et al. Megalencephalic leukoencephalopathy with cysts: The *Glialcam* -null mouse model. *Ann. Clin. Transl. Neurol.* **2017**, *4*, 450–465. [CrossRef]
37. Hoegg-Beiler, M.B.; Sirisi, S.; Orozco, I.J.; Ferrer, I.; Hohensee, S.; Auberson, M.; Gödde, K.; Vilches, C.; De Heredia, M.L.; Nunes, V.; et al. Disrupting MLC1 and GlialCAM and ClC-2 interactions in leukodystrophy entails glial chloride channel dysfunction. *Nat. Commun.* **2014**, *5*, 3475. [CrossRef]
38. Pérez-Rius, C.; Folgueira, M.; Elorza-Vidal, X.; Alia, A.; Hoegg-Beiler, M.B.; Eeza, M.N.H.; Díaz, M.L.; Nunes, V.; Barrallo-Gimeno, A.; Estévez, R. Comparison of zebrafish and mice knockouts for Megalencephalic Leukoencephalopathy proteins indicates that GlialCAM/MLC1 forms a functional unit. *Orphanet J. Rare Dis.* **2019**, *14*, 268. [CrossRef]
39. Schiza, N.; Sargiannidou, I.; Kagiava, A.; Karaiskos, C.; Nearchou, M.; Kleopa, K.A. Transgenic replacement of Cx32 in gap junction-deficient oligodendrocytes rescues the phenotype of a hypomyelinating leukodystrophy model. *Hum. Mol. Genet.* **2015**, *24*, 2049–2064. [CrossRef]
40. Wilson, C.S.; Dohare, P.; Orbeta, S.; Nalwalk, J.W.; Huang, Y.; Ferland, R.J.; Sah, R.; Scimemi, A.; Mongin, A.A. Late adolescence mortality in mice with brain-specific deletion of the volume-regulated anion channel subunit LRRC8A. *FASEB J.* **2021**, *35*, e21869. [CrossRef]
41. Yang, J.; del Carmen Vitery, M.; Chen, J.; Osei-Owusu, J.; Chu, J.; Qiu, Z. Glutamate-Releasing SWELL1 Channel in Astrocytes Modulates Synaptic Transmission and Promotes Brain Damage in Stroke. *Neuron* **2019**, *102*, 813–827. [CrossRef]
42. Qiu, Z.; Dubin, A.E.; Mathur, J.; Tu, B.; Reddy, K.; Miraglia, L.J.; Reinhardt, J.; Orth, A.P.; Patapoutian, A. SWELL1, a plasma membrane protein, is an essential component of volume-regulated anion channel. *Cell* **2014**, *157*, 447–458. [CrossRef] [PubMed]
43. Voss, F.K.; Ullrich, F.; Münch, J.; Lazarow, K.; Lutte, D.; Mah, N.; Andrade-Navarro, M.A.; Von Kries, J.P.; Stauber, T.; Jentsch, T.J. Identification of LRRC8 heteromers as an essential component of the volume-regulated anion channel VRAC. *Science* **2014**, *344*, 634–638. [CrossRef] [PubMed]
44. Dubey, M.; Brouwers, E.; Hamilton, E.M.C.; Stiedl, O.; Bugiani, M.; Koch, H.; Kole, M.H.P.; Boschert, U.; Wykes, R.C.; Mansvelder, H.D.; et al. Seizures and disturbed brain potassium dynamics in the leukodystrophy megalencephalic leukoencephalopathy with subcortical cysts. *Ann. Neurol.* **2018**, *83*, 636–649. [CrossRef]
45. Larsen, B.R.; Assentoft, M.; Cotrina, M.L.; Hua, S.Z.; Nedergaard, M.; Kaila, K.; Voipio, J.; Macaulay, N. Contributions of the Na^+/K^+-ATPase, NKCC1, and Kir4.1 to hippocampal K+ clearance and volume responses. *Glia* **2014**, *62*, 608–622. [CrossRef]
46. Kinboshi, M.; Ikeda, A.; Ohno, Y. Role of Astrocytic Inwardly Rectifying Potassium (Kir) 4.1 Channels in Epileptogenesis. *Front. Neurol.* **2020**, *11*, 1832. [CrossRef] [PubMed]
47. Sirisi, S.; Elorza-Vidal, X.; Arnedo, T.; Armand-Ugón, M.; Callejo, G.; Capdevila-Nortes, X.; López-Hernández, T.; Schulte, U.; Barrallo-Gimeno, A.; Nunes, V.; et al. Depolarization causes the formation of a ternary complex between GlialCAM, MLC1 and ClC-2 in astrocytes: Implications in megalencephalic leukoencephalopathy. *Hum. Mol. Genet.* **2017**, *26*, 2436–2450. [CrossRef]
48. Morales, J.E.; De, A.; Miller, A.A.; Chen, Z.; McCarty, J.H. Mlc1-Expressing Perivascular Astrocytes Promote Blood-Brain Barrier Integrity. *J. Neurosci.* **2022**, *42*, 1406–1416. [CrossRef]
49. MacAulay, N. Molecular mechanisms of K+ clearance and extracellular space shrinkage—Glia cells as the stars. *Glia* **2020**, *68*, 2192–2211. [CrossRef]
50. Gilbert, A.; Elorza-Vidal, X.; Rancillac, A.; Chagnot, A.; Yetim, M.; Hingot, V.; Deffieux, T.; Boulay, A.C.; Alvear-Perez, R.; Cisternino, S.; et al. Megalencephalic leukoencephalopathy with subcortical cysts is a developmental disorder of the gliovascular unit. *eLife* **2021**, *10*, e71379. [CrossRef]
51. May, D.; Tress, O.; Seifert, G.; Willecke, K. Connexin47 protein phosphorylation and stability in oligodendrocytes depend on expression of Connexin43 protein in astrocytes. *J. Neurosci.* **2013**, *33*, 7985–7996. [CrossRef]
52. Lanciotti, A.; Brignone, M.S.; Visentin, S.; De Nuccio, C.; Catacuzzeno, L.; Mallozzi, C.; Petrini, S.; Caramia, M.; Veroni, C.; Minnone, G.; et al. Megalencephalic leukoencephalopathy with subcortical cysts protein-1 regulates epidermal growth factor receptor signaling in astrocytes. *Hum. Mol. Genet.* **2016**, *25*, 1543–1558. [CrossRef] [PubMed]
53. Hirabayashi, Y.; Kim, Y.J. Roles of GPRC5 family proteins: Focusing on GPRC5B and lipid-mediated signalling. *J. Biochem.* **2020**, *167*, 541–547. [CrossRef] [PubMed]
54. Meyer, R.C.; Giddens, M.M.; Schaefer, S.A.; Hall, R.A. GPR37 and GPR37L1 are receptors for the neuroprotective and glioprotective factors prosaptide and prosaposin. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9529–9534. [CrossRef] [PubMed]
55. Smith, N.J. Drug discovery opportunities at the endothelin B receptor-related orphan g protein-coupled receptors, GPR37 and GPR37L1. *Front. Pharmacol.* **2015**, *6*, 275. [CrossRef] [PubMed]
56. Hu, G.M.; Mai, T.L.; Chen, C.M. Visualizing the GPCR Network: Classification and Evolution. *Sci. Rep.* **2017**, *7*, 15495. [CrossRef] [PubMed]

57. Valdenaire, O.; Giller, T.; Breu, V.; Ardati, A.; Schweizer, A.; Richards, J.G. A new family of orphan G protein-coupled receptors predominantly expressed in the brain. *FEBS Lett.* **1998**, *424*, 193–196. [CrossRef]
58. Donohue, P.J.; Shapira, H.; Mantey, S.A.; Hampton, L.L.; Jensen, R.T.; Battey, J.F. A human gene encodes a putative G protein-coupled receptor highly expressed in the central nervous system. *Mol. Brain Res.* **1998**, *54*, 152–160. [CrossRef]
59. Marazziti, D.; Golini, E.; Gallo, A.; Lombardi, M.S.; Matteoni, R.; Tocchini-Valentini, G.P. Cloning of GPR37, a gene located on chromosome 7 encoding a putative g- protein-coupled peptide receptor, from a human frontal brain EST library. *Genomics* **1997**, *45*, 68–77. [CrossRef]
60. Takahashi, R.; Imai, Y. Pael receptor, endoplasmic reticulum stress, and Parkinson's disease. *J. Neurol.* **2003**, *250* (Suppl. S3), 25–29. [CrossRef]
61. Zeng, Z.; Su, K.; Kyaw, H.; Li, Y. A novel endothelin receptor type-B-like gene enriched in the brain. *Biochem. Biophys. Res. Commun.* **1997**, *233*, 559–567. [CrossRef]
62. Imai, Y.; Soda, M.; Inoue, H.; Hattori, N.; Mizuno, Y.; Takahashi, R. An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of Parkin. *Cell* **2001**, *105*, 891–902. [CrossRef]
63. Marazziti, D.; Di Pietro, C.; Golini, E.; Mandillo, S.; La Sala, G.; Matteoni, R.; Tocchini-Valentini, G.P. Precocious cerebellum development and improved motor functions in mice lacking the astrocyte cilium-,patched 1-associated Gpr37l1 receptor. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16486–16491. [CrossRef] [PubMed]
64. Yang, H.-J.; Vainshtein, A.; Maik-Rachline, G.; Peles, E. G protein-coupled receptor 37 is a negative regulator of oligodendrocyte differentiation and myelination. *Nat. Commun.* **2016**, *7*, 10884. [CrossRef] [PubMed]
65. Keshet, Y.; Seger, R. The MAP kinase signaling cascades: A system of hundreds of components regulates a diverse array of physiological functions. *Methods Mol. Biol.* **2010**, *661*, 3–38.
66. Jolly, S.; Bazargani, N.; Quiroga, A.C.; Pringle, N.P.; Attwell, D.; Richardson, W.D.; Li, H. G protein-coupled receptor 37-like 1 modulates astrocyte glutamate transporters and neuronal NMDA receptors and is neuroprotective in ischemia. *Glia* **2018**, *66*, 47–61. [CrossRef]
67. Marazziti, D.; Mandillo, S.; Di Pietro, C.; Golini, E.; Matteoni, R.; Tocchini-Valentini, G.P. GPR37 associates with the dopamine transporter to modulate dopamine uptake and behavioral responses to dopaminergic drugs. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9846–9851. [CrossRef]
68. Giddens, M.M.; Wong, J.C.; Schroeder, J.P.; Farrow, E.G.; Smith, B.M.; Owino, S.; Soden, S.E.; Meyer, R.C.; Saunders, C.; LePichon, J.B.; et al. GPR37L1 modulates seizure susceptibility: Evidence from mouse studies and analyses of a human GPR37L1 variant. *Neurobiol. Dis.* **2017**, *106*, 181–190. [CrossRef]

Brain Disorders 11 (2023) 100079

Contents lists available at ScienceDirect

Brain Disorders

journal homepage: www.sciencedirect.com/journal/brain-disorders

Characterization of an MLC patient carrying two *MLC1* variants showing radiological improvement

Clara Mayayo-Vallverdú [a,1], Laura Ferigle [b,1], Marta Vecino-Pérez [a], Julián Lara [c], Virginia Nunes [a,*], Raúl Estévez [b,d,**]

[a] *Genes, Disease and Therapy Program, Molecular Genetics Laboratory-Genes, Disease and Therapy Program-IDIBELL, Genetics Section, Department of Physiological Sciences, Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain*
[b] *Physiology Unit, Department of Physiological Sciences, School of Medicine and Health Sciences, Genes, Disease and Therapy Program, Physiology and Pathology of the Functional Relationship Between Glia and Neurons-IDIBELL, Institute of Neurosciences, University of Barcelona, L'Hospitalet de Llobregat 08907, Spain*
[c] *Hospital Universitario Puerta de Hierro, Madrid, Spain*
[d] *Centro de Investigación en red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain*

ARTICLE INFO

Keywords:
Megalencephalic leukoencephalopathy with subcortical cysts
Myelin
Reverting phenotype
Mutations
MLC1

ABSTRACT

Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare type of vacuolating leukodystrophy. Approximately 75% of MLC patients have variants in *MLC1*, while the rest in *GLIALCAM, GPRC5B and AQP4*. From the *GLIALCAM* patients, a classical and a benign phenotype can be distinguished, in which a recessive and dominant inheritance is observed, respectively. Here, we report a new MLC patient harboring two variants in *MLC1* with radiological improvement. The patient is heterozygous for the variants c.597+37C>G and c.895–1G>*T* affecting both mRNA splicing, and the latest causes the deletion p.Pro299_Glu353del. By analyzing mRNA and protein obtained from patient's peripheral blood leukocytes, we could demonstrate the expression of a small amount of wild-type *MLC1* mRNA and protein in the patient. Thus, we suggest that the improvement of clinical and radiological abnormalities observed in all remitting MLC patients might be due to the presence of residual amounts of MLC1 protein.

Introduction

Megalencephalic leukoencephalopathy with subcortical cysts (MLC) is a rare type of leukodystrophy (inherited white matter disease) characterized by early-onset macrocephaly, epilepsy, late-onset mild mental deterioration and cerebral white matter edema [1]. Patient's MRI shows diffuse signal abnormalities of the cerebral hemispheric white matter and the presence of subcortical cysts, mostly located in the anterior temporal region. Based on the clinical phenotype, two groups of MLC patients have been defined: the classic phenotype, in which patients deteriorate, and the remitting phenotype, in which patients show an improvement over time [2].

Classic MLC is caused by recessive variants in the *MLC1* gene (MLC1, MIM 604,004) or recessive variants in the *GLIALCAM* gene (MLC2A, MIM613925) [3,4]. In contrast, patients with remitting MLC present dominant variants in *GLIALCAM* (MLC2B, MIM 613,926) [4]. Recent work identified two siblings with variants in *AQP4* in remitting MLC, and three patients with two *de novo* heterozygous variants in *GPRC5B* in classic MLC [5], which was also identified as a MLC1 and GlialCAM interacting protein [6]. Compared to classic MLC where patients develop neurological signs after the first years, in remitting MLC patients' neurological deterioration does not occur, and although initially MRI abnormalities resemble classic MLC, they improve or normalize [7]. Moreover, it has been described that two patients with MLC1 also presented a radiological improvement, but a detailed characterization is lacking [7]. Here, we identify a novel MLC1 patient that shows a remitting phenotype. We performed genetic and biochemical studies to characterize it.

Abbreviations: MLC, Megalencephalic leukoencephalopathy with subcortical cysts; MRI, Magnetic resonance imaging; PBL, peripheral blood leukocytes.

* Corresponding author.
** Corresponding author at: Centro de Investigación en red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain.
E-mail addresses: vnunes@idibell.cat (V. Nunes), restevez@ub.edu (R. Estévez).

[1] These two authors contributed equally to this work.

https://doi.org/10.1016/j.dscb.2023.100079

Available online 13 June 2023

Case report

Clinical details are summarized in Supplementary Table 1. The patient is now an eleven-year-old boy. His-father is healthy and has a slight macrocephaly. Since the first month of life, the child had evident macrocephaly, being in the percentile>98 group (the brain ultrasound was normal). He was diagnosed with familial macrocephaly due to a history of macrocephaly in the father. At 13 months, he presented a mild simple motor delay along with a delay in cognitive development. Thus, a brain MRI was performed, which showed diffuse signal abnormality and swelling of the cerebral white matter (Fig. 1A). As previously noticed in classic MLC patients [7], there was an abnormal signal in the posterior limb of the internal capsule (Fig. 1A) and slight signal abnormalities in the cerebellar white matter (Supplementary Fig. 1). In addition, the presence of at least one right temporal subcortical cyst was detected (Fig. 1B). A second MRI was performed at three years of age. Swelling was slightly reduced, and the subarachnoid spaces were more visible (Fig. 1A). The central structures including the internal capsule were relatively preserved. Temporal cysts were present on the right and more clearly on the left (Fig. 1B). No cystic lesions were found in other locations outside of both temporal lobes. Cerebellar white matter abnormalities were not present (Supplementary Fig. 1). A third MRI was performed at 9 years of age. Interestingly, a much more reduced swelling and signal abnormalities were observed (Fig. 1A). The patient has never had epilepsy. The motor, language and cognitive evolution has been favorable. He is educated in relation to his age getting good academic results. At the motor level, he can now run, jump, go up and down stairs without difficulty and ride a bicycle.

Molecular studies

Tethered sequencing of cDNA demonstrated that the father carries the c.597+37C>G and the mother carries the c.895–1G>*T* (p.Pro299_Glu353del) variants in the MLC1 gene (Supplementary Table 1). On one hand, the c.597+37C>G variant creates a splice acceptor in intron 7. A recent study identified this variant in another MLC patient [8] and demonstrated by a Minigene splicing assay that this variant results into the creation of 2 novel mRNA isoforms. On the other hand, the c.895–1G>*T* variant also affects splicing, in this case, the splice acceptor site of exon 11, which causes a large deletion in the protein sequence (p.Pro299_Glu353del) and resulted in a shorter RNA sequence (Fragment R3, Fig. 2A and 2B). Surprisingly, we found that the father also has this shorter mRNA sequence although with lower intensity. It has previously found [9] the presence of a MLC1 isoform lacking exon 11 in healthy controls with variable intensity, correlating with the presence of the benign polymorphism c.1059+16G>*A*. Nevertheless, a variant affecting the same nucleotide c.895–1G>C has been reported in a MLC1 patient [10]. Therefore, we concluded that these two variants are pathogenic by affecting splicing.

The improvement observed by MRI and the clinical exploration

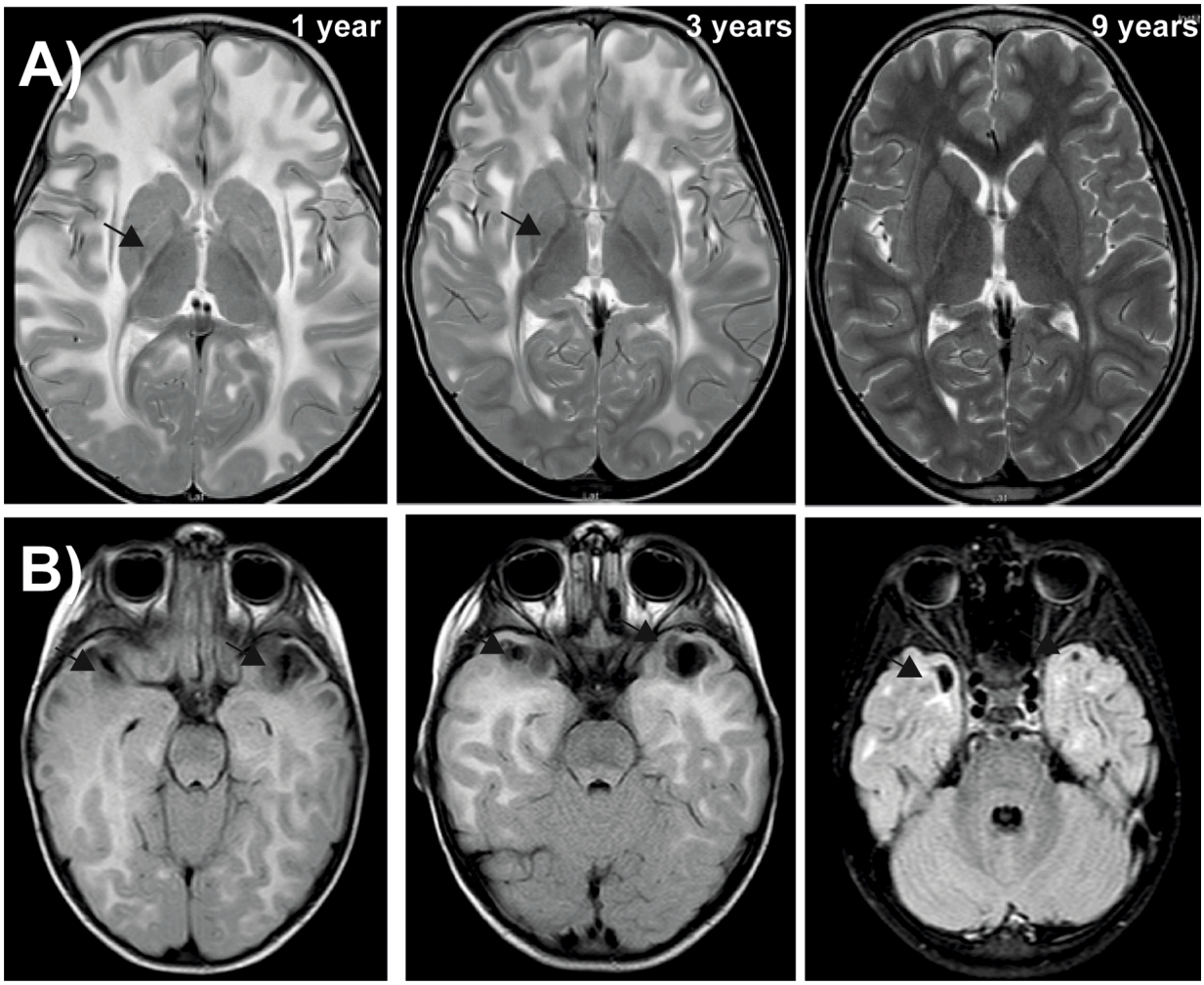


Fig. 1. MRI findings in a patient with MLC containing two MLC1 variants. at age 1, 3 and 9 years. (A) The T2-weighted axial images of the hemispheres show diffuse white matter abnormalities that improve with time. The signal abnormalities observed in the posterior limb of the capsule at the age of 1 and 3 years indicated with an arrow, are much more reduced at 9 years. (B) FLAIR sequences showing the evolution of temporal cysts at different ages. Arrows mark the cysts.

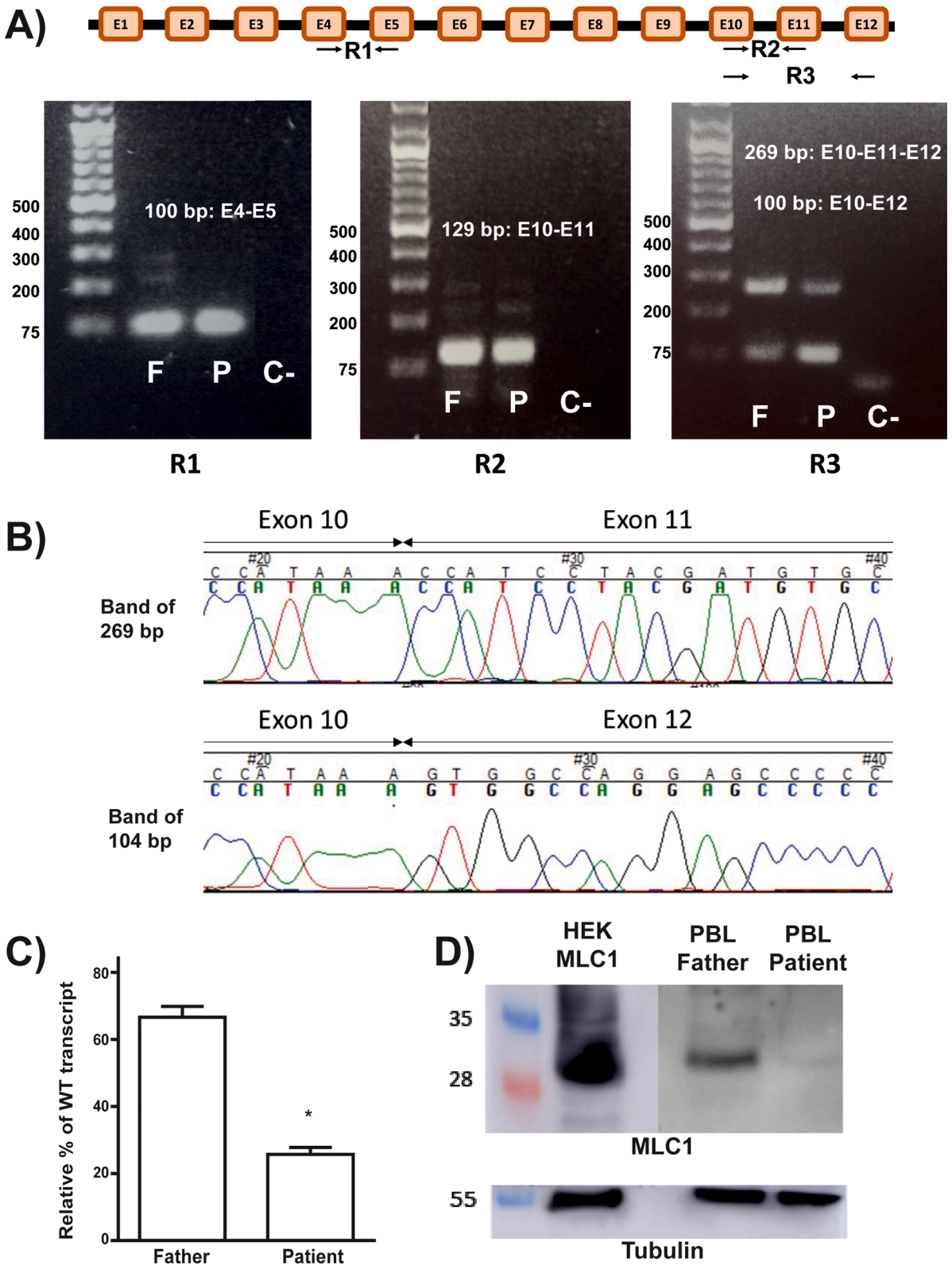


Fig. 2. Detection of wild-type *MLC1* mRNA and protein in PBLs from the patient and his father. (A) Scheme of the strategy with the primers used to detect and quantify the wild-type allele and to confirm the skipping of exon 11. cDNA amplification of the PCR products originated with the primers designed (regions R1, R2 and R3). *F* = Father, *P* = Patient and C- = Negative control. (B) Sequence of the two fragments obtained when amplifying exon 10 to 12, confirming the skipping of exon 11 in the shorter fragment. (C) Quantitative real time PCR data indicating the presence, although at reduced level of the wild-type *MLC1* transcript in the patient (which was statistically significant; **P*< 0.05 *t*-test versus father, *n* = 4). D) Detection of MLC1 in PBLs from both the patient and his father via Western blot. The detection of human MLC1 in transfected HEK293T cells was used as a control. Tubulin was used as a loading control.

could be reasoned by an incomplete penetrance of the splicing variant, which may lead to the expression of some wild-type MLC1 protein. Thus, from blood samples of both the patient and his father, cDNA was obtained. To quantify the level of WT transcript in the patient, we designed a primer pair covering the cDNA sequence from the end of exon 10 and the beginning of exon 11 (Fig. 2A, Fragment R2). As in the other *MLC1* allele the patient has a variant that induces the skipping of exon 11, amplifying this sequence assured us the specificity for the WT transcript avoiding the interference from the mRNA generated by the allele with the c.895–1G>*T* variant. Furthermore, as other authors proposed [11], to normalize the expression of the WT transcript, we amplified a common region of all the transcripts (the end of exon 4 and the beginning of exon 5, fragment R1) (Fig. 2A).

Quantitative PCR results revealed that the percentage of wild-type *MLC1* transcripts in the patient was 25%, while in the father was 66%, relative to each control *MLC1* gene region (Fig. 2C).

To verify that the presence of the wild-type *MLC1* transcript resulted in the presence of MLC1 protein, a solubilized extract from peripheral blood leukocytes (PBLs) from both the father and the patient was obtained and anti-human MLC1 antibodies were used to detect MLC1 protein [12]. A band corresponding to MLC1 wild-type protein was detected in the patient, which was about 10% of the protein detected in the father (Fig. 2D). We expect that the deletion of 54 aa (p.Pro299_Glu353del) will may result in an unstable protein which will be degraded by the proteasome system. Thus, based on the mRNA and protein data, we conclude that the patient expresses a small amount of wild-type MLC1.

Discussion

In this study, we investigated an MLC patient carrying two *MLC1* variants, but who showed an improvement in some MRI features, as typically observed in MLC2B patients [7]. MLC2B patients with the remitting phenotype show an absence of signal abnormalities of the internal capsule and cerebellar white matter. In contrast, in this patient, we observed that he displayed initially signal abnormalities at the posterior limb of the internal capsule and slights alteration in cerebellar white matter pointing to a classical phenotype. Possibly, the levels of expressed MLC1 protein may tune the observed radiological phenotype. We suggest that this patient, together with the other two already described, may belong to a new classification named MLC1B, in which patients with *MLC1* variants display a remitting phenotype.

What is the molecular basis that explains this reversibility? As recently discussed, MLC1 [13] protein expression in humans is higher at one year of age, then it progressively decreases and it is stabilized after 5 years of age. Therefore, we hypothesized that a minimal amount of MLC1 protein might be enough after 5 years of age, but not at earlier stages. In a similar manner, in *Mlc1* KO mice [14], it has been shown that the re-expression of minimal amounts of MLC1, even in adult animals, can abolish the vacuolation observed in the KO [15].

Thus, this work also suggests that a therapy aimed to correct RNA splicing as the use of antisense oligonucleotides or RNA editing could be beneficial to ameliorate the symptoms observed in patients with MLC splicing variants [16,17], as previously performed with one variant [9]. Importantly, these results suggest that just a minor correction would be sufficient to improve the clinical phenotype.

Author contributions

All the authors contributed to the conception and design of the study. RE conceived and planned the experiments in consultation with JL and VN. CM-V and MV-P performed mRNA expression and LF protein data. JL performed clinical evaluation of the patient. RE wrote the manuscript with support from all the authors, which critically reviewed and revised the manuscript draft and approved the final version for submission.

Funding

This research was supported by Ministerio de Ciencia, Innovación y Universidades-Agencia Estatal de Investigación- FEDER-EU. Proyectos I+D+i 2020 Retos Investigación under Grant Number PID2021–126246NB-I00 to RE. This work has been funded by the Instituto de Salud Carlos III through the project PI20/00200 to VN (Co-funded by European Regional Development Fund. ERDF, a way to build Europe), Generalitat de Catalunya Grant SGR2017–191 to VN. We also thank CERCA Programme/Generalitat de Catalunya for institutional support.

Ethics approval

This study was approved by the Hospital Universitario Puerta de Hierro, and written informed consent was obtained from the patient's parents.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

We are grateful to the patient and their family for providing samples and clinical histories and Marjo van der Knaap for her kindness. We thank Ana Maria Cosialls for the technical help to isolate peripheral blood leukocytes.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dscb.2023.100079.

References

[1] M.S. van der Knaap, I. Boor, R. Estévez, Megalencephalic leukoencephalopathy with subcortical cysts: chronic white matter oedema due to a defect in brain ion and water homoeostasis, Lancet Neurol. 11 (2012) 973–985.
[2] M.S. Van Der Knaap, V. Lai, W. Köhler, M.A. Salih, M.J. Fonseca, T.A. Benke, C. Wilson, P. Jayakar, M.R. Aine, L. Dom, et al., Megalencephalic leukoencephalopathy with cysts without MLC1 defect, Ann. Neurol. 67 (2010) 834–837.
[3] P.A. Leegwater, B.Q. Yuan, J. van der Steen, J. Mulders, A.A. Konst, P.K. Boor, V. Mejaski-Bosnjak, S.M. van der Maarel, R.R. Frants, C.B. Oudejans, et al., Mutations of MLC1 (KIAA0027), encoding a putative membrane protein, cause megalencephalic leukoencephalopathy with subcortical cysts, Am. J. Hum. Genet. 68 (2001) 831–838.
[4] T. Lopez-Hernandez, M.C. Ridder, M. Montolio, X. Capdevila-Nortes, E. Polder, S. Sirisi, A. Duarri, U. Schulte, B. Fakler, V. Nunes, et al., Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism, Am. J. Hum. Genet. 88 (2011) 422–432.
[5] P. EMJ, et al., Aquaporin-4 and GPRC5B: old and new players in controlling brain oedema, Brain 139 (2023) 16–17.
[6] M. Alonso-Gardón, X. Elorza-Vidal, A. Castellanos, G. La Sala, M. Armand-Ugon, A. Gilbert, C. Di Pietro, A. Pla-Casillanis, F. Ciruela, X. Gasull, et al., Identification of the GlialCAM interactome: the G protein-coupled receptors GPRC5B and GPR37L1 modulate megalencephalic leukoencephalopathy proteins, Hum. Mol. Genet. 30 (2021) 1649–1665.
[7] E.M.C. Hamilton, P. Tekturk, F. Cialdella, Di.F. Van Rappard, N.I. Wolf, C. Yalcinkaya, Ü. Çetinçelik, A. Rajaee, A. Kariminejad, J. Paprocka, et al., Megalencephalic leukoencephalopathy with subcortical cysts: characterization of disease variants, Neurology 90 (2018) E1395–E1403.
[8] A. Schlüter, A. Rodríguez-Palmero, E. Verdura, V. Vélez-Santamariá, M. Ruiz, S. Fourcade, L. Planas-Serra, J.J. Martínez, C. Guilera, M. Girós, et al., Diagnosis of genetic white matter disorders by singleton whole-exome and genome sequencing using interactome-driven prioritization, Neurology 98 (2022) E912–E923.
[9] C. Mancini, G. Vaula, L. Scalzitti, S. Cavalieri, E. Bertini, C. Aiello, C. Lucchini, R. A. Gatti, A. Brussino, A. Brusco, Megalencephalic leukoencephalopathy with subcortical cysts type 1 (MLC1) due to a homozygous deep intronic splicing mutation (c.895-226T>G) abrogated in vitro using an antisense morpholino oligonucleotide, Neurogenetics 13 (2012) 205–214.
[10] C. Patrono, G. Di Giacinto, E. Eymard-Pierre, F.M. Santorelli, D. Rodriguez, N. De Stefano, A. Federico, R. Gatti, V. Benigno, A. Megarbané, et al., Genetic

heterogeneity of megalencephalic leukoencephalopathy and subcortical cysts, Neurology 61 (2003) 534–537.
[11] J. Camacho Londoño, S.E. Philipp, A reliable method for quantification of splice variants using RT-qPCR, BMC Mol. Biol. 17 (2016).
[12] A. Duarri, O. Teijido, T. Lopez-Hernandez, G.C. Scheper, H. Barriere, I. Boor, F. Aguado, A. Zorzano, M. Palacin, A. Martinez, et al., Molecular pathogenesis of megalencephalic leukoencephalopathy with subcortical cysts: mutations in MLC1 cause folding defects, Hum. Mol. Genet. 17 (2008) 3728–3739.
[13] M. Dubey, M. Bugiani, M.C. Ridder, N.L. Postma, E. Brouwers, E. Polder, J. G. Jacobs, J.C. Baayen, J. Klooster, M. Kamermans, et al., Mice with megalencephalic leukoencephalopathy with cysts: a developmental angle, Ann. Neurol. 77 (2015) 114–131.
[14] M.B. Hoegg-Beiler, S. Sirisi, I.J. Orozco, I. Ferrer, S. Hohensee, M. Auberson, K. Gödde, C. Vilches, M.L. De Heredia, V. Nunes, et al., Disrupting MLC1 and GlialCAM and ClC-2 interactions in leukodystrophy entails glial chloride channel dysfunction, Nat. Commun. 5 (2014) 3475.
[15] A. Sánchez, B. García-Lareu, M. Puig, E. Prat, J. Ruberte, M. Chillón, V. Nunes, R. Estévez, A. Bosch, Cerebellar astrocyte transduction as gene therapy for megalencephalic leukoencephalopathy, Neurotherapeutics 17 (2020) 2041–2053.
[16] B.J. Booth, S. Nourreddine, D. Katrekar, Y. Savva, D. Bose, T.J. Long, D.J. Huss, P. Mali, RNA editing: expanding the potential of RNA therapeutics, Mol. Ther. (2023).
[17] S.F. Hill, M.H. Meisler, Antisense oligonucleotide therapy for neurodevelopmental disorders, Dev. Neurosci. 43 (2021) 247–252.